

УДК 81'25:81'322.4:81'373.43:004
 DOI <https://doi.org/10.24919/2308-4863/98-1-40>

Микита БОНДАР,
 orcid.org/0009-0002-3386-6233
 аспірант кафедри прикладної лінгвістики та перекладознавства
 Українського державного університету імені Михайла Драгоманова
 (Київ, Україна) m.o.bondar@udu.edu.ua

ЕМПІРИКО-КОГНІТИВНА МОДЕЛЬ ВІДТВОРЕННЯ АНГЛОМОВНОЇ ІННОВАЦІЙНОЇ ІТ-ЛЕКСИКИ: ВІД КОРПУСНОГО АНАЛІЗУ ДО АЛГОРИТМІЗАЦІЇ РІШЕНЬ

У сучасних умовах стрімкого розвитку інформаційно-комунікаційних технологій відтворення інноваційної лексики та вузькоспеціалізованих неологізмів є одним із найскладніших викликів фахової комунікації. Актуальність проблеми зумовлена виникненням системного лексикографічного дефіциту, нездатністю традиційної лексикографії синхронно фіксувати нові концепти та нагальною потребою в подоланні термінологічної ентропії, що виникає внаслідок використання суб'єктивного інтуїтивного перекладу чи недосконалих систем нейромережевого машинного перекладу (далі – NMT), схильних до генерації дослівних кальок та алгоритмічних «галюцинацій».

Метою статті є розроблення, емпіричне обґрунтування та репрезентація емпірико-когнітивної моделі відтворення інноваційної ІТ-термінології, а також окреслення ефективних підходів і перспектив її застосування в освітньому процесі та технологічній сфері.

У дослідженні застосовано емпіричні методи корпусного аналізу, автоматизованої обробки природної мови (NLP), а також теоретичні підходи фреймової семантики та теорії qualia-структур. Результати дослідження полягають у розробленні комплексної алгоритмічної архітектури емпірико-когнітивної моделі та обґрунтуванні її ролі в оптимізації процесу відтворення. З'ясовано, що функціонування моделі є багатоступеневим і охоплює етап екстракції «незарєєстрованої лексики» за допомогою спеціальних Python-скриптів (виокремлено цільовий масив із 255 інноваційних термінів на матеріалі паралельного англо-українського корпусу); етап об'єктивізації лексикографічного статусу неологізмів шляхом частотного аналізу за базу ГРАК-19а та автоматизованої фільтрації за словником ВЕСУМ; а також критично важливий етап когнітивного моделювання, який передбачає виділення еталонних фреймів та телічних і агентивних qualia-ролей досліджуваних одиниць. Аналіз основних механізмів генерації перекладацьких рішень засвідчив ефективність застосування системи із семи детермінованих трансформційних правил (R1–R7), вибір яких диктується когнітивними параметрами концепту. Доведено, що використання такого вбудованого когнітивного пре-фільтра для семантико-синтаксичної верифікації на прикладі репрезентативних кейсів (autocompletion, supersampling, worker, peter-panning) сприяє успішному усуненню синтаксичних конфліктів та гарантує збереження функціонального призначення спеціалізованої лексеми. Окреслено практичні шляхи інтеграції розробленої архітектури у процеси фахової підготовки перекладачів та в системи автоматизованого перекладу через механізм «двопрохідного перекладу» (Two-Pass Translation) із використанням непорушних семантичних «якорів».

У висновках наголошено на доцільності системного впровадження алгоритмізованих когнітивно-дискурсивних підходів з метою створення умов для безвартного перенесення функціонального ядра концептів, подолання термінологічної неконсистентності та технологічної оптимізації сучасних інструментів обробки природної мови.

Ключові слова: емпірико-когнітивна модель, інноваційна ІТ-лексика, qualia-структура, машинний переклад, алгоритмізація перекладу, корпусний аналіз.

Mykyta BONDAR,
 orcid.org/0009-0002-3386-6233
 postgraduate student at the Faculty of Foreign Philology
 Mykhailo Dragomanov Ukrainian State University
 (Kyiv, Ukraine) m.o.bondar@udu.edu.ua

EMPIRICAL-COGNITIVE MODEL OF RENDERING ENGLISH INNOVATIVE IT VOCABULARY: FROM CORPUS ANALYSIS TO DECISION ALGORITHMIZATION

In the modern conditions of rapid information and communication technology development, rendering innovative vocabulary and highly specialized neologisms is one of the most complex challenges in professional communication. The relevance of the problem is conditioned by the emergence of a systemic lexicographic deficit, the inability of traditional lexicography to synchronously record new concepts, and the urgent need to overcome the terminological entropy resulting

from the use of subjective intuitive translation or imperfect neural machine translation (hereinafter – NMT) systems, which are prone to generating literal calques and algorithmic “hallucinations”.

The aim of the article is to develop, empirically substantiate, and present an empirical-cognitive model for rendering innovative IT terminology, as well as to outline effective approaches and prospects for its application in the educational process and the technological sphere.

The study employs empirical methods of corpus analysis and natural language processing (NLP), alongside theoretical approaches of frame semantics and qualia structure theory. The results of the study consist in developing a comprehensive algorithmic architecture for the empirical-cognitive model and substantiating its role in optimizing the rendering process. It was found that the functioning of the model is multi-staged and encompasses the stage of extracting “unregistered vocabulary” using custom Python scripts (a target dataset of 255 innovative terms was isolated based on an English-Ukrainian parallel corpus); the stage of objectifying the lexicographic status of neologisms through frequency analysis using the GRAC-19a database and automated filtering against the VESUM dictionary; and the critically important stage of cognitive modeling, which involves extracting reference frames along with the telic and agentive qualia roles of the studied units. The analysis of the main mechanisms for generating translation decisions confirmed the effectiveness of applying a system of seven deterministic transformational rules (R1–R7), the choice of which is dictated by the cognitive parameters of the concept. It is proven that using such a built-in cognitive pre-filter for semantic-syntactic verification, as demonstrated by representative cases (autocompletion, supersampling, worker, peter-panning), contributes to successfully eliminating syntactic conflicts and guarantees the preservation of the specialized lexeme's functional purpose. Practical ways of integrating the developed architecture into the professional training of translators and into automated translation systems via the “Two-Pass Translation” mechanism using immutable semantic “anchors” are outlined.

The conclusions emphasize the expediency of the systemic implementation of algorithmized cognitive-discursive approaches in order to create conditions for the lossless transfer of the concepts' functional core, overcome terminological inconsistency, and technologically optimize modern natural language processing tools.

Key words: *empirical-cognitive model, innovative IT vocabulary, qualia structure, machine translation, translation algorithmization, corpus analysis.*

Постановка проблеми. Стрімкий розвиток інформаційно-комунікаційних технологій у сучасному глобалізованому суспільстві супроводжується безперервним формуванням значних масивів інноваційної лексики та вузькоспеціалізованих неологізмів. Традиційна академічна лексикографія та статичні термінологічні бази об'єктивно не здатні забезпечити синхронну фіксацію нових концептів, що спричиняє системний лексикографічний дефіцит у фаховому дискурсі. В умовах відсутності сталих нормативних еквівалентів перекладачі змушені покладатися переважно на власні когнітивні ресурси та професійну інтуїцію під час ухвалення перекладацьких рішень. Це нерідко призводить до термінологічної неконсистентності, руйнації вихідних смислових структур та виникнення семантичної ентропії в цільових текстах, що ускладнює процеси спеціалізованої міжмовної комунікації. З іншого боку, масове впровадження систем нейромережевого машинного перекладу (NMT) та архітектур на базі великих мовних моделей (LLM) не вирішує зазначеної проблеми повною мірою. Сучасні алгоритми автоматизованого перекладу функціонують здебільшого на основі обчислення статистичної ймовірності дистрибуції tokenів і не володіють механізмами глибинного когнітивного аналізу ситуативного контексту. Під час опрацювання нестабільних лексем ІТ-галузі чи одиниць із метафоричним перенесенням системи машинного перекладу часто утворюють дослівні

кальки, що порушують граматичну й синтаксичну логіку цільового речення, або генерують семантично нерелевантні алгоритмічні «галюцинації». Окреслені недоліки зумовлюють потребу в концептуальному перегляді традиційних практик локалізації та створенні нових, об'єктивних алгоритмізованих підходів, здатних імітувати когнітивні процеси фахівця за допомогою сучасних обчислювальних методів. Інтеграція когнітивного елемента в методологію перекладу, що спирається на структурований корпусний аналіз та об'єктивне моделювання фреймових структур концепту, створює передумови для формалізації перекладацьких стратегій. Усе це підтверджує своєчасність і високу науково-практичну значущість порушеної проблеми, розв'язання якої є необхідним як для вдосконалення компетентностей майбутніх фахівців-перекладачів, так і для оптимізації архітектур автоматизованого оброблення природної мови в умовах нових технологічних викликів.

Аналіз досліджень. Проблематика оптимізації перекладацького процесу шляхом поєднання когнітивних теорій та інструментарію автоматизованої обробки природної мови перебуває в центрі уваги багатьох сучасних науковців. Аналіз фахової літератури свідчить про два паралельні вектори досліджень: вивчення когнітивної динаміки перекладача та вдосконалення алгоритмів машинного перекладу. Зокрема, І. Бойко та В. Борщовецька доводять ефективність когнітивно-дискурсивних підходів, наголошуючи, що переробка

лексичної інформації має спиратися на глибинне розуміння концептуальних структур, а не на механічну міжмовну заміну (Бойко, 2021; Борщовецька, 2022). Важливість психолінгвістичного аналізу цих когнітивних аспектів підтверджують Н. Юхан та співавтори (Yuhan та ін., 2024). О. Ребрій і В. Демецька досліджують механізми адаптації, асоціації та аналогії як фундаментальні фактори інтуїтивного ухвалення перекладацьких рішень (Ребрій, Демецька, 2020). Розширюючи розуміння людського мислення, М. Карл моделює процеси перекладу через призму активного логічного висновування та досліджує часову динаміку емоційних і когнітивних станів фахівця (Carl, 2024; Carl, 2025). Водночас значний масив праць присвячено інтеграції когнітивних стратегій у нейромережевий переклад. З. Хе разом із колегами та С. Лю аналізують потенціал великих мовних моделей у відтворенні людиноподібних стратегій трансляції (He та ін., 2024; Liu та ін., 2025). На необхідності застосування концепторієнтованих підказок (prompting) для створення людиноцентричного машинного перекладу наголошують М. Цянь і Ц. Конг (Qian, Kong, 2024). Г. Сі та С. Янг переконують у безальтернативності колаборативних підходів (human-in-the-loop), де експертний когнітивний контроль компенсує алгоритмічні недоліки нейромережових систем (Si, 2024; Yang та ін., 2023). Важливу роль ситуативного контексту як вбудованої моделі для когнітивного розвитку підкреслюють Б. Халаф та співавтори (Khalaf та ін., 2022). Отже, попри ґрунтовне теоретичне опрацювання когнітивних аспектів перекладу та стрімкий розвиток інформаційних технологій, у науковому дискурсі залишається значною мірою нерозв'язаним питання алгоритмізації роботи з інноваційною лексикою. Очевидною є прогалина у створенні цілісних, формалізованих моделей, які б органічно поєднували автоматизований корпусний аналіз із когнітивною теорією (зокрема, фреймовою семантикою та теорією qualia-структур) для оптимізації ухвалення рішень. Відсутність інтегрованих інструментів, здатних об'єктивно моделювати синтаксичні валентності під час обробки неологізмів, актуалізує потребу в розробці комплексної емпірико-когнітивної моделі, яка б дозволила подолати обмеження як суб'єктивного інтуїтивного перекладу, так і суто статистичної машинної трансляції.

Мета статті – розроблення емпіричне обґрунтування та репрезентація емпірико-когнітивної моделі відтворення англомовної інноваційної ІТ-термінології із застосуванням алгоритмо-корпусних підходів, а також окреслити перспективи її засто-

сування. Відповідно до поставленої мети передбачено виконання таких завдань: 1) обґрунтувати методологію автоматизованої екстракції «незарєєстрованої лексики» та її когнітивних валентностей на матеріалі паралельних корпусів; 2) представити систему трансформаційних перекладацьких правил (R1–R7) як ключовий складник алгоритмічної архітектури моделі; 3) продемонструвати дієвість моделі шляхом аналізу репрезентативних лексичних кейсів із фокусом на збереженні еталонних qualia-структур; 4) окреслити шляхи практичної імплементації запропонованої моделі в процеси фахової підготовки перекладачів та в архітектуру сучасних систем машинного перекладу.

Виклад основного матеріалу. Для розв'язання окреслених завдань та побудови емпірико-когнітивної моделі відтворення англомовної інноваційної ІТ-лексики було застосовано комплексну методологію, що поєднує інструментарій автоматизованої обробки природної мови (NLP) із фундаментальними принципами фреймової семантики. На першому етапі дослідження було сформовано репрезентативний паралельний англо-український корпус текстів, матеріалом для якого слугували файли локалізації технічної документації та програмного забезпечення від провідних технологічних проєктів. До корпусу було включено 7 текстів ІТ-галузі (Comprehensive Rust, б. д.; Kirigami KDE, б. д.; Kuhn, б. д.; Linietsky та ін., б. д.; Oualline, б. д.; Pradeep, б. д.; Tougher, б. д.). За допомогою Python було здійснено автоматизовану обробку вихідних файлів, їх очищення від технічної розмітки та пореченнєве вирівнювання (alignment), що дозволило зберегти природне синтаксичне оточення кожної лексичної одиниці. Наступний крок передбачав ізоляцію так званої «незарєєстрованої лексики» – термінів та неологізмів, які де-факто ще не зафіксовані в нормативних академічних словниках. Для цього українську частину корпусу було піддано токенизації та лематизації з використанням NLP-моделі «uk_core_news_sm» бібліотеки «spaCy». Отриманий масив унікальних лем пройшов автоматизовану процедуру фільтрації: алгоритм послідовно звіряв кожну одиницю з реєстром великого електронного словника української мови (ВЕСУМ). Усі лексеми, знайдені в словнику були визначені як загальноживана лексика та, відповідно, автоматично відсікалися, у результаті чого було об'єктивно виокремлено цільовий масив із 255 унікальних інноваційних ІТ-термінів, які потребували подальшого когнітивно-перекладацького аналізу.

З метою об'єктивізації статусу виявлених неологізмів було застосовано метод частотного ана-

лізу. Відфільтровані одиниці було автоматизовано зіставлено з Генеральним регіонально анотованим корпусом української мови (ГРАК-19а). На основі отриманих кількісних даних було встановлено операційний поріг частотності: лексеми з показником входжень ≤ 50 кваліфікувалися як інноваційні або нестабільні концепти, тоді як одиниці з показником понад 50 визначалися як відносно сталі еквіваленти. Ця метрика стала основою для подальшої алгоритмізації перекладацьких рішень та вибору стратегій відтворення. Визначальним когнітивним елементом запропонованої методології стала автоматизована екстракція еталонних фреймів та qualia-структур досліджуваних термінів. Із застосуванням інструментів синтаксичної обробки (Dependency Parsing) з англomовного контексту виокремлювалися ключові дієслова-маркери, якими керує досліджуваний термін або які керують ним. Це дозволило ідентифікувати формальні, телічні (функціональне призначення об'єкта) та агентивні (спосіб виникнення або специфіка маніпуляції з об'єктом) qualia-ролі кожної одиниці. Таким чином, фокус перекладацького аналізу було зміщено з ізольованого слова на його синтаксичну валентність та функціональне навантаження в реченні, що створило надійну матрицю для обчислення семантичної еквівалентності та розробки системи перекладацьких правил.

На основі інтеграції результатів автоматизованого корпусного аналізу та когнітивного моделю-

вання було розроблено цілісну архітектуру емпірико-когнітивної моделі відтворення англomовної інноваційної ІТ-лексики. Запропонована модель функціонує як багатоступеневий алгоритмічний контур, що імітує логіку ухвалення рішень фахівцем. Процес розпочинається з ідентифікації оригінального терміна та екстракції його еталонного фрейму й qualia-ролей на матеріалі англomовного корпусу. Далі алгоритм здійснює перевірку частотності концепту в цільовій мові, на основі якої визначається його лексикографічний статус (сталій еквівалент або нестабільна інновація). Залежно від визначеного статусу та специфіки еталонного фрейму застосовується етап генерації перекладацьких кандидатів за системою детермінованих трансформаційних правил. Критично важливим компонентом моделі є етап семантико-синтаксичної верифікації, який виконує функцію когнітивного пре-фільтра. На цьому етапі згенерований кандидат зіставляється з еталонним фреймом для перевірки збереження телічної та агентивної валентностей: якщо запропонований варіант вступає в синтаксичний конфлікт із цільовими дієсловами-маркерами, алгоритм блокує механічну заміну та циклічно повертається до етапу генерації за альтернативним правилом. У разі успішної валідації верифікований концепт фіксується як фінальний переклад. Логіку функціонування та структурні взаємозв'язки розробленої моделі схематично відображено на рисунку 1.

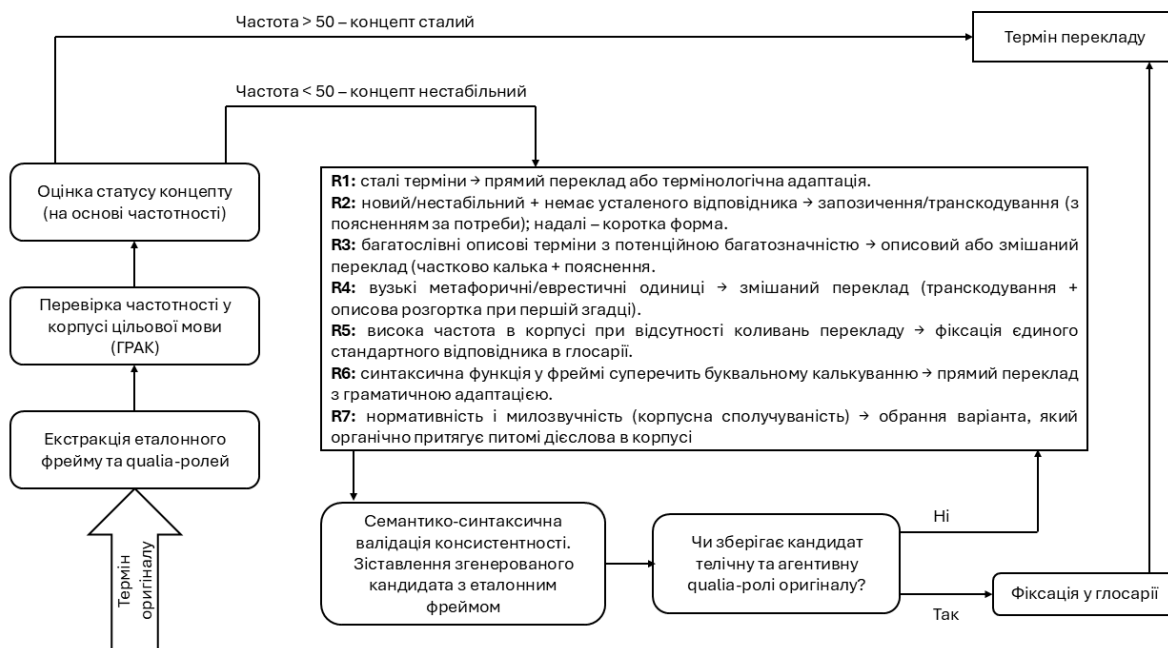


Рис. 1. Блок-схема емпірико-когнітивної моделі перекладу

Як було зазначено вище, ядром етапу генерації кандидатів є розроблена на основі емпіричних спостережень система трансформаційних приписів, що диктує вибір перекладацької стратегії залежно від когнітивних параметрів терміна. До цієї системи входять такі правила:

– правило R1 (сталі концепти): застосування прямого перекладу або прямого перекладу із термінологічною адаптацією;

– правило R2 (нестабільні інновації без усталеного відповідника): використання запозичення або транслітерації під час першої згадки (з обов'язковим поясненням у дужках) та перехід до короткої форми в подальшому тексті;

– правило R3 (багатослівні описові терміни з потенційною багатозначністю): застосування описового перекладу або змішаної стратегії, що поєднує часткову кальку з експлікацією;

– правило R4 (одиниці з глибоким метафоричним або евристичним перенесенням): змішаний переклад з транскодуванням та описовою розгорткою для збереження нетипового візуального чи асоціативного фрейму оригіналу;

– правило R5 (високочастотні одиниці за відсутності перекладацьких коливань): жорстка фіксація єдиного стандартного відповідника в глосарії з маркуванням інших варіантів як периферійних синонімів;

– правило R6 (синтаксичний конфлікт у цільовій мові): застосування прямого перекладу з обов'язковою граматичною адаптацією (зміна порядку слів або конверсія частини мови) для збереження органічної валентності цільових дієслів;

– правило R7 (нормативність та корпусна сполучуваність): превентивне уникнення неприродних кальок шляхом обрання варіанта, який об'єктивно підтверджує свою здатність утворювати органічні синтаксичні зв'язки в україномовному корпусі.

Для верифікації розробленої моделі та демонстрації механізмів її функціонування було здійснено зіставний аналіз 255 виокремлених лексичних інновацій. У межах цієї статті доцільно детально розглянути чотири репрезентативні кейси, які ілюструють застосування різних трансформаційних правил залежно від когнітивно-синтаксичних параметрів концепту (див. табл. 1).

Таблиця 1

Ілюстрація репрезентативних лексичних кейсів із застосуванням правил емпірико-когнітивної моделі

Застосоване правило	Термін оригіналу	Телічні та агентивні qualia-ролі оригіналу	Термін перекладу	Телічні та агентивні qualia-ролі перекладу
R1 – прямий переклад	autocompletion	complete, suggest, assist, enable	автодоповнення	доповнювати, забезпечувати, налаштувати
R2 – транскодування	supersampling	Smooth, antialias, calculate, perform	суперсемплінг	Згладжувати, Увімкнути, налаштувати, рендерити
R3 – часткове калькування	screen-space ambient occlusion (ssao)	darken, simulate, calculate, render	екранно-просторова амбієтна оклюзія (ssao)	затінювати, рендерити, використовувати, порівняти
R4 – змішаний переклад	peter-panning	detach, float, cause, decrease	пітер-панінг	від'єднуватися, спотворювати, спричиняти, встановити
R5 – фіксація єдиного варіанту	singleton	build, enable, access, add	синглтон (у корпусі також «сінглтон»)	дозволяти, реагувати, переглянути, використовувати
R6 – граматична адаптація	worker	Cache, load, configure, run	воркер	Кешувати, завантажувати, встановлювати, працювати
R7 – морфологічна адаптація	committed	save, track, ensure, version	закомітити	зберігати, фіксувати, додати, налаштувати

Джерело: сформовано автором.

Кейс правила R1 демонструє роботу алгоритму зі сталими термінологічними одиницями на прикладі концепту *autocompletion*. Корпусний аналіз англомовного фрейму «Text_creation, Assistance» виявив чіткі телічні та агентивні *qualia*-ролі, марковані дієсловами «complete», «suggest», «assist», «trigger», «enable». Зважаючи на високу частотність відповідника в українському корпусі (понад 50 входжень), було виділено використання правила R1 (прямий переклад), утворивши «автодоповнення». Етап семантико-синтаксичної валідації підтвердив абсолютну когнітивну еквівалентність: український термін органічно заповнює еталонні валентності, керуючи питомими предикатами доповнювати, забезпечувати, налаштувати в межах цільового фрейму. У разі обробки низькочастотних інновацій (кейс правила R2) було виділено інший підхід інший підхід, що ілюструється концептом «supersampling» (технологія згладжування в комп'ютерній графіці). Оскільки ГРАК-19а зафіксував низьку частотність цього неологізму, одиницю було класифіковано як нестабільний концепт. За відсутності усталеного перекладу, здатного передати специфічну телічну *qualia*-роль («render», «scale», «smooth») було використано правило R2. Відповідно до нього, було обрано стратегію транскодування («суперсемплінг»), яка дозволяє перенести інновацію в цільовий текст без руйнації її технічного фрейму, уможливаючи подальше обростання терміна українськими агентивними дієсловами (увімкнути, застосувати суперсемплінг).

Кейс правила R3 демонструє дію алгоритму під час опрацювання багатослівних описових термінів зі складною когнітивною структурою на прикладі концепту «screen-space ambient occlusion (ssao)». Аналіз англомовного фрейму виокремлює специфічні телічні та агентивні *qualia*-ролі, марковані предикатами «darken», «simulate», «calculate», «render». Зважаючи на багатокомпонентність та технічну специфіку одиниці, модель ініціювала правило R3 (змішана стратегія). У результаті було згенеровано варіант «екранно-просторова амбієнтна оклюзія (ssao)», у якому часткове калькування («екранно-просторова») поєднано з транслітерованим запозиченням. Цей підхід дозволив алгоритму успішно пройти етап синтаксичної валідації, зберігши здатність терміна органічно заповнювати цільові валентності («затінювати», «рендерити», «використовувати», «порівняти»), тоді як спроба застосувати питомі відповідники (наприклад, «навколишнє перекриття») призвела б до втрати фахової точності.

Кейс правила R4 репрезентує специфіку перекладу одиниць із глибоким метафоричним пере-

несенням на прикладі концепту «peter-panning» (графічний артефакт, коли тінь відривається від об'єкта). Обробка таких евристичних неологізмів вимагає збереження нетипового візуально-асоціативного фрейму. Спроба розкласти термін на складові («відривання тіні») руйнує авторську аналогію та конститутивну *qualia*-роль концепту. Згідно з розробленою матрицею, було застосовано правило R4 – змішаний переклад («пітер-панінг»), що передбачає транслітерацію оригінальної метафори з подальшою описовою розгорткою в тексті. Наведені результати емпіричної апробації переконливо доводять, що формалізований аналіз *qualia*-структур є дієвим інструментом, який дозволяє перетворити абстрактну категорію перекладацької адекватності на вимірюваний параметр синтаксичної сполучуваності.

Для усунення термінологічної ентропії під час роботи з високочастотними одиницями модель застосовує прескриптивний підхід, що ілюструється кейсом концепту «singleton» правила R5. У вихідному дискурсі цей термін репрезентує фундаментальний патерн проектування (еталонні *qualia*-ролі: «build», «enable», «access», «add»). Оскільки корпусні дані цільової мови часто фіксують перекладацькі коливання та орфографічні девіації (наприклад, паралельне вживання «синглтон» та «сінглтон»), алгоритм активує правило R5. Воно передбачає жорстку фіксацію єдиного нормативного відповідника «сінглтон» як сталого еквівалента. Це автоматично відтинає периферійні синоніми та гарантує консистентність цільового фрейму, у якому концепт стабільно керує питомими предикатами (дозволяти, реагувати, переглянути, використовувати).

Особливої уваги заслуговує кейс правила R6 – концепт «worker» (у контексті веб-програмування, Service Worker), який яскраво демонструє функцію когнітивного пре-фільтра (правило R6). У вихідному ІТ-дискурсі цей термін репрезентує неживий програмний механізм, телічна та агентивна ролі якого маркуються дієсловами «cache», «fetch», «execute», «run». Первинна спроба застосувати дослівний словниковий еквівалент «робітник» або «працівник» не пройшла етап валідації, оскільки в українській мові ці лексеми мають жорсткий антропоморфний фрейм і вступають у синтаксичний конфлікт із технічними предикатами (конструкція «робітник кешує файли» є семантичною хибою). Відтак, замість калькування було використано правило синтаксичної адаптації, згенерувавши запозичення «воркер». Цей варіант успішно пройшов валідацію, оскільки здатний безперешкодно утворювати органічні валентні зв'язки з

маркерними дієсловами цільової мови (запускати воркер, воркер кешує), зберігаючи функціональне ядро оригіналу.

Кейс R7 (нормативність та корпусна сполучуваність) репрезентує кейс морфологічної адаптації технічного процесу, маркованого дієсловом «committed». Оригінальні телічні та агентивні валентності цього концепту («save», «track», «ensure», «version») вказують на специфічну дію в системах контролю версій. Спроба механічного перекладу за допомогою описових конструкцій (наприклад, «здійснити фіксацію змін») порушує природну динаміку фахового IT-дискурсу. Для превентивного уникнення таких неприродних кальок когнітивний пре-фільтр згенерував адаптоване запозичення «закомітити». Цей варіант об'єктивно підтверджує свою здатність утворювати органічні синтаксичні зв'язки в україномовному середовищі, ідеально переймаючи на себе необхідне функціональне навантаження (зберігати, фіксувати, додати, налаштувати) без втрати семантичного ядра оригіналу.

Практична значущість запропонованої емпірико-когнітивної моделі не обмежується суто теоретичним обґрунтуванням перекладацьких трансформацій, а відкриває широкі перспективи для її імплементації в освітній процес та професійну діяльність. У дидактичному вимірі інтеграція такого підходу в підготовку майбутніх фахівців-перекладачів дозволяє здійснити концептуальний перехід від інтуїтивного пошуку словникових відповідників до структурованого аналізу синтаксичних валентностей та фреймів. Навчання студентів екстракції qualia-ролей формує навички об'єктивного оцінювання того, чи здатне обране цільове слово виконати ту саму функцію в реченні, що й оригінальний термін. Крім того, алгоритмічний підхід моделі закладає методологічну базу для укладання інноваційних динамічних глосаріїв, у яких вузькоспеціалізований неологізм фіксується не ізольовано, а в комплексі з його органічною синтаксичною рамкою (маркерними дієсловами).

У технологічному вимірі розроблена модель створює концептуальне підґрунтя для якісної оптимізації сучасних систем автоматизованого перекладу. Насамперед, застосовані Python-алгоритми та NLP-парсери можуть бути масштабовані для безперервного моніторингу фахових текстів із метою автоматичного виявлення нової «незарєєстрованої лексики» шляхом її постійного зіставлення з наявними лексикографічними базами. Для безпосереднього покращення якості нейромережевої генерації тексту пропонується

імплементація архітектури «двопрохідного перекладу» (Two-Pass Translation). Відповідно до цієї концепції, на першому етапі розроблена модель функціонує як інтелектуальний когнітивний пре-фільтр: вона ізолює інноваційні терміни у вихідному масиві, аналізує їхні фрейми, генерує оптимальні цільові еквіваленти за правилами R1–R7 та верифікує їхню синтаксичну сумісність. На другому етапі ці жорстко зафіксовані та перевірені терміни передаються базовій системі машинного перекладу (NMT або LLM) як непорушні семантичні «якорі». Базова нейронна мережа вибудовує граматичну структуру цільового речення навколо цих якорів, що зменшує вірогідність генерування дослівних кальок та семантичних «галюцинацій», гарантуючи збереження функціонального ядра концепту.

Висновки. У результаті проведеного дослідження було розроблено, емпірично обґрунтовано та репрезентовано емпірико-когнітивну модель відтворення англійської мови інноваційної IT-термінології із застосуванням алгоритмо-корпусних підходів. Обґрунтовано комплексну методологію автоматизованої екстракції «незарєєстрованої лексики» на матеріалі англо-українських паралельних корпусів. Доведено, що поєднання інструментарію автоматизованої обробки природної мови (NLP) із фреймовою семантикою дозволяє об'єктивно виокремлювати інноваційні концепти та визначати їхні синтаксичні й когнітивні валентності через екстракцію qualia-ролей. Представлено систему трансформаційних перекладацьких правил (R1–R7). Розроблений контур правил формалізує теорію qualia-структур, перетворюючи абстрактну категорію перекладацької адекватності на детермінований алгоритм. Емпірично продемонстровано дієвість створеної моделі шляхом аналізу репрезентативних лексичних кейсів. Підтверджено, що фокус на збереженні еталонних qualia-структур під час застосування правил дозволяє алгоритмічно усувати синтаксичні конфлікти та забезпечує безвратне перенесення функціонального призначення спеціалізованих концептів у цільовий дискурс. Окреслено ефективні шляхи практичної імплементації моделі. У процесі фахової підготовки перекладачів запропонований підхід сприяє концептуальному переходу від інтуїтивного пошуку словникових еквівалентів до об'єктивного аналізу валентностей. У технологічному вимірі архітектура моделі створює підґрунтя для оптимізації сучасних систем машинного перекладу шляхом її використання як інтелектуального пре-фільтра («двопрохідний переклад»), що запобігає виникненню алгорит-

мічних «галюцинацій». Перспективи подальших досліджень убачаємо у масштабуванні розробленого алгоритмічного контуру на інші галузі спеціалізованого дискурсу, а також у практичній програмній інтеграції представленої моделі в сучасні архітектури великих мовних моделей.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Бойко І. В. Когнітивно-дискурсивні моделі перекладу. *Вісник Луганського національного університету імені Тараса Шевченка. Філологічні науки*. 2021. № 7 (345). С. 135–144. URL: [https://doi.org/10.12958/2227-2844-2021-7\(345\)-135-144](https://doi.org/10.12958/2227-2844-2021-7(345)-135-144)
2. Борщовецька В. Д. Когнітивний підхід до процесу перекладу та переробки лексичної інформації: Сучасні підходи, виклики і перспективи. *Міжнародний філологічний часопис*. 2022. Т. 13, № 1. С. 41–49. URL: <https://doi.org/10.31548/philolog2022.01.041>
3. Ребрій О., Демецька В. Адаптація, асоціація та аналогія: три важливих механізми прийняття рішень в перекладі. *East European Journal of Psycholinguistics*. 2020. Т. 7, № 2. С. 231–242. URL: <https://doi.org/10.29038/ejpl.2020.7.2.reb>
4. Carl M. An Active Inference Agent for Modeling Human Translation Processes. *Entropy*. 2024. Vol. 26, № 8. P. 616. URL: <https://doi.org/10.3390/e26080616>
5. Carl M. Temporal Dynamics of Emotion and Cognition in Human Translation: Integrating the Task Segment Framework and the HOF Taxonomy. *Digital Studies in Language and Literature*. 2025. Vol. 2, № 2. P. 202–222. URL: <https://doi.org/10.1515/dsll-2025-0002>
6. Comprehensive Rust. URL: <https://github.com/comprehensive-rust/> (дата звернення: 10.04.2026).
7. Harrington Fernández O. Mindreading in Translation: Theory of Mind in a Shared Cognitive Translation Process. *Journal of Translation*. 2025. Vol. 21, № 1. P. 103–122. URL: <https://doi.org/10.54395/JOT-OHMIND>
8. He Z., Liang T., Jiao W., Zhang Z., Yang Y., Wang R., Tu Z., Shi S., Wang X. Exploring Human-Like Translation Strategy with Large Language Models. *Transactions of the Association for Computational Linguistics*. 2024. Vol. 12. P. 229–246. URL: https://doi.org/10.1162/tacl_a_00642
9. Khalaf B. K., Zin Z. M., Al-Abbas L. S. Contemporary Perspective on Cognitive Development: Reconceptualising Situational Context as Embedded Model. *International Journal of Instruction*. 2022. Vol. 15, № 1. P. 401–420. URL: <https://doi.org/10.29333/iji.2022.15123a>
10. Kirigami KDE. URL: <https://develop.kde.org/docs/> (дата звернення: 10.04.2026).
11. Kuhn M. UTF-8 and Unicode FAQ for Unix/Linux. URL: <https://www.cl.cam.ac.uk/~mgk25/unicode.html> (дата звернення: 10.04.2026).
12. Li W., Chen J., Li B., Liu B., Wen Z., Shan N., Liu X., Liu A., Liu H., Song H., Zhang L. TACTIC: Translation Agents with Cognitive-Theoretic Interactive Collaboration (Version 2). *arXiv*. 2025. URL: <https://doi.org/10.48550/ARXIV.2506.08403>
13. Li X., Yuan J., Xu Y. Investigating the Decision-making Style of Translation Raters in Large-scale Language Tests. *SAGE Open*. 2025. Vol. 15, № 3. P. 21582440251379569. URL: <https://doi.org/10.1177/21582440251379569>
14. Linietsky J., Manzur A., Godot Community. Godot Engine 4.6 Documentation. URL: <https://github.com/godotengine/godot-docs-110n> (дата звернення: 10.04.2026).
15. Liu S., Lyu C., Wu M., Wang L., Luo W., Zhang K., Shang Z. New Trends for Modern Machine Translation with Large Reasoning Models (Version 2). *arXiv*. 2025. URL: <https://doi.org/10.48550/ARXIV.2503.10351>
16. Oualline S. Vim Cookbook. URL: <http://www.oualline.com/vim-cook.html> (дата звернення: 10.04.2026).
17. Pradeep P. NCURSES Programming HOWTO. URL: <https://tldp.org/HOWTO/NCURSES-Programming-HOWTO/index.html> (дата звернення: 10.04.2026).
18. Qian M., Kong C. Enabling Human-Centered Machine Translation Using Concept-Based Large Language Model Prompting and Translation Memory. *Artificial Intelligence in HCI* / eds. H. Degen, S. Ntoa. Springer Nature Switzerland, 2024. Vol. 14736. P. 118–134. URL: https://doi.org/10.1007/978-3-031-60615-1_8
19. Shih C. Y.-Y. Problem-Solving and Decision-Making in Translation Revision: Two Case Studies. *Across Languages and Cultures*. 2015. Vol. 16, № 1. P. 69–92. URL: <https://doi.org/10.1556/084.2015.16.1.4>
20. Si G. Research on Optimized Path of Machine Translation from the Perspective of Cognitive Approach—Collaborative Development of Human Translation and Machine Translation. *Transactions on Social Science, Education and Humanities Research*. 2024. Vol. 10. P. 97–100. URL: <https://doi.org/10.62051/jqdzka20>
21. Supardi Moh. S. Cognitive Aspects of Translators' Activities in English-Indonesian Translation: An Analysis of Using Translog-II and Think-Aloud Protocol. *JURNAL BASIS*. 2025. Vol. 12, № 2. P. 367–376. URL: <https://doi.org/10.33884/basisupb.v12i2.10121>
22. Tiedemann J., Aulamo M., Bakshandaeva D., Boggia M., Grönroos S.-A., Nieminen T., Raganato A., Scherrer Y., Vázquez R., Virpioja S. Democratizing Neural Machine Translation With Opus-Mt. *Language Resources and Evaluation*. 2024. Vol. 58, № 2. P. 713–755. URL: <https://doi.org/10.1007/s10579-023-09704-w>
23. Tougher R. Xlib Programming in C++. URL: <https://linuxgazette.net/78/tougher.html> (дата звернення: 10.04.2026).
24. Uysal N. M. A Case for 'Situated Learning' in Translator Education: Translating Scientific Texts for a Digital Science Blog. *RumeliDE Dil ve Edebiyat Araştırmaları Dergisi*. 2021. № Ö9. P. 480–491. URL: <https://doi.org/10.29000/rumelide.981563>
25. Wang J. Cognitive Dynamics in Language Mediation Insights from Translation, Audiovisual Translation, and Interpreting Studies. *SHS Web of Conferences*. 2024. Vol. 187. P. 01022. URL: <https://doi.org/10.1051/shsconf/202418701022>

26. Wei Y., Hu X. Research on Transforming Translation Project Management with Large Language Models. *Journal of Basic and Applied Research International*. 2025. Vol. 31, № 4. P. 122–128. URL: <https://doi.org/10.56557/jobari/2025/v31i49560>

27. Yang X., Zhan R., Wong D. F., Wu J., Chao L. S. Human-in-the-loop Machine Translation with Large Language Model (Version 1). *arXiv*. 2023. URL: <https://doi.org/10.48550/ARXIV.2310.08908>

28. Yuhan N., Zhylin M., Antonyuk-Kyrychenko S., Popova N., Harbar M. Cognitive Aspects of Translation: The Latest Research in Psycholinguistics and Cognitive Science. *Forum for Linguistic Studies*. 2024. Vol. 6, № 4. P. 316–325. URL: <https://doi.org/10.30564/fls.v6i4.6775>

REFERENCES

1. Boiko I. V. (2021) Kohnityvno-dyskursyvni modeli perekladu [Cognitive-Discursive Models of Translation]. *Visnyk Luhanskoho natsionalnoho universytetu imeni Tarasa Shevchenka. Filolohichni nauky*, 7(345), 135–144. URL: [https://doi.org/10.12958/2227-2844-2021-7\(345\)-135-144](https://doi.org/10.12958/2227-2844-2021-7(345)-135-144) [in Ukrainian].

2. Borshchovetska V. D. (2022) Kohnityvnyi pidkhid do protsesu perekladu ta pererobky leksychnoi informatsii: Suchasni pidkhody, vyklyky i perspektyvy [Cognitive Approach to Translation and Vocabulary Processing: Modern Approaches, Challenges and Perspectives]. *Mizhnarodnyi filolohichni chasopys*, 13(1), 41–49. URL: <https://doi.org/10.31548/philolog2022.01.041> [in Ukrainian].

3. Carl M. (2024) An Active Inference Agent for Modeling Human Translation Processes. *Entropy*, 26(8), 616. URL: <https://doi.org/10.3390/e26080616>

4. Carl M. (2025) Temporal Dynamics of Emotion and Cognition in Human Translation: Integrating the Task Segment Framework and the HOF Taxonomy. *Digital Studies in Language and Literature*, 2(2), 202–222. URL: <https://doi.org/10.1515/dsll-2025-0002>

5. Harrington Fernández O. (2025) Mindreading in Translation: Theory of Mind in a Shared Cognitive Translation Process. *Journal of Translation*, 21(1), 103–122. URL: <https://doi.org/10.54395/JOT-OHMIND>

6. He Z., Liang T., Jiao W., Zhang Z., Yang Y., Wang R., Tu Z., Shi S., Wang X. (2024) Exploring Human-Like Translation Strategy with Large Language Models. *Transactions of the Association for Computational Linguistics*, 12, 229–246. URL: https://doi.org/10.1162/tacl_a_00642

7. Khalaf B. K., Zin Z. M., Al-Abbas L. S. (2022) Contemporary Perspective on Cognitive Development: Reconceptualising Situational Context as Embedded Model. *International Journal of Instruction*, 15(1), 401–420. URL: <https://doi.org/10.29333/iji.2022.15123a>

8. Li W., Chen J., Li B., Liu B., Wen Z., Shan N., Liu X., Liu A., Liu H., Song H., Zhang L. (2025) TACTIC: Translation Agents with Cognitive-Theoretic Interactive Collaboration (Version 2). *arXiv*. URL: <https://doi.org/10.48550/ARXIV.2506.08403>

9. Li X., Yuan J., Xu Y. (2025) Investigating the Decision-making Style of Translation Raters in Large-scale Language Tests. *SAGE Open*, 15(3), 21582440251379569. URL: <https://doi.org/10.1177/21582440251379569>

10. Liu S., Lyu C., Wu M., Wang L., Luo W., Zhang K., Shang Z. (2025) New Trends for Modern Machine Translation with Large Reasoning Models (Version 2). *arXiv*. URL: <https://doi.org/10.48550/ARXIV.2503.10351>

11. Qian M., Kong C. (2024) Enabling Human-Centered Machine Translation Using Concept-Based Large Language Model Prompting and Translation Memory. *Artificial Intelligence in HCI*, 14736, 118–134. URL: https://doi.org/10.1007/978-3-031-60615-1_8

12. Rebrii O., Demetska V. (2020) Adaptatsiia, asotsiatsiia ta analohiia: try vazhlyvykh mekhanizmy pryiniattia rishen v perekladu [Adaptation, Association, and Analogy: Triple A of the Translator's Decision-Making]. *East European Journal of Psycholinguistics*, 7(2), 231–242. URL: <https://doi.org/10.29038/eejpl.2020.7.2.reb> [in Ukrainian].

13. Shih C. Y.-Y. (2015) Problem-Solving and Decision-Making in Translation Revision: Two Case Studies. *Across Languages and Cultures*, 16(1), 69–92. URL: <https://doi.org/10.1556/084.2015.16.1.4>

14. Si G. (2024) Research on Optimized Path of Machine Translation from the Perspective of Cognitive Approach–Collaborative Development of Human Translation and Machine Translation. *Transactions on Social Science, Education and Humanities Research*, 10, 97–100. URL: <https://doi.org/10.62051/jqdzka20>

15. Supardi Moh. S. (2025) Cognitive Aspects of Translators' Activities in English-Indonesian Translation: An Analysis of Using Translog-II and Think-Aloud Protocol. *JURNAL BASIS*, 12(2), 367–376. URL: <https://doi.org/10.33884/basisupb.v12i2.10121>

16. Tiedemann J., Aulamo M., Bakshandaeva D., Boggia M., Grönroos S.-A., Nieminen T., Raganato A., Scherrer Y., Vázquez R., Virpioja S. (2024) Democratizing Neural Machine Translation With Opus-Mt. *Language Resources and Evaluation*, 58(2), 713–755. URL: <https://doi.org/10.1007/s10579-023-09704-w>

17. Uysal N. M. (2021) A Case for 'Situating Learning' in Translator Education: Translating Scientific Texts for a Digital Science Blog. *RumeliDE Dil ve Edebiyat Araştırmaları Dergisi*, (Ö9), 480–491. URL: <https://doi.org/10.29000/rumelide.981563>

18. Wang J. (2024) Cognitive Dynamics in Language Mediation Insights from Translation, Audiovisual Translation, and Interpreting Studies. *SHS Web of Conferences*, 187, 01022. URL: <https://doi.org/10.1051/shsconf/202418701022>

19. Wei Y., Hu X. (2025) Research on Transforming Translation Project Management with Large Language Models. *Journal of Basic and Applied Research International*, 31(4), 122–128. URL: <https://doi.org/10.56557/jobari/2025/v31i49560>

20. Yang X., Zhan R., Wong D. F., Wu J., Chao L. S. (2023) Human-in-the-loop Machine Translation with Large Language Model (Version 1). *arXiv*. URL: <https://doi.org/10.48550/ARXIV.2310.08908>

21. Yuhan N., Zhylin M., Antonyuk-Kyrychenko S., Popova N., Harbar M. (2024) Cognitive Aspects of Translation: The Latest Research in Psycholinguistics and Cognitive Science. *Forum for Linguistic Studies*, 6(4), 316–325. URL: <https://doi.org/10.30564/fls.v6i4.6775>
22. Comprehensive Rust. (n.d.) URL: <https://google.github.io/comprehensive-rust/> (Accessed: 10 April 2026)
23. Kirigami KDE. (n.d.) URL: <https://develop.kde.org/docs/> (Accessed: 10 April 2026)
24. Kuhn M. (n.d.) UTF-8 and Unicode FAQ for Unix/Linux. URL: <https://www.cl.cam.ac.uk/~mgk25/unicode.html> (Accessed: 10 April 2026)
25. Linietsky J., Manzur A., Godot Community. (n.d.) Godot Engine 4.6 Documentation. URL: <https://github.com/godotengine/godot-docs-110n> (Accessed: 10 April 2026)
26. Oualline S. (n.d.) Vim Cookbook. URL: <http://www.oualline.com/vim-cook.html> (Accessed: 10 April 2026)
27. Pradeep P. (n.d.) NCURSES Programming HOWTO. URL: <https://tldp.org/HOWTO/NCURSES-Programming-HOWTO/index.html> (Accessed: 10 April 2026)
28. Tougher R. (n.d.) Xlib Programming in C++. URL: <https://linuxgazette.net/78/tougher.html> (Accessed: 10 April 2026)

Дата першого надходження статті до видання: 17.04.2026

Дата прийняття статті до друку після рецензування: 05.05.2026

Дата публікації (оприлюднення) статті: 25.05.2026

Стаття поширюється на умовах
ліцензії відкритого доступу (CC BY 4.0)

