МОВОЗНАВСТВО. ЛІТЕРАТУРОЗНАВСТВО

УДК 81’322:81’33:821.161.2-3
DOI https://doi.org/10.24919/2308-4863/99-2-52

Світлана ЛУЦАК,
orcid.org/0000-0002-6862-5838
доктор філологічних наук, професор,
завідувач кафедри мовознавства
Івано-Франківського національного медичного університету
(Івано-Франківськ, Україна) slucak@ifnmu.edu.ua

Алла ТАРАН,
orcid.org/0000-0001-8091-1477
кандидат філологічних наук, доцент,
доцент кафедри українського мовознавства і прикладної лінгвістики
Черкаського національного університету імені Богдана Хмельницького
(Черкаси, Україна) alla__taran@ukr.net

Галина ЮРЧАК,
orcid.org/0000-0002-2740-9012
кандидат філологічних наук, доцент,
доцент кафедри мовознавства
Івано-Франківського національного медичного університету
(Івано-Франківськ, Україна) haliur@ifnmu.edu.ua

КОРПУС ТЕКСТІВ МОДЕСТА ЛЕВИЦЬКОГО ЯК ДОСЛІДНИЦЬКИЙ І НАВЧАЛЬНИЙ РЕСУРС

У статті досліджено корпус текстів українського письменника, лікаря та громадського діяча Модеста Левицького, що становить важливе джерело для вивчення мовних і культурних процесів кінця XIX - початку XX ст. Метою є створення корпусу текстів Левицького та його аналіз із застосуванням цифрових інструментів корпусної лінгвістики для визначення статистичних характеристик, стилістичних домінант та ключових мовних особливостей. У дослідженні застосовано комплекс корпусних методів аналізу із використанням платформи AntConc. Зокрема, проведено частотний аналіз, що дозволив визначити найуживаніші лексеми та виокремити стилістичне ядро текстів. Використання n-грамного аналізу забезпечило виявлення стійких словосполучень, які відтворюють характерні моделі синтаксичної організації та комунікативні практики доби. Колокаційний аналіз дав змогу простежити взаємозв’язки між ключовими словами та їхніми контекстними партнерами, що розкриває соціально-професійні та побутові аспекти мовлення. За допомогою конкордансного аналізу було детально досліджено функціонування окремих лексем, зокрема «спитав» і «лікар», у різних наративних та інституційних контекстах. Нарешті, контекстуальний аналіз дозволив інтерпретувати отримані статистичні дані у зв’язку з ідейно-тематичними домінантами прози М. Левицького. Результати показали, що корпус налічує понад 111 тисяч слововживань і близько 21 тисячі унікальних словоформ, що свідчить про значне лексичне багатство текстів. Виявлено домінування дієслів мовлення, які задають динаміку та драматизм оповіді, а також значний пласт соціально-побутової й професійної лексики, що репрезентує різні сфери суспільного життя. Особливе місце займають лексеми «спитав» і «лікар», які унаочнюють баланс між діалогічністю як стилістичною домінантою та медичною проблематикою як соціально значущим мотивом. Отже, використання корпусних методів дозволило виявити подвійність мовної організації у прозі Левицького, де поєднуються реалістичне відтворення народного мовлення і професійно-медичний дискурс. Це формує унікальний ідіостиль письменника й відкриває перспективи для подальших міждисциплінарних студій.

Ключові слова: *корпусна лінгвістика, ідіостиль, Модест Левицький, лінгвостатистичний аналіз, колокації, медична лексика. діалогічність.*

Svitlana LUTSAK,
orcid.org/0000-0002-6862-5838
Doctor of Sciences in Philology, Professor,
Head of the Department of Linguistics
Ivano-Frankivsk National Medical University
(Ivano-Frankivsk, Ukraine) slucak@ifnmu.edu.ua

Alla TARAN,
orcid.org/0000-0001-8091-1477
Candidate of Philological Sciences,
Associate Professor at the Department of Ukrainian Linguistics and Applied Linguistics
Bohdan Khmelnytsky National University of Cherkasy
(Cherkasy, Ukraine) alla__taran@ukr.net

Halyna YURCHAK,
orcid.org/0000-0002-2740-9012
Candidate of Philological Sciences,
Associate Professor at the Department of the Department of Linguistics
Ivano-Frankivsk National Medical University
(Ivano-Frankivsk, Ukraine) haliur@ifnmu.edu.ua

THE CORPUS OF MODEST LEVYTSKYI'S TEXTS AS A RESEARCH AND EDUCATIONAL RESOURCE

The article explores the text corpus of the Ukrainian writer, physician, and public figure Modest Levytskyi, which serves as an important source for studying linguistic and cultural processes of the late 19th – early 20th centuries. The purpose is to create a corpus of Levytskyi's texts and analyze it using digital tools of corpus linguistics to identify statistical characteristics, stylistic dominants, and key linguistic features. The study employed a set of corpus-based methods using the AntConc platform. In particular, frequency analysis was applied to identify the most frequent lexemes and determine the stylistic core of the texts. N-gram analysis revealed stable lexical patterns that reproduce characteristic syntactic structures and communicative practices of the period. Collocational analysis made it possible to trace the connections between key words and their contextual partners, highlighting both social-professional and everyday aspects of speech. Concordance analysis allowed a detailed examination of the use of specific lexemes, especially "spytav" ("asked") and "likar" ("doctor"), across different narrative and institutional contexts. Finally, contextual analysis provided an interpretive framework for relating the statistical results to the thematic and stylistic dominants of Levytskyi prose. The processed corpus contains more than 111,000 word tokens and about 21,000 unique word forms, indicating significant lexical richness. The analysis revealed the dominance of speech verbs, which create narrative dynamics and dramatism, as well as a substantial layer of socio-professional and everyday vocabulary reflecting various spheres of social life. Special attention is paid to the lexemes "spytav" and "likar", which illustrate the balance between dialogicity as a stylistic dominant and medical discourse as a socially significant theme. Thus, the application of corpus methods revealed the dual nature of Levytskyi language organization, where realistic representation of folk speech coexists with professional-medical discourse. This balance constitutes the uniqueness of the writer's idiostyle and opens new perspectives for interdisciplinary research.

Key words: *corpus linguistics, idiolect, Modest Levytskyi, linguostatistical analysis, collocations, medical vocabulary, dialogicity.*

Постановка проблеми. Сучасний розвиток цифрових технологій та корпусної лінгвістики відкриває нові можливості для вивчення мовної та літературної спадщини. В українському науковому просторі спостерігається інтенсивна динаміка зацікавлення корпусними дослідженнями, яка «виявляється в постійному оновленні та розширенні лінгвістичних корпусів, збільшенні кількості наукових праць» (Луцак, Юрчак, 2025: 34). У цьому контексті творчість Модеста Левицького – письменника, лікаря, педагога й громадського діяча – залишається малодослідженою з погляду корпусного аналізу. Його твори становлять цінне джерело для розуміння мовних, культурних та історичних процесів кінця XIX – початку XX століття. Створення корпусу текстів Левицького та його опрацювання за допомогою сучасних інструментів корпусної лінгвістики є важливим кроком для розвитку вітчизняних гуманітарних студій та інтеграції українських досліджень у світовий науковий простір.

Аналіз останніх досліджень і публікацій. У сучасному науковому дискурсі корпусна лінгвістика «посідає провідне місце серед інно-

ваційних підходів до аналізу мови, адже вона забезпечує поєднання емпіричної достовірності, масштабності даних і об'єктивності висновків» (Луцак, Юрчак, 2025:42) й утверджується як ефективний міждисциплінарний інструмент дослідження мови, що поєднує теоретичні засади, цифрові технології та статистичні методи. Теоретичні засади підходу до організації й дослідження лексики в межах корпусу докладно окреслено у працях О. Демської, В. Жуковської, Є. Карпіловської, де корпус розглядається як «новітній інформаційно-дослідницький інструментарій сучасного мовознавства» (Демська, 2011:36). В. Жуковська обґрунтовує методологічні основи й напрями застосування корпусної лінгвістики (Жуковська, 2020); D. Stetsenko та I. Okulska висвітлюють можливості StyloMetrix для автоматизованого аналізу українських текстів (Stetsenko, Okulska, 2023); T. Walkowiak аналізує стилометричні підходи до ідентифікації авторства та соціальних характеристик (Walkowiak, 2020); I. Bashir зі співавторами підкреслюють подвійний статус корпусної лінгвістики як теорії і методології та її значення у викладанні мов (Bashir, Yunus, Ibrahim, 2018); С. Голощук окреслює сучасний стан і перспективи галузі та стверджує, що «корпусний підхід до вивчення мови відкриває нові можливості як для лінгвістів, так і для фахівців інших галузей науки… на основі корпусної вибірки, можна чітко та зі статистичними даними простежити ту чи іншу словосполучуваність та підібрати найкращу з них» (Голощук, 2022: 251). Практичний вимір представлено у працях Р. Ворошилової – у контексті становлення корпусної прагматики (Волошилова, 2025), В. Степанова і А. Могильника – у дослідженні фразеології та сленгу (Степанова, Могильник, 2025), а Н. Плюта акцентує на репрезентативності корпусів і статистичній надійності результатів (Плюта, 2025). Отже, корпусна лінгвістика постає як динамічна галузь із широкими аналітичними та прикладними можливостями.

Попри наявність ґрунтовних напрацювань у галузі корпусної лінгвістики, окремі аспекти залишаються недостатньо опрацьованими, зокрема можливості застосування корпусних технологій до аналізу творчої спадщини українських письменників, передусім Модеста Левицького, а також специфіка діалогічності, контекстуальне функціонування ключових термінів і прагматичні особливості його мовлення.

Саме ці аспекти зумовили формування **мети** нашого дослідження – створення корпусу текстів Модеста Левицького та його аналіз за допомогою сучасних цифрових методів корпусної лінгвістики (зокрема інструментів платформи *AntConc*) для виявлення статистичних характеристик, домінантних лексем і стійких словосполучень, а також дослідження функціонування ключових термінів у контексті. Реалізація цієї мети дозволяє визначити особливості ідіостилю письменника, простежити соціально-історичні та професійні нашарування його мови й окреслити практичну цінність корпусу як ресурсу для мовознавчих, літературознавчих та міждисциплінарних студій.

Виклад основного матеріалу. Проведене корпусне дослідження текстів Модеста Левицького (Левицький, 2016) дало змогу виокремити низку кількісних і якісних характеристик, що відображають його мовну та стильову своєрідність. Загальний обсяг корпусу становить понад 111 тисяч слововживань, серед яких зафіксовано понад 21 тисячу унікальних словоформ. Така співвідносна висота показників свідчить про багатство лексики й виразну стилістичну варіативність, притаманну авторському письму. Щільність словникового запасу, що дорівнює 0,189, підтверджує різноплановість використаних мовних засобів та розмаїття лексичних рівнів – від побутової до професійної й суспільно-політичної термінології. Індекс читабельності, який становить 7,355, демонструє помірний рівень складності тексту: твори Левицького залишаються доступними для широкого читача, але водночас зберігають інтелектуальну глибину та стилістичне багатство. Середня довжина речення у 11,9 слова вказує на прагнення автора до динамічного викладу, що поєднує описову насиченість із діалогічністю. Статистичні параметри корпусу відображають цілісну картину ідіостилю письменника: багатий словниковий запас, поєднання простоти й інтелектуальної глибини, а також тяжіння до живої, наближеної до усного мовлення структури викладу. Корпусний аналіз «дає змогу об'єктивізувати висновки через статистичну репрезентативність» (Луцак, Юрчак, 2025: 44). Саме ці особливості формують основу для подальшого частотного, контекстуального та стилістичного аналізу.

Для визначення ключових лексем корпусу текстів М. Левицького застосовано інструмент AntConc, що забезпечує можливість автоматизованого підрахунку частотності слів і візуалізації результатів. Варто лише «ввести слово в пошукову систему відповідного корпусу, щоб отримати повний індекс його ж словоформ у найближчих і розширених контекстах із відповідною адресацією» (Таран, 2019: 100). Використання цього програмного забезпечення дозволило швидко ідентифікувати найуживаніші слова, простежити

закономірності їх функціонування та виділити лексеми, які формують стилістичне ядро текстів. Частотний аналіз дав змогу виявити домінування діалогічних дієслів, що створюють ефект динамічності й драматизму, а також велику кількість соціально-побутової та професійної лексики, котра віддзеркалює різні сфери суспільного життя та професійний досвід автора. Завдяки інструментам AntConc стало можливим не лише кількісно описати мовні особливості, а й визначити ключові контексти їх функціонування, що у підсумку сприяє глибшому розумінню ідіостилю письменника та відкриває нові перспективи для міждисциплінарних досліджень.

Вибір для детального корпусного аналізу саме лексем *«спитав»* та *«лікар/лікарня»* зумовлений їхньою особливою статистичною та стилістичною репрезентативністю. По-перше, дієслово *«спитав»* належить до найбільш уживаних у корпусі дієслів мовлення, що відіграють ключову роль у побудові діалогічної тканини текстів. Його регулярність, багатофункціональність і здатність маркувати зміну наративних позицій роблять цю лексему своєрідним «індикатором поліфонії» у прозі Левицького. По-друге, професійна медична лексика (*«лікар», «лікарь», «лікарня»*) формує інший важливий пласт стилістичного ядра, оскільки безпосередньо відображає соціальні практики та інституційні реалії доби. Її концентрація в корпусі дозволяє реконструювати специфіку медичного дискурсу та критику стану охорони здоров'я у кінці XIX – на початку XX століття.

Отже, обидва лексичні блоки – дієслово *«спитав»* і медичні терміни – репрезентують дві взаємодоповнюючі домінанти ідіостилю письменника: діалогічність і соціальну інституційність. Їхнє поєднання створює подвійне стилістичне ядро прози Левицького, що водночас забезпечує ефект живого мовлення й актуалізує суспільно значущі теми.

Аналіз корпусу із зафіксованими прикладами вживання дієслова *«спитав»* у прозі М. Левицького засвідчив його провідну роль у формуванні діалогічної тканини тексту. Частотність цієї лексеми у різних оповідних ситуаціях виявляє її універсальність і функціональну багатогранність: вона з'являється у побутових сценах, у родинному колі, у службово-офіційних контактах та у діалогах мандрівників. У корпусі виразно простежуються три провідні суб'єктні ролі, з якими пов'язане це дієслово: узагальнений персонаж («він»), що може належати до будь-якої соціальної групи; оповідач («я»), який інтегрується у комунікативний простір твору; і професійна постать лікаря, що унаочнює реалістичне відтворення спеціалізованої комунікації. Такий розподіл дозволяє розглядати *«спитав»* як індикатор зміни наративних позицій і засіб створення поліфонії.

Змістові контексти використання дієслова репрезентують широкий спектр тематичних полів: від дорожніх і побутових питань («Далеко ще до Берестечка?») та інтимно-родинних реплік («А чого б ви, дітки, хотіли собі на ялинку?») до медичних діалогів («Ну, як ви себе почуваєте?») і соціально-службових запитів («Від станового?»). Завдяки цьому *«спитав»* постає лексемою-маркером різних комунікативних практик, що відображають соціальну структуру і професійний досвід автора. Повторюваність цієї лексеми створює відчутний ритм і надає текстові ефекту живої розмови, тоді як у порівнянні з іншими дієсловами мовлення (*сказав, мовив, відповів*) саме *«спитав»* концентрує драматичну напругу і стає стилістичною домінантою.

У цьому контексті особливої ваги набувають лексеми *«лікарь»*, *«лікар»*, *«лікарі»*, *«лікарня»*, які формують інший полюс стилістичного ядра прози Левицького. Вони позначають важливий пласт соціально-побутової та професійної лексики, де медицина, хвороби та лікування постають як ключові елементи повсякденного життя громади. Образ лікаря реалізується у різних функціях: від учасника безпосереднього діалогу («Ну-ну, – перебив лікарь»; «Беріть, – сказав лікарь») до авторитетної постаті, що постає посередником між життям і смертю («Лікарь не сказав нічого, тільки розвів руками»). Згадування лікарів у множині («ходили по лікарях, і по шептухах та знахорях») підкреслює амбівалентність народних уявлень, де поряд із науковою медициною продовжують існувати традиційні форми лікування.

Не менш показовим є функціонування лексеми *«лікарня»*, що позначає не лише простір лікування, а й символізує інституційну недосконалість та соціальну проблематику. Описи земських лікарень, їхніх умов та побутових нарікань («у такій лікарні не можна держати хворих, хіба на погибель») відображають реалії медичної сфери другої половини XIX століття. У такий спосіб *«лікарня»* постає маркером державної опіки і водночас критики бюрократичної системи. У поєднанні з дієсловами мовлення (*«спитав»*, *«сказав»*, *«мовив»*), медична лексика формує подвійне стилістичне ядро: перший компонент задає динаміку діалогу і поліфонію, другий відтворює професійні та інституційні реалії. Це дозволяє зробити висновок, що корпус Левицького відзначається унікальною наративною стратегією, де індивідуальний голос

персонажа співіснує з критичним осмисленням соціальних практик, а медицина стає важливим лексико-стилістичним ресурсом у відображенні модернізаційних процесів доби. Як засвідчує сучасна наукова парадигма, «завдяки ефективному поєднанню технологічного та людського компонентів у лінгвістичному аналізі корпусна лінгвістика відкриває нові перспективи для мовознавчої науки» (Жуковська, 2020: 117).

Для кращої ілюстрації корпусних результатів подамо «діалог кластерів», де кожна сполука виступає як умовний «мовець» і презентує власну частоту та стилістичну функцію. Дані кластерного аналізу підтверджують провідну роль дієслова «спитав» у комунікативній організації прози М. Левицького (Левицький, 2016). Найчастотнішими виявилися сполуки «спитав він» (33 випадки) та «спитав я» (24 випадки), що свідчить про домінування двох основних наративних перспектив: з боку узагальненого персонажа та від імені оповідача. Їхня частка у загальній структурі корпусу є найбільшою (NormFreq 0.194 та 0.141 відповідно), а показник нормованого діапазону (NormRange 1.000) демонструє сталість цих конструкцій у всіх контекстах. Важливим є також виявлення поєднань із професійними чи соціальними маркерами, зокрема «спитав лікарь» (7 випадків), «спитав становий» (5 випадків), «спитав батько» (4 випадки). Це засвідчує тенденцію до чіткої соціальної конкретизації мовців, яка формує реалістичне багатоголосся тексту.

Менш численні кластери («спитав швець», «спитав соцький», «спитав суддя», «спитав священик», «спитав хлопець», «спитав циган» тощо) відображають соціальну різнорідність персонажів та їхні мовні практики. Навіть за одиничної частоти (1-2 випадки) вони зберігають важливе стилістичне навантаження, адже показують розширення спектра комунікації й підкреслюють етнічну та соціальну палітру доби. Отже, «спитав» виступає не лише як формальний маркер діалогу, а як ключова лексема для соціальної типізації та наративної організації тексту.

Особливу увагу привертає поєднання «спитав лікарь», яке демонструє перетин діалогічної й професійної лексики. Це підтверджує тезу, що медична тематика становить окремий пласт у творчості Левицького: лікарі та лікарні фігурують не лише як сюжетні деталі, а як символи соціальних інституцій та модернізаційних процесів. Кількісні дані корелюють із якісними спостереженнями: частотність «спитав» створює діалогічну динаміку, а поєднання з професійними та соціальними назвами конкретизує багатоголосся, формуючи реалістичну модель суспільної комунікації.

Тож, можна стверджувати, що корпусна статистика не лише підтверджує високу частотність дієслова «спитав», а й дозволяє виокремити його провідні семантико-стилістичні функції. Воно організовує діалогічний простір, структурує наратив, виступає маркером соціальної належності персонажів і водночас створює ефект автентичності живого мовлення. Отже, «спитав» разом із медичною термінологією формує стилістичне ядро текстів Левицького, де переплітаються побутові, службові та професійні дискурси, що унаочнюють ідіостиль письменника.

Отримані дані кластерного аналізу засвідчують, що медична лексика («лікарь», «лікар», «лікарня») формує окремий пласт стилістичного ядра прози М. Левицького. Найвищі показники частотності мають сполуки «лікарь а» та «лікарь і» (по 6 випадків), що свідчить про активне використання цієї лексеми у структурі діалогів та оповідних коментарів. Значущими є також кластери «лікарь знов» (4 випадки), «лікарь що», «лікарь як», «лікарь він» (по 3 випадки), які фіксують динаміку опису професійних дій, реакцій та мовлення персонажа-лікаря. Це дозволяє зробити висновок, що фігура лікаря у прозі Левицького має не лише сюжетну, а й структурно-організуючу роль: через неї вибудовуються діалоги, оцінюється стан хворих, втілюються критичні судження про соціальні умови. Лексема «лікарня» також посідає помітне місце у корпусі. Її кластери («лікарні у», «лікарні я», «лікарні а», «лікарні сказали», «лікарні хворий») демонструють сталість у позначенні простору хвороби й лікування, а також у відображенні соціальної проблематики. Частотність цих сполук хоч і нижча (переважно 1-2 випадки), проте семантично вони концентрують критику матеріальних умов, бюрократичних процедур та моральних дилем, пов'язаних із системою земських лікарень. Лексичні сполуки на кшталт «лікареві не», «лікарню тяжко», «лікарня місця» увиразнюють драматичне тло творів, акцентуючи увагу на безпорадності пацієнтів і недосконалості інституцій.

У семантико-стилістичному вимірі результати двох корпусних зрізів доповнюють один одного. Дієслово «спитав» (з найвищими показниками для кластерів «спитав він» – 33 випадки та «спитав я» – 24 випадки) формує комунікативний ритм, у той час як лексеми «лікарь / лікарня» конкретизують соціальний і професійний вимір діалогів. Завдяки аналізу можна виявити, «як мова функціонує в різних комунікативних контекстах, аналізуючи великі набори лінгвістичних даних» (Ворошилова, 2025: 6). Отримане співвідношення доводить, що Левицький вибудовує оповідь через поєднання діа-

логічної динаміки та професійної конкретики, що дозволяє досягти реалістичного відтворення багатоголосої соціальної дійсності. Корпусна статистика демонструє, що і дієслово «спитав», і медична лексика є взаємодоповнюючими компонентами стилістичного ядра прози письменника. Перше створює ефект живого мовлення, друге виводить на поверхню соціально значущі теми охорони здоров'я та модернізації інституцій. У результаті формується унікальна наративна стратегія, в якій через мовні маркери репрезентується структура суспільних відносин та культурний контекст доби (Луцак, Юрчак: 2023: 161).

Для поглибленого аналізу корпусу використано методику n-грам, що передбачає виявлення багатослівних сполук, у яких фіксується не лише частотність окремих лексем, а й закономірності їх поєднання в реальному мовному контексті. Такий підхід дозволяє простежити стійкі комунікативні формули, які відтворюють діалогічну динаміку та інституційну проблематику. Якщо частотний аналіз окремих слів лише вказує на домінування дієслів чи професійних термінів, то аналіз n-грам відкриває глибинний рівень взаємодії – показує, у яких комунікативних середовищах вони функціонують, як пов'язані із соціальними ролями персонажів і які культурні смисли закодовують.

Отже, корпусна статистика демонструє, що і дієслово «спитав», і медична лексика є взаємодоповнюючими компонентами стилістичного ядра прози письменника. Перше створює ефект живого мовлення, друге виводить на поверхню соціально значущі теми охорони здоров'я та модернізації інституцій. У результаті формується унікальна наративна стратегія, в якій через мовні маркери репрезентується структура суспільних відносин та культурний контекст доби. У корпусі М. Левицького діалогічні дієслова та професійна лексика функціонують у тісному взаємозв'язку: перші забезпечують поліфонічність і автентичність мовлення, другі – підкреслюють інституційні виклики та соціальну проблематику доби. У результаті формується подвійне стилістичне ядро, де приватний голос персонажа співіснує з публічною критикою соціальних практик.

Висновки n-грамного аналізу дають підстави стверджувати, що стилістика Левицького вирізняється двома взаємодоповнюючими осями: горизонтальною, яка відтворює живий діалог через численні формули із дієсловом «спитав», та вертикальною, що окреслює професійно-інституційний зріз через медичну лексику. Разом вони формують цілісну картину суспільного буття, у якій приватне й публічне переплітаються, а мова стає каналом вираження соціальної пам'яті, культурних цінностей і модернізаційних процесів.

Додаткову глибину дослідженню надає аналіз колокацій, який дозволяє виявити найближчих лексичних сусідів ключових слів та простежити закономірності їхнього функціонування у контексті. Для дієслова «спитав» найбільш частотними є поєднання з займенником «він» (що підтверджує його роль як універсального маркера діалогічності), а також із лексемами, що позначають соціальні ролі («панотець», «никодим», «шпак»). Показовими є й комбінації з питальними словами («чого», «думаєте»), які акцентують функцію дієслова як засобу ініціації комунікації. У такий спосіб «спитав» постає не лише як нейтральне дієслово, а як структуроутворюючий елемент наративу, що активізує поліфонію тексту.

Для лексеми «лікар» колокаційний аналіз засвідчив її тісний зв'язок із професійною та інституційною сферою. Поруч із нею з'являються дієслова мовлення («мовив», «казав»), що підкреслюють комунікативну роль лікаря як авторитетного співрозмовника. Водночас у контекстах простежується соціально-критичне навантаження: слова «трують», «земська», «розбудили» репрезентують як побутову практику лікування, так і інституційну реальність земських лікарень. Особливу увагу заслуговують випадки, де лікар асоціюється з емоційними реакціями («жалкую»), що виводить його образ за межі професійної функції й наділяє індивідуальними психологічними рисами.

Отже, колокаційний аналіз підтвердив і розширив результати частотних та n-грамних спостережень: «спитав» функціонує як універсальний маркер діалогу, що інтегрує різні соціальні голоси й створює поліфонічну тканину тексту; «лікар» же виступає не тільки як номінатор професії, а й як символ інституційного досвіду та культурних трансформацій доби. Поєднання цих двох полів показує, що корпусна стилістика Левицького ґрунтується на перехрещенні приватної комунікації та публічної інституційності, завдяки чому його проза віддзеркалює і щоденне життя, і проблематику модернізації суспільних практик.

Висновки. Проведене дослідження корпусу текстів Модеста Левицького дало змогу систематизувати його мовну спадщину та проаналізувати її з позицій сучасної корпусної лінгвістики. Створений масив обсягом понад 111 тисяч слововживань і понад 21 тисячу унікальних словоформ засвідчив багатство лексики та стилістичну варіативність письма автора. Статистичні параметри корпусу – щільність словникового запасу, індекс читабельності, середня довжина речення – під-

твердили поєднання доступності викладу з художньою глибиною й композиційною цілісністю його прози. Частотний аналіз виявив домінування діалогічних дієслів, соціально-побутової й професійної лексики, а також емоційно забарвлених конструкцій, що формують індивідуальний мовний портрет письменника. Дослідження сталих словосполучень показало значущість антропонімів, адміністративної та медичної термінології, а також експресивних діалогічних структур, які створюють реалістичний мовний колорит. Конкордансний та кластерний аналіз ключових лексем – *«спитав»* і *«лікар»* – продемонстрував подвійне стилістичне ядро прози Левицького: з одного боку, поліфонічність живого мовлення, з іншого – репрезентацію професійного досвіду та інституційних проблем доби.

Практична цінність корпусу полягає у його багатофункціональності. Він може бути використаний для вивчення архаїзмів, історизмів, професійної та медичної термінології, авторських новотворів, що збагачує не лише мовознавчі й літературознавчі студії, а й суміжні галузі знань. Корпус відкриває перспективи для лексикографічних розробок, є ефективним ресурсом для освітнього процесу й водночас слугує базою для міждисциплінарних досліджень на перетині мови, літератури, історії та медицини. Створений корпус текстів М. Левицького є вагомим внеском у розвиток української корпусної лінгвістики. Він не лише репрезентує мовні та культурні особливості доби, а й демонструє значний потенціал для подальших студій, спрямованих на поглиблене розуміння ідіостилю письменника й соціокультурного контексту його епохи.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Ворошилова Р. Корпусна прагматика – нова галузь корпусної лінгвістики. *Наукові записки Національного університету «Острозька академія»: серія «Філологія»*. 2025. Вип. 24(92). С. 3-7. URL: https://www.journals.oa.edu.ua/Philology/article/view/4248.
2. Голощук С. Корпусна лінгвістика: сучасний стан та перспективи досліджень. *Закарпатські філологічні студії*. 2022. Вип. 21, т. 1. С. 249-252. DOI: https://doi.org/10.32782/tps2663-4880/2022.21.1.47.
3. Демська О. Текстовий корпус: поняття і визначення. *Дивослово.* 2011. № 10. С. 35-37. URL: https://dyvoslovo.com.ua/wp-content/uploads/2016/03/7-1011.pdf
4. Жуковська В. Лінгвістичний корпус як новітній інформаційно-дослідницький інструментарій сучасного мовознавства. *Вчені записки Таврійського національного університету імені В. І. Вернадського. Серія: Філологія. Соціальні комунікації*. 2020. Т. 31 (70), № 3, ч. 1. С. 113-119. DOI: https://doi.org/10.32838/2663-6069/2020.3-1/20.
5. Левицький М. Вибране / упоряд. Н. Пушкар, О. Яблонська, Т. Данилюк-Терещук; Луцьк: Терези, 2016. 462 с. URL: https://shron1.chtyvo.org.ua/Levytskyi_Modest/Vybrane.pdf
6. Луцак С., Юрчак Г. Тенденції розвитку прикладної лінгвістики в контексті розвитку наукових галузей. Digital transformations in culture: Scientific monograph. Riga, Latvia: «Baltija Publishing», 2023. P. 255-268. DOI https://doi.org/10.30525/978-9934-26-319-4-15.
7. Луцак С., Юрчак Г. Теоретичні основи корпусної лінгвістики. *Прикарпатський вісник Наукового товариства імені Шевченка «Слово»,* 2025. (21 (77), с. 33-48. https://doi.org/10.31471/2304-7402-2025-21(77)-33-48.
8. Степанов В., Могильник А. Потенціал корпусних технологій у вивченні фразеології та сленгу. *Філологічні трактати*. 2025. Т. 17, № 1. С. 63-77. DOI: https://doi.org/10.21272/Ftrk.2025.17(1)-6.
9. Плюта Н. Експериментальне дослідження впливу способу формування корпусної вибірки на репрезентативність результатів. *Актуальні питання гуманітарних наук*. 2025. Вип. 88, т. 2. С. 214-220. DOI: https://doi.org/10.24919/2308-4863/88-2-29.
10. Таран А. Пошукова система й корпус як інструмент дослідження семантики слів. *Вісник Черкаського національного університету імені Богдана Хмельницького. Серія : Філологічні науки*. 2019. Вип. 1. С. 99-103. DOI: https://doi.org/10.31651/2076-5770-2019-1-99-103
11. Bashir I., Yunus K., Ibrahim B. Perspectives on Corpus Linguistics: The Methodological Synergy in Second Language Pedagogy and Research. *Arab World English Journal*. 2018. Vol. 9, No. 3. P. 84-97. DOI: https://dx.doi.org/10.24093/awej/vol9no3.6.
12. Stetsenko D., Okulska I. The Grammar and Syntax Based Corpus Analysis Tool For The Ukrainian Language. *arXiv preprint*. 2023. arXiv:2305.13530 [cs.CL]. DOI: https://doi.org/10.48550/arXiv.2305.13530.
13. Walkowiak T. Computer Based Stylometric Analysis of Texts in Ukrainian Language. Artificial Intelligence and Soft Computing. Springer International Publishing, 2020. P. 233-243. DOI: https://doi.org/10.1007/978-3-030-61534-5_20.

REFERENCES

1. Voroshylova R. (2025). Korpusna prahmatyka – nova haluz korpusnoi linhvistyky [Corpus pragmatics as a new branch of corpus linguistics]. Naukovi zapysky Natsionalnoho universytetu «Ostrozka akademiia»: seriia «Filolohiia». 24(92). 3-7. [in Ukrainian].
2. Holoshchuk S. (2022). Korpusna linhvistyka: suchasnyi stan ta perspektyvy doslidzhen [Corpus linguistics: Current state and research prospects]. Zakarpatski filolohichni studii. 21(1). 249-252. [in Ukrainian].
3. Demska O. (2011). Tekstovyi korpus: poniattia i vyznachennia [Text corpus: Concept and definition]. Dyvoslovo, 10. 35-37. [in Ukrainian].

4. Zhukovska V. (2020). Linhvistychnyi korpus yak novitnii informatsiino-doslidnytskyi instrumentarii suchasnoho movoznavstva [Linguistic corpus as a modern informational and research tool of contemporary linguistics]. Vcheni zapysky Tavriiskoho natsionalnoho universytetu imeni V. I. Vernadskoho. Seriia: Filolohiia. Sotsialni komunikatsii. 31(70) (3, Part 1). 113-119. [in Ukrainian].

5. Levytskyi M. (2016). Vybrane [Selected works] / uporiad. N. Pushkar, O. Yablonska, T. Danyliuk-Tereshchuk. Lutsk: Terezy. 462. [in Ukrainian].

6. Lutsak S., Yurchak H. (2023). Tendentsii rozvytku prykladnoi linhvistyky v konteksti rozvytku naukovykh haluzei [Trends in the development of applied linguistics in the context of the development of scientific fields]. Digital transformations in culture: Scientific monograph. 255-268. Riga, Latvia: Baltija Publishing. [in Ukrainian].

7. Lutsak S., Yurchak H. (2025). Teoretychni osnovy korpusnoi linhvistyky [Theoretical foundations of corpus linguistics]. Prykarpatskyi visnyk Naukovoho tovarystva imeni Shevchenka "Slovo". 21(77). 33-48. [in Ukrainian].

8. Stepanov V., Mohylnyk A. (2025). Potentsial korpusnykh tekhnolohii u vyvchenni frazeolohii ta slenhu [The potential of corpus technologies in the study of phraseology and slang]. *Filolohichni traktaty.* 17(1). 63-77. [in Ukrainian].

9. Pliuta N. (2025). Eksperymentalne doslidzhennia vplyvu sposobu formuvannia korpusnoi vybirk y na reprezentatyvnist rezultativ [An experimental study of the influence of corpus sampling method on the representativeness of results]. *Aktualni pytannia humanitarnykh nauk.* 88(2). 214-220. [in Ukrainian].

10. Taran A. (2019). Poshukova systema y korpus yak instrument doslidzhennia semantyky sliv [Search engine and corpus as tools for studying word semantics]. Visnyk Cherkaskoho natsionalnoho universytetu imeni Bohdana Khmelnytskoho. Seriia: Filolohichni nauky. 1. 99-103. [in Ukrainian].

11. Bashir I., Yunus K., Ibrahim B. (2018). Perspectives on corpus linguistics: The methodological synergy in second language pedagogy and research. Arab World English Journal, 9(3). 84-97. .

12. Stetsenko D., Okulska, I. (2023). The grammar and syntax based corpus analysis tool for the Ukrainian language. *arXiv preprint*. arXiv:2305.13530 [cs.CL]. .

13. Walkowiak T. (2020). Computer based stylometric analysis of texts in Ukrainian language. Artificial Intelligence and Soft Computing. Springer International Publishing. 233-243. .

Дата першого надходження статті до видання: 30.04.2026
Дата прийняття статті до друку після рецензування: 22.05.2026
Дата публікації (оприлюднення) статті: 30.05.2026

Стаття поширюється на умовах
ліцензії відкритого доступу (CC BY 4.0)