ISSN 2311-259X

https://doi.org/10.28925/2311-259x.2026.3.9
УДК 811.161.2'33'37

Наталія Дарчук
Київський національний університет імені Тараса Шевченка,
бульвар Тараса Шевченка, 14, м. Київ, 01601, Україна
https://orcid.org/0000-0001-8932-9301
n.darchuk@knu.ua

Оксана Зубань
Київський національний університет імені Тараса Шевченка,
бульвар Тараса Шевченка, 14, м. Київ, 01601, Україна
https://orcid.org/0000-0002-2644-3892
oxana.zuban@knu.ua

Валентина Робейко
Київський національний університет імені Тараса Шевченка,
бульвар Тараса Шевченка, 14, м. Київ, 01601, Україна
https://orcid.org/0000-0003-2266-7650
valentyna.robeiko@knu.ua

Юлія Цигвінцева
Інститут української мови НАН України
вул. Михайла Грушевського, 4, Київ, 01001, Україна
https://orcid.org/0000-0002-9684-3840
julivoznuk@gmail.com

ОЦІННИЙ ПОТЕНЦІАЛ ФРАЗЕОЛОГІЗМІВ ТА АВТОМАТИЗАЦІЯ ЙОГО ВИЗНАЧЕННЯ В МЕДІЙНОМУ ТЕКСТІ

В умовах зростання обсягів медійного контенту особливої актуальності набуває проблема автоматизованого розпізнавання тональності тексту, зокрема тих її складників, які реалізуються через ідіоматику. Наукова проблема полягає в браку тонально розміченої бази даних фразеологізмів, яка б об'єктивно відображала реальне сприйняття оцінного компонента фразеологічних одиниць носіями мови й могла слугувати підґрунтям для обчислення індексу тональності в автоматизованих системах аналізу тексту. Предмет дослідження – тональна розмітка фразеологічних одиниць шляхом краудсорсингу. Мета – визначити оцінний потенціал фразем у медійних українськомовних текстах на основі краудсорсингового анотування та сформувати тонально розмічену базу даних фразеологізмів. Методи: краудсорсинговий експеримент, дистрибутивний, компонентний, кількісно-статистичний аналіз. Емпіричною базою слугувала фразеологічна база даних (20 985 одиниць), укладена за матеріалами українських газетних і мережевих видань 2000–2024 рр. із Корпусу текстів української мови. Уперше апробовано методику тональної розмітки значного масиву фразеологічних одиниць шляхом краудсорсингу, що завдяки залученню групи незалежних експертів-філологів та подальшому обчисленню середнього арифметичного їхніх оцінок уможливило нівелювання суб'єктивності індивідуального оцінювання й забезпечило статистично обґрунтовану квантифікацію оцінного компонента фразеологічної семантики. Результати: висвітлено хід, висліди й проблеми краудсорсингового підходу до визначення оцінного компонента фразеологічної семантики; установлено домінування негативно конотованих одиниць у створеній базі даних; виявлено, що ступінь інтенсифікації оцінного компонента корелює з функціонально-стильовою належністю фраземи; підтверджено ефективність краудсорсингу як оптимального методу створення тонально розміченої бази даних фразеологізмів. Перспективи: розширення бази даних неофразеологізмами і фраземами, уживаними в інших функціональних стилях; поглиблений лінгвістичний аналіз бази даних та усунення суперечностей оцінювання; імплементація результатів у модуль визначення тональності системи комп'ютерного аналізу українськомовного тексту.

Ключові слова: українська мова; медійний текст; оцінка; тональність; фразеологічна одиниця; база даних фразеологізмів; краудсорсинг; тональна розмітка.

Вступ. Медійний простір завжди був чи не основною сферою формування суспільної думки, громадських настроїв та ціннісних орієнтирів авдиторії. Медіа перебувають у постійному пошуку ефективних вербальних способів та засобів увиразнення мовлення, які б привертали увагу споживача інформації. Серед них особливе місце посідають фразеологічні одиниці (*далі* —

© Н. Дарчук, О. Зубань, В. Робейко, Ю. Цигвінцева, 2026

Ліцензовано за CC BY 4.0

ФО, фраземи) — стійкі сполуки, власне фразеологізми, прислів'я, приказки, крилаті вислови, що акумулюють досвід культури, світогляд, ментальність конкретної групи мовців і мають значні впливові можливості, адже використовуються зі сформованими та відтворюваними в мовній практиці оцінними значеннями.

В. Ужченко та Д. Ужченко слушно зауважували, що «виникнення й уживання фразеологізму викликане постійним відчуттям лексичної недостатності, намаганням вербалізувати людські емоції, утілені в когнітивні моделі душевного стану» (Ужченко, & Ужченко, 2005, с. 6). Саме тому спостерігаємо таку прискіпливу увагу науковців до емоційного й оцінного потенціалу ФО, їхньої здатності репрезентувати позитивний чи негативний образ світу.

Категорія оцінки є одним із центральних понять сучасної лінгвістики. Як зазначає І. Онищенко, оцінка — це «складний феномен для наукового опису і теоретичного пояснення, що зумовлено онтологічною універсальністю категорії оцінки, адже будь-яка онтологія: природа, суспільство, суб'єктивний світ людини, мислення, мова тощо — усе пронизане категорією оцінки і розглядається через її призму» (Онищенко, 2004, с. 5). Сучасні українські лінгвістичні розвідки присвячені вивченню різних аспектів оцінки: засобів її вираження в українській мові (Космеда, 2001), зокрема й у публіцистичних та інформаційних текстах (Онищенко, 2004); функціонально-семантичної природи (Чобот, 2002); ядра та периферії функціонально-семантичного поля оцінки (Онищенко, 2011); суб'єктивної та об'єктивної оцінки (Степаненко, & Дейна, 2018); аксіологічних можливостей граматичних одиниць (Халіман, 2019); оцінного компонента лексем як основи створення системи автоматичного аналізу тональності (Дарчук, 2019); функціонально-стилістичних параметрів оцінки як лінгвістичної категорії (Коць, 2021); мовної концептуалізації аксіологічного простору людини (Деменчук, 2023) тощо.

Методологічним підґрунтям для розуміння механізмів образності та оцінності фразеологічного фонду стало дослідження В. Ужченка та Д. Ужченка «Фразеологія сучасної української мови» (2005), автори якого зазначають: «через фразеологію проходить не все людське життя, а тільки ті його сторони й грані, які вторинно відображаються через випукле дзеркало людських почуттів, сприймань і їх оцінок» (Ужченко, & Ужченко, 2005, с. 24). Особливу увагу привертають і праці про відображення категорії оцінки у фразеології (Краснобаєва-Чорна, 2015; Ганжа, & Онищенко, 2016), психолінгвістичний аспект мовленнєвого впливу через фразеологію (Зирка, 2004), оцінність та маніпулятивну природу фразеологізму (Сизонов, 2018; Сизонов, 2019) та ін. Наприклад, В. Зірка дійшла висновку, що «фразеологізм є знаряддям мовленнєвого впливу, адже за законом економії мовленнєвих зусиль він повідомляє адресату більше інформації, ніж нейтральні висловлювання» (Зирка, 2004, с. 126).

Серед закордонних досліджень вирізняємо праці про аксіологічну функцію фразеологізмів у масмедійному дискурсі. Зокрема, британська дослідниця Р. Амосси підкреслює, що сталі вирази є одним із найдієвіших риторичних інструментів, оскільки вони апелюють до колективної пам'яті аудиторії та її ціннісної системи, обходячи критичне сприйняття (Amossy, 2006). Польський дослідник Є. Бральчик наголошує, що стійкі вирази в медіатексті виконують функцію «ярликів» — спрощених оцінних суджень, які не потребують аргументації і тому є особливо ефективними у формуванні масових установок (Bralczyk, 2004).

Попри пильну увагу до оцінки як лінгвістичної категорії та докладність опису оцінного значення, з позицій сучасних завдань лінгвістичного аналізу ці проблеми остаточно не розв'язані. Співвідносячи оцінку — «закріплену в семантиці слова позитивну чи негативну кваліфікацію об'єкта за ознакою «добре/погано» з боку мовця-суб'єкта» (Краснобаєва-Чорна, 2015, с. 40) з тональністю, тобто позитивним, негативним або нейтральним відношення автора до об'єктів, явищ, осіб, згаданих у ньому, ми фокусуємо увагу на можливості автоматизації її визначення, адже включення фразеологізмів як окремих одиниць аналізу суттєво покращує результати автоматизованого визначення тональності тексту (Williams, Bannister, Arribas-Ayllon, Preece, & Spasić, 2015).

Мета нашого дослідження – визначити оцінний потенціал фразем у медійних українськомовних текстах на основі краудсорсингового анотування та сформувати тонально розмічену базу даних ФО сучасної української мови.

Досягнення поставленої мети передбачає розв'язання таких дослідницьких **завдань**: 1) з'ясувати співвіднесеність логіко-філософського та власне мовного аспектів категорії оцінки; 2) схарактеризувати базу даних фразеологізмів як основу для дослідження їхнього оцінного потенціалу; 3) застосувати метод краудсорсингу для обчислення оцінки фразем сучасної української мови, описати його хід та результати; 4) окреслити перспективи імплементування результатів експерименту в модуль визначення тональності системи комп'ютерного аналізу українськомовного тексту.

Об'єктом дослідження є ФО як засоби вираження оцінки в медійних текстах. **Предмет дослідження** – тональна розмітка фразеологічних одиниць шляхом краудсорсингу.

Джерельною базою слугували матеріали українських газет («Літературна Україна», «Факти», «Дзеркало тижня», «Економічна правда», «Тиждень» тощо), а також публікації в мережевих медіа («Київ», «Саме той») за 2000–2024 рр., розміщені в Корпусі текстів української мови (далі – КУМ) на платформі Mova.info (КУМ, 2003–2026).

Методи дослідження визначено характером поставлених завдань. Особливості фразеологічного вираження оцінки в медійних текстах встановлено шляхом дистрибутивного й компонентного аналізу, а також контекстно-ситуативного, класифікаційного та кількісно-статистичного методів. Основним для тональної розмітки бази даних фразем став метод краудсорсингу, що передбачав залучення групи людей для виконання завдань організаторів експерименту.

Виклад основного матеріалу. У нашому дослідженні оцінку розглядаємо як функціонально-семантичну категорію, яка є своєрідною аксіологічною константою і відображає позитивне чи негативне ставлення мовця до висловленого, зокрема й через уживання ФО. Фразеологічне вираження емоцій вважаємо не менш важливим, ніж лексичну номінацію, адже фразеологізм одночасно повідомляє про предмет, репрезентує думку мовця про повідомлюване, а також транслює емоційне ставлення до нього. Оцінний компонент у ФО може бути закріплений у денотативній (раціональна оцінка: *скласти подяку, доводити до абсурду*) або конотативній частині семантики (емоційна оцінка: *підкласти свиню, скакати в гречку, кінь не валявся*). Особливу роль у медійній комунікації відіграють прислів'я та приказки (*Ранні пташки росу п'ють, а пізні — слізки ллють, Волос сивіє, а голова шаліє, Далі в ліс — більше дров*), що акумулюють багатовіковий суспільний досвід. До кола оцінних засобів належать також фразеологізовані порівняльні звороти (*як бугай, як скельце, як баран (теля) на нові ворота, як церковна миша*), інші сполучення фразеологізованого типу (*як не крути, як не верти, хоч у воду*) та словосполуки, що лише фразеологізуються (*вистрибнути на коня, ловити мишей у темній хаті*), зокрема й різного роду неофразеологізми (*відправити вслід за російським кораблем, диванний експерт, ядерний шантаж, тепла ванна, епоха бідності*) тощо.

Аналізовані тексти з КУМ підтверджують, що в медійному мовленні функціонують ФО традиційні, усталені за структурою та семантикою, і модифіковані, трансформовані. Аналіз підкорпусу публіцистичних текстів показав переважання традиційних фразеологізмів. Вони створюють додаткову динаміку в сприйнятті будь-якого факту, загострюють увагу читача, мають здатність посилювати чи послаблювати ступінь вияву ознаки або інтенсивність протікання дії. У публіцистичних текстах домінують фразеологізми оцінного характеру, яким притаманна висока експресивність, напр.: *Олександр Ткаченко, відповідаючи на журналістські запитання, не обминув можливості* ***підкласти свиню*** *й Президентові Кучмі* (Україна молода, 4.02.1999, с. 3); ***Поляпають язиками****, і розповзаються по хатах, а почуття обов'язку в українців, на жаль, атрофоване майже* (Літературна Україна, 4.12.2001, с. 3). Досить часто автори вживають і ФО неоцінного характеру (*І після того, як приватизація ФК „Динамо" (Київ), про яку весь час говорили зовсім не більшовики, здійснювалася —* ***на повен зріст*** *постали нові реалії нової реформаторської доби* (Київ, 2002, №3, с. 158); *Ялта мала* ***розставити всі крапки над і*** (Факти. Підсумки дня, 29.01.2004)) та фразеологізми, які містять оцінку в денотаті (*Своєю діяльністю вони* (партії — І.О.) ***завдають клопоту*** *не тільки Верховній Раді, а й окремим людям* (Саме той, 24.04.2002); *Ющенківці, на мій погляд,* ***подають надії*** (Саме той, 12.05.2002)).

Різні типи фразеологізмів у медійних текстах характеризуються різним ступенем інтенсифікації оцінного компонента значень: розмовно-побутові ФО виконують яскраво виражену оцінну функцію (*сидіти на шиї, напускати туману*) та становлять домінанту фразеологічних засобів оцінки: *...Більшість населення перетворили в рабів, яких не було ніде по світу, а купка багатіїв* ***сіла*** *їм* ***на шию*** (Літературна Україна, 22.01.2004, с. 3); книжні мають середній ступінь оцінності (*сізіфова праця*, *олімпійський спокій*) та низький показник частотності: *А за ту* ***сізіфову працю,*** *окрім шишаків, нічого не отримаєш* (Літературне Придніпров'я, 2003, №1, с. 3); фольклорні та народнопоетичні (*Сказав пан, кожух дам. Тепле його слово, хоч з лиця воду пий*) утворюють крайню периферію оцінних засобів та виконують слабко виражену оцінну функцію: *Якщо добробут українського села триматиметься на чесному слові Азарова і Шуфрича, то справи кепські.* ***„Сказав пан, кожух дам. Тепле його слово..."*** (Нація і держава, 2.12.2003, №8, с. 2).

Функціонування ФО у трансформованому вигляді сприяє появі додаткового оцінного компонента в їхній семантиці або збільшенню ступеня інтенсифікації оцінки, раніше закріпленої за одиницею. Дослідження показало, що трансформація фразем відбувається за допомогою таких прийомів: а) поширення: *Це така собі* ***локшина на європейські вуха*** (Нація і держава, 16.12.2003, №10, с. 3) (пор.: *локшина на вуха → локшина на європейські вуха***)**; б) скорочення: *Повторюю, неточності й помилки можуть статись, але щоб створити суцільне звалище на тему "****чув дзвін****" — це все-таки треба мати особливий талант* (Літературна Україна,

19.02.2004, с. 8) (пор.: *чути дзвін, та не знати, де він → чути дзвін*); в) субституції: *Але **якщо взявся за гуж, то куди дінешся*** (Літературна Україна, 29.01.2004, с. 4) (пор.: *якщо взявся за гуж, то не кажи, що не дуж → якщо взявся за гуж, то куди дінешся*); г) фразеологічного натяку: *Хто винен? Коні, як відомо, поза докором і підозрами, вони ні при чому* (Київ, 2001, №9-10, с. 113) (пор.: *коні не винні*); д) контамінації: *Але ж, мабуть, і вони повинні підпорядковуватися вищій меті, спершу спільно підживити державу, **поставити з колін на тверді ноги**, а вже потім змагатися за право тримати в своїх руках кермо* (Київ, 1995, №6, с. 6) (пор.: *поставити з колін на (тверді) ноги ← поставити з голови на ноги + піднятися з колін*). Трансформовані фразеологізми (у публіцистичних текстах становлять 19% від загальної кількості фразеологічних засобів, у мережевих — 11%) є важливим джерелом експресивно-емоційної насиченості, мають яскраве прагматико-стилістичне спрямування — використовуються з метою викликати в адресата нові емоції, формування нових оцінних суджень, несподіваних асоціації тощо.

Для реалізації завдань було укладено базу даних фразем, що налічує 20 985 записів (рис. 1). За наявності у складі фраземи хоча б одного слова зі змінюваною граматичною формою, їй присвоювали маркер #, розміщений у препозиції, напр., *в одній упряжці #бути*. Якщо ФО мала варіанти, у яких було змінено порядок слів або вжито синонім — кожний варіант становив новий рядок у базі даних, напр., *#бути в одній упряжці* і *в одній упряжці #бути, голову #втратити* і *розум #утратити, фокуси #викидати* і *#викидати штуку, у свинячий голос* і *у чортів голос.* До кожної фраземи подано тлумачення, сформульоване укладачами за реєстром фразеологічних словників, зокрема «Фразеологічного словника української мови» у 2 книгах (Білоноженко, Винник, Гнатюк та ін., 1993) та словника «Нові слова та фразеологізми в медіа» (Шевченко, & Сизонов, 2017–2022), яке було основою для визначення тональності.

id	sortname	mainword	tlum
20	манна небесна	1	щось несподіване і дуже бажане, що приходить у момент гострої потреби а
21	наколоті апельсини	2	неправдива інформація
22	отримувати на горіхи	1	Хто-небудь буде покараний за якусь провину.
23	перший хлопець на селі	2	найкращий або найвідоміший чоловік у спільноті, зазвичай сільській.
24	повний абзац	2	Вираження захвату від когось, чогось або розчарування від невдачі, краху.
25	поливати брудом	1	несправедливо звинувачувати когось у чомусь; обмовляти, неславити, ганьб
27	політичний колобок	2	Зеленський - це політичний Колобок. Дивіться: - від баби (ЄС) - втік; - від дід
29	помаранчева "революція"	2	значні соціально-політичні зміни, які супроводжуються масовими протестам
30	помаранчевий" електорат	2	виборці, які підтримували політичні сили, що брали участь у Помаранчевій
31	"помаранчеві" сили	2	політичні сили, пов'язані з "Помаранчевою революцією" в Україні
32	права "рука"	2	найближчий помічник і соратник
33	"правити бал"	1	мати повну владу або контролювати ситуацію.
34	притча во язицех"	1	Те (рідше той), про що (про кого) всі говорять; предмет розмов.
35	притягнутий за вуха"	1	Використовувати що-небудь невиправдано, випадково або без достатніх під
36	п'ята колона"	2	те, що зайве; у політичній комунікації – ті, хто працює на знищення держави
37	робити із мухи слона"	1	перебільшувати; надавати чому-небудь занадто велике значення
38	рука Москви"	1	Про вплив та втручання російської влади у внутрішні справи інших країн.
39	свій #хлопець"	2	Близький за переконаннями, поглядами, настроями і т.ін.
40	святая святих"	1	Місце, відкрите для небагатьох.
42	сірий #кардинал"	2	Впливова людина, яка діє негласно і зазвичай не обіймає формальних посад
43	скелет у шафі"	1	Душевні таємниці, особисті секрети людини, ретельно приховувані від стор

Рис. 1 Фрагмент бази даних фразем

Матеріалом для проведення краудсорсингового експерименту слугувала фразеологічна база даних у повному її обсязі. Анотаторами стали викладачі й студенти Навчально-наукового інституту філології Київського національного університету імені Тараса Шевченка — усього 18 осіб, котрі оцінювали кожну фразему в діапазоні від **-2** до **+2** (дуже негативна, помірно негативна, нейтральна, помірно позитивна і дуже позитивна). Оцінки виставляли анонімно, а потім автори автоматично підрахували середнє арифметичне оцінок і кожній фраземі присвоїли відповідну вагу тональності.

За умовами краудсорсингу до уваги брали оцінний компонент, репрезентований і на денотативному, і на конотативному рівнях семантичної структури фразеологічних одиниць. Це зумовлено існуванням інтелектуально-логічної (раціональної) та емоційної оцінок. Взаємодія поняттєвого ядра з логічною оцінкою предметно-логічної частини семантики формує емоційну оцінку, яка відбиває емоційне ставлення суб'єкта до об'єкта. Оцінний компонент не виступає постійним для певної мовної одиниці. Це наочно демонструє табл. 1, у якій у першій колонці подано оцінку, тобто усереднене значення оцінок респондентів, у другій — кількість фразем, які отримали в результаті краудсорсингу таке значення, у третій наведено приклад ФО. Усереднене значення та його розкид — це результат сприйняття контекстуальної реалізації значення фраземи кожним індивідом зокрема.

ISSN 2311-259X

Табл. 1 Розподіл фразеологічних одиниць за усередненим показником оцінного компонента семантики (за результатами краудсорсингу)

Оцінка	К-сть	Приклад фраземи	Оцінка	К-сть	Приклад фраземи
2	30	*жива легенда*	-0,1	1	*аж кишить*
1,89	7	*гори зрушити*	-0,11	199	*бити на сполох*
1,88	7	*душа співає*	-0,13	68	*зіткнутися лицем в лице*
1,86	45	*зоряна година*	-0,14	317	*треба вже і честь знати*
1,83	26	*пишна красуня*	-0,17	145	*нагострити кігті*
1,78	18	*блиснути знаннями*	-0,2	14	*напхати грошима кишеню*
1,75	19	*жити, як у Бога за пазухою*	-0,22	177	*братися під сумнів*
1,71	68	*перевернути гори*	-0,25	47	*душа міру знає*
1,67	74	*вінчати лаврами*	-0,27	1	*«любі друзі»*
1,63	35	*голова на плечах*	-0,29	245	*тримати на достатній віддалі*
1,6	3	*піднімати дух*	-0,3	1	*а кіт ковбаску уминає, неначе й не до нього річ*
1,57	110	*розправляти крила*	-0,33	332	*закривати душу на замок*
1,56	74	*відданий друг*	-0,38	60	*злоба дня*
1,5	166	*набути ваги*	-0,4	8	*наче крізь сон пам'ятати*
1,44	98	*заживати слави*	-0,43	274	*язик до Києва доведе*
1,43	183	*поставити на твердий ґрунт*	-0,44	196	*два на третього не ждуть*
1,4	13	*аж надриватися із сміху*	-0,5	325	*який їхав, таку й стрів*
1,38	53	*знання — то сила*	-0,55	1	*«зелена» чорна скринька*
1,33	211	*добрий геній*	-0,56	207	*бути в Зе-шоколаді*
1,29	263	*рвати боки зо сміху*	-0,57	310	*тобі який клопіт*
1,27	2	*«зелений» бізнес*	-0,6	19	*мов грім з чистого неба*
1,25	74	*жива енциклопедія*	-0,63	60	*добуватись у відчинені двері*
1,22	139	*висловити свою пошану*	-0,67	442	*з двох бід вибрати меншу*
1,17	138	*наставляти на шлях істини*	-0,71	407	*усмішка авгура*
1,14	246	*іти своїм шляхом*	-0,75	76	*знай кобило, де брикати*
1,13	70	*заглибитися з головою*	-0,78	288	*диванний експерт*
1,11	140	*віддавати перевагу*	-0,8	31	*мов мильна бульбашка*
1,1	2	*вдихнули нове життя*	-0,83	268	*мов горохом об стіну*
1,09	1	*«коронований» вебінар*	-0,86	544	*ховати погляд*
1	706	*іти на лад*	-0,88	89	*за двома зайцями бігати*
0,91	6	*акції пішли вгору*	-0,89	292	*звалити в одну купу*
0,9	6	*аж гай гуде*	-1	1445	*насилу ноги тягти*
0,89	184	*бувала людина*	-1,11	359	*бити байдики*
0,88	62	*за безцінок купити*	-1,13	101	*утратити владу над собою*

0,86	303	*широке поле діяльності*	-1,14	529	*стінка на стінку*
0,83	158	*можна і треба; маленький українець*	-1,17	279	*мов загнаний звір*
0,8	12	*підвести голову*	-1,2	19	*манія переслідування*
0,78	165	*давати серцю волю*	-1,22	373	*не шукай горя — воно само знайде*
0,75	56	*за всіма правилами мистецтва*	-1,25	91	*утрачати владу над собою*
0,71	377	*приймати вогонь на себе*	-1,29	446	*тіньова пандемія*
0,67	308	*ділом довести*	-1,3	7	*вдарити по кишенях*
0,65	1	*«на повну котушку»*	-1,33	434	*наче ножем по серцю*
0,63	53	*дотримуватися букви закону*	-1,38	112	*жовта преса*
0,6	9	*місце під політично-бізнесовим сонцем*	-1,4	15	*підливати масла у вогонь*
0,57	312	*прошу любити і шанувати*	-1,43	297	*трясця тобі в бік*
0,56	187	*вести свою лінію*	-1,44	265	*безпросипний п'яниця*
0,5	291	*мовити словечко*	-1,5	396	*несосвітенна глупота*
0,44	184	*без пострілу взяти*	-1,55	1	*«зелене» політичне кумівство*
0,43	317	*правда-матінка*	-1,56	250	*немов з хреста знятий*
0,4	15	*на власні вуха почути*	-1,57	266	*шибениця плаче*
0,38	79	*заглядати у майбутнє*	-1,6	5	*пити гірку чашу*
0,33	370	*за сімома замками*	-1,63	98	*несосвітенна дурість*
0,3	6	*або пан, або пропав*	-1,67	222	*зайти в безвихідь*
0,29	361	*риба шукає, де глибше, а людина, де ліпше*	-1,71	118	*проливати крокодилові сльози*
0,25	87	*замовляти слівце*	-1,75	94	*знищити до кореня*
0,22	217	*бачити власними очима*	-1,78	56	*закатувати під асфальт*
0,2	16	*на все свій час*	-1,8	4	*абсолютний нуль*
0,17	185	*на гарячому вчинку зловити*	-1,83	50	*попити крові*
0,14	355	*хилити на сон*	-1,86	68	*пустити по миру*
0,13	75	*доводити до відома*	-1,88	50	*драти шкуру*
0,11	224	*викидати з голови*	-1,89	28	*вимотувати всі жили*
0,1	3	*а мурий їсть собі та їсть*	-1,91	1	*«зелені» чоловічки*
0	939	*не винось бруду з хати*	-2	117	*агент Кремля*
-0,07	1	*«помаранчеві» сили*			

Фразеологічні одиниці у медійних текстах є ефективним стилістичним засобом, який допомагає авторам цілісно відтворювати дійсність шляхом образного узагальнення, впливати на адресата, адже образність фразеологізмів дає змогу уникнути прямої констатації факту, дозволяє змалювати ситуацію опосередковано, через картину, яку читач домислює самостійно. У базі даних функціонують і фразеологізми без оцінного значення, і фразеологізми з оцінним значенням. За результатами краудсорсингового експерименту анотатори найбільшій кількості фразем приписали оцінку «помірно негативна» (1445 фразем); на другому місці — нейтральні, тобто позбавлені оцінки фраземи (939); на третьому —

ISSN 2311-259X

помірно позитивні фраземи (706). Ці оцінки утворюють три піки, між якими — зони розкиду значень, які, власне, відображають уявлення носіїв мови про оцінний потенціал ФО (рис. 2).

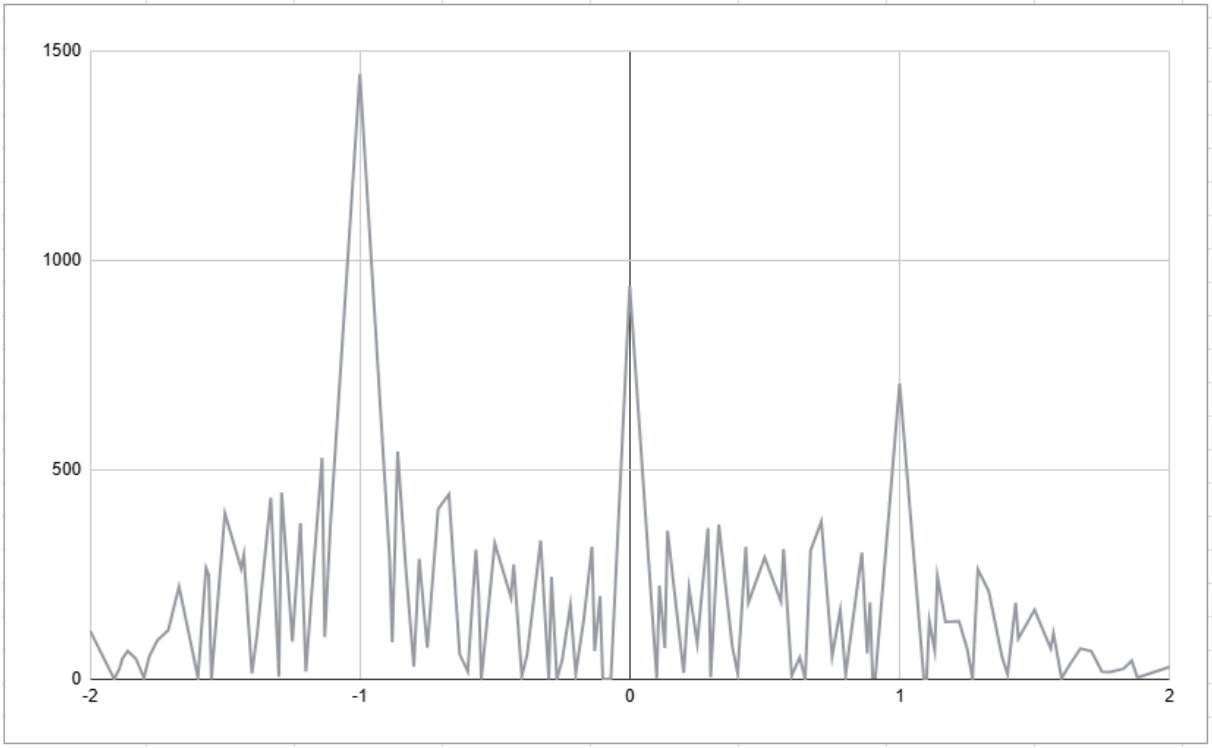


Рис. 2 Гістограма розподілу фразем за усередненими показниками: дуже негативна (-2), помірно негативна (-1), нейтральна (0), помірно позитивна (+1) і дуже позитивна (+2)

Виявилося, що більшість фразем оцінені як негативно конотовані, причому фразем зі значеннями в діапазоні (-2) — (-1) вдвічі більше, ніж зі знаком + (рис. 3).

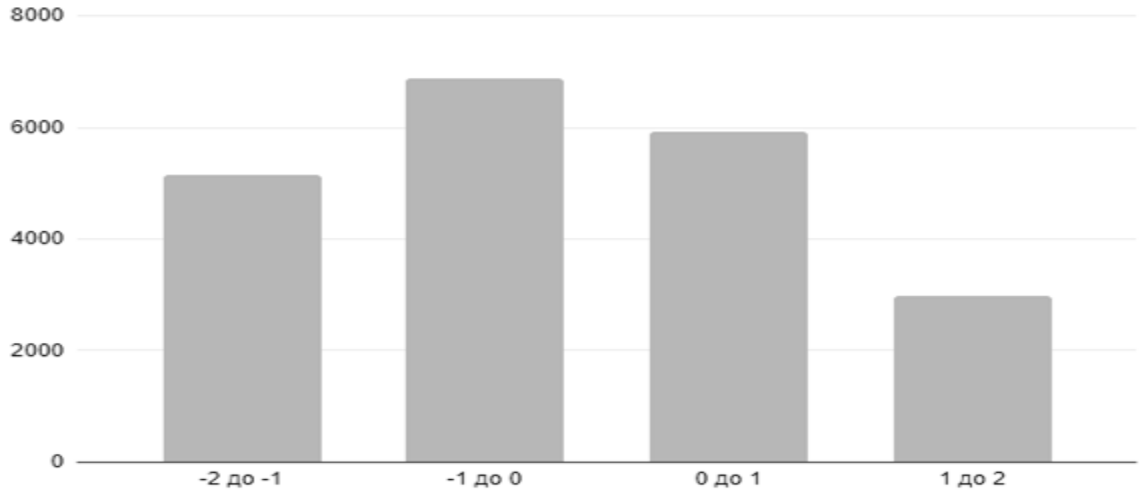


Рис. 3. Співвідношення загально позитивних і загально негативних фразем

Значна кількість фразем потрапила до інтервалів (-1) — (0) і (0) — (+1) з несуттєвою різницею у кількості, що свідчить про різний ступінь інтенсифікації оцінного компонента значень, як було зазначено вище: якщо розмовно-побутові фразеологізми виконують яскраво виражену оцінну функцію (*сидіти на шиї, напускати туману*) та становлять домінанту фразеологічних засобів, то книжні мають середній ступінь оцінності (*сізіфова праця, олімпійський спокій*), а фольклорні та народнопоетичні взагалі виконують слабко виражену оцінну функцію (*хоч з лиця воду пий*). Звідси висновуємо, що краудсорсинговий експеримент варто продовжити, збільшивши кількість анотаторів, щоб або підтвердити результати експерименту, або їх скоригувати. Наразі кожна фразема бази даних отримала

кількісний показник оцінки, яка в системі мови виражає сприймання її мовцями як такої, що має здатність посилювати чи послаблювати ступінь вияву ознаки або інтенсивність протікання дії.

Для удосконалення експерименту доцільно в майбутньому поглибити аналіз оцінного компонента фразеологічної семантики, диференціювавши фраземи за типом оцінки, яку вони експлікують, — раціональну чи емоційну. Така диференціація уможливить точніше пояснення розкиду усереднених значень, отриманих під час краудсорсингу, оскільки два типи оцінки спираються на різні психологічні механізми сприйняття й, відповідно, можуть по-різному впливати на стабільність і однорідність експертних оцінок. Раціональна оцінка ґрунтується на логіко-поняттєвому, інтелектуальному осмисленні позначуваної ситуації: мовець кваліфікує дію чи стан як доцільний / недоцільний, ефективний / неефективний, нормативний / ненормативний, спираючись на об'єктивні критерії – суспільну норму, логіку причиново-наслідкових зв'язків, користь та ін. Цей тип оцінки переважно позбавлений яскравого афективного забарвлення й апелює радше до розуму, ніж до почуттів реципієнта. Так, ФО *скласти подяку* реалізує позитивну раціональну оцінку через указівку на відповідність вчинку соціально схваленій нормі, а, наприклад, *доводити до абсурду* — негативну раціональну оцінку, що кваліфікує спосіб аргументації як логічно невиправданий, помилковий. До цієї групи ФО можна віднести й такі фраземи, як *дотримуватися букви закону, за всіма канонами жанру*, *віддавати данину,* що стосуються співвіднесення дії з критерієм правильності чи неправильності, а не пов'язані з безпосередньою емоційною реакцією мовця. Емоційна оцінка, на відміну від раціональної, спирається на чуттєвий компонент сприйняття й передає безпосередню емоційну реакцію — обурення, презирство, захоплення, відразу, симпатію тощо. Скажімо, ФО *підкласти свиню* викликає не стільки логічний висновок про порушення норми, невиправданість якогось учинку, скільки емоційний осуд підступності й зради, *скакати в гречку* передає не нейтральну констатацію подружньої зради, а зневажливо-іронічне ставлення мовця до такого вчинку. У цій групі варто аналізувати й фраземи *драти шкуру, попити крові* (вираження обурення жорстокістю), *рвати боки зо сміху* (передача яскравої позитивної емоції), *наче ножем по серцю* (експлікація гострого емоційного болю) та ін.

Виокремлення цих двох груп має суттєве методологічне значення для подальшого експерименту. По-перше, воно дозволить перевірити гіпотезу про те, що фраземи з емоційною оцінкою демонструють більший розкид значень, оскільки афективна реакція індивіда залежить від суб'єктивного емоційного досвіду, тоді як раціональна оцінка спирається на відносно стабільні, об'єктивні нормативні критерії і тому має тяжіти до меншої варіативності в оцінках респондентів. По-друге, така типологія уможливить точніше пояснення прагматичного потенціалу фразеологізмів у медіатекстах: одиниці з емоційною оцінкою, очевидно, будуть ефективнішим інструментом маніпулятивного впливу, оскільки діють на рівні афекту, минаючи критичне мислення, тоді як фраземи з раціональною оцінкою радше формують аргументативну, а не сугестивну стратегію переконання читача. По-третє, розмежування цих груп дасть змогу зіставити отримані результати з лінгвістичними теоріями оцінки, що підвищить теоретичну обґрунтованість дослідження та його інтеграцію в загальний контекст фразеологічної семантики.

Отримані внаслідок краудсорсингу дані загалом засвідчують спроможність респондентів розмежовувати позитивну й негативну оцінку чи її відсутність у структурі фразеологічної одиниці, водночас з поодинокими результатами тональної розмітки важко погодитися. Наприклад ФО *бути хазяїном свого слова* має вагу тональності 0,9, а синонімічна фразема *господар своєму слову* — 1,56; ФО *брати з повітря* за підсумками експерименту присвоєно вагу 0. Ці та деякі інші моменти потребують уточнення та докладнішого аналізу. Попри це, отримані результати вважаємо задовільними. Наступним кроком дослідження буде імплементація результатів краудсорсингу в алгоритм сентимент-аналізу системи «TextAttributor 1.0» та перевірка коректності роботи алгоритму, повноти бази даних, адекватності шкали оцінки та зручності підтримки системи, тобто можливості оперативно коригувати систему під тексти інших стилів, адже, оскільки під час її створення джерельною базою був дослідницький корпус медіатекстів, стилеметричне порівняння вона виконує лише для текстів медійного стилю, хоча загальний лінгвостатистичний аналіз доступний для українськомовних текстів будь-якого стилю.

Висновки. Проведений аналіз дозволяє висновувати, що ФО є потужним інструментом вираження оцінки в медійних текстах та ефективним засобом впливу на адресата. Краудсорсинговий експеримент на основі бази даних із 20 985 фразем показав, що більшість ФО української мови має негативну конотацію: фразем у діапазоні від -2 до -1 вдвічі більше, ніж фразем зі знаком + в аналогічному діапазоні. Це відображає особливості сприйняття людиною дійсності та загальну тенденцію до більшого розрізнення негативних станів і явищ порівняно з позитивними.

Результатом краудсорсингового анотування стала тональна розмітка сформованої бази даних, яка буде імплементована в модуль сентимент-аналізу системи «TextAttributor 1.0». Здобуті

шляхом краудсорсингу кількісні показники оцінного потенціалу ФО дають змогу, по-перше, статистично підтвердити виявлену кореляцію між ступенем інтенсифікації оцінки та функціонально-стильовою належністю фраземи; по-друге, використовувати отриману вагу тональності кожної фраземи як кількісний критерій розмежування емоційно забарвленого й нейтрального тексту в автоматизованих системах аналізу.

Дослідження підтвердило ефективність краудсорсингу як об'єктивного методу визначення оцінного компонента фразеологічної семантики та створення тонально розміченої бази даних фразеологічних одиниць за відносно короткий проміжок часу. Водночас застосована методика унаявнила деякі неточності, що можуть бути пов'язані і зі структуруванням самої бази даних (зокрема з дублюванням варіантів фразем, недостатньою диференціацією їхніх значень чи неоднорідністю формулювань дефініцій), і зі суб'єктивним сприйняттям деяких одиниць окремими анотаторами, зумовленим індивідуальним мовним досвідом, асоціаціями та розбіжностями в інтерпретації межі між раціональною й емоційною оцінкою. Виявлені неточності потребують подальшого докладного аналізу й усунення, що становитиме окремий етап удосконалення бази даних перед її імплементацією в систему «TextAttributor 1.0».

Перспективи дослідження пов'язуємо і з розширенням бази даних за рахунок нео-фразеологізмів, що активно виникають в сучасних суспільних умовах, фразеологізмів, активно вживаних в інших функціональних стилях української мови. На часі й докладніший аналіз оцінного потенціалу фразеологізмів у різних типах медіатекстів: традиційних друкованих, онлайн-видань і соціальних мереж. Окрім того, збільшення вибірки анотаторів у краудсорсинговому експерименті дасть змогу підвищити статистичну надійність отриманих усереднених показників оцінок за рахунок зменшення впливу випадкової варіативності індивідуальних суджень. Зростання маніпулятивного використання фразеологізмів в умовах інформаційної війни робить актуальним і завдання розроблення методів автоматичного розпізнавання маніпулятивних стратегій у медіатексті тощо.

Умовні скорочення

КУМ — Дарчук, Н., Сорокін, В., Сірук, О., Ходаківська, Я., Чейлитко, Н., Лангенбах, М. (2003–2026). *Корпус текстів української мови.* https://www.mova.info/corpus.aspx?l1=209

TextAttributor — Гончаренко, К., Дарчук, Н., Сорокін, В., Зубань, О., Робейко, В., Цигвінцева, Ю., Сажок, М., Костіков, М. (2024–2026). *TextAttributor 1.0: комп'ютерний аналіз українськомовного тексту.* http://ta.mova.info

Покликання

Білоноженко, В., Винник, В., Гнатюк, І. та ін. (Уклад.). (1993). *Фразеологічний словник української мови*. У 2 кн. Наукова думка.

Ганжа, С., & Онищенко, Г. (2016). Відображення категорії оцінки в сучасній фразеології. *Дослідження з лексикології і граматики української мови*, *17*, 34–40. http://nbuv.gov.ua/UJRN/dlgum_2016_17_6

Дарчук, Н., Зубань, О., Робейко, В., & Цигвінцева, Ю. (2024). Індексування негативного сентименту українськомовного тексту системою «TextAttributor 1.0». *Українське мовознавство*, *1*(54), 204-221. https://doi.org/10.17721/um/54(2024).204-221

Деменчук, О. (Голов. ред.). (2023). *Мовна концептуалізація аксіологічного простору людини: колективна монографія*. РДГУ.

Зирка, В. (2004). *Манипулятивные игры в рекламе: лингвистический аспект*. ДНУ.

Космеда, Т. А. (2001). *Аксіологічні аспекти прагмалінгвістики: засоби вираження категорії оцінки в українській та російській мовах* [Автореф. дис. д-ра філол. наук, Харківський державний педагогічний університет ім. Г. С. Сковороди].

Коць, Т. (2021). Оцінність як лінгвістична категорія: вербальна реалізація і стилістична диференціація. *Вісник Київського національного університету імені Тараса Шевченка. Літературознавство, мовознавство, фольклористика*, *1*, 18–20. http://nbuv.gov.ua/UJRN/VKNU_LMF_2021_1_6

Краснобаєва-Чорна, Ж. (2015). Семантичний простір оцінної фраземіки. *Мовознавство*, *5*, 39–51. https://nasplib.isofts.kiev.ua/server/api/core/bitstreams/bbd71589-092b-4146-85e0-5945add371f5/content

Онищенко, І. (2011). Лексичні одиниці у структурі функціонально-семантичного поля оцінки сучасної української мови. *Лінгвістичні студії*, *22*, 303–308.

Онищенко, І. В. (2005). *Категорія оцінки і засоби її вираження в публіцистичних та інформаційних текстах* [Автореф. дис. канд. філол. наук, Дніпропетровський національний університет].

Сизонов, Д. (2018). "Якщо куля в лоб...": про маніпулятивні тактики нових фразеологізмів ідіоматичного характеру. *Актуальні проблеми української лінгвістики: теорія і практика*, *36*, 76–90. https://doi.org/10.17721/APULTP.2018.36.76-90

Сизонов, Д. (2019). Оцінність як категорія сучасної фразеології. *Актуальні проблеми української лінгвістики: теорія і практика*, *37*, 39–53. https://doi.org/10.17721/APULTP.2018.37.39-53

Степаненко, М., & Дейна, Л. (2018). *Суб'єктивна та об'єктивна оцінка в українському щоденниковому дискурсі*. Дивосвіт.

Ужченко, В., & Ужченко, Д. (2005). *Фразеологія сучасної української мови*. Альма-матер.

Халіман, О. (2019). *Граматика оцінки: морфологічні категорії української мови*. Майдан.

Чобот, І. (2002). Оцінка як функціонально-семантична категорія. *Нова філологія*, *3*, 130–137.

Шевченко, Л., & Сизонов, Д. (2017–2022). *Нові слова та фразеологізми в медіа*. ВПЦ «Київський університет».

Amossy, R. (2005). The argumentative dimension of discourse. In F. van Eemeren, & P. Houtlosser (Eds.), *Argumentation in Practice* (pp. 87–98). John Benjamins Publishing Company.

Bralczyk, J. (2004). *Język na sprzedaż*. GWP.

Williams, L., Bannister, C., Arribas-Ayllon, M., Preece, A., & Spasić, I. (2015). The role of idioms in sentiment analysis. *Expert Systems with Applications, 42(21),* 7375–7385. 10.1016/j.eswa.2015.05.039

Legend

КУМ — Darchuk, N., Sorokin, V., Siruk, O., Khodakivska, Ya., Cheilytko, N., Lanhenbakh, M. (2003–2026). Korpus tekstiv ukrainskoi movy [Corpus of the Ukrainian language]. *Mova.info: linguistic portal.* http://www.mova.info/corpus.aspx.

TextAttributor — Honcharenko, K., Darchuk, N., Sorokin, V., Zuban, O., Robeiko, V., Tsyhvintseva, Yu., Sazhok, M.,

Kostikov, M. (2024–2026). *TextAttributor 1.0: kompiuternyi analiz ukrainskomovnoho tekstu* [TextAttributor 1.0: Computer Analysis of Ukrainian-Language Text]. http://ta.mova.info.

References (translated and transliterated)

Amossy, R. (2005). The argumentative dimension of discourse. In F. van Eemeren, & P. Houtlosser (Eds.), *Argumentation in Practice* (pp. 87–98). John Benjamins Publishing Company.

Bilonozhenko, V., Vynnyk, V., Hnatiuk, I. et al. (Eds.). (1993). *Frazeolohichnyi slovnyk ukrainskoi movy. U 2 kn*. [Phraseological Dictionary of the Ukrainian Language (2 vols.)].Naukova Dumka.

Bralczyk, J. (2004). *Jezyk na sprzedaz* [Language for sale]. GWP.

Chobot, I. (2002). Otsinka yak funktsionalno-semantychna katehoriia [Evaluation as a functional-semantic category]. *New Philology, 3,* 130–137.

Darchuk, N., Zuban, O., Robeiko, V., & Tsyhvintseva, Yu. (2024). Indeksuvannia nehatyvnoho sentymentu ukrainskomovnoho tekstu systemoiu «TextAttributor 1.0» [Indexing of the negative sentiment of the Ukrainian text by "TextAttributor 1.0"]. *Ukrainian Linguistics, 1(54),* 204–221. https://doi.org/10.17721/um/54(2024).204-221

Demenchuk, O. (Ed.). (2023). *Movna kontseptualizatsiia aksiolohichnoho prostoru liudyny* [Linguistic Conceptualization of the Axiological Space of a Person]. RSUH.

Hanzha, S., & Onyshchenko, H. (2016). Vidobrazhennia katehorii otsinky v suchasnii frazeolohii [The reflection of the category of evaluation in modern phraseology]. *Studies in Lexicology and Grammar of the Ukrainian Language*, 17, 34–40. http://nbuv.gov.ua/UJRN/dlgum_2016_17_6

Khaliman, O. (2019). *Hramatyka otsinky: morfolohichni katehorii ukrainskoi movy* [Grammar of Evaluation: Morphological Categories of the Ukrainian Language]. Maidan.

Kosmeda, T. A. (2001). *Aksiolohichni aspekty prahmalinhvistyky: zasoby vyrazhennia katehorii otsinky v ukrainskii ta rosiiskii movakh* [Axiological aspects of pragmalinguistics: means of expressing the category of evaluation in Ukrainian and Russian] (Abstract of doctoral dissertation, H. S. Skovoroda Kharkiv State Pedagogical University).

Kots, T. (2021). Otsinnist yak linhvistychna katehoriia: verbalna realizatsiia i stylistychna dyferentsiatsiia [Axiology as a linguistic category: verbal realization and stylistic differentiation]. *Bulletin of Taras Shevchenko National University of Kyiv. Literary Studies, Linguistics, Folklore Studies*, *1,* 18–20. http://nbuv.gov.ua/UJRN/VKNU_LMF_2021_1_6

Krasnobaieva-Chorna, Zh. (2015). Semantychnyi prostir otsinnoi frazemiky [Semantics cope of evaluative phrasemics]. *Linguistics*, 5, 39–51. https://nasplib.isofts.kiev.ua/server/api/core/bitstreams/bbd71589-092b-4146-85e0-5945add371f5/content. [in Ukrainian]

Onyshchenko, I. (2011). Leksychni odynytsi u strukturi funktsionalno-semantychnoho polia otsinky suchasnoi ukrainskoi movy [Lexical units in the structure of the functional-semantic field of evaluation in modern Ukrainian]. *Linguistic Studies, 22*, 303–308.

Onyshchenko, I. V. (2005). *Katehoriia otsinky i zasoby yii vyrazhennia v publitsystychnykh ta informatsiinykh tekstakh* [The category of evaluation and the means of its expression in publicistic and informational texts] (Abstract of Candidate of Philological Sciences dissertation, Dnipropetrovsk National University).

Shevchenko, L., & Syzonov, D. (2017–2022). *Novi slova ta frazeolohizmy v media* [New Words and Phraseological Units in the Media]. Publishing and Printing Center "Kyiv University".

Stepanenko, M., & Deina, L. (2018*). Subiektyvna ta obiektyvna otsinka v ukrainskomu shchodennykovomu dyskursi* [Subjective and objective evaluation in Ukrainian diary discourse]. Dyvosvit.

Syzonov, D. (2018). "Yakshcho kulia v lob...": pro manipuliatyvni taktyky novykh frazeolohizmiv idiomatychnoho kharakteru ["If the bullet in the forehead...": about manipulative tactics of new phraseologisms of an idiomatic character]. *Current Issues of Ukrainian Linguistics: Theory and Practice, 36,* 76–91. https://doi.org/10.17721/APULTP.2018.36.76-90

Syzonov, D. (2019). Otsinnist yak katehoriia suchasnoi frazeolohii [Evaluation as category of modern phraseology]. *Current Issues of Ukrainian Linguistics: Theory and Practice*, *37*, 39–53. https://doi.org/10.17721/APULTP.2018.37.39-53

Uzhchenko, V., & Uzhchenko, D. (2005). *Frazeolohiia suchasnoi ukrainskoi movy* [Phraseology of Modern Ukrainian]. Alma-mater..

Williams, L., Bannister, C., Arribas-Ayllon, M., Preece, A., & Spasić, I. (2015). The role of idioms in sentiment analysis. *Expert Systems with Applications, 42(21),* 7375–7385. 10.1016/j.eswa.2015.05.039

Zyrka, V. (2004). *Manipuliativnye igry v reklame: lingvisticheskii aspekt* [Manipulative Games in Advertising: A Linguistic Aspect]. DNU.

Nataliia Darchuk
Taras Shevchenko National University of Kyiv, Ukraine

Oksana Zuban
Taras Shevchenko National University of Kyiv, Ukraine

Valentyna Robeiko
Taras Shevchenko National University of Kyiv, Ukraine

Yuliia Tsyhvintseva
Institute of Ukrainian Language of the NAS of Ukraine, Ukraine

THE EVALUATIVE POTENTIAL OF PHRASEOLOGICAL UNITS AND THE AUTOMATION OF ITS DETERMINATION IN MEDIA TEXT

As the volume of media content continues to grow, the automated recognition of text sentiment — particularly those components realized through idiomatic expressions — has become an increasingly pressing concern. The scholarly problem addressed in this study is the lack of a sentiment-annotated database of phraseological units that would objectively reflect native speakers' actual perceptions of the evaluative component of idioms and could serve as a foundation for computing sentiment indices in automated text analysis systems. The study focuses on sentiment annotation of phraseological units via crowdsourcing. The aim is to determine the evaluative potential of idioms in Ukrainian-language media texts based on

ISSN 2311-259X

crowdsourced annotation and to compile a sentiment-annotated database of phraseological units. The methods employed include a crowdsourcing experiment, distributional, componential, and quantitative-statistical analysis. The empirical material comprises 20,985 phraseological units extracted from the Corpus of the Ukrainian Language and compiled from Ukrainian newspapers and online media published between 2000 and 2024. This study is the first to test a methodology for the sentiment annotation of a large corpus of phraseological units through crowdsourcing: by engaging a group of independent expert philologists and subsequently computing the arithmetic mean of their ratings, the approach neutralizes the subjectivity inherent in individual evaluation and yields a statistically grounded quantification of the evaluative component of phraseological semantics. The results shed light on the process, outcomes, and challenges of the crowdsourcing approach to determining the evaluative component of phraseological semantics; establish the predominance of negatively connoted units in the resulting database; reveal that the degree of intensification of the evaluative component correlates with the functional-stylistic register of the idiom; and confirm the effectiveness of crowdsourcing as an optimal method for building a sentiment-annotated database of phraseological units. Prospects for further research include expanding the database with neo-phraseologisms and idioms used in other functional styles, conducting a more in-depth linguistic analysis of the database to resolve discrepancies in evaluation, and implementing the findings in the sentiment-determination module of the system for computer analysis of Ukrainian-language text.

Keywords: Ukrainian language; media text; evaluation; sentiment; phraseological unit; database of phraseological units; crowdsourcing; sentiment annotation.

Стаття надійшла до редколегії 07.07.2026
Прийнято до публікації 22.09.2026
Опубліковано 30.10.2026