

**Literature:**

1. Francophonie. URL: <https://www.francophonie.org>.
2. Lingua Franca. URL: <https://www.britannica.com/topic/lingua-franca>.

**NLP-ТЕХНОЛОГІЇ: МОЖЛИВОСТІ ТА СФЕРИ  
ЗАСТОСУВАННЯ У ЛІНГВІСТИЦІ**

**Шуляк І.М.**

*кандидат філологічних наук, викладач  
кафедри англійської філології та методики навчання англійської мови  
Тернопільський національний педагогічний університет  
імені Володимира Гнатюка  
м. Тернопіль, Україна*

Сучасний лінгвальний простір маркований багатовекторністю та міждисциплінарним характером досліджень мовної діяльності особистості. Із розвитком інформаційних технологій відбувається вплив технологічного прогресу на вивчення мови у нових ракурсах. NLP (Natural Language Processing) у перекладі – обробка природної мови, як сучасні методи аналізу мовних одиниць, тексту і дискурсу загалом увірвалися у простір мовознавчої науки, пропонуючи нові можливості у реалізації лінгвістичних досліджень [1; 3; 4; 5; 6].

NLP-технології є частиною штучного інтелекту, який надає машинам здатність розуміти мову людини. Такий комп'ютеризований підхід до аналізу мовних категорій базується на низці теорій та наборі технологій. Український дослідник Олег Іванов визначає «аналіз природної мови як міждисциплінарну галузь науки, що охоплює методики обчислювальної лінгвістики та теорії штучного інтелекту, основним проблемним полем якої є забезпечення взаємодії мовленнєвих актів та комп'ютерних систем» [1].

Використовуючи лінгвістичні знання, статистичні методи та комп'ютер, розроблене програмне забезпечення не лише сприймає зміст вимовленого чи написаного, воно також може аналізувати й інтерпретувати контекстуальні відтінки значень, комунікативні інтенції та настрої мовця у тій мірі, як це роблять його співрозмовники наживо.

Обробка природної мови охоплює широкий спектр задач: від спам-фільтрів та визначення емоційного забарвлення тексту до голосового розпізнавання мовлення особистості та діалогових систем.

Відповідно до цього, бум розвитку NLP-технологій є результатом появи машинного перекладу (machine translation), сентимент аналізу (sentiment analysis), рівневої обробки текстових фрагментів за попередньо визначеними критеріями (морфологічний, лексичний, синтаксичний, семантичний, прагматичний рівні) (machine learning acts analysis), генерування/розпізнавання природної мови (speech recognition), перевірка грамотності текстів (grammar

correction), інтелектуальний аналіз тексту (text mining), векторне представлення (word embedding) тощо.

Розглянемо можливості NLP-технологій, а саме сентимент аналізу мовних одиниць в англomовному медіа дискурсі за допомогою застосунку Tweet Sentiment Visualization.

У наш час емоції захопили весь комунікативний простір особистості, стали одним із важливіших компонентів людського мислення та діяльності, тому з'являються нові ресурси для проведення досліджень вербальної взаємодії на визначення її емотивної складової. Аналіз тональності тексту як один з найбільш використовуваних методів NLP-технологій призначений для автоматизованого виявлення в текстах емоційно забарвленої мовної одиниці й оцінки мовцем об'єкту реального світу, події, процесу, тощо. Аналіз тональності тексту вилучає суб'єктивну інформацію з текстових файлів, наприклад, почуттів, думок, суджень або оцінок мовця щодо певної теми, події чи компанії з метою аналізу його настроїв.

Емоційна складова, виражена на рівні мовної одиниці або комунікативного фрагмента, називається лексичною тональністю (або лексичним сентиментом) [2].

У ході аналізу тональності першочергово визначається, чи є у реченні сентимент чи ні, і якщо він є, визначити, чи є емотивна оцінка позитивною, негативною чи нейтральною.

Програма *SentiStrength* (та її онлайн версія) створена для сентимент-аналізу та оцінки настрою тексту загалом чи мовної одиниці окремо, оцінюючи силу позитивних і негативних настроїв мовця до них.

Проілюструвати можливості такої програми пропонуємо на прикладі мовної одиниці *Ukraine*. За допомогою онлайн програмного застосунку Tweet Sentiment Visualization, який предствалений такими розділами, як: тональність (sentiment), тема (topics), теплокарта (heatmap), хмара тегів (tagcloud), шкала часу (timeline), твіт (tweets).

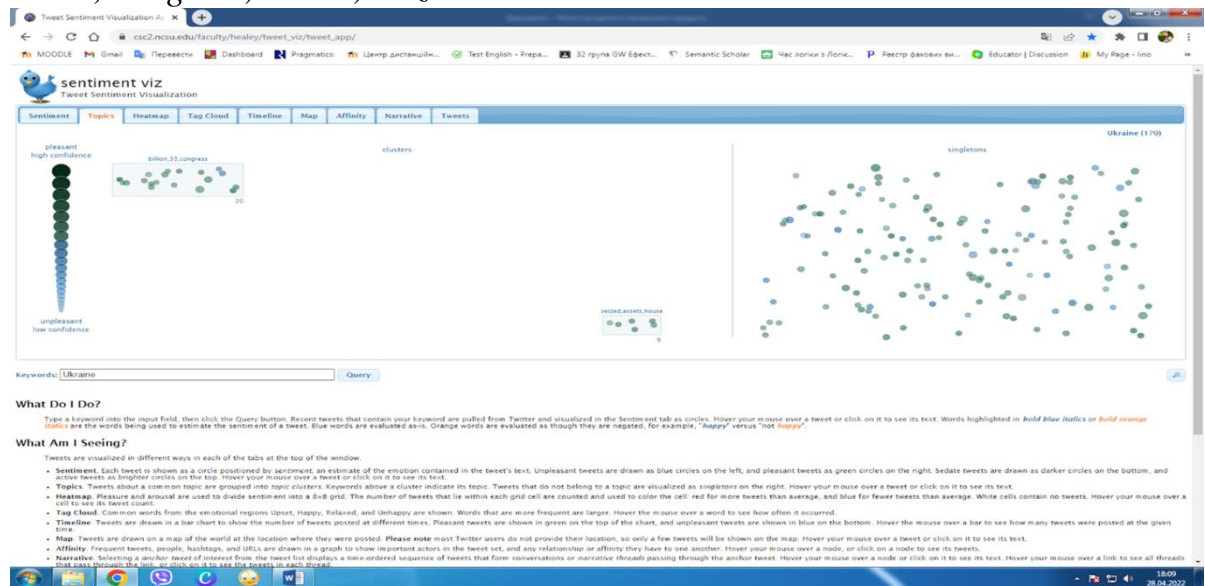
Тональність представлена записом у твіттері, який відображається у вигляді кола, розташованого за настроєм, оцінкою емоцій, що містяться в тексті твіту. Неприємні твіти малюються синіми кругом зліва, а приємні – зеленими – праворуч. Спокійні твіти представлені у вигляді темних кіл унизу, а активні твіти – у вигляді яскравіших кіл угорі. Якщо необхідно побачити його текст, то слід навести курсор миші на твіт або натисніть на нього. Вгорі рисунку можна полюбачити кількість твітів із мовної одиницею, а також дані змінюються щохвилини.

Наприклад, розглянемо рисунок 2, на якому представлено, сентимент мовної одиниці *Ukraine*, яка у твітах отримує позитивний характер.



**Рис. 2. Сентимент аналіз мовної одиниці Ukraine**

Категорія теми (topics) представлена твітами, які групуються в тематичні кластери за певною темою. Ключові слова над кластером вказують на його тему. Твіти, які не належать до теми, відображаються як одиночні повідомлення. Як бачимо на рисунку 4, основна тематика твітів за 28.04.2022р. представлена двома кластерами з такими ключовими словами, як: *humanitarian aid, billions, congress, assets, seized*.



**Рис. 3. Аналіз тематичних кластерів твітів з мовною одиницею Ukraine**

Хмара тегів представлена за допомогою чотирьох областей «Засмучений» (Upset), «Щасливий» (Happy), «Розслаблений» (Relaxed), «Нещасний»



них корисні відповіді. Обробка природної мови, поєднуючи комп'ютерні науки, штучний інтелект та методи й дослідження обчислювальної лінгвістики, формується у міждисциплінарну галузь та виступає медіатором у забезпеченні взаємодії між комп'ютером та людською мовою.

#### **Література:**

1. Іванов О. В. Класичний контент-аналіз та аналіз тексту: термінологічні та методологічні відмінності. Вісник Харківського національного університету імені В. Н. Каразіна, 2013. № 1045. С.72.

2. Обробка природної мови: веб-сайт. URL: <https://uk.wikipedia.org/wiki/%D0%9E%D> (дата звернення: 01.05.2022).

3. Jurafsky D., Martin J. Speech and Language Processing, 3rd edition, 2020. 623 p.

5. Bird, Klein, and Loper. Natural Language Processing with Python. Handbook of Natural Language Processing. 2009. 502 p.

6. Tatwadarshi P. Nagarhalli, Dr. Vinod Vaze, and Dr. N. K. Rana. Impact of Machine Learning in Natural Language Processing: A Review. The Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV 2021), 2021. pp. 1529–1534:

веб-сайт. URL: [https://www.researchgate.net/publication/354415641\\_Schizophrenia\\_Detection\\_Using\\_Machine\\_Learning\\_Approach\\_from\\_Social\\_Media\\_Content](https://www.researchgate.net/publication/354415641_Schizophrenia_Detection_Using_Machine_Learning_Approach_from_Social_Media_Content) (дата звернення: 01.05.2022).

7. Manning C. D. Foundations of Statistical Natural Language Processing, 1st Edition. 1999. 620 p.

## **ЛЕКСИЧНІ МАТРИЦІ В РАМКАХ ПСИХОЛІНГВІСТИЧНОЇ ТЕОРІЇ НА ПРИКЛАДІ ФРАНЦУЗЬКИХ КОДОВИХ ПЕРЕМИКАНЬ**

**Іваник О. О.**

*магістрантка*

*факультету іноземних мов*

*Тернопільський національний педагогічний університет*

*імені Володимира Гнатюка*

*м. Тернопіль, Україна*

**Науковий керівник:**

*доктор філологічних наук,*

*професор кафедри романо-германської філології*

**Косович О.В.**

Одна з моделей мовлення у психолінгвістиці належить В. Левелту. Він ділить процес мовлення на чотири стадії: 1) стадія підготовки довербального