

ПСИХОЛІНГВАЛЬНІ АСПЕКТИ ПЕРЕКЛАДУ СЛОВНИКА LIWC

Сергій Засєкін

sergiozasyekin@eenu.edu.ua

Наталія Безуглова

ibezuglova98@gmail.com

Анастасія Гапон

anastassi53@gmail.com

Владислав Матюшенко

matiushenko.vladyslav@gmail.com

Ольга Подольська

olhapodolska@gmail.com

Дар'я Зубчук

dariazubchuk@gmail.com

Східноєвропейський національний університет імені Лесі Українки, Україна

Received May 30, 2018; Revised June 8, 2018; Accepted June 14, 2018

Останні досягнення в галузі комп'ютерних лінгвістичних та психологічних інструментів дослідження актуалізують не лише потребу поширення сфер їхнього застосування, а й застосування з носіями різних мов. Тому актуальним видається вивчення особливостей перекладу на основі з'ясування культурних відмінностей. Дослідження сьогодення в галузі перекладознавства визнають критичну роль, що її відіграють перекладацькі трансформації як психолінгвальні логічні операції, які тривають у голові перекладача. Ця розвідка вивчає шляхи адекватного відтворення українською мовою вибраних слів зі словника програми Linguistic Inquiry and Word Count (LIWC) 2015 (Pennebaker et al., 2015) на основі встановлення їхніх семантичних, прагматичних та культурних аспектів значення. Найголовнішим висновком дослідження є те, що перекладачеві під час роботи слід удаватися до застосування низки перекладацьких трансформацій, як-от: диференціація, конкретизація, генералізація значення слова, а також транскодування, експлікація, калькування. Деякі мовні одиниці, наприклад, із категорії 'Leisure', 'Netspeak' не перекладалися взагалі, зберігаючи їхню оригінальну англійську форму, позаяк сьогодні носії англійської та української спільно вживають таку лексику в своєму мовленні. Перспективним видається дослідження способів перекладу словника LIWC на базі інших мов.

Keywords: *Linguistic Inquiry and Word Count (LIWC), перекладацькі трансформації, вихідна мова, цільова мова.*

Zasiekin, Serhii; Bezuglova, Nataliia; Hapon, Anastasiia; Matiushenko, Vladyslav; Podolska, Olha; Zubchuk, Darya. Psycholinguistic Aspects of Translating LIWC Dictionary.

Abstract. Recent developments in computerized linguistic and psychological research tools have heightened the need for broadening not only their areas of use, but also their applied languages. Therefore, there is an urgent need to address the translation problems caused by different target culture-bound issues. Existing research in the domain of translation studies

recognises the critical role played by translating transformations as psycholinguistic logical operations performed in the translator's mind. This paper explores the ways in which selected original Linguistic Inquiry and Word Count (LIWC) 2015 (Pennebaker et al., 2015) dictionary items can be reproduced in Ukrainian with due account of their semantic, pragmatic, and cultural load. The most obvious finding to emerge from this study is that the translator should apply a set of lexical translation transformations including differentiation, concretization, generalization of word meaning along with transcoding, explication, and calquing. Some linguistic items, e.g. from the category of 'Leisure', 'Netspeak' were not translated at all, i.e. retained their original form, due to their shared use today both by speakers of English and Ukrainian. Further research should be undertaken to investigate methods of reproducing LIWC vocabulary in other languages.

Keywords: *Linguistic Inquiry and Word Count (LIWC), translation transformations, source language, target language.*

1. Вступ

Ми можемо отримати багато інформації про людей, їхні особистості, переконання, страхи, спосіб мислення та соціальні відносини на основі аналізу шляхів використання ними слів повсякденному мовленні. Однією з ефективних сучасних методик психолінгвістичної діагностики емоційних, когнітивних та поведінкових компонентів психіки, відбитих у мовленні людей, є комп'ютерна програма *Linguistic Inquiry and Word Count (LIWC)*, автор якої – американський соціальний психолог Дж. Пеннебейкер. Ця програма засвідчила свою ефективність як надійного інструменту не лише лінгвістичного, а й перекладознавчого дослідження (Zasiiekin, Rosenhart, 2018). Перша версія LIWC була розроблена як частина дослідницького проекту вивчення кореляції мови й інформації. Друга (LIWC2001) і третя (LIWC2007) версії удосконалили оригінальну програму розширеним словником та сучаснішим дизайном програмного забезпечення (Pennebaker, Francis & Booth, 2001; Pennebaker, Booth & Francis, 2007).

Програмне забезпечення останньої версії LIWC 2015 має значно змінений словник та параметри програмного забезпечення (Pennebaker, Booth, Boyd, & Francis, 2015). Важливо те, що версія і програми, і словника 2015 року є новими, а не просто оновленою «декорацією» попередніх версій. Проте, як і попередні версії програми, вона складена таким чином, щоб швидко та ефективно аналізувати і окремі файли, і низку файлів у теці. Водночас, програма видається прозорою та гнучкою у своєму використанні, надаючи користувачам різні шляхи для дослідження лексики. Обидві версії програми, завантажувана та онлайн, спираються на внутрішній стандартний словник, який визначає, які слова слід фіксувати в цільових текстових файлах.

1.1. Основний модуль обробки текстів

LIWC2015 призначений для прийому письмового або транскрибованого тексту, що зберігався як цифровий файл, який в одному з декількох форматів, включаючи звичайний текст, PDF, RTF або стандартні файли Microsoft Word (наприклад, .doc та .docx). На відміну від попередніх версій, поточна версія може обробляти текст рядок за рядком у межах стовпців та між ними в кількох форматах електронних таблиць, включаючи ті, що зберігаються у форматах .xls, .xlsx та .csv. Кожен текстовий файл підлягає аналізу відносно 73 катего-

рій словника, як-от: аналітичне мислення, вплив, автентичність та емоційний тон; середня кількість слів у реченні, відсоток слів в тексті, довжина яких перевищує шість літер; стандартні лінгвістичні категорії, наприклад, відсоток у тексті функціональних слів – займенників, прийменників, артиклів та повнозначних слів допоміжних дієслів тощо. Крім того, словник уключає категорії слів, які стосуються психологічних станів, зокрема афектів, когнітивних процесів, біологічних процесів, мотивів; особистісні тематичні категорії, наприклад, робота, дім, дозвілля; соціолінгвальні маркери з повсякденного спілкування: лексичні одиниці зі значенням згоди, слова-паразити, нецензурна лексика, жаргон інтернет-користувачів.

1.2. Опис словника LIWC

Оригінальна версія словника LIWC2015 є центральною складовою стратегії аналізу тексту. За замовчуванням він складається майже з 6400 лем. Кожен словниковий запис додатково визначає одну або кілька категорій слів або словникових підтем. Наприклад, слово ‘cried’, через свою семантику, є складником п’яти підкатегорій: Sad, Negemo, Affect, Verb та FocusPast. Отже, якщо програма фіксує в тексті слово ‘cried’, воно ввійде до кожної із п’яти субкатегорій.

LIWC, як і всі сучасні комп’ютерні інструменти аналізу тексту, не позбавлений помилок під час виявлення та віднесення окремих слів до конкретної категорії. Наприклад, *mad* (божевільний) – слово, яке машина відносить до групи негативних емоцій (Negemo), що впливає на загальну оцінку тексту. Як правило, семантика *mad* містить гнів. Проте іноді *mad* може мати конотативні значення радості (*he’s mad for her*) або ж психічної нестабільності (*mad as a hatter*). Для програми LIWC такі речі не є серйозною проблемою, оскільки в основі програми – переваги ймовірнісних (статистичних) моделей використання мови. Так, у певному реченні слово “mad” може бути використане для вираження позитивних емоцій. Проте, якщо авторський текст «випромінює» позитив, програма спирається на відносно високі кількісні показники інших позитивно заряджених слів, попри наявність *mad*. Відтак, подібні до вищевказаної незначні класифікаційні помилки рідко впливають на висновки, позаяк перші компенсуються тим, які слова найчастіше використовує автор у тексті. Крім того, ці та подібні недоліки усуваються більшим за розміром текстом. Інакше кажучи, що більше слів у тексті, то надійнішими є результати. Тому, природно, результати аналізу тексту з 10 тис. слів дає набагато надійніші результати, ніж тексту обсягом 100 слів.

2. Методи дослідження

Належність англійської та української мов до різних гілок індоєвропейської родини мов (англійська – до германської, українська – до слов’янської), а також до різних структурних типів мов (англійська мова – аналітична, де граматичні відношення в реченні відтворюють граматичні морфеми, а українська мова – синтетична, що спирається на флексії), становить най-

більший комплекс граматичних проблем перекладу. У мові оригіналу (англійській) словника LIWC немає дієприслівників, категорії роду іменників і прикметників тощо, а в мові перекладу (українській) вони наявні, вимагаючи окремого перекладу в усіх формах. Натомість в українській відсутні артиклі, які в англійському мові сигналізують про особливості актуального членування інформації в реченні.

Слід також зазначити, що англійській мові властива граматична омонімія, тобто одні й ті само дієслова можуть бути смисловими або допоміжними, наприклад, “do”. Це теж викликає суттєві труднощі під час перекладу, адже потрібно обирати спосіб перекладу залежно від виконуваної словом функції.

Смислова, або семантична структура слова унікальна для кожної конкретної мови, а тому не завжди збігається в лексичних системах англійської мови (вихідної) та української (мови перекладу).

Під час перекладу подекуди неможливо скористатися буквальним словниковим відповідником, і перекладач удається до перетворення внутрішньої форми слова або ж його повної заміни, тобто до лексичної трансформації. Серед провідних лексичних трансформацій перекладознавці виокремлюють такі: диференціація, конкретизація, генералізація, транскодування (транскрибування і транслітерація), експлікація, модуляція (смисловий розвиток), описовий та антонімічний переклад (Максімов 2010; Карабан 2004).

Під час перекладу лексичних одиниць із значенням соціальних стосунків (Social) словника LIWC, застосовано здебільшого диференціацію, конкретизацію, генералізацію як лексичні заміни, а також транскрибування й транслітерацію. Диференціацію відзначає залежність від вибору перекладача на основі принципу синонімії. Наприклад, “champion” – «чемпіон», «переможець». Конкретизація – це процес, протилежний генералізації - заміна слова або словосполучення з широким предметно-логічним значенням на одиницю мови перекладу з більш вузьким значенням. У багатьох випадках застосування конкретизації пов’язане з тим, що у мові перекладу відсутнє слово з таким же широким значенням. Логічна операція, на якій ґрунтується ця трансформація – гіпонімія.

У результаті конкретизації одиниця мови-джерела, що виражає родові поняття, замінюється на одиницю мови-перекладу – вузьким, видовим поняттям, яке входить до складу данного роду. У багатьох випадках застосування конкретизації пов’язане з тим, що у мові перекладу відсутнє слово з таким же широким значенням. Наприклад, слово “dear” перекладали так: *любий, шановний, дорогий, милий*.

Генералізація як заміна мовної одиниці оригіналу, яка має більш вузьке значення, одиницею перекладу з ширшим значенням (Максімов, 2010:118), спирається на логічну операцію гіперонімії. Наприклад, “fiancee” наречена. В українській мові для позначення жінки, якій зробили пропозицію та жінки, яка виходить заміж використовують одне слово – наречена. Однак в англійській мові слово “fiancee” використовують лише стосовно жінки або дівчини, якій зробили пропозицію, в той час як у день весілля наречену називають “bride”.

Транскодування, яке охоплює транскрибування й транслітерацію, застосовують тоді, коли потрібно відтворити звукову та/або графічну форму слова вихідної мови. *Транскрибування і транслітерація* як форми транскодування, – способи перекладу лексичної одиниці оригіналу шляхом відтворення її форми за допомогою букв мови перекладу. Під час *транскрибування* відтворюється звукова форма іншомовного слова, а під час *транслітерації* – його графічна форма (буквений склад). Провідним способом у сучасній перекладацькій практиці є транскрипція зі збереженням деяких елементів транслітерації. Найчастіше такий спосіб перекладу застосовується для відтворення власних назв та імен, що не змінюють своєї форми в перекладній версії.

3. Обговорення результатів

У словнику LIWC зафіксовано велику кількість неологізмів, що ними послуговуються сучасні носії англійської початку ХХІ століття. Транскодування неологізмів відбувається в перекладі в тих випадках, коли в цільовій мові та культурі не ввійшло в обіг нове відповідне поняття і відповідний перекладний еквівалент, а перекладач не може підібрати слово або слова в мові перекладу, які б адекватно передавали зміст поняття і відповідали вимогам до термінотворення.

Особливо часто транскодування термінів відбувається в тих випадках, коли термін у мові перекладу складається з міжнародних терміноелементів латинського або давньогрецького походження.

Способом транскодування перекладено, наприклад, такі лексичні одиниці: *chat* – *чат*, *chorus* – *хорус*, *blog* – *блог*, *bowling* – *боулінг*, *cocktail-коктейль*, *cricket* – *крикет*, *football* – *футбол*, *croquet* – *крокет*, *frisbee* – *фрісбі*, *hip hop* – *гін-гон*, *hotel* – *готель*, *jazz* – *джаз*, *karaoke* – *караоке*, *marathon* – *марафон*, *liquor* – *лікер*, *motel* – *мотель*, *orchestra* – *оркестр*, *podcast* – *подкаст*, *selfie* – *селфі*, *soccer* – *сокер*, *symphony* – *симфонія*, *flirt* – *флірт*, *optimism* – *оптимізм* та ін. Навіть деякі слова-вигуки ми перекладали цим методом: *wow!* – *вау!*, *ha* – *ха*, *haha* – *ха-ха*, *hehe* – *хе-хе*, *heh* – *хех*.

Експлікація – лексико-граматична трансформація, під час якої лексична одиниця мови-оригіналу замінюється словосполученням, що дає пояснення або визначення цієї одиниці. Наприклад: *CD (compact-disc)* – *компакт-диск*; *e-cig (electronic cigarette)* – *електронна сигарета*; *esig (electronic signature)* – *електронний підпис*.

Слід зауважити, що більшість одиниць ВТ мала еквівалентний переклад, коли значення англійського слова цілковито збігалось зі значенням українського. Найбільше таких слів трапилося під час перекладу лексичних одиниць зі значенням афектів (Affect), хоча і при перекладі лексичних одиниць категорії дозвілля (Leisure) ми теж керувалися даним методом. Ось декілька прикладів еквівалентного перекладу: *attractive* – *привабливий*, *annoyed* – *роздратований*, *cheerful* – *життєрадісний*, *gorgeous* – *розкішний*, *hope* – *надія*, *hunting* – *полювання*, *jogging* – *пробіжка*, *solitaire* – *пасьянс*.

Додамо, що під час перекладу деяких лексичних одиниць були і неперекладні лексичні одиниці. Тут мається на увазі цілковите копіювання англомовної графічної форми слова. Не перекладали назви корпорацій, технологічні стандарти і назви програмних продуктів, та соціальних мереж, які за інших умов мали б перекладатися, наприклад: *Sony, Netflix, YouTube, Facebook, Instagram, Snapchat, Netflix, Nintendo, Spotify, Xbox*.

Водночас, провідним прийомом перекладу слів-термінів було використано добір лексичного еквіваленту – постійного лексичного відповідника, який точно збігається із значенням вихідного слова (Кияк, 2007). Проте лексичні одиниці мали бути подані у табличному форматі .xls, одне слово в одній комірці. Тому ми добирали найвлучніший відповідник згідно з вказаною тематикою лексичних одиниць. Наприклад, до лексичної одиниці *official* можна дібрати кілька варіантів перекладу: *офіційний, державний, службовий, посадовий, формальний*. Тому обмежувалися першими трьома, які розташовували в окремі вертикальні комірки.

Поширена була трансформація диференціації, яку характеризував добір слова, що не є словниковим відповідником, проте належно враховує контекст (маркований категорією LIWC, у межах якої здійснювали переклад) Основний принцип, який лежить в основі цієї трансформації – синонімія. Наведемо приклади застосування цієї трансформації: *Age – вік, літа, старість, час, період, епоха*. Як відповідник використано одну лексичну одиницю *вік*. *Approve – схвалювати, випробовувати, доводити, затверджувати, підтверджувати*. У перекладі дібрано відповідник *схвалювати*.

Під час виконання перекладу було також застосовано адаптацію. Загалом, цей тип трансформації часто застосовується при перекладі фразеологізмів, сленгу, усталених фраз, які передбачає суспільна традиція (Корунець, 2007). Приклади лексичних одиниць, переклад яких було здійснено за допомогою адаптації, наприклад:

underdog – невдах*, moron* – недоум*, kingpin – цабе, homeless* – безхатченк*, madam – пані, madame* – добродійк*.*

Антонімічний переклад – це лексична трансформація, при якій створюється заміна ствердної форми в оригіналі на негативну форму в перекладі або, навпаки, негативної на ствердну супроводжується заміною лексичної одиниці вихідної мови, на одиницю мови перекладу, з протилежним значенням. Характерне часте застосування у англо-українських перекладах, якщо негативна форма оригіналу слова утворена за допомогою префікса з негативним значенням (Максімов, 2010).

Прикладами застосування цієї трансформації є такі лексичні одиниці:

undo – скасо, minor – незначн*, infamous – ганебн*.*

Під час перекладу лексичних одиниць зі значенням біологічних процесів (Віо) словника LIWC з англійської мови на українську виникають труднощі, пов'язані з правильним вибором українського терміна або загальноновживаного слова. Це стається в тих випадках, коли в англійській мові для кількох понять існує один багатозначний термін (слово), а в українській мові йому відповідає

не один термін (слово), а декілька. Щоб уникнути таких помилок, слід уточнювати значення подібних слів у словниках, а також звертати особливу увагу на контекст (biology), в якому перебувають ці слова. Наприклад, *vital* в межах біологічної тематики має декілька значень: *суттєв**, *життєво необхідн**, *енергійн**, *важлив**, *фатальн**.

Вибір неуніверсального відповідника може суттєво відобразитись на точності перекладеного повідомлення. Оскільки в українській мові є відповідники, що мають вужче значення, в межах певного контексту вони можуть суперечити один одному, навіть протиставлятися. Відтак, метод генералізації під час перекладу слів із галузі біології, дає можливість досягти бажаного результату, хоча й не завжди існує збалансований відповідник. Тому більшу ефективність у таких випадках має метод транскодування, оскільки він дає можливість відтворити оригінальне лексичне значення слова в межах біологічного контексту. Тому, наприклад, окрім варіанта *суттєвий*, ми додали ще й *вітальний*.

Особливе місце посідають аббревіатури, зокрема ті, яким притаманний збіг зі скороченнями, поширеними в інших галузях. Оскільки структурна організація лексичного матеріалу словника полягала в розподілі слів за тематичними категоріями, для користувача оригінальної версії LIWC ця проблема вирішена. Проте від українського перекладача цей аспект потребує підвищеної пильності. Наприклад, аббревіатура *VD* має декілька значень, що задовольняють обраним темам: *venereal disease (STD is widely preferred)*; *vapor density*; *v.d. various dates*; *vasodilator*.

Попри те, що в мові оригіналу перевагу надано скороченню *STD (sexually transmitted disease)*, *VD* у значенні *venereal disease* має більшу вагу. Для якісного перекладу таких аббревіатур необхідно зважати на частотність їхнього вживання. З метою забезпечення прозорості при перекладі надавалась перевага розшифрованим експлікованим аналогам скорочень із подальшим адаптацією (*domestication*) й обмеженням одним словом, наприклад: *vd – венеричн* захворюван* – венеричн**; *isi – відділ інтенсивн* терапії – реанімаці**.

4. Висновки

Лексика, яку люди застосовують у повсякденному мовленні, виявляє їхні думки, почуття, особистісні риси та мотиви. Інакше кажучи, слова, які використовують у мовленні люди, мають високу психологічну цінність. Сучасні комп'ютерні технології дають змогу здійснювати психолінгвістичну діагностику текстів й робити узагальнення щодо авторських мотивів. Однією із найбільш точних програм, призначених для аналізу тексту є *Linguistic Inquiry and Word Count (LIWC)*, розроблена техаськими науковцями під керівництвом Дж. Пенебейкера. Розробка її української версії потребувала від перекладачів долання розбіжностей вихідних та цільових слів у лінгвокультурному плані. Незважаючи на те, що більшість слів були

загальновідомими та загальноновживаними, під час перекладу лексичних одиниць зі словника LIWC виникала потреба перетворення внутрішньої форми слова або ж його повної заміни, тобто застосування лексичних трансформації. Найпоширенішими серед них виявилися диференціація та генералізація значень поряд із транскодуванням – транскрипцією у комбінації з транслітерацією.

References

Література

- Корунець, І. В. Теорія та практика перекладу. Вінниця: Нова книга, 2007.
- Карабан, В. І. Переклад англійської наукової і технічної літератури [Текст]: у 2-х ч. Ч. 1., Ч. 2. Вінниця: Нова книга, 2004.
- Кияк Т. Р. Функції та переклад термінів у фахових текстах// Вісник Житомирського державного університету імені Івана Франка. № 32. С. 104-108.
- Максимов С. Є. Практичний курс перекладу. К.: Ленвіт, 2010.
- Pennebaker, J. W., Booth, R.J., Francis, M. E. (2007). *Linguistic Inquiry and Word Count (LIWC): LIWC2007*. Austin, TX: University of Texas at Austin.
- Pennebaker, J.W., Boyd, R.L., Jordan, K., & Blackburn, K. (2015). *The Development and Psychometric Properties of LIWC2015*. Austin, TX: University of Texas at Austin.
- Pennebaker, J. W., Francis ME, Booth RJ. (2001). *Linguistic Inquiry and Word Count (LIWC): LIWC2001*. Mahwah: Lawrence Erlbaum Associates.

References (translated and transliterated)

- Korunets', I. (2007). *Teoriya ta Praktyka Perekladu [Theory and Practice of Translation]*. Vinnytsia: Nova Knyha.
- Karaban, V. (2004). *Pereklad Angliyskoyi Naukovoyi i Tekhnichnoyi Literatury [Translation of English Scientific and Technical Literature]*. Volumes 1-2. Vinnytsia: Nova Knyha.
- Kyiak, T. (2007). *Funktsii ta pereklas terminiv u fakhovykh tekstakh [Functions and translation of terms in specialized texts]*. *Visnyk of Zhytomyr State University*, 32, 104-108.
- Maksimov, S. (2010). *Praktychnyi Kurs Perekladu [Practical Course of Translation]*. Kyiv: Lenvit.
- Pennebaker, J. W., Booth, R.J., Francis, M. E. (2007). *Linguistic Inquiry and Word Count (LIWC): LIWC2007*. Austin, TX: University of Texas at Austin.
- Pennebaker, J.W., Boyd, R.L., Jordan, K., & Blackburn, K. (2015). *The Development and Psychometric Properties of LIWC2015*. Austin, TX: University of Texas at Austin.
- Pennebaker, J. W., Francis ME, Booth RJ. (2001). *Linguistic Inquiry and Word Count (LIWC): LIWC2001*. Mahwah: Lawrence Erlbaum Associates.