

и средства словообразования онимной и апеллятивной лексики, которые также являются показателями их национально-культурной семантики в двух языках. Имена, которые мотивированы названиями одежды и ее элементов, образовано средствами семантической деривации в английском языке и лексико-семантическим способом в сочетании с аффиксацией – в украинском. Было установлено, что для украинского языка характерно образование фамилий путем добавления прилагательного (колороним) и существительного, которым является наименование одежды, в то время как в английском языке таких названий не было зафиксировано. В справочных источниках украинского языка находим также гидронимы, происхождение которых так или иначе связано с названиями одежды (названия рек и ручьев), которые образованы путем прямого метафорического переноса названия. В то же время в английском языке среди гидронимов таких названий не наблюдалось. Апеллятивную лексику (общие названия), происхождение которой связано с названиями предметов одежды и ее элементов, образовано преимущественно средствами семантической деривации (метафора) в двух исследуемых языках. В результате было обнаружено, что не все названия одежды имеют одинаковый культурный потенциал обозначать предметы других лексических сфер в двух сопоставляемых языках.

Ключевые слова: номинации одежды, внешняя мотивация, онимная и апеллятивная лексика, антропонимы, топонимы, гидронимы.

УДК 811. 111

Оксана Скобнікова  
(Київ)

### АНАЛІЗ КОНЦЕПТУ FAMILY НА БАЗІ АМЕРИКАНСЬКИХ НАЦІОНАЛЬНИХ КОРПУСІВ

Об'єктом дослідження, представленого в статті, є концепт FAMILY, один з найбільш значущих концептів американської культури. Під концептом ми розуміємо сукупність знань, інформації про певну матеріальну або ідеальну субстанцію; це виступає в якості основної одиниці наївної картини світу. Метою нашого дослідження є аналіз частотності лексичної одиниці, яка репрезентує концепт FAMILY, представлений у текстах. Матеріалом для нашого дослідження було обрано статичний корпус американського варіанту англійської мови Brown University Standard Corpus of Present-Day American English, або Brown Corpus (Браунівський корпус), що містить близько 1 мільйона слів та динамічний корпус Corpus of Contemporary American English (Корпус сучасної американської англійської мови), що містить більше 560 мільйонів слів. У нашому дослідженні ми використовували корпусний менеджер *AntConc*, розроблений Лоуренсом Ентоні. Ця програма є однією із найзручніших безкоштовних комп'ютерних програм для роботи з корпусом, що може використовуватися у навчальних і в дослідницьких цілях, оскільки має зручний і зрозумілий інтерфейс та пропонує опції потужного конкордансера, генератора частотного списку, аналізатора сполучуваності, а також візуалізує входження досліджуваних слів у досліджуваному масиві. Програма надає можливість пошуку за окремими словами, словосполученнями, з урахуванням чи без урахування регістру клавіатури, а також можливість розширення пошуку за допомогою регулярних виразів. Загальний кількісний аналіз частотності використання лексеми *family*, яка репрезентує концепт FAMILY, показав, що вона широко представлена в англійській мові. Відсоток частотності вживання лексеми *family* відповідно до загальної кількості токенів становить 0.035% (345) у ВС та 0.048% (272121) у СОСА. В обох випадках лексема *family* найчастіше зустрічається в текстах, отриманих з журналів та газет. Значне кількісне вираження в мові говорить про

те, що концепт *FAMILY* є дуже важливим для американської ментальності, оскільки він широко представлений в досліджуваних корпусах.

Ключові слова: корпусна лінгвістика, корпусний аналіз, корпус, колокат, концепт, ім'я концепту.

Будь-який концепт потрібно інтерпретувати з урахуванням контексту, тобто ситуації, в якій він реалізується. Одним з найважливіших факторів, який слід взяти до уваги при аналізі концепту, є контекстуальне поле та частотність, тобто кількість зареєстрованих в текстах випадків вживань лексичної одиниці, що номінує концепт. Це ми можемо зробити за допомогою методів корпусної лінгвістики, що представляє собою нову стратегію вивчення мови за допомогою обробки текстів природної мови.

Нині спостерігається стрімкий розвиток корпусних студій у всьому світі, що підтверджується як величезною кількістю досліджень у галузі, так і різноманіттям корпусів, що постійно з'являються. Корпусна лінгвістика привертає увагу дослідників багатьох країн світу (В. В. Жуковська, В. А. Широков, В. В. Риков, Є. А. Карпіловська, В. П. Захаров, С. Ю. Богданова, Н. Б. Гвішиані, О. М. Карпова, Д. Бібер, Т. МакЕнері, А. Уілсон, С. Йохансон, К. Майер, В. Крофт, С. Адольфс).

Термін «корпусна лінгвістика» міцно ввійшов до наукового вжитку лише в останні десятиліття XX століття з публікацією у 1983 році збірника наукових праць «Corpus Linguistics: Recent Developments in the Use of Computer Corpora in English Language Research» [2, с. 8]. Корпусна лінгвістика як напрямок прикладної лінгвістики визначає загальні принципи побудови лінгвістичних корпусів даних із використанням сучасних комп'ютерних технологій, окреслює методику відбору письмових та усних текстів та шляхи їх збереження та аналізу [6, с. 4].

У наш час існує значна кількість визначень поняття «корпус», запропонованих вітчизняними та зарубіжними дослідниками. Наприклад, у визначенні, наведеному в підручнику Е. Фінегана, зазначено, що корпус – це репрезентативне зібрання текстів, що зазвичай представлене в електронному форматі і містить інформацію про ситуацію, в якій текст був створений, тобто інформацію про мовця, автора, адресата або аудиторію [7]. Т. МакЕнері та Е. Вілсон наводять таке визначення корпусу – це зібрання мовних фрагментів, відібраних за чіткими мовними критеріями для використання в якості моделі мови [8]. В. В. Риков визначає корпус текстів як певне зібрання текстів, в основі якого лежить логічний задум, логічна ідея, яка об'єднує ці тексти і втілена в правилах організації текстів в корпус, алгоритмі і програмі аналізу корпусу текстів, пов'язаною з цим ідеологією і методологією [4]. В. П. Захаров під корпусом текстів розуміє значний за обсягом, представлений в електронному вигляді, уніфікований, структурований, розмічений, філологічно компетентний масив мовних даних, створений для вирішення конкретних лінгвістичних завдань [3, с. 3]. О. М. Демська-Кульчицька визначає корпус як машиночитане, стандартно організоване зібрання репрезентативних для певної мови, діалекту або іншої підмножин(и) мов(и) писемних або усних текстів, призначених для лінгвістичного аналізу й опису, відібраних і впорядкованих згідно з експліцитними екстра- та інтралінгвальними критеріями [1].

Корпуси надають лінгвістам багатий репрезентативний емпіричний матеріал. Корпуси можуть у принципі дати три типи даних, які можуть бути використані в ході лінгвістичних досліджень: емпірична підтримка, інформація щодо частотності, екстралінгвістична інформація, або метайнформація (вік або стать мовця, жанр тексту, часова або просторова інформація щодо походження тексту тощо), яка дозволяє порівнювати різні типи текстів чи різні групи мовців.

Саме тому в нашому дослідженні доцільним буде залучення даних із корпусу американського варіанту англійської мови, який являє собою зібрання письмових та усних текстів різних жанрів, адже об'єктивація досліджуваного концепту буде яскраво виражена в корпусах текстів за допомогою різних лексичних засобів. Корпусний аналіз дозволить

визначити частотність вживання слів і словосполучень, за допомогою яких вербалізується концепт. Відповідно до частотності вживання словосполучень, їх наявності або відсутності, можна визначити найбільш характерні слова, сталі вирази і фразеологізми, що об'єктивують досліджуваний концепт.

Робота з корпусом здійснюється за допомогою спеціальних програмних засобів – конкордансерів і корпусних менеджерів, що надають різноманітні можливості по одержанню необхідної інформації з корпусу. Корпусний менеджер – це спеціальна пошукова система, що включає програмні засоби для пошуку даних у корпусі, отримання статистичної інформації й надання результатів користувачеві в зручній формі [2]. У нашому дослідженні ми використовували корпусний менеджер AntConc, розроблений Лоуренсом Ентоні. Ця програма є однією із найзручніших безкоштовних комп'ютерних програм для роботи з корпусом, що може використовуватися у навчальних і в дослідницьких цілях, оскільки має зручний і зрозумілий інтерфейс та пропонує опції потужного конкордансера, генератора частотного списку, аналізатора сполучуваності, а також візуалізує входження досліджуваних слів у досліджуваному масиві. Програма надає можливість пошуку за окремими словами, словосполученнями, з урахуванням чи без урахування реєстру клавіатури, а також можливість розширення пошуку за допомогою регулярних виразів.

Матеріалом для нашого дослідження було обрано один з найбільш авторитетних та репрезентативних корпусів американського варіанту англійської мови Brown University Standard Corpus of Present-Day American English, або Brown Corpus – «Браунівський корпус», далі BC. Він містить близько 1 мільйона слів, а саме 500 фрагментів текстів по 2 тисячі слів в кожному, які були опубліковані англійською мовою в США в 1961 році. Тексти були відібрані із п'ятнадцяти найбільш масових жанрів англійської друкованої прози США, що поділялись на дві категорії – науково-публіцистична (374 тексти) та художня (126 тексти).

Корпус BC є статичним, тобто тексти в ньому не оновлюються. Варто зазначити, що спочатку всі корпуси текстів створювалися як статичні утворення, тобто такі, що відображають певний часовий зріз мовної системи. Однак значна частина чисто лінгвістичних і не тільки лінгвістичних завдань вимагає можливості виявлення історичних змін у функціонуванні мовних явищ, наприклад, змін значення слів, частотності використання тих чи інших синтаксичних конструкцій тощо. Для цього була розроблена спеціальна технологія побудови та експлуатації так званого динамічного корпусу текстів. Особливість збірки динамічних корпусів полягає в тому, що вони не містять раз і назавжди заданий набір текстів. Протягом заздалегідь фіксованого проміжку часу відбувається оновлення і доповнення текстів корпусу.

До числа динамічних відноситься Corpus of Contemporary American English, Корпус сучасної американської англійської мови, далі COCA, що було нами обрано задля співставлення даних. У ньому міститься понад 560 мільйонів слів (по 20 мільйонів слів у кожному році, починаючи з 1990 і закінчуючи 2017рр., оновлення корпусу відбувається один раз або двічі на рік). У корпусі відображаються не тільки письмові джерела (художня, публіцистична, наукова проза), а й записи радіо і телепередач, тобто джерела живої мови і живого мовлення.

Проведений нами загальний кількісний аналіз частотності вживання лексеми *family*, яку ми визначили іменем концепту FAMILY, дозволяє судити про те, що концепт FAMILY досить широко представлений в англійській мові, так як вербалізуються, як правило, найбільш значущі для людини поняття. Значне кількісне вираження в мові дає підстави вважати, що концепт FAMILY дуже важливий для американської ментальності. Лексема *family* зустрічається 345 і 272121 рази в корпусах BC і COCA відповідно:

*The bushes can be cleaned and sterilized by boiling and are detachable so that every member of the **family** can have his own.* (BC)

*Specifically, we examined how growing up in a two-parent **family** with one or more alcoholic parents affects womens' coping methods.* (COCA)

До того ж, корпуси дають змогу структурувати отримані дані за жанрами текстів, представлених у корпусах:

Spoken: *And it wasn't long before the Sommer **family** got into financial trouble.* (COCA)

Fiction: *His name was George Needham and he, too, had come from a good **family**.* (BC)

Magazine/general prose: *Every member of the **family** must have a vital place in its life.* (BC)

Newspaper/press: *Furhmann 's faculty is proud that this has been a spontaneous effort, started largely among the students themselves , because of fondness for Vicky and sympathy for her entire **family** , Pohly said .* (BC)

Academic/learned: *It is important to consider cases where catastrophic illnesses or other unexpected financial calamities that have impacted a **family** or individual to the point where their debts are too heavy to carry a loan.* (COCA)

В обох корпусах лексема *family* найчастіше зустрічається в текстах, отриманих з журналів та газет. На думку спеціалістів з корпусної лінгвістики, це зумовлено специфікою періодики публіцистичного характеру охоплювати художню і наукову літературу, а також тим, що саме в текстах цього стилю найбільш широко представлена нова лексика загального вживання [5].

Корпуси дозволяють досліджувати значення слів шляхом використання конкордансів. Почати дослідження значень слова можна з аналізу його колокатів – слів, з якими аналізоване слово зустрічається разом. Виділяючи найбільш частотні колокації слова, можна ефективно і надійно аналізувати коло значень слова, виділяючи їх ієрархічний взаємозв'язок.

У корпусі COCA найчастішими лівими колокатами іменника *family* є артиклі (*the, a*) та займенники (*his, my, her, your, our*), також досить часто вживається сполучник *and*, прийменник *of* та іменники в присвійному відмінку. Найчастішими правими колокатами є сполучник *and*, іменник *members*, прийменники (*of, in, to*) дієслова (*is, was, has, had*), також досить часто іменник *family* вживається в присвійному відмінку. Наприклад:

*It's my contention that just about every **family** has at least one psychopath, if not two.*

*Jon has another **family member** who is about to have a major impact on his life: Daenerys.*

У корпусі BC спостерігаємо схожі результати – найуживанішими лівими колокатами також виявились артиклі, займенники, сполучник *and*, прийменники *of, for* та іменники в присвійному відмінку. На відміну від COCA, у першу десятку потрапив прикметник *entire*. Щодо правих колокатів, картина також подібна до даних корпусу COCA. Тут також найчастіше зустрічаються сполучник *and*, прийменники, дієслова, іменник *family* також часто вживається в присвійному відмінку. До того ж, до списку іменників потрапили іменники *business, life* та *planning*. Наприклад:

*We'll do everything we can to support Mike, his wife Kia and the **entire family** over this tragedy.*

***Family planning** is encouraged, so that parents will be able to provide properly for their offspring.*

Таким чином, спираючись на числові дані корпусного аналізу і їх інтерпретацію, можна з достатньою точністю говорити про численну та багатогранну вербалізацію концепту FAMILY в американському варіанті англійської мови. Виходячи з кількісних і якісних характеристик проаналізованих засобів вираження концепту, можна зазначити, що величезна кількість репрезентацій на лексичному рівні говорить про значущість даного концепту в свідомості носіїв американського варіанту англійської мови.

Наступним кроком у нашому дослідженні ми плануємо розглянути синтаксичні зв'язки та сполучуваність слова *family* з прикметниками та іменниками, що його описують, тобто щодо кластерів *adjective + noun* та *noun + noun*, які найяскравіше ілюструють значення досліджуваної лексеми.

## ЛІТЕРАТУРА

1. Демська-Кульчицька О. М. Базові поняття корпусної лінгвістики / О. М. Демська-Кульчицька // Українська мова. – 2003 – №1. – С. 42– 47.
2. Жуковська В. В. Вступ до корпусної лінгвістики: навчальний посібник / В. В. Жуковська – Житомир: Вид-во ЖДУ ім. І. Франка, 2013. – 142 с.

3. Захаров В. П. Корпусная лингвистика: учебник для студентов гуманитарных вузов/ В. П. Захаров, С. Ю. Богданова. – Иркутск : ИГЛУ, 2011. – 161 с.
4. Рыков В. В. Корпус текстов как реализация объектно-ориентированной парадигмы // Труды Международного семинара Диалог-2002. – М.: Наука, 2002. – С. 59– 61.
5. Шимкова М. Репрезентативность корпуса как лингвистическая проблема // MegaLing 2005: Прикладная лингвистика в поиске новых путей. Материалы международной конференции. С.-Пб.: Осипов, 2005. – С. 130 – 139.
6. Biber D. Corpus Linguistics: Investigating language structure and use/ D. Biber, S. Conrad, R. Reppen. – Cambridge University Press, 2004.
7. Finegan E. LANGUAGE: its structure and use. – N.Y.: Harcourt Brace College Publishers, 2004.
8. McEnery T., Wilson A. Corpus Linguistics. – Edinburgh: Edinburgh University Press, 2001.

Oksana Skobnikova

#### ANALYSIS OF THE CONCEPT *FAMILY* ON THE BASIS OF AMERICAN NATIONAL CORPORA

The object of our study is the concept *FAMILY*, one of the most meaningful concepts in American culture. A concept is understood as a set of knowledge, information about a certain material or ideal substance; it acts as the main unit of the naive picture of the world. The purpose of our study is to analyze the frequency of the lexical unit that nominates the concept *FAMILY* registered in the texts. As the material for our study, we chose the Brown University Standard Corpus of Present-Day American English, or Brown Corpus (BC), containing about 1 million words and the Corpus of Contemporary American English (COCA) containing more than 560 million words, which are static and dynamic correspondingly. In our study we used the AntConc corpora manager, developed by Lawrence Anthony. This program is one of the most convenient freeware computer applications, which can be used for educational and research purposes, since it has a convenient and understandable interface and offers options for a powerful concorder, frequency generator, compatibility analyzer, and also visualizes the entry of the subjects words in the studied array. The program provides the ability to search for individual words, phrases, with or without the case of the keyboard, as well as the ability to expand the search using regular expressions. The general quantitative analysis of the frequency of use of the lexeme family, which nominates the concept *FAMILY*, has shown that it is widely represented in the English language. The percentage of the lexeme family to the general number of tokens is 0.035% (345) in BC and 0.048% (272121) in COCA. In both cases, the lexeme family is most often found in texts derived from magazines and newspapers. Moreover, the concept name *FAMILY* has a tendency for steady and strong frequency increase in diachronic perspective that allows us to conclude about the growing importance of a family as a social unit. Consequently, a significant quantitative expression in the language suggests that the concept *FAMILY* is very important for American mentality, since it is widely verbalized in the corpora.

Key words: corpus linguistics, corpus analysis, corpus, collocate, concept, name of the concept.

Оксана Скобникова

#### АНАЛИЗ КОНЦЕПТА *FAMILY* НА БАЗЕ АМЕРИКАНСКИХ НАЦИОНАЛЬНЫХ КОРПУСОВ

Объектом исследования, представленного в статье, является концепт *FAMILY*, один из самых значимых концептов американской культуры. Под концептом мы понимаем совокупность знаний, информации об определенной материальной или идеальной субстанции; которая выступает в качестве основной единицы наивной картины мира. Целью нашего исследования является анализ частотности лексической единицы, которая представляет концепт *FAMILY*, представленный в текстах. Материалом для нашего исследования было избрано статический корпус американского варианта английского языка Brown University Standard Corpus of Present-Day American English, или Brown Corpus (Брауновский корпус), содержащий около 1 миллиона слов и динамичный корпус Corpus of

Contemporary American English (Корпус современного американского английского языка), содержащий более 560 миллионов слов. В нашем исследовании мы использовали корпусный менеджер AntConc. Эта программа является одной из самых удобных бесплатных компьютерных программ для работы с корпусом, который может использоваться в учебных и в исследовательских целях, так как имеет удобный и понятный интерфейс и предлагает опции мощного конкордансера, генератора частотного списка, анализатора сочетаемости, а также визуализирует репрезентацию исследуемых слов в исследуемом массиве. Программа предоставляет возможность поиска по отдельным словам, словосочетаниям, с учетом или без учета регистра клавиатуры, а также возможность расширения поиска с помощью регулярных выражений. Общий количественный анализ частотности использования лексемы *family*, которая представляет концепт FAMILY, показал, что она широко представлена в английском языке. Процент частотности употребления лексемы *family* в соответствии с общим количеством токенов составляет 0.035% (345) в ВС и 0.048% (272 121) в СОСА. В обоих случаях лексема *family* чаще всего встречается в текстах, полученных из журналов и газет. Значительное количественное выражение в языке говорит о том, что концепт FAMILY является очень важным для американской ментальности, поскольку он широко представлен в исследуемых корпусах.

Ключевые слова: корпусная лингвистика, корпусный анализ, корпус, коллокат, концепт, имя концепта.