

Lysak H. O. Formation of New Nominative Units in Modern English Under the Influence of Innovative Technologies.

Summary

The article deals with nominative processes that have appeared under the influence of new technologies. That provoked the development of new lexical units. The concepts of “nomination” and “nominative unit” have been analyzed. The term nomination is considered to be very complicated concept. Under the nomination scientists understand the formation of linguistic units, that serve for naming new concepts in form of linguistic units – words, word combinations, phraseological units. Main types of nomination have been considered: primary and secondary nomination. It has been defined that the result of nomination process is the nominative unit.

In modern English there are a lot of examples of forming new lexical units dealing with computers, information and communication technologies. In this case, innovations provoke the nominations for well-known phenomena and technologies. It has been considered that the cognitive basis for nominations is formed by the principles of analogy and contrast. It has been defined that the most productive ways of new concept nominations are semantic derivation, conversion and word formation.

Key words: nomination, nominative unit, word formation, conversion.

УДК 81'42'322.2'324'06-13:004

Лотоцька Н. Я.,
аспірант кафедри прикладної лінгвістики
Національного університету «Львівська політехніка»
E-mail: nata07lototska@gmail.com

ДОСЛІДЖЕННЯ КОЛОКАЦІЙ У СУЧАСНОМУ МОВОЗНАВСТВІ: КОРПУСНО-СТАТИСТИЧНИЙ ПІДХІД

Статтю присвячено дослідженню поняття колокація в сучасному мовознавстві. Описано критерії інтерпретації цього терміна крізь призму контекстно-орієнтованого, семантико-синтаксичного, корпусного та корпусно-статистичного підходів. Особливу увагу зосереджено на важливості корпусів текстів та статистичних методів під час дослідження цих словосполук. Розглянуто та проаналізовано найпоширеніші в лінгвістиці статистичні показники асоціації для обчислення ступеню зв'язку між компонентами колокації.

Ключові слова: колокація, нод, колокат, корпус текстів, статистичні показники асоціації.

Об'єктом вивчення традиційного мовознавства довгий час було слово. Завдяки працям Ф. де Сосюра науковий інтерес змістився зі слова на синтагматичні зв'язки, оскільки в мовленні слово не є ізольованою одиницею, а через асоціації завжди комбінується з іншими словами [14, с. 67]. Нині лінгвісти зацікавлені в аналізі високочастотних стійких синтагматичних послідовностей – колокацій – регулярно відтворюваних конструкції, що видобуваються з пам'яті як готові утворення безпосередньо під час мовлення, подібно до простих слів [2, с. 14].

Інформація про стійкі сполуки є актуальною та необхідною для вирішення різних прикладних завдань у лінгвістиці: укладання словників, тезаурусів, баз даних, автоматичного аналізу тексту, машинного перекладу, інформаційно-пошукових систем, зокрема, потреби розв'язання завдань автоматичного аналізу тексту [2, с. 18].

Окремі загальнотеоретичні аспекти колокацій як прояву сполучуваності були об'єктом дослідження лінгвістів ще на початку ХХ ст. З появою праць Дж. Фьорза зароджується теоретичне дослідження явища колокації й традиція вивчення цього поняття в різних аспектах. Аналіз фахової літератури показує, що основи теоретичних інновацій у концепції колокації закладено працями Дж. Синклера [22]. Ш. Еверт, В. Левицького, М. Стабс та ін. [9; 17; 23] застосовують квантитативні методи для вивчення параметрів сполучуваності, досліджуючи колокації за допомогою статистичних показників. Проблемі вилучення колокацій із текстового масиву й встановленню критеріїв їхньої ідентифікації присвячено праці Т. Бобкової, М. Хохлової, В. Захарова та ін. [2; 7; 12].

Мета цієї праці – описати та проаналізувати підходи дослідження поняття колокація в сучасному мовознавстві. Досягнення поставленої мети передбачає виконання таких завдань: 1) описати критерії інтерпретації поняття колокація в сучасному мовознавстві; 2) розглянути термін колокація в аспекті корпусного та корпусно-статистичного підходів; 3) визначити важливість

статистичних методів для дослідження сполучуваності слів; 4) проаналізувати статистичні показники асоціації для обчислення ступеню зв'язку між компонентами словосполучки.

У науковий обіг термін *колокація* ввів Дж.Р. Фьорз у межах британського контекстуалізму. За Дж. Р.Фьорзом, *колокації* – це типове оточення слова, зі семантичним зв'язком між її компонентами [18].

Залежно від інтерпретації концепції колокації у традиціях певної лінгвістичної школи виокремлюють контекстно-орієнтований, семантико-синтаксичний, корпусний та корпусно-статистичний підходи.

У межах контекстно-орієнтованого підходу поняття *колокація* визначено як типове й постійне оточення слова [15; 18; 21]. Цей термін застосовують до частотних синтагматичних комбінацій слів, навіть до таких, між якими немає особливих семантичних зв'язків. Таке розуміння *колокації* характерне для досліджень із корпусної лінгвістики і лексикографії.

Представники семантико-синтаксичного підходу [1; 3; 6; 19; 20; 22] трактують *колокації* як комплексні семантико-синтаксичні одиниці, яким властива семантична, синтаксична й дистрибутивна регулярність, з проміжним статусом між вільними й стійкими сполучками.

З появою праць Дж. Синклера було відкрито шлях до нового розуміння словосполучки загалом і сформульовано теоретичні засади *корпусного підходу* вивчення колокацій [2; 5; 11; 22]. У межах цього підходу *колокації* розглядають як словосполучки, що трапляються в тексті частіше поряд, ніж за випадковою вірогідністю окремо [12, с. 115–116]. Ш. Еверт додає, що «*колокація* – словосполучка, семантичні і синтаксичні властивості якої не можуть бути повністю передбачені зі значень її окремих компонентів» [17, с. 17]. Отож, лексична одиниця набуває певного сенсу не ізольовано, а лише в контексті, оскільки значення зумовлено колокатами – релевантними словами [24, с. 137–138].

Корпуси текстів (КТ) й інструментарій корпусної лінгвістики дають змогу виявити й суттєво розширити лексичний фонд стійких словосполук різного типу й особливості їх використання [7].

З появою електронних корпусів з'явилася можливість статистичного опрацювання великих масивів текстів. У лінгвістиці широко застосування набув *корпусно-статистичний підхід* [2; 5; 9; 11; 12], базований на аналізі статистичної зв'язності компонентів словосполучки.

Так, В. Левицький, один із представників *корпусно-статистичного підходу*, трактує колокацію як «звичне асоціювання слова з іншими словами в реченнях» і наголошує на «новому способі аналізу синтагми, який ґрунтується на лексичній очікуваності користувачів мови стосовно ймовірності появи тих чи інших комбінацій лексичних одиниць, які творять разом синтагму» [8, с. 28].

Статистичні методи в лінгвістиці дозволяють отримати достовірні статистичні дані сполучуваності лексичних одиниць на базі КТ, вивчати лексичні одиниці в контексті, отримувати дані про частоту словоформ, лексем, граматичних категорій, спільну появу лексичних одиниць, особливості їхньої сполучуваності тощо. До того ж результати пошуку можна ранжувати за різними параметрами та задавати порогові значення, що уможливило отримання значущої інформації [12, с. 66].

Треба зазначити, що в статистичній лінгвістиці термін *колокація* позначає «статистично стійкі поєднання» слів [12, с. 8], а дослідження цих конструкцій базуються на математичних критеріях – статистичних показниках асоціації, основою яких є теорія ймовірності та математична статистика. «Мовна система є по своїй суті ймовірнісною, і частота в тексті є ілюстрацією граматичної ймовірності» [19, с. 31]. Це означає, що слова в мовленні підпорядковуються граматичним правилам мови й не можуть вживатися довільно в мовному потоці.

Одним з основних інструментів роботи з корпусними даними для вивчення колокацій є конкорданс – рядки КТ заданої довжини, що представляють досліджуване слово у контексті. Рядки конкордансу є джерелом інформації про закономірності вживання досліджуваного слова (нода) і зв'язку між іншими словами (колокатами).

Структура колокацій представлена залежними зв'язками між компонентами словосполучки: основного елемента (нода) і залежного (колоката). Колокації можуть мати кілька варіантів колокатів, що належать ноду, до того ж зв'язок між нодом і колокатом досить гнучкий, їм властива умовність, семантична самостійність і відсутність лексикалізації словосполучки [22].

Комбінації слів мають різний ступінь стійкості зв'язку між компонентами, тим часом корпусно-статистичний підхід уможливило встановлення об'єктивних критеріїв дослідження колокацій і визначення певного зв'язку між смислом і структурою мови [2, с. 19].

Оцінка сили зв'язку залежить від типу одиниць (лем або словоформ), статистичні показники яких, використовують для розрахунків. Частота спільної появи сполучень слів є одним із способів визначення їхньої стійкості, що своєю чергою пов'язана з частотою окремих компонентів колокації.

У розпорядженні сучасної лінгвістики є кілька способів для обчислення ступеня синтагматичної близькості між компонентами колокації, об'єднаних терміном статистичні показники асоціації – числові індекси, що обчислюють силу статистичної залежності двох або більше квалітативних змінних» [16, с 241], як-от: MI-score, t-score log-likelihood, Dice, gmean та ін. [11, с 278]. Ш. Еверт вважає, що за допомогою статистичних критеріїв можна обрахувати величину асоціації між компонентами сполуки» [17, с. 75]. Дослідження цього напряму ґрунтуються на знаходженні n-грам (біграм або триграм тощо) у межах заданого діапазону та порівнянні їхніх частот.

«Величини, обчислені показником асоціації, можуть бути інтерпретовані в різний спосіб: вони можуть бути використані для оцінки величини асоціації між компонентами парного типу; для ранжування парних типів у наборі даних; для оцінки парних типів із визначеним першим або другим компонентом» [17, с 75].

Коефіцієнти асоціації застосовують до стрижневого слова (нода), потім список кандидатів розподіляють за кількісними значеннями. У верхній частині списку розміщують словосполучення з найсильнішими зв'язками – наймовірнішими кандидатами колокації. Потрібно враховувати частоту спільної появи стрижневого слова (нода) і його колоката, аналізуючи наскільки випадкова сила зв'язку між сусідніми словами.

Розгляньмо найпоширеніші статистичні показники асоціації в статистичній лінгвістиці. Наприклад, показник VI (показник взаємної інформації, MI, MI-score) демонструє притягання між двома словами, або ймовірність того, що якщо одне слово з'являється в рядку, слово, яке складає з ним колокацію, з'явиться в найближчій околиці досліджуваного слова [4]. Показник VI дає змогу виокремити стійкі словосполучення, власні назви, спеціальні терміни. Його широко використовують для інформаційного пошуку, що уможливує отримання інформації про сполучуваність термінів, що характерно для різних сфер знань. Недоліком використання цього коефіцієнта є завищення результатів для рідковживаних словосполучень [23].

T-показник (t-score, t-test, t-критерій, T-value) відображає частоту спільної появи стрижневого слова і його колоката, відповідаючи на питання, наскільки не випадковою є сила асоціації між колокатами. За допомогою цього коефіцієнта можна виміряти вірогідність появи слова як парного в поєднанні з яким-небудь іншим словом. На відміну від показника взаємної інформації (VI), t-показник вимірює тяжіння слова, яке сполучається, з досліджуваним словом, але не навпаки. Отже, список таких слів буде містити велику кількість граматичних слів, які важливі для існування досліджуваного слова [10, с 303].

Недоліками застосування t-score є виокремлення колокацій з частотними словами, зокрема, зі службовими словами. Для цього критерію необхідно складати стоп-листи, щоб «відкинути» найчастотніші слова (прийменники, займенники або сполучники) [23].

Жоден із перерахованих вище способів і показників не можна виокремити як найкращий у дослідженні колокацій [4]. Е. Ягунова, Л. Пивоварова дійшли висновку про те, що принципово відрізняються списки колокацій, отримані за допомогою VI та t-score: VI найкраще дає змогу виокремити низькочастотні сполуки – назви об'єктів, терміни, складні номінації; t-score, навпаки, краще працює для вилучення високочастотних конструкцій – «загальномовних стійких сполук» (похідних службових слів, дискурсивних слів) й «стійких конструкцій» [13]. MI-score є кращим для знаходження схожості, тоді як t-score – для встановлення відмінностей.

У результаті дослідження з'ясовано, що колокації – це мовні одиниці, широко розповсюджені в лексико-семантичній системі, які вживаються за суворими правилами комбінаторики й реалізують свій потенціал у процесі утворення мовних конструкцій. КТ є надійним матеріалом для статистичного аналізу будь-яких мовних або мовленнєвих одиниць, здатним виявити, які слова сполучаються з досліджуваною лексичною одиницею, у який спосіб, як часто це відбувається і що можна очікувати від випадкового розподілу. Для отримання достовірних показників сполучуваності слів запропоновано використовувати статистичні показники асоціації.

Список використаної літератури

1. Ахманова О. Очерки по общей и русской лексикологии / Ахманова О. – М. : Просвещение, 1957. – 362 с.
2. Бобкова Т. В. Теоретико-методологічні підходи до вивчення колокацій / Т. В. Бобкова // Вісник Київського нац. лінгв. ун-ту. Серія: Філологія. – 2014. – Т. 17. – № 2. – С. 14–22.
3. Влавацкая М. В. Понятия коллокации и коллигации в диахроническом рассмотрении / М. В. Влавацкая // Актуальные проблемы филологии и методики преподавания иностранных языков: мат-лы Междунар. науч.-практ. конф. / отв. ред. Ю. Н. Тамберг. – Новосибирск : Изд-во НГПУ, 2011. – С. 19–26.
4. Горина О. Г. Применение методов корпусной лингвистики для определения контекстносpezifических слов и коллокаций / О. Г. Горина // Вестник ЛГУ им. А. С. Пушкина. – 2011. – № 3. – С. 27–36.
5. Дарчук Н. П. Статистичні характеристики лексики як відображення структури тексту // Мовознавчі студії. – К. : Наукова думка, 1976. – 102 с.
6. Єрмоленко С. Я. Українська мова: короткий тлумачний словник лінгвістичних термінів / С. Я. Єрмоленко, С. П. Бибик, О. Г. Тодор; [за ред. С. Я. Єрмоленко]. – К. : Либідь, 2001. – 224 с.
7. Захаров В. П. Сочетаемость через призму корпусов // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Москва, 27 – 30 мая 2015 г.). – Вып. 14 (21) : в 2 т. – Т. 1. – М. : Изд-во МГГУ, 2015. – С. 667–682.
8. Левицкий В. В. Семасиология. – Винница : Нова книга, 2006. – 512 с.
9. Левицкий В. В., Огуй О. Д., Кійко С. В., Кійко Ю. Є. Апроксимативні методи вивчення лексичного складу : навчальний посібник. – Чернівці : Рута, 2000. – 136 с.
10. Левченко О. Гендерні особливості української художньої прози (за даними аналізу частотності колокацій у текстах 60-80-х та 2000-х років) / О. Левченко, Я. Бутельський // XX–XXI століття: жанрово-стильові й лінгвістичні метафорози в українській мові та літературі : монографія. – Оломоуць, 2016. – С. 299–334.
11. Ситар Г. В. Синтаксичні фразеологізми в розрізі конструкційної граматики : монографія : [наук. і відп. ред. А. П. Загнітко] / Г. В. Ситар. – Винница : ТОВ «Нілан-ЛТД», 2017. – 458 с.
12. Хохлова М. В. Исследование лексико-синтаксической сочетаемости в русском языке с помощью статистических методов (на базе корпусов текстов): дисс. на соискание ученой степени канд. филол. наук: спец. 10.02.21 «Прикладная и математическая лингвистика» / М. В. Хохлова / Санкт-Петербургский государственный университет. – Санкт-Петербург, 2010. – 211 с.
13. Ягунова Е. В. Природа коллокаций в русском языке. Опыт автоматического извлечения и классификации на материале новостных текстов / Е. В. Ягунова, Л. М. Пивоварова // Сб. НТИ. – Серия 2. – № 5. – 2010. – [Электронный ресурс]. – Режим доступа : http://jf.spbu.ru/upload/files/file_1357769727_3479.pdf
14. Bally Ch. Traité de stylistique française / Ch. Bally. – Heidelberg, 1909.
15. Coseriu E. Lexikalische Solidaritäten / H. Coseriu // Geckeler H. Strukturelle Bedeutungslehre. – Darmstadt, 1967. – P. 239–253.
16. Everitt B. S. The Cambridge Dictionary of Statistics. 2nd edition / B. S. Everitt. – Cambridge : Cambridge University Press, 2002. – 410 p.
17. Evert S. The Statistics of Word Cooccurrences: Word Pairs and Collocations : PhD dissertation, IMS, University of Stuttgart, 2004 (Published in 2005) / S. Evert. – 353 p. — Free PDF available from <http://purl.org/stefan.evert/PUB/Evert2004phd.pdf>
18. Firth J. R. Modes of Meaning / J. R. Firth // Papers in Linguistics 1934-1951. – London : Oxford University Press, 1957. – P. 190–215.
19. Halliday M. A. K. Current ideas in systemic practice and theory / M. A. K. Halliday. – London : Pinter, 1991. – 157 p.
20. Hausmann F. J. Was sind eigentlich Kollokationen? / F. J. Hausmann // Steyer K., Porzig W. Wesenhafte Bedeutungsbeziehungen // Schmidt L. Wortfeldforschung. Zur Geschichte und Theorie des sprachlichen Feldes. – Darmstadt, 1973. – P. 78–103.
21. Palmer H. E. Second Interim Report on English Collocations / H. E. Palmer. – Tokyo : Kaitakusha, 1933.
22. Sinclair J. Corpus. Concordance and Collocation / J. Sinclair. – Oxford : Oxford University Press, 1991. – 137 p.
23. Stubbs M. Collocations and semantic profiles: On the cause of the trouble with quantitative studies / M. Stubbs // Functions of Language. – 1995. – Vol. 1 – P. 23–55.
24. Teubert W. Linguistique de Corpus: un Alternative / W. Teubert // Semen. – Besanzon : PUFC, 2009. – Т. 27. – P. 130–152.

Лотоцкая Н. Я. Исследование коллокации в современном языкознании: корпусно-статистический подход.

Аннотация

Статья посвящена исследованию понятия коллокация в современном языкознании. Описаны критерии интерпретации этого термина сквозь призму контекстно-ориентированного, семантико-синтаксического, корпусного и корпусно-статистического подходов. Особое внимание сосредоточено на важности корпусов текстов и статистических методов при исследовании этих словосочетаний. Рассмотрены и проанализированы наиболее распространенные в лингвистике статистические показатели ассоциации для вычисления степени связи между компонентами коллокации.

Ключевые слова: коллокация, нод, коллокат, корпус текстов, статистические меры ассоциации.

Lototska N. Ya. Collocation research in modern linguistics: corpus and statistical approaches.

Summary

The article deals with linguistic phenomena of collocation denoting the process of combining words into phrases governed by a complex system of rules in the language.

The main purpose of the research is to study and analyze the concept of collocation in modern linguistics using corpus and statistical approaches.

The criteria of interpreting the concept of collocation in modern linguistics are described.

Many modern authors and most corpus linguists understand collocations as statistically determined set phrases. The study of the connectivity of lexical units should base on the corpus texts, because regularly generated syntagmatic structures play a significant role in the text construction. In corpus linguistics collocation is the word combination used in the text together more often, than used separately at random probability.

The importance of statistical methods in studying the compatibility of words is considered. Statistical methods provide reliable quantitative data on the compatibility of lexical units based on corpus texts, to study lexical units in context, to get the results on frequency of lexical forms, lexemes, grammatical categories, to investigate the co-occurrence of lexical units, the peculiarities of their compatibility, etc.

The most common statistical association measures to calculate the degree of syntagmatic closeness between the components of collocation, based on probability theory and mathematical statistics, are examined.

Key words: *collocation, node, collocates, corpus texts, statistical association measures.*

УДК 81'271: (316.647.8-055.2):(051 «Жінка»)

Мамич М. В.,

доктор філологічних наук, доцент,

завідувач кафедри прикладної лінгвістики

Національного університету «Одеська юридична академія»

E-mail: miroslavamiros@ukr.net

**ГЕНДЕРНІ СТЕРЕОТИПИ У ДОЗВІЛЬНО-РОЗВАЖАЛЬНИХ МЕДІАТЕКСТАХ НА
СТОРІНКАХ ЖІНОЧОГО ЖУРНАЛУ «ЖІНКА»**

У статті проаналізовано текстовий матеріал дозвільно-розважальних медіажанрів – ‘кросворд’, ‘гра в шахи’, ‘конкурс’, ‘анкета’, ‘тест’, ‘астрологічний прогноз’ як джерел різноманітної соціокультурної інформації, зокрема країнознавчої, соціально-психологічної, соціологічної, культурологічної. Систематизовано мовні знаки, що репрезентують у цих медіатекстах образ жінки і пов'язані з ним основні гендерні стереотипи.

Ключові слова: *дозвільно-розважальний медіатекст, гендерний стереотип, гедоністична функція, лінгвокультурологічний індикатор цінностей.*

Вивчення дозвільно-розважальної функції медіатексту нині має посідати одне з центральних місць, оскільки такі тексти міцно утвердилися в соціальних комунікаціях. Т.В. Радзівєвська підкреслює, що розважальність – важлива стратегічна лінія у текстотворенні [5, с. 45]. Більшість газетно-журнальних видань гендерної тематики у своєму контенті мають дозвільно-розважальний складник, вивчення якого зумовлене потребою пізнання специфіки контенту конкретного ЗМК, його лінгвоформатів, мовно-стилістичних особливостей, що і є головною метою даного дослідження.

Жіночий журнал «Жінка» як визнане якісне інформаційно-аналітичне та художньо-публіцистичне видання завжди мало і має нині дозвільно-розважальний складник з характерними для нього мікрожанрами, більшість з яких ще не була об'єктом медіастилістики взагалі або лише почала входити в поле зору дослідників медій саме із сучасного погляду. Вага дозвільно-розважальних медіатекстів (далі – ДРМТ) визначена спрямуванням мас-медіа знімати психоемоційну, інтелектуальну, соціальну напругу, допомагати читачеві на певний період створити ілюзію власного простору, індивідуалізуватися, самоактуалізуватися у просторі, запропонувати певний аспект вирішення проблем, розважити аудиторію, дати їй відчуття свободи від професійних, соціальних обов'язків в адекватному для типу мас-медіа лінгвоформаті.

Розваги й форми дозвілля в медіа не обмежуються лише ігровими формами. Текстові матеріали, які виконують дозвільно-розважальну функцію і розповсюджуються в ЗМІ, підсилюють пізнавальну мотивацію читачів, розширюють здобуті знання. ДРМТ – це концептуалізовані