

**Анастасія КИРИСЮК,**

*orcid.org/0000-0001-7522-8235*

*аспірантка кафедри германської і фіно-угорської філології*

*Київського національного лінгвістичного університету*

*(Київ, Україна) anastasiia.kyrysiuk@knl.u.edu.ua*

## **РЕСУРСИ КОРПУСНОГО ІНСТРУМЕНТАРІЮ ДЛЯ ПОБУДОВИ КОГНІТИВНИХ МАТРИЦЬ ЛСП ПАНДЕМІЯ (НА МАТЕРІАЛІ ГЕНЕРАЛЬНОГО РЕГІОНАЛЬНО АНОТОВАНОГО КОРПУСУ УКРАЇНСЬКОЇ МОВИ)**

*Стаття присвячена вивченню можливостей лінгвістичних корпусів текстів, як «банку прикладів», що надають змогу дослідити мовні та позамовні знання представників певної лінгвокультури про досліджуване явище чи предмет. У статті розкрито значення базових понять сучасної корпусної лінгвістики – «ресурси корпусу» й «інструменти корпусу». Під час дослідження з'ясовано, що під ресурсами корпусу розуміють тексти, а до інструментів відносять засоби перегляду текстів, запиту даних і засоби збагачення корпусу аналітичною інформацією (анотацією, або метарозміткою). Завдяки ресурсам та інструментам лінгвістичного корпусу дослідники мають змогу всебічно вивчити будь-яку мовну одиницю, а саме простежити всі її зв'язки і відношення для розв'язання різноманітних лінгвістичних завдань.*

*Основну увагу у статті приділено аргументації доцільності застосування ресурсів корпусного інструментарію при дослідженні семантики, зокрема для розширення побудованих когнітивних матриць ЛСП ПАНДЕМІЯ на основі синонімів до ключового слова, які були відібрані зі словників. Для пошуку нових значень вживання лексеми пандемія в українській мові залучено ресурси й інструменти Генерального регіонально анотованого корпусу української мови, який є українським корпусним онлайн-проєктом, що охоплює період з 1816 р. по теперішній час. У результаті дефініційного аналізу встановлено нові значення лексеми пандемія, зокрема назви хвороб, що набувають характеру пандемії та словосполучення, у яких аналізована лексема вжита не у значенні інфекційної хвороби. За допомогою застосування методики компонентного аналізу з визначенням семного складу лексичних значень слова пандемія виокремлено нові диференційні семи, яких не було виявлено під час дослідження на основі лексикографічних значень, – 'несмертоносна хвороба / захворювання', 'штучна хвороба / захворювання'.*

**Ключові слова:** корпусна лінгвістика, корпус текстів, ресурси корпусу, інструменти корпусу, ЛСП ПАНДЕМІЯ.

**Anastasiia KYRYSIUK,**

*orcid.org/0000-0001-7522-8235*

*PhD Student at the Department of Germanic and Finno-Ugrian Philology*

*Kyiv National Linguistic University*

*(Kyiv, Ukraine) anastasiia.kyrysiuk@knl.u.edu.ua*

## **RESOURCES OF CORPUS TOOLS FOR COGNITIVE MATRICES CONSTRUCTION OF LEXICO-SEMANTIC FIELD ПАНДЕМІЯ (LIT. PANDEMIC) (BASED ON THE MATERIALS OF THE GENERAL REGIONALLY ANNOTATED CORPUS OF THE UKRAINIAN LANGUAGE)**

*The article studies the possibilities of linguistic corpora of texts as a «bank of examples» that make it possible to investigate the linguistic and non-linguistic knowledge of representatives of a particular linguistic culture about the phenomenon or subject under investigation. The article reveals the meaning of the basic concepts of modern corpus linguistics – «corpus resources» and «corpus tools». During the research, it was found that the resources of the corpus mean texts, and the tools include text viewing and data query tools and means of enriching the corpus with analytical information (annotation or meta mark-up). The resources and tools of the linguistic corpus may help scholars to be able to comprehensively study a linguistic unit, i.e. to trace all the links and relations in order to solve different linguistic objectives.*

*The primary attention in the article is devoted to the argumentation of the expediency of using corpus toolkit resources in the study of semantics, in particular, to expand the constructed cognitive matrices of LSF ПАНДЕМІЯ (lit. PANDEMIC) based on synonyms for the keyword, which were selected from dictionaries.*

*To search for new meanings of the use of the lexeme пандемія (lit. pandemic) in the Ukrainian language, we use the resources and tools of the General Regionally Annotated Corpus of the Ukrainian Language, which is a Ukrainian online corpus project covering the period from 1816 to the present. As a result of the definitional analysis, new meanings*

*of the lexeme пандемія (lit. pandemic) were established, in particular, the names of diseases that acquire the character of a pandemic and phrases in which the analyzed lexeme isn't used in the direct meaning of disease. In the process of application of the method of component analysis with the determination of the seminal composition of the lexical meanings of the word пандемія (lit. pandemic), new differential semes were singled out. They were not identified during the research based on lexicographic meanings – 'non-lethal disease / disease', 'artificial disease / disease'.*

**Key words:** *corpus linguistics, the corpus of texts, corpus resources, corpus toolkit, LSF ПАНДЕМІЯ (lit. PANDEMIC).*

**Постановка проблеми.** Сучасна комп'ютерна лінгвістика, розділом якої є корпусна, – новий самостійний розділ науки про мову, інструментарій якого все більше застосовують під час досліджень, що пов'язано із зростанням кількості оброблюваної інформації. За словами О. Демської-Кульчицької (2005: 68), нині всі необхідні дані для повсякденного життя і опанування мовами, є в електронному форматі. Навіть колишні друковані видання, зокрема музейні та архівні манускрипти, рукописи мають електронні відповідники. У зв'язку з цифризацією писемного й усного матеріалу, на думку Г. Кривенко (2017: 51), постало питання розробки пошукової системи, яка змогла б ефективно і мобільно забезпечити опрацювання великих об'ємів різноманітної інформації. Однією з таких пошукових систем стали *лінгвістичні корпуси текстів*, які мають системи кодування і зручного пошуку інформації. За словами О. Демської-Кульчицької (2004), корпуси текстів надають змогу усвідомити когнітивно-інформативну / інформаційну сутність людської мови, а також окремої мови як її представника.

**Аналіз досліджень.** Як показує попередній огляд наукової літератури (Н. Дарчук, О. Демська-Кульчицька, В. Жуковська, М. Леміш, С. Петрасова, Н. Хайрова, J. Asmussen, L. Bowker, A. Hardie, A. Kilgarriff, I. Kosem, T. MacEnery, J. Pearson та ін.), завдяки корпусним ресурсам і корпусному інструментарію у дослідників з'являється можливість вивчити будь-яку мовну одиницю багатовимірно, тобто простежити всі її зв'язки і відношення для розв'язання лінгвістичних завдань різного характеру.

**Мета статті** – дослідити ресурси корпусного інструментарію Генерального регіонально анотованого корпусу української мови (далі – ГРАК) для побудови когнітивних матриць ЛСП ПАНДЕМІЯ.

**Виклад основного матеріалу.** Традиційно під *корпусом* розуміють тіло тексту, проте його можна описати як велику колекцію автентичних текстів, зібраних в електронній формі згідно з певним набором критеріїв (Bowker, Pearson, 2002). Схожої думки притримується Дж. Асмуссен (Asmussen, 2007) і розуміє під корпусом значне за обсягом цифрове зібрання текстів / текстових уривків, що застосовується як репрезентативна вибірка для

обмеженого використання мови, а в загально-мовному лексикографічному контексті становить цільову вибірку з мови загалом.

У роботах інших зарубіжних вчених (McEnery, Hardie, 2012: 230) лінгвістичний корпус розуміють як «набір машиночитаних текстів, які є надійною основою для вивчення певних дослідницьких питань». При цьому, на їхню думку, основним дискусійним питанням є розкриття можливостей корпусу при дослідженні мовних явищ у діячності, чи він має тексти лише для аналізу мови у синхронії. Ми переконані, що це залежить від площини конкретного корпусу і мети його укладання.

У певному сенсі об'єктом вивчення корпусу стає мовлення, яке не зводиться до норм літературної мови, судженням про правильність / неправильність в мові, заснованим виключно на інтуїції дослідника. Підтвердження цієї думки знаходимо в О. Смашнюк (2008: 63), яка зазначає, що корпуси надають можливість вивчати писемне і спонтанне мовлення, мовлення певної вікової, гендерної, соціальної та етнічної групи, інформацію про особливості певного діалекту тощо.

Ідеєю будь-якого корпусу є укладання їх на основі текстів, а не окремих лексем, бо «досвід навколишнього світу значною мірою є текстовим опосередкованим досвідом, і, з огляду на таке припущення, люди живуть у світі текстів» (McEnery, Hardie, 2012: 230). Відповідно до цього твердження корпусний підхід порушує питання щодо перегляду статусу слова як основної мовної одиниці. Попри те, що корпусний підхід ґрунтується не на окремих лексемах, а на розробці принципів укладання текстів та їхньої розмітці, ключовим питанням цього напряму є здатність корпусу надавати досліднику можливість вимірювати частоту вживання певних слів і фраз, визначати способи їхньої сполучуваності, валентні зв'язки, комбінаторні можливості, композиційну семантику, стилістичні реєстри окремих слів тощо (там само).

Отож, базовими поняттями корпусної лінгвістики є *ресурси корпусу* і *корпусний інструментарій*.

Ресурси корпусу, за В. Жуковською (2013), є знаряддям у дослідженні, адже це повнотекстові бази даних. Доступ до ресурсів більшості корпусу

сів надається безкоштовно через мережу Інтернет. При цьому обсяг ресурсів не обмежений, бо корпуси постійно поповнюються новими джерелами, що є перевагою для дослідників.

А. Кілгарріфф та І. Косем (Kilgarriff, Kosem, 2012) наголошують, що корпус, крім ресурсів, повинен мати комплекс інструментів для роботи з ними. Ці інструменти можна поділити на дві категорії: 1) засоби перегляду текстів і запити даних; 2) засоби збагачення корпусу аналітичною інформацією, яка називається анотацією, або метарозміткою.

Щодо засобів перегляду текстів, то у більшості сучасних корпусів можна шукати різноманітну інформацію (конкретні словоформи; словоформи за лемою (усі форми однієї лексеми, які трапилися у тексті); нерозривні та розривні словосполучення), що дає змогу побудувати *конкорданс* – список усіх вживань певного слова в контексті. До кожного прикладу надається інформація про джерело, звідки його взято. Також передбачена можливість отримання статистичної інформації про запитувану лексему чи словосполучку: відносну частоту по всьому корпусу; розподіл за жанрами; часовими зрізами; частоту його сполучуваності (Петрасова, Хайрова, 2020).

Важливим є інструмент частотного аналізу, бо він, як доводить В. Жуковська (2013), допомагає простежити еволюцію та динаміку функціонування будь-якої лексеми як на синхронному зрізі, так і в діахронії, а також виявити відсоткове співвідношення цієї частоти в різних типах дискурсу.

А. Габідулліна й О. Колесніченко (2019) наголошують, що корпусний інструментарій текстів повинен містити додаткову лінгвістичну і металінгвістичну інформацію, тобто *метарозмітку*, яка є відображенням екстралінгвістичної інформації про текст, зокрема його назву, автора, жанрову належність тощо. Наявність метарозмітки дозволяє користувачеві створювати підкорпус текстів, які можна відібрати за жанром, тематикою, часом написання тощо. Також у сучасних корпусах є розмітка, яка відповідає усім рівням лінгвістичного опису – фонетична, морфологічна, синтаксична та інші, за допомогою яких дослідник може знайти, наприклад, необхідний набір морфологічних ознак: усі словосполучення виду ‘прийменник + іменник у місцевому відмінку’ та ін. інформацію, що відповідає рівням лінгвістичної розмітки, представлених у корпусі (там само).

Отож, ресурси та інструменти сучасних електронних корпусів є потужними і гнучкими, дозволяють ставити питання про склад мови і миттєво отримувати на них відповіді у вигляді сотень

речень. Тому часто корпус використовують як «банк прикладів», тобто намагаються знайти емпіричну підтримку для своїх гіпотез, принципів і правил, над якими вони працюють (Петрасова, Хайрова, 2020). Приклади, звичайно, можуть бути придумані або ж знайдені випадково, але підхід корпусної лінгвістики забезпечує репрезентативність і збалансованість мовного матеріалу, а також пошуковий інструмент, який зазвичай дає можливість хорошої вибірки в певному корпусі. Завдяки таким можливостям, на основі корпусів проводять дослідження на різних мовних рівнях і визначають: стилеметричні ознаки морфемних структур слів у поетичному мовленні Т. Шевченка (Зубань, 2014); ключові лексеми-номінативи до концептів, зокрема концепту *VIPA* (Яницька, 2016); лексичну і граматичну варіативність української мови, зокрема вживання варіативних форм іменників-назв неістот у знахідному відмінку однини в літературних текстах регіонів України XIX–XXI ст. (Шведова, 2018a) і динаміку синтаксичних конструкцій на позначення шляху при дієсловах руху (Шведова, 2018b); семантику фразово-дієслівних емотивів (Бобер, 2020).

Звертаємо увагу, що дослідження семантики за допомогою ресурсів корпусу вже проведені, тому припускаємо, що корпуси текстів нададуть широкі можливості також і для дослідження ЛСП, зокрема допоможуть розширити побудовані когнітивні матриці ЛСП ПАНДЕМІЯ на основі синонімів до ключового слова, які були відібрані зі словників (Кугусяк, 2022), проте, на наш погляд, не розкривають усіх особливостей значень вживання цього слова в мові представників певної лінгвокультури. Для пошуку нових (не словникових) значень вживання лексеми *пандемія* в українській мові залучаємо ГРАК, який є українським корпусним онлайн-проектом, що охоплює період з 1816 р. по теперішній час, має обсяг понад 860 мільйонів токенів, містить 90 тисяч різнодискурсних текстів і 26 тисяч авторів.

Скориставшись ГРАК, звертаємося до вебсторінки «Пошук у ГРАКу», де натискаємо на ГРАК-15, який відсилає користувача до NoSketch Engine з відповідним інтерфейсом.

Натискаємо на опції “Concordance” і в “Query type” обираємо “simple”. До поля abc вводим українськомовну лексему *пандем\** і натискаємо клавішу “GO”. Зауважимо, що використовуємо зірочку (\*) для того, щоб охопити більшу кількість символів, наприклад, у результаті пошуку отримаємо не лише результат із словом *пандемія*, а і його словозміну – *пандемії, пандемію, пандемією, пандемій, пандеміям, пандеміями, пандеміях*. Так



конкорданс показує 1 086 випадків уживання українськомовної лексеми *пандемія* у текстах різних дискурсів. Інструментарій ГРАКа дозволяє аналізувати конкорданс за роками, тому звертаємо увагу, що перша згадка лексеми діагностована в 1889 р. у газеті «Дѣло» (27.12.1889), а друга – аж через 72 роки, в 1961 р. – в журналі «Мистецтво», № 4. Загалом ГРАК містить 610 згадок лексеми до 2017 р. включно. У 2018 р. – 23, а в 2019 р. відзначаємо незначне збільшення кількості вживання лексеми *пандемія* – зафіксовано 30 згадок. У 2020 р. досліджувану лексему згадано в ГРАК 346 разів у різножанрових текстах: це – офіційно-діловий дискурс: стенограми засідань Верховної Ради України (304); медійний дискурс (30); науковий – дисертації (9); художній дискурс (5); науково-популярний дискурс із соціології (7), із психології (1). У 2021 р. лексему *пандемія* згадано 77 разів. Високу частоту вживання лексеми з 2020 р. пояснюємо спалахом захворюваності на коронавірусну хворобу, спричинену SARS-CoV-2, та поширенням COVID-19 загалом у світі.

Під час корпусної діагностики вдалося виявити 381 значення лексеми *пандемія*, з яких у результаті проведеного компонентного аналізу виокремлено такі семи:

1. Архісема *‘хвороба / захворювання’*: «*хвороба*» (7), «*захворювання*».

1.1. Інтегральна сема *‘тип хвороби / захворювання’*:

1.1.1. Диференційна сема *‘інфекційна хвороба / захворювання’*: «*інфекційна хвороба*» (2).

1.1.2. Диференційна сема *‘інфекційна хвороба / захворювання (людей)’*: «*передаватиметься від людини до людини*», «*швидко поширюються серед населення*», «*спроможна швидко передаватися від людини до людини*», «*передавання вірусу від людини до людини*».

1.1.3. Диференційна сема *‘інфекційна хвороба / захворювання (тварин)’*: «*не передається від свиней людям*», «*свині масово хворіли на грип*».

1.2. Інтегральна схема *‘якість хвороби / захворювання’*:

1.2.1. Диференційна сема *‘смертоносна хвороба / захворювання’*: «*смертоносна*» (4), «*забрала життя десятків мільйонів людей*» (3), «*найсмертоносніша*», «*винищила майже все грабарське населення Землі*», «*звела в могилу удвічі більше людей, ніж Перша світова війна*», «*дуже висока смертність населення*», «*забрала багато життів*», «*померло 70 000 людей*», «*число загиблих склало 33 800 чоловік*», «*може забрати навіть не мільйони, а мільярди життів*», «*відправила на той світ, за різними підрахунками, від 20 до*

*100 млн хворих*», «*вмерли від пандемічного грипу 2555 осіб*», «*забрала із собою в могилу близько половини всього населення тих територій, на які поширилася*», «*вбила, за різними оцінками, від 50 до 100 млн людей*», «*створюють загрозу життю і здоров’ю значних верств населення*», «*може загинути від 2 до 7,4 млн чоловік*».

1.2.2. Диференційна сема *‘несмертоносна хвороба / захворювання’*: «*перетворює людей на живих мерців*», «*число хворих сягає критичної маси*», «*вже прийшла майже в кожную сім’ю*», «*вражали мільйонів людей*», «*жертвам стали близько 100 мільйонів осіб у всьому світі*».

1.2.3. Диференційна сема *‘штучна хвороба / захворювання’*: «*штучно викликана*» (3), «*штучно влаштована*», «*штучно створена*».

1.2.4. Диференційна сема *‘циклічність хвороба / захворювання’*: «*циклічна*», «*виникають кожні 30–40 років*», «*три-чотири рази на століття*», «*можуть виникати в середньому три-чотири рази на сторіччя*».

1.3. Інтегральна сема *‘збудник хвороби / захворювання’*: «*вірус*» (7), «*вірус грипу*» (4), «*мутація вірусу*» (3), «*вірус H1N1*» (2), «*вірус H5N1*», «*вірус типу А*», «*вірус А / H1N1*», «*вірус грипу А / H1N1*», «*вірус грипу А H3N2*», «*пташиний вірус*», «*супервірус*», «*мутація двох типів вірусу – людського і пташиного*», «*вірус мутує*», «*вірус-мутант грипу птиці і людини*», «*варіант вірусу*», «*штам вірусу*», «*штам вірусу грипу*», «*генетично модифікована*», «*вірусна інфекція*», «*грипозна інфекція*», «*виникає, коли вірус людського грипу докорінно змінюється*», «*спричинена поширенням гострої респіраторної хвороби COVID-19, викликаной коронавірусом SARS-CoV-2*», «*новий вірус, який раніше не спричиняв захворювання у населення*».

1.4. Інтегральна сема *‘простір поширення хвороби / захворювання’*: «*охопила весь світ*» (2), «*вразила весь світ*», «*котиться із заходу до центру і на схід України*», «*не тільки українська проблема*», «*прогресує по всьому світу у шаленій динаміці*», «*поширилася більше ніж в одній країні, а то й у цілому світі*», «*швидко поширюються серед населення*», «*покотилася хвиля пандемії по світу*», «*поширення по всій земній кулі*», «*у всіх країнах світу*», «*охопила всі округи*».

1.4.1. Диференційна сема *‘простір поширення хвороби / захворювання (людей)’*: «*створюють загрозу життю і здоров’ю значних верств населення*», «*масове захворювання серед людей у всьому світі*», «*захоплює націю*», «*передавання вірусу від людини до людини в двох регіонах планети*».

1.5. Інтегральна сема 'час поширення хвороби / захворювання': «...сталося...це буквально декілька днів», «за шість днів винищила...», «тривала півтора року».

1.6. Градуальна сема 'ступінь поширення хвороби / захворювання': «глобальна» (4), «світова» (8), «катастрофа» (2), «глобальна проблема» (2), «глобальна катастрофа», «катастрофічна пандемія», «катастрофічних масштабів», «найбільша небезпека», «страшна проблема», «найбільша за всю історію людства», «шоста фаза поширення грипу за класифікацією ВООЗ», «масовий характер».

Під час аналізу виявлено такі назви хвороби, що набувають статусу пандемії: «грип» (75), «пандемія коронавірусу» (54), «свинячий грип» (21), «пташиний грип» (18), «пандемія COVID-19» (18), «грип А / H1N1» (17), «пандемія коронавірусної хвороби» (6), «коронавірусна пандемія» (6), «пандемія «іспанки»» (5), «пандемія чуми» (4), «пандемія COVID» (3), «пандемія СНІДу» (3), «пандемія іспанського грипу» (2), «пандемія діабету» (2), «пандемія туберкульозу» (2), «каліфорнійський грип» (2), «грип АНІШ», «грип H1N1», «грип свинячого каліфорнійського H1N1», «людський грип», «епідемія», «пташиний грип H5N1», «захворюваність на COVID-19», «вірусного характеру COVID-19», «пандемія коронавірусної інфекції», «пандемія ВІЛ і СНІД», «пандемія «чорної смерті» (чуми)», «пандемія цукрового діабету», «пандемія ожиріння», «пандемія холери».

Також зафіксовано 5 словосполук, де лексема пандемія вжита не у прямому значенні інфекційної хвороби: «алкогольна пандемія», «футбольна

пандемія», «пандемічна корупція», «пандемія самотності», «пандемія серед комп'ютерних систем».

**Висновки і перспективи подальших досліджень.** Підводячи підсумок викладених положень, зазначимо, що за допомогою ресурсів корпусного інструментарію дослідники можуть не тільки швидше й ефективніше вирішувати поставлені завдання, а й ставити нові, які раніше здавалися в лінгвістиці нездійсненними через трудомісткість. До таких завдань відносимо і побудову когнітивних матриць ЛСП ПАНДЕМІЯ, зокрема їхні значенні та семантичні поля.

За допомогою корпусного інструментарію вдалося виявити нові значення аналізованого ключового слова пандемія, зокрема назви хвороби, що набувають характеру пандемії та словосполучення, у яких аналізована лексема вжита не у значенні інфекційної хвороби. У результаті компонентного аналізу виокремлено диференційні семи 'несмертоносна хвороба / захворювання', 'штучна хвороба / захворювання', яких не було під час дослідження на основі лексикографічних значень. Робимо висновок, що ресурси корпусної лінгвістики з високим індексом верифікаційного потенціалу математичної обробки даних разом із семантичним аналізом забезпечують новий рівень об'єктивної верифікації мовних фактів у межах нашого підходу до аналізу мовного матеріалу.

Перспективами подальших досліджень є детальне вивчення складу та організації слів-синонімів ключового слова ЛСП ПАНДЕМІЯ за допомогою корпусного інструментарію ГРАК та їх графо-матричному конструюванні.

#### СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Бобер Н. М. Матричне профілювання семантики фразово-дієслівних емотивів у Британському національному корпусі : дис. ... канд. філ. наук : 10.02.21. Київ, 2020. 228 с.
2. Габідулліна А. Р., Колесніченко О. Л. Методологія сучасних лінгвістичних досліджень : навч. посібник. Слов'янськ : Видавництво Б. І. Маторіна, 2019. 322 с.
3. Дарчук Н. П. Корпусна лінгвістика: проблеми, методи, перспективи. Київ : Київський національний університет імені Тараса Шевченка, 2013. 11 с.
4. Демська-Кульчицька О. М. Що таке корпус текстів?. Культура слова. 2004. № 64. С. 35–38.
5. Демська-Кульчицька О. М. Основи національного корпусу української мови. Київ : Інститут української мови НАНУ, 2005. 219 с.
6. Жуковська В. В. Ресурси корпусної лінгвістики у дослідженні історичної динаміки мови. Слово і речення : синтактика, семантика, прагматика / за ред. В. О. Огнев'юк, Л. Л. Хоружа, О. Є. Бондарева. Київ, 2013. С. 151–156.
7. Зубань О. М. Стилеметричні ознаки морфемних структур слів у поетичному мовленні Т. Шевченка (на матеріалі Корпусу української мови). Мовні і концептуальні картини світу, 2014. № 48. С. 165–178.
8. Кривенко Г. Л. Корпусні дослідження дискурсу: становлення, стан і перспективи. Вісник Київського національного лінгвістичного університету. Серія Філологія. 2017. Вип. 20(1). С. 51–63.
9. Петрасова С. В., Хайрова Н. Ф. Сучасні інформаційні технології в лінгвістиці. Харків : ФОП Панов А. М., 2020. 124 с.
10. Смашнюк О. І. Маркери емоційності у спонтанній комунікації (на матеріалі Британського національного корпусу текстів : дис. ... канд. філ. наук : 10.02.04. Київ, 2008. 238 с.
11. Шведова М., Вальденфельс Р. фон, Яригін С., Рисін А., Старко В., Ніколаєнко Т. та ін. Генеральний регіонально анований корпус української мови (ГРАК). 2017–2022. URL: <http://uacorus.org/> (дата звернення: 21.12.2022).

12. Шведова М. Динаміка вживання варіативних форм іменників-назв неістот у знахідному відмінку однини в літературних текстах регіонів України XIX-XXI ст. Науковий вісник Херсонського державного університету. Серія «Лінгвістика». 2018а. № 31. С. 88–94.
13. Шведова М. Динаміка синтаксичних конструкцій на позначення шляху при дієсловах руху : корпусне дослідження. Українська мова. 2018b. № 83. С. 67–79.
14. Яницька О. М. Корпусний аналіз для визначення ключових лексем-номінантів концепту VIRA на матеріалі англійської, української та французької мов. Науковий вісник Міжнародного гуманітарного університету. Серія: Філологія. 2016. Вип. 25 (1). С. 160–162.
15. Asmussen J. Korpuslinguistische Verfahren zur Optimierung lexikalisch-semantischer Beschreibungen. Sprachkorpora – Datenmengen und Erkenntnisfortschritt / Hrsg. von W. Kallmeyer, & G. Zifonun. Berlin, Boston : De Gruyter, 2007. P. 123–151.
16. Bowker L., Pearson J. Working with specialized language: A practical guide to using corpora. London : Routledge, 2002. 256 p.
17. Kilgarriff A., Kosem I. Corpus tools for lexicographers. Electronic Lexicography / eds. S. Granger and M. Paquot. Oxford : Oxford University Press, 2012. P. 31–56.
18. Kyrysiuk A. Lexico-Semantic Graphs as Formalized Schemes of Reflection of Fragment of Reality in Human Consciousness. Language. Culture. Discourse : monograph / ed. R. Vasko. Kharkiv : PC TECHNOLOGY CENTER, 2022. P. 5–21.
19. MacEnery T., Hardie A. (2012). Corpus Linguistics : Method, Theory and Practice. Cambridge University Press.

## REFERENCES

1. Bober N. M. Matrychne profilyuvannya semantyky frazovo-diyeslivnykh emotyviv u Brytans'komu natsional'nomu korpusi. [Matrix Profiling of the Semantics of Phrasal Emotive verbs in the British National Corpus]. Phd thesis. Kyiv, 2020. 228 p. [in Ukrainian]
2. Habidullina, A. R., Kolesnichenko, O. L. Metodolohiya suchasnykh linhvistychnykh doslidzhen'. [Methodology of modern linguistic research] Textbook manual. Slavyansk : Publishing House B. I. Matterin, 2019. 322 p. [in Ukrainian]
3. Darchuk N. P. Korpusna linhvistyka: problemy, metody, perspektyvy. [Corpus linguistics: problems, methods, perspectives]. Kyiv : Taras Shevchenko Kyiv National University, 2013. 11 p. [in Ukrainian]
4. Dems'ka-Kul'chyts'ka O. M. Shcho take korpusu tekstiv? [What is a corpus of texts?]. Word culture. 2004. № 64. pp. 35–38. [in Ukrainian]
5. Dems'ka-Kul'chyts'ka O. M. Osnovy natsional'noho korpusu ukrayins'koyi movy. [Fundamentals of the national corpus of the Ukrainian language]. Kyiv : Institute of the Ukrainian Language of the National Academy of Sciences, 2005. 219 p. [in Ukrainian]
6. Zhukovs'ka V. V. Resursy korpusnoyi linhvistyky u doslidzheni istorichnoyi dynamiky movy. [Resources of corpus linguistics in the study of historical dynamics of language]. Word and sentence : syntax, semantics, pragmatics / za red. V. O. Ohnev'yuk, L. L. Khoruzha, O. YE. Bondareva. Kyiv, 2013. pp. 151–156. [in Ukrainian]
7. Zuban' O. M. Stylemetrychni oznaky morfemnykh struktur sliv u poetychnomu movlenni T. Shevchenka (na materiali Korpusu ukrayins'koyi movy). [tylometric signs of morpheme structures of words in T. Shevchenko's poetic speech (based on the material of the Ukrainian Language Corpus)]. Linguistic and conceptual pictures of the world, 2014. № 48. pp. 165–178. [in Ukrainian]
8. Kryvenko H. L. Korpusni doslidzhennya dyskursu: stanovlennya, stan i perspektyvy. [Corpus studies of discourse: formation, state and prospects.]. Messenger of Kyiv National Linguistic University. Series Philology. 2017. Issue 20(1). pp. 51–63. [in Ukrainian]
9. Petrasova S. V., Khayrova N. F. Suchasni informatsiyi tekhnolohiyi v linhvistytsi. [Modern information technologies in linguistics]. Kharkiv : FOP Panov A. M., 2020. 124 p. [in Ukrainian]
10. Smashnyuk O. I. Markery emotsiynosti u spontanniy komunikatsiyi (na materiali Brytans'koho natsional'noho korpusu tekstiv. [Emotions markers in spontaneous speech (on the basis of the British National Corpus)]. Phd thesis. Kyiv, 2008. 238 p. [in Ukrainian]
11. Shvedova M., Val'denfel's R. fon, Yaryhin S., Rysin A., Starko V., Nikolayenko T. ta in. Heneral'nyy rehional'no anotovanyy korpus ukrayins'koyi movy (HRAK). 2017–2022. URL: <http://uacorpora.org/> (date of application: 21.12.2022). [in Ukrainian]
12. Shvedova M. Dynamika vzhivannya variativnykh form imennykiv-nazv neistot u znakhidnomu vidminku odnyny v literaturnykh tekstakh rehioniv Ukrayiny XIX-XXI st. [The dynamics of the use of variant forms of nouns-names of non-beings in the accusative singular in literary texts of the regions of Ukraine in the 19th-21st centuries]. Scientific Bulletin of Kherson State University. Series «Linguistics». 2018а. № 31. pp. 88–94. [in Ukrainian]
13. Shvedova M. Dynamika syntaksychnykh konstruktiv na poznachennya shlyakhu pry diyeslovakh rukhu : korpusne doslidzhennya. [Dynamics of syntactic constructions for path marking in verbs of motion : a corpus study]. Ukrainian language. 2018b. № 83. pp. 67–79. [in Ukrainian]
14. Yanyts'ka O. M. Korpusnyy analiz dlya vyznachennya klyuchovykh leksem-nominantiv kontseptu VIRA na materiali anhliys'koyi, ukrayins'koyi ta frantsuz'koyi mov. [Corpus analysis to determine the key lexemes-nominants of the concept of VIRA on the material of English, Ukrainian and French languages]. International Humanitarian University Herald. Philology. 2016. Vol. 25 (1). P. 160–162. [in Ukrainian]

15. Asmussen J. Korpuslinguistische Verfahren zur Optimierung lexikalisch-semanticer Beschreibungen. Sprachkorpora – Datenmengen und Erkenntnisfortschritt [Corpus linguistic methods to optimize lexical-semantic descriptions. Speech corpora – amounts of data and progress in knowledge] / Hrsg. von W. Kallmeyer, & G. Zifonun. Berlin, Boston : De Gruyter, 2007. pp. 123–151. [in German]

16. Bowker L., Pearson J. Working with specialized language: A practical guide to using corpora. London : Routledge, 2002. 256 p.

17. Kilgarriff A., Kosem I. Corpus tools for lexicographers. Electronic Lexicography / eds. S. Granger and M. Paquot. Oxford : Oxford University Press, 2012. pp. 31–56.

18. Kyrysiuk A. Lexico-Semantic Graphs as Formalized Schemes of Reflection of Fragment of Reality in Human Consciousness. Language. Culture. Discourse : monograph / ed. R. Vasko. Kharkiv: PC TECHNOLOGY CENTER, 2022. pp. 5–21.

19. MacEnery T., Hardie A. (2012). Corpus Linguistics : Method, Theory and Practice. Cambridge University Press.