

УДК 81'373:578.27

DOI <https://doi.org/10.32782/2522-4077-2024-209-6>

КОРПУСНЕ ДОСЛІДЖЕННЯ МЕДІАОБРАЗУ «УКРАЇНА» В АМЕРИКАНСЬКИХ НОВИННИХ ТЕКСТАХ

MEDIA IMAGE OF "UKRAINE" IN AMERICAN NEWS TEXTS: A CORPUS-BASED STUDY

Бондаренко К. Л.,

orcid.org/0000-0002-2964-5123

кандидат філологічних наук,

доцент кафедри перекладу, прикладної та загальної лінгвістики

Центральноукраїнського державного університету імені Володимира Винниченка

Бурлака В. С.,

orcid.org/0000-0002-3055-6881

вчитель англійської мови комунального закладу

«Центральноукраїнський науковий ліцей-інтернат Кіровоградської обласної ради»

Єршова А. А.,

orcid.org/0009-0000-9429-3462

учениця 10 класу комунального закладу

«Центральноукраїнський науковий ліцей-інтернат Кіровоградської обласної ради»

Мета дослідження – встановлення лексичних особливостей формування медійного образу України у новинних текстах. Досліджень лінгвістів медіаобразу України в американських медіа під час президентських виборів в США недостатньо, незважаючи на його вагомість як політичного інструменту, яка стала очевидно на початку 2024 року. Ретельний аналіз медіаобразу вимагає значного часового ресурсу та великого обсягу релевантних даних. Комп'ютерна обробка текстів дозволяє лінгвістам зібрати необхідну інформацію. Дослідження базується на корпусному аналізі присвячених Україні новинних текстів, опублікованих у передвиборчий період джерелами, які орієнтуються на прихильників демократів (The New York Times) і республіканців (The Wall Street Journal). Зібраний корпус статей складає 291565 лексичних одиниць в 194 публікаціях, на основі яких було проаналізовано тематику новин. Використовуючи статистичний аналіз у програмі WordSmith 8.0 для обробки виділених текстів, було визначено найчастіше вживані (ключові) слова: *war, Russia, military, forces, defense, drone, weapon, time, soldiers, president, officials, Biden, government, Zelensky* тощо. Згодом за допомогою корпусного аналізу в програмі Sketch Engine було проаналізовано контексти, в яких вжито ці ключові слова. Виділені ключові слова (93 одиниці) за спільними семантичними ознаками і з урахуванням контекстів було об'єднано у лексико-семантичні групи, що вказують на фрагменти дійсності, актуальні для американських медіа: 1. ЛСГ «Назви військових дій» (22,6% від кількості ключових слів), «Назви актів підтримки» (21,5%), «Назви представників влади (адміністрації)» (17,2%), «Назви зброї» (13,9%), «Назви супротивників» (12,8%), «Назви територій» (7,5%), «Назви громадян» (7,5%).

Кількісні та якісні параметри ЛСГ свідчать про їхню значимість у контексті висвітлення цих аспектів, що, в свою чергу, підкреслює їх важливість для створення медіаобразу України. Розуміння ролі та важливості мовних засобів у медіатекстах може сприятливо вплинути на створення або корекцію образу країни за її межами.

Ключові слова: медіатекст, медіаобраз, корпусний аналіз, аналіз частотності, лексико-семантичне поле, лексико-семантична група.

The objective of the research is to identify the lexical features that shape the media image of Ukraine in news texts. There is a lack of linguistic studies on Ukraine's media image in the American media during the U.S. presidential elections, despite its significance as a political tool, which became evident at the beginning of 2024. Thorough analysis of the media image requires significant time resources and a large volume of relevant data. Computer processing of texts allows linguists to collect the necessary information. The study is based on the corpus analysis

of news texts dedicated to Ukraine, published during the pre-election period by sources catering to the supporters of the Democrats (The New York Times) and the Republicans (The Wall Street Journal). The compiled corpus of articles consists of 291,565 lexical units in 194 publications, which were used to analyze the thematic content of the news. Using statistical analysis in WordSmith 8.0 to process the extracted texts, the most frequently used (key) words were determined: *war, Russia, military, forces, defense, drone, weapon, time, soldiers, president, officials, Biden, government, Zelensky*, etc. Subsequently, the contexts in which these key words were used were analyzed using corpus analysis in the program Sketch Engine. The identified key words (93 units) were grouped into lexical-semantic groups (LSGs) based on common semantic characteristics and considering the contexts, which reflect the reality fragments that are current for the American media: 1. LSG «*Names of military actions*» (22.6% of the number of key words), «*Names of support acts*» (21.5%), «*Names of government (administration) representatives*» (17.2%), «*Names of weapons*» (13.9%), «*Names of adversaries*» (12.8%), «*Names of territories*» (7.5%), «*Names of citizens*» (7.5%). The quantitative and qualitative parameters of the LSGs highlight their significance in the context of highlighting these aspects, which in turn emphasizes their importance for creating the media image of Ukraine. Understanding the role and significance of linguistic means in media texts can positively influence the creation or correction of the country's image abroad.

Key words: media text, media image, corpus analysis, frequency analysis, lexical-semantic field, lexical semantic group.

Постановка проблеми. Питання медійного образу України на шпальтах світової преси з початком повномасштабного вторгнення набуло особливої актуальності. Сприйняття російсько-української війни як події загальносвітового масштабу спричинило активізацію публікацій української тематики, як у Європі, так і за океаном. Так у «The Wall Street Journal» за період з листопада 2019 р. по 23 лютого 2022 р., було опубліковано 81 таку статтю, а в період з 24 лютого 2022 р. по 30 листопада 2023 р. – 3078. Станом на початок 2024 року констатуємо, що українське питання в США стало аргументом в боротьбі за президентське крісло. Враховуючи те, що медіа – це джерела, що формують ставлення потенційних виборців до проблеми, припускаємо, що системний аналіз медійних джерел, яким надають перевагу прихильники тієї чи іншої політичної сили, може висвітлити особливості того медіаобразу, який сформовано або формується у читачів. Сформований у читачів-виборців медіаобраз країни – надійне підґрунтя для подальшої реалізації претендентом на президентську посаду зовнішньополітичної стратегії. Корпусний аналіз допомагає визначити продуктивність мовних елементів у масивах текстів, а інтерпретація даних, які мають статистичне значення, за допомогою традиційних методів, таких як теорія полів, дозволяє виділити найважливіші для медіаобразу фрагменти дійсності, діагностувати чи спрогнозувати ставлення до цих цільової аудиторії, а відтак, і політиків, які втілюють очікування цієї аудиторії.

Аналіз останніх досліджень і публікацій. Питання медіаобразу образу України в ЗМУ давно вийшло поза межі журналістики і стало предметом уваги лінгвістів. Увага дослідників була зосереджена на аналізі медійного образу України, створеного ЗМІ Німеччини [1], Індії [2], Іспанії [3] тощо. Проте медійний образ України у американських ЗМІ під час президентських виборів у США, не був достатньо детально досліджений лінгвістами, хоча на початку 2024 року стало очевидним, що тема підтримки України використовується як важливий політичний аргумент у боротьбі за посаду президента США. Формування медіаобразу країни потребує часу, а аналіз засобів його формування – репрезентативного масиву відповідної інформації. Комп'ютерна обробка вербальної інформації дає змогу аналізувати релевантні текстові масиви та вибирати ключові елементи для детального аналізу, надаючи лінгвістам точні кількісні дані для обґрунтування або спростування припущень [4]. Усвідомлюючи потужну силу ЗМІ, які можуть як висвітлювати реальні події, так і маніпулювати свідомістю через перекручування або приховування фактів, дослідники намагаються вивчити механізми та вербальну репрезентацію поточних подій у нашій країні із застосуванням корпусних інструментів [5; 6], які дозволяють аналізувати різні контексти, виводити з них відтінки значення слова, моделювати смислові поля [7].

Метою дослідження було визначення особливостей вербалізації медіаобразу «Ukraine» в американських новинних виданнях. Досягнення мети передбачало виконання низки завдань: формування корпусу новинних текстів «The New York Times» (NYT) та «The Wall Street Journal» (WSJ), присвячених Україні; визначення ключових слів за допомогою інтегрованої програми WordSmith 8.0. та онлайн-інструменту SketchEngine в новинних текстах NYT та WSJ. Вибір джерел дослідження був умотивований необхідністю показати особливості висвітлення подій в Україні виданнями, які користуються прихильників різних політичних сил (республіканців та демократів). За даними дослідницького центру П'ю (Pew Research Center), серед друкованих новинних видань найвищий рівень довіри серед демократів має газета «The New York Times» (53% опитуваних абсолютно довіряє і 41% більше довіряє, ніж не довіряє), серед республіканців – газета «The Wall Street Journal» (24% опитаних абсолютно довіряє і 57% більше довіряє, ніж не довіряє) [8]. Матеріалом дослідження став корпус загальною кількістю 291565 лексичних одиниць, який був створений шляхом суцільної вибірки тематичних статей про Україну у газетах The New York Times (152 статті обсягом 152551 слово) та The Wall Street Journal (142 статті обсягом 139014 слів), опублікованих за вересень-жовтень-листопад 2023 року на офіційних сайтах видань [9, 10]. За нашою гіпотезою, визначена на основі корпусного аналізу частотність уживання лексичних одиниць в опрацьованому матеріалі свідчить про важливість фрагментів дійсності, номінованих цими одиницями. Встановлення сполучуваності цих одиниць у кожному окремо взятому корпусі визначає специфіку висвітлення зазначених фрагментів дійсності в ідеологічно-відмінних медіа.

Виклад основного матеріалу. Медіаобраз – це «одиниця виразних засобів ЗМІ, яка представляє собою художню конструкцію, спрямовану на моделювання не стільки фактичної, як емоційної складової медіаподії» [11, с. 86]. Вивчення специфіки медіаобразу має враховувати не лише встановлення специфіки мовних засобів, які використовують для висвітлення проблем, а й наміри творців (редакційної колегії, власників видань тощо) та реакцію аудиторії, на яку здійснюється вплив [12]. Очевидно, що в свідомості творців і споживачів медіаобраз може бути відмінним, зважаючи на ефективність подачі, активність формування медіаобразу в медіа, неоднорідність джерел, завдяки яким формується медіаобраз у споживачів медійного продукту тощо. В подальшому під сформованим у споживачів медіаобразом країни розумітимемо набір переконань про державу, що формується під впливом інформації, яку люди отримують з медійних джерел.

Медійними джерелами цієї розвідки стали 194 новинних тексти двох авторитетних ЗМІ – «The New York Times» (далі – NYT) та «The Wall Street Journal» (далі – WSJ) опубліковані у вересні- листопаді 2023 року [the NEW, The Wall]. Обрані для аналізу новинні тексти мають різну цільову аудиторію (республіканців і демократів), що дозволяє припускати неоднакове висвітлення подій в Україні і відмінність медіаобразу України, сформованого цими виданнями. Мовні засоби, за допомогою яких формується образ, варто досліджувати з урахуванням ключових (найрозповсюдженіших) слів, контекстів, у яких їх використано, семантичного оточення тощо [13]. Автоматизований аналіз корпусу медіатекстів, що надає статистичні дані, на основі яких може здійснюватися інтерпретація мовних фактів, – достовірний спосіб встановити особливості медіаобразу України в сучасному американському медіадискурсі.

Ідея виокремлення та аналізу ключових слів ґрунтується на гіпотезі, що часте використання (рекурентність) певних способів вираження понять і ідей дозволяє розкрити розуміння соціального світу [14, с. 597]. Т. ван Дейк наголошує на тому, що вибір словникових одиниць має вирішальне значення в контексті новинного дискурсу, адже саме через лексику виявляються приховані інтенції та ідеологія [15, с. 176]. У контексті досліджень ЗМІ вивчення ключових слів дозволяє тлумачити ідеологічний сенс (ideological meaning), що має бути доведений до певної аудиторії [16, с. 565]. В українському мовознавстві існують певні традиції аналізу ключових слів корпусів текстів новин [17; 18], які довели свою продуктивність.

Виділення ключових слів у аналізованих нами корпусах дозволяє визначити не просто тематику (aboutness), тобто розуміння основних концептів, тем та поглядів [19], а й простежити формування системи переконань читачів стосовно висвітлених тем.

Грунтовний статичний аналіз дозволяють провести інструменти програми WordSmith 8.0. Аналіз частотності слів (WordList) дозволяє оцінити важливість обговорюваної тематики, їхню значущість у досліджуваному корпусі. Зрозуміло, що найчастіше уживаними словами виявилися службові частини мови (артиклі, прийменники і сполучники тощо) та найпоширеніші дієслова (said, has, was, are). В подальшому за рекомендацією розробників програми такі слова нами було видалено вручну.

За результатами аналізу було виявлено, що з-поміж загальної кількості слів (291565) найуживанішими є Ukraine (3521 випадок вживання – 2,53% тут і далі від загальної кількості слів), Russian (1874 – 1,35%), Russia (1698 – 1,22%), Ukrainian (1638 – 1,18%), war (1277 – 0,92%), military (1022 – 0,74%), forces (650 – 0,47%), drones (574 – 0,41%), Kyiv (560 – 0,4%), support (519 – 0,37%), President (483 – 0,35%), missiles (479 – 0,34%), defense (467 – 0,34%), aid (412 – 0,30%), weapons (383 – 0,28%), invasion (285 – 0,20%). Такі результати свідчать, що головна тема для обговорення у ЗМІ – війна в Україні: сторони протистояння, підтримка українського захисту, атаки загарбників, зброя.

Надалі для об'єктивності процесу аналізу розробники програми пропонують вдатися до лематизації – групування форм слова для аналізу як тотожного елементу. Тому як леми нами розглядалися слова: *Ukraine/Ukraine's/Ukrainian/Ukrainians, drone /drones /drone's; child /children; Biden/ Biden's; change /changes /changing /changed; war /war's/wars; begin /begins /begun* тощо. Використання лематизації призвело до усунення «шумів» – частотних, але семантично не насичених одиниць, типу артиклів, прийменників, форм дієслова *to be* тощо.

Word	Freq	%	Tests	%	Dispersion
THE	18.081	6.12%	6	100.00%	0.82
TO	18.952	6.49%	6	100.00%	0.82
OF	9.221	3.16%	6	100.00%	0.82
AND	7.123	2.41%	6	100.00%	0.82
A	6.986	2.37%	6	100.00%	0.82
IN	6.951	2.36%	6	100.00%	0.82
THAT	3.867	1.31%	6	100.00%	0.82
FOR	2.953	1.00%	6	100.00%	0.83
IS	2.710	0.92%	6	100.00%	0.82
ON	2.663	0.90%	6	100.00%	0.81
#	2.632	0.89%	6	100.00%	0.81
SAID	2.536	0.86%	6	100.00%	0.81
IS	2.526	0.86%	6	100.00%	0.81
IT	1.920	0.65%	6	100.00%	0.81
RUSSIAN	1.874	0.63%	6	100.00%	0.81
AS	1.853	0.63%	6	100.00%	0.82
WITH	1.826	0.62%	6	100.00%	0.82
HAS	1.803	0.61%	6	100.00%	0.81
UKRAINIAN	1.638	0.55%	6	100.00%	0.80
FROM	1.429	0.49%	6	100.00%	0.81
HAVE	1.467	0.50%	6	100.00%	0.82
HE	1.460	0.49%	6	100.00%	0.82
BY	1.459	0.49%	6	100.00%	0.81
WAS	1.390	0.47%	6	100.00%	0.82
ARE	1.349	0.46%	6	100.00%	0.82
AT	1.290	0.44%	6	100.00%	0.82
WAR	1.277	0.43%	6	100.00%	0.81
BUT	1.268	0.43%	6	100.00%	0.81
ITS	1.211	0.41%	6	100.00%	0.81
RUSSIA	1.197	0.41%	6	100.00%	0.82
AM	1.166	0.39%	6	100.00%	0.81
BE	1.116	0.38%	6	100.00%	0.81
THEY	1.080	0.37%	6	100.00%	0.83

Рис. 1. Результати пошукового запиту щодо частотності вживання слів до і після лематизації в Sketch Engine

За допомогою інструменту KeyWords було виділено ключові слова, які засвідчили, що увага у статтях акцентується в першу чергу на Україні і її громадянах (6941 випадок вживання – 4,99% від загальної кількості слів), природно, що наступним за частотністю є питання війни, її учасників та зброї: *war* (1379 – 0,99%), *Russia* (4139 – 2,97%), *military* (1041 – 0,75%), *forces* (783 – 0,56%), *defense* (705 – 0,51%), *drone* (595 – 0,43%), *weapon* (383 – 0,28%), *time* (399 – 0,29%), *soldiers* (354 – 0,25%), *counteroffensive* (222 – 0,16%) тощо. Також в центрі уваги – ключові фігури політикуму: *president* (483 випадки вживання – 0,35%), *officials* (457 – 0,33%), *Biden* (414 – 0,30%), *government* (401 – 0,29%), *Zelensky* (375 – 0,27%). Однією з найбільш обго-

(447), *Kyiv* (337 випадків вживання у значенні центру прийняття рішень), *authority* (111), *Zaluzhny* (28).

4. ЛСГ «Назви зброї» 13,9% від 93 ключових слів (11% від усіх випадків уживання в корпусі): *drones* (574), *missiles* (479), *weapons* (383), *ammunition* (113).

5. ЛСГ «Назви супротивників» 12,8% від 93 ключових слів (17,7% від усіх випадків уживання в корпусі): *Russian* (1874), *Russia* (1698), *Moscow* (385), *Putin* (376), *Kremlin* (122).

6. ЛСГ «Назви територій» 7,5% від 93 ключових слів (3% від усіх випадків уживання в корпусі): *Ukraine* (1253 випадків вживання у значенні місця подій), *Kyiv* (223 випадків вживання у значенні місця подій), *Eastern* (166), *Southern* (141), *central* (57), *Western* (37).

7. ЛСГ «Назви громадян» 7,5% від 93 ключових слів (4,5% від усіх випадків уживання в корпусі): *Ukrainians* (298), *citizen* (38), *local* як іменник (10).

Окремо зауважимо означення, які вживаються у сполученні зі словом *Ukraine*: *occupied* (93), *strong* (90), *free* (47), *future* (26), *democratic* (25), *war-torn* (12), *sovereign* (9), *targeted* (5), *thriving* (3), *modern-day* (3), але відмінність семантики не дозволяє сформуванню з них ЛСГ. Поєднання означень із ключовими іменниками у корпусі досліджуваних текстів є вкрай важливим, адже саме вони можуть нести оцінне значення та конотації, які впливають на емоції читача і формують уявлення про певні події та факти.

Розглядаючи лексико-семантичне поле і групи, що його формують, як «мовну модель реальної картини світу», на основі аналізу лексико-семантичних груп лексико-семантичного поля «UKRAINE» в обох аналізованих джерелах доходимо висновку, що в медіафокусі – війна та люди, що беруть у ній участь або страждають від її наслідків; меншою мірою – зброя та гроші.

Висновки. Встановлення особливостей медіаобразу передбачає виокремлення фактів та характеристик, що оприлюднюються медіа і впливають на формування ставлення чи переконання цільової аудиторії до фрагменту дійсності. Визначені в результаті опрацювання корпусу ключові слова загалом спільні для корпусів обох видань. Використання елементів компонентного аналізу і врахування контексту, в яких ужито ту чи іншу лексему, уможливило викремлення низки лексико-семантичних груп, кількісні та якісні характеристики яких указують на актуальність для висвітлення, а відтак – важливість для формування медіаобразу України. Розуміння ваги і значущості мовних засобів в медіатекстах може мати позитивний вплив на формування або корегування медіаобразу країни за кордоном.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:

1. Данилевич А. Інформаційний імідж і образ України у друкованих ЗМІ Німеччини: на межі понять. Медіапростір, вип. 10, 2017. С. 85–91.
2. Польовик С.М. Образ України у світі: аналіз азійських ЗМІ. Вісник ОНУ. Серія: Бібліотекознавство, бібліографознавство, книгознавство, т. 23, вип. 1, 2018. С. 97–106.
3. Бондаренко К., Балуба О., Грабовська А. Медійний образ України в сучасних іспанських ЗМІ (за даними корпусного аналізу). Наукові записки. Серія: Філологічні науки, вип.175, 2019. С. 518–523.
4. Жуковська В. В. Вступ до корпусної лінгвістики: навчальний посібник. Житомир: Вид-во ЖДУ ім. І. Франка, 2013. 142 с.
5. Леврінц М. Корпусний аналіз репрезентування війни в Україні у міжнародній пресі. Вісник ХНУ імені ВН Каразіна. Серія: Іноземна філологія. Методика викладання іноземних мов, (96), 2022. С. 61–68.
6. Біляк І.В. Прагматика англомовного новинного медіатексту: маніпулятивний потенціал. Вчені записки Таврійського Національного університету ім. В.І. Вернадського. Серія Філологія. Соціальні комунікації, том 25 (64), №1, часть 1, 2012. С. 221–226.
7. Швайбер Д.І., Романишин, Н.І. Використання конкордансів та корпусних технологій для дослідження концепту ЗДОРОВ'Я / HEALTH. Молодий вчений, вип. 11 (87), 2020. С. 508-512.
8. Democrats report much higher levels of trust in a number of news sources than Republicans. Pew Research Center. [Електронний ресурс]. URL: <https://www.pewresearch.org/journalism/2020/01/24/democrats-report-much-higher-levels-of-trust-in-a-number-of-news-sources-than-republicans/#sortable-tables> (дата звернення: 01.11.2023).

9. The New York Times. [Електронний ресурс]. Режим доступу: <https://www.nytimes.com/international/>.
10. The Wall Street Journal. [Електронний ресурс]. Режим доступу: <https://www.wsj.com/>.
11. Балалуєва І. А. Медіаобраз і соціообраз: процеси взаємоуправління в інформаційному суспільстві. *Право і управління. XXI століття*, №4 (33), 2014. С. 86–91.
12. Tarasheva E. The Image of a Country created by International Media: The Case of Bulgaria. Newcastle upon Tyne, 2014. 185 p.
13. Махачішвілі Р.К., Білик К.М. Дослідження функціонально-лінгвістичних аспектів заголовків текстів новин у сучасному французькому медіадискурсі. *Науковий вісник Міжнародного гуманітарного університету. Серія: Філологія*, № 48 (том 1), 2021. С.139–143.
14. Ädel, A. How to use corpus linguistics in the study of political discourse. In *The Routledge Handbook of Corpus Linguistics*, eds. A. O’Keeffe, M. McCarthy. London and New York: Routledge, 2010. С. 591–604.
15. Van Dijk, T.A. *News as Discourse*. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc., 1988. 200 p.
16. O’Halloran K. How to use corpus linguistics in the study of media discourse. In *The Routledge Handbook of Corpus Linguistics*, eds. A. O’Keeffe, M. McCarthy. London and New York: Routledge, 2010. С. 563–577.
17. Ділай М. П. Ключові слова англійськомовного новинного дискурсу у 2021 році: корпусний підхід. *Закарпатські філологічні студії*, редкол.: І. М. Зимомря (гол. ред.), М. М. Палінчак, Ю. М. Бідзіля та ін. Ужгород: Видавничий дім «Гельветика», 2021. Т.1, Вип. 19. С. 124–133.
18. Білик К.М. Жанр «надзвичайні новини» у сучасному європейському медіадискурсі. дис. ... доктора філософії: 035. Київський університет імені Бориса Грінченка, Київ, 2023. 263 с.
19. Phillips M. *Lexical Structure of Text. Discourse Analysis Monograph*, No. 12. English Language Research, University of Birmingham, 1989. 86 p