

## ОСОБЛИВОСТІ ЗАСТОСУВАННЯ ЗАСОБІВ КОРПУСНОЇ ЛІНГВІСТИКИ ДЛЯ ПОКРАЩЕННЯ ЯКОСТІ ПЕРЕКЛАДУ ТА ОПТИМІЗАЦІЇ ЙОГО ПРОЦЕСУ

**Варешкіна Н. В.**

*кандидат педагогічних наук, доцент,  
доцент кафедри іноземних мов*

*Національний університет «Одеська юридична академія»*

*вул. Фонтанська дорога, 23, Одеса, Україна*

*orcid.org/0000-0002-2184-1007*

*vareshkinanv@gmail.com*

**Гроза І. М.**

*старший викладач кафедри іноземних мов*

*Національний університет «Одеська юридична академія»*

*вул. Фонтанська дорога, 23, Одеса, Україна*

*orcid.org/0000-0003-2449-4378*

*grozairinal2@gmail.com*

**Ключові слова:** *корпусна лінгвістика, корпус, корпусний інструментарій, паралельний корпус, порівняльний корпус, переклад.*

Статтю присвячено аналізу можливостей засобів корпусної лінгвістики у прикладних і дослідницьких цілях, розгляду основних труднощів, з якими стикаються перекладачі у перекладацькому процесі, а також наданню пропозицій щодо уникнення перекладацьких помилок, покращення якості та оптимізації процесу перекладу за допомогою альтернативного підходу – застосування корпусного інструментарію. Цей підхід ґрунтується на використанні електронних корпусів текстів – величезних масивів природних текстів іноземною мовою, зібраних на магнітному носії, які належним чином упорядковані та розмічені для швидшого пошуку певної лінгвістичної інформації. Сутність проблеми полягає в тому, що питання щодо використання корпусного підходу до перекладознавчих досліджень залишається одним із найактуальніших у царині сучасної лінгвістики, оскільки комп'ютерні технології, прийшовши в лінгвістику, запропонували свої технічні можливості для опрацювання, збереження й відбору мовного матеріалу.

У статті розглядається поняття корпусів текстів, визначаються основні їх види, відмінні риси, можливості, переваги й недоліки використання основних видів перекладацьких корпусів у вивченні мови на їх базі та застосуванні у процесі перекладу. Надається визначення можливостей використання корпусу текстів у процесі вирішення тієї чи іншої перекладацької проблеми для підвищення якості результату перекладу та оптимізації його процесу. Стаття містить аналіз перекладацького процесу із застосуванням корпусного інструментарію, результати якого дають широкі можливості для вивчення корпусу як альтернативного рішення для покращення якості та оптимізації перекладу. Пропонується практичне застосування корпусу на прикладах окреслення основних його функцій і можливостей, що сприяють досягненню найповнішого рівня еквівалентності та адекватного перекладу. У статті приділено увагу перевагам застосування корпусу у процесі перекладу, наприклад вищій репрезентативності, реалістичності та зручності. Робота відкриває перспективи для подальшого вивчення корпусного інструментарію з метою виявлення нових корисних можливостей для полегшення роботи перекладача та підвищення якості перекладацького продукту.

## PECULIARITIES OF APPLICATION OF CORPUS LINGUISTICS MEANS TO IMPROVE THE QUALITY OF TRANSLATION AND OPTIMIZE THE TRANSLATION PROCESS

**Varieshkina N. V.**

*Candidate of Pedagogical Sciences, Associate Professor,  
Associate Professor at the Department of Foreign Languages  
National University "Odessa Academy of Law"  
Fontanska doroha str., 23, Odesa, Ukraine  
orcid.org/0000-0002-2184-1007  
vareshkinanv@gmail.com*

**Hroza I. N.**

*Senior Lecturer at the Department of Foreign Languages  
National University "Odessa Academy of Law"  
Fontanska doroha str., 23, Odesa, Ukraine  
orcid.org/0000-0003-2449-4378  
grozairinal2@gmail.com*

**Key words:** *corpus linguistics, corpus, corpus tools, parallel corpus, comparative corpus, translation.*

This article analyzes the possibilities of corpus linguistics means for applied and research purposes, considers the main difficulties faced by translators during the translation process, as well as suggestions for avoiding translation errors, improving the quality and optimizing the translation process with an alternative approach – the use of corpus tools. This approach is based on the use of electronic corpora of texts – huge arrays of natural texts in a foreign language, collected on magnetic media, which are properly organized and marked to find quickly certain linguistic information. The essence of the problem is that the issue of using a corpus approach to translation studies remains one of the most relevant in the field of modern linguistics, because computer technologies, having come to linguistics, offered its technical capabilities for processing, preserving and selecting language material. The article considers the concepts of the texts corpora, the main types of corpora, identifies the distinctive features, capabilities, advantages and disadvantages of using the main types of translation corpora in language learning on their basis and application in the translation process. The definition of using the corpus of texts opportunities in the process of solving the translation problems to improve the quality of the translation result and optimize its process is provided. The presented article contains an analysis of the translation process using corpus tools, the results of which provide the wide opportunities to study the corpus as an alternative solution to improve the quality and optimize translation. It is proposed the practical application of the corpus on the examples of outlining its main functions and capabilities, which contribute to the achievement of the fullest level of equivalence and adequate translation. Attention is paid to the advantages of using the corpus during the translation process, such as higher representativeness, actuality and convenience. The work develops the prospects for further study of the corpus tools to identify the new useful opportunities to facilitate the work of the translator and improve the quality of the translation product.

**Вступ.** У новому тисячолітті переклад як один із засобів спілкування починає відігравати дедалі важливішу роль. У зв'язку зі стрімким розширенням масштабів інформаційного обміну збільшу-

ється обсяг перекладів з одних мов іншими. Поява нових галузей науки та знання, а також постійне зростання інформації, що генерується і споживається, є наслідком науково-технічного прогресу. За

помітного зменшення часу, що виділяється замовником на виконання перекладацького завдання, та значного підвищення вимог до якості перекладу спостерігається ситуація створення для перекладача дуже жорстких умов роботи.

Багато дослідників (наприклад, Х. Крингс, Р. А. Усов, П. Бейкер, Т. МакЕнері, Дж. Ліч, Дж. Аартс, О. М. Демська-Кульчицька, В. В. Жуковська, В. В. Захаров та інші) вивчали ефективні методи пошуку адекватних перекладацьких рішень і навичок, що забезпечують значне підвищення якості перекладу.

Однак, незважаючи на досить значну наукову базу даних та велику кількість наукових експериментів, присвячених процесу письмового перекладу, на периферії перекладознавчого дослідження залишається така важлива проблема, як можливості застосування засобів корпусної лінгвістики як один зі способів вирішення проблеми підвищення якості перекладу й удосконалення процесу перекладу.

**Актуальність дослідження** зумовлена стрімким розвитком електронних та інформаційних технологій, а також науки про мову, що дає змогу використовувати оптимальні рішення у практичній галузі перекладу для вдосконалення процесу перекладу та отримання якісних результатів.

**Мета роботи** полягає в тому, щоб дослідити можливі способи використання корпусного інструментарію у процесі вирішення перекладацьких завдань.

Наші основні **завдання** полягають в аналізі корпусного інструментарію і можливостей його застосування в перекладацькій діяльності та лінгвістичних дослідженнях; у визначенні можливостей використання корпусу текстів у процесі вирішення тієї чи іншої перекладацької проблеми для підвищення якості результату перекладу та оптимізації цього процесу.

**Об'єктом дослідження** є переклад текстів як процес і результат мовної комунікації, спрямований на досягнення комунікативної мети між представниками різних країн та культур.

**Предмет дослідження** – подолання труднощів у процесі перекладу текстів, підвищення якості результату перекладу шляхом застосування програмних засобів корпусної лінгвістики.

**Виклад основного матеріалу.** За останні декілька десятиліть завдяки інтенсивній комп'ютеризації суспільства отримав розвиток такий новий і досить перспективний напрям у лінгвістичній науці, як корпусна лінгвістика. Залучення комп'ютерів і спеціальних телекомунікаційних мереж дало змогу не тільки зберігати великі обсяги текстів в електронному вигляді, а й здійснювати пошук у них, обробляти їх. Завдання зі створення зібраних текстів в електронному вигляді, або кор-

пусів, є настільки значущим для сучасної лінгвістики, що ці зібрання електронних текстів стають об'єктом досліджень спеціального розділу прикладної лінгвістики – корпусної лінгвістики, яка займається розробленням загальних принципів побудови й використання лінгвістичних корпусів за допомогою комп'ютерів [1, с. 4].

Статус корпусної лінгвістики в сучасній мовознавчій парадигмі однозначно ще не визначений. Так, Т. МакЕнері, Е. Вілсон, Р. Ксіао та інші лінгвісти характеризують її як особливу «методологію» дослідження мовного матеріалу, на противагу таким традиційним галузям лінгвістики, як фонетика, синтаксис, семантика чи граматики. Зазначені мовознавчі галузі досліджують окремий аспект мовної системи, натомість корпусна лінгвістика не обмежується вивченням лише одного аспекту мовної реалізації. Більше того, корпусні методи можуть використовуватися для вивчення мовних одиниць будь-якого мовного рівня. Корпусна лінгвістика як система методів і принципів використання корпусів для вивчення мови та для навчання мови має теоретичне підґрунтя, однак саме по собі воно ще не є науковою теорією. Через це корпусна лінгвістика визнається методологією вивчення мови із широкими можливостями застосування в багатьох лінгвістичних напрямках і теоріях [2, с. 6–7].

Центральним поняттям корпусної лінгвістики є поняття корпусу. Корпус текстів, на думку Т. МакЕнера та А. Харді, являє собою «збирання мовних фрагментів, відібраних відповідно до чітких мовних критерій для використання як моделі мови» [3]. Корпус як величезний масив даних характеризується репрезентативністю та наповнюваністю, а також певною структурою побудови. На відміну від одиничних текстів чи паперових видань, корпус текстів дає можливість моментально здійснити пошук необхідних елементів, виявити їхні кількісні та статистичні характеристики.

Сучасне корпусне мовознавство має у своєму дослідницькому арсеналі значну кількість корпусів різних типів і розмірів. Таке розмаїття наявних корпусів визначається широким спектром дослідницьких і прикладних завдань, для розв'язання яких вони створюються, та особливостями мовного матеріалу, покладеного в основу [4, с. 153–157]. Дослідники-лінгвісти можуть як створювати власні корпуси, а потім проводити необхідні дослідження на їх базі, так і використовувати загальнодоступні корпуси, створені іншими дослідниками чи колективами.

Для дослідження процесу перекладу в корпусній лінгвістиці використовується декілька видів корпусів, наприклад паралельний та порівняльний корпуси текстів.

Паралельний корпус текстів деякі дослідники інтерпретують також як перекладацький корпус (*parallel/translational corpus*), він складається з набору текстів мовою оригіналу та відповідних текстів мовою перекладу [5, с. 223–243]. Відмінною рисою паралельного корпусу є розмітка, яка зазвичай у великих обсягах корпусів виконується автоматичним способом. У процесі підготовки такого корпусу кожному реченню вихідної мови повинне відповідати речення мови перекладу, однак оскільки у процесі перекладу таке співвідношення в багатьох випадках порушується, то іноді варто вдаватися до ручного редагування, щоб не допустити розбіжність речень мовою оригіналу з реченнями мовою перекладу під час розмітки. Такий корпус дає можливість встановити перекладні відповідності між шуканими елементами корпусу, виявити їх частотність та представити список контекстів, у яких вживаються досліджувані одиниці.

Важливим видом корпусу для дослідження перекладу є також порівняльний корпус (*comparable corpus*). Цей тип корпусів включає в себе два різні набори текстів однією мовою: перший складається з оригінальних текстів, а другий – з перекладених текстів цією мовою [5]. Таким чином, порівняльний корпус – це насамперед одномовний корпус, що складається з двох підкорпусів однією мовою. Такий корпус дає змогу розкрити внутрішню структуру мови, адже він сприяє вичленуванню й опису одиниць за умов природного їх функціонування в мові.

С. Гренжер зазначає, що, незважаючи на явну перевагу порівняльного корпусу, представленого двома порівняними одномовними підкорпусами, такий вид корпусу має й недолік – складнощі виявлення параметрів зіставлення. Деякі види текстів мають культурні особливості, немає точних відповідників під час їх перекладу [6, с. 17–29]. Незважаючи на те, що порівняльний корпус текстів може не містити перекладні відповідники, такий вид корпусу вкрай корисний для контрастивних досліджень. Зіставлення оригінальних і перекладних текстів дає можливість виявити різні умови для опису функціонального та прагматичного використання елементів, що входять у корпус, а також може дати змогу розглядати одиниці мови, які мають значно більшу чи меншу частотність вживання в перекладних текстах, ніж в оригінальних.

Поява електронних корпусів суттєво змінила процеси перекладу та спосіб дослідження мовних явищ. Репрезентативні, збалансовані, значні за обсягом корпуси надають інформацію про динаміку процесів зміни лексичного складу мови, функціонування граматичних і синтаксичних одиниць, дають можливість здійснювати аналіз вживання мовних одиниць у різних стилях

і жанрах, різними авторами тощо. Корпуси надають не доступні для безпосереднього спостереження дані про реалізацію мовної системи, що своєю чергою відображається у словниках, граматиках та навчальних матеріалах. Якщо попередні корпусні дослідження зосереджувалися здебільшого лише на вживаності мовних одиниць (наприклад, частоті вживання певних іменників, дієслів і прикметників), то з появою потужних комп'ютерних технологій науковці отримали змогу визначити й аналізувати більш складні лексичні, граматичні та синтаксичні явища. Використання корпусу в лексикографічній практиці замістило паперові картки автоматичним пошуком слів у корпусі та суттєво переорієнтувало підхід до відбору слів до словника, замінивши відбір слів, зумовлений особистими перевагами лексикографа, більш об'єктивними критеріями, а саме даними про частоту та дистрибуцію слів у певних текстах чи жанрах. Більшість сучасних словників є «корпусоінформованими», а не «укладачеінформованими» [7, с. 33–71]. Одним з останніх нововведень у лексикографічній практиці стало розміщення у словниковій статті, крім традиційної інформації про значення та семантичні відношення одиниць, також інформації про її сполучуваність і дистрибуцію.

Необхідно підкреслити, що питання підвищення якості перекладу є одним із найбільш значущих у сучасному перекладознавстві в тому сенсі, що його вирішення дасть змогу перекладачам досягти успішності у своїй діяльності. Особливу актуальність це питання має під час письмового перекладу іноземною мовою, оскільки результат діяльності перекладача в цьому випадку (тобто текст перекладу іноземною мовою) призначений для сприйняття носієм цієї мови, який через свою мовну компетенцію бачить і помічає навіть найменші неточності й помилки, допущені в тексті. Отже, особливість перекладацької діяльності полягає в тому, щоб забезпечити такий тип міжмовної комунікації, коли текст, який створюється мовою реципієнта, міг би бути повноцінною комунікативною заміною оригіналу та ототожнюватися реципієнтами перекладу з оригіналом у функціональному, структурному і змістовому відношенні. Функціональне ототожнення оригіналу та перекладу полягає в тому, що переклад приписується авторові оригіналу, публікується під його іменем, обговорюється, цитується тощо, наче він і є оригіналом, тільки іншою мовою. Передбачається, що перекладач точно передає структуру та порядок викладу змісту в оригіналі, не дозволяє собі будь-що змінити, видалити або додати від себе [8, с. 43–44].

Варто зауважити, що основні проблеми, з якими перекладач стикається під час передачі

вихідного тексту, – це неспівпадіння кола значень, властивих одиницям мови оригіналу та мови перекладу. Найбільш поширеним випадком під час зіставлення лексичних одиниць двох мов є часткова відповідність, коли одному слову в мові оригіналу відповідає не один, а декілька семантичних еквівалентів у мові перекладу. Більшість слів будь-якої мови характеризується багатозначністю, адже система значень слова в одній мові, як правило, не співпадає повністю із системою значень слів в іншій мові. При цьому можуть спостерігатися різні випадки. Так, іноді коло значень слова в мові оригіналу виявляється ширшим, ніж у відповідного слова в мові перекладу, і навпаки. Тобто можлива ситуація, коли у слова в мові оригіналу (або в мові перекладу) є всі ті самі значення, що й у слова в мові перекладу (відповідно, мові оригіналу), однак у нього, крім того, є значення, які в іншій мові передаються іншими словами. Так, українське слово «характер», як і англійська лексема “character”, мають такі однакові значення: 1) «сукупність психічних особливостей людини»; 2) «тверда воля, завзятість у досягненні мети» (*Він людина без характеру – He has no character*); 3) «властивість, якість, своєрідність чого-небудь». Однак в англійського слова “character” є ще значення, яких немає в українській лексемі «характер» та які передаються в українській мові іншими словами, а саме: 4) «репутація»; 5) «письмова рекомендація, характеристика»; 6) «відмінна риса, ознака, якість»; 7) «фігура, особистість (часто дивна, оригінальна)»; 8) «літературний образ, герой, діюча особа у п’єсі»; 9) «печатний знак, буква, символ» (напр., *Chinese characters – китайські ієрогліфи*).

Усе вищевикладене означає, що під час вирішення тієї чи іншої перекладацької проблеми має місце звернення до зовнішніх інформаційних джерел.

**Висновки.** Звернемо увагу на те, що кожен перекладач вільний сам вирішувати, яким способом йому шукати чи перевіряти лексичні еквіваленти, уточнювати семантику, шукати граматичні відповідності, лексико-граматичну сполучність, а також чи перевіряти її взагалі. Однак ми вважаємо, що корпуси є більш зручним і достовірним засобом порівняно зі словниками з низки причин. По-перше, корпуси текстів (особливо так звані on-line) – це не застигли структури, як традиційні словники, а бази даних, які постійно поповнюються, завдяки чому перекладач отримує можливість бути в курсі останніх тенденцій у мові на основі аналізу вживання того чи іншого слова в корпусі. По-друге, корпус є більш репрезентативним джерелом слововживання, ніж словник, оскільки він набагато більший за обсягом, тому інформація про слово, яке можна отримати з корпусу, більш об’єктивна

й точна. По-третє, працювати з корпусом набагато зручніше, ніж зі словником: корпус розміщується на машині (комп’ютері), яка є частиною робочого місця перекладача, і шляхом нескладних маніпуляцій із клавіатурою та «мишею» перекладач уже через декілька секунд отримує інформацію, яка його цікавить.

Додамо, що ще однією важливою перевагою корпусу текстів є його реалістичність, адже до складу корпусу входять оригінальні, природні тексти, які виникають для задоволення потреб у координації діяльності людей у тій чи іншій референтній ситуації. Таким чином, застосування електронних корпусів текстів у процесі письмового перекладу іноземною мовою забезпечує високу якість перекладу за рахунок виключення неточностей і слабких місць, які викликають сумніви в перекладача.

Наостанок варто зазначити, що подальші дослідження в цій галузі допоможуть виокремити додаткові можливості корпусів текстів, а також визначити будь-які додаткові обмеження їх застосування. Звісно, зі збільшенням обсягів та вдосконаленням технічного оснащення дослідницький потенціал корпусів невпинно зростає, що робить їх цінним лінгвістичним ресурсом як для вузькопрофільних спеціалістів у галузі мовознавства, так і для всіх тих, хто зацікавлений у покращенні процесу перекладу та якості перекладацького продукту.

## ЛІТЕРАТУРА

1. Захаров В. П. Корпусная лингвистика : учебно-методическое пособие. Санкт-Петербург : СПбЕУ, 2005. 48 с.
2. McEnery T., Xiao R., Tono Y. Corpus-based Language Studies: an Advanced Resource Book. London; New York : Routledge, 2006. 386 p.
3. McEnery T., Hardie A. Corpus Linguistics: method, theory and practice. New York : Cambridge University Press, 2012. 292 p.
4. Демська-Кульчицька О. М. Дещо про класифікацію текстових корпусів. *Наукові записки Тернопільського національного педагогічного університету. Серія «Мовознавство»*. 2004. Вип. 1 (11). С. 153–157.
5. Baker M. Corpora in Translation Studies. An Overview and Suggestions for Future Research. *Target*. 1995. № 7 (2). P. 223–243.
6. Granger S. The Corpus Approach: a Common Way Forward for Contrastive Studies. *Corpus-based Approaches to Contrastive Linguistics and Translation Studies* / S. Granger, J. Lerot, S. Petch-Tyson (eds.). Amsterdam : Rodopi, 2003. P. 17–29.
7. McEnery T., Gabrielatos C. English Corpus Linguistics. *The Handbook of English Linguistics* /

B. Aarts, A. McMahon (eds.). Oxford : Blackwell, 2006. P. 33–71.

8. Комиссаров В. Н. Теория перевода (лингвистические аспекты) : учебник для институтов и факультетов иностранного языка. Москва : Высшая школа, 1990. 253 с.
9. Жуковська В. В. Вступ до корпусної лінгвістики : навчальний посібник. Житомир : Вид-во ЖДУ ім. І. Франка, 2013. 142 с.
10. Kennedy G. An Introduction to Corpus Linguistics. London; New York : Routledge, 2014. 328 p.

#### REFERENCES

1. Zakharov, V. P. (2005). Korpusnaya lingvistika : uchebno-metodicheskoe posobie [Corpus linguistics : educational and methodological manual]. Saint Petersburg : SPbEU. [in Russian]
2. McEnery, T., Xiao, R., Tono, Y. (2006). Corpus-based Language Studies: an Advanced Resource Book. London; New York : Routledge. [in English]
3. McEnery, T., Hardie, A. (2012). Corpus Linguistics: method, theory and practice. New York : Cambridge University Press. [in English]
4. Demaska-Kulchytska, O. M. (2004). Deshcho pro klasyfikatsiiu tekstovykh korpusiv [Something about the classification of text corpus]. *Naukovi zapysky Ternopilskoho natsionalnoho pedahohichnoho universytetu. Seriiia "Movoznavstvo" – Scientific notes of*

*the Ternopil National Pedagogical University. Series "Linguistics", iss. 1 (11), pp. 153–157. [in Ukrainian]*

5. Baker, M. (1995). Corpora in Translation Studies. An Overview and Suggestions for Future Research. *Target*, no. 7 (2), pp. 223–243. [in English]
6. Granger, S. (2003). The Corpus Approach: a Common Way Forward for Contrastive Studies. *Corpus-based Approaches to Contrastive Linguistics and Translation Studies / S. Granger, J. Lerot, S. Petch-Tyson* (eds.). Amsterdam: Rodopi, pp. 17–29. [in English]
7. McEnery, T., Gabrielatos, C. (2006). English Corpus Linguistics. *The Handbook of English Linguistics / B. Aarts, A. McMahon* (eds.). Oxford : Blackwell, pp. 33–71. [in English]
8. Komissarov, V. N. (1990). Teoriya perevoda (lingvisticheskie aspekty) : uchebnyk dlya institutov i fakul'tetov inostrannogo yazyka [Translation theory (linguistic aspects) : textbook for institutes and departments of a foreign language]. Moscow : Vysshaya shkola. [in Russian]
9. Zhukovska, V. V. (2013). Vstup do korpusnoi linhvistyky : navchalnyi posibnyk [Introduction to corpus linguistics : study guide]. Zhytomyr : Ivan Franko Zhytomyr State University Publishing House. [in Ukrainian]
10. Kennedy, G. (2014). An Introduction to Corpus Linguistics. London; New York : Routledge. [in English]