

комплексно, що безумовно привертає увагу лінгвістів і дослідників суміжних дисциплін, а саме: соціологів, політологів, журналістів, спеціалістів у галузі PR і т. п.

Упродовж останніх років особливий інтерес для лінгвістів становить питання вивчення візуального компонента фуд-блогів, головно фотографій, оскільки саме візуальний контент є головним інструментом для привертання уваги фоловерів до будь-якого блога й передавання емоційного плану повідомлення та його аргументативного потенціалу. Такі візуальні складники блогів на кулінарну тематику, як загальний тон фотографії, використання пастельної, темної чи різнокольорової гама на зображеннях, удадо підібране спрямоване світло та застосування різноманітного кулінарного реквізиту в поєднанні із вербальним повідомленням створюють цілісну картину сприйняття рецептури певної страви, доповнюючи та поглиблюючи зміст один одного. Саме завдяки фуд-фотографіям високого рівня майстерності читачі фуд-блогів можуть побачити не лише відображення об'єктивної дійсності, але й досягнути особливості світосприйняття автора цих фотографій.

#### Список використаних джерел:

1. Белова А. Д. Реклама автомобилей Mercedes-Benz: принципы и составляющие / А. Д. Белова, О. А. Васильева // Лінгвістика XXI століття: нові дослідження і перспективи. – К.: Логос, 2008. – С. 37–45.
2. Білюк І. Л. Вербальні і невербальні засоби брендингу Олімпіади ХХХ / І. Л. Білюк // Лінгвістика ХХІ століття: нові дослідження і перспективи. – К.: Логос. – 2012. – С. 60–67.

K. Karpova, Phd., assistant  
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

### MULTIMEDIA OF ENGLISH CULINARY BLOG

*The article focuses on the study of visual aspect in present-day English culinary blogs. In particular the research deals with nonverbal peculiarities of food photography namely general tone of the photograph, color terms, directional artificial light and culinary props as means of multimedia content in computer-mediated communication. Moreover the article studies patterns of effective coordination of verbal and nonverbal (visual) means of information transmission in modern English food blogs.*

*Key words: multimedia, visualization, blogosphere, culinary blog, food photographer, food photography.*

Е. Карпова, канд. філол. наук, асист.  
Київський національний університет імені Тараса Шевченка, Київ, Україна

### МУЛЬТИМЕДИА АНГЛОЯЗЫЧНОГО КУЛИНАРНОГО БЛОГА

*В статье исследуется визуальный аспект подачи информации в современных англоязычных кулинарных блогах. В частности рассматриваются невербальные особенности (общий тон фотографии, цветообозначение, использование направленного искусственного света и кулинарного реквизита) фуд-фотографии как единицы мультимедийного контента сетевой коммуникации. Кроме того, проанализированы примеры эффективного слияния вербальных и невербальных (визуальных) средств трансляции информации в современных англоязычных фуд-блогах.*

*Ключевые слова: мультимедиа, визуализация, блогосфера, кулинарный блог, фуд-фотограф, фуд-фотография.*

УДК 81'322.2

Л. Клименко, канд. філол. наук, наук. співроб.  
Київський національний університет імені Тараса Шевченка, Київ

### КОРПУСНА І КОГНІТИВНА ЛІНГВІСТИКА: ПЕРСПЕКТИВИ КОНВЕРГЕНЦІЇ МЕТОДІВ

*Здійснено стислий огляд анованих корпусів французької мови із зазначенням особливостей їх функціоналів. Висвітлено сучасний стан французької корпусної лінгвістики, а також напрями розробок, здійснюваних на масивах франкомовних текстів. Окреслено перспективи застосування напрацьовань комп'ютерної лінгвістики в дискурсивно-когнітивних дослідженнях, можливості розробки інтегрованої міждисциплінарної методики.*

*Ключові слова: корпусна лінгвістика, корпус, когнітивна лінгвістика, французька мова, метод.*

Бурхливий розвиток науки і техніки останніх десятиліть відкриває для дослідників нові можливості, а сучасні інформаційні технології сприяють швидкій, об'єктивній і неупередженій обробці великих масивів даних, надаючи можливість дослідження мовних явищ не тільки із застосуванням традиційного мовного матеріалу (словників, художньої літератури, наукових текстів тощо), але і шляхом використання в ролі матеріалу значних за обсягом вибірок текстів, уведених у комп'ютер. У лінгвістичних дослідженнях дедалі ширше застосовують електронні

3. Курбасова Е. Б. Роль фотографии в становлении диалога культур [Електронний ресурс] / Е. Курбасова. – Режим доступу: <http://iculture.spb.ru/index.php/stucult/article/viewFile/397/395>.

4. Савчук В. В. Философия фотографии / Савчук Валерий Владимирович. – СПб.: Издательство СПб. гос. ун-та, 2005. – 256 с.

5. Симакова С. И. О визуализации журналистских материалов [Електронний ресурс] / С. Симакова. – Режим доступу: <http://elar.urfu.ru/bitstream/10995/31093/1/journ-2015-prep-37.pdf>.

6. Сохань И. В. Визуальные репрезентации гастрономической культуры / И. В. Сохань // Праксема. Проблемы визуальной семиотики. – Томск: Изд-во Томск. гос. ун-та, 2014. – № 1(1). – С. 88–96.

7. Малішевський Р. О. Food-дизайн – новий напрям у дизайні ХХІ століття / Р. О. Малішевський, А. С. Свириденко, Л. М. Страшко [Електронний ресурс]. – Режим доступу: [http://dSPACE.puet.edu.ua/bitstream/123456789/5629/1/збірник\\_тез\\_любченко.pdf#page=208](http://dSPACE.puet.edu.ua/bitstream/123456789/5629/1/збірник_тез_любченко.pdf#page=208).

8. Gluten Free Chocolate Pancakes From The UnDiet Cookbook. – Online. Retrieved from: <https://yumuniverse.com/gluten-free-chocolate-pancakes-undiet-cookbook>.

9. Red Cabbage Tahini Salad Bowls. – Online. Retrieved from: <http://www.foodheavenmadeeasy.com/red-cabbage-tahini-salad-bowls>.

10. Late Summer Abundance Bowl. – Online. Retrieved from: <https://www.mynewroots.org/site/2014/09/late-summer-abundance-bowl>.

11. Mushroom Confit By Andrew Janjigian. – Online. Retrieved from: <http://kokblog.johannak.com/category/mushrooms>.

12. Senyei K. Food Blogging For Dummies / Kelly Senyei. – New York: Published by John Wiley and Sons Inc., 2012. – 320 p.

13. The Best Homemade Chocolate (Dairy-Free). – Online. Retrieved from: <https://yumuniverse.com/best-homemade-chocolate-dairy-free>.

14. Young Nicole S. Food Photography: From Snapshots to Great Shots / Nicole Young. – Pearson Education. Peachpit Press, 2012. – 272 p.

15. <http://www.pantryconfidential.com/features/johanna-kindvall-illustrator>.

16. <https://yumuniverse.com/>

Надійшла до редколегії 30.10.17

- використання великих за обсягом, структурованих колекцій природних текстів (корпусів) як основи для аналізу;

- широке залучення комп'ютерних технологій для дослідження мовного матеріалу;

- застосування квалітативних і квантитативних аналітичних методик, із суттєвою перевагою останніх (вивчення частоти вживання лінгвістичних одиниць, статистичні дослідження сполучуваності тощо) [1–3].

Традиційно корпус визначають як масив анотованих і структурованих письмових або усних текстів в електронному вигляді, що створюється для розв'язання прикладних лінгвістичних завдань. Корпуси мають низку типових технічних характеристик, серед яких:

- стандартизоване представлення даних (sampling), що дозволяє застосування комп'ютерних методів обробки;

- репрезентативність (representativeness) – перелік параметрів, відповідно до яких корпус було укладено;

- визначений обсяг текстів (finite-tize)

- розміщення ресурсів на електронному носії (machine-readable forms) [5, p. 114–115].

Корпусній лінгвістиці властива певна двовекторність, оскільки вона вивчає як теорію та практику створення корпусів, так і самі мовні корпуси, спираючись на дискурс, який становить її матеріал.

Саме спільна природа матеріалу досліджень поєднує корпусну і когнітивну лінгвістику. Обидві напрями, вивчаючи мову в дії, знизу – вгору, роблять висновки на основі дискурсу, так званого живого мовлення [1]. Обидві лінгвістичні парадигми займаються пошуками еталонової (узусу) і відхиленя від нього. До кола як когнітивних, так і корпусних досліджень входить вивчення дистрибутивних особливостей значення слова (напр., для концептуального аналізу або встановлення сітки значень полісемантів).

Брак формалізованих методів тестування й апробації результатів концептуального аналізу у когнітивній лінгвістиці об'єктивно зумовлює і диктує необхідність залучення корпусних методів до когнітивної лінгвістики.

**Метою** статті є стислий огляд анотованих корпусів французької мови із зазначенням особливостей їхніх функціоналів, висвітлення сучасного стану розробок у галузі корпусної лінгвістики на масивах франкомовних текстів, визначення векторів її розвитку, окреслення **перспектив** застосування напрацювань у комп'ютерній лінгвістиці до дискурсивно-когнітивних досліджень, можливості розробки інтегрованої міждисциплінарної методології.

Здійснено огляд корпусних розробок на матеріалі французької мови, що наразі існують, і з'ясуємо їхні функціональні можливості. Оскільки французька лінгвістика зі значним запізненням звернулася до укладання корпусів французької мови [4], то застосування методів корпусної лінгвістики і корпусного аналізу, і, відповідно, стан розробленості галузі, значно відстає порівняно з аналогічними у Великій Британії та США – флагманах методики, а також із корпусними дослідженнями на базі корпусів інших європейських мов. Так, нині не існує жодної французької граматики, заснованої на даних корпусного аналізу, немає граматики ефективного вжитку французької мови, тоді як англійською на основі корпусних даних уже понад 15 років тому було укладено і широко вживається відома *The Grammar of Spoken English*. На базі корпусів англійської створено перелік зі 100 тис. частотних слів, перелік 500 найчастотніших фразеологізмів тощо. Жодний сучасний англійський тлумачний словник більше не укладають без застосування Британського національного корпусу (BNC).

Аналогічною є ситуація із застосуванням методів корпусної лінгвістики у французькій лексикології, де майже відсутні корпусні дослідження особливостей жанрової диференціації текстів. Відтак звернення до проблем розвитку сучасної французької корпусної лінгвістики вбачається наразі надзвичайно актуальним. Більш того, оскільки одним з основних викликів XXI ст. є тенденція до інтеграції різних сфер знань і дисциплін, то розробка новітніх наукових комплексних методик, можливість застосування алгоритмів прикладних досліджень в інших сферах може відкривати нові наукові горизонти, зокрема в царині когнітивістики.

Серед **письмових** корпусів **нині** найбільшим є *Fran-text*, який умовно можна назвати національним корпусом французької мови. Заснований та підтримуваний французькою організацією l'ATILF-CNRS із штаб-квартирою в Ансі, корпус об'єднує переважно літературні та філософські тексти (90 %), а також незначну кількість наукових і технічних (10 %), а його обсяг становить близько 220 млн слів. Укладання корпусу було розпочато у 70-х рр. минулого століття з метою наповнення 16-томного словника *le Trésor de la Langue Française* текстовими прикладами (обсяг становив 1000 зразків текстів). WEB-базована версія корпусу побачила світ у 1998 р. Наразі корпус формують 5 116 текстів (станом на грудень 2016), з них 1026 текстів становлять його "сучасну" частину і датовані часовим проміжком після 1950 р. Є корпуси скромнішого обсягу, укладені для специфічних цілей, такі як *Varilex*, *Manulex*, *Emolex*. Використання цих корпусів як матеріалу, вочевидь, не відкриває особливих перспектив для аналізу сучасних видів дискурсів, усе ж уможливорює розвідки, присвячені певній часовій добі, або ж діахронічному аналізу.

У галузі публіцистики і художньої літератури є досить цікавий корпус *Sketchengine FrTenTen*, укладений на основі текстів із WEB, який містить 12 млрд слововживань. Корпус дивує своїм розміром, але не може похизуватися жанровою варіативністю представлених текстів [6]. Особливий інструментарій, застосований при укладанні цього ресурсу, забезпечує диференційоване завантаження вибірки текстів з усіх можливих ресурсів і доменів, які відповідають зазначеній мові, слідкуючи за тим, щоб фрагменти нових текстів не дублювали ті, що вже увійшли до корпусу (використовується спеціальна пошукова програма-робот (англ. crawler).

Інші розробки письмових корпусів, як не дивно, належать не французам, а німецьким фахівцям. Серед них:

*Leipziger Corpus Francais* [8] – створений дослідниками лейпцизького університету, нараховує 7 млн слів і містить переважно публіцистичні і тексти з WEB. Недоліком цього корпусу є те, що він дозволяє здійснювати лише дуже прості пошукові операції. Так, його можливості обмежуються отриманням переліку прикладів на слово-запит розміром в одне речення із посиланням на джерело, але без дати публікації, а також переліку правосторонніх та лівосторонніх колокатів, що вживаються з цим словом.

*Kölnner romanistische korpus datenbank / Varietext* – дві платформи текстового аналізу, диференційовані за жанрами, що відкривають ширший дослідницький діапазон і дозволяють створювати конкорданси, робити підрахунки частотної специфіки, аналіз лексико-синтаксичних входжень.

Загалом у письмових корпусах французької мови переважають тексти літературних і публіцистичних творів, а їхні загальні функціональні можливості значно поступаються корпусам англійської мови.

Дещо кращою є ситуація із французькими **усними корпусами**, серед яких можна згадати:

- *Traitement de corpus oraux en français* (TCOF): 70 тис. слів;
- *L'Enquête Socio-Linguistique à Orléans* (ESLO 1 et 2): 6,5 млн слів;
- *Corpus de Langue Parlée en Interaction* (CLAPI): 100 тис. слів;
- *Communication langagière chez le jeune enfant* (CoLaJe): 3 млн слів.

Існують також інші корпуси, доступ до яких повністю обмежений:

- *Corp-Aix-2* (Cresti&Moneglia 2005): 300 тис. слів;
- *Ciel-F*.

Загальний обсяг усієї сукупності усних корпусів становить близько 10 млн слововживань. Їхньою беззаперечною перевагою, порівняно з письмовими, є можливість синтаксичного і фонологічного аналізу, проте функціонал не дозволяє проводити лексико-граматичні статистичні дослідження на предмет повторюваності сегментів, колокацій, валентностей. Зведено одночасне використання кількох корпусів також, найчастіше, неможливе з причини відсутності координованості серед науковців-розробників.

У 2016 р. було анонсовано розробку корпусу сучасної французької мови з робочою назвою *Le corpus de référence du français contemporain* (CRFC) [7], над яким наразі співпрацюють науковці із ФРН і Франції (Dirk Siepmann (Universität Osnabrück, Institut für Anglistik/Amerikanistik), Christoph Bürgel (Universität Paderborn, Institut für Romanistik) та Sascha Diwersy (лабораторія Praxiling, UMR 5267, Université Paul-Valéry Montpellier 3). Станом на 2018 р. корпус формують тексти французькою мовою Франції (*le français de France*) 1945–2014 рр. із відсотковою часткою 90 % текстів двох останніх десятиліть. Корпус містить приблизно 310 млн слів (його обсяг було свідомо зведено до такого числа), з них 60–65 млн спонтанного мовлення офіційного та неофіційного характеру. Це перший французький корпус анотований за жанрами (14 жанрів), який відрізнятиметься від інших за рахунок того, що це буде:

- найбільший за обсягом французький корпус, розміщений у мережі Інтернет;

- перший французький корпус, що міститиме вагому частку даних усного мовлення і псевдоусного мовлення;
- корпус, який уперше згрупує колекцію гетерогенних текстів.

Після закінчення робіт над корпусом і його відкриття загалу, планують другий етап із його поповнення текстами, датованими періодом після 2014 р.

Структуру корпусу значного мірою скальовано із двох корпусів англійської мови – British National Corpus (BCN) та Corpus of Contemporary American English, з тією відмінністю, що жанрова варіативність найбільша серед усіх корпусів. Щодо джерел текстового матеріалу: усну частину корпусу формує широка гама мовленнєвих фрагментів офіційного спілкування (30–35 млн слів), а також тексти неформального мовлення – 30–35 млн слів, з яких 75 % – транскрипти телевізійних передач 70-ти різних типів (загальна тривалість 3 тис. год звучання), 25 % – інтерв'ю з газети *le Figaro*, а також дані інших корпусів.

Джерелами письмового корпусу стали тематичні WEB-сторінки, форуми з обговорення, сайти, на яких розміщені скрипти художніх фільмів, тексти із сайтів літературних видавництв, сайти газет, журналів, у тому числі наукового і науково-популярного характеру, корпус SMS-спілкування.

Цікаво, що вперше за історію корпусної лінгвістики при розмітці було здійснено окреме виділення *псевдоусного масиву корпусу*. Віднесення текстів до категорій усного або ж псевдоусного мовлення відбувалося на основі застосування комунікативної шкали за параметрами наближеності vs віддаленості між комунікантами. Так, до категорії псевдоусних було віднесено фільми, театральні вистави, чати, форуми з обговорення як такі, що більше тяжіють до розмовного мовлення за параметром комунікативної наближеності. Такі ж жанри, як листування, еротичні журнали, блоги, електронне листування, навпаки, було за ознакою віддаленості між комунікантами параметризовано як письмові. На рис. 1 наведено структурне розподілення контенту за жанрами і їхні кількісні характеристики

	Секція	Розмір
Усний	Офіційне мовлення	30 млн
	Неофіційне мовлення	30 млн
Псевдоусний	Театральні вистави, сценарії до фільмів	30 млн
	Субтитри до фільмів, телесеріалів	2,5 млн
	SMS/чати	2,5 млн
	Форуми спілкування	60 млн
		155 млн
Письмовий	Наукові і технічні тексти	30 млн
	Інші нелітературні тексти	30 млн
	Романи, оповіді, казки	30 млн
	Газети	45 млн
	Журнали	10 млн
	Журнали і блоги інтимного характеру	5 млн
	Листування (у т. ч. електронне)	1 млн
	Інші тексти	4 млн
	155 млн	

Рис. 1. Структура корпусу CRFC

Корпус CRFC є анотованим, розмітка полягає у приписуванні текстам та їхнім компонентам спеціальних лінгвістичних й екстралінгвістичних тегів. Метадані корпусу розподіляються за такими параметрами: назва, автор, джерело, рік видання, тип (*genre*) (письмовий,

усний, псевдоусний), супер-жанр (*super-genre*), де їх розрізняють 16, серед яких: наукова проза, ненаукова проза, особистісне спілкування (*conversation privée*), публічне спілкування (*conversation publique*), художня література, фільми тощо. Далі розробники

виокремлюють два типи піджанрів. Піджанр 1 означає тип текстів або мовленнєвих пасажів, наприклад, енциклопедії, журнали, інтерв'ю, новини, телефонне спілкування, роман, телесеріал, ток-шоу тощо. У піджанрі II, виокремлення якого важливе особливо для художньої літератури, розрізняють детектив, роман, наукову фантастику тощо.

Подальший тематичний розподілення здійснено за аналогією до BNC. Розрізняють 16 тематичних доменів, таких як мистецтво, бізнес, економіка, політика, державне управління, юриспруденція, інформатика,

точні науки, довілля, охорона здоров'я, медицина, розваги, спорт, дозвілля.

Розмітка і функціонал корпусу дозволяє дослідити як лексико-граматичні особливості структури кожної з виокремлених сфер, так і когнітивно-прагматичні особливості їхньої організації.

Так, на базі корпусу CRFC було виокремлено перелік фразем, найбільш частотних для французької мови (понад 5 тис. вживань). Потому окремо було про-скановано корпусні масиви на предмет встановлення частотності вживань виокремлених фразем у кожному з них, результати якого містить зведена таблиця (рис. 2).

	Liste de l'oral		Liste de l'écrit		Liste complète
1.	un peu	1.	de plus	1.	un peu
2.	comme ça	2.	par exemple	2.	parce que
3.	parce que	3.	ainsi que	3.	par exemple
4.	quelque chose	4.	à partir de	4.	en plus
5.	quand même	5.	en effet	5.	quelque chose
6.	tout le monde	6.	nombre de	6.	alors que
7.	en plus	7.	en plus	7.	beaucoup de
8.	tout de suite	8.	alors que	8.	comme ça
9.	c'est vrai	9.	en place	9.	quand même
10.	non plus	10.	sans doute	10.	en effet
11.	tout à l'heure	11.	lors de	11.	ainsi que
12.	pas mal	12.	par rapport à	12.	peu de
13.	au moins	13.	grâce à	13.	nombre de
14.	pas encore	14.	au moins	14.	tout le monde
15.	pour que	15.	au sein de	15.	au moins
16.	bien que	16.	de plus en plus	16.	même si
17.	eh bien	17.	afin de	17.	plus tard
18.	par exemple	18.	en particulier	18.	non plus
19.	tu sais	19.	faire partie de	19.	de plus en plus
20.	pas du tout	20.	à travers	20.	de même

Рис. 2. Частотність вживання фразем у корпусних масивах

Найчастіше вживаною є фразема *un peu*. Цей результат дозволяє зробити припущення щодо причин частотності вживання цього маркера, яку можна пояснити високими вимогами французької традиції до ввічливості й екологічності міжособистісної комунікації. Висловлення стає менш безпосереднім, менш агресивним, що робить спілкування більш ввічливим. З дискурсивного погляду за допомогою фраземи *un peu* реалізується стратегія ввічливості спілкування, що спрямована на пом'якшення висловлення, зменшення його ілюктивної сили. Можливість робити завдяки функціоналу корпусних платформ висновки такого характеру відкриває перспективи переходу досліджень на основі корпусних даних із лексико-граматичної до когнітивно-комунікативної площини.

Інтерес для когнітивістів може становити функціонал платформи *SketchEngine*, розроблений на основі гіпотези дистрибутивної семантики про те, що лінгвістичні одиниці, які зустрічаються в подібних контекстах, мають близькі значення [9, с. 35]. Функціонал платформи уможлиблює автоматичне створення словесного простору предметної області для кожного окремого слова у складі корпусу завдяки автоматичній ідентифікації слів у контекстах, схожих на ті, де зустрічається цільове слово [1]. Це дозволяє вирахувати ступінь семантичної наближеності між одиницями на основі їхньої дистрибуції у великих масивах лінгвістичних даних. Так, при здійсненні запиту для слова *simulacre*, було отримано відповідь системи, яку зображено на рис. 3.

**simulacre** (noun) French Web 2012 (frTenTen12) freq = [16,464](#) (1.43 per million)

Lemma	Score	Freq
<a href="#">supercherie</a>	0.229	18,760
<a href="#">mascarade</a>	0.226	26,521
<a href="#">subterfuge</a>	0.201	11,368
<a href="#">machination</a>	0.185	13,115
<a href="#">stratagème</a>	0.162	20,834
<a href="#">semblant</a>	0.162	80,714
<a href="#">dissimulation</a>	0.151	16,750
<a href="#">tricherie</a>	0.150	17,370
<a href="#">faux-fuyant</a>	0.146	1,876
<a href="#">superficialité</a>	0.142	5,666
<a href="#">duperie</a>	0.140	5,161
<a href="#">fabulation</a>	0.136	2,099
<a href="#">faux-semblant</a>	0.134	6,431
<a href="#">mystification</a>	0.134	8,626
<a href="#">manigance</a>	0.127	7,987
<a href="#">traîtrise</a>	0.124	2,706
<a href="#">canular</a>	0.123	14,490
<a href="#">tromperie</a>	0.123	27,292
<a href="#">trucage</a>	0.120	10,533
<a href="#">duplicité</a>	0.120	8,374
<a href="#">dupe</a>	0.119	9,471
<a href="#">fausseté</a>	0.112	12,179
<a href="#">Désirs</a>	0.107	1,502
<a href="#">artifice</a>	0.105	89,634
<a href="#">pastiche</a>	0.104	8,306
<a href="#">sophisme</a>	0.102	9,378
<a href="#">leurre</a>	0.102	51,882
<a href="#">mirage</a>	0.101	33,570
<a href="#">scooterroues</a>	0.100	69
<a href="#">traîtrise</a>	0.100	5,849
<a href="#">illusions</a>	0.097	1,218



Рис. 3. Словесний простір предметної області слова *simulacre*

Дані в лівій частині таблиці демонструють числове вираження семантичної наближеності слів до ключового слова *simulacre* від найбільшого показника до найменшого. Праворуч у графічному вигляді слова представлені як крапки на площині, причому ті слова, що більш наближені до ключового за змістом, зображені більшими за розміром літерами ближче до

центру, а решта за параметром смислового віддалення – маленькими літерами ближче до периферії.

Якщо в разі запиту для окремого слова система вибудовує словесний простір предметної області, то при застосування функціоналу до словосполучення предметна область може бути представлена сукупністю понять [2], як на прикладі зі словосполученням *fake news*, яке дедалі частіше з'являється в медіапросторі (див. рис. 4).

Use WebBootCaT with selected words			
Multi-word	Score	F	Reff
<input type="checkbox"/> <a href="#">fake news</a>	1,735.92	38	0
<input type="checkbox"/> <a href="#">fausse information</a>	479.36	12	0
<input type="checkbox"/> <a href="#">fausses informations</a>	436.39	14	5,354
<input type="checkbox"/> <a href="#">élections indiennes</a>	365.88	8	21
<input type="checkbox"/> <a href="#">dernier hôpital</a>	319.00	7	0
<input type="checkbox"/> <a href="#">fausse info</a>	267.19	6	0
<input type="checkbox"/> <a href="#">fausses infos</a>	264.62	6	453
<input type="checkbox"/> <a href="#">liberté de la presse</a>	248.14	6	0
<input type="checkbox"/> <a href="#">fausses nouvelles</a>	218.57	5	571
<input type="checkbox"/> <a href="#">industrie de la désinformation</a>	183.62	4	0
<input type="checkbox"/> <a href="#">cellule secrète</a>	181.99	4	0
<input type="checkbox"/> <a href="#">fins de propagande</a>	169.71	4	943
<input type="checkbox"/> <a href="#">médias traditionnels</a>	144.66	5	6,697
<input type="checkbox"/> <a href="#">conseiller de campagne</a>	137.69	3	33
<input type="checkbox"/> <a href="#">manipulation des émotions</a>	137.28	3	62
<input type="checkbox"/> <a href="#">listes de diffusions</a>	134.34	3	0
<input type="checkbox"/> <a href="#">sites de référence</a>	126.92	3	999
<input type="checkbox"/> <a href="#">printemps arabes</a>	125.20	3	0
<input type="checkbox"/> <a href="#">réseaux sociaux</a>	112.07	23	95,899
<input type="checkbox"/> <a href="#">ère numérique</a>	109.76	3	0
<input type="checkbox"/> <a href="#">médias occidentaux</a>	92.91	3	0

Рис. 4. Предметна область словосполучення *fake news*

Отриманий завдяки автоматичній обробці корпусу перелік слововживань у складі предметної області словосполучення *fake news* можна взяти за основу для здійснення семантичного моделювання концептосфери *FAKE NEWS* і встановлення множинності відношень між компонентами, визначення семантики поняття в контексті інших понять.

Однією із проблем, що виникає при автоматичному наповненні тезаурусу на основі корпусу є певна кількість "шуму" (у нашому випадку, це, вочевидь, словосполучення *élections indiennes*, *dernier hospital*, *conseiller de campagne*, які не мають відношення до предметної області), який потрібно ефективно відсіювати шляхом ручної обробки отриманого матеріалу для отримання даних більшої точності.

Незважаючи на це, цілком очевидно, що функціонал корпусної лінгвістики надає унікальну можливість моделювання окремих предметних областей – частин реального світу, що розглядаються в межах даного контексту. На основі моделей предметних областей стає реальним розв'язання складних когнітивних завдань, що полягають у фіксуванні експертом усіх семантичних, лексичних, контекстних та інших типів відносин, здійсненні їхньої наступної класифікації за певними ознаками, їхньому порівнянні, доповненні опису значення новими семантичними ознаками тощо, і подальше концептуальне моделювання.

#### Список використаних джерел:

1. Боярская Е.Л. Разработка методологии когнитивного анализа многозначного слова // Методы когнитивного анализа семантики слова:

компьютерно-корпусный подход / Е.Л. Боярская ; под общ. ред. В.И. Заботкиной. – М. : Языки славянской культуры, 2015. – С. 81–119.

2. Крижановский А.А. Автоматизированное построение списков семантически близких слов на основе рейтинга текстов в корпусе с гиперссылками и категориями / А.А. Крижановский // Компьютерная лингвистика и интеллектуальные технологии : труды междунар. конф. Диалог'2006 "Компьютерная лингвистика и интеллектуальные технологии". – М. : РГГУ, 2006. – С. 297–302.

3. Саломатина Н.В. О возможностях автоматического выявления связей между терминами предметной области (на примере катализа) / Н.В. Саломатина // Компьютерная лингвистика и интеллектуальные технологии : труды Междунар. конф. Диалог'2010. / Саломатина Н.В., Гусев В.Д., Ильина Л.Ю. и др. – М. : Изд-во РГГУ, 2010.

4. Deulofeu H.-J. Une tâche à accomplir pour la linguistique française du XXIe siècle: élaborer une grammaire des usages du français / H.-J. Deulofeu, J.-M. Debaisieux // Langue française. – 2012. – № 176. – P. 27–46.

5. Gries Stefan Th. Exploring variability within and between corpora: some methodological considerations / Stefan Th. Gries // Corpora. – 2006. – № 1(2). – P. 109–151.

6. Jakubíček M. The TenTen Corpus Family / M. Jakubíček, A. Kilgarriff, V. Kovář et al. // Proceedings of the 7th International Corpus Linguistics Conference. – Lancaster, 2013. – P. 125–127.

7. Le corpus de référence du français contemporain (CRFC) [Electronic source]. – Access mode : [https://www.researchgate.net/publication/273847494\\_Le\\_corpus\\_de\\_reference\\_du\\_francais\\_contemporain\\_CRFC](https://www.researchgate.net/publication/273847494_Le_corpus_de_reference_du_francais_contemporain_CRFC) [accessed Oct 18 2018C.

8. Leipziger Corpus Francais [Electronic source]. – Access mode : [http://wortschatz.uni-leipzig.de/ws\\_fra](http://wortschatz.uni-leipzig.de/ws_fra).

9. Sahlgren M. The Distributional Hypothesis. From context to meaning / M. Sahlgren // Distributional models of the lexicon in linguistics and cognitive science (Special issue of the Italian Journal of Linguistics), Rivista di Linguistica. – 2008. – Vol. 20, no. 1. – P. 33–53.

10. Semino E. Corpus Stylistics : Speech, Writing and Thought Presentation in a Corpus of English Writing / E. Semino, M. Short. – L. ; N. Y. : Routledge, 2004. – 272 p.

Надійшла до редколегії 12.12.17

L. Klymenko, PhD, Researcher

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

## CORPUS AND COGNITIVE LINGUISTICS: METHODS CONVERGENCE PROSPECTS

The article provides a brief overview of the French language annotated corporas and highlights their functionals. The actual status of the French corpus linguistics and its scientific development directions are defined. The perspectives of applying computer science linguistics in discourse-cognitive studies, the possibility of an integrated interdisciplinary methodology developing are determined.

**Keywords:** corpus linguistics, corpora, cognitive linguistics, French language, methods.

Л.В. Клименко, канд. филол. наук, науч. сотр.

Київський національний університет імені Тараса Шевченка, Київ, Україна

## КОРПУСНАЯ И КОГНИТИВНАЯ ЛИНГВИСТИКА: ПЕРСПЕКТИВЫ КОНВЕРГЕНЦИИ МЕТОДОВ

Осуществлен краткий обзор аннотированных корпусов французского языка с указанием особенностей их функционалов. Освещен современный статус французской корпусной лингвистики, а также направления разработок, осуществляемых с использованием массивов франкоязычных текстов. Определены перспективы применения разработок компьютерной лингвистики в дискурсивно-когнитивных исследованиях, возможности разработки интегрированной междисциплинарной методики.

**Ключевые слова:** корпусная лингвистика, корпус, когнитивная лингвистика, французский язык, метод.

УДК 811.112.2'81'37

О. Материнська, д-р филол. наук, доц.

Київський національний університет імені Тараса Шевченка, Київ

## СТРУКТУРА ТА СЕМАНТИКА ТЕМПОРАЛЬНИХ МЕРОНІМІВ У НІМЕЦЬКІЙ І АНГЛІЙСЬКІЙ МОВАХ

Статтю присвячено дослідженню структурно-семантичних типів іменників – найменувань сегментів часу (темпоральних меронімів) у німецькій та англійській мовах. У дослідженні розроблено тематичну класифікацію темпоральних меронімів, визначено їхні структурно-семантичні типи, висвітлено шляхи розвитку полісемії за рахунок метафоричних і метонімичних переносів. Простежуються кореляції між структурною та семантичною організацією досліджуваних лексичних одиниць, що дозволяє створити модель їх системного опису. Результати дослідження демонструють розбіжності у сегментації темпоральних понять навіть у близькосторідних мовах: німецькій і англійській. Емпіричний матеріал сприяє систематизації даних про наявні у німецькій та англійській мовах стереотипи щодо впорядкування часу, способів його позначення та мовної репрезентації на матеріалі іменників двох досліджуваних мов.

**Ключові слова:** темпоральні мероніми, структурно-семантичний тип, полісемія, меронім, метафора, метонімія.

Відмінності у сформованій мовній картині світу відбиваються на семантиці лексичних одиниць, що проявляється у різній концептуалізації носіями конкретних

мов певних фрагментів об'єктивної дійсності, побудові її лінгвоспецифічних моделей та обумовлює у мовному плані різну логіко-семантичну сегментацію навіть базових