

УДК 81'33

ББК 81

## ТЕОРЕТИКО-МЕТОДОЛОГІЧНІ ПІДХОДИ ДО ВИВЧЕННЯ КОЛОКАЦІЙ У СУЧАСНОМУ МОВОЗНАВСТВІ

БОБКОВА Т.В.

*Київський національний лінгвістичний університет*

У статті проаналізовано основні лінгвістичні підходи і корпусний підхід до вивчення колокацій. Окреслено проблеми визначення терміна “колокація” і здійснено аспектуалізацію цього поняття у широкому й вузькому розумінні. Встановлено типологію та розроблено теоретико-методологічні положення підходів до аналізу явища колокації. У якості концептуального підґрунтя формування корпусного підходу взято положення британського контекстуалізму. Обґрунтовано необхідність застосування комбінованого підходу до аналізу колокацій. Розглянуто сучасні українські дослідження колокацій, представлені лексико-граматичним і корпусним підходами.

**Ключові слова:** колокація, британський контекстуалізм, лексико-граматичний і корпусний підходи.

The article deals with the main linguistics and corpus-based approaches toward collocations study. The main problems of “collocation” term definitions are represented and its aspectualization is performed with the consideration of wide and narrow understanding. The approaches to the typology and the theoretical and methodological principles of collocation study are established. The ideas of London structural linguistics school and British contextualism are determined as the conceptual basis of corpus-based approach toward collocation study. The necessity of combining approach use in collocation study is proved. The modern Ukrainian collocation studies are presented by lexical, semantic and corpus-driven approaches.

**Key words:** collocation, British contextualism, lexical, semantic and corpus-driven approaches.

Прикладні системи автоматичного аналізу тексту потребують дослідження великих обсягів природно-мовних текстів. Результати спостережень корпусів текстів засвідчують той факт, що у процесі комунікації використовуються не окремі слова, а регулярно відтворювані конструкції – колокації. Зазначені природно-мовні структури не створюються учасником комунікації безпосередньо під час мовлення, а цілком видобуваються з пам’яті як готові утворення, подібні до простих слів. Регулярність і частотність появи колокацій у текстах становить значний інтерес для прикладних лінгвістів і визначає актуальність дослідження основних підходів до опису цього феномена.

**Мета статті** – дослідити основні підходи до вивчення колокацій у сучасному мовознавстві. Досягнення поставленої мети передбачає виконання таких завдань:

- здійснити аспектуалізацію поняття “колокація” з урахуванням широкого й вузького розуміння терміна;
- встановити типологію підходів до витлумачення колокацій;
- сформулювати теоретико-методологічні положення щодо опису колокацій у межах основних підходів;
- дати визначення терміна “колокація” з огляду на корпусний підхід;
- навести аргументи на користь як корпусного підходу, так і застосування комбінованого багатометодного вектору у вивченні колокацій.

У науковий обіг термін “колокація” вперше був уведений Дж. Р. Фьорзом у межах контекстуальної теорії значення Лондонської школи структурної лінгвістики [26, р. 9]. Основною заслугою представників британського контекстуалізму (J.R. Firth, M.A.K. Halliday, A. McIntoshas,)

було усвідомлення необхідності емпіричного дослідження мови безпосередньо у функціонуванні: встановлення значення можливо саме через аналіз вживання форми [13, с. 344]. У 1930 рр. під колокаціями розуміли типове оточення – звичні або звичайні позиції даного слова [18]. І хоча в межах британського контекстуалізму не було сформульовано чіткого визначення колокації, концептуальні засади напрямку послуговували підґрунтям для контекстно-орієнтованого й корпусного підходів до вивчення феномена. Незважаючи на відносно недавнє входження аналізованого терміна в широкий науковий обіг (1990 рр.), історія вивчення стійких сполучень нараховує майже століття. Окремі загальнотеоретичні аспекти колокацій як прояву синтагматичних відношень становили об'єкт дослідження лінгвістів з початку ХХ ст. (Ю.Д. Апресян, Ш. Баллі, К.Г. Борисова, В.В. Виноградов, І.Р. Вихованець, К.Г. Городенська, В.О. Горпинич, Н.П. Дарчук, А.П. Загнітко, Л.М. Іорданська, Н.Ф. Клименко, М.П. Кочерган, В.В. Левицький, Р. Марчинкевічене, І.О. Мельчук, В. Порциг, С.М. Різник, Дж. Синклер, О.С. Снітко, В.М. Телія, Л. Теньєр, Дж. Фьорз).

Дослідження в галузі корпусної лінгвістики й автоматичного аналізу тексту сприяли поширенню терміна в інші сфери мовознавства – лексикологію, дериватологію, синтаксис, психолінгвістику, лексикографію, теорію перекладу й лінгвістику тексту [24, р. 77] й уможливили окреслення основних теоретичних проблем щодо визначення колокацій, їх суттєвих ознак і статусу серед словосполучень. Певною мірою поширеність терміна пояснюється тими значеннями, які вкладають в поняття колокації різні лінгвісти [25, р. 315]. Безперечно, термін не є одновимірним, оскільки покликаний відобразити комплекс лексичних, граматичних і статистичних ознак. Дискусійний характер точного визначення терміна “колокація” спонукав до аспектуалізації аналізованого поняття й виділення широкого лінгвістично мотивованого значення й вузького – у статистичному аспекті [19, р. 47; 21, р. 109–110; 23, р. 40; 26, р. 26]. На сучасному етапі в широкому сенсі під колокацією розуміється поєднання двох слів або звичне використання певних лексем разом [23, р. 40; 27, р. 115–116]. У вузькому значенні колокація – це слова, які зустрічаються у тексті частіше поруч, ніж за випадковою вірогідністю окремо [27, р. 115–116]. Залежно від широкого або вузького розуміння аналізованого поняття визначається якісний і кількісний склад колокацій [5, с. 18].

Слід зауважити, що у слов'янській лінгвістичній традиції введення терміна “колокація” до цього часу не набуло остаточного закріплення, тому в сучасних працях часто спостерігається співіснування конкурентних найменувань на позначення аналізованого поняття: *стійке сполучення* [3, с. 77; 15, с. 224; 16, с. 31; 22, с. 19], *стабільний контекст*, *стабільна сполука* [11, с. 205], *невільне сполучення* [1, с. 25; 4, с. 235; 6, с. 120; 14, с. 10;], *фразеологізм* [13, с. 345], *лексико-функціональний вираз* [9, с. 239], *фразема* [9, с. 239; 13, с. 345], *неоднослівна лексична одиниця*, *неоднослівна номінація* [1, с. 27], *конструкція* [10, с. 45; 17, с. 399] або *багатослівна конструкція* [1, с. 26; 12, с. 158]. Зазначена неусталеність термінології пояснюється співіснуванням у сучасній лінгвістиці давно сформованих традицій і парадигм корпусної лінгвістики [17, с. 398; 19, р. 47]. З іншого боку, це зумовлено неоднорідністю, багатоаспектністю феномена [14, с. 4; 17, с. 398; 20, р. 133; 24, р. 89], певною мірою невизначеним статусом колокацій, що, у свою чергу, потребує їх досліджень в аспекті статистичної, формальної і семантичної зв'язаності. На сьогодні питання щодо статусу колокацій з їх “характерним регулярно спостережуваним напівідиоматичним зв'язком” елементів [7, с. 7] є дискусійним. За ступенем фразеологізації, або “жорсткості зв'язку” елементів [6, с. 120], колокації традиційно зараховуються до стійких сполучень [3, с. 77; 9, с. 239; 22, р. 19] або об'єднуються в один клас із вільними сполученнями [див. 5, с. 18], або характеризуються проміжним між вільними й фразеологізованими сполуками статусом невільних сполучень [1, с. 25; 4, с. 235; 5, с. 16; 6, с. 120; 14, с. 4; 18; 19, р. 47].

Залежно від інтерпретації концепції колокації у традиціях певної школи в сучасній лінгвістиці виділяють **семантико-синтаксичний, контекстно-орієнтований і корпусно-орієнтований підходи** [5, с. 16; 13, с. 344]. Однак аналіз робіт за останні п'ять років [15; 18; 20; 24] дозволяє на додаток до окреслених вище виокремити **когнітивний і психолінгвістичний підходи**. У традиційних лінгвістичних дослідженнях колокацій найбільш поширеними є підходи, за якими зв'язність визначається семантичною сумісністю елементів словосполучення і/або певною синтаксичною моделлю [14, с. 3; 26, р. 27]. У загальному вигляді типологія лінгвістичних підходів до вивчення колокацій може бути представлена у вигляді таксономії (див. Схема 1).

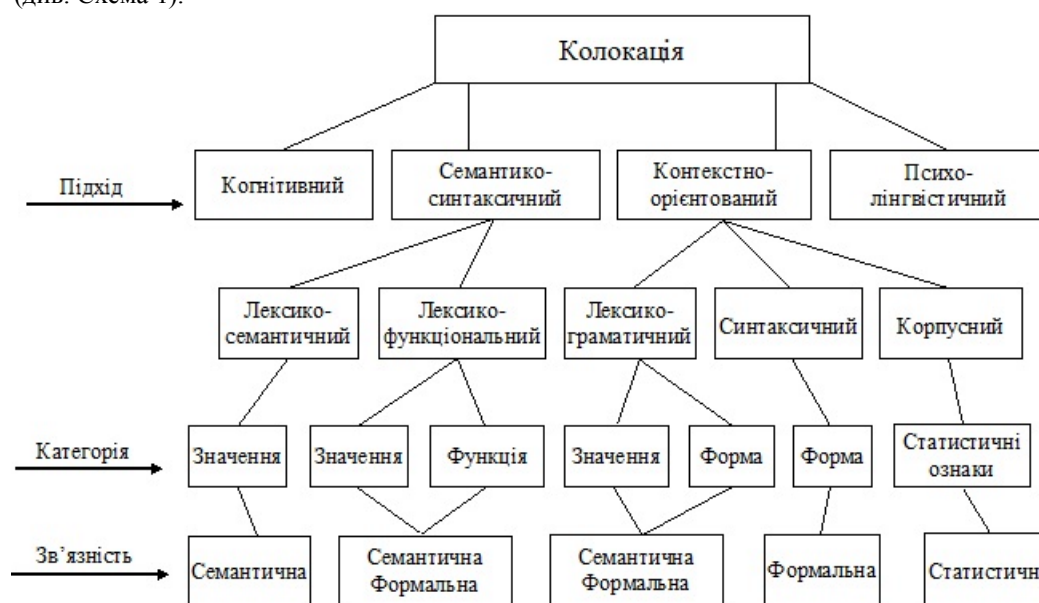


Схема 1. Типологія лінгвістичних підходів до вивчення колокацій

**Семантико-синтаксичний підхід** передбачає аналіз феномена в аспекті теорії стійких сполучень і граматики валентності. У якості основних методів дослідження колокацій застосовуються контекстологічний, структурний (дистрибутивний), метод моделювання і квантитативні методи [11, с. 166–167, 190–206]. В основу вивчення колокацій покладено лексико-семантичні поняття: аналізуються такі параметри, як семантична сумісність лексем, сполучуваність семантичних класів, перенос значення, словозміна стійких сполук і граматичні обмеження сполучуваності [14, с. 9]. У межах семантико-синтаксичного підходу слід виокремити власне *лексико-семантичний, лексико-функціональний, лексико-граматичний і синтаксичний підходи*.

**Лексико-семантичний підхід** вважається найбільш традиційним, оскільки в лінгвістичних дослідженнях з визначення колокації частіше постулюється “лексичний пріоритет” [24, р. 88]. Під **колокацією** розуміються сполучення слів, у складі якого один компонент (базовий або зв'язаний) визначається змістом, темою повідомлення й фразеологізує сполучення (*приймати рішення, брати до уваги*), а інший компонент (вільний) характеризує базове слово і є семантично ключовим [3, с. 77]. У зазначених словосполученнях базовий компонент є аргументом, а інший (або інші, якщо йдеться про багатослівну колокацію) – колокатом, у певному розумінні “функцією” від даного аргумента [14, с. 4–5; 11, с. 192]. Причому лексичне значення колоката відрізняється від основного значення (або основних значень) цього слова в лексичній системі мови, іншими словами – значення колокації не дорівнює сумі значень складових.

Загалом традиційне вивчення колокацій здійснюється в аспекті аналізу лексичної сполучуваності й семантичної зв'язності: під колокацією розуміється певний тип стійкого сполучення, зв'язність якого зумовлена семантичними відношеннями. До основних здобутків лексико-семантичного підходу до вивчення сполучень слів можна віднести розробку класифікації фразеологізмів на підставі семантичної структури (Ш. Баллі, В.В. Виноградов, В.М. Телія, М.М. Шанський, Д.М. Шмельов) і категорії зв'язності значення (В.В. Виноградов, В.М. Телія). За лексико-семантичним підходом, у якості основних ознак колокацій виділяють ідіоматичність, стійкість уживання, відтворюваність у мовленні, неможливість передбачення значення сполучення на основі значень компонентів і наявність самостійного значення залежного компонента [1, с. 78; 22, с. 19–20].

Лексико-функціональний підхід теорії “Смисл-Текст” передбачає аналіз колокацій як підкласу невідільних сполучень, або фразем. При цьому колокація визначається як словосполучення, одна складова якого виступає семантичною домінантою і зумовлює вибір іншої для передачі значення всього словосполучення: *land a job, stand a comparison* [13, с. 345]. Вагомим внеском у розвиток концепції колокації є укладання тлумачно-комбінаторного словника (О.К. Жолковський, І.О. Мельчук), теорія лексичної функції (Ю.Д. Апресян, І.О. Мельчук), класифікація лексико-семантичних відношень (Ю.Д. Апресян) і фразем (Л.М. Іорданська, І.О. Мельчук).

У якості основного методу використовуються лінгвістичне моделювання: побудована в межах теорії “Смисл-Текст” функціональна модель призначена для опису природної мови. Поставлені завдання синтезу тексту – побудови й комбінування обраних засобів у фразі за допомогою тлумачно-комбінаторного словника привели до необхідності розв'язання питань лексичної сполучуваності й, зрештою, до розробки теорії лексичної функції. Визначена в межах теорії “Смисл-Текст” лексична функція ставить у відповідність певному слову або словосполученню  $X$  інше слово, або словосполучення  $Y$ , певним чином зв'язане з  $X$  за змістом [2, с. 43]. Власне, це функція в математичному сенсі – головна лексема  $X$  є її аргументом, а множина можливих (для  $X$ ) виражень смислу  $f$ , тобто  $Y$  – її значенням: якщо  $X = \text{преступление}$ , смисл  $f = \text{делать}$  виражається через  $Y = \text{совершать}$ , а якщо  $X = \text{лекція}$ , то ідентичний смисл виражається через  $Y = \text{читать}$  [22, с. 28]. Теорія лексичної функції широко використовується в сучасних прикладних дослідженнях для опису значення, синтаксичної, актантної структури колокацій і визначення функціональної залежності колоката від основи [7, с. 8; 19, р. 46; 22, с. 28–31;].

У лексико-граматичному підході (К.І. Большакова, К.Г. Борисова, О. Митрофанова, Ф. Хаусманн) колокації розглядаються як семантико-синтаксичні одиниці, або лексично визначені елементи граматичних структур [13, с. 344]. При цьому часто колокації ототожнюються з синтаксичними конструкціями й визначаються як “послідовності таких слів, сумісна вживаність яких істотно перебільшує випадковий рівень” [1, с. 22]. Аналізований підхід зорієнтований на вивчення колокацій через установлені лексико-граматичні шаблони – структурні моделі мовних конструкцій із зазначенням істотних граматичних ознак множини лексем, які входять до складу мовних виразів певного класу, й синтаксичних умов їх вживання [6, с. 120; 14, с. 10–11]. Використання лексико-синтаксичних шаблонів дозволяє виявляти в тексті стійкі сполучення, власні імена, терміни (*міцний чай, Великі Луки, корпусна лінгвістика*) і здійснити класифікацію колокацій за типом відношень між елементами [14, с. 19–20]: *атрибутивні* –  $N+N_2$ , *об'єктні* –  $V+N_2$ , *компаративні відношення* –  $N+Adjcomp+N_2$  та ін.

У цьому розумінні в українському мовознавстві дослідницькі пошуки в основному спрямовано на коло питань, що стосуються системного вивчення лексико-синтаксичної сполучуваності, взаємодії внутрішніх форм слів, а також сполучуваності в аспекті лексико-морфологічної та синтаксичної деривації (І.Р. Вихованець, К.Г. Городенська, В.О. Горпинич, Н. Гуйванюк, Н.П. Дарчук, А.П. Загнітко, Н.Ф. Клименко, М.П. Кочерган, О.С. Снітко). У прикладних

дослідженнях з автоматичного аналізу тексту колокації визначаються як невідільні сполучення, еквівалентні слову, або “фразеологічні звороти, значення яких з погляду сучасної мови зовсім не мотивується й не виводиться із суми значень слів-компонентів”, що зумовлює їх проміжний статус між вільними й стійкими сполуками [6, с. 120].

Синтаксичний підхід базується на звуженому понятті колокації як синтагматично обмеженої лексичної сполучуваності синтаксично пов’язаних елементів (К.І. Большакова, В.А. Гладка, Д. Круз, В. Серетан, А. Тьютін, О.В. Ягунова) і спирається на граматику конструкцій Ч.Дж. Філлмора [18; 1, с. 23]. Дослідження колокацій, за синтаксичним підходом, базується на врахуванні двох ключових ознак колокацій: дотриманні певного порядку слідування компонентів і відсутності синтаксичних трансформацій, що, власне, й відрізняє колокацію від вільних сполучень [5, с. 17]. У такий спосіб, під колокацією розуміється поєднання, або послідовність слів, яка утворює синтаксичну конструкцію (*іменник + прикметник*, *дієслово + іменник (об’єкт)*, *іменник + прийменник + іменник*) [5, с. 17; 26, р. 13], або ідіоматичну сполуку, семантика якої ширше семантики складових [18]. В основу визначення колокацій покладено ознаку синтаксичної стійкості, контактного розташування в тексті й наявності зв’язку підпорядкування [16, с. 31]. Отже, головною метою вивчення колокацій, за синтаксичним підходом, є встановлення синтаксичних та структурних моделей, які забезпечують формальну синтаксичну зв’язність.

Однак спостереження фактичного мовного матеріалу доводять можливість варіацій певних синтаксичних моделей, що є підставою для їх об’єднання в один клас з вільними сполученнями [5, с. 18] або зумовлює точку зору щодо проміжного статусу колокацій. Результати дослідження колокацій, за синтаксичним підходом, широко використовуються в розробці сучасних лексикографічних систем [16, с. 32; 26, р. 28] й автоматичного аналізу тексту. Зокрема на основі синтаксично анотованої бази даних колокацій (80 млн. записів) базується автоматичне укладання тезаурусу в системі Sketch Engine [10, с. 44]. Урахування синтаксичних конструкцій колокацій використовується для морфологічної, поверхневої синтаксичної і семантичної розмітки, зняття граматичної омонімії і розв’язання багатозначності [6, с. 119; 10, с. 45; 17, с. 398]. У межах семантико-синтаксичного підходу колокації розглядаються як комплексні семантико-синтаксичні одиниці, які характеризуються семантичною, синтаксичною і дистрибутивною регулярністю.

Витоки контекстно-орієнтованого підходу коріняться в емпіричній традиції британського контекстуалізму Дж. Фьорза (J.R. Firth, M.A.K. Halliday, A. McIntosh, J. Sinclair) й обґрунтуванні контекстуалізації як прийому контекстологічного [11, с. 167] або синтагматичного аналізу [24, р. 75]. Результатом застосування контекстуалізації на лексичному рівні є встановлення колокацій – типового й постійного оточення слова. Фьорзіанський підхід, або контекстуальна теорія значення, ґрунтується на визнанні колокації цілісною одиницею смислу, дослідження якої передбачає виокремлення складових [24, р. 43]. Під колокаціями, у такий спосіб, розуміються характерні стійкі сполучення слів, поява яких поруч зумовлюється регулярним характером взаємного очікування внаслідок дії семантичних факторів. Основним здобутком фьорзіанства вважаються встановлені ознаки формальної зв’язності: 1) колокація – супровід слова, або матеріал, в якому воно найчастіше зустрічається, і 2) колігація – взаємозв’язок граматичних категорій у синтаксичній конструкції [28, р. 209]. Окреслений вище фьорзіанський підхід працює в рамках контекстуальної теорії значення, за якою ідентифікація сенсу слова або виразу базується на формалізації контекстної моделі.

Концептуальною близькістю до фьорзіанського підходу визначення колокацій є теорія синтагматичних, або елементарних полів В. Порцига [11, с. 202–206; 24, р. 71], розвинута в українській семасіології В.В. Левицьким і Е.Ф. Скороходьком. Зокрема, В.В. Левицький розглядав елементарні поля як різновид стабільних сполучень, відповідних колокаціям [11, с. 202–206]. Застосування квантитативних методів для вивчення параметрів широти,

інтенсивності й селективності сполучуваності дозволило встановити, що перехід від вільних словосполучень до стійких має “градуйований характер”, і високий ступінь зв’язку між компонентами є не випадковим [11, с. 205–206]. Незважаючи на усвідомлення перспектив статистичного дослідження колокацій, зокрема для вирішення проблем лексикографії, теорії і практики (автоматичного) перекладу й лінгводидактики, В.В. Левицький заперечував наявність зв’язку між частотою сумісної вживаності двох слів і ступенем їх синтагматичної зв’язності [11, с. 191]. Отже, традиційні підходи мало що могли запропонувати в аспекті оцінювання ступеня зв’язності.

Поява в 1980-90 рр. електронних корпусів і необхідність статистичного опрацювання великих обсягів текстів вимагали застосування нового підходу на основі розвинутої концепції колокації. В основу формування корпусного підходу, базованого на аналізі статистичної зв’язності, було покладено положення **британського контекстуалізму**. Подібно до контекстно-орієнтованого корпусний підхід визначається опорою на контекст, який розуміється як колекція текстів або одиничні тексти з колекції [1, с. 23; 18]. Відповідно до цього колокація визначається як невипадкове сполучення двох і більше одиниць, характерне як для мови в цілому, так і для текстів певного типу або навіть для (під)вибірки текстів. Завдяки статистичній інтерпретації контекстуальна теорія значення Дж. Фьорза набуває нового виміру в корпусному підході. За корпусним підходом колокація розуміється як послідовність слів, що зустрічаються разом частіше, ніж можна було б очікувати випадково [1; 14; 18; 27].

Основи теоретичних інновацій у концепції колокації було закладено роботами Дж. Синклера “з перевизначення одиниці значення в аспекті корпусних даних” [29, р. 214]. На підставі спостереження корпусних даних щодо невідповідності меж значень і лексичних одиниць було розроблено “корпусну модель значення” [29, р. 214], суть якої полягає в тому, що значення не зосереджується в межах однієї лексичної одиниці, а розповсюджується на послідовність слів у словосполученні – колокації. При цьому, за Дж. Синклером, моделлю вважається сполучення слів, яке двічі зустрілось у тексті [27, р. 57]. У такий спосіб завдяки теоретичним інноваціям Дж. Синклера в галузі корпусної лінгвістики було відкрито шлях до нового розуміння словосполучення в цілому [24, р. 44] і сформульовано теоретичні засади нового підходу до вивчення колокацій. Фактично заслугою Дж. Синклера було застосування засад контекстуальної теорії значення в лексикографічній практиці (проект COBUILD, 1980 р.), що, власне, й завершило виокремлення напрямку нео-фьорзіанства в сучасному мовознавстві [28, р. 137].

За корпусним підходом, феномен колокації розглядається як вияв принципу ідіоматичності, який разом з принципом необмеженого вибору зумовлює побудову тексту на рівні структури й значення [27, р. 110]. При цьому під ідіоматичністю, або фразеологічністю, розуміється наявність у будь-якій мові великої кількості готових фраз або їхніх напівфабрикатів, використовуваних мовцем відповідно до індивідуальних преференцій [24, с. 60]. На основі узагальнення спостережень корпусних даних Дж. Синклером встановлено характерні ознаки колокації як вияву ідіоматичності [27, р. 112–113], подані нижче в узагальненому вигляді:

- 1) невизначеність обсягу великої кількості фраз у мові (*set eyes on*);
- 2–4) можливість внутрішньої лексичної (*in some cases / in some instances*), синтактичної варіативності (*not / hardly / scarcely*) і зміни порядку слів (*recriminate isn't in his nature / isn't in his nature of an academic*);
- 5–6) зумовленість появи високочастотних слів у колокаціях (*hard work, hard evidence*), тенденції до сумісної вживаності слів у певних граматичних конструкціях (*set about testing*) і в певному семантичному оточенні (*happen – accident*).

Отже, на відміну від аналізованих вище підходів у межах корпусного підходу поняття колокації та її традиційних ознак зв’язності та ідіоматичності [1, с. 128] переосмислюється й певним чином формалізується: статистичні чинники вважаються визначальними для синтагматичної зв’язності. Саме тому корпусний підхід доцільно кваліфікувати як статистичний: колокації

визначаються як статистично стійкі словосполучення [14, с. 4–5], або навіть “суто статистичний феномен” [19, р. 47]. Зняття обмежень щодо якісного й кількісного складу забезпечує переваги корпусного підходу в розпізнанні колокацій у тексті, виявленні граматично некоректних конструкцій (*I can't stand being alone, She bothered to do it*) [там само, р. 47] і, зрештою, встановленні лінгвістично релевантних колокацій. Статистичні міри асоціації (MI, score, t-score і log-likelihood) визнаються показниками сили смислового [18] і синтагматичного зв'язку [14, с. 4] між елементами словосполучення. У такий спосіб використання статистичного апарату становить перевагу корпусного підходу, оскільки уможливорює встановлення об'єктивних критеріїв дослідження колокацій і визначення певного зв'язку між смислом і структурою мови. Це забезпечує вихід корпусної моделі на перший план зарубіжної корпусної лексикографії: колокація визнається основною одиницею значення й ужитку, яку неможливо описати традиційними категоріями [29, р. 206] (див. Схема 1.). При цьому певні статистично стійкі сполуки можуть бути як вільними, так і фразеологізованими, тому розглядаються як певний континуум, шкала без чітких меж посеред різних типів сполук [14, с. 10; 19, р. 48]

На противагу інтерпретації, базованій на лінгвістичній інтуїції дослідника, корпусний підхід завдяки застосуванню статистичних методів опрацювання природно-мовного тексту сприяє наближенню до “точного наукового розрахунку значення” лексичної одиниці, подібно до того, як у фізиці можна розрахувати рух [28, р. 137]. Відповідно до концепції колокації лексична одиниця набуває певного сенсу не в ізольованому положенні, а лише в контексті, оскільки значення зумовлюється колокатами – релевантними словами. У такий спосіб, формування певного лексичного значення є результатом колокації – поєднання слів у даному контексті, регулярно повторюваному в корпусі текстів [там само, р. 137–138]. Саме дослідження колокацій на базі статистичних методів покладено в основу встановлення семантико-синтаксичних ознак лексичних одиниць й укладання корпусних словників [8; 27].

З огляду на особливості впровадження корпусного підходу слід зазначити, що в сучасній українській лінгвістиці дослідження колокацій ґрунтуються на засадах семантико-синтаксичного (словник фразеологізмів [6, с. 174], колокаційна лексикографічна система [16, с. 32]) й статистичного (лексикон багатослівних сполук [12, с. 163], словник колокацій для бази знань криміналістичних інформаційних систем [8, с. 184]) підходів. При цьому увагу дослідників в основному зосереджено не на розв'язанні лінгвістичних проблем фразеологізації еквівалентних слову стійких сполучень, а на ефективності автоматичного виявлення цих сполук у текстах [6, с. 119–120; 8, с. 184; 12, с. 158–159]. “Непередбачуваність” природи колокацій зумовлює їх включення до реєстру особливого словника, призначеного не скільки для розпізнання, а для кодування тексту, як і більшість традиційних словників [19, р. 42]. Зокрема, в системі АГАТ колокації поруч з “безумовними ідіомами” заносяться до інвентарних одиниць електронного словника фразеологізмів і при здійсненні автоматичного синтаксичного аналізу “згортаються” в один вузол [6, с. 120–121].

Отже, перспектива об'єктивного дослідження колокацій стає можливою лише з появою корпусу текстів зі статистичним програмним забезпеченням, здатним виявити, які слова сполучаються з аналізованою лексичною одиницею, у який спосіб, наскільки часто це відбувається і що можна очікувати від випадкового розподілу. У цьому розумінні кінцева мета корпусного підходу дослідження колокацій, а саме: встановлення закономірностей і формулювання узагальнень, що викладаються у вступі до словників [28, р. 138], збігається з метою лексикографічного аналізу. Запропонована Дж. Синклером концепція колокації стає для європейських корпусних лінгвістів моделлю, яка не тільки обґрунтовує використання, а й розширює теорію природної мови [24, р. 62, 75].

Однак, слід розуміти, що в межах одного підходу або різних, часом конкуруючих шкіл, формування всеохоплюючої концепції колокації вкрай ускладнюється [19, р. 47]. Саме тому застосування лише корпусного підходу для опису колокацій, безперечно, недостатньо: виявлені

в корпусі зазвичай “невидимі, очікувані й характерні моделі” лише окреслюють контури теорії мови [28, р. 144]. Поєднання корпусного підходу з традиційними створює перспективи виходу не тільки за межі простої статистики, а й деяких звичних уявлень про мову [24, р. 53–57]. Вивчення колокацій має ґрунтуватись на багатоаспектному підході, який передбачає використання корпусного підходу як доповнення до традиційних, а в жодному разі не як єдино вірної теоретико-методологічної основи дослідження. Можливості корпусного підходу сприяють його поєднанню з іншими в сучасних дослідженнях колокацій: синтаксичного й корпусного [26, 1, 16], семантико-синтаксичного й корпусного [14, 6], когнітивного й корпусного з метою вивчення метафоричної природи феномена [24; 15], корпусного й психолінгвістичного – для вивчення ментального лексику, до якого включають колокації [20].

Здійснене дослідження дозволяє зробити такі висновки: 1. Аспектуалізація поняття “колокація” передбачає виділення широкого лінгвістично мотивованого значення й вузького – статистичного. 2. У межах семантико-синтаксичного підходу колокації визначаються як стійкі сполучення і характеризуються проміжним статусом між вільними й стійкими сполуками. 3. За корпусним підходом колокації розглядаються як певний континуум між вільними й стійкими сполуками. 4. Основними ознаками колокацій слід уважати: стійкість уживання сполук, наявність у складі ядра й колоката, непередбачуваність значення цілого сполучення, тенденція до появи колокацій у певних граматичних конструкціях і в певному семантичному оточенні, можливість лексичної і синтаксичної варіативності. 5. Феномен колокації вимагає дослідження статистичної, формальної і семантичної зв’язаності з урахуванням мовних і статистичних чинників.

#### ЛІТЕРАТУРА

1. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика : [учеб. пособ.] / [ред. кол. Е.И. Большакова, Э.С. Клышинский, Д.В. Ландэ и др.]. – М. : МИЭМ, 2011. – 272 с.
2. Апресян Ю.Д. Избранные труды / Юрий Дереникович Апресян // Том I. Лексическая семантика. – М. : “Языки русской культуры”, 1995. – 472 с.
3. Борисова Е.Г. Фразеологическое значение в устойчивых словосочетаниях / Е.Г. Борисова, О.В. Захарова // Филологические науки. – 1994. – № 4. – С. 77–84.
4. Влавацкая М.В. Критерии определения типа комбинаторного словаря / М.В. Влавацкая // Мир науки, культуры, образования. – 2011. – № 3 (28). – С. 235–241.
5. Гладка В.А. Структурно-синтаксичний підхід у вивченні колокацій (на матеріалі французької мови) / В.А. Гладка // Наукові записки національного університету “Острозька академія”. Серія “Філологічна”. – 2013. – Вип. 39. – С. 16–20.
6. Дарчук Н. Комп’ютерне анотування українського тексту: результати і перспективи : [монографія] / Наталія Дарчук. – К. : Освіта України, 2013. – 544 с.
7. Дяченко П.В. Разработка компьютерных методов обучения владению языком с помощью аппарата лексических функций : автореф. дис. на соиск. уч. степ. канд. тех. наук : спец. 05.13.17 “Теоретические основы информатики” / П.В. Дяченко. – М., 2008. – 20 с.
8. Зацеркляний М.М. Об’єктно-орієнтований тезаурус і словник колокацій для бази знань криміналістичних інформаційних систем / М.М. Зацеркляний, Д.Ю. Узлов // Системи обробки інформації. – 2013. – Вип. 2. – С. 183–186.
9. Иорданская Л.Н. Смысл и сочетаемость в словаре / Л.Н. Иорданская, И.А. Мельчук. – М. : “Языки славянских культур”, 2007. – 672 с.
10. Кобрицов Б.П. Снятие лексико-семантической омонимии в новостных и газетно-журнальных текстах : поверхностные фильтры и статистическая оценка / Б.П. Кобрицов, О.Н. Ляшевская, О.Ю. Шеманаева // Интернет-математика 2005. Автоматическая обработка веб-данных. – 2005. – С. 38–57.



11. Левицкий В.В. Семасиология / Виктор Васильевич Левицкий. – Винница : Нова Книга, 2006. – 508 с.
12. Романюк А. Розпізнавання багатослівних конструкцій / А. Романюк, Г. Кваснюк, М. Романишин // Вісник Національного університету “Львівська політехніка”. Серія Комп’ютерні системи проектування. Теорія і практика. – 2011. – № 711 – С. 158–165.
13. Хохлова М.В. Экспериментальная проверка методов выделения коллокаций / М.В. Хохлова // Slavica Helsingiensia 34. Инструментарий русистики : корпусные подходы. – 2008. – С. 343–357.
14. Хохлова М.В. Исследование лексико-синтаксической сочетаемости в русском языке с помощью статистических методов (на базе корпусов текстов) : автореф. дис. на соиск. уч. степ. канд. филол. наук : спец. 10.02.21 “Прикладная и математическая лингвистика” / М.В. Хохлова. – СПб., 2010. – 26 с.
15. Шилихина К.М. Дискурсивная практика иронии: когнитивный, семантический и прагматический аспекты : дис. ... доктора филол. наук : 10.02.19 “Теория языка” / Ксения Михайловна Шилихина. – Воронеж, 2014. – 399 с.
16. Шкурко В.В. Лексикографічний агент екстракції колокацій у природномовному тексті / В.В. Шкурко // Вісник Київського національного університету імені Тараса Шевченка. Серія : Літературознавство. Мовознавство. Фольклористика. – 2012. – № 28. – С. 31–35.
17. Ягунова Е.В. Исследование контекстной предсказуемости единиц текста с помощью корпусных ресурсов / Е.В. Ягунова // “Корпусная лингвистика–2008” : [труды междунар. конф.], (6–10 октября 2008 г., г. Санкт-Петербург). – СПб., 2008. – С. 396–403.
18. Ягунова Е.В. [Электронный ресурс] / Е.В. Ягунова, Л.М. Пивоварова // Acta linguistica petropolitana. Труды Института лингвистических исследований РАН. – 2011. – Режим доступа : [http://www.webground.su/data/lit/pivovarova\\_yagunova/Ot\\_kollokatsiy\\_k\\_konstruktsiya.pdf](http://www.webground.su/data/lit/pivovarova_yagunova/Ot_kollokatsiy_k_konstruktsiya.pdf)
19. Fontenelle Th. What on earth are collocations? / Th. Fontenelle // English today : the International Review of the English Language. – 1994. – Vol. 10 (40). – No. 4. – P. 42–48.
20. Durrant Ph. Are high-frequency collocations psychologically real? / Ph. Durrant, A. Doherty // Corpus Linguistics and Linguistic Theory. – Berlin–New York : De Gruyter, 2010. – Vol. 6. – No 2. – P. 125–155.
21. Čermák F. Linguistic Units and Text Entities: Theory and Practice / F. Čermák // EURALEX’98 Proceedings. – Liege, 1998. – P. 109–118.
22. Kapstad M. Faste uttrykk i russisk og norsk med henblikk på russiskundervisning for nordmenn : Masteroppgave i russisk språk ved. / Margarita Kapstad. – Veron, 2006. – 124 с.
23. Marcinkevičienė R. Kolokacija : tyrimo kryptys, metodai, mokyklos / R. Marcinkevičienė // Lituania. – 1995. – No 2. – P. 40–54.
24. Marcinkevičienė R. Lietuvių kalbos kolokacijos : [monografija] / R. Marcinkevičienė . – Kaunas : Vytauto Didžiojo universiteto leidykla, 2010. – 212 p.
25. Meer V.G. Collocations as One Particular Type of Conventional Word Combinations. Their Definition and Character / V.G. Meer // EURALEX’98 Proceedings. – Liege, 1998. – P. 313–322.
26. Seretan V. Syntax-Based Collocation Extraction / Violeta Seretan. – Berlin : Springer Science & Business Media, 2011. – 232 p.
27. Sinclair J. Corpus, Concordance, Collocation / J. Sinclair. – Oxford : Oxford University Press, 1991. – 200 p.
28. Teubert W. Linguistique de Corpus: un Alternative / W. Teubert // Semen. – Besançon : PUFC, 2009. – T. 27. – P. 130–152.
29. Tognini-Bonelli E. Corpus Classroom Currency / E. Tognini-Bonelli // Darbai ir Dienos. – 2000. – No 24. – P. 205–243.

*Дата надходження до редакції 18.11.2014*