

УДК 81'33

МАНДЗІЙ Христина – студентка кафедри прикладної лінгвістики, Національний університет «Львівська політехніка», вул. Митрополита Андрея, 5, Львів, 79016, Україна (khrystyna.mandziy@gmail.com)

ORCID: <https://orcid.org/0000-0002-1216-9538>

ДІЛАЙ Маріанна – кандидат філологічних наук, доцент кафедри прикладної лінгвістики, Національний університет «Львівська політехніка», вул. Митрополита Андрея, 5, Львів, 79016, Україна (marianadilai@gmail.com)

ORCID: <https://orcid.org/0000-0001-5182-9220>

DOI: <https://doi.org/10.24919/2522-4565.2021.47.16>

Бібліографічний опис статті: Мандзій, Х., Ділай, М. (2021). Особливості побудови та перспективи використання англійсько-українського паралельного корпусу ІТ-текстів. *Проблеми гуманітарних наук: збірник наукових праць Дрогобицького державного педагогічного університету імені Івана Франка. Серія «Філологія», 47, 120–127, doi: <https://doi.org/10.24919/2522-4565.2021.47.16>*

ОСОБЛИВОСТІ ПОБУДОВИ ТА ПЕРСПЕКТИВИ ВИКОРИСТАННЯ АНГЛІЙСЬКО-УКРАЇНСЬКОГО ПАРАЛЕЛЬНОГО КОРПУСУ ІТ-ТЕКСТІВ

Анотація. У статті обґрунтовано актуальність створення англійсько-українського паралельного корпусу текстів сфери інформаційних технологій та окреслено перспективи його використання передусім у перекладацьких і мовознавчих студіях. Особливу увагу приділено опису основних етапів побудови паралельного корпусу, зокрема збору даних, вичитці й унормуванню текстів, вилученню нетекстової інформації, трансформуванню текстів у файли необхідного формату, завантаженню корпусу в корпусний менеджер і текстовий аналізатор «Sketch Engine», додаванню метатекстової розмітки, наданню індивідуального доступу користувачам.

Хоч в українській корпусній лінгвістиці вже наявні деякі корпуси (Український національний лінгвістичний корпус (УНЛК), Генеральний регіонально анотований корпус української мови (ГРАК) тощо), нестача порівняльних корпусів, зокрема англійсько-українських, усе ж відчутна. З огляду на потужність і швидкість розвитку в наші дні сфери ІТ і на труднощі перекладу ІТ-текстів, зокрема їх термінів, актуально створити англійсько-український паралельний корпус саме для текстів сфери інформаційних технологій.

Основну цінність англійсько-українського паралельного корпусу ІТ-текстів убачаємо в тому, що він може бути джерелом великої кількості перекладацьких рішень, що дасть змогу полегшити процес перекладу ІТ-текстів, покращити їх якість, уникнути стихійності перекладів. Показано також, як побудований корпус текстів можна застосувати для здійснення мовознавчих досліджень ІТ-дискурсу та дійти обґрунтованих висновків стосовно вживання в ньому мовних одиниць, визначення особливостей ІТ-термінології, аналізу структурних моделей і способів перекладу термінів, стосовно виокремлення особливостей стилю текстів сфери ІТ з оперттям на реальні контексти і статистичні параметри.

Ключові слова: паралельний корпус, ІТ-тексти, технічна термінологія, труднощі перекладу.

MANDZIY Khrystyna – Student at the Department of Applied Linguistics, Lviv Polytechnic National University, 5, Mytropolita Andreia str., Lviv, 79016, Ukraine (khrystyna.mandziy@gmail.com)

ORCID: <https://orcid.org/0000-0002-1216-9538>

DILAI Marianna – PhD in Philology, Associate Professor at the Department of Applied Linguistics, Lviv Polytechnic National University, 5, Mytropolitya Andreia str., Lviv, 79016, Ukraine (mariannadilai@gmail.com)

ORCID: <https://orcid.org/0000-0001-5182-9220>

DOI: <https://doi.org/10.24919/2522-4565.2021.47.16>

To cite this article: Mandziy, K., Dilai, M. (2021). Osoblyvosti pobudovy ta perspektvyvy vykorystannia anhliisko-ukrainskoho paralelnoho korpusu IT tekstiv [Specificity of compiling and perspectives of using of English-Ukrainian parallel corpus of IT texts]. *Problemy humanitarnykh nauk: zbirnyk naukovykh prats Drohobyskoho derzhavnoho pedahohichnoho universytetu imeni Ivana Franka. Seriiia "Filolohiia" – Problems of Humanities. "Philology" Series: a collection of scientific articles of the Drohobych Ivan Franko State Pedagogical University*, 47, 120–127. doi: <https://doi.org/10.24919/2522-4565.2021.47.16> [in Ukrainian].

SPECIFICITY OF COMPILING AND PERSPECTIVES OF USING OF ENGLISH-UKRAINIAN PARALLEL CORPUS OF IT TEXTS

Summary. *The article deals with the topicality of creating an English-Ukrainian parallel corpus of texts in the field of information technology and outlines the prospects for its use primarily in translation and linguistic studies. Particular attention is paid to the main stages of building a parallel corpus, including data collection, proofreading and normalization of texts, deletion of non-textual information, transformation of texts into files of the needed format, uploading the corpus to the corpus manager and text analyzer Sketch Engine, adding corpus markup, providing individual access to users.*

Although some corpora already exist in Ukrainian corpus linguistics (Ukrainian National Linguistic Corpus, General Regional Annotated Corpus of the Ukrainian Language, etc.), there is still a lack of parallel corpora, including English-Ukrainian ones. And, given how powerfully and rapidly the IT sector is developing these days, as well as taking into account the peculiarities and difficulties of translating IT discourse and, in particular, terminology, it will be useful to create an English-Ukrainian parallel corpus of IT texts.

The main value of the English-Ukrainian parallel corpus of IT texts is that it can be a source of a large number of translation solutions, which will facilitate the process of translating IT texts, as well as improve their quality, avoid spontaneous translations. At the same time, it is shown how the constructed corpus of texts can be used for linguistic research into IT discourse. It allows us to make valid conclusions about the use of language units based on real contexts and statistical parameters (identify features of IT terminology, analyze frequently used structural models and translation approaches, highlight features of the style of texts in the field of IT).

Key words: *parallel corpus, IT texts, technical terminology, translation difficulties.*

Постановка проблеми. У сучасному світі значення перекладацької діяльності постійно зростає, але разом із цим виникають різноманітні перекладацькі проблеми. Одним із сучасних інструментів їх вирішення є створення корпусу текстів, серед яких виділяють паралельні тексти двох або більше мов. Попри те що корпусна лінгвістика в Україні сьогодні перебуває у фазі розвитку, досі існує велика прогалина в цій галузі, особливо це стосується паралельних українсько-іноземних корпусів текстів.

Сфера інформаційних технологій в Україні теж активно розвивається. Важливим питанням є доступність і зрозумілість інформації для всіх зацікавлених верств населення. Тут варто зважати й на те, що всі новачки при-

ходять до нас із-за кордону, тож більшість цінної інформації доступна лише іноземними мовами. Виникає потреба створення якісних перекладів цієї інформації українською мовою.

Аналіз останніх досліджень і публікацій. Корпусна лінгвістика – один із найперспективніших розділів сучасного мовознавства, який, окрім усього, вирізняється швидкими темпами свого розвитку. Її фокусом є побудова та опрацювання текстових корпусів, а також використання останніх як джерела фактичного матеріалу для досліджень мовної різноманітності.

Перевагою корпусів текстів є те, що їх утворюють реальні немодифіковані уривки усного або писемного мовлення. Завдяки

цьому корпус стає емпіричною категорією, а його матеріал може розглядатися як емпірична база для досліджень природної мови (Демська-Кульчицька, 2004).

Хоч не всі вчені повністю довіряють методам корпусної лінгвістики, її важливість підтверджена великою різноманітністю розвідок, у яких наведено чимало аргументів на користь того, що сьогодні вже неможливо проводити лінгвістичні дослідження так, щоб вони були повністю поза межами корпусної лінгвістики: Д. Байбер, Дж. Сінклер, В. Тойберт, В.А. Плуґян (Бобер, 2018).

Українські мовознавці теж присвячують праці корпусній лінгвістиці, зокрема вивченню її теоретичних засад, методів, технічних проблем і практичному застосуванню корпусів текстів. Серед них згадаємо передусім Н. Дарчук, О. Ванівську, І. Данилюка, Т. Бобкову, І. Кульчицького, В. Широкова.

Усе зазначене підтверджує важливість розвитку корпусної лінгвістики та, зокрема, актуальність укладання англійсько-українського паралельного корпусу текстів сфери ІТ, яка є однією з провідних галузей сучасності. Такий корпус був би першим в українській корпусній лінгвістиці й може стати корисним інструментом для перекладачів, фахівців зі сфери машинного перекладу, а також тих, хто вивчає англійську або українську мови чи цікавиться сферою інформаційних технологій.

Мета статті – описати особливості й етапи побудови англійсько-українського паралельного корпусу текстів сфери ІТ та окреслити перспективи його використання.

Реалізація мети передбачає виконання таких завдань: аналіз напрацювань щодо створення й застосування паралельних корпусів текстів; обґрунтування необхідних комп'ютерних засобів для створення паралельного корпусу текстів; виокремлення перспектив використання створеного корпусу, зокрема, для перекладу та лінгвістичного аналізу.

Виклад матеріалу. За останні десятиліття українська корпусна лінгвістика зробила значний крок уперед, позаяк створено завдяки кропіткій праці вітчизняних мовознавців уже кілька корпусів української мови: Український національний лінгвістичний корпус (УНЛК) (http://unlc.icybcluster.org.ua/virt_unlc/), Дослідницький корпус сучасної української

мови (режим доступу на лінгвістичному порталі [Mova.info](http://mova.info): <http://mova.info/corpus.aspx>), Генеральний регіонально анований корпус української мови (ГРАК) (<http://uacorpus.org>). Проте деякі її напрями досі потребують більшої уваги, відчутним є також брак деяких видів корпусів, які були б дуже цінними для вітчизняних дослідників. Ідеться, зокрема, про необхідність розроблення паралельних корпусів, у яких тексти українською мовою зіставлені з текстами іншими мовами. Звичайно, варто згадати паралельні корпуси, які вже є у відкритому доступі в мережі Інтернет: російсько-український підкорпус Національного корпусу російської мови (<http://ruscorpora.ru/search-para-uk.html>) та польсько-український корпус (<http://domeczek.pl/~polukr>).

Отже, одним із перспективних та актуальних напрямів розвитку корпусної лінгвістики в Україні є укладання паралельних корпусів, які могли б широко застосовувати мовознавці, перекладачі, фахівці сфери штучного інтелекту й машинного перекладу та ін.

Сфера інформаційних технологій теж належить до перспективних напрямів. Зараз циркулює велика кількість інформації, пов'язаної із цією сферою, і кількість зацікавлених людей теж зростає. Це породжує підвищений інтерес до перекладів тематичних текстів. Попри те що велика частина комп'ютерної термінології стала сьогодні загальновідомою, цей шар лексики дуже динамічно зростає, з'являються нові вузькоспеціальні терміни. Окрім цього, самі способи творення цієї лексики, лінгвістичні й екстралінгвальні чинники, що впливають на неї, ще мало вивчені (термінологію сфери ІТ досліджували Н. Виноградова, Н. Грицик, Е. Балюта, А. Бічева). Усе це можна порівняти з вибухом – неконтрольованим процесом, який спричиняє все нові й нові зміни (Сидор, Нанівський, 2019). У такому середовищі перекладачам щоразу доводиться приймати рішення стосовно того, до яких технік перекладу вдатися та як правильно перекласти той чи той термін. Згадані вище чинники стимулювали ідею щодо створення паралельного англійсько-українського корпусу текстів зі сфери ІТ як цінного підґрунтя для проведення лінгвістичних досліджень, розв'язання проблем перекладу чи для реалізації інших цілей.

Робота над створенням корпусу – це процес, який передбачає поступове виконання певного списку завдань у чітко визначеній послідовності. Планування процесу укладання англійсько-українського паралельного корпусу ІТ-текстів відбувалося відповідно до виділених В. Жуковською основних етапів побудови корпусу текстів (Жуковська, 2013).

Перший етап створення корпусу – це відбір даних, а саме текстів, які становитимуть корпус. Для створюваного англійсько-українського паралельного корпусу необхідно було вибрати тести, що стосуються сфери ІТ і доступні в мережі Інтернет англійською як мовою оригіналу й українською як мовою перекладу. Цей етап виявився доволі проблематичним, тому що практика перекладу текстів потрібної тематики з англійської українською доволі непопулярна. Такі переклади комерційно не вигідні, позаяк зазвичай більшість фахівців сфери ІТ вільно володіють англійською мовою. Крім того, частина літератури перекладена з англійської російською мовою, і нею часто користуються вітчизняні спеціалісти. Тож кількість якісних матеріалів для корпусу обмежена. У доступі є деякі підручники з програмування, а також незначна кількість перекладеної технічної документації. Варто також зазначити, що якість цих текстів не завжди висока, тому що нерідко такі матеріали перекладають не професійні перекладачі, а технарі-ентузіастки. На момент написання статті базу корпусу становили тексти чотирьох підручників із програмування («Програмування мовою Java для дітей, батьків, дідусів та бабусь» Якова Файна (<http://myflex.org/books/java4kids/java4kids.htm>), «Підручник мови Python» Гвідо ван Россума й Фреда Л. Дрейка Молодшого (<https://docs.python.org/3/tutorial/>), «Пориньте у Python 3» Марка Пілігрима (<https://diveintopython3.net/>), «Вивчить собі Хаскела на велике щастя!» Мірана Ліповачі (<http://learnyouahaskell.com/>)), двох сайтів із технічною документацією для фреймворку «Bootstrap» (<https://getbootstrap.com/docs/3.3/>) та скриптової мови програмування php (<https://github.com/php/doc-uk>).

Наступним кроком є вчитка й унормування тестів, пов'язані з процедурами усунення непотрібного форматування та деяких невідповідностей, видалення нетекстової

інформації (формул, шматків програмного коду, таблиць, рисунків тощо). Крім того, створення паралельного корпусу передбачає подання його в таблиці, де речення оригіналу та речення перекладу мають бути чітко зіставлені. Таку роботу майже неможливо автоматизувати: дослідник змушений вручну усувати розбіжності на рівні речень. Результатом цього етапу є файл з усіма реченнями корпусу, зіставленими за логікою «оригінал – переклад». Станом на сьогодні корпус містить 9854 таких пар речень обсягом понад 300 тис. слів.

Коли тексти корпусу вчитані та внормовані, їх можна завантажити в застосунок із функціоналом для здійснення пошуку й аналізу в корпусі. Для створення англійсько-українського паралельного корпусу ІТ-текстів ми обрали корпусний менеджер і текстовий аналізатор «Sketch Engine» (<https://www.sketchengine.eu>). Цей програмний продукт дає змогу створювати власний паралельний корпус текстів, будувати на його основі паралельні конкорданси, частотні списки тощо. Користувач може здійснювати пошук за словами або фразами й отримувати результат у вигляді паралельного конкордансу, де всі речення однієї мови з пошуковим словом або фразою зіставлені з відповідними реченнями іншою мовою (такий пошук працює як з англійської українською, так і з української англійською мов). Доступне, окрім цього, фільтрування результатів за умовою «містить/не містить».

«Sketch Engine» уможливорює додавання до корпусу метатекстової розмітки. Вона передбачає опис завантажених текстів за низкою параметрів, що стосуються бібліографічного опису творів, а також даних про особу автора та перекладача. Метатекстова розмітка є важливим блоком корпусної розмітки, що дає змогу стежити за збалансованістю складу текстів, а також надає користувачеві можливість добирати і групувати текстовий матеріал за різноманітними параметрами. На рис. 1 зображено 6 категорій метарозмітки, представлених у корпусі, за якими користувач має змогу фільтрувати результати: ім'я автора, ім'я перекладача, оригінальна назва твору, назва твору в перекладі, рік видання оригіналу, рік видання перекладу.

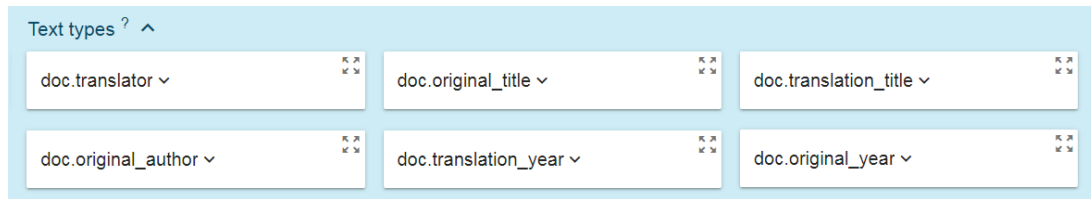


Рис. 1. Категорії метарозмітки в англійсько-українському паралельному корпусі ІТ-текстів

Анотація додає корпусу цінності, значно розширюючи діапазон дослідницьких питань, що можуть бути простежені на його матеріалі.

Основна цінність англійсько-українського паралельного корпусу ІТ-текстів полягає в тому, що він може бути джерелом великої кількості перекладацьких рішень. Попри те що в ІТ-сфері тексти часто містять багато усталених термінів, до яких неважко дібрати переклад, стрімкий розвиток сфери сприяє виникненню нових понять і термінів. Перекладачеві, щоб якісно виконати своє завдання, важливо добре зрозуміти суть та історію походження кожного невідомого терміна, а після цього дібрати найоптимальніший спосіб перекладу. У такій ситуації дуже цінною буде інформація про те, які переклади цього терміна вже наявні. Нижче (рис. 2) наведено приклад пошуку в корпусі фрази «instance declaration», що повертає 5 результатів, на

основі яких перекладач може обрати варіант для свого перекладу.

Як бачимо, словосполучка «instance declaration» перекладена двома способами: як «утілення» або «instance-означення». З погляду програмування в конструкціях із ключовим словом «instance» означають тіла функцій і в такий спосіб означають утілення. Перекладач повинен прийняти рішення, який варіант доречніший для конкретного випадку. Наприклад, якщо в реченні йдеться про використану конструкцію, то потрібно обрати варіант «instance-означення», а якщо йдеться про саме надане значення, то – «утілення».

Паралельний корпус текстів може бути базою для мовознавчих досліджень ІТ-дискурсу. Дослідимо, наприклад, як часто в текстах трапляються окличні речення. Робимо це для того, щоб переконатися, що ІТ-дискурсу справді не притаманна емоційна забарвленість. Виконати таке завдання дуже просто –

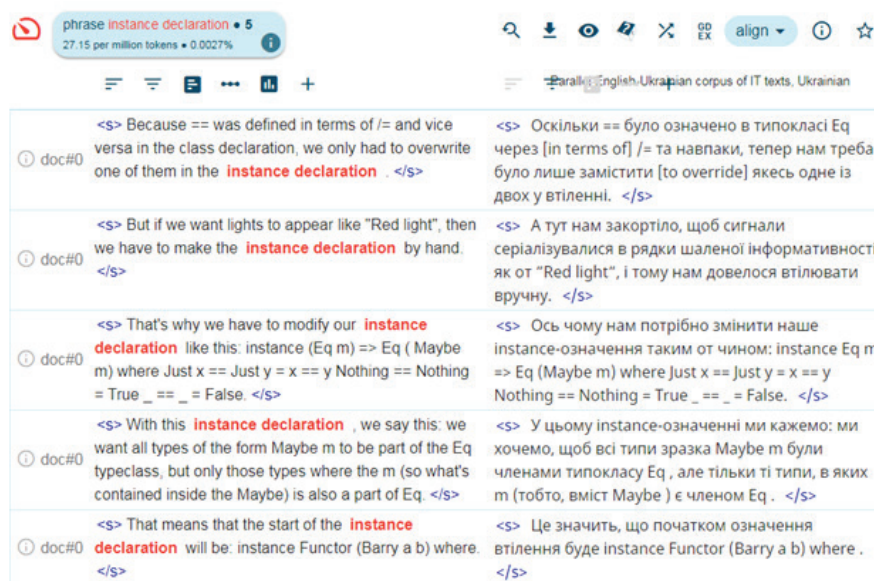


Рис. 2. Результат пошуку фрази «instance declaration» в англійсько-українському паралельному корпусі ІТ-текстів

здійснюємо пошук за одним символом – знаком оклику (!). Пошук видає 304 окличних речень, що становить усього 3% від усіх речень корпусу. Хоч і в невеликій кількості, але експресивні речення в текстах представлені. Як зазначалося вище, наш корпус формують тексти чотирьох підручників і технічної документації з двох сайтів. Можемо припустити, що окличні речення використовують автори підручників для привернення уваги та зацікавлення читача. Щоб це підтвердити, профільнуємо результати так, щоб отримати лише окличні речення з технічної документації (метарозмітка це допускає). За таким запитом отримуємо лиш один результат (рис. 3.). Отже, окличні речення справді мало притаманні IT-дискурсу, а тим більше технічній документації.

Ще одним прикладом застосування корпусу для лінгвістичного аналізу є визначення найпоширеніших способів перекладу термінів комп'ютерної сфери. Це дослідження зосередимо на ста найчастотніших термінах англомовної частини корпусу (перші з них – «function», «class», «method», «file», «value»). За допомогою пошуку в корпусі ми отримали всі варіанти україномовних перекладів для цих термінів. Серед них виділимо такі структурні моделі для обох мов: іменник (N), дієслово (V), прикметник (Adj), словосполучення (Word combination). Також трапляються комбінації цих моделей. У таблиці 1 представлені результати такого аналізу – усі наявні моделі перекладу (для одного терміна їх може бути більше ніж одна) та відсоткове значення для кожної із цих моделей.

Отже, найпоширенішими моделями перекладу для нашої вибірки є моделі «іменник – іменник» та «іменник/дієслово – дієслово». Щодо деяких термінів у межах цих моделей застосовано різні способи перекладу, наприклад, калькування та пошук еквіва-

лентів (це ілюструють приклади, наведені в таблиці 1).

Окрім цього, корпус може бути чудовою базою знань для людей, які хочуть глибше осягнути поняття сфери IT (як перекладачі, так і студенти). Наприклад, слово «performance» часто трапляється в IT-дискурсі й означає, «наскільки добре щось працює». Проте з першого погляду не завжди зрозуміло, про що саме йдеться. Користувач корпусу може швидко проаналізувати результати, отримані за пошуковим словом «performance»: серед результатів переважає переклад «продуктивність», трапляється також «ефективність»: «*However, sometimes there is a need for alternative implementations with different performance trade-offs*» – «*Але інколи потрібно застосовувати альтернативні рішення, які можуть вплинути на ефективність виконання програми*». Ці переклади дають розуміння значення досліджуваного слова, але можуть залишати певну невизначеність. У деяких результатах трапляється слова «швидкість» або «швидко»: «*The timeit module quickly demonstrates a modest performance advantage*» – «*Модуль timeit швидко демонструє, що традиційний спосіб є набагато швидшим*». Так користувач може дійти висновку, що йдеться про те, наскільки швидко програма виконує свої завдання. Такі вправи є особливо корисними для перекладачів, які хочуть зробити свої переклади максимально точними та природними.

Висновки. Попри наявність деяких українськомовних корпусів (Український національний лінгвістичний корпус, дослідницький корпус сучасної української мови, російсько-український підкорпус Національного корпусу російської мови та польсько-український корпус), питання створення паралельних корпусів актуальне й потребує додаткової уваги науковців. З огляду

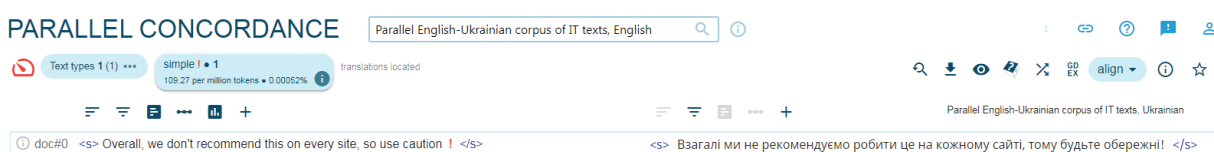


Рис. 3. Результат пошуку окличних речень у технічній документації

Таблиця 1

Структурні моделі перекладу термінів сфери ІТ

Структурна модель англійського терміна	Структурна модель українського терміна	Кількість прикладів	Приклади
N	N	64 (60,38%)	installer – інсталятор/установник; bug – помилка/проблема/неполадка
N/V	V	21 (19,82%)	import – імпортувати; click – клікнути/натиснути
N	Adj + N	9 (8,49%)	dropdown – спадне меню; modal – модальне вікно
N	N + N	4 (4,72%)	environment – середовище розробки
N/V	V + N	2 (1,89%)	underflow – утрачати значущість
N	Word combination	5 (4,7%)	carputer – комп'ютер, вбудований в авто

на те, наскільки потужно й швидко в нашій дні розвивається сфера ІТ, варто створити англійсько-український паралельний корпус ІТ-текстів. Це нелегке завдання, зокрема, через обмежену кількість потрібних текстів. Сьогодні створено корпус із 9854 пар речень, зіставлених за логікою «оригінал – переклад». Доступ до корпусу здійснено через корпусний менеджер і текстовий аналізатор «Sketch Engine». Корпус містить

метатекстову розмітку. Подальша робота над корпусом передбачає розширення кількості текстів, що забезпечить репрезентативність корпусу, а також подальше додавання розмітки, що допоможе автоматизувати роботу з корпусом. Цей корпус може бути дуже цінним джерелом інформації для перекладачів, які шукають найкращі перекладацькі рішення для своїх текстів, і мовознавців, які аналізують ІТ-дискурс.

ЛІТЕРАТУРА

- Бобер Н.** Корпусний підхід у контексті досліджень прикладної лінгвістики. *Актуальні питання іноземної філології*. 2018. № 8. С. 34–40.
- Паралельний корпус текстів ПарКУМ / Н.П. Дарчук, М.О. Лангенбах, В.М. Сорокін, Я.В. Ходаківська. *Науковий часопис НПУ імені М. П. Драгоманова. Серія 9 «Сучасні тенденції розвитку мов»*. 2017. № 15. С. 28–35.
- Демська-Кульчицька О.** Дещо про класифікацію текстових корпусів. *Наукові записки. Серія «Мовознавство»*. 2004. № 1 (11). С. 153–157.
- Жуковська В.В.** Вступ до корпусної лінгвістики : навчальний посібник. Житомир : ЖДУ ім. І. Франка, 2013. 142 с.
- Сидор А.Р., Нанівський З.С.** Урахування лексичних особливостей сфери інформаційних технологій під час перекладу з англійської мови українською. *Закарпатські філологічні студії*. 2019. № 7. С. 47–51.
- Широков В.А., Бугаков О.В., Грязнухіна Т.О.** Корпусна лінгвістика. Київ : Довіра, 2005. 471 с.

REFERENCES

- Bober, N.** (2018). Korpusnyi pidkhd u konteksti doslidzhen prykladnoi lnhvistyky [Corpus Approach in the Context of Applied Linguistics Research]. *Aktualni pytannia inozemnoi filolohii – Current Issues of Foreign Philology*, 8, 34–40 [in Ukrainian].
- Darchuk, N., Lanhenbakh, M., Sorokin, V., Khodakivska, Ya.** (2017). Paralelnyi korpus tekstiv ParKUM [Parallel corpus of texts ParKUM]. *Naukovyi chasopys NPU imeni M. P. Drahomanova. Seriiia 9. Suchasni tendentsii rozvytku mov – Scientific journal of The National Pedagogical Dragomanov University. Series 9. Modern trends in language development*, 5, 28–35 [in Ukrainian].
- Demska-Kulchytska, O.** (2004). Deshcho pro klasyfikatsiiu tekstovyykh korpusiv [Something about the Classification of Text Corpora]. *Naukovi zapysky. Seriiia : Movoznavstvo – Scientific Notes. Serie s: Linguistics*, 1(11), 153–157 [in Ukrainian].
- Zhukovska, V.** (2013). *Vstup do korpusnoi lnhvistyky : navchalnyi posibnyk [Introduction to Corpus Linguistics]*. Zhytomyr : Zhsu im. I. Franka [in Ukrainian].
- Sydor, A., Nanivskyi, Z.** (2019). Urakhuvannia leksychnykh osoblyvostei sfery informatsiinykh tekhnolohii pid chas perekladu z anhliskoi movy ukrainskoiu [Taking Into Account The Lexical Features Of

The Field Of Information Technology When Translating From English Into Ukrainian]. *Zakarpatski filologichni studii – Transcarpathian philological studies*, 7, 47–51 [in Ukrainian].

Shyrov, V., Buhakov, O., Hriaznukhina, O. (2005). *Korpusna linhvistyka [Corpus Linguistics]*. Kyiv : Dovira [in Ukrainian].