

Тотрова Д. Б. О понятиях «лексико-тематическая группа» и «лексико-семантическая группа». URL: http://www.viu-online.ru/tl_files/nauka/1%20-37%20Totrova.doc. (дата звернення: 21.01.2015).

Туровська Л. В. Терміни та номени у науково-технічній сфері. *Українська термінологія і сучасність* : зб. наук. праць. Київ, 2005. С. 225-229

Шведчикова Т. В. Лексико-семантическая группа как отражение языковой системы. URL: <http://journal.sakhgu.ru/archive/2005-04-7.doc>. (дата звернення: 15.02.2015)

COCA – Corpus of Contemporary American English. URL: <http://corpus.byu.edu/coca/>. (access date: 28.12.2017).

(Матеріал надійшов до редакції 11.02.19. Прийнято до друку 24.03.19)

УДК : 81'255: 004

DOI: <https://doi.org/10.26661/2414-1135/2019-76-11>

ЛАЗЕБНА Н. В.

(Національний університет «Запорізька політехніка»)

СЕМАНТИЧНА НЕОДНОЗНАЧНІСТЬ УРБАНІСТИЧНОЇ ТЕРМІНОЛОГІЇ (УКРАЇНСЬКО-АНГЛІЙСЬКИЙ ПЕРЕКЛАД GOOGLE TRANSLATE VS ЛЮДСЬКИЙ ПЕРЕКЛАД)

Стаття присвячена подоланню семантичної неоднозначності урбаністичної термінології, виявленню семантичних та синтаксичних помилок у процесі машинного перекладу Google Translate, та їх порівнянню із людським перекладом. Проаналізовано 33000 перекладених друкованих знаків з пробілами з української англійською мовою. Людський фактор перекладача у цьому випадку «перемагає» нейронну мережу Google Translate.

Ключові слова: семантика, синтаксис, машинний переклад, нейронна мережа, фаховий переклад, людський фактор.

Lazebna N. Semantic Ambiguity of Urbanistic Terminology (Ukrainian-English Google Translate vs Human Translation).

The problem of semantics is one of the key concerns both for machine and human translation. Despite the fact that neural networks have been used as a central background for Google Translate, a human factor has contributed much to the final version of the translation. The optimal solution to the problems of semantics is the use of interactive mechanisms (dialogue, semiautomatic), which provide a mutual decision to the problem by a human being and the computer. The human factor of the translator is a decisive tool for quality translation achievement. Finally, the target text sounds as a source text, following all pragmatic intentions of the source text author.

This research considers the translation of an urban design article from Ukrainian into English (Google Translate vs human translation). Based on the analysis of 33,000 printed symbols, the semantic errors were identified and corrected. In addition to the semantic charges of words that Google Translate tries to correct using neural networks, the structure of the target language sentences does not meet the grammatical requirements of the English language in the analyzed passage. Machine translation seems to copy the structure of the original sentence and translates it not into English, but into "UkrEnglish". The translator based on his 'sense of language' restructures the sentence and in the process of translation he chooses the necessary lexical translation equivalents. Obviously, there are more bugs than Google Translate developers believe. Of course, the translation that appears is understandable in many cases but sounds unnatural in terms of the target language norms. Thus, in the passage examined, the translator made 244 corrections as can be seen from the Word automated comparison of the professional translation text and Google translate. Structural adjustments, semantic correspondences, a large number of pronouns, prepositions, and even articles have overloaded the Google Translate target text.

Keywords: semantics, syntax, machine translation, neural networks, professional translation, human factor.

Глобальне павутиння все більше поглинає сучасне суспільство. Автоматизація людської діяльності, домінування роботів та механізмів над людьми стирає будь-які межі можливостей технологій. Нейронна мережа, яку використовує онлайн-перекладач Google Translate, призвела до створення перекладів із мов, які не були спочатку включені до цієї системи [Melvin 2017]. На думку сучасних дослідників, ця внутрішня репрезентація всередині мережі є універсальною мовою людства. Скоріше, це певний машинний код, з яким відбувається робота на внутрішньому проміжному рівні нейромережі. Це машинна мова-посередник між будь-якими мовними парами всіх людських мов, підтримуваних системою.

Проте, існування такої універсальної бази смислів, яка об'єднує всі мови світу, є непередбаченим та незбагненим полем діяльності для лінгвістів, перекладачів, програмістів та інших професіоналів і зумовлює **актуальність** досліджень такого напрямку. Лінгвісти та

перекладачі існують на перетині двох потужностей: нейронних мереж Google Translate та людських знань.

Можливо, розробники створили штучне уявлення про текст як щось цілісне. Мова інтерлінгва – універсальна, орієнтована на її використання в технологіях автоматичного перекладу. Фахівці компанії Google використовують zero-shot translation, який є більш складним методом перекладу. Оцінити текст перекладу як дещо ціле і є тією особливістю або перевагою людини перед роботом-системою перекладу.

Метою дослідження є виявити семантичну неоднозначність урбаністичної термінології під час машинного перекладу Google Translate з української англійською мовою та порівняти її рівень після перекладача. **Завданням** є виявлення семантичних помилок під час машинного перекладу обраною програмою україномовного тексту з урбаністичної тематики англійською мовою.

Предметом дослідження є семантичні помилки обраного уривку та його перекладу. **Об'єктом** дослідження є якість машинного перекладу Google Translate у порівнянні із перекладом, зробленим людиною-перекладачем.

Сучасна урбаністична термінологія передає ідею «щільності забудов» на тлі простору, що зменшується. Відповідно, семантичне навантаження такої термінології передбачає підхід «місто скрізь». Будівлі в обіймах міського пейзажу, тривимірність міста, дивні форми і конструкції, різноманітність і множинність інноваційних будівель визначають сутність новаторських урбаністичних термінів. З'явилося "вертикальне передмістя" і вплинуло на семантику термінів. У результаті основна частина використовуваної термінології стосується опису вертикалі елементів будівель і споруд (вертикальний брус, тривимірні, викривлені балкони тощо). Сучасні тенденції в архітектурі зробити акцент на монументальності будь-яких видів будівель. Таким чином, соціальне житло стало шедевром з оригінальними ідеями та концепціями. Семантика таких термінів головним чином стосується заплутаного характеру сучасних тенденцій урбанізації.

У такому ключі і мають віднаходити відповідники перекладачі. На відміну від людини-перекладача, штучний інтелект оперує штучною мовою, яка не розбиває речення на фрази або окремі слова, а перекладає все речення повністю не випускаючи нічого із контексту. Однак, аналізований уривок демонструє необхідність перебудови речень, добору адекватних відповідників, виправлення граматичних помилок тощо.

Google Translate vs людський переклад.

Нейронна мережа постійно самовдосконалюється, збагачується, і може працювати відразу з багатьма парами мов (англійська-японська; англійська-корейська). Унікальність машинного перекладу Google Translate не можна не підкреслити. Наприклад, Гугл зможе виконати переклад японсько-корейський хоча спочатку не було закладено механізмів перекладу між цими двома мовами. Більш того, англійська мова вже не буде проміжною між ними включаючи ті, які не були задіяні в процесі навчання. Наприклад, якщо систему навчили перекладу між англійською та японською мовами, і англійською і корейською, то вона вільно може виконати переклад між японською і корейською, але вже без використання англійської мови в якості проміжної. «За останні десять років система автоматичного перекладу текстів Google Translate зросла з кількох мов до 103, а зараз вона перекладає 140 млрд слів щодня» [Castelvecchi 2017].

Глибинні механізми сервісу Google Translate вдосконалюють переклад, спираючись на контекст. Нейромережа Google для машинного перекладу називається Google Neural Machine Translation (GNMT). Інноваційний рівень, на якому оперує ця система, є новим, і спирається на «розподіл тексту» за «словниковими сегментами». Останні результати, отримані в основному за допомогою глибинних нейронних мереж, показують, що можна поліпшити якість алгоритмів прогнозування, якщо переходити до аналізу великих і надвеликих навчальних вибірок [Schuster 2016; Surafel et al 2017]. Для великої вибірки вкрай складно

отримати її т.зв. «повну розмітку» (коли досить багато об'єктів і дуже складно виміряти значення прихованих змінних). У той же час виявляється, що на практиці для багатьох завдань можливо набрати достатньо велику кількість об'єктів, для яких невідоме точне значення прихованої змінної, але відома підмножина її можливих значень [Ефремова 2017].

Відповідно, згідно з латентною семантичною моделлю, ця задача є типовим прикладом завдання зі слабкою розміткою на відміну від задачі прогнозування слова за контекстом, де емпіричним матеріалом може слугувати будь-який текст. «Під об'єктом в цьому випадку розуміється просто позиція слова в тексті; змінна - це контекст певного слова, а прихована змінна - це саме слово, яке намагатимуться спрогнозувати за контекстом, і прихований компонент відомий для будь-якої позиції в тексті» [Ветров].

Реалізацію того чи іншого значення слова здійснює контекст або ситуація, загальна тематика мови. Контекст обумовлює конкретне значення багатозначного слова, і в певних умовах він може створювати семантичну дифузність, тобто сумісність окремих лексичних значень, коли їх розмежування не відбувається. Не тільки лексична сполучуваність і словотворчі особливості характеризують різні значення слів, але також в ряді випадків і особливості граматичної сполучуваності [Іваницька 2017].

У прикладних задачах комп'ютерної лінгвістики не розмежуються омонімічні і полісемічні значення слів. Це пов'язано з тим, що в переважній більшості прикладних задач важлива не стільки етимологія слова, скільки його семантика. Оптимальним вирішенням проблем семантики є інтерактивні механізми (діалогові, напівавтоматичні), які передбачають спільне рішення задачі людиною і комп'ютером, і зводяться до того, що комп'ютер надає користувачу набір альтернатив, з якого він повинен вибрати один варіант.

Людський фактор перекладача на сьогодні є вирішальним у досягненні якісного перекладу, який буде звучати так само як і оригінал, досягне такого ж комунікативно ефекту та прагматичних настанов. Спираючись на «класичні» визначники якості перекладу, звернемося до таких понять як «адекватність» та «еквівалентність».

Адекватність – це таке співвідношення вихідного і кінцевого текстів, при якому послідовно враховується мета перекладу. Терміни «адекватність» і «адекватний» орієнтовані на переклад як процес, тоді як терміни «еквівалентність» і «еквівалентний» мають на увазі відношення між вихідним і кінцевим текстами, що виконують подібні комунікативні функції в різних культурах. На відміну від адекватності, еквівалентність орієнтована на результат.

Критерієм адекватності є ступінь відступу від еквівалентності, обумовлений об'єктивною необхідністю, а не сваволею перекладача. Семантико-стилістична адекватність, як впливає зі сказаного вище, визначається через оцінку відношення текста перекладу до текста оригіналу, а точніше кажучи, через оцінку семантичної і стилістичної еквівалентності мовних одиниць, що складають текст перекладу і текст оригіналу. Поняття семантико-стилістичної адекватності релевантно для оцінки власне лінгвістичного змісту, або аспекту, перекладацької діяльності.

Для прикладу розглянемо переклад статті з урбаністичного дизайну англійською мовою та порівняємо його із версією перекладу Google translate.

	Вихідний текст	Людський переклад	Google translate
	У такому підході до дизайну міста формується образ міста і міського середовища, що означає автотель культури, інструмент його самопізнання і самоосмислення.	This approach creates the image of the city and urban environment design defining the auto model of culture and the tool of its self-cognition and reflection.	In this approach to the design of the city, an image of the city and the urban environment is formed, which means the car model of culture, a tool for its self-knowledge and self-reflection.

<p>Душа міста входить в склад міста не повністю, образ не включає ту її частину, яка не відображається в свідомості і почуттях людей.</p>	<p>The soul of the city is not its integral constituent. The image does not reflect a subconscious part of the city soul, an unperceived component for minds and feelings of people.</p>	<p>The soul of the city parts the city not completely, the image does not include that part which is not visible for minds and feelings of the people.</p>
---	--	--

У загальній вибірці проаналізовано 33000 друкованих знаків з пробілами з української англійською мовою. Окрім семантичних зв'язків, які намагається виправити Гугл застосовуючи нейромережі, у цьому уривку структура речень мови перекладу не відповідає граматичним вимогам англійської мови. Здається, що машинний переклад копіює структуру вихідного речення і перекладає не на англійську, а «укранглійську». Перекладач одразу бачить необхідність у перебудові структури речення і в процесі перекладу вже підбирає необхідні лексичні перекладні еквіваленти. Тож, помилки очевидно існують і їх набагато більше, ніж вважають розробники Гугл. Звісно, переклад, який видається, є зрозумілим у багатьох випадках, але звучить неприродно з точки зору мови перекладу.

Застосовуючи популярні метрики пошуку помилок у тексті, можливо проаналізувати які ж найбільш частотні помилки виникають у Google translate. Такими метриками є WER, BLEU, NIST і METEOR [Ветров]. Однак, перекладач може виявити помилки тексту перекладу за довший час, але викоринити їх та замінити на правильний варіант на відміну від автокорегувань.

Можливо класифікувати синтаксичні помилки: (1) неправильний порядок слів, 2) неправильна форма дієслова, 3) відсутні або неправильні артиклі, 4) відсутні або неправильні займенники.

Під (1) неправильним порядком слів маються на увазі слова, розміщені системою машинного перекладу в невірному порядку, таким чином, що перекладне речення є неграмотно написаним англійською мовою. (2) Неправильна форма як помилка виникає коли системі машинного перекладу не вдається розрізнити форму дієслова. 3) Відсутні або неправильні артиклі та 4) відсутні або неправильні займенники перевантажують речення на мові перекладу англійською мовою.

Семантичні помилки такі як (1) пропущені слова, (2) неправильні перекладні еквіваленти, (3) додаткові слова. 1) Пропущені слова можуть порушити зміст речення. (2) Неправильні переклади зустрічаються, якщо система машинного перекладу не може знайти потрібний переклад вихідного слова. В цьому випадку перекладне слово просто може бути абсолютно неправильним або багатозначним. Під (3) додатковими словами маються на увазі слова, відсутні у вихідному реченні, але присутні в перекладеному реченні.

Висновки і перспективи дослідження. Отже, у дослідженому уривку, перекладачем було виконано 244 виправлення, як видно із автоматизованого порівняння фахового тексту перекладу та варіанту Google translate. Більшість речень перекладено і варіант Гугл перекладача відкинуто. Структурні перебудови, семантичні відповідники, велика кількість займенників, союзів та навіть артиклів перевантажили текст перекладу Google translate.

Література

- Ветров Д.* Латентная семантическая модель. URL: <https://postnauka.ru/video/49258>
- Ефремова Н.* Нейронные сети : практическое применение. 2017. URL: <https://habr.com/ru/post/322392/> doi:10.1038/nature.2016.20696
- Іваницька Н.* Повнозначне слово української мови в сучасних категорійних вимірах : монографія. Нілан-ЛТД, 2017. 266 с.
- Castelvecchi D.* Deep learning boosts Google Translate tool. 27 September 2016. URL: <https://www.scientificamerican.com/article/deep-learning-boosts-google-translate-tool/>
- Lewis-Kraus G.* The Great A. I. Awakening. Dec. 14, 2016. URL: <https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html>

Melvin Johnson et al. Google's Multilingual Neural Machine Translation System : Enabling Zero-Shot Translation. 2017. URL: <https://www.aclweb.org/anthology/Q17-1024/>
 SCHUSTER M. ET AL. Zero-shot translation with GOOGLE's multilingual neural machine translation system. November 22, 2016. Url: <https://ai.googleblog.com/2016/11/zero-shot-translation-with-googles.html>
 Surafel M. Lakew, Quintino F. Lotito, Matteo N., Turchi, M., Federico, M. Improving Zero-Shot Translation of Low-Resource Languages. 2017. URL: <https://arxiv.org/abs/1811.01389>

(Матеріал надійшов до редакції 12.02.19. Прийнято до друку 18.03.19)

УДК: 811.111'25'38

DOI: <https://doi.org/10.26661/2414-1135/2019-76-12>

ЛЕЩЕНКО Г. А., БЛАКИТА І. О.

(Національний університет «Запорізька Політехніка»)

ОСОБЛИВОСТІ ПЕРЕДАЧІ ЕЛЕМЕНТІВ ВЕРТИКАЛЬНОГО КОНТЕКСТУ МОВОЮ ПЕРЕКЛАДУ (НА МАТЕРІАЛІ ПЕРЕКЛАДУ ПОВІСТІ О. КОБИЛЯНСЬКОЇ «В НЕДІЛЮ РАНО ЗІЛЛЯ КОПАЛА»)

Стаття присвячена впливу вертикального контексту на повноту розуміння задуму автору тексту. У ході дослідження визначено саме поняття вертикального контексту, та як він впливає безпосередньо на текст та його інтерпретацію різними читачами. Проведено аналіз передачі елементів тексту з національними реаліями іншою мовою оскільки розуміння позатекстової інформації впливає на цей процес.

Ключові слова: вертикальний контекст, горизонтальний контекст, експлікація, калькування, транскодування.

Leshchenko H. A., Blakita I. O. Peculiarities of Reproducing the Elements of the Vertical Context in the Target Language (on the Materials of O.Kobylianska's Story *On Sunday Morning She Gathered Herbs*). The article intends to show how vertical context influences understanding what the author implied in the text. During the research, the concept of vertical context has been defined as well as how it directly influences the text and its interpretation by different readers. The analysis of reproducing the source text elements with national realia in the target language and the extent of understanding extra-textual information influences this process has been made.

The article encompasses both theoretical background of the topic in question and practical one by first delving into the very meaning of context in general and vertical context in particular. The article then proceeds to present the results of research that has been conducted by comparing the work of folklore by O. Kobylianska in original and a version translated into English by M. Skrypniuk. Kobylianska's work is abundant with various contextual and intertextual elements requiring the reader to actually pay close attention and possess a certain amount of background knowledge indirectly related to the said text. However, the conducted study proved national realias and various author's implicit messages to be lost during the translation process. The translation process itself offers a great deal of tools such as loan translation, transcoding and explication at the translator's disposal to make it easier for a foreign reader to understand the text however at the cost of neglecting the initial message of the original text, thus requiring footnotes for the reader explaining various peculiarities alien to their culture.

Key words: vertical context, horizontal context, explication, loan translation, transcoding.

Міжкультурна комунікація здійснюється не тільки через усне або письмове спілкування, але й за допомогою літературних творів, які створюються носіями одних й читаються носіями інших культур. І найчастіше нестача відомостей про місце й час описаних подій веде до неповного або навіть невірною розуміння того, що прагнув виразити автор. Тому для адекватного розуміння твору недостатньо одного лише знання мови – важливо знати в цілому обстановку місця й епохи, у яких створювався, й які описує текст. Ситуація ускладнюється тим, що автор літературного твору далеко не завжди звертає пильну увагу на вищезгадані особливості. Для позначення такої культурологічної, історичної й іншої інформації, яка не виражена в тексті явно, але потрібна для його повного розуміння, використовується термін «вертикальний контекст».

Актуальність теми вертикального контексту та його впливу на безпосередньо сприйняття тексту читачем не можна переоцінити. Вертикальний контекст відзначається широкими текстотворчими можливостями та впливом на правильність та повноту сприйняття тексту читачем. Культурологічний сегмент змісту літературного твору, появу якого генерують компоненти вертикального контексту, потребує комплексного лінгвістичного аналізу, що й зумовило вибір теми наукової статті.