

Алла Титок

Київський національний університет імені Т. Шевченка

ДОСЛІДЖЕННЯ КОНЦЕПТІВ
УКРАЇНА, УКРАЇНСЬКИЙ, УКРАЇНЕЦЬ, РОСІЯ,
РОСІЙСЬКИЙ, РОСІЯНИН
У ХРОНОЛОГІЧНИХ ЗРІЗАХ НА МАТЕРІАЛІ
НАЦІОНАЛЬНОГО ЛІНГВІСТИЧНОГО КОРПУСУ

Створення та продуктивне використання лінгвістичних корпусів є одним із основних напрямів роботи сучасного мовознавства. У роботі проаналізовано ряд актуальних концептів із використанням матеріалів та інструментів Національного лінгвістичного корпусу української мови; визначено особливості використання кожної із лексем у певний історичний період; наведено основні кількісні характеристики.

Ключові слова: корпусна лінгвістика, концепт, Національний лінгвістичний корпус, Україна, український, українець, росія, російський, росіянин.

The creation and productive use of linguistic corpora is one of the main areas of work in modern linguistics. The work analyzes several relevant concepts using materials and tools of the National Linguistic Corpus of the Ukrainian Language; the peculiarities of the use of each lexeme in a certain historical period are determined; main quantitative characteristics are presented.

Key words: corpus linguistics, concept, National Linguistic Corpus, Ukraine, Ukrainian, Russia, Russian.

Розвиток сучасного мовознавства неможливо уявити без комп'ютерних технологій. Елементи комп'ютерної галузі займа-

ють провідне місце у мовознавчих дослідженнях. Створення та використання лінгвістичних корпусів є одним із пріоритетних завдань сучасних мовознавчих розвідок.

Актуальність проблеми полягає у відсутності в науковому просторі опрацювання концептів *Україна* та *росія* та похідних від них, в аналізі частотності та умотивованості вживання відповідних лексем у художніх текстах у діячності, у необхідності уваги мовознавчих досліджень до живих творів та записів, що інтерпретують справжню думку народу певного періоду.

Філологія впродовж кількох десятиліть не функціонує ізольовано. Можливості корпусної лінгвістики спрощують здійснення необхідних пошуків, а години гортання книжок у пошуках потрібного слова тепер змінилися на кілька тисячних секунд роботи комп'ютера.

Під терміном «**корпусна лінгвістика**» розуміємо галузь мовознавства, що об'єднує його теоретичну та прикладну частини, здійснює обробку корпусів тексту, здійснює користувачу запити на вимогу [2, с. 9]. Корпусна лінгвістика використовує тексти з письмової та усної сфер функціонування, тож цим розширює сферу лінгвістичних досліджень [1, с. 9].

Під час проведення дослідження ми звертали увагу на художні тексти різних століть і географічне походження авторів. На початковому етапі було визначено перелік письменників, що хронологічно брав початок від XIX століття, але від початку проведення аналізу було виявлено, що досліджувані концепти не використані безпосередньо авторами у текстах до середини XX століття. Натомість, мають місце синонімічні відповідники досліджуваних лексем, які теж було проаналізовано під час дослідження.

Зокрема, як матеріали дослідження були використані поетичні, прозові та драматичні твори українських письменників із застосуванням Українського національного лінгвістичного корпусу (УНЛК), що створений Українським мовно-інформаційним

фондом НАН України (режим доступу: https://svc.ulif.org.ua/UNLC/virt_unlc_4.5/). Сьогодні корпус налічує більше 200 мільйонів слововживань, об'єднує тексти різних стилів та періодів, постійно розширюється та наповнюється новими джерелами [3, с. 249].

Порівняння визначених нами концептів *Україна, український, українець, росія, російський, росіянин* дало такі результати: 1) найчастотнішою є лексема *Україна* – 4208 слововживань, дещо менше *український* – 3940 прикладів. Найменша частота використання лексеми *росіянин* – всього 1221 слововживання, навіть разом зі словоформою *росіянка* (247) не вдається подолати поріг до вищої сходинки, де опинилася лексема *українець* – 1941 слововживання. Інші місця розділили лексеми *російський* та *росія* – 2775 та 2232 слововживання відповідно.

У текстах лексеми майже ніколи не використовуються відокремлено, часто помічаємо їх поряд із назвами інших країн, народів чи належності до них: «я не розумію, що таке *українці*, хто вони такі: євреї, татари, вірмени» (М. Куліш); «потім греки, група німців, там три турки, там три поляки, там *росіяни*, але це все, як мозаїка, було вкраплене в основне море *українців*» (Іван Багряний); «то не литовський акцент у *російській*, то просто мова така окрема є. На «У» починається» (І. Карпа);

Зафіксовано поширені словосполучення: «свідомі *українці*» (Б. Антоненко-Давидович, Ю. Винничук, Остап Вишня, У. Самчук та ін.); «*українці*, старші слов'яни, ніж руські» (В. Винниченко); «*щирий українець*» (М. Андрусак, В. Винниченко, Б. Антоненко-Давидович, М. Дочинець, П. Загребельний та ін.); «*російський* полон» (Ю. Андрухович, О. Гончар, В. Шкляр та ін.); «*російський* агент» (А. Кокотюха, Л. Кононович).

Відчутний контраст вбачаємо у поєднанні лексем із іншими словами. Для певних – це піднесені та позитивні характеристики, для інших – ненависть та темрява. Порівнюємо сполучуваність: *ненька України, вільна, велика, автономія України, Слобід-*

ська Україна на противагу федеративна Росія, перекиривляє російську вимову, найзапущеніший російський міф, крадькома анексованими Росією.

Найдавніший досліджуваний етап — кінець XIX століття — майже не має прикладів для демонстрації концептів. Кілька згадок — творчість П. Куліша: «за здоров'є великої, одностайної України» та І. Нечуя-Левицького: «Україні не сподіватись добра од Москви». Результати дослідження показують, що на лексеми Україна та спільнокореневі кілька разів можемо натрапити, але ніде не помічаємо росію чи пов'язане з нею.

Візьмемо до уваги слово *руський*. До лексеми знайдено в корпусі 1076 слововживань. У різний період, у різних контекстах помічаємо неоднакове тлумачення. У Я. Головацького «*руський з руським*» — українець з росіянином, інші письменники (О. Кобилянська, О. Довженко, Олесь Гончар, Кузьма Скрябін, С. Жадан, І. Карпа) переважно мають на увазі «російські»: «*руськi романи та українські пісні*», «*донька українсько-руського народу*». Ю. Яновський закладав у лексему значення «український»: «*Проклинаю тебе моїм руським серцем...*».

Акцентуємо увагу також на відповідниках до лексем, які використовували у певний час для означення країни чи людини. Концепти *малорос* і *великорос* мають 270 та 169 слововживань відповідно, звертаємося і до лексем *хохол* та *кацап*, що мають 276 та 146 згадувань у корпусі. Часто лексеми не лише поєднані одна з одною у синтаксичному контексті, а й існують поряд з іншими, більш офіційними назвами, часто додають емоційного забарвлення: «*мішала великоруські слова з українськими*» (І. Нечуй-Левицький), «*Так, я — малорос, або, як модно тепер говори-ти, українець*» (В. Винниченко), «*Ми теж ділимося на українців, малоросів і хохлів*» (І. Жиленко), «*“ти українчику”, а “ти каца-пе!”*» (Б. Лепкий). Не широко поширеною, але все ж існує і назва *Кацапія*: «*десь там в кацапії землі багато і вона дуже дешева*» (У. Самчук), «*Він на Україні, а не в Кацапії*» (В. Винниченко).

Звертаємося також до лексем *москаль* (733 слововживання) та *Московія* (264), які теж популярні на сторінках літератури. У творчості М. Кропивницького знаходимо: «А вона й почала мені товквувати, що хахол – це насмішка, все рівно, як *москаля* звуть *кацапом* або *єврея – жидом...*», Ліна Костенко пише: «Всяк *московит* *Московію* *трубить*».

Не можемо оминати і згадки слова *козак* та спільнокореневих щодо належності українцям у контексті зі вже згаданими: «Щоб моя дочка, *козацька дитина*, та стала дружиною *кацапа!*..» (О. Кониський), «*засяють на весь степ козачий*, на всю нашу Україну» (М. Куліш).

Звертаються письменники і до означення української мови як *солов'їної*, до того ж, характеристику знаходимо не лише у фольклорних текстах і не лише у порівняннях: «*наша солов'їна мова*» (Ю. Андрухович). Є й інші, менш вживані означення до мови: «*Тарасошевченківська*» (М. Куліш). Протиставлення знаходимо у формі «*кацапсько-більшевицький молитвеник*» (В. Винниченко) на позначення російської мови.

Отже, зважаючи на дані корпусу та результати дослідження, можемо окреслити основне поле використання концептів *Україна*, *український*, *українець*, *росія*, *російський*, *росіянин* впродовж кількох останній століть. Суспільні та політичні події мають неабиякий вплив на літературу, тож бачимо, як змінюється кількість та значення лексем у художньому стилі. Від лексем *руський*, *малорос*, *українець*, *козак*, *кацап* на початку ХХ століття переходимо до *Україна*, *росія*, *українець*, *московія*, *російський* у ХХІ столітті з різноманітними поєднаннями у проміжний період. Звертаючи увагу на кількісні показники представлення концептів в УНЛК, помічаємо, що важливішими та актуальнішими для українців є все ж лексеми на позначення України та всього, що може її стосуватися.

Перспективу подальших досліджень вбачаємо у розширенні аналізу концептів поза межами українського корпусу та просте-

женні активності й виразності вживання концептів у інших стилях мови.

Список використаних джерел

1. Жуковська В. В. Вступ до корпусної лінгвістики : навчальний посібник. Житомир : ЖДУ ім. І. Франка, 2013. 142 с.
2. Лінгвістично-інформаційні студії : праці Українського мовно-інформаційного фонду НАН України : у 5 т. / В. А. Широков та ін. Київ : Український мовно-інформаційний фонд НАН України, 2018. Т. 4 : Корпусна та когнітивна лінгвістика. 246 с.
3. Семенов О., Надутенко М., Тимошук Р. Національні лінгвістичні ресурси Українського мовно-інформаційного фонду НАН України – ефективний засіб формування академічної культури дослідника *Педагогічні науки: теорія, історія, інноваційні технології*. 2020. № 10 (104). С. 245–258.
4. Український національний лінгвістичний корпус. URL: https://svc.ulif.org.ua/UNLC/virt_unlc_4.5/