

## ТЕОРІЯ М.А.К. ХАЛЛІДЕЯ ТА ПРАКТИКА МАШИННОГО ПЕРЕКЛАДУ

Сергій Щербина

Полтавський університет економіки і торгівлі, вул. Коваля, 3, Полтава, 36014, Україна e-mail: [sdsh@ua.fm](mailto:sdsh@ua.fm)

Здійснено спробу спростувати розповсюджений погляд, ніби сучасна теорія прикладної лінгвістики і практика машинного перекладу, зокрема перекладу з використанням ТМ (Translation Memory) мають дуже мало спільного, що свідчить, з одного боку, про кризу прикладної лінгвістики, і з другого боку, про швидкий розвиток прикладних перекладацьких програм (Trados, PROMT) силами фахівців комп'ютерної справи. Мета статті - на прикладі порівняльного аналізу теоретичних положень М. А.К. Халлідея, зокрема його теорії реєстру, стилю (tenor) та режиму (mode) і практики використання програм машинного перекладу з використанням ТМ (Trados, MultiTerm) спростувати або принаймні поставити під сумнів зазначену думку.

*Ключові слова:* теорія прикладної лінгвістики, практика машинного перекладу, перекладацька програма, криза, стиль, режим.

У вступі до однієї зі своїх останніх праць *APPLIED LINGUISTICS AS AN EVOLVING THEME* (2002) М.А.К. Халлідей писав: "Тож, тоді як розвиток *прикладної лінгвістики* був значною мірою процесом поширення в нові сфери, також існував рух виходу з тих сфер, які початково вважалися *центральними*. Звичайно, це не означає, що такі сфери діяльності більше не є на меті, це правда, що машинний переклад таки вийшов із моди в багатьох головних центрах, але певні форми обробки природної мови за допомогою комп'ютера продовжуються з того часу. Тільки-но це набуло окремої форми сфери дослідження і розвитку з її власними інституційними структурами і власним дискурсом, головне тому, що технологічні досягнення трансформували як наявні ресурси, так і знання спеціалістів, що потрібні для використання цих ресурсів. Отож, якщо ми говоримо про прикладну лінгвістику як про розвиток, то це не означає, що вона просто збільшується... В цьому сенсі лінгвістика, яку розуміють як низку принципів і методів дослідження і пояснення мови, може вважатися дисципліною, і якщо ми візьмемо це за відправний пункт, то можемо сказати, що прикладна лінгвістика розвивалася як використання результатів лінгвістики, щоб звернутися до інших проблем - чи інших об'єктів вивчення, якщо ми думаємо про дослідницьке застосування, чи якісь інші практичні проблеми, такі цілеспрямовані види діяльності як викладання мови і машинний переклад" [3, с. 9]. Є всі підстави вважати, що вчений мав на увазі кризові явища в основному руслі прикладної лінгвістики, проте коли писав про "ресурси" і "знання спеціалістів", то йшлося про сучасні ПК і нову генерацію програмістів, яким зовсім необов'язково знатися на лінгвістиці для створення програм машинного перекладу. Про які "окремі форми сфери дослідження і розвитку з її власними інституційними структурами і власним дискурсом" йдеться в статті лінгвіста? У Як робочу гіпотезу ми можемо висунути здогадку, що однією з таких форм є технічний переклад із застосуванням ТМ. Як свідчить практика, саме цей вид машинного перекладу набув надзвичайної популярності в різноманітних галузях господарства, будівництва і промисловості. Пропозиції виконання перекладу з використанням програми Trados рясніють в Інтернеті, причому замовники охоче надають перекладацькі бази і вимагають дотримуватись їх найсуворіше. Здавалося б, практика комп'ютерного перекладу набагато випередила теоретичні розробки прикладної лінгвістичної теорії. Проте саме теоретичні положення М. А.К. Халлідея стали основою і дали ключ до практичної реалізації цього напрямку перекладацької діяльності.

Центральне місце в нашому порівняльному аналізі займає теорія реєстру, одним з авторів якої є М. А.К. Халлідей. Використання терміна "реєстр" започаткував лінгвіст Томас Бертран Рід у 1956 р, але в загальний ужиток він увійшов 1961 р. Ввели його лінгвісти, які хотіли розрізнити зміни в мові залежно від користувача і визначалися такими змінними: соціальні умови, географія, стать і вік, зміни залежно від використання (в тому сенсі, що в кожного мовця є діапазон варіантів і вибору між ними в різний час). Увага сфокусована на способі, за допомогою якого

мова використовується в окремих ситуаціях: юридичний жаргон, мова науково-дослідницької лабораторії чи газетного повідомлення. М.А.К. Халлідей інтерпретує реєстр як "мовні риси, які зазвичай асоціюються з конфігурацією ситуативних особливостей, - з особливим значенням поля, режиму і стилю" [5]. Халлідей обговорював термін "реєстр" у деталях. Цей термін належить до взаємозв'язку між мовою (та іншими семіотичними формами) і функцією контексту [2, с. 97]. Цей зв'язок характеризують за допомогою описових категорій поля, стилю (tenor) і режиму (mode). Реєстри змінюються. Є ключі або індекси мови, які допоможуть нам передбачити, яким є реєстр цього тексту.

Змінні реєстру визначають відносини між мовою і функцією форми мови [2, с. 112]. Описова функція надає актуальну інформацію. Ви можете перевірити цю інформацію, а потім прийняти або відхилити її. Виразні функції дають інформацію про належність мовця. Він використовує висловлювання, щоб висловити свої почуття, незалежно від будь-якої відповіді. Соціальна функція виявляє особливий зв'язок між мовцем і слухачем. Щодо інформаційної функції, то основною інформаційною функцією мови є зовнішні ситуації, факти, теми, реальність поза мовою, в тім числі ідеї або теорії.

Поле - це особливості ситуації, що спонукають до використання мови певного виду. Мова заповнена словами, пов'язаними з об'єктами у середовищі, що нас оточує, процесами діяльності і таке інше. На думку Халлідея [2, с. 125], поле передає те, що відбувається з соціальною природою дій, які відбуваються.

Режим (mode) - це роль самої мови, в цьому випадку - текстові функції, тісний зв'язок між словами у всій їхній взаємодії. Ця роль "працює", щоб згуртувати весь текст, тобто обидва учасники спільно будують значущий комунікативний акт з використанням мови.

Стиль (tenor) за Халлідеем [2, с. 137] - це відносини між учасниками, що також є окремим випадком використання мови, репрезентує міжособистісні функції мови. Може простежуватись вибір референта, використання імператива тощо. Отож, стиль підпорядкований тому, хто з ким спілкується.

Підручники лінгвістики можуть вживати термін "тенор", проте вони надають більшу перевагу терміна "стиль" - "ми характеризуємо стилі як варіанти мови, що розглядаються з погляду формальності" [4, с. 21], визначаючи реєстри більш вузько як використання мови спеціалістами, пов'язане з особливою діяльністю на кшталт академічного жаргону.

Якщо під кутом теорії Халлідея подивитися на феномен перекладу, то можна зробити низку певних узагальнень. Переклад є явищем контекстуальним, це міжкультурна комунікація, комунікативний акт, що намагається точно передати контекстне значення, так що і зміст, і мова стають прийнятними для читачів. Він повинен бути зрозумілим, простим, соціальним і справляти на свого читача такий ефект, який максимально близький до ефекту, який спричинює читання оригіналу. Він відповідає вимогам особливого реєстру мови, стосується головно потреб реципієнта в контексті мови та різноманітних реєстрів і точно відтворює загальні риси та специфічні особливості оригіналу.

Тепер повернемося до машинного перекладу з застосуванням ТМ. Будь-який замовник на питання про популярність саме цього типу перекладу відповідь, що такий переклад максимально близький до оригіналу, тож є практично еквівалентним йому. Тож головним критерієм перекладу, особливо для технічної сфери, є еквівалентність. Не можемо замовчувати той факт, що другим критерієм для замовника є швидкість перекладу, оскільки давно відомо, що будь-який переклад потрібен "не на завтра, а на вчора". Цілком очевидно, що замовник цінує здатність комп'ютерної програми швидко обробляти великі масиви текстів. Про цю важливу якість машинного перекладу ми поговоримо дещо пізніше.

Отже, розглянемо еквівалентність з погляду теорії, і з погляду машинного перекладу. Еквівалентність може бути наявною на багатьох рівнях: рівні слова, рівні словосполучень, граматична еквівалентність, текстова еквівалентність (рівень теми і реми). Проте нас цікавить саме та еквівалентність, яка може бути "цеглинкою" в побудові еквівалентних (з огляду SL і TL) речень, а на їхній основі - текстів, при машинному перекладі. Дуже корисним, на наш погляд, є постулат М.А.К. Халлідея, що еквівалентність одиниць і речей втрачається тоді, коли ми спускаємося нижче рівня речення. Проте сам лінгвіст застерігав, що морфема перекладу не підлягає. Отож, предметом перекладу може бути речення, словосполучення та слово. Не треба проводити

спеціальних лінгво-статистичних досліджень, щоб виявити досить невелику частку речень, що складаються з одного слова ("Сутеніло" (укр.), *Ojo!* (ісп.), то навряд чи програміст закладав таку можливість у створення комп'ютерної програми-перекладача. Перепорою для програмування перекладу розповсюджених речень є передусім конфлікт між граматичною, а отже, текстовою еквівалентністю речень, що належать до різних мовних груп (непридатність до еквівалентного перекладу структур типу *He will make a good husband* чи стосунків актив / пасив). Отже, залишаються словосполучення, які деякі лінгвісти розглядають як своєрідні речення, зокрема О. М. Пешковський.[1, с. 163]

Тепер, коли ми значно звузили коло об'єктів дослідження, можна переходити до суто прагматичних проблем машинного перекладу, а саме: якими є суттєві складові частини забезпечення еквівалентності словосполучень при машинному перекладі. Під час аналізу перекладу зі слов'янських мов на германські, зокрема, на англійську і навпаки, визначимо ті властивості словосполучень, які типологічно подібні в обох мовах і чия парадигматика не становить складнощів для програмування.

Передусім візьмемо групу нерозкладних словосполучень, які мають структурну схему, певне граматичне значення, а їхні компоненти об'єднані одним із способів підрядного зв'язку (узгодження, керування, прилягання). Наше завдання полегшуємо тим, що в них головний компонент не має семантичної повноти, чітко вираженого лексичного значення, воно конкретизується або доповнюється значенням залежного слова, яке несе основне семантичне навантаження, внаслідок чого формується словосполучення з нечленованою, цілісною семантикою. Це й пов'язує обидва компоненти в нероздільне ціле. Якщо виключити з цієї групи фразеологізми і прислів'я, які практично ніколи не вживаються в технічних текстах, то можна досягнути практично стовідсоткової еквівалентності при перекладі таких словосполучень.

Базисним у написанні програми є програмування трьох видів підрядного зв'язку - узгодження, якщо форми числа, роду і відмінка прикметників, дієприкметників, порядкових прикметників, займенникових прикметників уподібнюються відповідним формам іменників; керування, якщо залежне слово використано в формі непрямого відмінка без прийменника чи з прийменником; прилягання, якщо залежне слово має нульову форму. Враховуючи специфіку англійської мови, ми не будемо включати параметрів роду, а параметри відмінка будуть в дещо обмеженому, порівняно зі слов'янськими мовами, варіанті.

Приблизно за такою ж схемою відбувається і другий крок: треба запрограмувати три основні типи семантично-синтаксичних відношень: 1) атрибутивні (означальні) відношення - відношення ознаки до предмета; 2) субстанціальні відношення - відношення предмета до дії, процесу, стану та інших ознак; 3) адвербіальні відношення - відношення обставини до дії, процесу, стану, якості тощо. Виходячи з суто прагматичних міркувань потреб технічного перекладу, основна увага програміста буде зосереджена на побудові парадигми атрибутивних і субстанціальних відносин і значно менше - на адвербіальних відносинах.

Отже, ми маємо реєстр, або специфічний для якоїсь певної галузі набір словосполучень (наприклад, перекладацька база для виконання перекладу з англійської на російську нафтогазової тематики), маємо змінні реєстру, які реалізуються як при програмуванні - передача, скажімо, атрибутивних відносин, - так і при практичному застосуванні - в межах нафтогазової тематики нас цікавлять тільки ті значення, що належать до офшорного буріння, тобто змінюється поле значень. Наскільки це важливо у разі забезпечення еквівалентності перекладу, можна побачити з такого прикладу: словник із нафти та газу перекладає словосполучення *unit stress* як "напруга на одиницю перетину", словник з офшорного буріння дає переклад "питоме навантаження", що зовсім не одне й те ж саме.

Важливість звертати увагу на стиль саме в розумінні його М.А.К. Халлідеєм може бути проілюстрована таким прикладом. Якщо взяти словосполучення *unit rule* і спробувати визначити його реєстр *per se*, то це може бути або перекладацька база з загальноправничої тематики, або з охорони праці. Режим, тобто текстове оточення скеровує нас на пошук значення в царину американського законодавства. І тільки *tenor*, або погляд на словосполучення як на окремий випадок, а не, скажімо, як вільне словосполучення, дає нам змову виявити, що належить воно до термінології політичного устою партійної системи США і перекладається як "правило

одностайності", тобто обов'язок усіх членів делегації штату голосувати на партійному з'їзді єдиним блоком.

Проте знаходження точного перекладацького еквівалента терміна зовсім не гарантує (при звичайному, не машинному перекладі), що натрапивши на цей термін через півроку-рік перекладач зможе відтворити процес знаходження еквівалентності, який одного разу привів його до успіху. Причин цього може бути декілька: втрата відповідної довідкової літератури, брак часу, та інші людські чинники. Машинний переклад дає змогу подолати ці недоліки шляхом створення перекладацьких баз, де один знайдений еквівалент зберігається необмежено довго. Такою базою можна ділитися з менш досвідченим перекладачем, що на практиці й роблять замовники, надсилаючи перекладацькі бази виконавцям перекладу, які користуються тим самим реєстром. До того ж, перекладацька база може витримувати будь-які модифікації: оновлення, виправлення, побудова нової бази на базі старої при зміні поля чи режиму (наприклад, маємо загальний реєстр нафтогазової тематики, проте нам потрібна база з офшорного буріння; частина перекладацьких еквівалентів залишається придатною для нового напрямку, а частину ми змінюємо, зберігаємо цю базу під новою назвою).

Перекладацькі програми *Trados i Multiterm*, у яких реалізовані зазначені функції, дають змогу задовольнити ще одну вимогу до перекладу, особливо технічного - вимога однаковості термінології впродовж усього тексту. Програма "запам'ятовує" переклад словосполучення, якщо його на якомусь етапі ввели в перекладацьку базу і "підказує" перекладацький еквівалент кожного разу, коли цей вираз трапляється. Всі ці наведені якості перекладацьких програм допомагають забезпечити високу швидкість перекладу, що, як зазначено раніше, теж є важливою вимогою замовника.

Саме поєднання теоретичних положень М. А.К. Халлідея і практичних навичок машинного перекладу з використанням зазначених програм, надає перекладачеві ключ до розуміння, а отже, й удосконалення комп'ютерного перекладу технічних текстів. Щодо програміста, який працює над удосконаленням перекладацького ПЗ, то теорію М. А.К. Халлідея він може взяти як одну з методологічних засад і підказати певний алгоритм програмування саме тих рис і відносин, що суттєво підвищать еквівалентність перекладу.

1. Пешковский А.М. Вопросы методики родного языка, лингвистики, стилистики. - М.-Л.: Работник просвещения, 1923. 2. Halliday, M. A. K. and Hasan, R. Cohesion in English., London, Longman, 1976 3. Language and Education. The Collected Works of M. A. K. Halliday, Edited by Jonathan J. Webster, Continuum 2007. 4. Trudgill, P., Introducing language and society, London: Penguin, 1992 5. [http://en.wikipedia.org/wiki/Register\\_\(sociolinguistics\)](http://en.wikipedia.org/wiki/Register_(sociolinguistics))

## **ТЕОРИЯ М.А.К. ХАЛЛИДЕЯ И ПРАКТИКА МАШИННОГО ПЕРЕВОДА**

**Сергей Щербина**

*Полтавский университет экономики и торговли, ул. Коваля, 3, Полтава, 36014, Украина e-mail: [sdsh@ua.fm](mailto:sdsh@ua.fm)*

Исследовано наличие корреляции между центральными положениями теории М. А.К. Халлидея о регистре, режиме (*mode*), стиле (*tenor*) и практикой перевода технических текстов с использованием программного обеспечения на основе "переводческой памяти" (TRADOS, Multiterm).

*Ключевые слова:* регистр, режим, стиль, практика перевода, технический текст, переводческая память.

## **THE THEORY OF M.A.K. HALLIDAY AND THE PRACTICE OF MACHINE TRANSLATION**

**Serhiy Shcherbyna**

*The University of Economy and Commerce in Poltava, 3, Koval St, Poltava, 36014, Ukraine e-mail: [sdsh@ua.fm](mailto:sdsh@ua.fm)*

The article touches upon the issues of the relevance of the main principles of M.A.K. Halliday's

theory concerning the register, mode and tenor for the day-to-day practice of the translators that resort to the memory-based computer software (TRADOS, Multiterm).

*Key words:* correspondence, register, mode, tenor, technical texts, translation memory .