

DOI: <https://doi.org/10.32839/2304-5809/2020-79.1-17>

УДК 378.016:[811.161.2+821.161.2]

Скрипник Н.І.

Комунальний заклад вищої освіти  
«Вінницький гуманітарно-педагогічний коледж»

## ЗАСТОСУВАННЯ КОРПУСНИХ ТЕХНОЛОГІЙ У НАВЧАННІ МАЙБУТНІХ УЧИТЕЛІВ УКРАЇНСЬКОЇ МОВИ ТА ЛІТЕРАТУРИ

**Анотація.** У публікації окреслено основні проблеми визначення дидактичних термінів «корпусна лінгвістика», «корпусні технології», «корпус текстів». Зібрано ґрунтовний матеріал щодо наукових розвідок у цій царині. Розглянуто теоретичні засади корпусної лінгвістики. Обґрунтовано використання ресурсів корпусної лінгвістики в процесі підготовки фахівця з української мови та літератури. Сучасна корпусна лінгвістика тісно пов'язана із когнітивною лінгвістикою, які разом зорієнтовані на вивчення мовленнєвої діяльності. Ми вважаємо, що з появою електронних корпусів суттєво змінився й спосіб дослідження мовних фактів. Сучасні комп'ютерні технології дозволяють студентам визначати та аналізувати більш складні лексичні, граматичні та синтаксичні явища, а використання різних типів корпусів стає дедалі популярнішим у навчанні та вивченні мови.

**Ключові слова:** лінгводидактика, корпусна лінгвістика, корпусні технології, корпус текстів, методика викладання української мови, мовленнєві конструкції в аутентичних текстах.

Skrypnyk Nadiia

Communal Higher Education Institution  
«Vinnitsia Humanities Pedagogical College»

## APPLICATION OF CORPORATE TECHNOLOGIES IN TEACHING FUTURE TEACHERS IN UKRAINIAN LANGUAGE AND LITERATURE

**Summary.** The publication outlines the main problems of defining didactic terms «corpus linguistics», «corpus technologies», «corpus texts». Substantial material on scientific exploration in this field has been collected. The theoretical principles of corpus linguistics are considered. The use of corpus linguistics resources in the preparation of a specialist in Ukrainian language and literature is substantiated. Linguistics, which expounds the theoretical and practical aspects of language learning, operates today with a number of techniques, including traditional ones; classical (based on different types of speech activity); audio-lingual (contributes to the formation of not only spoken language, but also the ability to understand the text); intensive (dominance of different types of communication, trainings, webinars); linguocultural (through the culture of the people – to his language), the method of the 25th frame, emotional-semantic, etc. Each of these techniques has been tested over time and in some ways transformed and complemented by elements of the other. However, a common feature for them is that the primary means of language learning remains text. Units of all linguistic levels function in the text, and corpus-related techniques are inherent in the text, which are based on the use of computer technologies and the creation of electronic text-enclosures that can be used for linguistic-educational purposes. Modern corpus linguistics is closely linked to cognitive linguistics, which together are focused on the study of speech. V. Rykov rightly points out that corpus linguistics made it possible to clarify the results and conclusions drawn earlier, as well as to conduct new linguistic studies, more systematic and wider in volume of empirical material. The focus of corpus linguistics is the linguistic personality, that is, its speech activity, mass communication, the problem of its description. A very important feature of the corpus of texts is its representativeness, which is determined by phonetic, morphological, syntactic, stylistic parameters. We believe that with the advent of electronic enclosures, the way in which language facts are investigated has also changed significantly. Modern computer technologies allow students to identify and analyze more complex lexical, grammatical and syntactic phenomena, and the use of different types of corpora is becoming increasingly popular in language learning and learning.

**Keywords:** linguistics, corpus linguistics, corpus technologies, corpus of texts, teaching methodology of Ukrainian, speech constructions in authentic texts.

**Постановка проблеми.** На сьогодні зміст вищої педагогічної освіти, зокрема і філологічної, має бути зорієнтований на розвиток, не лише студентів університету, а й гуманітарно-педагогічного коледжу, здатності до творчості, самоорганізації, розробки й упровадження інноваційних технологій навчання. Система освіти України зазнає кардинальних змін, що спричинено її виходом до європейського та світового простору. Ці зміни стосуються всіх навчальних предметів освітнього процесу, натомість значною мірою вони охоплюють мовно-мовленнєву підготовку, а саме вивчення мови та літератури як засобів міжкультурного

спілкування. У зв'язку з цим виникає необхідність підготувати професійних фахівців нового типу, в нашому випадку, студента-філолога зі сформованими креативними, дослідницькими та мовленнєвими вміннями.

**Аналіз останніх досліджень і публікацій.** Аналіз останніх лінгводидактичних досліджень визначеної проблематики дає змогу виокремити наукові здобутки таких авторів як С. Бук [1], І. Гладка, П. Грабовий, Н. Дарчук, О. Демська-Кульчицька [8], В. Жуковська, Є. Карпіловська, А. Корольова, Л. Коцюк, Н. Леміш, С. Лендау, В. Риков, Н. Саєнко, В. Старко, М. Davies, D. Biber, D. Kies, T. Johns, R. Reppen та ін.

**Виокремлення мети та основних завдань досліджень.** Мета розвідки – обґрунтувати доцільність використання корпусів українських текстів як невід'ємного елементу професійних компетентностей майбутніх учителів та окреслити можливі шляхи дидактичного застосування корпусних ресурсів у закладі вищої освіти.

Досягнення поставленої мети передбачає виконання наступних завдань: 1) дати визначення понять «корпусна лінгвістика», «корпусні технології», «корпус текстів»; 2) окреслити історичний розвиток формування корпусної лінгвістики; 3) висвітлити теоретичні та практичні засади корпусної лінгвістики.

**Виклад основного матеріалу дослідження.** Сучасна корпусна лінгвістика є галуззю прикладного мовознавства, яка займається формуванням комп'ютерних корпусів текстів у різних мовах і спрямована на максимально об'єктивний аналіз мовних явищ в умовах реальної живої комунікації [12, с. 667]. Корпусна лінгвістика вивчає мову як соціальне явище, що описується заснованими на досвіді даними, тобто в певному мовленнєвому акті. У текстах мова проявляє себе як соціальний феномен, оскільки саме в них її можна записати, описати й проаналізувати. Для корпусних розвідок головним є постулат, що мова як об'єкт дослідження може бути вивчена лише у формі писемних та усних текстів. Більшість текстів зустрічаються у вигляді мовленнєвих актів, тобто як соціальна взаємодія між членами певного мовного суспільства [16, с. 112]. Центральною дефініцією корпусної лінгвістики є корпус мовленнєвої реалізації мови, що кваліфікується як сформована за певними вимогами вибірка мовленнєвого матеріалу, який може використовуватися для опису й дослідження мови як системи [12, с. 668].

Зазначимо, що термін «корпусна лінгвістика» ввійшов до наукового вжитку наприкінці ХХ ст. завдяки публікації у 1984 році збірника наукових праць «Corpus Linguistics: Recent Developments in the Use of Computer Corpora in English Language Research» за матеріалами конференції ICAME «Conference on the Use of Computer Corpora in English Language Research» у м. Неймеген (Голландія) у 1983 р. [13].

Під поняттям «корпус текстів» розуміють зібрання текстів певною мовою в електронному вигляді, відібраних за визначеним критерієм, яке певним чином впорядковане, але основною особливістю корпусу є його розмітка (анотування, маркування) [2, с. 209]. Це уніфікований, структурований, розмічений, філологічно компетентний масив мовних даних, створений для вирішення конкретних лінгвістичних завдань [9, с. 3].

У сучасному освітньому процесі вивчення української мови активно використовуються інноваційні методи і підходи. Одним із таких підходів є корпусний методологічний підхід, що має чимало переваг [14, с. 9]. Текстовий корпус покликаний надавати дані користувачам, які б вони не змогли отримати з традиційних підручників і посібників. Існування корпусів текстів дає можливість значно розширити й автоматизувати аналіз мовного матеріалу, який є найважливішою базою будь-якого лінгвістично-методичного дослідження. Чим більше матеріалів аналізу-

ється, тим вища значущість висновків і рівень їх достовірності. Сучасні комп'ютерні програми дозволяють знаходити потрібні приклади з корпусів текстів, які зберігаються в електронному вигляді на комп'ютері. Це економить значну кількість часу в порівнянні з традиційною технологією збору прикладів вручну.

Мовні корпуси базуються на текстах, відібраних з «реальних» автентичних джерел: книг, газет, журналів, інтерв'ю, промов, отожд подаючи мову у «природньому» середовищі її функціонування. Важливим є не те, які тексти використовуються (фантастика, документальна проза, журнали, наукові статті, українська література) або інші акти мовленнєвої комунікації (розмови, телефонні дзвінки), а їх аутентичність.

Прикладне мовознавство має у своєму дослідницькому арсеналі значну кількість корпусів різних типів і розмірів на різних мовах. Багато з них наявні в мережі Інтернет у вільному доступі. Так, наприклад, в Україні корпус текстів української мови розроблений співробітниками лабораторії комп'ютерної лінгвістики Інституту філології Київського національного університету імені Тараса Шевченка під керівництвом Н.П. Дарчук. У корпусі зберігаються тексти, опрацьовані автоматичним лінгвістичним аналізатором. Це означає, що кожній одиниці тексту (морфемі, слову, словосполученню, реченню) приписана певна супровідна інформація: частини мовна належність, граматична форма, синтаксична функція, контекст тощо [10].

Існує також Генеральний регіонально анотований корпус (ГРАК) української мови обсягом понад 235 млн слівовживань, де здійснено розмітку текстів за регіонами – це велика, репрезентативна, структурована колекція текстів українською мовою у супроводі програми, яка дозволяє будувати на базі корпусу власні підкорпуси, шукати слова, граматичні форми та їх сполучення, а також обробляти результати пошуку, сортувати, робити збалансовані вибірки і одержувати різну статистичну інформацію. Корпус охоплює період з 1818 по 2018 р. і містить понад 25 тисяч текстів близько 3850 авторів [3].

Однією з найважливіших якісних ознак будь-якого корпусу є його здатність коректно відображати предметну галузь чи той мовний рівень, що його задекларовано корпусом. Під предметною галуззю розуміється рівень реалізації мовної системи, яка містить феномени, що підлягають лінгвістичному описові [7, с. 101–102].

Корпуси як джерела емпіричних даних відіграють вагомий роль у підготовці фахових навчальних матеріалів. Адже емпіричний процес надає змогу вивчати зразки мовленнєвих конструкцій в автентичних текстах. Можливість опрацювання величезної кількості емпіричних мовних даних відкриває величезні перспективи: створення спеціалізованих корпусів та корпусів текстів [5, с. 224].

На базі мовознавчих корпусів можна сформувати списки автентичної активної лексики студентів, списки термінів для використання у професійному вживанні.

Текстовий корпус відкриває широкі можливості його застосування, оскільки кожен студент-словесник може вільно використовувати назва-

ний ресурс як у стінах навчального закладу, виконуючи завдання під час лабораторних чи практичних занять, так і за його межами, готуючи домашнє завдання, ІНДЗ чи творчі проекти.

Завдяки використанню корпусу мови студенти мають можливість знаходити слова та вирази, притаманні тому чи іншому стилю та тематиці тексту, а також граматичні конструкції, які використовуються для реалізації тої чи іншої комунікативної функції. Такий підхід допомагає студенту-філологу оволодіти автентичними лексичними та граматичними конструкціями. У свою чергу, викладачі мають змогу створювати завдання, необхідні для досягнення дидактичних цілей, використовувати матеріали при укладанні підручників та посібників з української мови, забезпечуючи реалізацію принципів автентичності та частотності мовного матеріалу.

За допомогою різних вправ студенти вчать спостерігати за даними лінгвістичного корпусу, висувати гіпотези та самостійно формувати правила вживання тих чи інших одиниць (індуктивний підхід) або перевіряти на даних корпусу дійсність правил з навчальних граматик чи підручників (дедуктивний підхід). Через застосування корпусних технологій студенти набувають (або принаймні удосконалюють) такі важливі вміння як уміння прогнозувати, спостерігати, помічати, думати, міркувати, аналізувати, інтерпретувати, рефлексувати, досліджувати, робити висновки (індуктивно чи дедуктивно), фокусуватися, здогадуватися, порівнювати, диференціювати, теоретизувати, висувати гіпотези та перевіряти їх [15, с. 277].

За дидактичними дослідженнями П. Грабового, шляхи використання ресурсів корпусної лінгвістики можемо умовно поділити на такі групи:

1. Використання в освітньому процесі при здійсненні власної професійної підготовки (навчання лексики, граматики, аудіювання, читання, письма тощо) та при підготовці навчально-методичних матеріалів, підручників, посібників з української мови, у лексикографічній практиці.

2. Використання при викладанні спеціалізованих предметів для формування професійних умінь та навичок («Історія української мови», «Українська діалектологія», «Лексикологія», «Стилістика», «Культура мовлення», «Риторика», «Літературна освіта в профільній школі» та ін.).

3. В організації дослідницької діяльності студентів (написання курсових та випускних кваліфікаційних робіт різних рівнів) [4].

**Висновки та перспективи подальших досліджень.** Підсумовуючи все вище викладене, слід зазначити, що на сьогодні в Україні простежується позитивна динаміка щодо застосування корпусних технологій у навчанні студентів-словесників. Такий спосіб вивчення наукового мовлення сприяє розширенню словникового запасу студентів та підвищує рівень опанування ними різними мовленнєвими моделями.

Підготовка майбутніх фахівців з української мови та літератури у гуманітарно-педагогічному коледжі націлена на формування висококваліфікованих спеціалістів, які мають ґрунтовні знання з предмета та здатні до власного креативного пошуку.

## Список літератури:

1. Бук С. Корпус текстів Івана Франка: спроба визначення основних параметрів. *Прикладна лінгвістика та лінгвістичні технології: MegaLing-2006 : Зб. наук. пр.* / НАН України. Укр. мовн.-інформ. фонд, Таврійськ. нац. ун-т ім. В.І. Вернадського; за ред. В.А. Широкова. Київ : Довіра, 2007. С. 72–82.
2. Войтко Г. Корпусна лінгвістика: історія формування і перспективи розвитку. *Науковий вісник Східноєвропейського національного університету імені Лесі Українки*. 2014. Вип. 5. С. 208–212.
3. Генеральний регіонально анований корпус української мови (ГРАК). Київ, Осло, Йена, 2017–2018. URL: <http://uacorporus.org>
4. Грабовий П. Використання ресурсів корпусної лінгвістики в процесі підготовки фахівця з англійської мови. *Збірник наукових праць. Частина 1*. 2013. С. 48–54.
5. Данилюк І. Корпус текстів для вивчення граматичної службовості. *Лінгвістичні студії*. 2013. Вип. 26. С. 224–229.
6. Данчевська Ю. Корпуси текстів у лінгводидактиці: здобутки та перспективи. *Нова педагогічна думка*. 2014. № 1. С. 58–60.
7. Демська-Кульчицька О. Деяко про класифікацію текстових корпусів. *Наукові записки. Серія : Мовознавство*. 2004. 1(11). С. 153–157.
8. Демська-Кульчицька О. Основи національного корпусу української мови : монографія. Київ : Інститут української мови НАН України, 2005. 218 с.
9. Захаров В.П. Корпусная лингвистика : Учебно-метод. пособие. СПб., 2005. 48 с.
10. Корпус української мови. URL: <http://www.mova.info/corpus.aspx>
11. Рыков В.В. Корпусная лингвистика. 2002. URL: <http://rykov-cl.narod.ru/c.html>
12. Селіванова О.О. Сучасна лінгвістика: напрями і проблеми : підручник. Полтава : Довкілля-К, 2008. 712 с.
13. *Corpus Linguistics: Recent Developments in the Use of Computer Corpora in English Language Research* / ed. by J. Aarts and W. Meijs. Amsterdam : Rodopi, 1984. 229 p.
14. Gena B. Using Corpora in the Language Learning Classroom: *Corpus Linguistics for Teachers*. University of Michigan Press ELT, 2010. P. 9–11.
15. O'Sullivan I. Enhancing a process-oriented approach to literacy and language learning: the role of corpus consultation literacy. *ReCALL* 19(3). 2007. P. 269–286.
16. Teubert W. *Corpus linguistics and lexicography. Text Corpora and Multilingual Lexicography*. Amsterdam/Philadelphia : John Benjamins Publishing Company, 2007. P. 109–134.

## References:

1. Buk, S. (2007). Korpus tekstiv Ivana Franka: spробa vyznachennia osnovnykh parametriv [The body of Ivan Franko texts: an attempt to determine the basic parameters]. *Prykladna linhvistyka ta linhvistychni tekhnolohii: MegaLing 2006: Zb. nauk. pr.* [Applied Linguistics and Linguistic Technologies: MegaLing 2006] / NAN Ukrainy. Ukr. movn.-inform. fond, Tavriisk. nats. un-t im. V.I. Vernadskoho; za red. V.A. Shyrokov. Kyiv: Dovira, p. 72–82. (in Ukrainian)

2. Voitko, H. (2014). Korpusna linhvistyka: istoriia formuvannia i perspektyvy rozvytku [Corpus Linguistics: History of Formation and Prospects for Development]. *Scientific Bulletin of Lesya Ukrainka Eastern European National University*, vol. 5, pp. 208–212. (in Ukrainian)
3. (2017–2018). Heneralnyi rehionalno anotovanyi korpus ukrainskoi movy (HRAK) [General Regional Annotated Corps of the Ukrainian Language (HRAK)]. Kyiv, Oslo, Yena. URL: <http://uacorpora.org> (in Ukrainian)
4. Hrabovyi, P. (2013). Vykorystannia resursiv korpusnoi linhvistyky v protsesi pidhotovky fakhivtsia z anhliiskoi movy [The use of corpus linguistics resources in the preparation of an English specialist]. *Collection of scientific works*. Part 1, pp. 48–54. (in Ukrainian)
5. Danyliuk, I. (2013). Korpus tekstiv dlia vyvchennia hramatychnoi sluzhbovosti [A body of texts for the study of grammatical servitude]. *Linguistic Studios*, vol. 26, pp. 224–229. (in Ukrainian)
6. Danchevska, Yu. (2014). Korpusy tekstiv u linhvodydaktytsi: zdobutky ta perspektyvy [Text corpora in linguodidactics: achievements and perspectives]. *New pedagogical thought*, no. 1, pp. 58–60. (in Ukrainian)
7. Demaska-Kulchytska, O. (2004). Doshcho pro klasyfikatsiiu tekstovykh korpusiv [Something about the classification of text corpora]. *Proceedings. Series: Linguistics*, no. 1(11), pp. 153–157. (in Ukrainian)
8. Demaska-Kulchytska, O. (2005). *Osnovy natsionalnoho korpusu ukrainskoi movy: monohrafiia* [Fundamentals of the National Corpus of the Ukrainian Language: Monograph]. Kyiv: Instytut ukrainskoi movy HAH Ukrainy, 218 p. (in Ukrainian)
9. Zakharov, V.P. (2005). *Korpusnaia linhvistyka: Uchebno–metod. posobye* [Corpus linguistics: A textbook]. SPb., 48 p. (in Ukrainian)
10. Korpus ukrainskoi movy [Corps of Ukrainian Language]. URL: <http://www.mova.info/corpus.aspx> (in Ukrainian)
11. Ryikov, V.V. (2002). *Korpusnaya lingvistika* [Corpus linguistics]. URL: <http://rykov-cl.narod.ru/c.html> (in Russian)
12. Selivanova, O.O. (2008). *Suchasna linhvistyka: napriamy i problemy: pidruchnyk* [Modern linguistics: directions and problems: a textbook]. Poltava : Dovkillia-K, 712 p. (in Ukrainian)
13. J. Aarts and W. Meijs (ed). (1984). *Corpus Linguistics: Recent Developments in the Use of Computer Corpora in English Language Research*. Amsterdam: Rodopi, 229 p.
14. Gena, B. (2010). Using Corpora in the Language Learning Classroom: Corpus Linguistics for Teachers. University of Michigan Press ELT, pp. 9–11.
15. O’Sullivan, I. (2007). Enhancing a process-oriented approach to literacy and language learning: the role of corpus consultation literacy. *ReCALL* 19(3), pp. 269–286.
16. Teubert, W. (2007). *Corpus linguistics and lexicography. Text Corpora and Multilingual Lexicography*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 109–134.