

## Образливі висловлювання в українському й міжнародному контексті

**Євгенія Ковкіна**

кандидат філологічних наук, завідувач лабораторії почеркознавчих, лінгвістичних, психологічних та мистецтвознавчих досліджень, Національний науковий центр «Інститут судових експертиз ім. Засл. проф. М. С. Бокаріуса» Міністерства юстиції України, м. Харків, Україна, ORCID: <https://orcid.org/0000-0001-8693-9868>, e-mail: [evkovkina@gmail.com](mailto:evkovkina@gmail.com)

**Ігор Лущик**

старший науковий співробітник відділу науково-методичної роботи та післядипломної освіти Київського відділення, Національний науковий центр «Інститут судових експертиз ім. Засл. проф. М. С. Бокаріуса» Міністерства юстиції України, м. Київ, Україна, ORCID: <https://orcid.org/0000-0001-9766-4628>, e-mail: [fizedu@ukr.net](mailto:fizedu@ukr.net)

*Розглянуто національну й міжнародну проблематику лінгвістичних експертиз щодо висловлювань образливого характеру та основні виклики, з якими стикаються фахівці під час визначення такого контенту.*

**Ключові слова:** лінгвістична експертиза; образливі висловлювання; інвективна лексика.

## Offensive Remarks in the Ukrainian and International Contexts

**Yevheniia Kovkina, Ihor Lushchik**

*This paper examines national and international issues related to linguistic analysis of offensive statements and the main challenges experts face while identifying such content.*

**Keywords:** linguistic analysis; offensive statements; abusive language.

Виявлення висловлювань образливого характеру є не лише актуальним завданням лінгвістичної експертизи, а й предметом пильної уваги з боку всієї міжнародної спільноти. Сьогоднішній світ важко уявити без соціальних мереж, месенджерів та інших засобів глобальної комунікації, а залученість і залежність від цих явищ вимагають від свідомої громадськості дієвих механізмів забезпечення користувачів від зловживань у цій сфері.

Зважаючи на складну політичну, економічну й суспільну ситуацію в Україні внаслідок збройної агресії РФ, першочерговими завданнями є виявлення й дослідження таких висловлювань у кримінальному контексті.

Висловлювання образливого характеру — ключовий об'єкт семантико-текстуальної (лінгвістичної) експертизи, яку зазвичай призначають під час розгляду справ щодо захисту честі, гідності (ст. 97 Цивільного кодексу України [1] (далі — ЦК України), відшкодування моральної шкоди (ст. 23 ЦК України) і ділової репутації (ст. 299 ЦК України); наклепів або образ (ст. 151, 277 Кримінального кодексу України [2] (далі — КК України), у справах про захист прав споживачів, трудові спори,

публічні виступи, дописи в соціальних мережах або засобах масової інформації, а також рідше в кримінальних справах про погрози (наприклад, погроза вбивством за ст. 129 КК України), заклики (зокрема, публічні заклики до вчинення терористичного акту за ст. 258-2 КК України; публічні заклики до вчинення дій, що загрожують громадському порядку, за ст. 295 КК України; публічні заклики до вчинення дій, що мають ознаки жорстокого поводження з тваринами, або поширення матеріалів із закликами до вчинення таких дій за ст. 299 КК України; публічні заклики до геноциду за ст. 442 КК України; публічні заклики до насильницької зміни чи повалення конституційного ладу або до захоплення державної влади за ст. 109 КК України та ін.).

Головне завдання експерта-лінгвіста в таких справах — визначити, чи містить наданий на дослідження мовленнєвий об'єкт негативну інформацію образливого характеру та в якій формі вона виражена.

Із правового погляду, образа — це умисне приниження честі та гідності особи, виражене в непристойній формі [3]. Основні ознаки образи: непристойна форма (уживання



брутальної, лайливої лексики (інвектив), непристойних жестів або міміки); суб'єктивність (на відміну від наклепу, який є повідомленням неправдивих фактів, образа є емоційним оцінним судженням, яке неможливо перевірити на правдивість) і спрямованість на приниження особистого почуття гідності або соціальної репутації людини.

Натомість у лінгвістиці образу розглядають не просто як слово, а як специфічний мовленнєвий акт, мета якого — дискредитувати адресата. Основними параметрами лінгвістичного визначення є уживання інвективної лексики (слів або висловів, які в певній культурі сприймають як лайливі, непристойні або такі, що принижують гідність людини — нецензурна лексика, зоосемантичні метафори з негативними конотаціями, слова з негативним оцінним значенням щодо хвороб, фізичних вад або розумових здібностей); форми вираження (типологія) у вигляді прямої (вираженої експліцитно (явно) за допомогою брутальних слів) або непрямой (імпліцитної) образи (уживання літературних слів у контексті, що принижує гідність людини) і ознаки образливого вислову для аналізу адресності (вислів має бути спрямований на конкретну особу або групу осіб), негативної оцінки (лексема має містити негативну характеристику особистих якостей людини) і непристойної форми (чи є слово стилістично зниженим (вульгарним, брутальним) і чи виходить воно за межі етичних норм спілкування). Саме тому для того, щоб зрозуміти, чи було висловлювання образливим, необов'язковим є вживання непристойної (нецензурної) лексики: експерт зазвичай аналізує загальний контекст, комунікативний намір автора (інтенцію) і стилістичне забарвлення слів.

На відміну від українських національних пріоритетів міжнародна спільнота опікується проблемами фільтрування образливих висловлювань у соціальних медіа й у навчанні великих мовних моделей, що охоплюють численні аудиторії користувачів. 2021 р. польські науковці запропонували для аналізування образливих, агресивних і ненависницьких висловлювань під час створення анотацій у наборах даних для навчання мультимодальних моделей перейти від підходу, орієнтованого на дані, до підходу, орієнтованого на людину.

Аналізуючи образливий контент і мову ворожнечі, вони дійшли висновку про необхідність застосування персоналізованих підходів до ідентифікації образливого контенту, запропонувавши групову та індивідуальну моделі його сприйняття. Вони експериментально довели, що узгодженість між окремими особами та групами сильно корелює з якістю розпізнавання образливого контенту, тому слід застосовувати персоналізовані моделі, що адаптуються до особистих переконань користувачів, а не до загальноприйнятого розуміння образливості [4].

Образливу лексику на форумах у засобах масової інформації аналізували O. Dontcheva-Navratilova і R. Povolná, які досліджували три проблемні сфери: розмежування між здійсненими мовленнєвими актами, лексичну реалізацію конкретних властивостей об'єкта, а також ідентифікацію й категоризацію неявних образ. Особливу увагу під час ідентифікації мовленнєвих актів вони приділяли контекстуальним підказкам, що вказували на властивості об'єкта, і типам семантичних зсувів у неявних виразах образливості. Вони дійшли висновку, що можуть виникнути проблеми щодо адекватної інтерпретації досліджуваних мовних засобів, якщо для виявлення образливої мови не користуватися контекстною інформацією [5].

Активна популяризація соціальних мереж, яка відбулася останнім часом, сприяла надмірному поширенню образливого й ненависницького контенту, насамперед завдяки певній анонімності, яку вони нібито надають. Це мало серйозні негативні наслідки для багатьох людей та спільнот і викликало занепокоєння громадськості. Тому на популярних міжнародних платформах соціальних медіа (наприклад, у твіттері, ютубі або фейсбуці) з'явилися методики виявлення образливих висловлювань спочатку в одномовних (насамперед англomовних) середовищах, але, зважаючи на міжнаціональний і багатокультурний формат соціальних мереж, виникла потреба в більш ґрунтовному їх модеруванні з огляду на їх багатомовні та міжмовні аспекти. Тому Kh. Mnassri зі співавт. прове-

ли розгалужене дослідження з виявлення образливих висловлювань у багатомовних текстах. У їх дослідженні ретельно проаналізовано багатомовні набори даних, описано глобальні ініціативи та проекти, спрямовані на боротьбу з образливими висловлюваннями (наприклад, німецький *PROZHATERS: PROactive PROFiling of HATE speech spreaders* або американський *ProPublica's Documenting Hate project*), а також підкреслено важливість загальнодоступних інструментів з відкритим кодом для сприяння подальшим дослідженням та їх практичному застосуванню [6].

Одним із перших досліджень, у якому окреслено проблему створення автоматизованих систем розпізнавання образливої лексики в соціальних медіа для неанглійської мови, стала робота *G. I. Sigurbergsson* і *L. Derczynski*, які створили данські набори даних для чотирьох систем автоматичної класифікації, що містять коментарі користувачів із різних соціальних мереж і фіксують різні типи та об'єкти образливих виразів як для англійської, так і для данської мов [7].

Ідею виявлення образливої лексики на основі глибокого навчання для багатомовних даних зі змішаним кодом досліджували *K. Shanmugavadeivel* зі співавт. Вони розглядають виявлення образливих висловлювань необхідним не взагалі, абстраговано, а щодо наявності в коментарях небажаного тексту, що зачіпає окрему особу або групу. З іншого боку, для подолання обмежень на висловлювання, визначені в публічних політиках соціальних медіа і для функціонування яких адміністратори залучають значні людські та програмні ресурси, користувачі соціальних платформ почали користуватися незаконними способами (так звані дані зі змішаним кодом) шляхом поєднання слів і фраз із двох або більше різних мов в одному тексті, що значно ускладнило їх ідентифікацію [8].

Зрозуміло, що масштаб тез не дає змогу висвітлити всі наявні проблеми в цій галузі, це потребує докладнішої уваги й передбачає подальші дослідження. Але можемо констатувати, що лінгвістична експертиза образли-

вих висловлювань посідає в українському інформаційному просторі важливе місце, оскільки напрям забезпечує національні інтереси та дотримання законності у громадянському суспільстві.

### Перелік джерел посилання

1. Цивільний кодекс України від 16.01.2003 р. № 435-IV (зі змін. та допов.). URL: <https://zakon.rada.gov.ua/laws/show/435-15#Text> (дата звернення: 14.04.2026).
2. Кримінальний кодекс України від 05.04.2001 р. № 2341-III (зі змін. та допов.). URL: <https://zakon.rada.gov.ua/laws/show/2341-14#Text> (дата звернення: 14.04.2026).
3. Ємельяненко В., Оробець К. Образа честі і гідності військовослужбовця як форма об'єктивної сторони кримінального правопорушення за ст. 435-1 КК України. *Науковий вісник Ужгородського національного університету. Серія: Право*. 2023. Т. 2. № 75. С. 93—99. DOI: 10.24144/2307-3322.2022.75.2.15 (дата звернення: 14.04.2026).
4. Kocoń J., Figas A., Gruza M., Puchalska D., Kajdanowicz T., Kazienko P. Offensive, aggressive, and hate speech analysis: From data-centric to human-centered approach. *Information Processing & Management*. 2021. Vol. 58. Is. 5. DOI: 10.1016/j.ipm.2021.102643 (дата звернення: 14.04.2026).
5. Dontcheva-Navratilova O., Povolná R. Offensive language in media discussion forums: A pragmatic analysis. *Lodz Papers in Pragmatics*. 2023. Vol. 19 (2). Pp. 223—238. DOI: 10.1515/lpp-2023-0012 (дата звернення: 14.04.2026).
6. Mnassri Kh., Farahbakhsh R., Chalehchaleh R., Rajapaksha P., Jafari A. R. et al. A survey on multi-lingual offensive language detection. *PeerJ Computer Science*. 2024. DOI: 10.7717/peerj-cs.1934 (дата звернення: 14.04.2026).
7. Sigurbergsson G. I., Derczynski L. Offensive Language and Hate Speech Detection for Danish. 2023. URL: <http://arxiv.org/1908.04531v2> (дата звернення: 14.04.2026).
8. Shanmugavadeivel K., Sathishkumar V. E., Raja S. et al. Deep learning based sentiment analysis and offensive language identification on multilingual code-mixed data. *Scientific Reports*. 2022. Vol. 12. DOI: 10.1038/s41598-022-26092-3 (дата звернення: 14.04.2026).