

Отримано: 19 квітня 2023 р.

Прорецензовано: 1 травня 2023 р.

Прийнято до друку: 5 травня 2023 р.

e-mail: ekaterina.vukolova17@gmail.com

DOI: 10.25264/2519-2558-2023-17(85)-10-13

Вуколова К. В. Використання лінгвістичних корпусів у вивченні словобудови діалектів американського варіанту сучасної англійської мови. *Наукові записки Національного університету «Острозька академія»: серія «Філологія»*. Острог: Вид-во НаУОА, 2023. Вип. 17(85). С. 10–13.

УДК: 811.111:81'28

Вуколова Катерина Володимирівна,
кандидат філологічних наук,
Білоцерківський національний аграрний університет

ВИКОРИСТАННЯ ЛІНГВІСТИЧНИХ КОРПУСІВ У ВИВЧЕННІ СЛОВОБУДОВИ ДІАЛЕКТІВ АМЕРИКАНСЬКОГО ВАРІАНТУ СУЧАСНОЇ АНГЛІЙСЬКОЇ МОВИ

Стаття досліджує засади використання концептів корпусної лінгвістики у процесі вивчення американської діалектної англійської мови.

Розглянуто витоки появи корпусної лінгвістики американсько-діалектної англійської мови, досліджено корпусні напрацювання У. Френсіса та Г. Кучери (Brown Corpus, 1964).

Визначається поняття «корпусної лінгвістики» та описано деякі її проблемні аспекти. Означено цenzi використання корпусно-лінгвістичного підходу при вивченні діалектно-граматичних особливостей американської англійської. Окреслено граматичні відмінності між англійською мовою для Британії та англійською мовою для США. Означено термін «розмовна англійська» та визначено, як остання діалектизувалась в США. Розрізнено американо-англійську діалектну та американську базову (Course Book English) варіанти англійської мови.

Акцентовано увагу на порівняльно-правових напрацюваннях щодо корпусної лінгвістики британської та американської англійської мов (Браунівський та Ланкастерський корпуси). Згадано про дискографічні дослідження розмовної американської англійської (діалектної): Lancaster/IBM SEC та Corpus of Spoken American English). Окреслено особливості використання Інтернаціонального корпусу англійської мови (International Corpus of English), що вивчає британський та американський варіанти англійської мови у філолого-методологічній єдності. Означено процеси розширення мовного корпусу американської англійської мови у період 1990 р. – сьогодення (крізь призму статистики від The Bank of English).

Означено основні методичні розробки щодо систематизації корпусу американської англійської мови з точки зору теорії корпусної лінгвістики: програмне корпусне забезпечення; проблемно-параметричні галузеві відбори; текстова транскрипція; пошуково-системна робота; дескрипційне кодування «збережених» діалектних слів та словотворень.

Ключові слова: лінгвістичний корпус, корпусна лінгвістика, розмовна англійська, американський діалект англійської мови, американсько-діалектна слово будова.

Kateryna Vukolova,
Ph.D. in Linguistics,
Bila Tserkva National Agrarian University

USAGE OF LANGUAGE CORPUSES IN STUDYING WORD FORMATION OF MODERN AMERICAN ENGLISH DIALECTS

The article deals with the principles of using the concepts of corpus linguistics in studying American English dialects.

The origins of corpus linguistics of American English dialects are considered, and developments in corpus linguistics by W. Francis and G. Kuchera are investigated (Brown Corpus, 1964).

The concept of «corpus linguistics» is defined, and some problematic aspects are described. The qualifications for using the corpus-linguistic approach in the study of dialect grammatical features of American English are defined. The grammatical differences between British and American English are outlined. "Spoken English" is defined as how the latter was dialectical in the United States. The American dialects and Standardized American English are distinguished.

Attention is focused on comparative and legal developments in corpus linguistics of British and American English (Brown and Lancaster Corps). Discographic studies of spoken American (dialectal) are mentioned: Lancaster/IBM SEC and Corpus of Spoken American English. The peculiarities of using the International Corpus of English, which studies British and American English in philological and methodological unity, are outlined. The processes of expanding the language corpus of the American English language from 1990 till the present (through the prism of statistics from The Bank of English) are determined.

The main methodological developments in the systematization of the corpus of American English from the point of view of the theory of Corpus Linguistics are indicated: software corpus; problem-parametric industry selections; text transcription; search and system work; description encoding of "preserved" dialect words and word formations.

Keywords: linguistic corpus, corpus linguistics, spoken English, American dialect of English, American-dialect word structure.

Постановка проблеми. Використання лінгвістичних корпусів – довідко-статистична та джерельно-інформаційна активність щодо уживаності певного слова або словотворення у мовному, мовленнєвому, словоформному та лексемно-часотному контекстах (інша назва – «конкорданс»).

Особливості застосування лінгвістичного корпусу американської англійської мови (переважно – США, American English або АЕ) надає змогу систематизувати основні відмінності даної мовної течії від британської англійської (British English або ВЕ) та визначити аспекти (концепти) автентичності American English.

Визначення особливостей розвитку мовного-корпусної лінгвістики сучасної American English демонструє, що від 1989 р. до 2022 р. запас загальноновживаних словотворень зріс від 20 млн. до 650 + млн. слів (за даними The Bank of English). Відтак, важливо визначити мірила та критерії мовно-лінгвістичної асиміляції (глобалізація, міграційні процеси, інформатизація) American English в умовах сьогодення та закономірності її (асиміляції) відображення у лінгвістично-корпусних дослідженнях щодо АЕ.

Аналіз останніх досліджень і публікацій. Варто відзначити відносно невисокий рівень наукової розробленості проблематики використання лінгвокорпусів у процесі вивчення АЕ. Через це, доцільно згадати про дотичні вітчизняні та іноземні науково-лінгвістичні розробки.

Вітчизняний парк лінгвістично-корпусних розробок щодо АЕ представлений працями Г. Войтко (до питання перспектив розвитку корпусної лінгвістики, окреслення закономірностей реформаційно-розвиткової ролі АЕ); В. Жуковської (загальне означення закономірностей корпусної лінгвістики); О. Демської-Кульчицької (до питання аспектного застосування корпусної лінгвістики); О. Міщенко (щодо граматичних наративів вивчення АЕ із використанням корпусно-лінгвістичного підходу); В. Тітової (до питання вокабулярних запозичень до американського варіанту англійської мови – АЕ) тощо.

Іноземний вітчизняний апарат дослідження аспектів використання лінгвістичних корпусів представлений науковими напрацюваннями А. Бултона (лінгвокорпуси у викладанні, навчанні та щоденному використанні); А. Франкенберг-Гарсії (позитивні та негативні аспекти лінгвістично-корпусного вивчення іноземної мови); П. Жударські (роль та місце лінгвістичного корпусу в процесах вивчення британського та американського варіантів англійської мови).

Мета та завдання дослідження – дослідити особливості використання лінгвістичних корпусів у вивченні словобудови діалектів американського варіанту сучасної англійської мови через виконання наступних завдань: розглянути витoki появи корпусної лінгвістики американсько-діалектної англійської мови; визначити поняття «корпусної лінгвістики» та розгляд деяких її проблемних аспектів; розрізнити американську діалектну та американську базу (Course Book English) варіанти англійської мови; розглянути порівняльно-правові напрацювання щодо корпусної лінгвістики британської та американської англійської мов (Браунівський та Ланкастерський корпуси); окреслити особливості використання Інтернаціонального корпусу англійської мови (International Corpus of English); визначити основні методичні розробки щодо систематизації корпусу американської англійської мови з точки зору теорії корпусної лінгвістики.

Виклад основного матеріалу. Корпусна лінгвістика – методологія, в якій мовний шаблон вивчається в корпусі (великій частині лінгводаних, лінгвоінформації). В свою чергу, лінгвістичний корпус – текстово-письмова та усно-доступна для аналізу сукупність іноземного тексту, слів, словотвору та словосполучень, що у сукупності формують т.зв. «мовний діалект» (Boulton, 2016).

Історична поява американського діалекту англійської мови (АЕ) носить generic-характер. Інакше її називають «видозміненою» англійською мовою – у якості об'єднуючого елемента, що систематизував постійний іммігрантський рух до США європейців, австралійців та представників народів Африки. З точки зору лексикографічної та лінгвістично-корпусної логіки, АЕ – спрощений від ВЕ варіант, ідеологічною основою якого є «різнорідна масова культура без певної національності». Логічно, що міжнародним проєктно-аналітичним мовним стартапом The Bank of English (COBUILD Corpus) відмічено стрімке наповнення мовно-лінгвістичного, мовно-сингулярного та мовно-арсенального апаратів АЕ наприкінці ХХ – початку ХХІ ст. : так, у 1997 р. лінгвістично-мовний корпус АЕ нараховував до 300 млн. діалектичних словотворень та усталених зворотів, у 2005 р. – більше 500 млн., станом на 2022 р. – понад 650 млн. фразеологічно-діалектичних зворотів (розмовно-звичасвих висловлювань) (LTW. The Bank of English (COBUILD Corpus)).

Етимологічні основи використання та вивчення лінгвістичних корпусів АЕ в осучасненому форматі заклали У. Френсіс та Г. Кучери у 1963 р., сформувавши т.зв. «Корпус Брауна» (Brown Corpus). Завданнями корпусу визначалось «корпусний, систематизований текстовий набір». Обсяг – до 1 млн. діалектичних словотворень, слововживань та словосполучень. Враховуючи машинно-носіїний формат аплікації Brown Corpus 1964, його прийнято вважати першопроходцем корпусно-лінгвістичної теорії, що досліджувала аспекти генерації АЕ-діалектів (Francis, Kuchera, Brown, 1979).

Аспекти асиміляції АЕ в розмовний простір США, в свою чергу, призвели до появи двовимірного «похідного англомовного середовища» – формування розмовної (day-to-day) АЕ та специфічної АЕ, що вивчається на рівні освітнього забезпечення США (т.зв. Course Book English). Якщо для «носіїв» американської англійської (АЕ) даний процес є опануванням академічного курсу АЕ на основі уже здобутих навичок комунікації на неформальному АЕ-рівні, то для іммігрантів до США (представників інших національностей) першочерговим є процес осягнення академічної АЕ-англійської мови, внаслідок чого адаптація до неформального мовного середовища набуває ознак утрудненості. Дана конотація, однак, має й позитивні риси : процес пристосування до АЕ у «неносія» мови відбувається шляхом набуття фонетично-розмовних компетентностей (звуки, слова та речення вимовляються чітко, повільно та зрозуміло для співрозмовника), що є передумовою формування нового способу лексикографічного та корпусно-лінгвістичного спілкування – «інтернаціоналізації АЕ».

Логічним продовженням лексичних та лінгвістично-корпусних напрацювань У. Френсіса та Г. Кучери у Brown Corpus 1964 стала поява Ланкастер-Осло-Берген корпусу (Lancaster-Oslo-Bergen Corpus) у 1985 р. (анотована версія) (Johansson, 1978). Структурна ідентичність із Корпусом Брауна дозволила розширити коло науково-лінгвістичних та морфемно-фразеологічних досліджень АЕ до формату англо-американського мовного порівняння (АЕ та ВЕ у застосовчій мовній сукупності) за різножанровою моделлю та із використанням комп'ютерного моделювання.

Як наслідок, у період 1985 – 1990 рр. у лінгвістично-корпусній теоретичній доктрині запровадили дослідження АЕ-діалектів за допомогою дискографії. Розмовна англійська була об'єктом напрацювань на базі Lancaster/IBM SEC та Corpus of Spoken American English ((Knowles, 1996: 278). З точки зору науки, поява нових корпусів засвідчила, по-перше, проникнення технічного прогресу в лінгвістично-словотворчі колокації, і по-друге – розширення лінгвістичних корпусів розмовної АЕ внаслідок недостатнього рівня предметного впливу корпусів 1-го рівня (Brown Corpus 1964) та поступовий перехід до технологічного розвитку корпусів шляхом обробки текстів та формування новітніх корпусів обсягом 100 млн словотворчих концептів (т.зв. «корпуси другого покоління»).

Оскільки, як зазначалося раніше, асимілятивний ефект АЕ має соціально-навчальну сутність (соціальна – розмовна АЕ, навчальна – опанування академічної АЕ), актуальності набуло питання узагальненої методології викладання. Звернемо увагу : вивчення АЕ – найчастіше «перший вибір» під час академічного навчання поза межами США (країни-члени ЄС, країни Східної Європи, країни Африки тощо), внаслідок чого застосовуються усталені проформи вимови, вокабулярності, академічно-писемної та академічно-розмовного мислення. Результатами науково-лінгвістичної та морфемно-корпусної

експансії АЕ стала поява двох АЕ-корпусів ХХ ст. – лінгводидактичного корпусу (International Corpus of Learner's English) та Міжнародного корпусу англійської мови (International Corpus of English) (Greenbaum, 1991: 3-7).

Даний процес, однак, не нівелює т.зв. «негативної» (похідної) мовно-вокабулярної та лінгвістично-корпусної асиміляції, внаслідок якої з'являються нові діалекти: африканський (ПАР, Нігерія та ін.), латино-американський (Аргентина, Бразилія, Чилі та ін.), азійський (Китай, Японія та ін.). Крім того, варто пам'ятати про наявність австралійського та канадського діалектів АЕ, що є похідними (видозмінними) від неї.

Доцільно відмітити, що корпусно-лінгвістичний апарат АЕ наділений ідеологічним компонентом. На думку Г. Войтко, США як основний носій АЕ-культури спілкування пропонує «американську мрію» не лише як засіб державно-соціального курсу, а й як спосіб інтернаціонального приваблення. Масова імміграція до США, американська різнокультурність та політика автономної єдності – спосіб уніфікації мовного питання шляхом «формування симпліфікованих лінгвістично-корпусних зворотів, граматичних та синтаксичних формул, прийнятих для опанування різними верствами населення – незалежно від соціально-етнічного походження та належності до певної мовної групи (Войтко, 2014: 208-212).

Американська англійська мова (АЕ) у іноземній лінгвістично-корпусній теорії відноситься до «природних мов». Це означає, що АЕ-лінгвокорпуси пропонують широкі опції навчально-наукового осягнення розмовних стандартів. На думку польського науковця П. Жударські, АЕ-корпусна лінгвістика визначає такі методи вивчення англійської мови, як освіта (включно із наданням освітніх послуг), контекст культурної асиміляції та техніко-теоретичний (техніко-технологічний) прогрес як спосіб Е-осягнення АЕ-варіанту англійської мови (Szudarski, 2010-2011: 37-47).

Корпуси АЕ-англійської мови також використовуються для систематизації автентичних мовних даних (інформації). Уже сформовані звичай формального та неформального спілкування виступають основою для розробки широкого спектру галузевих мовних матеріалів (словників, підручників). Як зазначають Р. Картер та М. Маккарті (Carter, McCarthy, 2014: 254), нерідко навіть носій мови є неточним при оцінці вживання лінгвістичних елементів та певних корпусно-лінгвістичних зворотів. Наразі лексикографія та лінгвістичне розробництво набуло ознак суб'єктивності: бази даних та Е-корпуси нерідко «пропускають» важливі для мовно-лінгвістичного балансу та запасу морфемно-колокаційні структури.

На рис. 1 представлено узагальнений концепт мовно-лінгвістичної та лінгво-корпусної АЕ-структуризації в ретроспективно-термінальному аспекті.

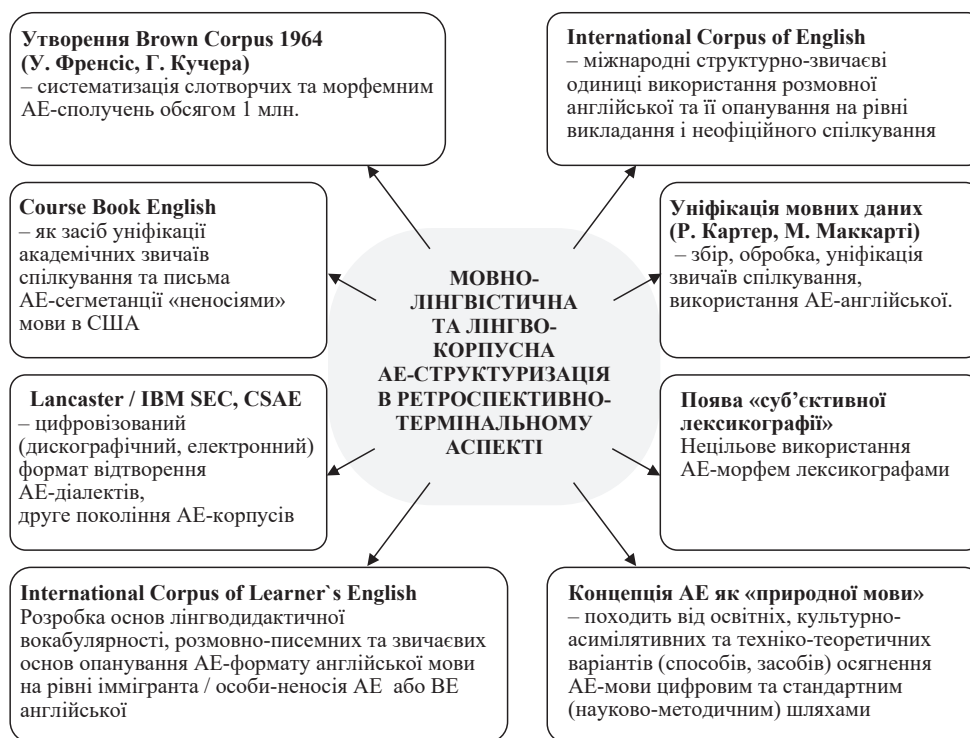


Рис. 1 Мовно-лінгвістична та лінгво-корпусна АЕ-структуризація в ретроспективно-термінальному аспекті

Не дивлячись на наявні приклади цифрового дослідження АЕ-вокабуляру науковцями та педагогами (зібрання фраз-зворотів, висловів-термінів), подекуди не вирішеним залишається аспект розгалуження лінгвістичних корпусів в умовах «дистанційнізації освіти». Приклад позитивного стартапу даної сфери на сьогодні – лише Compleat Lexical Tutor (засновник – Т. Кобб), що дозволяє проводити регулярні мовні заняття із використанням АЕ та ВЕ-лексики, формувати завдання «на базі» певного опрацьованого корпусу, створювати списки вокабулярних частот (частин) із текстів або формувати «закритий» текст із опцією виконання задля здобуття фазових АЕ-компетентностей.

В свою чергу, технологізація навчально-соціального середовища створює прецедент використання Інтернет-мережі як окремого «лінгвістичного корпусу». Згідно позиції Ф. Келлера та М. Лапати, створення «онлайн лінгвокорпусу» має двовимірну структуру: по-перше, цифрова систематизація наявних АЕ-джерел значно спрощує доступ до останніх та надає можливість їхнього вивчення «на руках» (в режимі реального часу); по-друге – власне worldwide web має певні сленгові,

розмовні, писемні ідіоми, слововживання та елементи словотвору, котрі використовуються для позначення емоцій та комунікації на інтернаціональному рівні (наприклад, соціальна мережа на кшталт Facebook або Twitter), через що можуть бути частиною корпусно-лінгвістичного дискурсу за умови правильного оформлення.

При цьому, цифровізація корпусно-лінгвістичної англійської мови може мати результатом не лише раніше згадану умову «інтернаціоналізації» (присотосування носія АЕ-мови до мовного середовища та імплікація спрощеного варіанту спілкування – наприклад, в США), але й інтернаціонально-витокового дослідження АЕ з точки зору її зв'язку із лексикографією, наприклад, ірландсько-шотландської або валлійської мовної традиції, а також впливом на остаточне формування АЕ не лише процесу адаптації носіїв (non-native speakers), а рецепції (увібрання) в АЕ комунікативно-лексичних мотивно-морфемних мотивів італійської, чеської та ін. мов.

Висновки. Дослідження аспектів використання лінгвістичних корпусів у вивченні словобудови діалектів американського варіанту сучасної англійської мови дозволило дійти наступних умовиводів.

По-перше, АЕ-лінгвістичний корпус є де-факто масовим симпліфікованим зібранням мовно-розмовних, звичаєвих та комунікативно-ситуативних варіантів використання усталених зворотів, ідіом, слововживань та словосполучень у похідних морфемно-стилістичних варіаціях.

По-друге, процес АЕ-імплікації має двовимірну структуру : соціально-розмовну (Brown Corpus 1964 від У. Френсіса та Г. Кучери) та навчально-методичну (Course Book English) інтеграцію, що корелюють між собою (носієм мови інтегрується у мовне середовище завдяки Course Book English, носій мови завдяки Course Book English «шліфує» уже набуту мовно-соціальні комунікативні здібності).

По-третє, цифровізація АЕ-вокабуляру може допомогти у створенні двоаспектного онлайн АЕ-середовища : на рівні т.зв. «АЕ-Інтернету» (вислови, підходи, Інтернет-сленги, Інтернет-фрази, що використовуються для комунікації) та на рівні систематизації наявних АЕ-висловів та АЕ-словотворень.

Література:

1. Boulton A. Using Corpora in Language Teaching, Learning and Use. *Pour une formation durable en Lansad*. 2016. Vol. 35 (2). URL : <https://journals.openedition.org/apliut/5433>
2. LTW. The Bank of English (COBUILD Corpus). LTW official website. URL : https://www.lt-world.org/kb/resources-and-tools/language-data/ltw_x3alanguage_x5fdata_2010-09-23.5678579368
3. Francis W. N., Kuchera H. Brown Corpus Manual. Providence, Rhode Island : Department of Linguistics Brown University, 1979. URL: <http://icame.uib.no/brown/bcm.html>
4. Johansson S. Lancaster-Oslo-Bergen corpus of modern English (LOB), 1978. Oxford Text Archive. URL: <https://ota.bodleian.ox.ac.uk/repository/xmlui/handle/20.500.12024/0167>
5. Knowles G. et al. A Corpus of Formal British English Speech. The Lancaster/IBM Spoken English Corpus. Routledge, 1996. 278 p.
6. Greenbaum S. ICE: the International Corpus of English. Cambridge University Press, 1991. P. 3–7.
7. Войтко Г. Корпусна лінгвістика: історія формування і перспективи розвитку. *ВНУ : Розділ IV. Теоретичні засади лінгвістичних досліджень*. 2014. № 5. С. 208–212. [Voitko, H. Korpusna linhvistyka : istoriia formuvannia i perspektyvy rozvytku. VNU : Rozdil IV. Teoretychni zasady linhvistychnykh doslidzhen, 2014. № 5. S. 208–212.]
8. Szudarski P. The Role of Language Corpora in Teaching English as a Foreign Language in Poland. *Language, Communication, Information*. 2010-2011. Vol. 5. P. 37–47.
9. Carter R., McCarthy M. Vocabulary and Language Teaching. Routledge, 2014. 254 p.