

ЛІНГВОДИДАКТИКА

УДК 811.161.2:81'34

О. Ю. Плахотнікова

ДИДАКТИЧНИЙ ПОТЕНЦІАЛ КОРПУСУ ТРАНСКРИБОВАНОГО УКРАЇНСЬКОГО УСНОГО МОВЛЕННЯ

У статті продемонстровано основні практичні можливості Корпусу українського транскрибованого усного мовлення для дослідження фонетичних особливостей українського мовлення та обґрунтовано доцільність використання цього Корпусу в навчальному процесі (у викладанні курсу української мови для студентів-філологів). З'ясовано, що навчання студентів на основі Корпусу дало їм змогу отримати й закріпити практичні навички роботи з аудіозаписами усного мовлення та комп'ютерними програмами Praat і ELAN, а також поглибити й практично застосувати теоретичні знання з української фонетики з метою набуття фахових навичок науково-дослідницької роботи.

Ключові слова: лінгводидактика, українське мовлення, корпусна лінгвістика, корпус усного мовлення, Корпус українського транскрибованого усного мовлення, навчання на основі даних.

Plakhotnikova O. Didactic Potential of the Corpus of Transcribed Ukrainian Speech. The article deals with The Corpus of the Transcribed Ukrainian speech and its possible use in teaching Ukrainian language to philology students. The relevance of the research is determined by the fact that in modern Ukrainian higher education computer-based teaching methods and interactive linguistic corpora are being introduced which enable to conduct the speech research. The aim of the study is to justify the usefulness of The Corpus of the Transcribed Ukrainian speech in teaching Ukrainian language to philology students. The corpus-based studies were introduced into the training internship on experimental phonetics of the second-year philology students (specialty "Applied linguistics and English"), and into special courses for Master's students of the Institute of Philology, Taras Shevchenko National University of Kyiv (specialty "Ukrainian language and literature, English language"), as well as for Master's students of Kyiv National Linguistic University (specialty "Applied linguistics"). It turned out that teaching the philology students on the basis of the Corpus of the Transcribed Speech allowed them to acquire and consolidate practical skills in working with speech audio recordings, and computer programmes Praat and ELAN. These corpus-based studies also helped the students to deepen and apply their theoretical knowledge of Ukrainian phonetics in order to acquire professional skills of

scientific and research work. Consequently, we can make a conclusion that the above-mentioned Corpus can be used in the following fields: 1) in Ukrainian language teaching (in courses of Ukrainian phonetics and orthoepy, such as Computer Tools for Transcription and Transliteration; Applied Phonetics; Analysis and Synthesis of Speech; Anthropophonetics and Psychophonetics: Interaction Aspects; 2) in scientific phonetic studies of Ukrainian standard speech.

Key words: *language teaching, Ukrainian speech, corpus, linguistics, speech corpus, Corpus of the Transcribed Ukrainian Speech, data-driven learning.*

Вступ

Сучасний стан розвитку інформаційних технологій сприяє залученню в навчальний процес інноваційних форм та методів. Зміни в системі освіти України в контексті сучасних європейських і світових тенденцій потребують оновлення методик викладання навчальних предметів, зокрема й української мови. На сьогодні корпусний підхід до вивчення мовних засобів є одним із провідних, він дає змогу розширити можливості власне наукового пошуку й аналізувати значний обсяг мовного матеріалу, а також відкриває нові перспективи для застосування корпусів текстів з дидактичною метою в процесі навчання / вивчення мови.

Актуальність проблематики статті зумовлена потребою модернізації освітньої системи України й, відповідно, якісними змінами вимог до набутих знань, навичок і фахової компетенції здобувачів вищої освіти за напрямом підготовки «Філологія». Дієвим засобом модернізації української освіти є впровадження в навчальний процес ресурсів корпусної лінгвістики та навчання на основі корпусних даних.

Сучасна лінгводидактика послуговується кількома методиками вивчення (навчання) мови, а в основі тієї чи тієї методики — сукупність методів (когнітивних, креативних, організаційних або комунікативних). Текст (усний чи писемний) є основним засобом навчання мови, оскільки в ньому функціонують одиниці всіх мовних рівнів. Саме з текстом пов'язані корпусні методики вивчення мови, які на сьогодні активно використовують у лінгводидактиці (Кочан, 2008: 15). Найбільше досліджень і здобутків у використанні корпусів у лінгводидактиці належить саме американським ученим (корпусна лінгвістика зародилася в США), а також науковцям з Великобританії (Данчевська, 2014: 58). Проблемі використання корпусних ресурсів

у методиці викладання мови присвячено багато лінгвістичних праць, зокрема й в українському мовознавстві. Ідеться про статті С. Бук (2007), І. Кочан (2008), П. Грабового (2013), Ю. Данчевської (2014), В. Жуковської (2018), А. Синявської та Т. Лісної (2019), Н. Скрипник (2020) та ін.

За спостереженнями П. Грабового, у закладах вищої освіти ресурси корпусної лінгвістики можна використовувати:

1) у навчальному процесі під час здійснення мовної підготовки (вивчення лексики, граматики тощо) та в процесі підготовки навчально-методичних матеріалів до курсів;

2) у викладанні спеціалізованих предметів та спецкурсів для формування професійних навичок і вмінь («Історія мови», «Діалектологія», «Стилістика» тощо);

3) в організації дослідницької діяльності студентів (написання курсових і кваліфікаційних робіт) (Грабовий, 2013: 52).

Окрім розробок навчально-методичних і дидактичних матеріалів та добору ілюстративних прикладів з корпусів, корпуси текстів можуть бути використані як викладачами для підготовки до занять, так і студентами під час аудиторних занять. Сьогодні цей метод, що передбачає взаємодію між студентом та корпусом або між викладачем і корпусом, відомий під назвою «навчання на основі даних» (з англ. – data-driven learning (Данчевська, 2014: 59)). Таке навчання зумовлює розробку вправ та завдань з використанням даних корпусу, і тут розрізняють два підходи: прямий (студенти самостійно аналізують корпусний матеріал за допомогою програмного забезпечення) і непрямий (студенти ознайомлюються з прикладами, дібраними з корпусу). Звісно, навчання на основі даних має свої переваги й недоліки; серед переваг — можливість вивчення природного мовлення, а серед недоліків прямого підходу — труднощі роботи з інтерфейсом корпусу, неправильна інтерпретація студентами досліджуваного матеріалу тощо (Жуковська, 2018).

Мета статті — розглянути практичні можливості застосування Корпусу українського транскрибованого усного мовлення (далі — Корпусу) в навчальному процесі. Завдання дослідження — обґрунтувати доцільність використання Корпусу у викладанні курсу української мови для студентів-філологів; провести аналіз дидактичних можливостей Корпусу, виявлених у викладанні курсу української

мови; продемонструвати основні переваги використання корпусних ресурсів у навчальному процесі.

Методи дослідження

Для розв'язання поставлених у роботі завдань і досягнення основної мети дослідження застосовано загальнонаукові методи. На основі описового методу й методу узагальнення здійснено огляд стану проблеми та спостереження над фактичним матеріалом. За допомогою аналітичного методу проведено аналіз можливостей використання Корпусу у викладанні курсу української мови в закладах вищої освіти.

Виклад основного матеріалу

Звичайно, корпусна методика навчання мови в процесі підготовки фахівця-філолога в Україні ще перебуває на етапі становлення. Проте впровадження до навчального процесу ресурсів корпусної лінгвістики передбачає досягнення якісно нового рівня підготовки за напрямом «Філологія» (Грабовий, 2013: 50).

На сьогодні в Україні вже розроблено декілька *корпусів писемного й усного українського мовлення* (Корпус української мови, Генеральний регіонально анотований корпус української мови, Акустичний корпус українського ефірного мовлення), але необхідна більша кількість фонетично анотованих баз даних усного мовлення, що дають можливість опрацьовувати звукову реалізацію записаних текстів у відкритому доступі. Запропонований Корпус задовольняє потребу у фонетично анотованих відкритих корпусах усного українського мовлення для здійснення спеціальних фонетичних досліджень і для використання цих корпусів з навчальною метою.

Корпус українського транскрибованого усного мовлення розроблено О. Плахотніковою на базі навчальної лабораторії експериментальної фонетики Інституту філології Київського національного університету імені Тараса Шевченка. Корпусне дослідження українського усного мовлення стало джерельною базою дисертаційної роботи О. Плахотнікової «Корпус українського усного мовлення: теоретичні засади побудови й основи практичного втілення» (Плахотнікова, 2017b). Матеріалом Корпусу слугували аудіотексти читаного й спонтанного українського природного монологічного усного мовлення. На сьогодні Корпус містить фрагменти аудіозаписів загальною

тривалістю 93 хв. 6 сек. (28 аудіозаписів монологічного мовлення у *.wav-форматі, записаних 10 дикторами). Для створення анотаційних файлів аудіозаписів Корпусу використано комп'ютерну програму ELAN, розроблену в Інституті психолінгвістики імені Макса Планка (The Language Archive, м. Неймеген, Нідерланди). Ця програма апробована у світовій науковій практиці: на її основі вже створено чимало мультимедійних корпусів (Плахотнікова, 2017а). Корпус українського транскрибованого усного мовлення є фонетично анотованим і містить індивідуалізовані транскрипційні записи аудіотекстів, адаптовані до комп'ютерної програми ELAN, з використанням 1) спрощеної фонетичної індивідуалізованої транскрипції на основі кириличної графіки; 2) транскрипції за стандартами Міжнародного фонетичного алфавіту (ISO 15924) — спрощеної фонетичної індивідуалізованої транскрипції на основі латинської графіки. Анотаційні файли Корпусу, розроблені на базі комп'ютерної програми ELAN, пристосовані до пошукових можливостей цієї програми. Опції пошуку задовольняють основні потреби досліджень фонетичної специфіки українського мовлення: забезпечено пошук усіх основних фонетичних одиниць у кириличній та латинській графіці — від алофона до цілої синтагми (Плахотнікова, 2017b: 3-4).

На сьогодні Корпус українського транскрибованого усного мовлення *доступний онлайн* для широкого кола користувачів у мережі Інтернет, адже частина корпусних матеріалів (17 анотованих аудіозаписів тривалістю 60 хв. 28 сек.) представлена на сайті The Language Archive Інституту психолінгвістики імені Макса Планка (м. Неймеген, Нідерланди (Plakhotnikova)). Ці 17 анотаційних файлів також містять переклад орфографічного запису аудіотекстів англійською мовою, необхідний для використання Корпусу іноземцями, що не володіють українською мовою. Отже, відкритий фонетично анотований корпус українського мовлення дає можливість досліджувати фонетичні явища сучасного українського мовлення та порівнювати переклад синтагм англійською мовою з оригінальним українським текстом.

Розроблену методику сегментації й аналізу мовленнєвого сигналу на базі комп'ютерних програм ELAN і Praat було апробовано в навчальній практиці з експериментальної фонетики студентів 2 курсу спеціальності «Прикладна лінгвістика та англійська мова» Інституту філології Київського національного університету імені Тараса

Шевченка. Вищевказана методика була вміщена в програму навчальної практики для студентів II курсу спеціальності «Прикладна лінгвістика та англійська мова» (Зубань, Дудник, Бас-Кононенко, & Плахотнікова, 2016). Також навчання на основі даних Корпусу було впроваджено в спецкурси «Аналіз і синтез усного мовлення» студентів-магістрантів Інституту філології Київського національного університету імені Тараса Шевченка (спеціальність «Українська мова і література, іноземна мова») та «Комп'ютерні засоби транскрипції і транслітерації» студентів-магістрантів Київського національного лінгвістичного університету (спеціальність «Прикладна лінгвістика»).

Навчальна практика студентів-філологів 2 курсу спеціальності «Прикладна лінгвістика та англійська мова» Інституту філології передбачала роботу з Корпусом українського транскрибованого усного мовлення після проходження навчальних курсів «Прикладна фонетика», «Антропофонетика і психофонетика: аспекти взаємодії» та «Комп'ютерні засоби транскрипції і транслітерації» з метою поглиблення і практичного застосування теоретичних знань з української фонетики, а також набуття студентами фахових навичок науково-дослідницької роботи. Студенти-філологи мали змогу отримати й закріпити практичні навички роботи з аудіозаписами усного мовлення та комп'ютерними програмами Praat і ELAN; оволодіти методиками фонетичного аналізу мовленнєвого сигналу та опису явищ сегментного й просодичного рівнів мовлення дикторів; сформувати вміння сегментувати акустичний сигнал та анотувати сегменти усного мовлення за акустичними параметрами (Зубань, Дудник, Бас-Кононенко, & Плахотнікова, 2016: 5-6). Студенти-філологи змогли на власному досвіді простежити повний цикл створення анотаційного файлу для аудіозапису Корпусу — від сегментації акустичного сигналу на фрази й до створення спрощеної фонетичної індивідуалізованої транскрипції на основі кириличної графіки, а також спрощеної фонетичної індивідуалізованої транскрипції за стандартами Міжнародного фонетичного алфавіту (на основі латинської графіки).

Отже, науково-дослідна робота студентів-філологів з Корпусом українського транскрибованого усного мовлення передбачала проходження таких етапів:

- 1) ознайомлення з роботою комп'ютерної програми ELAN;

2) створення анотаційного файлу конкретного аудіозапису диктора з розширенням .eaf у програмі ELAN; цей етап передбачав визначення типів і структури рівнів анотаційного файлу; сегментацію аудіотексту на фрази на основі фізичних пауз; ознайомлення з протоколом використання транскрипційних символів і подальше створення транскрипційних записів у режимі транскрипції та введення орфографічного запису тексту в режимі розмітки;

3) створення характеристики мовлення диктора із зазначенням типологічних та індивідуальних особливостей його мовлення, а також оформлення звіту на основі вказаної інформації.

Висновки

Отже, Корпус українського транскрибованого усного мовлення є відкритим інтерактивним інтернет-ресурсом, доступним для широкої аудиторії мережі Інтернет, і використовується як з навчальною метою (у курсі вивчення української фонетики), так і для подальших лінгвістичних досліджень фонетичної варіативності сучасного українського мовлення. Перевагою цього корпусу є те, що він містить орфографічні записи усних текстів українською та англійською мовами, а це значно розширює його дидактичні можливості. Таким чином, використання Корпусу українського транскрибованого усного мовлення в навчальному процесі дає змогу вдосконалити спеціалізовану лінгвістичну підготовку студентів-філологів, а також забезпечити сприятливі умови для організації науково-дослідної роботи студентів.

Перспективою подальших досліджень є створення так званого «учнівського підкорпусу» (learner subcorpus, 'підкорпусу, що містить мовні дані осіб, які вивчають цю мову як іноземну' (Бук, 2007:20)) українського мовлення студентів-іноземців в межах Корпусу українського транскрибованого усного мовлення для аналізу особливостей засвоєння студентами іноземної мови й, відповідно, застосування цього ресурсу у викладанні української мови як іноземної.

ЛІТЕРАТУРА

1. Бук, С. (2007). Учнівські корпуси в методиці викладання іноземної мови. *Теорія і практика викладання української мови як іноземної*, 2, 19–23.
2. Грабовий, П. (2013). Використання ресурсів корпусної лінгвістики в процесі підготовки фахівця з англійської мови. *Зб. наук. пр. Уман. держ. пед. ун-ту імені Павла Тичини*, 1, 48–54.
3. Данчевська, Ю. (2014) Корпуси текстів у лінгводидактиці: здобутки та перспективи.

Нова педагогічна думка, 1, 58–60. 4. Жуковська, В. В. (2018). Застосування корпусних технологій у навчанні та вивченні іноземної мови. *Актуальні проблеми сучасної лінгвістики та методики викладання мови і літератури* (с. 39–50). Житомир: Вид-во ЖДУ ім. Івана Франка. 5. Зубань, О. М., Дудник, З. В., Бас-Кононенко, О. В., & Плахотнікова, О. Ю. (2016). *Програма навчальної практики для студентів II курсу спеціальності «Прикладна лінгвістика та англійська мова»*. Київ: ВЦ «Київ. ун-т». 6. Кочан, І. (2008). Нові освітні технології в практиці викладання української мови як іноземної. *Теорія і практика викладання української мови як іноземної*, 3, 14–20. 7. Плахотнікова, О. Ю. (2017). Корпус українського транскрибованого усного мовлення: засади створення. *Наук. вісн. Дрогоб. держ. пед. ун-ту імені Івана Франка. Серія «Філологічні науки»*. Мовознавство, 7, 145–148. 8. Плахотнікова, О. Ю. (2017). Корпус українського усного мовлення: теоретичні засади побудови й основи практичного втілення. (Дис. ... канд. філол. наук). Київ: КНУ імені Тараса Шевченка. 9. Сиявська, А., & Лісна, Т. (2019). Дидактичний аспект використання лінгвістичних корпусів у процесі навчання іноземної мови. *Теоретичні й прикладні проблеми сучасної філології*, 9 (Ч. 2), 66–73. Слов'янськ: Вид-во Б. І. Маторіна. 10. Скрипник, Н. І. (2020). Застосування корпусних технологій у навчанні майбутніх учителів української мови та літератури. *Молодий вчений*, 3.1, 67–70. <https://doi.org/10.32839/2304-5809/2020-79.1-17>. 11. Plakhotnikova, O. Corpus of the Transcribed Ukrainian Speech. *The Language Archive. MPI Nijmegen*. Retrieved from https://archive.mpi.nl/tla/islandora/object/tla%3A1839_00_1390CD7C_579F_45F7_9B00_65A31FC980DF.

REFERENCES

1. Buk, S. (2007). Uchnivski korpusy v metodytsi vykladannia inozemnoi movy [Learner corpora in the techniques of teaching a foreign language]. *Teoriia i praktyka vykladannia ukrainskoi movy yak inozemnoi — Theory and practice of teaching Ukrainian as a foreign language*, 2, 19–23 [in Ukrainian]. 2. Hrabovyi, P. (2013). Vykorystannia resursiv korpusnoi linhvistyky v protsesi pidhotovky fakhivtsia z anhliiskoi movy [Using the resources of corpus linguistics in teaching for English Majors]. *Zb. nauk. pr. Uman. derzh. ped. un-tu imeni Pavla Tychyny — Collection of scientific works of Uman State Pedagogical University named after Pavel Tychyna*, 1, 48–54 [in Ukrainian]. 3. Danchevska, Yu. (2014). Korpusy tekstiv u linhvodydaktytsi: zdobutky ta perspektyvy [Text corpora in language teaching: achievements and perspectives]. *Nova pedahohichna dumka — New pedagogical thought*, 1, 58–60 [in Ukrainian]. 4. Zhukovska, V. V. (2018). Zastosuvannia korpusnykh tekhnolohii u navchanni ta vyvchenni inozemnoi movy [Application of corpus technologies in teaching and learning a foreign language]. *Aktualni problemy suchasnoi linhvistyky ta metodyky vykladannia movy i literatury — Current problems of modern linguistics and methods of teaching language and literature* (pp. 39–50). Zhytomyr: Vyd-vo ZhDU im. Ivana Franka [in Ukrainian]. 5. Zuban, O. M., Dudnyk, Z. V., Bas-Kononenko, O. V., & Plakhotnikova, O. Yu (2016). *Prohrama navchalnoi praktyky dlia studentiv II kursu spetsialnosti «Prykladna linhvistyka ta anhliiska mova»* [The program of training internship for the 2nd course students of specialty «Applied linguistics and English language»]. Kyiv: «Kyiv. un-t» [in Ukrainian]. 6. Kochan, I. (2008). Novi osvitni tekhnolohii v praktytsi vykladannia ukrainskoi movy yak inozemnoi [New educational technologies in the practice of teaching ukrainian as a foreign language]. *Teoriia i praktyka vykladannia*

ukrainskoi movy yak inozemnoi — Theory and practice of teaching Ukrainian as a foreign language, 3, 14–20 [in Ukrainian]. 7. **Plakhotnikova, O. Yu.** (2017). Korpus ukrainskoho transkrybovanoho usnogo movlennia: zasady stvorennia [The Corpus of Transcribed Ukrainian Speech: principles of construction]. *Nauk. visn. Drohob. derzh. ped. un-tu imeni Ivana Franka. Seria: «Filolohichni nauky»*. *Movoznavstvo — Scientific Bulletin of Drohobych Ivan Franko State Pedagogical University. Philological Sciences Series. Linguistics*, 7, 145–148 [in Ukrainian]. 8. **Plakhotnikova, O. Yu.** (2017). Korpus ukrainskoho usnogo movlennia: teoretychni zasady pobudovy y osnovy praktychnoho vtilennia [Ukrainian speech corpus: the theoretical basis of construction and practical implementation]. *Candidate's thesis*. Kyiv: KNU imeni Tarasa Shevchenka [in Ukrainian]. 9. **Syniavska, A., & Lisna, T.** (2019). Dydaktychnyi aspekt vykorystannia linhvistychnykh korpusiv u protsesi navchannia inozemnoi movy [Didactic aspect of linguistic corpora usage in the process of foreign language studies]. *Teoretychni y prykladni problemy suchasnoi filolohii — Theoretical and applied problems of modern philology*, 9 (Ch. 2), 66–73. Sloviansk: Vyd-vo B. I. Matorina [in Ukrainian]. 10. **Skrypnyk, N. I.** (2020). Zastosuvannia korpusnykh tekhnolohii u navchanni maibutnikh uchyteliv ukrainskoi movy ta literatury [Application of corpus technologies in teaching future teachers of Ukrainian language and literature]. *Molodyi vchenyi — A young scientist*, 3.1, 67–70. <https://doi.org/10.32839/2304-5809/2020-79.1-17> [in Ukrainian]. 11. **Plakhotnikova, O.** *Corpus of the Transcribed Ukrainian Speech. The Language Archive. MPI Nijmegen*. Retrieved from https://archive.mpi.nl/tla/islandora/object/tla%3A1839_00_1390CD7C_579F_45F7_9B00_65A31FC980DF [in English].

Плахотнікова Олена Юрїївна — кандидат філологічних наук, асистент кафедри мовної підготовки, Навчально-науковий центр «Інститут біології та медицини», Київський національний університет імені Тараса Шевченка; проспект Академіка Глушкова, 2, Київ, 03127, Україна.

Tel.: +38-099-629-29-72

E-mail: olenaplahotnikova@gmail.com

<https://orcid.org/0000-0003-4274-8554>

Plakhotnikova Olena Yuriivna — Ph.D. in Philology, Assistant Professor at the Department of Language Training, Educational and Scientific Center «Institute of Biology and Medicine», Taras Shevchenko National University of Kyiv; Hlushkova Av., 2, Kyiv, 03127, Ukraine.

Надійшла до редакції 28 серпня 2020 року

CITATION

ДСТУ 8302:2015: Плахотнікова О. Ю. Дидактичний потенціал корпусу транскрибованого українського усного мовлення. *Лінгвістичні дослідження: зб. наук. пр. Харк. нац. пед. ун-ту імені Г. С. Сковороди*. Харків, 2020. Вип. 53. С. 254–262. DOI: <https://doi.org/10.34142/23127546.2020.53.23>

АРА: Плахотнікова, О. Ю. (2020). Дидактичний потенціал корпусу транскрибованого українського усного мовлення. *Лінгвістичні дослідження*, 53, 254–262. DOI: <https://doi.org/10.34142/23127546.2020.53.23>