

УДК 177

DOI: <https://doi.org/10.30970/PPS.2022.42.15>

ЕТИЧНІ ФЕНОМЕНИ СОЦІАЛЬНИХ МЕРЕЖ: МОВА ВОРОЖНЕЧІ, КІБЕРНЕНАВИСТЬ ТА КІБЕРБУЛІНГ

Анастасія Олейник

*Національна металургійна академія України,
гуманітарний факультет, кафедра філософії та політології
проспект Гагаріна, 4, 49600, Дніпро, Україна*

Актуальність проблеми. У статті описано та деталізовано окремі аспекти феномену кіберненависті та кібербулінгу у соціальних мережах як просторі для інформаційної взаємодії. Підкреслено, що користувачі соціальних мереж постійно зіштовхуються з кіберненавистю та мовою ворожнечі, перестаючи відчувати себе захищеними від впливу на фізичне та ментальне здоров'я. Різноманітність подібних жертв та водночас їхня схожість провокує актуалізацію питання про етичні норми у кіберпросторі після та у період світових змін, що вивели соціальні мережі на новий рівень.

Мета. Стаття присвячена дослідженню основних аспектів впливу глобальних світових змін на мову ворожнечі у соціальних мережах.

Методи. Було застосовано методи аналізу статистичного матеріалу, компаративного аналізу, порівняння цифрових даних організаційних звітів (Юнісеф тощо) та фокус-груп (наведено у аналізованих дослідженнях). Дослідження прикладів кіберненависті та мови ворожнечі було здійснено методом контент-аналізу у мережі Фейсбук та Інстаграм.

Результати дослідження. Кіберненависть, або мова ворожнечі, може спонукати до ксенофобії та політичним конфліктам між людьми. Така мова поділяє світ на «хороших» та «поганих», надаючи перевагу «хорошим», що автоматично ставить одних значно вище за інших. Така мова ворожнечі порушує не лише морально-етичні норми, але й міжнародний закон, утискаючи інших людей лише через їхню приналежність. Позаяк, мова ворожнечі є розповсюдженим способом вираження власних думок та свободи слова, така риторика не завжди несе у собі агресивний характер та не спонукає девіантну поведінку стосовно іншої людини. Мова ворожнечі це мова, що може слугувати як заклик до ненависті або агресії або спробою захистити себе та власну думку. Мова ворожнечі також може бути заборонена жорсткою цензурою, відбираючи у людини її право на слово, що представлено не тільки законами країни, а й на міжнародному рівні. Мова ворожнечі може бути способом самовираження, розкриттям власного світогляду, що робить людину унікальною серед інших. Самовиражаючись таким чином у людини є можливість бути поміченою, вона може знайти односторонній та повернути до себе увагу.

Ключові слова: кібербулінг, кіберненависть, кібер-етика, хайп, етика у соціальних мережах.

Вступ. У кіберпросторі кожна людина може відчувати певну свободу. Свободу слова, вираження власних думок та сприймання різноманітної інформації. Соціальні мережі – як частина світової павутини – дозволяють людям спілкуватися один із одним навіть на великій відстані. Знайомитися та підтримувати соціальний зв'язок та висловлювати власні думки. Саме соціальні мережі у період COVID-19 та карантину, дозволили нам залишатися у соціумі. Соціальні мережі можна розглядати як окремих простір, де стикається велика кількість людей різного віку, раси, релігії, культури та ін. Тому, саму у цей самий час активна діяльність у соціальних мережах привертає увагу та, що очевидно, може викликати певну реакцію. Ця реакція може бути неоднозначною та негативною, що провокує між людьми конфлікт. Тож, певна дозволеність та свобода має іншу, менш позитивну сторону і їх можна використати для вираження негативу або агресивної девіантної поведінки.

Такі феномени як мова ворожнечі, кіберненависть та кібербулінг – те, з чим має справу кіберетика, вона вирішує усі дилеми, що зустрічаються у кіберпросторі. Ці феномени характерні не лише для світової павутини, але частіше всього в Інтернеті знаходять свій осередок. Користувачі соціальних мереж кожен день зіштовхується з кібербулінгом, кіберненавистю та мовою ворожнечі, перестаючи відчувати себе захищеними від їхнього впливу на своє фізичне та ментальне здоров'я. Різноманітність «жертв» та водночас їхня схожість змушує адміністраторів соціальних мереж пильно слідкувати ситуацію, виявляючи порушників і застосовуючи до них «покарання».

Мова ворожнечі, кіберненависть та кібербулінг виражають сексизм, расизм, шовінізм, ейджизм, гомофобію, ксенофобію, газлайтинг та ін. Через велику кількість утиску людей такі явища привертають увагу вчених філософів, психологів та соціологів, які намагаються сприяти тому, щоб їхній вплив зменшився. Примітними є дослідження ряду вчених та фахівців-практиків: Хоссеймарді Х., Кастель М., Ватанабе Х., Бернап П., Четті Н., Хосравінік М., Лунструм Е., МакАвені С., Вільямс М. Л., Бернап П., Вахс С., Райт М. Ф.

Мета та завдання. Ціль даного дослідження було дослідити як впливають світові зміни на мову ворожнечі у соціальних мережах. Перед нами були поставлені задачі розглянути позитивні та негативні аспекти мови ворожнечі у соціальних мережах.

Методи дослідження. Нами було застосовано методи аналізу статистичного матеріалу, компаративного аналізу, порівняння цифрових даних організаційних звітів (UNICEF тощо) та фокус-груп (наведено у аналізованих дослідженнях). Дослідження прикладів кіберненависті та мови ворожнечі було здійснено методом контент-аналізу у мережі Facebook та Instagram.

Результати. Події пов'язані з пандемією COVID-19, змусили людей все більше зануритись у соціальні мережі та проводити час за спілкуванням онлайн. Саме тому, проблеми взаємодії, з якими довелося зіштовхуватися, стали найбільш актуальними. Це стосується усіх без винятку людей, не зважаючи на їх вік, стать расу або інші риси, притаманні великим групам людей.

Мова ворожнечі є основною складовою кібербулінгу. Саме використовуючи мову ворожнечі у мережі, опонент може виражати власну девіантну поведінку. Окрім мови ворожнечі існують також інші види девіантної поведінки у соціальних мережах такі як розпалювання ненависті, підбурювання до тероризму та геноциду [8].

Згідно з даними Мінцифри, з проблемою кібербулінгу зустрічається кожна п'ята людина в Україні. Кожний другий підліток відчуває на собі тиск кібербулінгу, так, серед опитаних 80% дівчат піддаються кібербулінгу у соціальних мережах. За тими ж даними, 70% хлопчиків вдаються до кібербулінгу. Приблизно саме таку статистику ми можемо спостерігати у Європі, де згідно з опитуванням, 80% людей відчувають на собі гніт кібербулінгу [2].

Синтезуючи теорію свободи слова та теорію критичної раси, Ендрю Селларс виділяє вісім ознак, які вказують на мову ворожнечі: 1 – мова спрямована на групу або індивіда як члена групи; 2 – зміст повідомлення виражає ненависть; 3 – мовлення завдає шкоди; 4 – оратор має намір заподіяти шкоду; 5 – промова спонукає до шкідливих дій за межами самої мови; 6 – промова є публічною або спрямована на члена групи; 7 – контекст робить можливим насильницьку відповідь; 8 – мова не має спокутної мети. Загалом, сучасні юридичні визначення мови ворожнечі відійшли від заборони лише образливих висловів і більше зосередилися на підбурюванні чи погрозі, тобто на формах висловлювання, які вже заборонені більшістю національних правових систем [12].

Первинні дослідження показали, що індивіди з девіантною та агресивною поведінкою формують власні групи у соціальних мережах. Соціальні мережі сприяють створенню

та розширенню спільнот однодумців. Цей аспект також стосується мови ворожнечі, оскільки кіберпростір має демократичний характер і дає можливість кожній людині долучитися до мови ворожнечі. Таким чином, вони, об'єднуючись, знаходять власний «спільний інтерес». Особливістю цих груп стає можливість залишатися інкогніто онлайн порівняно з офлайн-середовищем [6]. Провівши порівняльний аналіз А. Браун стверджує, що мова ворожнечі у кібер просторі може мати набагато гірші наслідки, ніж у офлайн-середовищі тому що привертає більшу аудиторію та зберігає можливість залишитися анонімним відправником ненависних повідомлень і коментарів [5]. Той факт, що мова ворожнечі поширилася у кібер просторі, означає, що її характер і наслідки набагато легше вивчити, ніж тоді, коли вона була обмежена традиційними медіа. Справді, суспільствознавці почали виявляти кореляцію між ворожими мовами в Інтернеті та офлайн-наси́льством [4; 9].

У 2020 році Твіттер почав позначати твіти президента Дональда Трампа міткою перевірки фактів після того, як він неправдиво стверджував, що бюлетені, надіслані поштою, були шахрайськими. У 2021 році у Твіттер заблокували аккаунт Трампа на два роки через заклики у соціальній мережі до насильства так як вважав, що він програв через шахрайство з бюлетенями. Тобто, минулий президент США використовував мову ворожнечі, яка могла спонукати до реального насильства у реальному житті, а не тільки згруповувати людей онлайн.

Соціологи Едвардс і Рашин, знайшли докази того, що вони називають «ефектом Трампа» – кореляції між виборами Трампа та зростанням злочинів на ґрунті ненависті в Сполучених Штатах протягом 2016 року. Використовуючи панельні регресії, вони виявили, що «округи, які голосували за президента Трамп з найбільшим відривом на президентських виборах також зазнали найбільше збільшення кількості зареєстрованих злочинів на ґрунті ненависті». Вони приходять до висновку, що вибори Трампа могли призвести до приблизно 410 додаткових злочинів на ґрунті ненависті [11].

Протидію мові ворожнечі онлайн надає кібер поліція та модератори соціальних мереж, а також штучний інтелект, який блокує яскраво виражену мову ворожнечі у соціальних мережах. Але, навіть модератори та штучний інтелект зіштовхується із проблемами, коли мова стосується мови ворожнечі. Основною проблемою стає упередженість. Мається на увазі, що штучний інтелект, який спершу навчається алгоритмам людиною також стає упередженим, адже людина має власну думку, яку і програмує у алгоритм. Тобто, штучний інтелект блокує повідомлення або коментарі лише у тому випадку, коли вважає це порушенням прав гідності людини, образою, заклику до насильства тощо. М. Кокельберг стверджує, що упередження часто є ненависним, але може виникнути на різних етапах процесу машинного навчання та науки про дані [7].

На думку М. Кокельберга, проблема упередженості виникає разом з тим, що люди – творці алгоритмів – не готові брати на себе відповідальність у повному обсязі і тому за недоліки у алгоритмах несе відповідальність лише штучний інтелект та не завжди навіть модератори. Окрім того, перед нами постає питання, чи можуть самі модератори або адміністратори соціальних мереж вирішувати, що є образою, а що ні?

Найчастіше модератори у соціальних мережах це звичайні люди, чия освіта не пов'язана з вирішуванням питань моральності інших людей. Якщо виникають питання, що пов'язані з образами чи булінгом у соціальних мережах, то модератори у першу чергу звертають увагу на кількість «репортів» – скарг і на те, чи були раніше скарги на цей профіль, а також на «показник адекватності». Але насправді при отриманні скарги модератори роздивляючись її зі своєї точки зору згідно з текстом можуть виявити образника, прийнявши заходи блокування коментарів або самого профілю.

З іншої точки зору, яку наводить Дж. Мілл надто емоційні висловлювання, що можуть трактуватися як мова ворожнечі, мають позитивні сторони. Ці висловлювання можуть допомогти знайти істину у суперечці або висловити власну думку без втручання цензури. Вони також розширюють коло ідей, які можуть звзунти через цензуру; на думку вченого, істина стає значно сильнішою коли стикається зі спротивом [10]. Критично ставлячись до цього підходу, К. Едвін Бейкер висловився, що думка про те, що істина виявляється через змагання з брехнею помилково передбачає, що істина є об'єктивною, а люди раціональні у своїх цілях [3].

Саме відштовхуючись від думки Дж. Мілла, ми погоджуємося, що власна думка не має піддаватися цензурі, адже згідно з Конституцією України кожна людина володіє правом на свободу слова та думки. Згідно з цим, людина в Україні може самовиражатись у рамках найвищого закону. У цьому випадку соціальні мережі пропонують найкращий шлях до самовираження та свободи у висловлюванні власних думок.

Маємо звернути увагу на те, що сказано вище: найчастіше мові ворожнечі піддаються люди певної національності, послідовники політичних поглядів, релігій, тощо. Яскравим прикладом подібного являється мова ворожнечі у період глобальних змін у світі до яких неможливо адаптуватися ні морально, ні фізично. У цей час, вираження власних думок може бути процесом, що розкриває істину поміж медійної та політичної брехні, що також має місце у соціальних мережах.

Тож, у період змін, соціальні мережі стають осередком різноманітної інформації, яку важко фільтрувати та яка може стати сміттям у пошуках власної істини. Істина, про яку іде мова, може також фільтруватися цензурою або ставати джерелом для обвинувачень з боку право чи кібер-охоронних органів.

Звернемо увагу на те, що у світі нещодавніх подій у країні, цензура у соціальних мережах, також медіа зменшується з кожним днем. Більш того, навіть власники соціальних мереж таких як Facebook та Instagram дозволили мову ворожнечі. Мова ворожнечі може стосуватися лідерів ворожих країн та солдат, але частіше за все штучним інтелектом розпізнати мову ворожнечі, ненависть до непричетних людей майже неможливо. Тому, соціальні мережі поповнюються ненависними коментарями, провокуючи на ксенофобію. Незважаючи на це, представники Meta – назва компанії, що володіє Facebook, Whatsapp і Instagram запевнили, що не допустять жодної ненависті по відношенню звичайних людей [13].

Meta запевнили, що вони проти ксенофобії, але цей процес може стати неминучим у майбутньому. Сотні людей піддаються критиці та ненависті через своє мовчання або як це можна розцінити «лояльність» до свого найвищого керівництва. Саме тому, щоб не піддавати людей ненависті та мові ворожнечі, очевидним стає заборона Meta для користувачів [1]. Певним чином, ця заборона можна стати цензурою для екстремізму або спроба захистити людей від ненависті у соціальних мережах.

У деяких випадках, коли недостатньо лише «забанити» людину, необхідно також розглянути мову ворожнечі як спосіб самозахисту. Мова ворожнечі може не нести у собі насильницькі заклики або спробу грубо образити опонента або опонентів, але вона теж може бути одним із способів захистити себе.

На рисунку 1. запропоновані приклади суперечки, де фігурує мова ворожнечі, але вона не виступає як агресія. Дана дискусія була знайдена у соціальних мережах під постом опису книги «Прозорість зла» Ж. Бодрійяра і ці коментарі були одними з найперших. Ми можемо спостерігати, як різність думок та поглядів стосовно цього твору посприяло суперечці і навіть можемо відстежити пасивну мову ворожнечі за сторони одного з коментаторів.

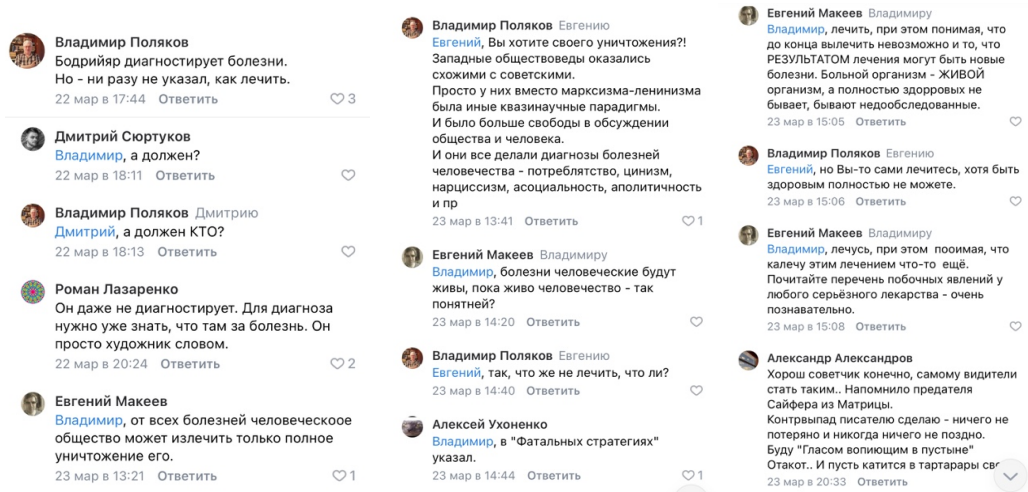


Рис. 1. Приклади мови ворожнечі без агресивної конотації

Джерело: Facebook

Подібні коментарі ми можемо зустріти майже під кожної публікацією, але при цьому не звернути увагу на суттєву різницю між мовою ворожнечі – агресивною, спонукаючою до насильницьких дій, провокуючи ненависть, та мовою ворожнечі, що захищає власне самовираження та сприйняття світогляду.

Висновки. Мова ворожнечі являється морально-етичним феноменом у соціальних мережах. Вона привертає увагу великої аудиторії та поширюється з кожної миттю все більше, варто лише натиснути на кнопку «post». Соціальні мережі – окремий простір, де діють закони, що підкріплені законами країн на юридичному рівні, але це саме й те місце, де слідкувати за їх виконанням дуже складно. Мова ворожнечі це мова, що може слугувати як заклик до ненависті або агресії або спробою захистити себе та власну думку. Мова ворожнечі також може бути заборонена жорсткою цензурою, відбираючи у людини її право на слово, що представлено не тільки законами країни, а й на міжнародному рівні.

Розглядаючи мову ворожнечі як феномен, що притаманний до соціальних мереж, нам вдалося дослідити яким чином вона може впливати на людину, або навіть цілий народ. Мова ворожнечі може несвідомо (чи навпаки) спонукати до ксенофобії та політичним конфліктам між людьми. Така мова ворожнечі здатна поділити світ на «хороших» та «поганих», надаючи перевагу «хорошим», що автоматично ставить одних значно вище за інших. Така мова ворожнечі порушує не лише морально-етичні норми, але в порушує міжнародний закон, утискаючи інших людей лише через їхню приналежність.

Мова ворожнечі являється актуальним способом вираження власних думок та свободи слова. Свобода слова – важливий компонент соціальних мереж. Така мова ворожнечі не несе у собі агресивний характер та не спонукає девіантну поведінку стосовно іншої людини. Мова ворожнечі може бути способом самовираження, розкриттям власного світогляду, що робить людину унікальною серед інших. Самовиражаючись таким чином у людини є можливість бути поміченою, вона може знайти однодумців та привернути до себе увагу.

Список використаної літератури

1. Берест Марина. У Росії суд остаточно заборонив Facebook та Instagram. *Мета* (21 березня 2022). 2022. URL: <https://meta.ua/uk/news/tech/42810-u-rosiyi-sud-ostatochno-zaboroniv-facebook-ta-instagram/> (дата звернення: 27.05.2022).
2. Мінцифра та ЮНІСЕФ в Україні презентували освітній серіал про кібербулінг для підлітків. 2020. URL: <https://www.unicef.org/ukraine/en/press-releases/ministry-digital-transformation-and-unicef-presented-edutainment-series-young-people> (дата звернення: 27.05.2022).
3. Baker Edwin C. Human Liberty and Freedom of Speech. New York, N.Y. : Oxford University Press. 1989. 385 p. URL: https://conservancy.umn.edu/bitstream/handle/11299/166001/08_01_BR_Stick.pdf;jsessionid=B3221046A49DF9B2B27AD7729BB5A3E0?sequence=1 (accessed 27 May 2022).
4. Baldauf J., Ebner J. and Guhl J. Hate Speech and Radicalisation Online: *the OCCI Research Report 5*; Foreword by Prof. Dr. Peter Neumann. 2019. URL: <https://www.isdglobal.org/wp-content/uploads/2019/06/ISD-Hate-Speech-and-Radicalisation-Online-English-Draft-2.pdf> (accessed 27 May 2022).
5. Brown A. What is hate speech? Part 1: the myth of hate. 2017. URL: https://www.researchgate.net/publication/315473313_What_is_hate_speech_Part_1_The_Myth_of_Hate (accessed 27 May 2022).
6. Brown A. What is so special about online (as compared to offline) hate speech?. *Ethnicities*. 2018. Vol. 18(3). P. 297-326. URL: https://ueaeprints.uea.ac.uk/id/eprint/64133/1/Accepted_manuscript.pdf (accessed 27 May 2022).
7. Coeckelbergh M. Artificial Intelligence: Some ethical issues and regulatory challenges. *Technology and Regulation*. 2019. Vol. 2019. URL: <https://techreg.org/article/view/10999> (accessed 27 May 2022).
8. Ghanea N. The Concept of Racist Hate Speech and its Evolution over time. 2012. URL: <https://www.ohchr.org/Documents/HRBodies/CERD/Discussions/RacistHatespeech/NazilaGhanea.pdf> (accessed 27 May 2022).
9. Jacobs J. Does Online Hate Drive Anti-Migrant Violence? *Financial Times* (Sept. 24, 2018). 2018. URL: <https://www.ft.com/content/a3d4b800-9bf3-11e8-88de49c908b1f264> (accessed 27 May 2022).
10. Mill John Stuart On Liberty; *Reprinted in*: John Arthur & William Shaw eds. (1992). Social & Political Philosophy. Englewood Cliffs, NJ : Prentice Hall, 1992. P. 283–84.
11. Muller K., Schwarz C. From Hashtag to Hate Crime: Twitter and Anti-Minority Sentiment. 2019. URL: https://warwick.ac.uk/fac/soc/economics/staff/crschwarz/hashtag_to_hatecrime.pdf (accessed 27 May 2022).
12. Sellars A. Defining Hate Speech. Berkman Klein Center Research Publication No. 2016-20, Boston Univ. *School of Law, Public Law Research Paper*. 2016. No. 16-48. URL: <https://ssrn.com/abstract=2882244> (accessed 27 May 2022).
13. Vengattil M., Culliford E. Facebook allows war posts urging violence against Russian invaders. *Reuters* (March 11, 2022). 2022. URL: <https://www.reuters.com/world/europe/exclusive-facebook-instagram-temporarily-allow-calls-violence-against-russians-2022-03-10/> (accessed 27 May 2022).

ETHICAL PHENOMENA OF SOCIAL NETWORKS: THE LANGUAGE OF THE ENEMY, CYBER HATE AND CYBER BULLYING

Anastasiia Oleinyk

*National Metallurgical Academy of Ukraine,
Faculty of Humanities, Department of Philosophy and Political Science
Gagarin Avenue 4, 49600, Dnipro, Ukraine*

Problem. This article describes and details some aspects of the phenomenon of cybernation and cyber bullying in social networks as a space for informational interaction. It has been stressed that social media users are constantly confronted with cyber hate and hate speech, ceasing to feel protected from the impact on physical and mental health. The diversity of similar victims and, at the same time, their similarity leads to the actualization of ethical norms in the cyber space after and during the period of world changes, which led the social media to a new level.

Purpose. This article is devoted to researching the main aspects of the influence of global world changes on the language of words in the social networks.

Methods. Methods of statistical material analysis, comparative analysis, comparison of digital data of organizational reports (Unicef, etc.) and focus groups (listed in the analyzed studies) were used. The research of applications of cybernation and the language of the enemy was carried out by means of content analysis on Facebook and Instagram.

Results. Hate, or the language of thieves, can lead to xenophobia and political conflicts between people. Such a language divides the world into «good» and «bad», giving preference to the «good», which automatically makes some people much higher than others. Such language violates not only moral and ethical norms, but also international law, by discouraging other people only because of their affiliation. After all, the speech of a thief is a popular way of expressing one's own thoughts and freedom of speech, such rhetoric is not always aggressive in nature and does not provoke deviant behavior towards other people. The language of enmity is a language that can serve as by an appeal to the hatred or aggression or attempt to defend oneself and own opinion. The language of enmity also can be forbidden by hard censorship, selecting for a man her right for a word, that it is presented by not only the laws of country, but also at an international level. The language of enmity can be a way of self-expression, a disclosure of one's own vision, which makes a person unique among others. By expressing themselves in this way, people have the opportunity to be noticed, they can find like-minded people and attract attention to themselves.

Key words: cyber bullying, cyber hate, cyber ethics, hype, ethics in social networks.