

УДК 811.161.2'33:004.85:528.9

DOI <https://doi.org/10.52726/as.humanities/2025.3.16>

О. І. РЕДЬКО

здобувачка другого (магістерського) рівня вищої освіти

кафедри прикладної лінгвістики,

Національний університет «Львівська політехніка», м. Львів, Україна

Електронна пошта: olena-maryna.redko.mflpl.2024@lpnu.ua

<https://orcid.org/0009-0005-4229-7679>

Т. В. ШЕСТАЕВИЧ

кандидат технічних наук, доцент,

доцент кафедри інформаційних систем та мереж,

Національний університет «Львівська політехніка», м. Львів, Україна

Електронна пошта: tetiana.v.shestakevych@lpnu.ua

<https://orcid.org/0000-0002-4898-6927>

ЕМОЦІЙНИЙ АНАЛІЗ СВІДЧЕНЬ ОЧЕВИДЦІВ РОСІЙСЬКО-УКРАЇНСЬКОЇ ВІЙНИ ЗА ДОПОМОГОЮ МОДЕЛЕЙ ОБРОБКИ ПРИРОДНОЇ МОВИ ТА КАРТОГРАФІЧНОЇ ВІЗУАЛІЗАЦІЇ

У статті досліджено поєднання сучасних методів автоматизованої обробки природної мови та картографічної візуалізації для аналізу емоційного забарвлення свідчень очевидців війни. Матеріалом став корпус із 1437 текстів, зібраних у 2022–2023 роках у межах проєкту «Narratives of War: Virtual Exhibit of Written Testimonies of the Russia-Ukraine War». Для емоційної класифікації використано багатоміткову модель «EmoBench-UA», адаптовану для української мови. В її основі лежить багатомовний трансформерний енкодер, який розпізнає шість базових емоцій (радість, страх, гнів, сум, відраза, здивування) та нейтральну категорію. Модель враховує співіснування кількох емоцій у тексті.

Аналіз показав, що домінантними емоціями є сум і страх, які разом становлять понад 60% від усіх маркерів. Це відображає травматичний характер досвіду війни, виражений втратами, невизначеністю й постійними загрозами. Сум у багатьох випадках перевищує 40–50%, тоді як страх становить 20–40% і в окремих регіонах сягає понад 60%. Нейтральні свідчення у середньому охоплюють 20–30% корпусу та свідчать про описовий характер викладу без емоційності. Інші емоції проявляються рідше: радість пов'язана з надією та солідарністю, гнів і відраза – з реакцією на агресію та воєнні злочини, здивування має епізодичний характер.

Картографічна візуалізація дала змогу простежити просторові закономірності. Отримані результати засвідчили переважання суму у центральних і західних областях та страху – у північних і східних. Дослідження доводить ефективність інтеграції природної обробки мови та мовознавчої картографії. Водночас нерівномірність кількості свідчень у певних регіонах обмежує репрезентативність висновків. Перспективним напрямом є розширення корпусу, залучення додаткових джерел і параметрів, що сприятиме глибшому вивченню емоційного ландшафту війни.

Ключові слова: емоційний аналіз; емоційне забарвлення; свідчення війни; обробка природної мови; картографічна візуалізація; EmoBench-UA.

Поставлення проблеми. Сучасні методи обробки та аналізу природних мов стають дедалі досконалішими збільшуючи швидкість, ефективність та точність проведення лінгвістичного аналізу тексту. Раніше робота з великими обсягами текстової інформації вимагала значних людських та часових ресурсів. Однак сьогодні цей процес суттєво пришвидшився, адже за допомогою натренованих мовних моделей можливо отримати попередній аналіз

тексту практично миттєво. У такий спосіб суттєво скорочується час від збору матеріалу до отримання інтерпретованих результатів.

Водночас отримані цифрові результати самі по собі не завжди забезпечують повне розуміння результатів аналізу. Значно більшої цінності вони набувають тоді, коли представлені у наочній формі. Одним із найефективніших способів візуалізації є картографування. Карти дозволяють відобразити розподіл певних мовних

явищ або характеристик текстів у просторі, що робить результати досліджень зрозумілими не лише для вузького кола лінгвістів, але й для ширшої аудиторії. Візуальне представлення на карті дає можливість одразу побачити територіальні відмінності, виявити закономірності, що залишилися б непоміченими у таблицях чи графіках, та робить висновки більш переконливими. Попри очевидні переваги такого підходу, спеціалізовані програмні засоби для картографічного розміщення результатів лінгвістичних досліджень майже відсутні.

Отже, актуальною постає проблема розробки та апробації методик, які поєднують сучасні досягнення у сфері обробки природних мов із картографічною візуалізацією, що визначає наукову новизну та підкреслює практичну цінність обраної теми. В межах цього дослідження було розроблено програмний інструмент, який дає змогу автоматично опрацьовувати свідчення очевидців війни за допомогою натренованих мовних моделей, визначати емоційний спектр кожного тексту та відображати результати на карті України.

Аналіз останніх досліджень і публікацій.

Питання визначення емоційного забарвлення текстів та картографування даних отриманих у результаті лінгвістичних досліджень аналізуються у широкій кількості праць українських та зарубіжних науковців. Нещодавня вкрай значуща робота проведена українськими дослідниками стала першою в Україні повноцінною еталонною розміткою для задачі розпізнавання емоцій у текстах. У межах дослідження створено емоційно анотовану базу (англійську схему адаптовано для української), використано краудсорсинг, а також протестовано як лінгвістичні базові підходи, так і великі мовні моделі [Dementieva, Fraser, Babakov].

Поєднання емоційного аналізу з просторовою візуалізацією постає як новий перспективний напрямок досліджень. У мовознавчій картографії вже напрацьовано потужний методологічний апарат, адже мовна карта виконує кілька функцій: документування, інтерпретацію та «візуальну економію». Вона фіксує мовні дані та створює нову когнітивну реальність у користувача завдяки поєднанню топографічних та лінгвістичних елементів [Girnth : 101–102]. Цей підхід можна розглядати

як теоретичну основу для візуалізації не лише мовних, але й емоційних даних, оскільки в обох випадках постає питання про баланс між точністю, узагальненням та наочністю. Сучасні підходи у сфері геоінформаційних систем відкривають нові можливості для реплікованості та інтеграції даних: традиційні мовні карти були значною мірою суб'єктивними, тоді як ГІС забезпечують алгоритмічну відтворюваність і дозволяють пов'язувати мовні чи емоційні дані з соціально-демографічними змінними [Goodchild : 3681].

У новітніх дослідженнях усе більше уваги приділяється власне поняттю емоційної картографії. Так, Н. Маркович пропонує осмислювати карти як семантичні простори, що здатні передавати «густину емоційних переживань», виходячи за межі традиційного географічного підходу [Marković : 1–2]. К. Андерсон, у свою чергу, систематизує напрями застосування емоцій у картографії, виокремлюючи карти емоцій, карти як інструменти збору емоційних даних та афективні реакції на самі картографічні продукти [Anderson : 1]. Також підтверджується валідність використання великих мовних моделей для емоційної класифікації свідчень. У роботі присвяченій емоційному аналізу корпусу інтерв'ю з очевидцями Голокосту автори зазначають, що «досягнення в галузі штучного інтелекту та обробки природної мови дозволили дослідникам отримати платформу та технологію для більш глибокого дослідження записів свідчень» [Ezeani et al : 95].

Мета статті полягає у дослідженні емоційного забарвлення свідчень про війну за допомогою моделей обробки природної мови та подальшої картографічної візуалізації результатів. В умовах війни свідчення очевидців набувають значення не лише як документальні джерела, а й як унікальні маркери колективних та індивідуальних емоційних переживань, що відображають соціально-психологічний стан суспільства. Саме тому актуальним є вивчення емоційного виміру текстів свідчень, яке поєднує досягнення комп'ютерної лінгвістики та методів геоінформаційного аналізу.

У межах поставленої мети передбачено здійснити автоматизований збір і попереднє опрацювання текстів свідчень; застосувати сучасну модель обробки природної мови для

класифікації емоційного забарвлення текстів українською мовою; побудувати картографічну візуалізацію результатів емоційного аналізу, що дозволить відобразити просторовий розподіл емоційних переживань та виявити потенційні закономірності їхнього прояву у різних регіонах. У ширшій перспективі дослідження окреслює можливості використання емоційної картографії у міждисциплінарних студіях.

Виклад основного матеріалу. Матеріалом дослідження постали 1437 свідчень очевидців війни у межах проєкту «Narratives of War: Virtual Exhibit of Written Testimonies of the Russia-Ukraine War» зібраних у 2022–2023 роках [Narratives]. Цей корпус є частиною ширшої ініціативи NoW (Narratives of War), у межах якої було зібрано та структуровано масштабну колекцію свідчень, що поєднують власне тексти з психометричними показниками респондентів. Дослідники зазначають, що такий ресурс відкриває нові можливості для лінгвістичного, психологічного та історичного аналізу, зокрема для завдань автоматизованої обробки мовлення та емоційної класифікації [Zasiekin et al : 2416].

Тексти були завантажені у хмарне середовище для програмування Google Colab, що надало гнучкі можливості для їхньої обробки та аналізу. На початковому етапі було здійснено автоматизований збір та структурування корпусу документів у форматі PDF. Для кожного тексту виконано попередню обробку, зокрема витягнення вмісту, що дало змогу перейти до

якісного аналізу. Значну увагу приділено підготовці метаданих, зокрема приведенню географічних назв до уніфікованого формату – це спростило подальше співставлення тексту з просторовими одиницями.

Ключовим етапом дослідження став автоматизований емоційний аналіз текстів. Для цього застосовано модель «EmoBench-UA», яка є багатомітковим класифікатором базових емоцій для української мови. В основі моделі лежить попередньо натренований генеративний енкодер (multilingual-E5-large), адаптований для багатоміткової класифікації емоційних станів: Joy (Радість), Fear (Страх), Anger (Гнів), Sadness (Сум), Disgust (Відраза), Surprise (Здивування) та None (Відсутність емоцій) [Dementieva, Fraser, Babakov : 1, 8]. Модель ідентифікує присутність кожної з цих емоцій у тексті шляхом оцінювання відповідних ймовірностей. Такий метод дозволяє ефективно аналізувати довгі тексти, оскільки один текст може містити відразу кілька емоційних маркерів.

Результати класифікації систематизовано у структурованій графічній та табличній формах, що надало змогу глибше аналізувати розподіл емоційних проявів.

Аналіз показує, що найпоширенішими емоційними станами у свідченнях стали сум (Sadness) та страх (Fear). У більшості областей вони разом складають понад 60% від усіх виявлених емоцій. Це відображає травматичний характер досвіду війни, що супроводжується

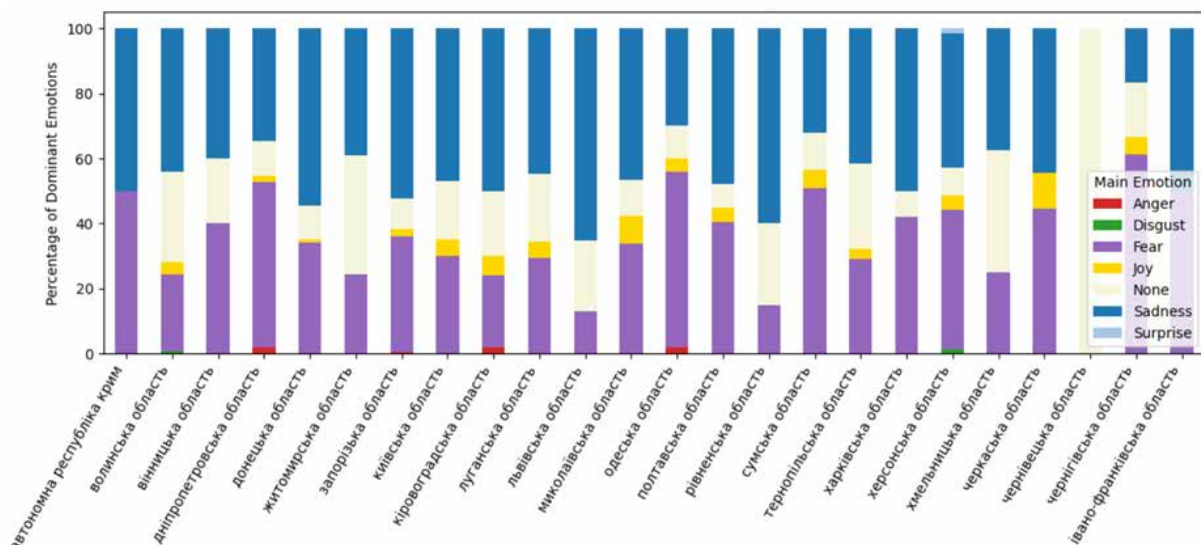


Рис. 1. Розподіл емоцій за областями у відсотках

втратою, невизначеністю та екзистенційною загрозою. Частка емоції сум (Sadness) у багатьох випадках перевищує 40–50%, тоді як страх (Fear) у середньому коливається в межах 20–40%, а у певних регіонах також перевищує 50–60%. У центральних та східних областях (Запорізька, Донецька, Харківська, Київська) частка суми (Sadness) особливо висока. Це пояснюється як великою кількістю свідчень, так і тим, що ці регіони зазнали активних бойових дій та масових руйнувань.

Відсутність емоційного маркування (None) також посідає суттєве місце у структурі – близько 20–30% у ряді регіонів. Це може свідчити про нейтральний або описовий характер частини свідчень, коли очевидці передають факти, уникаючи емоційності.

Інші емоції виявляються значно рідше. Радість (Joy) трапляється переважно у формі згадок про надію, вдячність чи солідарність, її питома вага зазвичай не перевищує 5–10%. Гнів (Anger) та відраза (Disgust) з'являються епізодично, однак їхня присутність важлива, оскільки вони відображають реакцію на агресора та воєнні злочини. Емоція здивування (Surprise) має найнижчу представленість і проявляється лише точково.

Варто підкреслити, що дані мають суттєву нерівномірність у кількості свідчень по

областях. Найбільший масив матеріалу надійшов із Запорізької області (311 свідчень) та Волинської області (156 свідчень). Водночас у низці регіонів кількість свідчень мізерна – Чернівецька область (1), АР Крим (2), Івано-Франківська область (4). Така диспропорція впливає на надійність висновків: відсоткові показники для областей з малою кількістю текстів не можна вважати репрезентативними, вони скоріше мають ілюстративний характер.

Порівняємо розглянуту графічну візуалізацію з картографічною.

На карті представлено домінуючі емоції у свідченнях очевидців війни з розподілом за областями України. Вона демонструє, який емоційний стан найчастіше зустрічається в текстах певного регіону, і якою є його частка в відсотках.

На відміну від стовпчикових діаграм, які показують співвідношення кількох емоцій у межах одного регіону, карта концентрується на провідній емоції. Завдяки цьому вона значно простіша для сприйняття і надає можливість побачити емоційний ландшафт України загалом. Один погляд на карту дозволяє відстежити просторові закономірності: які області переважно “забарвлені” сумом, де емоційна домінанта – страх, а де простежується відсутність емоційного маркування. Знову ж бачимо

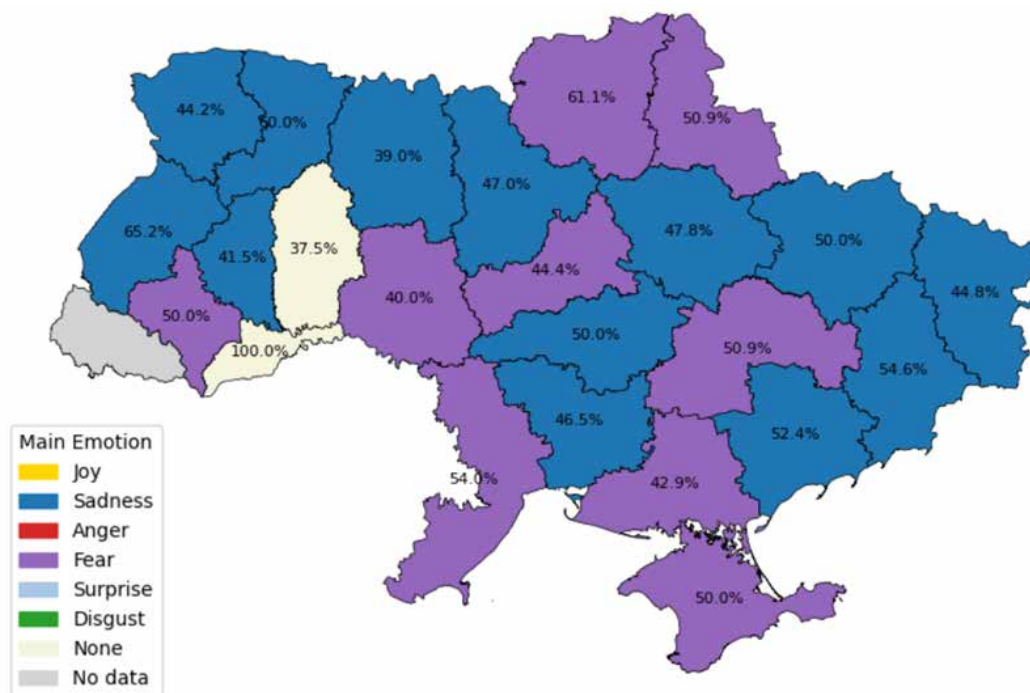


Рис. 2. Карта розподілу домінантних емоцій за областями у відсотках

абсолютне домінування емоцій сум (Sadness) з найвищим показником у Львівській області – 65,2% та страх (Fear) з найвищим показником у Чернігівській області – 61,1%.

Порівнюючи карту з попередньою діаграмою спостерігаємо, що стовпчикова діаграма відображала багатокомпонентний розподіл емоцій у межах кожної області. Вона була інформативною, але складнішою для сприйняття, особливо коли йшлося про порівняння багатьох регіонів одночасно. Натомість карта має більш наочний вигляд. Вона дозволяє легко простежити просторові закономірності, а також є інтуїтивно зрозумілою для широкої аудиторії, адже навіть без попереднього пояснення можна зробити висновок про емоційний стан свідчень у різних частинах України.

Висновки. Проведене дослідження засвідчило, що поєднання сучасних методів обробки природної мови та картографічної візуалізації відкриває нові перспективи для аналізу емоційного забарвлення текстів. Результати роботи підтвердили доцільність інтеграції автоматизованих лінгвістичних інструментів із просторовими методами представлення даних, адже така комбінація дозволяє не лише швидко та ефективно обробляти значні масиви текстів, але й робить висновки більш наочними та доступними для ширшого кола користувачів.

Застосована модель «EmoBench-UA» продемонструвала здатність ефективно розпізнавати базові емоційні стани у текстах українською мовою. Її багатомітковий підхід дозволяє враховувати складну природу людських переживань, адже у межах одного свідчення можуть співіснувати кілька різних емоцій. Аналіз показав, що серед свідчень очевидців домінують сум (Sadness) та страх (Fear), які разом формують понад половину емоційного поля. Це безпосередньо корелює з травматичним досвідом війни, що супроводжується втратами, руйнуваннями та відчуттям небезпеки. Відсутність емоцій (None) свідчить про прагнення частини

свідків уникати емоційної оцінки та зосереджуватися на фактах, що створює важливий контраст із більш емоційно насиченими текстами.

Інші емоції: радість, гнів, відраза та здивування, представлені значно меншою мірою, однак їхня поява вказує на багатовимірність досвіду війни. Зокрема, радість проявляється у вигляді солідарності та надії, гнів та відраза – як реакція на агресію та воєнні злочини, тоді як здивування часто пов'язане з неочікуваністю подій. Таким чином, навіть нечисленні прояви цих емоцій є важливими, оскільки вони відображають спектр психологічних реакцій суспільства.

Картографічна візуалізація результатів дозволила виявити просторові закономірності у поширенні емоцій. Карта дозволяє інтуїтивно зрозуміти емоційний ландшафт України та зробити порівняння між регіонами.

Водночас дослідження показало низку обмежень. Найбільшою проблемою є нерівномірність розподілу свідчень: у ряді областей їхня кількість є мізерною, що не дозволяє робити статистично значущі висновки. Іншим обмеженням є залежність результатів від якості самої моделі: попри високу точність «EmoBench-UA», автоматизовані класифікатори завжди мають певну похибку, особливо у випадках з іронією, багатозначністю чи складними культурними конотаціями.

Попри ці обмеження, результати дослідження мають важливе значення як для розвитку лінгвістики, так і для суміжних гуманітарних і соціальних наук. По-перше, вони демонструють потенціал використання цифрових інструментів для аналізу великих корпусів текстів. По-друге, показують, як інтеграція емоційного аналізу та картографії може покращити розуміння соціокультурних процесів. По-третє, створюють передумови для подальших міждисциплінарних досліджень, що поєднують лінгвістику, психологію та соціологію у спільному аналізі воєнного досвіду.

ЛІТЕРАТУРА

1. Anderson C. Quantifying Emotion: Survey Methods and Sentiment Analysis in Cartographic Design Research. *Abstracts of the ICA*. 2019. No. 1. P. 1–2. URL: <https://doi.org/10.5194/ica-abs-1-8-2019> (date of access: 18.08.2025).
2. Dementieva, D., Fraser, A., Babakov N. EmoBench-UA: A Benchmark Dataset for Emotion Detection in Ukrainian. 2025. 23 p. (Preprint. 2505.23297). URL: <https://doi.org/10.48550/arXiv.2505.23297> (date of access: 18.08.2025).
3. Girth H. Mapping language data. *Language and Space: Language Mapping : an International Handbook of Linguistic Variation*. Göttingen, 2010. P. 98–118.

4. Goodchild M. F. Mapping Language in the Age of GIS. *Handbook of the Changing World Language Map*. Cham, 2020. Vol. 6. P. 3673–3682.
5. Marković N. How to read ‘Emotional Cartographies’: Rethinking (Carto)graphic Representation and Semantics. *Abstracts of the ICA*. 2019. No. 1. P. 1–2. URL: <https://doi.org/10.5194/ica-abs-1-239-2019> (дата звернення: 18.08.2025).
6. Narratives of War: Virtual Exhibit of Written Testimonies of the Russia-Ukraine War. *Narratives of War*. URL: <https://now.omeka.net/> (date of access: 24.08.2025).
7. The Geography of ‘Fear’, ‘Sadness’, ‘Anger’ and ‘Joy’: Exploring the Emotional Landscapes in the Holocaust Survivors’ Testimonies / I. Ezeani et al. *CEUR Workshop Proceedings*. 2024. No. 3671. P. 93–103. URL: <https://ceur-ws.org/Vol-3671/paper8.pdf> (date of access: 17.08.2025).
8. The narratives of war (NoW) corpus of written testimonies of the Russia-Ukraine war / S. Zasiakin et al. *Lang Resources & Evaluation*. 2025. No. 59. P. 2415–2426. URL: <https://doi.org/10.1007/s10579-025-09813-8> (date of access: 18.08.2025).

REFERENCES

1. Anderson, C. (2019). Quantifying emotion: Survey methods and sentiment analysis in cartographic design research. *Abstracts of the ICA*, no. 1, pp. 1–2. Available at: <https://doi.org/10.5194/ica-abs-1-8-2019>
2. Dementieva, D., Fraser, A., & Babakov, N. (2025). *EmoBench-UA: A Benchmark Dataset for Emotion Detection in Ukrainian*. 23 p. (Preprint. 2505.23297). Available at: <https://doi.org/10.48550/arXiv.2505.23297>
3. Girnth, H. (2010). Mapping language data. In P. Auer & J. E. Schmidt (Eds.), *Language and Space: Language Mapping: An International Handbook of Linguistic Variation*, pp. 98–118. Göttingen: De Gruyter Mouton.
4. Goodchild, M. F. (2020). Mapping Language in the Age of GIS. In S. Brunn & R. Kehrein (Eds.). *Handbook of the Changing World Language Map*, vol. 6, pp. 3673–3682. Cham: Springer. Available at: https://doi.org/10.1007/978-3-030-02438-3_70
5. Marković, N. (2019). How to read “Emotional cartographies”: Rethinking (carto)graphic representation and semantics. *Abstracts of the ICA*, no. 1, pp. 1–2. <https://doi.org/10.5194/ica-abs-1-239-2019>
6. Narratives of War. (n.d.). *Narratives of War: Virtual exhibit of written testimonies of the Russia-Ukraine war*. Available at: <https://now.omeka.net/>
7. Ezeani, I., et al. (2024). The geography of “fear”, “sadness”, “anger” and “joy”: Exploring the emotional landscapes in the Holocaust survivors’ testimonies. *CEUR Workshop Proceedings*, no. 3671, pp. 93–103. Available at: <https://ceur-ws.org/Vol-3671/paper8.pdf>
8. Zasiakin, S., et al. (2025). The narratives of war (NoW) corpus of written testimonies of the Russia-Ukraine war. *Language Resources & Evaluation*, no. 59, pp. 2415–2426. Available at: <https://doi.org/10.1007/s10579-025-09813-8>

O. I. REDKO

Master’s Degree Student at the Department of Applied Linguistics,
Lviv Polytechnic National University, Lviv, Ukraine
E-mail: olena-maryna.redko.mfpl.2024@lpnu.ua
<https://orcid.org/0009-0005-4229-7679>

T. V. SHESTAKEVYCH

Candidate of Technical Sciences, Associate Professor,
Associate Professor at the Department of Information Systems and Networks,
Lviv Polytechnic National University, Lviv, Ukraine
E-mail: tetiana.v.shestakevych@lpnu.ua
<https://orcid.org/0000-0002-4898-6927>

EMOTIONAL ANALYSIS OF TESTIMONIES OF THE RUSSIA-UKRAINE WAR USING NATURAL LANGUAGE PROCESSING AND CARTOGRAPHIC VISUALIZATION

The article examines the integration of natural language processing (NLP) and cartographic visualization for analyzing the emotional dimension of war testimonies. The material includes 1437 texts collected within the project “Narratives of War: Virtual Exhibit of Written Testimonies of the Russia-Ukraine War” (2022–2023). Testimonies

were processed and classified using the EmoBench-UA model, a multi-label classifier adapted for Ukrainian. Based on a multilingual transformer encoder, the model identifies six basic emotions: Joy, Fear, Anger, Sadness, Disgust, Surprise, and a neutral category (None). Its multi-label approach captures the coexistence of several emotions in one text, reflecting the multidimensionality of human experience.

The analysis revealed the predominance of Sadness and Fear, which together form over sixty percent of emotional markers. This corresponds to the traumatic nature of war, linked to loss, uncertainty, and existential threat. Sadness often exceeded 40–50%, while Fear ranged from 20–40% and in some regions surpassed 60%. Neutral narration (None), averaging 20–30%, points to descriptive testimonies without explicit emotions. Other emotions were rarer: Joy appeared in contexts of solidarity and hope, Anger and Disgust reflected reactions to aggression, while Surprise occurred sporadically.

Cartographic visualization enabled the identification of regional patterns. Compared to charts, maps provided a more intuitive overview of Ukraine's emotional landscape, clearly showing dominant emotions by region.

The study highlights both potential and limitations. The uneven regional distribution of testimonies affects representativeness, pointing to the need for corpus expansion and integration of temporal and socio-demographic data for deeper insights and analysis.

Key words: emotional analysis; emotional tone; war testimonies; natural language processing; cartographic visualization; EmoBench-UA.

Дата першого надходження рукопису до видання: 17.08.2025

Дата прийнятого до друку рукопису після рецензування: 15.09.2025

Дата публікації: 30.10.2025