

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Київський національний лінгвістичний університет

Кафедра східної і слов'янської філології

Кваліфікаційна робота магістра з польської філології

на тему:

**УКРАЇНСЬКІ ЗАПОЗИЧЕННЯ В ТЕКСТАХ ПОЛЬСЬКИХ НОВИН:
КОРПУСНЕ ДОСЛІДЖЕННЯ**

Студента групи Мпол58-22

факультету східної і слов'янської філології

денної форми навчання

Освітньої програми

Сучасні філологічні студії (польська мова і друга іноземна мова): лінгвістика та перекладознавство

Спеціальності 035 Філологія

Спеціалізації 035.033 Слов'янські мови та літератури (переклад включно),

перша – польська

Сови Євгенія Станіславовича

Науковий керівник:

доктор. філол. наук, проф. Палатовська О.В.

Допущена до захисту

« ____ » _____ 2023 року

Завідувач кафедри

(підпис) (ПІБ)

Національна шкала _____

Кількість балів _____

Оцінка ЄКТС _____

Київ – 2023

ЗМІСТ

ВСТУП	4
РОЗДІЛ 1. ТЕОРЕТИЧНІ ЗАСАДИ ВИВЧЕННЯ ЗАПОЗИЧЕНОЇ ЛЕКСИКИ	8
1.1. Дослідження іншомовних запозичень у лінгвістиці: коротка історія питання	8
1.2. Поняттєво-термінологічний апарат вивчення запозиченої лексики	17
1.3. Чинники входження та особливості освоєння іншомовних слів	24
Висновки до розділу 1	29
РОЗДІЛ 2. МЕТОДОЛОГІЯ ВИКОРИСТАННЯ МОВНИХ КОРПУСІВ У ЛІНГВІСТИЧНИХ ДОСЛІДЖЕННЯХ	32
2.1. Мовні корпуси як інструментарій новітніх лінгвістичних досліджень	32
2.2. Частотність та квантитативні методи у дослідженнях лексики	38
2.3. Методика та етапи корпусного дослідження українських запозичень польської мови	51
Висновки до розділу 2	55
РОЗДІЛ 3. НОВІТНІ УКРАЇНІЗМИ В ПОЛЬСЬКІЙ МОВІ ЯК МОВНЕ ВІДДЗЕРКАЛЕННЯ РОСІЙСЬКОЇ ВОЄННОЇ АГРЕСІЇ	57
3.1. Українські запозичення в польській мові: соціально-історичні та культурні передумови проникнення	57
3.2. Тематичні групи українізмів воєнного часу в текстах польських новин	61
3.2.1. <i>Символи героїзму та національної ідентичності</i>	66
3.2.2. <i>Назви і види зброї</i>	71
3.2.3. <i>Новітні реалії, породжені війною</i>	73
3.2.4. <i>Іронічне зображення риторики російської пропаганди</i>	74

	3
3.3. Новий погляд на українські культури79	79
3.4. Неологічні топоніми в текстах польських ЗМІ85	85
Висновки до розділу 390	90
ВИСНОВКИ92	92
РЕЗЮМЕ96	96
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ99	99
СПИСОК ДЖЕРЕЛ ІЛЮСТРАТИВНОГО МАТЕРІАЛУ104	104

ВСТУП

Актуальність теми. У сучасному науковому лінгвістичному дискурсі мова розглядається як динамічна система, яка, в сукупності з позалінгвальними факторами, є продуктом соціально-культурної та суспільно-політичної взаємодії. В умовах глобалізації та пов'язаних з нею інтеграційних процесів у сфері економіки, культури, політики, технологій, інфраструктури, туризму будь-яка мова функціонує як відкрита система, постійно розвиваючись та збагачуючись новими елементами. У процесі глобальної взаємодії на різних рівнях мовної системи відбувається конкуренція, що призводить до витіснення одних елементів системи та прийняття інших, зокрема внаслідок мовних *запозичень*. Саме завдяки постійній взаємодії між мовними системами різних культур відбувається їхній розвиток і адаптація до сучасних реалій.

Вивченню процесів іншомовних запозичень, ступеня адаптації запозичених мовних одиниць та особливостям їх функціонування у системі мови-реципієнта присвячено праці багатьох українських і зарубіжних мовознавців, зокрема Б. Ажнюка, М. Банько, Л. Блумфілд, А. Віталіш, Е. Гаугена, О. Єсперсена, Н. Жмаєвої, М. Клапер, Ф. Коетсема, В. Лабова, Л. Лисиченко, Ю. Молоткіної, Й. Рончашек-Леонарді, Д. Свободової, Е. Сепір, О. Стишова, М. Тат'євського, М. Хаспельмат, Б. Хофера та ін. Відносно небагато досліджень присвячено вивченню чинників проникнення українських запозичень до польської мови протягом різних історичних епох (Н. Совтис, І. Кононенко, О. Палатовська, В. Русанівський, А. Фаловські, С. Дубіш та ін.). Водночас залишається недостатньо вивченим питання запозичення неологічних українізмів до польської мови на сучасному етапі їх розвитку.

Актуальність теми дослідження зумовлена передусім необхідністю дослідження новітніх змін у польській мові під впливом неологічних українізмів, які в умовах воєнних реалій активно входять до повсякденного лексикону населення, політиків та текстів засобів масової інформації, набуваючи нового значення та конотаційного

забарвлення. Недостатній ступінь вивчення цього питання зумовив написання роботи за відповідною тематикою.

Мета роботи – ідентифікувати та проаналізувати новітні українізми, що були запозичені до текстів польських медіа у зв'язку з воєнними реаліями в Україні.

Для досягнення поставленої мети у ході написання роботи було сформульовано та вирішено такі ключові **завдання**:

- дослідити сутність та різновиди лексичних запозичень;
- узагальнити чинники входження та особливості освоєння іншомовних слів;
- дослідити соціально-історичні та культурні передумови проникнення українських запозичень до польської мови;
- систематизувати та описати методологію використання мовних корпусів у лінгвістичних дослідженнях;
- розробити методику дослідження українських запозичень польською мовою на основі корпусного дослідження та квантитативних методів у лінгвістиці;
- виявити і описати тематичні групи українізмів, які були запозичені польською мовою в умовах воєнного часу;
- за допомогою методів корпусного дослідження перевірити, які неологічні українізми були запозичені до польської мови протягом новітнього періоду з лютого 2022 р. до січня 2023 р.

Об'єктом дослідження є лексичні запозичення. **Предметом** дослідження у межах визначеного об'єкту є запозичені українізми у текстах польських засобів масової інформації.

Матеріал дослідження. Інформаційною базою дослідження слугував моніторинговий корпус польських новин MoncoPL, що постійно доповнюється новими матеріалами. Станом на початок вересня 2023 р. корпус містив близько 447 млн речень, що еквівалентно майже 8,4 млрд слів. Протягом останніх десяти років середньомісячний приріст складав близько 2,7 млн речень.

Перед проведенням корпусного дослідження було сформовано початкову вибірку, що містила близько 300 слів з різною тематичною спрямованістю. За результатами корпусного дослідження з використанням інструментарію на порталі MoncoPL було сформовано фінальну вибірку, що містить понад 100 українізмів, переважно неологічних, які були запозичені до польського медіа-дискурсу протягом лютого 2022 р. – січня 2023 р.

Методи дослідження. Основним інструментарієм проведеного дослідження є методологія мовних корпусів, що охоплює набір різноманітних квантитативних та якісних методів лінгвістичного аналізу. Аналіз *частотності* дозволив визначити динаміку вживання обраних лексем протягом досліджуваного історичного періоду, на основі якого було виявлено неологічні українізми, що увійшли до польської мови протягом 2022 р. – початку 2023 р.. Аналіз *колокацій* дозволив визначити контекстні комбінації обраних лексем, що поглибило розуміння їх конотаційного забарвлення та сфер уживання. Аналіз *конкордансу* було використано для виявлення найбільш типових випадків вживання досліджуваних виразів, що сприяло розкриттю конкретного змісту запозичень українського походження у різних контекстах. Окрім цього, було проведено *якісний аналіз дискурсу* з метою визначення основних наративів і конотацій, що зустрічалися у медійних текстах, а також застосовано методи *лінгвістичного опису*, *лексико-семантичного* і *контрастивного* аналізу при дослідженні спільних і відмінних рис у семантичних структурах польської та української мов.

Наукова новизна. *Уперше* визначено неологічні українізми за різними тематичними групами, які були запозичені польською мовою в новітню епоху, зокрема під час повномасштабних воєних дій. *Удосконалено* підходи до дослідження іншомовних запозичень з використанням методу лінгвістичних корпусів. *Дістало подальшого розвитку* застосування квантитативних методів аналізу медійного дискурсу.

Теоретичне значення дослідження полягає в тому, що воно є певним внеском до теорії і методики дослідження лексичних запозичень на основі мовних корпусів з використанням квантитативних методів.

Практичне значення. Сформовані в роботі теоретико-методологічні положення та практичні підходи до дослідження запозиченої лексики, одержані результати та висновки можна використовувати при викладанні таких навчальних дисциплін, як корпусна лінгвістика, методологія лінгвістичних досліджень та аналіз медійних текстів.

Апробація результатів. Результати кваліфікаційної роботи апробовані на двох міжнародних конференціях: XX Міжнародній конференції з актуальних проблем лінгвістичних досліджень «Мова, думка, людина» (Харківський національний педагогічний університет імені Г.С. Сковороди. Харків, 23-24 березня 2023 р.), присвяченій пам'яті й з нагоди 95-річчя від дня народження доктора філологічних наук, професора Л.А. Лисиченко; Міжнародній науково-практичній відеоконференції “Наративи сучасної України у світовій геополітиці” (Київський національний лінгвістичний університет. Київ, 18–19 травня 2023 р.).

Проблематика, теоретичні положення та практичні результати роботи висвітлені в публікаціях:

Палатовська, О. В., & Сова, Є. С. (2023). Неологічні українізми в мові польських новин: січень 2021 – січень 2023 (корпусне дослідження). *Лінгвістичні дослідження*, (58), 268-279.

Sowa E.S., Palatovska O.V. *Українське неологічне топонімія в польських медіях. "Ad orbem per linguas. До світу через мови". Матеріали Міжнар. наук.-практ. відеоконференції "Наративи сучасної України у світовій геополітиці"*. Київ: Вид. центр КНЛУ, 351-353.

Структура роботи. Робота складається зі вступу, трьох розділів з висновками до кожного з них, висновків до всієї роботи, списку використаних джерел, списку джерел ілюстративного матеріалу та резюме.

РОЗДІЛ 1

ТЕОРЕТИЧНІ ЗАСАДИ ВИВЧЕННЯ ЗАПОЗИЧЕНОЇ ЛЕКСИКИ

1.1. Дослідження іншомовних запозичень у лінгвістиці: коротка історія питання

На сучасному етапі розвитку лінгвістичних наук мова розглядається не як статичний і формалізований об'єкт наукового пізнання, а як динамічна система, що постійно розвивається та є продуктом соціальної, культурної та когнітивної взаємодії. Поруч із постійним оновленням старих елементів новими, у мовній системі відбувається конкуренція між різними елементами за визнання їх нормативними (Жмаєва, 2011, с.48). Саме завдяки постійним перетворенням відбувається розвиток та адаптація мов до сучасних реалій.

У епоху глобалізації одним із основних каналів збагачення мови є її взаємодія з іншими мовами, у процесі чого формуються іншомовні *запозичення*. На думку Jespersen (1923), запозичені слова демонструють домінування народу, мова якого слугує джерелом нових слів для мови-реципієнта, у певній сфері (наприклад, культурі, науці). Це пояснює той феномен, що більшість сучасних запозичень походять саме з англійської мови, а у попередніх століттях, залежно від епохи та домінування тієї чи тієї культури й народу, – з французької, латинської, грецької тощо.

У науковому дискурсі останніх десятиліть розуміння терміну «запозичення» було предметом численних дискусій, внаслідок чого було сформовано різні підходи до його трактування. Для кращого розуміння наукової глибини та сфери застосування даного поняття, розглянемо деякі з загальноприйнятих дефініцій.

У «Сучасному лінгвістичному словнику» А. Загнітко (2020) мовні запозичені визначено як «елемент чужої мови (слово, морфема, синтаксична конструкція тощо), перенесений із однієї мови в іншу в результаті мовних контактів і міжособистісної,

міжнародної комунікації, а також сам процес переходу елементів однієї мови в іншу» (Загнітко, 2020, с. 211).

За Н. Жмаєвою (2011), цей термін у лінгвістиці охоплює комплекс процесів і результатів запозичення, зокрема «факт імплантації», процес адаптації, результати засвоєння мовних елементів, систему усіх етапів мовного запозичення (Жмаєва, 2011, с.51). На основі проведеного аналізу автор формулює цей термін як явище прийняття лінгвістичного матеріалу з інших мов внаслідок дії різноманітних екстралінгвістичних факторів (там само, с. 55). Окрім цього, Н. Жмаєва (2011) розрізняє два основні підходи до розуміння змісту й результату запозичень (там само, с.51-52):

- запозичення як власне перенесення елементів різних рівнів (лексичний, морфологічний, синтаксичний тощо) з однієї мовної системи до іншої;
- запозичення як творча імітація або будь-яка зміна, яка відбувається у мові-реципієнті під впливом зовнішніх імпульсів, наприклад, внаслідок контактів з іншими мовами.

А. Кравчук (2011) зазначає, що термін «запозичення» можна вживати як у широкому (на позначення різних іншомовних елементів у рідній мові, зокрема слів, фразеологізмів, синтаксичних конструкцій, дериватів) й вузькому (стосується лише лексичних запозичень і семантичних кальок) значеннях (с. 137). У широкому сенсі автор розрізняє зовнішні, що виникли в результаті взаємодії з іншими мовами, і внутрішні запозичення, що утворюються під впливом різноманітних традицій народної мови (Кравчук, 2011, с. 137).

Б. Хофер (2002) розглядає запозичення як процес імпортування елементів з однієї мовної системи до іншої в рамках контактів двох культур протягом певного періоду часу (Hoffer, 2001, с.1).

З іншого боку, на думку Е. Гауген (1950), термін «запозичення» неадекватно відображає поняття, для позначення якого він вживається у лінгвістичних науках. Автор аргументує своє твердження тим, що у даному процесі не потрібен дозвіл або інформування «позикодавця» (тобто носіїв мови-донора), а у народу-реципієнта

відсутнє зобов'язання щодо «погашення позики» чи плати за її використання, як у випадку матеріальних предметів (Naugen, 1950, с.211). Серед слів-замінників, рекомендованих автором, можна виділити наступні: адаптація та дифузія. Хоча використання поняття «дифузія» було б не зовсім точним, адже об'єктом запозичення є окремі елементи мови, а не змішування мовних систем. Загалом науковець вирішує дотримуватися традиційного терміну «лінгвістичні запозичення» з огляду на його поширення в науковому дискурсі, наголосивши однак на важливості його коректного розуміння у контексті лінгвістики (там само, с.211-212).

Важливо також зауважити, що групи лексики залежно від сфери вживання суттєво різняться за ступенем запозичення. Наприклад, базова лексика, що стосується ключових сфер людської життєдіяльності, зокрема природи (*дощ, місяць*), дій і станів (*спати, їсти, робити, бути, мати*) тощо, є більш стійкою до впливу іншомовних слів (Hock & Joseph, 1996; цитовано у Zabawa, 2012).

Як описано вище, більшість авторів дотримуються думки, що поняття «запозичення» охоплює не лише результат, а й процес переходу елементів однієї мовної системи до іншої, а власне процес запозичення нагадує не одномоментну подію, а досить тривалий та ітераційний процес адаптації окремих елементів мовою-реципієнтом.

У межах проведеного теоретичного аналізу було виявлено, що термін «запозичення» найчастіше вживається у вузькому значенні, тобто у контексті саме лексичних адаптацій, що є цілком природнім, враховуючи вагому частку даного виду запозичень. У той же час синтаксис і морфологія є більш стійкими до запозичень (Zabawa, 2012). Іменники запозичуються частіше, ніж інші частини мови, адже вони легше адаптуються до структури іншої мови (Haspelmath, 2009; Myers-Scotton, 2006).

Перші відомі дослідження лінгвістичних запозичень сягають ще XIX ст. – першої половини XX ст. Якщо у XVII ст. науковців цікавили передусім внутрішні зміни у мові, то з XVIII ст. учені зосереджувалися на вивченні процесу та результатів запозичень лексики, звукової системи, граматики з використанням методів історичної лінгвістики й порівняльно-історичного мовознавства (Hoffer, 2002).

Один із основоположників порівняльно-історичного методу був датський мовознавець Р. Раск (період наукової діяльності припадає на першу половину XIX ст.). У контексті розробки критеріїв порівняння мов, розрізнення питомої і запозиченої лексики автор дійшов висновку, що у разі наявності одного й того ж слова у декількох мовах, вважається, що воно належить тій мові, де має критично необхідне, більш загальне, давнє і матеріальне значення (Пампура, 2013, с.29). Багато праць даної епохи було присвячено дослідженню походження, взаємодії та шляхів запозичення лексики основних європейських мов, зокрема германських, романських та слов'янських мов, фінно-угорської сім'ї, увага приділялася також турецькій і монгольській мовам (Hoffer, 2002).

Деякі дослідження раннього періоду зосереджувалися на тому, як типи соціальної взаємодії впливають на обсяг запозичень. Було виявлено, що випадкові або пов'язані з комерційною діяльністю контакти не призводили до появи значної кількості запозичень; у той час як тісні контакти між двома народами протягом десятиліть або століть призводили до появи численних запозичень, як у випадку скандинавів в Англії, починаючи з XI століття (Hoffer, 2002, с.3)

Вагомий внесок у розвиток лінгвістичної науки в цілому, та у дослідження запозичень того періоду зокрема, було зроблено засновником структуралізму та одним із засновників сучасної лінгвістики Ф. де Сосюр, важливі наукові й педагогічні здобутки якого узагальнено у праці «Курс загальної лінгвістики». В рамках вивчення лінгвістичної системи автор розрізняв її внутрішні й зовнішні елементи. Наведемо деякі з ідей відомого лінгвіста, що пов'язані із запозиченнями (Сосюр, 1998, с.34-36):

- до поняття «зовнішня лінгвістика» Ф. де Сосюр відносить географічне поширення мов, взаємодію мови з етнологією і політичною історією, запозичення іншомовних слів;
- загалом автор стверджує, що зовнішні чинники є важливою сферою наукових досліджень. Попри їхнє вагоме значення при вивченні історичного розвитку

мови, іншомовні запозичення не є вкрай необхідними при дослідженні внутрішніх аспектів певної мовної системи;

- на думку Ф. де Сосюра, запозичене слово вже не вважається таким, коли вивчається в рамках мовної системи. Як і будь-яке питоме слово, запозичення існують у системі лише завдяки своїм зв'язкам з іншими пов'язаними елементами структури.

Е. Сепір приділив особливу увагу психологічним аспектам, що впливають на інтенсивність запозичення іншомовних слів. Науковець зауважив, що історичні події та культурні взаємозв'язки є важливими, проте не виключними факторами, які визначають обсяг запозичень у мові. Іншим важливим чинником входження іншомовних слів є психологічне відношення мови-реципієнта до лінгвістичного матеріалу. Автор наводить приклад англійської та німецької мов: англійська мова сприяла засвоєнню нових слів із французької та латинської, адже вона є схильною до ємних, збалансованих слів, що відображають завершену ідею, незалежно від кількості складів; у той же час запозичені французькі й латинські слова не закріпилися у німецькій мові, враховуючи її тенденції до аналізу й логічного розкладання складних слів на окремі елементи, що часто відбувається на підсвідомому рівні (Sapir, 2021).

Досить авторитетною у галузі мовознавства є книга Л. Блумфілда «Мова» (англ. *Language*), видана у 1933 році, частина розділів якої присвячена саме запозиченням. Серед іншого автор розглядав особливості інтеграції іншомовних слів до системи мови-реципієнта, зокрема афіксів. Було виявлено, що деякі слова співіснують у паралельних формах – з питомими і запозиченими афіксами. Наприклад, в англійській мові існують дві форми множини для запозиченого слова «*cactus*»: питома – «*cactuses*», запозичена з оригіналу – «*cacti*» (Hoffer, 2002, с.4).

Іншою сферою досліджень Л. Блумфілда були запозичення при взаємодії двох мов всередині одного суспільства чи території (в оригіналі «*intimate borrowing*»). На думку автора, подібна ситуація частіше траплялася через завоювання, ніж внаслідок добровільної міграції. В результаті процес запозичення найчастіше відбувається в односторонньому порядку: від «вищої» або привілейованої мови, якою користується

народ-завойовник (або країна, до якої прибули мігранти), до «нижчої» мови поневоленого народу (або мови іммігрантів). Варто зазначити, що в даному випадку йдеться саме про запозичення «нижчою» мовою за зразком мови-донора тих іншомовних форм і слів, які не пов'язані безпосередньо з культурними аспектами як, наприклад, «*spaghetti*», «*hamburger*», «*delicatessen*». Якщо у «нижчій» мові існують поняття, явища або культурні особливості, які були відсутні у «вищій» мові до їхнього контакту, то процес запозичення може відбуватися у зворотному напрямку – тобто «вища» мова збагачується новими словами з «нижчої» мови (Bloomfield, 1933).

Отже, на ранньому етапі вивчення іншомовних запозичень наукові праці зосереджувалися передусім на порівняльно-історичних аспектах розвитку та особливостей взаємодії різних мов. Важливим напрямом дослідження були також основні культурні, політичні, соціальні і психологічні чинники, які сприяли появі нових запозичень, або навпаки стійкості мови до іншомовних слів.

Широко відома праця Е. Гаугена (1950) під назвою «Аналіз лінгвістичних запозичень» («*The analysis of linguistic borrowing*») є основною точкою відліку, що відображає перехід від раннього етапу дослідження проблеми запозичень (XIX ст. – перша половина XX ст.) до сучасного періоду (з другої половини XX ст.). Ця стаття стала базисом для подальших досліджень у даному напрямі протягом другої половини XX ст. – початку XXI ст. Основною метою автора було розробити більш точний термінологічний апарат для лінгвістичних досліджень запозичень, а також сформулювати основні гіпотези щодо процесу запозичень (Haugen, 1950; Hoffer, 2002). Як було зазначено вище, Е. Гауген (1950) критикує поняття «запозичення», зазначивши, що воно не відображає справжню сутність даного явища й відповідного процесу. Враховуючи широке розповсюдження цього терміну у лінгвістичній науці, автор доходить висновку, що його можна використовувати в наукових працях, але за умови точного пояснення сутності, щоб уникнути невідповідних асоціацій.

Важливим досягненням у праці Е. Гаугена (1950) було пояснення ролі білінгвів у стимулюванні процесу лінгвістичних запозичень. Учений зазначає, що аналіз процесу

запозичень має розпочинатися з аналізу поведінки білінгвів, адже значні лінгвістичні запозичення можливі передусім за умови наявності суттєвої групи людей, які вільно володіють щонайменше двома мовами як своїми рідними та активно ними користуються.

У. Вайнрайх (1953) також розглядає процес лінгвістичних запозичень через призму білінгвізму, який спричинює таке явище як інтерференція, тобто контактування двох систем у мовленні білінгвів, що зумовлює відхилення від мовних норм. Через вільне володіння двома мовами суб'єкти мовлення продукують нові запозичення за рахунок перестановки елементів двох лінгвістичних систем. Автор зазначає, що інтерференція відображає відхилення від поточних норм певної мови, зокрема її фонологічного, лексичного, морфологічного або синтаксичного рівнів (Weinreich, 1953).

Аналіз досліджень, зосереджених на мовних запозиченнях, демонструє, що вони є досить розгалуженими за тематикою та сферою наукових інтересів. Протягом останніх десятиліть запозичення стали розповсюдженим предметом досліджень не лише різноманітних лінгвістичних напрямів, а й суміжних наук, зокрема діахронічного мовознавства, нормативної лінгвістики, соціолінгвістики, археології й культурології (Klapper, Bańko, Svobodová, Rączaszek-Leonardi & Tatjewski, 2016, с.13-14):

- *діахронічне мовознавство*: запозичення є свідченням взаємодії між певними народами, розкривають контекст і період відповідних мовних контактів, відображають історичні зміни у мовних системах;
- *нормативна лінгвістика*: вивчення запозичень у контексті критики загрози, яку вони становлять для розвитку рідної мови внаслідок витіснення питомої лексики, а також їх негативного впливу на стилістику мовлення, проведення ефективної комунікації тощо;
- *соціолінгвістика*: розглядаються явища, які пов'язані з процесом запозичення, зокрема колективна й індивідуальна мовна інтерференція, змішування культурних кодів. Предметом соціолінгвістичних досліджень можуть слугувати, наприклад, мова емігрантів, професійні жаргонізми (мова інвестиційних аналітиків,

інженерів, IT-спеціалістів тощо), взаємодія діалектів однієї мови, використання запозичень із «престижних» мов (латинської, англійської, французької та ін.) для здобуття соціального схвалення;

- *археологія й культурологія*: мовні й літературні джерела документують появу і розповсюдження мовних запозичень, які відображають взаємодію між різними етносами на культурному й побутовому рівнях, поглиблюють розуміння процесу розвитку цивілізацій протягом різних історичних епох.

Протягом другої половини ХХ ст. спостерігалось значне розширення тематики наукових праць та спеціалізація досліджень у сфері запозичень у межах різних лінгвістичних напрямів. Б. Хофер (2002) узагальнив основні напрями таких досліджень і ключові питання, які вони розглядали:

1. Історична лінгвістика й компаративістика: з-поміж інших розглянуто питання змішаних мов або феномен «перемикання мовних кодів» (англ. «codeswitching»); досліджено питання появи й розвитку піджини, від англ. «pidgin», для якого характерні універсалізація і гнучкість синтаксису (слова можуть використовуватися у довільній послідовності без втрати змісту), а також використання обмеженої кількості найбільш поширених і корисних слів з різних мов, тому особливістю піджини є відносно висока частка запозичень.

2. Історії розвитку мов: внаслідок міжнародної торгівлі, міграцій, подорожей, а також діяльності християнських місіонерів європейські мови, передусім англійська, іспанська, французька, нідерландська, німецька мови, впливали одна на одну та інші мови, що призвело до появи англіцизмів, германізмів, інтернаціоналізмів тощо. Дослідженню запозичень з перелічених мов та впливу інтернаціоналізмів присвячена велика кількість праць, адже дана тематика залишається актуальною й донині.

3. Систематизація термінології у дослідженні лінгвістичних запозичень, зокрема поняття «перемикання кодів»: протягом другої половини ХХ ст. було здійснено чимало спроб щодо розробки критеріїв для розрізнення таких понять як перемикання

кодів, змішування кодів, запозичення, перенесення (трансфер), інтерференція та багато інших, зокрема у М. Хеллер (1988), К. Скоттон (1988), С. Поплак (1988).

4. Соціальні та етнічні діалекти: запозичення між діалектами й соціальними класами у контексті взаємодії мови й соціальної поведінки досліджував В. Лабов (1966) на прикладі соціальних і стилістичних варіацій англійської мови серед жителів Нью-Йорка. Автор виявив постійні закономірності у фонологічних змінних залежно від соціальних, стилістичних, етнічних й індивідуальних факторів. Як приклад однієї із соціолінгвістичних закономірностей у В. Лабова (1966) наводиться частота акцентування звуку «r» залежно від ситуації й соціального класу. Він розглядає діалект з невисокою частотою вживання звуку «r», у якому цей звук замінюється подовженням голосного «a», наприклад, у слові «bar» ([baa]). У повсякденних неформальних частка вживання звуку «r» виявилася низькою серед усіх соціальних класів, а при зростанні статусу й формальності мовленнєвої ситуації частка «r» зростала. Цікавим спостереженням було те, що частка «r» в офіційному мовленні у представників менш заможного класу була вищою, ніж у більш заможних мешканців міста (у той час, як менш заможні носії цього діалекту майже не використовували звук «r» у неформальному мовленні). У такий спосіб спостерігалось явище гіперкорекції, коли представники менш заможного класу прагнули використовувати більш престижний стиль мови у формальних ситуаціях, коли приділяли увагу своєму мовленню.

5. Дискурс-аналіз: взаємне збагачення психології, соціології, комунікативної лінгвістики та інших суміжних наук в рамках міждисциплінарного підходу сприяли розвитку такого напрямку як аналіз дискурсу. Протягом останніх десятиліть дослідження мовних контактів перейшли від рівня окремих речень до аналізу особливостей дискурсу в контексті лінгвістичних запозичень. Наукові праці даного напрямку сфокусовані на вивченні дискурсивних маркерів, сигналів, слів-зв'язок, які запозичуються з англійської мови в рамках мовних контактів або при перемиканні мовного коду у середовищі білінгвів.

6. Невербальна комунікація: у центрі уваги науковців даного напрямку знаходяться жести, погляд, проксеміка, дотики та інші форми невербального мовлення, які супроводжують комунікацію. Внаслідок глобалізаційних процесів та стрімкого розвитку телекомунікаційних технологій та індустрії розваг чимало жестів інтерналізуються та запозичуються представниками різних мов і культур (наприклад, жест «ОК»). Хоча у багатьох випадках використання запозичених або інтернаціональних форм невербального мовлення сприяє міжнародній комунікації, іноді це може призводити до непорозуміння чи навіть конфліктів (наприклад, при використанні нетипових форм комунікації для певної культури, зокрема надуживання жестів і дотиків).

Хоча наведений вище перелік не є вичерпним, проте він надає загальне уявлення про ступінь розвитку та різноманіття досліджень у сфері лінгвістичних запозичень, особливо, починаючи з другої половини ХХ ст., що характеризується міждисциплінарним підходом та розгалуженою проблематикою наукових інтересів.

1.2. Поняттєво-термінологічний апарат вивчення запозиченої лексики

Як критично зазначили Е. Гауген (1950), О. Єсперсен (1923), Ф. Коетсем (1988) та інші дослідники, використання термінології у науковій літературі, яка стосується іншомовних запозичень, характеризується непослідовністю та несистемністю. Навіть термін «запозичення» неточно описує сутність цього поняття.

У цій частині дослідження наведено поширені класифікації запозичень за типами, а також представлено ключовий поняттєво-термінологічний апарат, що лежить в основі дослідження цієї тематики.

Перед тим, як перейти до опису видів запозичень, розглянемо ключові елементи даного процесу.

Розуміючи деякі обмеження, пов'язані з поняттям «запозичення» і його конотацією, ми все ж використовуємо цей термін у більшості випадків, враховуючи те,

що він є широко прийнятим у науковій літературі. Для уникнення некоректного розуміння його сутності та деяких метафоричних асоціацій, у пункті 1.1. досить ґрунтовно розкрито зміст цього поняття на підґрунті аналізу фахової наукової літератури. Окрім цього, слово «запозичення» у роботі використовується у широкому сенсі на позначення як власне процесу трансферу елементів однієї мови до іншої, їх адаптації, так і результат даного процесу.

У процесі запозичення можемо виділити такі три основні елементи:

1. *Мова-реципієнт* – та мова, яка адаптує нові елементи іншої мови, передусім лексику.
2. *Мова-донор* – мова, елементи якої переносяться до іншої мови.
3. *Носії* мови-реципієнта та *носії* мови-донора, які взаємодіють у мовних контактах.

Ці елементи лежать в основі категоризації запозичень за Coetsem (1988), що подана нижче.

У процесі запозичень особливу роль відіграють білінгви, які є одночасно носіями мови-реципієнта та мови-донора. *Білінгвізм* – це «практика індивідуального або колективного використання двох мов у межах однієї державної чи соціальної спільноти у відповідних комунікативних сферах» згідно з одним із визначень у словнику А. Загнітко (2020). З білінгвізмом досить тісно пов'язане явище «перемикання кодів» (від англ. «codeswitching»), коли носії декількох мов поєднують їхню граматику, вимову, лексику тощо, що призводить до нових запозичень (Hoffer, 2002).

Унаслідок контактів у мовних системах можуть відбуватися певні зміни, відхилення від поточних норм або «перенесення особливостей однієї мови на іншу», які називаються інтерференцією (Загнітко, 2020).

Мовні контакти можуть:

- 1) сприяти конвергентному розвитку мов;
- 2) зумовити виникнення допоміжних спрощених мов (піджинів або креолізованих мов);

3) Призвести до асиміляція (тобто поглинання) однієї мовної системи іншою. Розглянемо кожен з перелічених вище випадків окремо.

Внаслідок конвергентного типу розвитку мов з'являються *мовні союзи* – типи ареально-історичної спільності мов, яким притаманні достатня кількість подібних ознак і структурних особливостей, що з'явилися внаслідок тривалих та інтенсивних мовних контактів на певній території (Загнітко, 2020; Кочерган 2008).

Якщо при контактах (наприклад, торгівля, міжнародна комунікація, міжетнічні контакти в епоху колоніалізму) відбувалося погане засвоєння мови, то виникали *піджини*, тобто спрощені варіанти поширених мов та/або змішані мови на основі англійської, французької, португальської, іспанської тощо і мов корінного населення певної території. За визначенням А. Загнітко, піджин – це «міжетнічна мова, утворена внаслідок спрощення структури та лексики контактних мов, виконує обмежені функції та не є рідною для жодного мовця» (Загнітко, 2020, с. 556). Піджинам характерний досить невеликий словниковий запас, відсутність багатьох граматичних категорій (рід, число, відмінок). Якщо піджин засвоюється та широко використовується певною спільнотою як рідна мова, розширюючи функції та свою структуру, то можна стверджувати про утворення *креольської мови*, тобто мови, що постала внаслідок креолізації піджину (там само, с. 434-435).

Внаслідок дії соціальних тенденцій (наприклад, престижність певної мови), цілеспрямованої політики або інших чинників домігантна мова може поступово витіснити інші і навіть призвести до їх втрати. Як зазначено у підручнику М. Кочергана (2008), асимільовані або втрачені мови не зникають безслідно. Серед відповідних слідів мови розрізняють субстрат та суперстат (Кочерган, 2008; Загнітко, 2020):

1. Субстрат – мова-підоснова, елементи якої розчинилися в мові, яка над нею нашарувалася; або сліди витісненої мови, яка була поширеною на певній лінгвогеографічній території до приходу носіїв нової мови.

2. Суперстрат – мова-надоснова, елементи якої розчинилися у підлеглий мові, над якою вона нашарувалася. Насамперед це проявляється на фонетичному і граматичному рівнях. Прикладом є латинська мова, яка стала суперстатом для більшості мов Західної Європи.

Сутність поняття «мовні запозичення» розкривається глибше через аналіз класифікації і типології запозичень, а також відповідних прикладів.

Ф. Коетсем (1988) запропонував розділити загальне поняття «*трансферу*» елементів певної мовної системи до іншої на дві підкатегорії, зокрема імпозицію та запозичення, залежно від напрямку запозичення та суб'єкта мовлення. Якщо суб'єкт є носієм мови-реципієнта та використовує іншомовні слова у спілкуванні рідною мовою, то відповідний трансфер лінгвістичного матеріалу є *запозиченням*. Наприклад, англіцизми, використані українцем при спілкуванні рідною мовою є запозиченням за класифікацією Ф. Коетсем (1988). З іншого боку, якщо суб'єктом є носій мови-донора, то перенесення матеріалу з його рідної мови до системи мови-реципієнта визначено у Ф. Коетсем (1988) як *імпозиція*. Наприклад, якщо носій німецької мови розмовляє польською та привносить свої артикуляційні й стилістичні звички або лінгвістичний матеріал до польської мови.

Один з найбільш відомих підходів до класифікації запозичень представлений у праці відомого лінгвіста Е. Гаугена. Згідно з підходом Е. Гаугена (1950), враховуючи інтерпретацію Н. Гудими (2009) та А. Віталіш (2016), іншомовні слова можна розподілити за трьома граматичними типами залежно від ступеня морфологічної субституції:

1. *Запозичені слова або лексичні запозичення* (loanwords) – слова без морфологічної субституції зі збереженим морфологічним складом.
2. *Напівкальки* (loanblends) – слова з частковою морфологічною субституцією та частково збереженим морфологічним складом.
3. *Кальки* (loanshift) – слова з повною морфологічною субституцією, які зберігають лише значення прототипа.

Схематичне представлення даної класифікації подано на рис. 1.1.

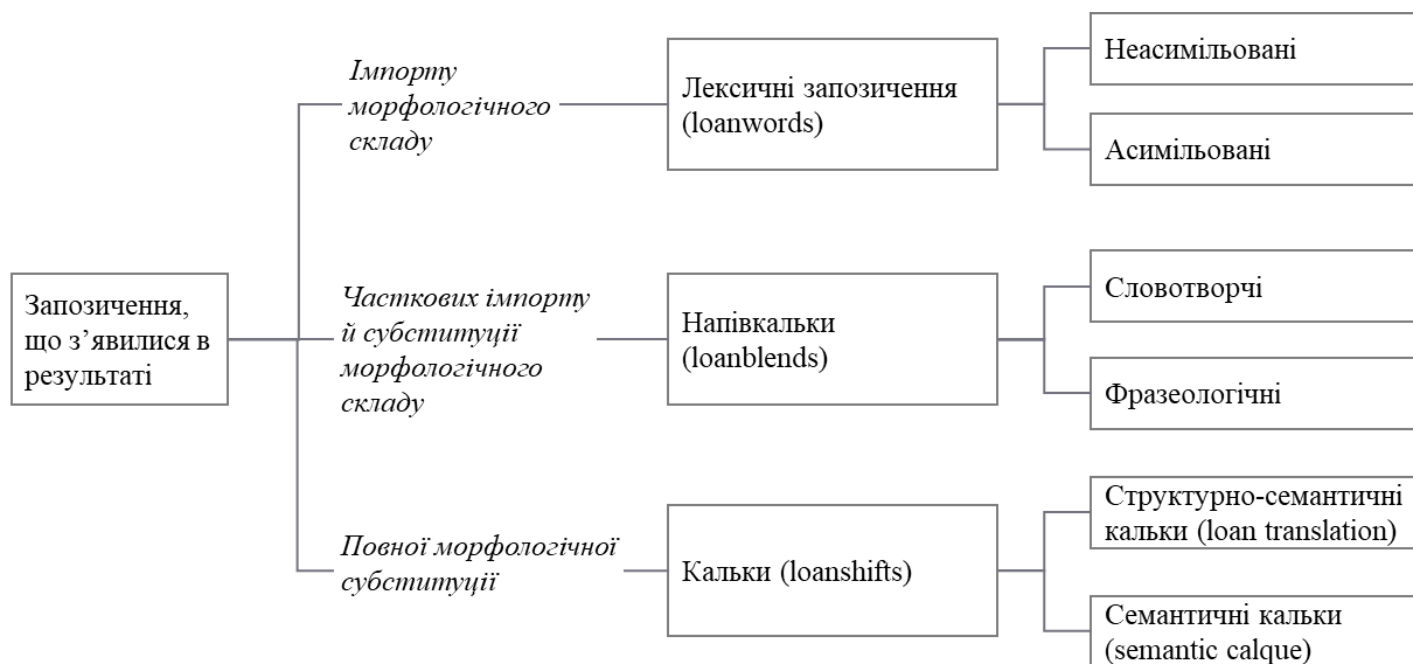


Рис. 1.1. Класифікація запозичень на основі моделі *Naugen (1950)* та доповнення *Witalisz (2016)*

Для кращого розуміння відмінностей між видами запозичень за класифікацією Е. Гаугена (1950) залежно від ступеня субституції, розглянемо деякі приклади для кожного з них:

1. Лексичні запозичення (loanwords) відображають імпортовані слова з найменшим ступенем змін. Серед *неасимільованих* лексичних запозичень у румунській мові, тобто тих, які не зазнали жодних змін, порівняно з оригіналом: *weekend, showroom, all-inclusive, bestseller, business, desktop, rating, freelancer, offshore, knowhow*. Хоча *асимільовані* лексичні запозичення є дуже подібними до оригінального варіанту слова, вони представлені у дещо зміненій формі внаслідок фонологічної інтеграції до мови-реципієнта. Прикладами асимільованих англiцизмів у румунській мові, що були запозичені без морфологічної субституції, є наступні: *administrație, cocteil, jeanși, interviu, fotbal, picur, tramvai, stres, trening* (Greavu, 2014, с.3-4).

2. Напівкальки (loanblends), тобто запозичення із частковою субституцією слова питомою морфемою. Наприклад, при запозиченні деяких англійських слів румунською

мовою суфікс «-у» замінюється на «-ig»: *bossy* → *bossig*, *funny* → *funnig*. У свою чергу суфікс «-ing» замінюється на «-are»: *branding* → *brandare*, *targeting* → *targetare*, *upgrading* → *upgradare*. Іншим прикладом напівкальок є заміна однієї із частин складених слів питомим відповідником: *red brick house* (англ.) → *Redbrickhaus* (нім.); *front garden* (англ.) → *Frontgarten* (нім.); *lunchtime* (англ.) → *Lunchzeit* (нім.); *team-leader* (англ.) → *team-lider* (рум.), *hair-stylist* (англ.) → *hair-stilist* (рум.), *business plans* (англ.) → *business-planuri* (рум.) (Greavu, 2014, с.5-7).

3. Внаслідок повної субституції у мовних кальках (loanshifts) може змінюватися не лише форма, а й первинне значення. Серед типових прикладів семантичних кальок: *magasin* (від нім. на позначення складу) → *magazine* (у англ. мові має декілька значень, у т.ч. періодичне видання, картридж, склад боєприпасів); *fattoria* (від італ. на позначення ферми) → *factory* (від англ. на позначення фабрики). Іншою підкатегорією даного типу запозичень є структурно-семантичні кальки, які відображають іноземні слова, що були повністю перекладені мовою-реципієнтом. Наприклад, англійське слово *skyscraper* (перекладається на українську як *хмарочос*) було запозичено різними мовами з повною морфологічною субституцією: *gratteciel* (франц.), *rascalielos* (исп.), *Wolkenkratzer* (нім.). Прикладами кальок з англійської у румунській мові є наступні: *first lady* → *prima doamnă*, *number one* → *numărul unu*, *second hand* → *la mâna a doua* (Greavu, 2014, с.9).

Хоча підхід до класифікації лексичних запозичень на основі ступеня морфологічної субституції за Е. Гауген (1950) є традиційним та досить розповсюдженим, існують також й інші види критеріїв.

Залежно від мови, з якої походить відповідне слово, розрізняють такі види запозичень: англіцизми (напр., *бестселер*, *лідер*), богемізми (*архієпископ*, *келих*, *робот*, *гаубиця*), галліцизми (*шарм*, *бюлетень*, *фаворит*, *десерт*), іспанізми (*какао*, *танго*, *сієста*, *карамель*), японізми (*карате*, *суші*, *ніндзя*, *соя*, *караоке*) (Sorbet, 2012).

Лексичні запозичення також можна розділити на *прямі*, *непрямі* й *інтернаціоналізми* залежно від близькості до мови-джерела у серії послідовних

запозичень. М. Клапер, М. Банько, Д. Свободова, Й. Рончашек-Леонарді, М. Татйєвський (2016) наводять приклад польського слова *kościół* (церква), з чеської – *kostel*, з німецької – *kástel*, з латини – *castellum*. У вказаній послідовності слів *kościół* (пол.) - *kostel* (чес.) є прикладом прямого або безпосереднього запозичення, а *kościół* (пол.) - *castellum* (лат.) відображає запозичення за посередництвом інших мов (чеської і німецької). До речі цікавим спостереженням у даному випадку є зміна значень запозиченого слова: якщо у польській і чеській мовах при безпосередньому запозиченні значення слова є ідентичним (на позначення церкви), то з латинської мови (першоджерело) воно перекладається як «фортеця». Зазвичай до інтернаціоналізмів відносять слова зі спільним походженням та однаковим значенням, що притаманні багатьом, навіть неспорідненим, мовам (Загнітко, 2020). Джерельною базою інтернаціоналізмів є переважно слова грецького і латинського походження, часто – слова з англійської, французької та німецької мов. Розповсюдження інтернаціоналізмів тісно пов'язано з політичним, економічним, технологічним і цивілізаційним домінуванням відповідних культур протягом різних історичних періодів.

У дослідженнях запозичених слів розрізняють також культурні й базові запозичення (Myers-Scotton, 2006). До *культурних* запозичень відносять слова на позначення нових об'єктів чи концептів, яких досі не існувало у цій культурі чи було нещодавно винайдено. Наприклад, *еспресо, суші, веб-сайт, комп'ютер, автомобіль*. *Базовими (core)* називають ті запозичення, які певною мірою дублюють значення вже наявних питомих слів у мові-реципієнті: наприклад, слово «ОК» замість властивих українській мові слів «*гаразд*» чи «*добре*». Часто базові запозичення виникають у мові-реципієнті за посередництвом білінгвів у процесі переключення мовних кодів (Haspelmath, 2006).

Окрім власне лексичних розглядають й інші різновиди запозичень, вживаючи цей термін у широкому сенсі. М. Клапер, М. Банько, Д. Свободова, Й. Рончашек-Леонарді, М. Татйєвський (2015), а також А. Віталіш (2016) виділяють такі додаткові види запозичень:

- 1) фонемні і просодичні властивості (наприклад, гучність, темп і ритм мовлення, інтонаційні особливості);
- 2) семантичні та структурно-семантичні запозичення;
- 3) морфологічні запозичення;
- 4) прагматичні запозичення – звороти, що мають чітко окреслену прагматичну функцію у конкретній ситуації за зразком іншомовного відповідника. Наприклад: *Як я можу допомогти?* (англ. *How can I help you?*).

5) особливості орфографії та інтерпункції. Іноді використання даного виду запозичень призводить до помилок, наприклад, у випадку, коли наприкінці листа після фрази «з повагою» ставлять кому, що не є характерним для ділової української мови, а є калькою з англійської мови. Іншим прикладом є формат чисел: використання коми для розділення тисяч у числі є некоректним в рамках української стилістики (наприклад, *1,000,000*); правильний варіант написання передбачає використання пробілу (наприклад, *1 000 000*).

6) фразеологічні звороти (наприклад, «бути на одній сторінці» з англ. “*to be on the same page*”);

7) особливості творення синтаксичних конструкцій;

8) мовленнєві шаблони (наприклад, *резюме, кореспонденція, платіжні документи*);

9) моделі мовної поведінки (наприклад, розповсюджені формули ввічливості).

Очевидно, що представлений перелік класифікацій запозичень за різними критеріями не є вичерпним, адже будь-яка категоризація є досить умовною і повинна використовуватися з огляду на предмет дослідження.

1.3. Чинники входження та особливості освоєння іншомовних слів

Значна кількість прямих і опосередкованих запозичень у мовах європейських народів протягом історії була пов’язана передусім з тісними та тривалими контактами

між ними внаслідок наступних причин: військового захоплення територій; мирного співіснування різних етносів в межах однієї держави; різноманітних торгово-економічних, політичних і культурних відносин між країнами (Кочерган, 2008).

Як зазначено у Б. Хофера (2002), при вивченні різних типів соціальної взаємодії на обсяг запозичень було виявлено, що випадкові або пов'язані з комерційною діяльністю контакти не призводили до появи значної кількості запозичень; у той час як тісні й тривалі контакти між двома народами протягом десятиліть або століть призводили до появи численних запозичень, як у випадку скандинавів в Англії, починаючи з XI століття.

Відповідно, мовні контакти є одним з первинних рушіїв входження й освоєння мовних запозичень, а також зовнішнім стимулом інших важливих процесів розвитку мови. М. Кочерган виділяє наступні ключові процеси, що активізуються в результаті мовних контактів:

- 1) лексичні й фразеологічні запозичення з різним ступенем морфологічної субституції;
 - 2) засвоєння артикуляційної специфіки іншої мови;
 - 3) зміна наголосу;
 - 4) зміни у граматичній системі, зокрема синтаксисі;
 - 5) запозичення суфіксів і префіксів, що спричинює зміну в словотворі
- (Кочерган, 2008).

Упродовж XX та XXI ст. внаслідок стрімкого розвитку технологій і глобалізаційних процесів спостерігалось подальше прискорення темпів збагачення мов за рахунок зовнішніх запозичень. Як зазначено у Б. Хофера (2002), О. Стишова (2017), Н. Жмаєвої (2011), серед ключових чинників екстралінгвального характеру можна виділити наступні:

- 1) винайдення радіо й телебачення, що сприяло появі нового типу культурної взаємодії, відносно швидкого поширення та засвоєння комунікативних елементів різних мов;

2) глобалізація економіки й культури в умовах переходу до інформаційного суспільства, у т. ч. завдяки впливу медіа. Одним із вираженням цього процесу є реклама й маркетинг, за посередництвом яких споживачі дізнаються про іншомовні назви, поняття, явища;

3) загальна демократизація суспільно-політичного життя, зниження впливу цензури у роботі ЗМІ;

4) активізація економічної, наукової, політичної, культурної і технологічної співпраці між Україною, США, європейськими та іншими країнами, зокрема в межах євроінтеграційних процесів, діяльності міжнародних корпорацій, бізнес-проектів;

5) доступність швидких видів транспорту широким верствам населення, зокрема літаків й автомобілів, що, разом з пом'якшенням адміністративних обмежень, сприяло розвитку ділових подорожей і туризму. Стрімкий розвиток індустрії подорожей та відповідної інфраструктури стимулювали більш тісні та інтенсивні контакти між носіями різних мов і культур, що, у свою чергу, сприяло появі нових лінгвістичних запозичень. Окрім цього, з кінця ХХ ст. спостерігалася активізація міграційних процесів, передусім з менш розвинених до більш розвинених країн. У свою чергу це сприяє зростанню кількості білінгвів, адже нові покоління у сім'ях мігрантів часто володіють щонайменше двома мовами як своїми рідними;

6) популяризація і підвищення загального рівня володіння іноземними мовами (передусім англійською) у суспільствах;

7) мовні експерименти й мода на екзотичні слова, що дозволяють суб'єктам мовлення виражати свою особистість. Іншомовний варіант слова нерідко сприймається як більш престижний;

8) формування термінологічної системи, яка використовується представниками певного професійного середовища чи економічної галузі по всьому світу.

Окрім чинників входження іншомовних слів до мови-реципієнта, важливо також розглянути й внутрішні причини використання запозиченої лексики замість питомих

слів суб'єктами мовлення. Хоча у багатьох випадках лінгвістичні запозичення є необхідністю, наприклад, при відсутності відповідних понять у культурі іншого народу, серед інших типових причин *використання* непитомої лексики є наступні: прагнення людей відповідати трендам певної епохи чи соціальної групи; недостатній рівень володіння певною мовою, а відтак і необхідність заміни слів іншою мовою; уникнення додаткових перекладацьких зусиль при цитуванні або використанні вузькоспеціалізованої лексики (Jespersen, 1923).

Н. Жмаєва (2011) наводить приклад впливу глобалізаційних та євроінтеграційних процесів на формування єдиного культурного й комунікаційного простору в межах співтовариства, що призводить до поступової конвергенції словникового запасу європейських мов на основі запозичень з англійської мови.

При вивченні процесу і чинників освоєння лексичних запозичень важливо розглянути й деякі аспекти, що визначають *ступінь адаптації* іншомовних слів до системи мови-реципієнта. Відповідно до М. Хаспельмат (2009), основними чинниками є наступні: ступінь знайомості з мовою-джерелом серед носіїв мови-реципієнта, ставлення суб'єктів мовлення до мови-донора, період запозичення (нещодавно прийняті слова або відносно давні запозичення). Якщо мова, з якої походить слово, є досить відомою для більшості суб'єктів мовлення та позитивно ними сприймається, то дуже імовірно, що запозичене слово буде збережено у близькому до оригіналу вигляді (і навпаки). Це стосується як фонологічної, так і орфографічної адаптації. Прикладом можуть слугувати слова грецького або латинського походження, запозичені англійською мовою: *phenomenon/phenomena*, *curriculum/curricula*, *data/datum*. Іншим важливим чинником є новизна: нещодавно запозичені слова часто приймаються в оригінальній формі, адже вони ще не пройшли етап часткової або повної адаптації. Як правило, такі слова краще сприймаються молодим поколінням, у той час як старше покоління може критично ставитися до відповідних інновацій (Haspelmath, 2009). Загалом *лексико-семантичні* запозичення мають один з найвищих рівнів проникнення, порівняно з іншими типами запозичень (Кочерган, 2008).

Також досить високим ступенем проникнення характеризується й *синтаксис*, що пояснює значну подібність синтаксичних систем різних мов світу. Відносно нижчий рівень адаптації характерний для *фонологічних* запозичень. Зазвичай іншомовні слова пристосовуються до фонетики мови-реципієнта внаслідок використання питомих звуків-замінників. Однак при масовому запозиченні слів і тривалих контактів з мовою-донором можуть статися й більш суттєві зміни у фонологічній системі мови-реципієнта, зокрема й поява нових звуків. Яскравим прикладом є грецький звук [ф], який не був притаманний мовам східних слов'ян у період прийняття ними християнства. На ранньому етапі освоєння запозичених грецьких слів, використовувалися звуки-замінники: [п], [т], [хв]. Коли запозичення слів грецького походження стало масовим, то іншомовний звук [ф] почав поступово засвоюватися спочатку представниками інтелігенції та освіченими людьми, а дещо пізніше – іншими верствами населення.

Морфологічному рівню притаманна найвища стійкість до впливу інших мов. Вважається, що внаслідок мовних контактів морфологічна система навпаки схильна до спрощення, що ілюструється прикладом піджинів і креолізованих мов.

Отже, ступінь проникності має зворотній зв'язок зі ступенем системності мови або певного її структурного рівня. Дана закономірність пояснює, чому лексико-семантичний є менш стійким до впливу інших мов і, відповідно, більш відкритий до іншомовних запозичень (Кочерган, 2008).

Відповідно до М. Клапера, М. Банько, Д. Свободової, Й. Рончашек-Леонарді, М. Татйєвського (2016), адаптація іншомовного слова мовою-реципієнтом є досить тривалим процесом, що по-різному відображається на різних рівнях мови. Найшвидше адаптація запозичених слів до правил мовної системи відбувається на фонетичному рівні, дещо повільніше – на орфографічному й граматичному. Протягом певного часу можуть співіснувати кілька форм запозичених слів з різним ступенем адаптації. Особливо це притаманно для слів, які містять літери, що не мають відповідники у мові-реципієнті. Загалом автори зазначають, що процес запозичення може бути описаний у вигляді параболи: на початку запозичене слово є максимально близьким до оригіналу за

вимовою й написанням; пізніше з'являється досить велика кількість фонетичних, граматичних і орфографічних варіантів з різним ступенем адаптації; на завершальному етапі варіативність значно зменшується під впливом кодифікації норм вживання лексичних одиниць у словниках. Іноді адаптоване слово може бути прийняте до нормованої мови у декількох варіантах, що різняться ступенем офіціальності й конотації. Наприклад, слово «*jazz*» мало два варіанти вимови у польській мові після запозичення: [джаз] та [джез]. Наразі у польській мові використовується другий варіант вимови. Окрім фонетичної варіативності, існувало три варіанти написання даного слова: *jazz*, *dżez*, *dżaz*. Відповідно до словника польської мови PWN, на сучасному етапі розвитку мови допускається вживання двох варіантів – *jazz* та *dżez*, що узгоджується з фонетичним варіантом.

Н. Жмаєва (2011) також розглядає запозичення як тривалий процес в контексті соціалізації прийнятої лінгвістичної інновації в рамках різних мовленнєвих середовищ. При первинній соціалізації мовна інновація набуває узуального статусу. У разі її активного поширення внаслідок міжколективної взаємодії та подальшого розповсюдження у широкому комунікативному просторі, прийнята інновація може змінюватися й асимілюватися у різних підсистемах. Внаслідок процесу поширення й асиміляції може з'явитися декілька фонетичних, лексико-семантичних і граматичних варіантів відповідного запозичення.

Отже, входження й освоєння іншомовних слів є багатограними та тривалими процесами, що залежать від низки чинників, зокрема ступеня подібності між мовою-донором і мовою-реципієнтом, ступеня стійкості елементів певного мовного рівня до впливу інших систем, низки соціолінгвістичних факторів.

Висновки до розділу 1

На думку багатьох дослідників, мова є складною та динамічною системою, що постійно розвивається та є результатом, соціальної, когнітивної та культурної взаємодії. Одним із основних каналів збагачення мовної системи є іншомовні запозичення. Хоча питання обґрунтованості терміну «запозичення» було досить дискусійним в лінгвістичному дискурсі ХХ ст. через те, що це поняття не відображає його сутність, все ж більшість дослідників використовує саме термін «запозичення слів» або його замінники з огляду на їх поширеність у науковому середовищі.

У межах аналізу наукової літератури з теми дослідження було узагальнено різні трактування сутності лінгвістичних запозичень: цей термін відображає не лише *результат*, а й *процес* прийняття й адаптації елементів однієї мовної системи (мови-донора) до іншої (мови-реципієнта). Існує безліч різних підходів до класифікації мовних запозичень залежно від походження, близькості до мови-джерела, ступеня морфологічної субституції тощо. Хоча більшість запозичень охоплюють лексико-семантичний рівень, трансферу підлягають елементи на різних мовних рівнях, зокрема фонема, орфографія й інтерпункція, фразеологічні звороти, синтаксичні особливості тощо.

Основними чинниками входження запозичень є тривалі й тісні мовні контакти, а також мовленнєва діяльність білінгвів. Протягом останніх десятиліть масштаб входження та темпи освоєння запозичень прискорилися з огляду на низку екстралінгвістичних чинників, зокрема інтеграційні процеси у сфері економіки, політики і культури, глобалізаційні процеси в інформаційно-культурному просторі, стрімкий технологічний розвиток, загальну демократизацію суспільно-політичного життя, розвиток транспортної інфраструктури й популяризацію туризму тощо.

На підґрунті аналізу лінгвістичних досліджень з'ясовано, що різні елементи мовної системи характеризуються різним ступенем стійкості до зовнішнього впливу, зокрема й прийняття іншомовних запозичень. Отже, запозичення іншомовних слів є тривалим процесом, в межах якого нові елементи, що характеризуються високим

ступенем варіативності вживання, проходять декілька етапів адаптації до мови-реципієнта під дією низки соціолінгвістичних чинників.

РОЗДІЛ 2

МЕТОДОЛОГІЯ ВИКОРИСТАННЯ МОВНИХ КОРПУСІВ У ЛІНГВІСТИЧНИХ ДОСЛІДЖЕННЯХ

2.1. Мовні корпуси як інструментарій новітніх лінгвістичних досліджень

Корпусна лінгвістика є відносно новим напрямом у прикладному мовознавстві, що досліджує великі масиви структурованих текстових даних з метою виявлення певних мовних закономірностей, тестування лінгвістичних гіпотез, вияву своєрідних властивостей певної мови, що зумовлені стилістичними нормами, особливостями мовленнєвого вживання тощо. Термін «корпус» у даному значенні з'явився у 60-ті роки ХХ ст., що у т.ч. пов'язано зі стрімким розвитком комп'ютерної техніки і, відповідно, появою перших комп'ютеризованих корпусів, зокрема в Університеті Брауна, обсягом близько одного мільйона слововживань (Загнітко, 2020). Однак корпусна лінгвістика зародилася як сукупність методів лінгвістичного дослідження ще наприкінці ХІХ ст. На ранньому етапі розвитку, дослідники займалися формуванням доелектронних корпусів, що використовувалися для вивчення текстів, аналізу граматичної структури мови, складання словників. Узагальнивши результати різних теоретичних досліджень, Т. Бобкова (2014) виділяє період ранньої корпусної лінгвістики (кінець ХІХ ст. – 60-ті роки ХХ ст.) і сучасний період комп'ютеризованої корпусної лінгвістики (період з 60-тих років ХХ ст.).

У широкому значенні під терміном «корпус» в лінгвістиці розуміють будь-яке зібрання текстів, що використовуються з метою дослідження мови. У вузькому значенні корпус – це електронна база текстів визначеного спрямування, що характеризує певну мову (Бобкова, 2014).

Для забезпечення об'єктивності, надійності та цілісності результатів дослідження необхідно, щоб лінгвістичний корпус відповідав низці ознак, зокрема: збалансованість, репрезентативність, порівнюваність, автентичність, тобто відображення природної

мови. Хоча перераховані вище ознаки допомагають відрізнити лінгвістичний корпус від збірки електронних текстів і бібліотек, однак багато сучасних корпусів не відповідають одночасно всім цим характеристикам (Жуковська, 2020).

На ранніх етапах розвитку корпусна лінгвістика формувалася переважно як методика, а не окремий напрям мовознавства, що було зумовлено її практичною орієнтацією та меншим зосередженням на розробці теоретичних положень. У процесі розвитку цей напрям збагатився науковим інструментарієм з інших дисциплін, що мають суміжний предмет дослідження. Сучасну корпусну лінгвістику можна охарактеризувати як *міждисциплінарну* галузь науки, яка поєднує власне лінгвістичну теорію, когнітивістику, інформаційні та комп'ютерні технології, статистику й математику тощо. Як зазначає Т. Бобкова (2014), за об'єктом дослідження корпусну лінгвістику можна віднести до гуманітарних наук, а за методологією – до точних.

Завдяки міждисциплінарності та розвиненому комп'ютерному і математичному інструментарію корпусні лексикологічні дослідження охоплюють широкий спектр тем, серед яких, згідно з Г. Лук'янець (2012):

- 1) аналіз частотності вживання лексичних одиниць;
- 2) вивчення контекстів та значення слів;
- 3) дослідження екстралінгвістичних характеристик тексту, зокрема соціальних, психологічних, професійних, географічних;
- 4) встановлення відношень між значеннями слів;
- 5) дослідження змін у словниковому запасі мови, зокрема під впливом запозичень і неологізмів.

Іншим важливим аспектом корпусного методу є те, що він має на меті досягнути *об'єктивного* вивчення мовного матеріалу, абстрагувавшись від суб'єктивного ставлення науковця до предмета дослідження. У центрі уваги дослідника перебуває не формалізована мова як система, а процес комунікації та його результату у природному вигляді з усіма особливостями та недоліками. Відповідно, серед основних методів

дослідження – *кількісний аналіз*, зокрема: аналіз частотності, колокацій, конкордансу, кореляції, варіативності слововживання, а також статистична перевірка гіпотез тощо.

Особливістю наряду є його *емпіричний* характер, адже аналіз мовних корпусів дозволяє відтворити лінгвістичний експеримент, мінімізуючи суб'єктивізм, а також соціальні, наукові, політичні й когнітивні упередження (Жуковська, 2020).

Окрім очевидних переваг корпусного методу, важливо також враховувати ключові обмеження, до яких, на думку Дж. Беннет (2010) і А. Стефановіч (2020), відносять:

1. Корпуси *не містять негативних доказів*, тобто вони не можуть підтвердити, що якийсь об'єкт (приклад слововживання, конструкція, вираз) є неможливим або неправильними, а лише демонструють, які об'єкти наявні у даному корпусі. Варто також звернути увагу, що й, у протилежній ситуації, не завжди наявність певних об'єктів в корпусі підтверджує їх коректність з боку лінгвістики.

2. Методика корпусної лінгвістики дозволяє прослідкувати лише факт існування певних мовних явищ. Вона не має на меті пояснити, *чому* щось існує у заданій формі. Для встановлення причинно-наслідкових зв'язків необхідно застосувати інші інструменти лінгвістичного аналізу, зокрема й професійне судження або інтуїцію. Зокрема для аналізу когнітивних, семантичних тощо аспектів мови, дослідник має самостійно їх інтерпретувати.

3. Будь-який корпус (навіть найбільші відомі корпуси у світі) містить *обмежений обсяг* даних і технічно не може відобразити реальність у повному обсязі, наприклад, усі випадки вживання певної граматичної конструкції чи словосполучення, або всі формальні та неформальні варіанти певного виразу. Наприклад, А. Стефановіч (2020) наводить приклад конструкції з британського варіанту англійської мови («it doesn't matter the ...»), яка не зустрічається жодного разу в Британському Національному Корпусі, що містить близько 100 мільйонів слів.

Одним із центральних завдань корпусної лінгвістики є визначення лінгвістичних закономірностей та перевірка наукових гіпотез на основі вибірки реальних даних із заданими характеристиками. Розглянемо декілька важливих видів аналізу на основі

методів корпусної лінгвістики, зокрема: аналіз частотності (описано детально у пункті 2.2), колокації, усталені комбінації лексичних одиниць, конкорданс.

Аналіз колокацій дозволяє виявити статистичні тенденції та імовірність сполучуваності мовних одиниць з іншими словами. Дж. Беннет (2010) пояснює колокацію на прикладі англійської мови, використовуючи такі слова, як «deal» (укр. «угода», «домовленість», «справа»). У дужках подано невичерпний перелік значень даних слів. На основі аналізу колокацій дослідник виявив, що слово «deal» у формі іменника найчастіше вживається з такими словами, як «big» (*big deal* – подія чи ситуація, що мають велике значення), «great» (*great deal* – часто стосується великої кількості), «good» (*good deal* – зазвичай стосується перемовин, торгів). Г. Кеннеді (1991), використовуючи корпусний метод, дослідив колокації таких близьких за значенням слів як «between» (укр. «між») та «through» (укр. «через», «між»). Автор визначив, що слово «between» найчастіше вживається після наступних слів: *difference, relationship, distinction, gap, agreement* тощо; а слово «through» після таких слів: *go, pass, come, be, and, get* тощо.

У межах аналізу усталених комбінацій (від англ. «lexical bundles») досліджують повторювані послідовності трьох та більше слів. Дж. Беннет (2010) наводить приклад фраз «*Do you want me to*» (укр. «хочеш, щоб я»), «*I don't know what*» (укр. «я не знаю, що»). На відміну від ідіоматичних виразів, усталені комбінації слів: мають значення, яке можна встановити зі змісту їхніх складових; є незавершеними фразами; можуть бути статистично визначеними. Одним з типів усталених комбінацій є граматичні конструкції що складаються зі сталого фрейму та набору змінних, які його оточують. На основі корпусу сучасної англійської мови в американському варіанті, Дж. Беннет (2010) встановив, що для позначення майбутнього часу найчастіше використовують такі конструкції слів: *is ... to, is going to, is likely to, is expected to, is supposed to, is due to*. Відповідно, в цьому випадку *is...to* є фреймом, а *going to, expected to, supposed to* тощо виконують роль змінних, які використовуються у комбінації з усталеною конструкцією *is...to*.

Фразеологічний аспект корпусної лінгвістики також охоплює вивчення *конкордансу*, тобто найбільш типових випадків вживання слів у певних контекстах, залежно від яких розкривається конкретний зміст певного слова. Дослідження конкордансу має не лише теоретичну, а й практичну користь, зокрема у процесі вивчення мови. С. Ханстон (2022) наводить приклад англійських слів «interested» і «interesting», які часто некоректно вживаються. Дослідивши випадки вживання даних слів у різних контекстах, автор визначив, що слово «interested» найчастіше використовується у конструкції «someone is interested in something» (укр. «хтось зацікавлений у чомусь»), а слово «interesting» майже завжди вживається перед іменником у конструкції «an interesting thing» («цікава річ»).

Іншим напрямом застосування методології корпусних досліджень є лексикограматика. Одна із ключових ідей лексикограматики – це існування єдності між лексичними одиницями й граматичними моделями. Наприклад, в англійській мові слова «know» (знати), «matter» (означати, значити), suppose (припускати) вживаються у теперішньому часі у близько 80% випадків; слова «know» (знати), «think» (думати) найчастіше вживають разом зі словом «that» (що); «like» (подобатися), «want» (бажати, хотіти), «seem» (здаватися) зазвичай вживають з прийменником «to» (Bennet, 2010; Viber et al., 1999).

Не менш цікавими питаннями є типи лінгвістичних корпусів та сфери застосування інструментарію корпусної лінгвістики на практиці.

Відповідно до класифікації С. Ханстон, серед найбільш поширених типів лінгвістичних корпусів можна виділити наступні (Hunston, 2022, p.14-16):

1. *Спеціалізований корпус* репрезентує певний тип текстової інформації, зокрема статті у журналах, академічні статті у певній галузі знань, лекції, студентські есе, неформальні діалоги між колегами. Мовознавці часто самостійно формують корпус із обраним ступенем спеціалізації для відображення досліджуваних стилістичних, структурних і змістових характеристик тексту, враховуючи ряд параметрів, зокрема часовий проміжок, тематику, соціальний статус суб'єктів мовлення тощо.

2. *Загальний (референційний) корпус* – великий за обсягом (зазвичай містить десятки мільйонів слів) та варіативністю набір текстових даних, що репрезентує широкий спектр мовного матеріалу різного спрямування на обраному синхронічному зрізі. Він часто слугує основою для створення граматик, словників, навчальних матеріалів, перекладу, а також використовується як база порівняння з більш спеціалізованими корпусами.

3. *Зіставні корпуси* набору мов або різних варіантів однієї мови. Такі корпуси мають подібну структуру, наприклад, з однаковими пропорціями публіцистичних текстів, романів, діалогів, наукових статей. Вони використовуються для порівняння особливостей, спільних і відмінних рис обраних мов, зокрема в рамках перекладацької чи педагогічної діяльності.

4. *Паралельні корпуси* – тексти, що були одночасно створені або перекладені кількома мовами (наприклад, регуляторні акти ЄС, що видаються мовами країн-членів). Такі корпуси є корисними при вивченні еквівалентних виразів у різних мовах або відмінностей між ними.

5. *Навчальний корпус* – набір текстів (наприклад, творів), що був створений групою людей, які вивчають певну мову. Основною метою дослідження такого корпусу є визначення відмінностей між різними учнями, а також між учнями й носіями мови, що є також корисним при розробці навчально-методичних матеріалів.

6. *Педагогічний корпус* – набір усіх навчальних матеріалів, зокрема підручників, статей, книг, відеозаписів тощо, які використовувалися студентами у різних контекстах. Такий корпус аналізують для перевірки, наскільки навчальні матеріали відповідають природнім характеристикам мови, що вивчається, а також ступінь ефективності засвоєння іншомовного матеріалу.

7. *Історичний або діяхронічний корпус* – збірка текстів з різних історичних періодів, що репрезентує розвиток певних аспектів мови протягом обраного часового проміжку.

8. *Моніторинговий корпус* – корпус, що регулярно актуалізується новим матеріалом з певною періодичністю (щоденно, щомісячно, щорічно), та використовується з метою вивчення поточних змін у мові.

Корпусний підхід надає важливий інструментарій для практичних лінгвістичних досліджень. Висновки, що базуються на основі аналізу лінгвістичних корпусів, є більш надійними й об'єктивними порівняно з суб'єктивними висновками на основі інтуїції носіїв мови, незалежно від їхнього індивідуального досвіду. Наприклад, часто лінгвісти або вчителі не можуть пояснити, чому певне словосполучення є більш доречним і природним, ніж інше, пояснюючи свій вибір переважно «відчуттям мови», «кращим звучанням», «натуральністю або питомістю». Отже, використання лише мовної інтуїції є менш ефективним порівняно із комп'ютеризованим корпусним аналізом у судженнях щодо таких питань як: колокації, частотність вживання, дослідження конотаційних забарвлень лексем, фразеологічних особливостей сталих зворотів. З іншого боку, не можна недооцінювати мовну інтуїцію та експертні судження, адже вони дозволяють екстраполювати узагальнення, що були одержані на підґрунті мовного корпусу.

2.2. Частотність та квантитативні методи у дослідженнях лексики

Як напрям мовознавства, що характеризується міждисциплінарним підходом, корпусна лінгвістика покладається на кількісні та емпіричні методи дослідження, запозичені зі статистики, комп'ютерних наук, психології тощо.

Існує чимало технік і методів, які застосовуються до лінгвістичних корпусів залежно від предмету й цілей дослідження. Розглянемо традиційні та найпоширеніші інструменти для аналізу тексту корпусу, як узагальнено на рис. 2.1.

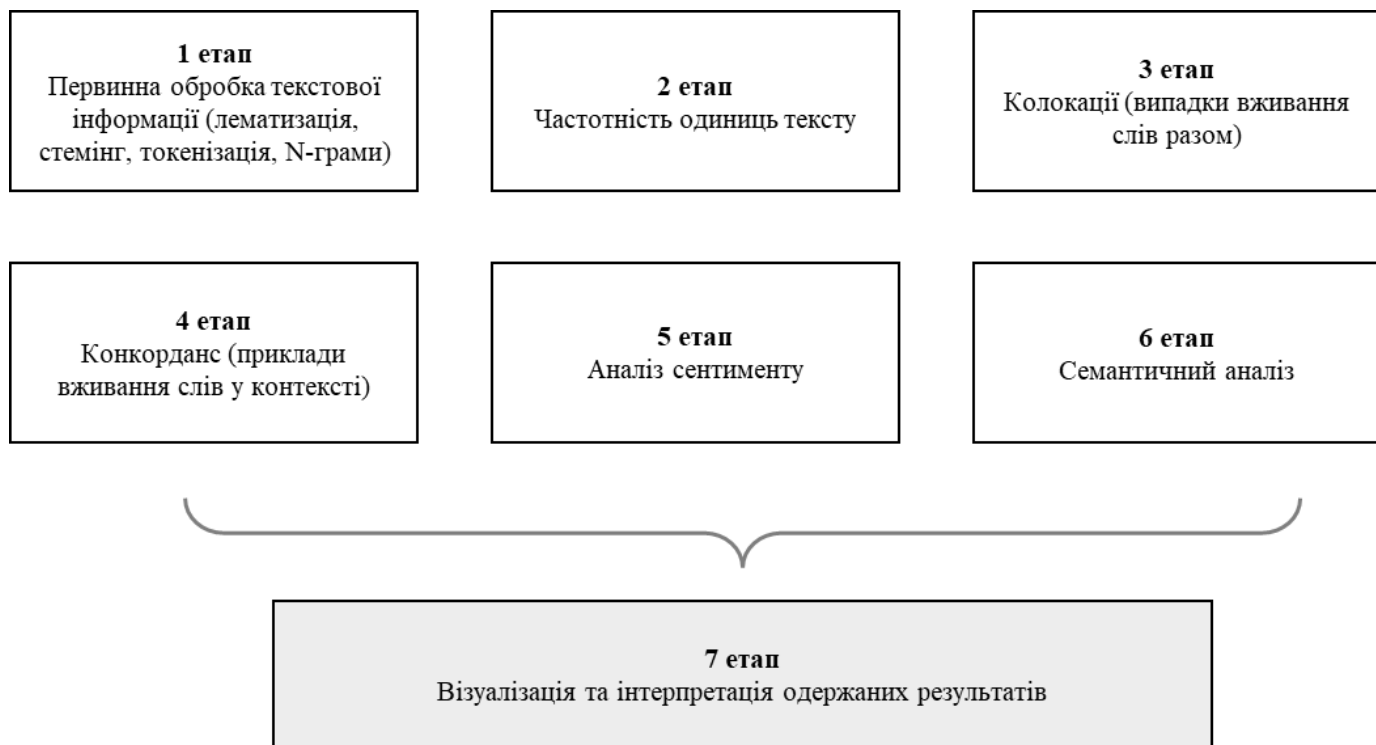


Рис. 2.1. Основні аналітичні методи, що використовуються у корпусних дослідженнях.

Для аналізу мовних корпусів з використанням комп'ютерних методів необхідно спочатку здійснити первинну обробку тексту та його нормалізацію. Цей процес складається з декількох взаємопов'язаних етапів:

1. *Токенизація* – поділ тексту на окремі елементи (зазвичай проводиться на рівні орфографічних слів).

2. *Зміна регістру літер*: усі великі (заголовні) літери замінюються на малі, тобто усувається розрізнення між словами залежно від розміру літер. Наприклад, словосполучення «Мовний корпус», «МОВНИЙ КОРПУС», «Мовний Корпус» зводяться до еквівалентних, адже після відповідної трансформації набувають вигляду «мовний корпус».

3. Ідентифікація та видалення *stop-слів*, до яких зазвичай відносять артиклі (характерно для багатьох західноєвропейських мов), службові частини мови, присвійні

займенники тощо, які не є носіями семантичного значення для цілей корпусного аналізу тексту, застосування алгоритмів машинного навчання. Варто зауважити, що, залежно від цілей, методів і предмету аналізу, одне й те саме слово може бути класифіковано як «стоп-слово» або як важливий елемент тексту із семантичним значенням.

4. *Стемінг* – алгоритмічний процес скорочення слів до їхньої основи шляхом відкидання допоміжних частин слова, зокрема суфіксів і закінчень (Глибовець & Точицький, 2017). Наприклад, слово «швидкий» зводиться до основи «швидк». Після алгоритмічної обробки на основі методу стемізації слово може відрізнитися від свого морфологічного кореня.

5. *Лематизація* – алгоритмічний процес зведення словоформ до словникової форми або морфологічного кореня (Загнітко, 2020). Відповідно, відбувається процес групування відмінюваних форм слова за спільною лемою (словниковою формою) з метою їх подальшого аналізу як єдиного елемента. Алгоритми лематизації зазвичай дозволяють одержати більш точні результати, ніж стемінг, проте потребують значних ресурсів, адже спираються на лексикографічний аналіз, враховують ширший контекст слова.

Послідовності з n елементів (токенів) з певної вибірки тексту називаються N -грамами, що використовуються у різноманітних видах комп'ютеризованого лінгвістичного аналізу на основі мовних корпусів, зокрема при аналізі настрою.

У такий спосіб нормалізація (зокрема стемізація або лематизація) дозволяють зменшити повторюваність слів у різних формах, згрупувавши їх за певним алгоритмом, та одночасно зберегти семантичну структуру тексту. Трансформований текст формує основу для подальших етапів аналізу, зокрема частотності, колокації, конкордансу тощо.

У багатьох сучасних комп'ютеризованих корпусах із розвинутим користувацьким інтерфейсом частина технічних процедур, описаних вище, вже вбудована у відповідну програму та реалізується автоматично, що значно спрощує для дослідників процес власне лінгвістичного аналізу тексту.

Базовим і одним з найбільш розповсюджених інструментів корпусної лінгвістики є *списки частотності* токенів/слів, що представлені у формі таблиці. У найпростішому вигляді таблиці частотності складаються з двох стовпців: в одному стовпці представлено перелік найбільш уживаних токенів, а у другому – частотність їх уживання у відповідному тексті. Показник частотності може ґрунтуватися на *абсолютній* кількості випадків уживання певного токена або бути розрахований у *відносному* вимірі як питома частка вживання певного токена від загальної кількості випадків. Показник відносної частотності дозволяє порівнювати розподіл токенів за вживанням у різних за обсягом корпусах. Ілюстративний приклад списку частотності представлено у табл. 2.1.

Таблиця 2.1.

Ілюстративний приклад таблиці частотності

Токен	Відносна частотність, %	Н-грама з двох елементів	Відносна частотність, %
такий	18	такий що	20
Що	15	до того	15
До	14	можна зробити	12
можна	12	бути як	11
Бути	11	робити це	10
робити	10	йти до	9
Йти	7	говорити про	8
говорити	6	окремий випадок	6
окремий	5	зокрема такі	5
зокрема	2	може бути	3
Всього	100	Всього	100

Джерело: розроблено автором на основі прикладу Gries (2009).

Примітки: дана таблиця є ілюстративною з умовним розподілом слів за частотністю.

Відповідно до Ш. Гріес (2009), таблиці частотності можуть бути також побудовані у вигляді переліку слів з оберненим порядком літер, що дозволяє проаналізувати розподіл слів із певним суфіксом (наприклад, «-ly»), як представлено у табл. 2.2.

Таблиця 2.2.

Ілюстративний приклад таблиці частотності зі зворотним порядком літер у слові

Слово	Слово у зворотному порядку	Абсолютна частотність
dreadfully	yllufdaerd	18
peacefully	yllufecaep	15
gracefully	yllufecarg	14
resourcefully	yllufecruoser	12
gleefully	yllufeelg	11
woefully	yllufeow	10
fully	ylluf	7
hopefully	yllufepoh	6
carefully	ylluferac	5
purposefully	yllufesoprup	2
dreadfully	Всього	100

Джерело: побудовано автором на основі прикладу Gries (2009).

Окрім цього, аналіз частотності може бути здійснено у динаміці протягом певного періоду часу для досліджуваного токєну або у вигляді графіку відношення частотностей набору токєнів, як представлено на рис. 2.2.

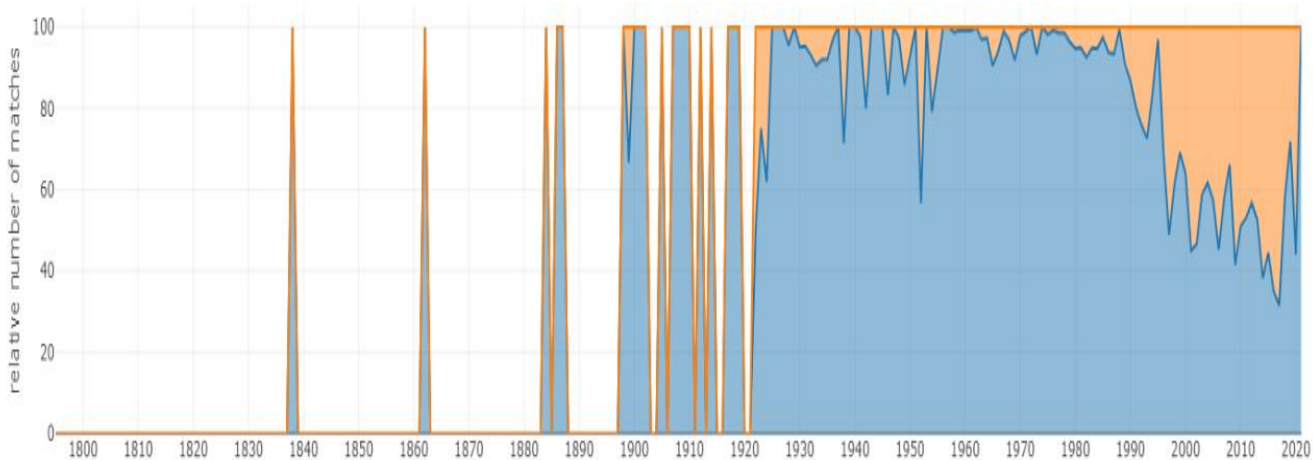


Рис. 2.2. Графік відношення кількості згадувань слів «слухавка» (помаранчевий колір на графіку) та «трубка» (синій колір на графіку) протягом історичного періоду на основі Генерального регіонально анотованого корпусу української мови.

Джерело: М. Шведова та ін. (2017).

Аналіз на основі таблиць частотності дозволяє визначити найбільш розповсюджені слова у заданому корпусі чи вибірці тексту, а також виокремити менш важливі токени. Розподіли токенів за частотністю є корисними при вивченні іноземних мов, у педагогічних цілях тощо. Аналіз на основі графіків частотності одного токена чи набору слів (N-грамів) є важливим інструментом перевірки лінгвістичних гіпотез, зокрема при дослідженні змін у вживанні запозичень чи різних форм певного слова протягом історичного періоду.

Показник *розсіювання/неоднорідності* токена дозволяє проаналізувати розподіл певного слова у лінгвістичному корпусі та ступінь однорідності такого розподілу, що може суттєво доповнити аналіз частотності. Ілюстрація розподілу певної лексичної одиниці в тексті наведено на рис. 2.3. За допомогою даного інструменту можна краще дослідити структуру тексту – як розподіл окремих фраз, тематичних полів, так і схему розвитку окремих ідей чи стилістичних елементів (Olejniczak, 2018).



Рис. 2.3. Ілюстративний приклад графіку розподілу лексичної одиниці у тексті.
Джерело: Olejniczak (2018).

Попри широку розповсюдженість аналіз частотності та дослідження розподілу слів у структурі тексту не є єдиними інструментами у корпусній лінгвістиці, адже вони не враховують асоціативну міру між словами і контекст, у якому відповідні токени можуть вживатися. Як зазначають І. Ділай та М. Ділай (2022), деякі слова можуть бути високочастотними лише у вузьких контекстах. Відповідно, виникає необхідність у таких методах аналізу, які б, окрім розрахунку частотності токенів, дозволили б враховувати асоціацію між ними та контекст спільного вживання.

Саме тому однією із центральних методологічних концепцій у корпусній лінгвістиці, окрім власне частотності, є аналіз *колокацій* на основі розрахунку асоціативної міри між словами, що разом зустрічаються у тексті.

Відповідно до Ш. Гріес (2009), колокації – це випадки спільного уживання слів, що називаються колокатами. Іншими словами, аналіз колокацій зосереджений на виявленні найчастіше уживаних слів-колокатів, які оточують цільове слово у близьких позиціях зліва та справа. Результати аналізу колокації зазвичай представляються у вигляді таблиць подібно до прикладу у табл. 2.3., де представлено колокації для слова «alphabetic». Побудувавши подібну таблицю для слова «alphabetical», можна проаналізувати граматичні і семантичні відмінності у їхньому вживанні.

Таблиця 2.3.

Ілюстративний приклад таблиці колокацій

1 слово зліва	Частота – 1 слово зліва	Слово	1 слово справа	Частота – 1 слова справа
of	8	alphabetic	literacy	7
the	6		writing	5
an	5		order	3
in	2		character	3
such as	2		and	2
our	2		system	2
when	2		characters	2
widespread	1		culture	2
systems	1		in	1
varying	1		.	1

Джерело: побудовано автором на основі прикладу Gries (2009).

Дослідження колокацій є важливим інструментом семантичного аналізу, у лексикографії, інших сферах наукових досліджень, при викладанні мов тощо. Окрім цього, у деяких ситуаціях може бути складно розрізнити близькі за значенням слова або визначити стилістичні, граматичні та інші особливості вживання слова в різних конструкціях, послуговуючись лише інтуїцією та лінгвістичним досвідом. Наприклад, носіям інших мов, які вивчають українську, непросто розрізнити відтінки слів, зокрема таких, як *швидкий, прудкий, жвавий, бурхливий, хуткий, невпинний, стрімкий*. Аналіз

колокаційної таблиці в таких випадках дозволить об'єктивно визначити для кожного слова, у яких фразах і ситуаціях воно найчастіше вживається.

Хоча аналіз колокацій є корисним інструментом, у деяких випадках він може бути субоптимальним, адже сфокусований передусім на спільному уживанні слів у сусідніх позиціях без урахування ширшого контексту. Ш. Гріес (2009) наводить приклад фрази «the **big** and mean guy», де найближчими колокатами слова «big» є *the* та *and*, які не надають достатньої кількості інформації щодо семантики слова «big».

Для вирішення цієї проблеми у корпусній лінгвістиці використовують більш комплексний метод, що називається «конкорданс». Він дозволяє дослідити більш широкий контекст, у якому згадується слово. Контекст цільового слова може бути визначений як відповідне речення, або як набір слів заданого обсягу, що оточує цільове слово з обох сторін (Gries, 2009). Результати застосування даного інструменту узагальнюються у табличному вигляді, як представлено в ілюстративному прикладі на рис. 2.4.

#			Data
1		Ukraina	atakuje cele w Rosji i na Krymie . 31/8/2023
2	Relacja w tvn24.pl Samoloty F-16 , które	Ukraina	ma otrzymać po wielu miesiącach lobbowania , po... 31/8/2023
3	W ubiegłym roku , w związku z agresją Rosji na	Ukrainę	i pojawieniem się ukraińskich dzieci i młodzieży w ... 7/8/2023
4	stwa , od 24 lutego 2022 r . , czyli od napaści Rosji na	Ukrainę	, do 31 sierpnia 2023 r . możliwe jest zatrudnienie ... 7/8/2023
5	Zobacz także : Atak Rosji na	Ukrainę	. 31/8/2023
6	ciu tysięcy części i muszą one zostać dostarczone na	Ukrainę	, by w razie konieczności wymiany jakiegoś eleme... 31/8/2023
7	generał James B . Hecker powiedział , że F-16 trafią na	Ukrainę	najwcześniej w przyszłym roku . 31/8/2023
8	z szczególnie w kontekście obecnej sytuacji wojennej na	Ukrainie	. 8/8/2023
9	Kwestie związane z trwającą na	Ukrainie	wojną , też , jak się uważa , będą jednym z temató... 16/8/2023
10	Waszyngton przy okazji konfliktu na	Ukrainie	testuje możliwości wojny z rywalem strategicznym ... 16/8/2023
11	Ale motywacja musi obejmować też techników i inżynierów , jest bowiem mało prawdopodobne , by amerykańscy technicy znaleźli się na	Ukrainie	– podkreśla CNN . 31/8/2023
12	Warunkiem jest co najmniej jeden uczeń z	Ukrainy	uczący się w tej szkole . 7/8/2023
13	Początku marca 2022 r . ustawy o pomocy obywatelom	Ukrainy	w związku z konfliktem zbrojnym na terytorium teg... 7/8/2023
14	uczyciela w oddziale przygotowawczym dla uczniów z	Ukrainy	nauczyciela pobierającego świadczenie kompensa... 7/8/2023
15	Zgodnie zaś z uchwaloną w grudniu 2022 r . nowelizacją ustawy o pomocy obywatelom	Ukrainy	, możliwe jest do 31 sierpnia 2023 r . zatrudnienie bez utraty świadczenia kompensacyjnego nauczyciela języka polskiego lub nauczyciela prowadzącego dodatkową naukę języka polskiego . 7/8/2023

Рис. 2.4. Ілюстративний приклад конкордансу.

*Джерело: побудовано на основі запиту автора (запит: Ukrain. *) за період серпня 2023 р. за допомогою комп'ютеризованого моніторингового корпусу польських новин на порталі monco.frazeo.pl (Pezik, 2020).*

Конкорданс дозволяє дослідити набагато ширший контекст виразу, ніж у випадку колокацій. Однак цей метод потребує і значно більших інвестицій часу та зусиль дослідника. Оскільки детальне ознайомлення з десятками, сотнями або мільйонами контекстів є технічно неможливим чи непрактичним, то існують спрощені методи застосування конкордансу: дослідник може сформувати випадкову вибірку достатнього обсягу, припускаючи, що вона відображає загальні тенденції усього набору даних; в межах альтернативного підходу дослідник може зосередитися на аналізі визначених патернів уживання цільового слова, звуживши у такий спосіб обсяг вибірки.

Одним із етапів корпусного аналізу є також дослідження *сентименту тексту* – напрям комп'ютерної лінгвістики, що активно розвивається протягом останніх десятиліть. Мета даного аналізу – визначити тональність тексту, що відображає відтінок відношення автора до об'єкту повідомлення (позитивне, негативне, нейтральне тощо).

Ця тема є особливо актуальною у сфері медіа-комунікації, що характеризується інтенсивною конкуренцією ЗМІ за увагу читачів, високим ступенем політичного й ідеологічного впливу, широким залученням емоційно-психологічних та маніпулятивних технік. Наприклад, за допомогою аналізу сентименту у мас-медіа і соціальних мережах існує можливість вимірювати соціальну напругу, політичні вподобання суспільства, аналізувати очікування домогосподарств щодо зміни економічної ситуації, перевіряти наукові гіпотези з боку лінгвістики, психології та інших гуманітарних і соціальних наук.

Важливе практичне значення аналіз сентименту має також у сфері онлайн-маркетингу. На основі дослідження відгуків і коментарів споживачів у соціальних мережах, на офіційних веб-сторінках, форумах тощо компанії можуть краще зрозуміти переваги й недоліки власного продукту або сервісу. Окрім цього, це дозволяє оперативно реагувати на негативні відгуки чи ворожі інформаційні кампанії у публічному просторі, аналізувати зміну сентименту споживачів залежно від рішень менеджменту щодо

запуску нових і модифікації чинних продуктів, соціальної корпоративної відповідальності компанії тощо.

Очевидно, що величезний масив текстових даних на порталах новин і соціальних мережах, які генеруються щосекунди по всьому світу, неможливо проаналізувати без допомоги автоматичних комп'ютеризованих методів. Як зазначає Н. Дарчук (2019), способи автоматичної оцінки тональності тексту можна умовно розділити на дві групи:

1. Методи, *засновані на правилах*, сформульованих дослідником. Суть методу полягає в розробці низки критеріїв щодо перевірки лексем, ланцюжків слів, словоформ, граматичних конструкцій із використанням спеціально укладених словників. На основі перевірки правил одиницям тексту, зокрема повідомленню, коментарю чи новині, присвоюють оцінку настрою за обраною шкалою (наприклад: дуже негативно, негативно, нейтрально, позитивно, дуже позитивно). При формулюванні правил враховують лексичні, граматичні, морфологічні та інші мовні засоби вираження емоційності й експресивності (у т.ч. вживання комбінацій слів у переносному значенні, використання афіксів з емотивним значенням, наявність оцінного компонента), позитивні чи негативні конотації слів на основі аналізу колокацій. Важливими факторами впливу на тональність повідомлення також є інтенсифікатори-прислівники (наприклад, *дуже, занадто, ледь, заледве, недостатньо*) та модифікатори оцінки (заперечна частка *не*).

2. Методи на основі *алгоритмів машинного навчання*. Спочатку дослідник створює розмітку для певної вибірки текстових даних (тренувальна вибірка даних) на основі обраної шкали оцінки настрою. За допомогою комп'ютерних алгоритмів модель оцінки настрою дослідним шляхом «навчається» на тренувальній вибірці даних, визначаючи оптимальні параметри. На основі оціненої моделі алгоритми машинного навчання здатні автоматично визначити настрій нових зразків тексту з досить високою точністю, що значно полегшує роботу дослідника. Відповідні моделі для настрою-аналізу тексту можуть бути загального спрямування (оцінені на основі тренувальної вибірки, що містила корпуси тексту з великої кількості різних джерел) та

спеціалізованими (оціненими з використанням корпусу тексту, створеного однією мовою або обмеженого лише певною галуззю знань). Хоча моделі загального спрямування є більш універсальними, спеціалізовані моделі можуть давати більш точні результати для вирішення вузьконаправлених завдань. Наразі існує чимало онлайн-ресурсів, що пропонують комерційні та дослідницькі інструменти, які здатні визначати сентимент тексту з використанням штучного інтелекту.

Важливим квантитативним методом є також *комп'ютеризований семантичний аналіз* текстів, який базується на основі математичного алгоритму прихованого розподілу Діріхле (LDA – Latent Dirichlet Allocation), що описаний і використовується зокрема у Н. Chen, X. Wang, S. Pan та F. Xiong (2019). Сукупність методів комп'ютеризованого семантичного аналізу сфокусовані на вирішенні багатьох практичних завдань, зокрема:

а) тематичного моделювання (від англ. *topic modelling*) – статистичного методу, що має на меті визначити найбільш імовірну тематику кожного тексту в корпусі, розглядаючи текст як набір тем, а тему як набір слів. Н. Chen, X. Wang, S. Pan та F. Xiong (2019) на основі корпусу наукових статей на тему великих даних (big data), що містить близько 5450 прикладів наукових досліджень, ідентифікували і категоризували найбільш поширені теми в цій галузі, як представлено на рис. 2.5.

TABLE I
TOPIC MODELING RESULT FOR BIG DATA RESEARCH

Topic Label	Topic Name	Topic Description
T01-Workfl	Scientific Workflow	scientific, workflow, interdisciplinary, computing, cyberinfrastructure
T02-Epidem	Epidemiology	population, observational, epidemiology, diabetes, prevalence
T03-Opti	Optimization	optimization, partitioning, NoSQL, scalability, genetic-algorithm
T04-Annota	Annotation	annotation, graphical, curation, repository, biological
T05-Genome	Genome Data	genome, DNA, NGS (next-generation sequencing), RNA, microarray
T06-Para	Parallelism	scalability, large-scale, parallelism, processors, multicore
T07-Govern	Data Protection	collection, protection, governance, transparency, reuse
T08-Archte	Architecture	architecture, deployment, enterprise, bandwidth, scalable
T09-LarSca	Large-Scale Data	dataset, large-scale, collections, ecological, biodiversity
T10-DecMak	Decision Making	analytics ,decision-making, GPU, competitive, opportunities
T11-Simula	Simulations	simulations, large-scale, observations, frequencies, calculations
T12-3D	3D Localization	3D, localization, interpolation, automata, instrumentation
T13-Relat	Relationship	relationship, correlation, subgraph, behavioral, differences
T14-BenchM	Benchmarking	historical, productivity, benchmarking, socio, disciplines
T15-GIS	Geographic IS	geographic, information-system, geospatial, popularity, behaviour
T16-Clust	Clustering	clustering, decomposition, dimension, unsupervised, pattern-recognition
T17-ScoNet	Social Network	social-network, opportunities, tweets, facebook, decision-making

Рис. 2.5. Ілюстративний приклад кластеризації корпусу статей на тему великих даних.

Джерело: уривок таблиці з дослідження Н. Chen, X. Wang, S. Pan та F. Xiong (2019).

б) визначення ступеня пов'язаності тем або імовірності співуживання тем, що визначається на основі тематичного моделювання.

У межах поглибленого аналізу Н. Chen, X. Wang, S. Pan та F. Xiong (2019) визначили силу зв'язків між виявленими категоріями тем, проілюструвавши їх за допомогою діаграми.

Як представлено на рис. 2.6., тема епідеміології (T02-Epidem), наприклад, найтісніше пов'язана з охороною здоров'я (T31-Health) та даними геному (To5-Genome), що є цілком очікуваним.

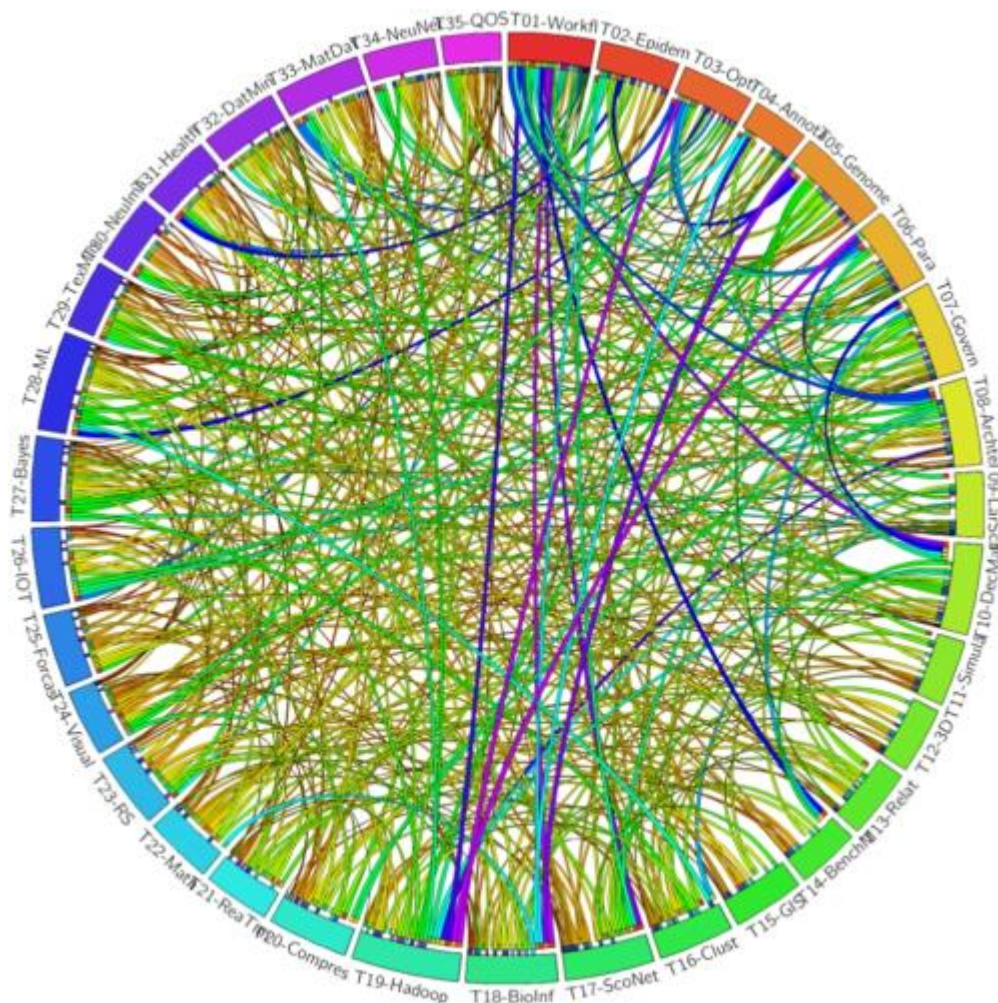


Рис. 2.6. Ілюстративний приклад візуалізації взаємозв'язків між категоріями тем на основі аналізу їх співживання.

Джерело: за результатами дослідження Н. Chen, X. Wang, S. Pan та F. Xiong (2019).

Примітка: товщина лінії характеризує щільність семантичного зв'язку між визначеними темами.

Хоча семантичний аналіз може проводитися й на основі знань і експертного судження дослідника, саме комп'ютерно-статистичні методи дозволяють з відносно нижчими витратами часу й ресурсів, а також на підґрунті об'єктивного підходу визначити приховані тематичні поля тексту та його взаємозв'язок з іншими текстами. Без застосування даних методів здійснення семантичного аналізу великих за обсягом текстів (що містять сотні тисяч або мільйони слів) було б практично недоцільним або неможливим.

Отже, використання квантитативних методів сприяє об'єктивності одержаних результатів, одержаних в межах емпіричного аналізу, а також розширює можливості дослідника щодо аналізу великого обсягу мовного матеріалу, зокрема із застосуванням великих лінгвістичних корпусів.

2.3. Методика та етапи корпусного дослідження українських запозичень польської мови

Як зазначено в меті роботи, наше дослідження зосереджене на аналізі українських запозичень у текстах польських засобів масової інформації з початку 2022 року.

Для досягнення визначеної мети дослідження необхідно, щоб обраний лінгвістичний корпус відповідав низці критеріїв:

- 1) Наявність у корпусі сегменту новин, що публікуються у польських мас-медіа.
- 2) Покриття тривалого історичного періоду.
- 3) Охоплення значної кількості медіа-ресурсів, зокрема найбільш популярних та авторитетних у Польщі.
- 4) Регулярне наповнення бази корпусу новими матеріалами.
- 5) Доступність на онлайн-платформі корпусу базового набору сучасних інструментів корпусної і квантитативної лінгвістики, зокрема інструментів для аналізу колокацій та конкордансу.
- 6) Відкритий онлайн-доступ до платформи корпусу та зручність у використанні.
- 7) Можливість завантаження результатів запитів.

Відповідно, було обрано *моніторинговий* корпус польських новин MoncoPL, адже він відповідає визначеним критеріям, як описано нижче.

Сервіс MoncoPL надає відкритий доступ до моніторингового корпусу польських новин за період з 2010 року, який щоденно актуалізується новими матеріалами (Peżik, 2010). У табл. 2.4. представлено 25 найбільших джерел новин, що входять до корпусу,

серед яких широко відомі польські ресурси GAZETA.PL, TVN24, Onet, Wirtualna Polska, Rzeczpospolita, wPolityce, Wprost, TVP Info, RadioMaryja та багато інших.

Таблиця 2.4.

Ключові джерела, на основі яких сформований корпус MoncoPL

Джерело	Кількість речень
gazeta.pl	32,070,212
tvn24.pl	18,767,638
interia.pl	16,674,710
onet.pl	15,610,960
gazetaprawna.pl	12,291,068
wp.pl	11,225,160
wpolityce.pl	9,722,532
rp.pl	9,044,507
wnp.pl	8,784,582
rmf24.pl	8,729,262
dziennikwschodni.pl	8,242,316
fakt.pl	6,493,759
wprost.pl	6,350,963
weszlo.com	6,176,676
naszemiasto.pl	5,654,825
news.google.com	5,396,675
money.pl	5,006,833
forsal.pl	4,649,259
tvp.info	4,249,032
bankier.pl	4,187,953
dziennik.pl	4,081,317
radiomaryja.pl	3,993,416
wykop.pl	3,538,553
pomorska.pl	3,413,279
trojmiasto.pl	3,335,650
Інші	227,239,763

Джерело: побудовано на основі даних на платформі MoncoPL (<http://monco.frazeo.pl/>).

Примітка: статистичні дані подано станом на 7 вересня 2023 р.

Станом на 8 вересня 2023 р. корпус MoncoPL містить близько 447 млн речень або майже 8,4 млрд слів та постійно доповнюється новими матеріалами. За даними порталу

MoncoPL протягом останніх десяти років щомісячний приріст складав в середньому близько 2,7 млн речень, як представлено на рис. 2.7.

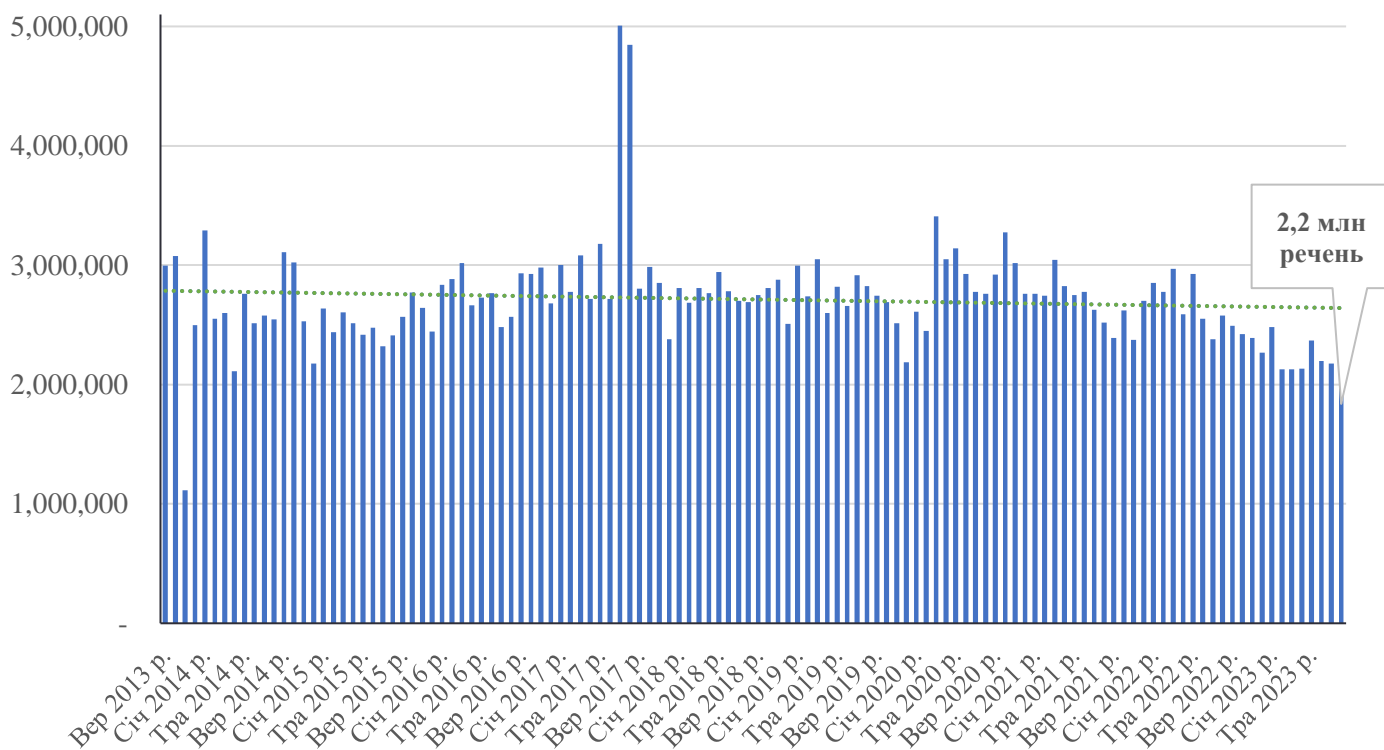


Рис. 2.7. Статистика щомісячного приросту даних корпусу MoncoPL протягом останніх 10 років.

Джерело: побудовано на основі даних на платформі MoncoPL (<http://monco.frazeo.pl/>).

Цей комп'ютеризований корпус містить основні інструменти, що необхідні для проведення корпусного аналізу текстів у засобах масової інформації, зокрема:

- 1) аналіз частотності словоформ;
- 2) аналіз конкордансу;
- 3) аналіз колокацій;
- 4) аналіз циклічності вживання слів;
- 5) розподіл слова за джерелами, де воно зустрічалося найчастіше;

б) можливість формулювання морфосинтаксичних виразів (наприклад, вираз «<tag=ADJ.*> zwycięstwo» дозволяє знайти прикметники, які вживалися у корпусі разом зі словом *перемога*) та визначення необхідних часових рамок з точністю до одного дня;

7) можливість вивантаження текстової інформації для більш детального аналізу.

Проведене емпіричне корпусне дослідження ґрунтується на аналізі інформації у відкритих веб-ресурсах, передусім із використанням інструментарію комп'ютеризованого корпусу MoncoPL, і складається з кількох взаємопов'язаних етапів, що представлені на рис. 2.8.:

1. Формування словника з українськими лексемами, що могли бути запозичені польськими медіа-ресурсами, зокрема слова-неологізми й вирази, що набули популярності в українському і світовому медіа-просторі з початку 2022 р. Для пошуку таких слів і виразів беруться до уваги тексти українських і міжнародних новин, промови українських і світових лідерів, що пов'язані з Україною, результати експрес-досліджень та думок експертів.

2. Розподіл за тематичними групами обраних лексем здійснюється залежно від тематики, лексичного значення та сфери їх вживання.

3. За допомогою інструментів, доступних на порталі MoncoPL, аналізується частотність вживання кожної з лексем, починаючи з січня 2021 р. по січень 2023 р., а також зміна частотності протягом узятого для аналізу періоду.

4. Якщо слово не вживалося у польських медіа до лютого 2022 р. або вживалося з незначною частотністю (нерегулярні, випадкові згадування), але почало активно згадуватися протягом 2022 р. та/або початку 2023 р., то воно включається до фінальної вибірки. Після відбору слів на основі аналізу частотності з вибірки було виключено ті лексеми, які не підпадають під визначення українських запозичень на основі аналітичних суджень.

5. Для кожного зі слів, що увійшли до фінальної вибірки, проводиться більш детальний аналіз колокацій і конкордансу.



Рис. 2.7. *Етапи корпусного дослідження українських запозичень у польських медіа. Джерело: побудовано автором.*

Висновки до розділу 2

У лінгвістиці під корпусом розуміють велике зібрання текстів певної тематики, що використовується з метою емпіричного дослідження мовних явищ, закономірностей, особливостей слововживання та перевірки наукових гіпотез. Як окремий мовознавчий напрям корпусна лінгвістика сформувалася та набула активного розвитку відносно

нещодавно, що пов'язано передусім з появою персональних комп'ютерів у другій половині ХХ ст.

Корпусна лінгвістика є *міждисциплінарною галуззю*, адже поєднує власне лінгвістичну теорію, когнітивістику, інформаційні та комп'ютерні технології, статистику і математику. Серед основних методів дослідження – кількісний та емпіричний аналіз. Відповідно, особливістю корпусної лінгвістики є об'єктивне вивчення мовного матеріалу у його природній формі з усією специфікою та недоліками на відміну від суб'єктивного підходу багатьох інших напрямів у лінгвістиці.

Водночас корпусній лінгвістиці властиві й певні обмеження. Наприклад, будь-який корпус, хоч і максимально наближається, проте не може відобразити реальність у повному обсязі, наприклад, усі випадки вживання певного виразу чи граматичної конструкції. Корпус підтверджує факт існування певного мовного елементу чи явища, але не обґрунтовує його лінгвістичну коректність або причину функціонування.

Серед основних етапів і методів сучасного корпусного дослідження можна виділити такі: аналіз частотності; порівняння динаміки частотностей для набору виразів протягом певного історичного періоду; аналіз колокацій та конкордансу; дослідження варіативності слововживання та розподілу виразу у певному тексті; аналіз сентименту; семантичний аналіз.

У рамках нашого дослідження використовується інструментарій комп'ютеризованого корпусу MoncoPL, що станом на вересень 2023 р. містить близько 8,4 млрд слів із найбільш популярних медіа-ресурсів у Польщі та щоденно доповнюється новими матеріалами.

РОЗДІЛ 3

НОВІТНІ УКРАЇНІЗМИ В ПОЛЬСЬКІЙ МОВІ ЯК МОВНЕ ВІДДЗЕРКАЛЕННЯ РОСІЙСЬКОЇ ВОЄННОЇ АГРЕСІЇ

3.1. Українські запозичення в польській мові: соціально-історичні та культурні передумови проникнення

Процес імпортування елементів однієї мовної системи до іншої, а також їх адаптація у мові-реципієнті залежать від низки чинників, серед яких:

- 1) тривалі мовні контакти між двома народами;
- 2) активність і мовленнєва діяльність білінгвів-носіїв обох мов;
- 3) тісні культурні й політичні зв'язки, наприклад, внаслідок співіснування двох народів у межах однієї держави протягом певного історичного періоду;
- 4) ступінь природної схильності мови до запозичення іншомовного лексичного матеріалу;
- 5) інші екстралінгвістичні чинники, зокрема інтеграційні процеси у сфері економіки, демократизація інформаційно-культурного простору, технологічний розвиток, глобалізаційні процеси, розвиток транспортної інфраструктури.

Іншим важливим аспектом є політичний, соціальний та/або культурний статус обох мов, що визначає напрям запозичень. Наприклад, в історичних реаліях мови країн-завойовників або культурних лідерів були схильними домінувати в культурному та суспільно-політичному просторі, пригнічуючи певним чином мови захоплених народів. Відповідно, більшість запозичень імпортувалися від мови умовно домінантного народу (мова-донор) до мови тих народів, що знаходилися під впливом іншомовної культури (мови-реципієнти).

Саме тому для дослідження механізму й результату запозичень між українською та польською мовами необхідно враховувати культурно-історичні аспекти їх спільного розвитку й функціювання.

Близькість польської та української можна пояснити як їхнім спільним походженням (обидві мови належать до сім'ї індоєвропейських мов та слов'янської мовної групи), так і тривалими культурними контактами, що активно відбувалися протягом століть і тривають донині (Кононенко, 2012).

Унаслідок численних та тісних контактів протягом багатьох століть відбувалися запозичення як з польської до української, так і з української до польської мови.

Враховуючи, що тривалий час значна частина українських земель перебувала у складі Королівства Польського, а пізніше Речі Посполитої, де основною державною мовою була польська, то значна частина слів імпортувалися саме з польської до української мови. М. Філон та О. Хомік (2010) виділили декілька тематичних груп запозичень з польської мови, що належать передусім до суспільно-економічної, побутової та адміністративної лексики:

- 1) військова справа: *полковник, рота*;
- 2) суспільно-політичне життя: *герб, урядник*;
- 3) виробництво: *маляр, тесля, ковадло*;
- 4) культура: *мазурка, краков'як*;
- 5) їжа та побут: *повидло, шинкувати*.

Серед основних ознак полонізмів М. Філон та О. Хомік (2010) визначили такі: звукосполучення [дл] і [тл] (*тлумачити, повидло, підлога*), суфікс *-изн* (*білизна, вітчизна*).

З іншого боку, ідентифікація та вивчення запозичених *українізмів* у польській мові є не менш важливим та цікавим напрямом з високим дослідницьким потенціалом, який досі залишається вкрай актуальним. Вивченню цього питання присвячено праці І. Кононенко (2012; 2015-2016), Н. Совтис (2005), А. Фаловські (2023) та інших українських і польських вчених. На основі аналізу п'яти словників та інших джерел,

А. Фаловські (2023) виявив близько 366 українізмів у польській мові, зазначивши, що цей перелік не є вичерпним. Н. Совтис (2005) наводить подібні дані, зазначивши, що в літературі зафіксовано близько 300 українізмів. Варто зазначити, що серед ідентифікованих слів частина відображає лексику зі східнослов'янських, східних та інших мов, що була активізована в польській мові за посередництвом української.

Отже, можна виділити декілька основних чинників, що сприяли поширенню українізмів у польській мові, зокрема:

1. Внаслідок творчої діяльності деяких польських письменників, які походили з т. зв. Східних Кресів та були носіями декількох мов (назва земель, частина яких є територією сучасної України). Яскравим прикладом є творчість Ю. Словацького, що наводиться у Кононенко (2015-2016): ... *Płonie chaty niska strzecha... Ledwo błysnął świt na niebie, Już po stepach, już po błoniu Leci Ruńko na swym koniu*. Особливо активно українізми проникали в польську мову в епоху романтизму, коли письменники відображали у своїй творчості фольклорні, історичні, природні та інші елементи української мови із емоційним забарвленням.

2. Тривалі суспільно-політичні, культурні та економічні контакти між представниками обох народів. Згідно з дослідженням Н. Совтис (2005), більшість українізмів у польській мові прямо чи опосередковано пов'язані з людиною та її діяльністю (зокрема назви осіб та груп людей, флористична, анімалістична, побутова й військова лексика).

Отже, за інформацією, наведеною І. Кононенко (2015-2016), перші запозичення з української мови потрапили до польської у XIV ст. Водночас найбільше слів було імпортовано протягом XVI – XIX ст. Н. Совтис (2005) наводить детальний розподіл кількості українізмів за періодом запозичення на основі аналізованої вибірки: 54% лексичних елементів у сучасній польській мові було запозичено протягом XVI–XVII ст., а 46% – XVIII–XIX ст., як представлено у табл. 3.1.

Таблиця 3.1.

Розподіл українізмів за періодом походження протягом XV-XIX ст.

XV ст.	XVI ст.	XVII ст.	XVIII ст.	XIX ст.
Ataman	Baskak	Bajdak	Choodziec	Bajdykowa
Derka	Bohater	Bandura	Czupurny	Bakun
Dubas	Braha	Bobak	Derkacz	Baabajka
Huba	Burzan	Buhaj	Duha	Baraban
Knia	Chata	Czereda	Hajdamaka	Basztan
Mikina	Czajka	Kompaczyk	Korowaj	Bodiak
Proskura	Duha	Komysz	Pohaniec	Chandra
Rohatyna	Horodyszcze	Mierzyn	Sadyba	Chmyz
...	Hospodyn	Opryszek	Sicz	Czerep
Інші	Hramota	Pohany	Step	Czumak
	Hreczka	Poroh	Wesela	Czwani si
	Hultaj	Semen	...	Halawa
	Kobza	Soboda	Інші	Hopak
	Kurchan	Wiedma		Kaczan
	Oczeret	...		Kutia/kucja
	Szarawary	Інші		Oczajdusza
	Wataga			Orowa

	Інші			Інші

Джерело: побудовано на основі інформації, поданої у Н. Советис (2005).

I. Кононенко зазначає, що в процесі адаптації до системи польської мови українізми часто зазнавали значної семантичної і стилістичної адаптації, виділивши такі групи:

1. Українізми зі збереженим значенням та нейтральним забарвленням (переважно небагатозначні слова на позначення об'єктів природи), наприклад, *czereпча, step, kukurydza*.

2. Запозичення з майже еквівалентним до оригіналу значенням, але з модифікованим конотативним статусом, наприклад, *blahu, kajdany, horowód*.

3. Українізми, що увійшли до функційно обмеженої лексики, зокрема історизми (*hajdamak, oseledec, sahajdaczny*), застарілі слова (*halaburda, kalyta*),

етнографізми (*bandura, hutor, hopak*), діалектизми (*bałakać, hreczka, kutia*), терміни (*oczeret, buhaj, derkacz*), жаргонізми (*lulka, brechać*).

4. Слова зі зміненим чи розширеним значенням. Одним з прикладів може слугувати слово *bohater* (у значенні героя, що відомий завдяки своїм незвичайним особистісним якостями чи видатним діям; герой твору або фільму) ← *богатыр*: оспівана у народній творчості фізично сильна й відважна людина-герой (Кононенко, 2015-2016).

Отже, внаслідок тісних мовних і культурних контактів польського й українського народів, обидві мови імпортували й запозичували елементи своїх систем протягом століть. Більшість українізмів були адаптовані польською мовою у період з XVI ст. до XIX ст. завдяки тривалим суспільно-політичним, культурним та економічним контактам, за посередництвом літературних творів письменників польсько-українського походження тощо. Українізми в польській мові еволюціонували, що нерідко призводило до суттєвих семантичних і конотативних змін. Водночас у фаховій літературі немає достатньо згадувань про запозичення українізмів польською мовою протягом XX ст. - початку XXI ст.

З початком російського повномасштабного вторгнення в Україну у лютому 2022 року в українській мові відбувається масштабний процес появи неологізмів, пов'язаних із воєнними реаліями, водночас суттєво активізуються мовні контакти між українцями і поляками, польські ЗМІ кожного дня відбивають події в Україні. Відповідно, можна припустити, що під впливом зазначених чинників частина неологізмів українського походження буде запозичена польською мовою.

3.2. Тематичні групи українізмів воєнного часу в текстах польських новин

Важливою складовою даного дослідження є емпіричний аналіз запозичень, що нещодавно увійшли до дискурсу польських засобів масової інформації за посередництвом української мови.

Корпусне дослідження охоплює період з січня 2021 р. по січень 2023 р. з метою виявлення груп слів і виразів, які почали активно використовуватися у текстах польських ЗМІ з початком повномасштабних воєнних дій (лютий 2022 р. – січень 2023 р.) та порівняно рідко вживалися або не вживалися взагалі за один рік до цього (січень 2021 р. – січень 2022 р.). Обрання відповідного періоду аналізу було зумовлено тим, що саме у період воєнного часу, під впливом політичних, національних, психологічних, економічних та інших суспільних чинників, українська мова стрімко розвивається, поповнюючи свій лексичний склад або змінюючи конотаційне забарвлення наявних слів.

Як описано у попередньому підрозділі, інші мови, зокрема й польська мова, запозичують нові лексеми або вирази, створені або переосмислені носіями української мови, внаслідок тісних контактів з українськими мігрантами й партнерами, а також в рамках глобального медіа-дискурсу. Стрімкому поширенню запозичень сприяє також і розвиток сучасних комунікаційних технологій, за допомогою яких новини з новітньою лексикою активно розповсюджуються і проникають до медіа-простору більшості країн світу.

Основним завданням цього дослідження було ідентифікувати лексеми, що були запозичені або змінили свою конотацію під впливом української мови протягом періоду воєнного часу.

На підготовчому етапі аналізу було сформовано первинну вибірку слів, які з'явилися в українському медіа-просторі з 2022 р. до початку 2023 р. (а також загально уживаних слів, що почали використовуватися в іншому контексті чи змінили конотативне забарвлення). Для пошуку слів використовувалися тексти українських новин, промови політиків, інтерв'ю з мовознавцями тощо. До первинної вибірки було включено близько 300 слів, що згадувалися у текстах українських і польських медіа.

Наступний етап дослідження полягав у перевірці кожного з попередньо відібраних слів з використанням інструментарію комп'ютеризованого корпусу MoncoPL у такий спосіб:

- лексичні одиниці, що не вживалися у текстах польських новин протягом січня 2021 р. – січня 2023 р., *не були обрані* до фінальної вибірки;
- лексичні одиниці, що активно вживалися без суттєвої зміни семантики або конотації як протягом січня 2021 р. – січня 2022 р., так і протягом лютого 2022 р. – січня 2023 р., також *не були обрані* до фінальної вибірки, адже вони вже широко використовувалися у польській мові протягом певного часу, тому не можуть бути визначені як новітні запозичення з української мови;
- лексичні одиниці, що почали активно вживатися у польських ЗМІ переважно у період з лютого 2022 р. до січня 2023 р., але не використовувалися протягом січня 2021 р. – січня 2022 р. (або згадувалися з набагато меншою частотою у цей період, порівняно з 2022 р., та/або в інших контекстах), *були обрані* до фінальної вибірки. Наприклад, такі вирази як *HIMARS*, *wojna na pełną skalę*, *Lend-Lease* містилися у корпусі за період 2021 р., проте на той момент вживалися у досить вузьких контекстах (професійні, галузеві, історичні, регіональні тощо аспекти), тому й не можуть вважатися як такі, що були широко розповсюдженими у польському медіа-дискурсі серед більшості соціальних груп: *Polska zdecydowała się na zakup w Stanach Zjednoczonych myśliwców F-35, zestawów rakietowych Patriot i modułów ogniowych HIMARS; Z powodu Nord Stream 2 może wybuchnąć wojna na pełną skalę; bez amerykańskiej pomocy w ramach programu Lend-Lease ZSRS przegrałby wojnę z Niemcami.*
- окрему категорію становили українські культурери, які були запозичені польською мовою декілька століть тому. До аналізу культуререм застосовувався окремий підхід. Основною метою було визначити, чи вживалися відповідні слова у новітньому дискурсі польських новин (з лютого 2022 р. до січня 2023 р.).

За результатами корпусного дослідження частотності й історичної динаміки вживання відповідних лексичних одиниць у польських ЗМІ було сформовано фінальну

вибірку, що містить понад 100 слів. Усі слова було класифіковано за однією з таких категорій, як представлено у табл. 3.1.:

- символи героїзму та національної ідентичності;
- назви й види зброї;
- власні назви, породжені війною;
- іронічне зображення риторики пропаганди;
- культурами;
- топоніми;
- інші потенційні запозичення.

За результатами корпусного дослідження частотності й історичної динаміки вживання відповідних виразів у польських ЗМІ було сформовано фінальну вибірку, що містить понад 100 слів. Усі слова було класифіковано за однією з таких категорій, як представлено у табл. 3.2.:

Таблиця 3.2.

Тематичні групи виразів у польській мові, які були запозичені за посередництвом української мови або набули нового конотативного забарвлення

Тематична група	Приклади виразів в базовій формі	Усього слів у групі
Символи героїзму та національної ідентичності	<i>czerwona kalina; Dobry wieczór, jesteśmy z Ukrainy; pies Patron; Siły Zbrojne Ukrainy (ZSU); palanycsa; oborona terytorialna</i>	11
Назви і види зброї	<i>Bajraktar, Kindzał, HIMARS, Stuhna, NASAMS, Patriot, szahed, dron kamikadze, Neptun</i>	32
Новітні реалії, породжені війною	<i>brudna bomba, filtracja, denazyfikacja, Mordor, orkowie, zombirowani</i>	11
Іронічне зображення риторики російської іпропаганди	<i>biolaboratoria, bojowe komary, gest dobrej woli, nie mający analogów, czmobiiki, druga armia świata</i>	10

Продовження табл. 3.2.

Культуреми	<i>mereżka, chmara, step, hopak, wataha, koromysło, bandura, kureń</i>	11
Топоніми	<i>Awdjiwka, Bachmut, Bałaklija, Borodzianka, Bucza, Hostomel, Kachowka, Melitopol, Swatowe</i>	23
Інше	<i>alarm powietrzny, umowa zbożowa, wojna na pełną skalę, przesiedleńcy/ uchodźcy z Ukrainy, wojsko informacyjne</i>	6
Всього		104

Джерело: побудовано на основі даних на платформі MoncoPL (<http://monco.frazeo.pl/>).

При проведенні корпусного дослідження важливо врахувати й деякі практичні обмеження:

1. Точність емпіричних результатів передусім залежить від якості пошукового запиту та особливостей вживання слова, що аналізується, у корпусі.
2. Навіть дуже комплексні запити, що враховують велику кількість випадків та обмежень, не можуть досконало виявити всі випадки слововживання з максимальним рівнем точності. Типовим є випадок, коли слово, що вказано у запиті, згадується не лише у властивих йому контекстах, а й у складі власних назв (прізвища, назви населених пунктів й вулиць, назви організацій). У такому випадку результати аналізу частотності можуть бути дещо зміщеними (або містити зайві згадування, або, у разі застосування набору обмежень до пошукового запиту, не містити деякі випадки слововживання).
3. Відповідно, показники частотності слововживання доцільно сприймати та аналізувати як *індикативні* величини *відносної* частоти згадувань, а не як *абсолютні* величини.
4. Власне частотність уживання слова у корпусі новин вказує на його потенціал до адаптації польською мовою. Зазвичай слова з більшою частотою й тривалістю вживання запозичуються й адаптуються до системи мови-реципієнта з вищою

імовірністю. Водночас варто враховувати, що навіть дуже популярні слова можуть зникнути з медійного дискурсу у середньостроковій перспективі.

Проаналізуємо детальніше кожен тематичну групу українізмів воєнного часу, вживання яких активізувалося у текстах польських новин з лютого 2022 р. по січень 2023 р., тобто у перший рік після початку повномасштабних військових дій.

3.2.1. Символи героїзму та національної ідентичності

Після початку повномасштабних військових дій у лютому 2022 р. в українській мові з'явилися нові вирази-символи, що демонструють героїзм, жертвність та стійкість збройних сил та громадянського суспільства. На основі проведеного аналізу в корпусі польських новин було виявлено понад 10 таких виразів українського походження, як представлено у табл. 3.3. Хоча окремі елементи цих виразів існували ще до початку 2022 року, їхня сукупність набула нового значення та відповідні конотації в умовах воєнного часу. Завдяки засобам масової інформації більшість таких фраз набули стрімкого поширення у світовій пресі в контексті висвітлення боротьби українського народу за незалежність.

Таблиця 3.3.

Символи героїзму та національної ідентичності, що були запозичені з українського суспільного дискурсу до текстів польських новин в умовах воєнного часу

Вираз	Частотність вживання			Частка згадувань, %
	січень 2021 р. – січень 2022 р.	лютий 2022 р. – січень 2023 р.	Всього	лютий 2022 р. – січень 2023 р.
<i>Azowstal</i>	-	13 902	13 902	100
<i>Sily Zbrojne Ukrainy (ZSU)</i>	111	6 300	6 411	98
<i>Chwała Ukrainie!</i>	43	875	918	95
<i>Czerwona kalina</i>	1	167	168	99
<i>Pies Patron</i>	-	105	105	100

Продовження табл. 3.3.

<i>Azowcy</i>	-	143	143	100
<i>Palanysa</i>	-	28	28	100
<i>Dobry wieczór, jesteśmy z Ukrainy! / Dobry wieczór , my z Ukrainy!</i>	-	23	23	100
<i>Wierzymy w ZSU</i>	-	4	4	100
<i>Oborona terytorialna</i>	-	4	4	100
<i>Traktorowe wojska</i>	-	3	3	100

Джерело: побудовано на основі даних на платформі MoncoPL (<http://monco.frazeo.pl/>).

Одним із найбільш наочних прикладів набуття нового конотаційного забарвлення та статусу національного символу в умовах воєнного часу є словосполучення *червона калина* з пісні «Ой, у лузі червона калина... », що була написана ще на початку 1914 р. за мотивами козацьких пісень. Хоча образ калини вже протягом століть у народній творчості символізує Україну, а також мужність і жертвність її народу, він набув особливого символічного значення та широкої популярності в Україні та за її межами лише після виконання цієї пісні Андрієм Хливнюком (лідер гурту Boombox), а також переробленої версії пісні під назвою «Hey Hey Rise Up» гуртом Pink Floyd. Завдяки популярності пісні образ калини набув нового конотаційного забарвлення, символізуючи волю до перемоги, непохитність українського народу, який тримається і виборює свободу, незважаючи на усі складнощі війни, а також його віру у світле майбутнє й відродження.

Хоча іноді ця пісня згадувалася у польських новинах у контексті історичних подій першої половини ХХ ст., останнім часом словосполучення *czzerwona kalina* вживалося в польських ЗМІ в контексті висвітлення боротьби українського народу з російською агресією в умовах сучасної війни. Наприклад:

Jednym z symboli obecnej wojny jest piosenka "Czerwona kalina" (dorzeczy.pl); Uczestnicy marszu, idąc ulicami Warszawy, wznosili hasła wyrażające solidarność z Ukrainą oraz podkreślające sprzeciw wobec działań wojennych ... Protestujący odśpiewali hymn Ukrainy, a później popularną pieśń "Czerwona Kalina" (tvn24.pl); Najmłodszy ... przywitały odegranym na skrzypcach hymnem Ukrainy oraz zaśpiewały pieśń "Czerwona kalina" (wnp.pl); ... pisaliśmy o wspólnym wykonaniu piosenki "Czerwona kalina" przez młodzieżową grupę teatralną Zamojskiego Domu Kultury oraz młodzież ukraińską, która z powodu wojny uczy się w zamojskich szkołach (wiadomosci.onet.pl).

Показовим прикладом часткової зміни значення та контексту вживання слова є назва одного з найбільших металургійних комбінатів в Україні «Азовсталь» (до того, як завод зазнав руйнування у 2022 р.). Якщо до повномасштабних військових дій ця назва згадувалася передусім у контексті металургійного виробництва, то, внаслідок героїчної та самовідданої оборони як цього об'єкта, так і загалом Маріуполя, вироблений на комбінаті метал став символом незламності українських військових. Новий контекст уживання слова *Azowstal* можна спостерігати і в текстах польських новин, які запозичили це слово-символ із українського медіа-простору вже переважно з новим значенням. На рис. 3.1. продемонстровано топ-10 найбільш пов'язаних колокатів слова *Azowstal* (за індексом пов'язаності Соренсена). За результатами аналізу 10 найбільш пов'язаних колокатів можна дійти до висновку, що лише два слова з десяти стосуються безпосередньо діяльності підприємства (*metalurgiczny* та *kombinat* - із загальною часткою близько 21,5%). Водночас інші колокати пов'язані з обороною заводу та, відповідно, містом Маріуполь, де він розташований.

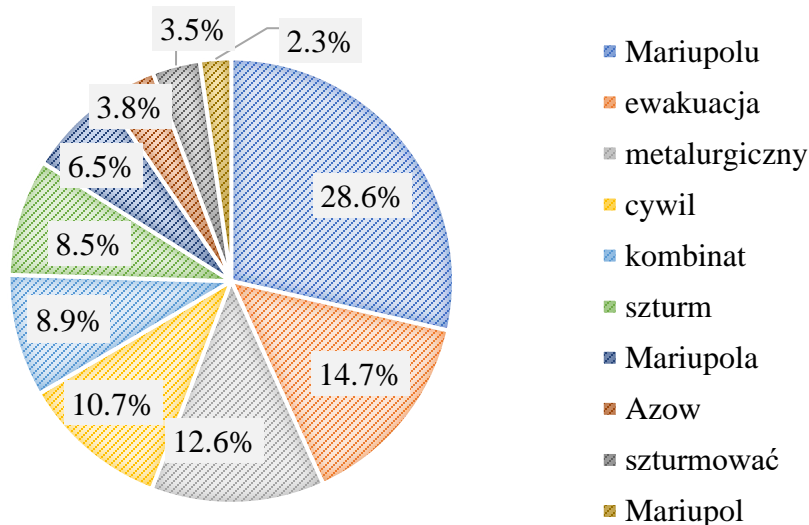


Рис. 3.1. Топ-10 найбільш пов'язаних колокатів слова «Azowstal» у корпусі польських новин за період з лютого 2022 р. по січень 2023 р.

Джерело: побудовано на основі даних на платформі MoncoPL (<http://monco.frazeo.pl/>).

Ілюстрацією українського запозичення до польського медіа-дискурсу, яке досі не згадувалося, оскільки було створено під час війни, є образ собаки-сапера, якого звать *Пес Патрон* (пол. *pies Patron*). На сьогодні він є одним із символів України, де навіть тварини відчайдушно рятують свою країну, наближаючи перемогу. Про це може свідчити характеристика собаки в текстах польських ЗМІ, зокрема: «*Pies-saper Patron został na festiwalu filmowym w Cannes laureatem nagrody dla psów Palm Dog-manitarian. Nagroda ta ... jest "najbardziej prestiżową nagrodą filmową na świecie, przyznawaną czworonogom" Ponieważ Patron rzeczywiście stał się w ciągu ostatnich paru miesięcy gwiazdą światowych kronik filmowych*» (tvn24.pl); «*Pies Patron, który służy w armii ukraińskiej, został odznaczony przez prezydenta Wołodymyra Zełenskiego. ... Ten pies rasy Jack Russell Terrier stał się czymś w rodzaju internetowego hitu*» (i.pl); «*Pies Patron jest bohaterem narodowym Ukrainy. ... Piesek Patron właśnie po to został twarzą ... kampanii rozminowywania, by pokazać światu wagę problemu*» (oko.press).

Історична динаміка найбільш використовуваних українських виразів у текстах польських новин, що символізують героїзм та національну ідентичність України (з

частотністю понад 100 згадувань протягом лютого 2022 р. – січня 2023 р.), представлена на рис. 3.2.

Інтенсивність вживання слів протягом періоду аналізу не була однаковою кожного місяця. Тренди згадувань запозичених українських виразів у публікаціях польських засобів масової інформації прямо корелюють з відповідними подіями, що є цілком природнім і очікуваним.

На рис. 3.2. також можна помітити, що більшість із досліджуваних запозичень не зникали з медіа-простору після первинного згадування, а періодично поверталися до медіа-дискурсу, що свідчить про їхню поступову адаптацію до мови-реципієнта.

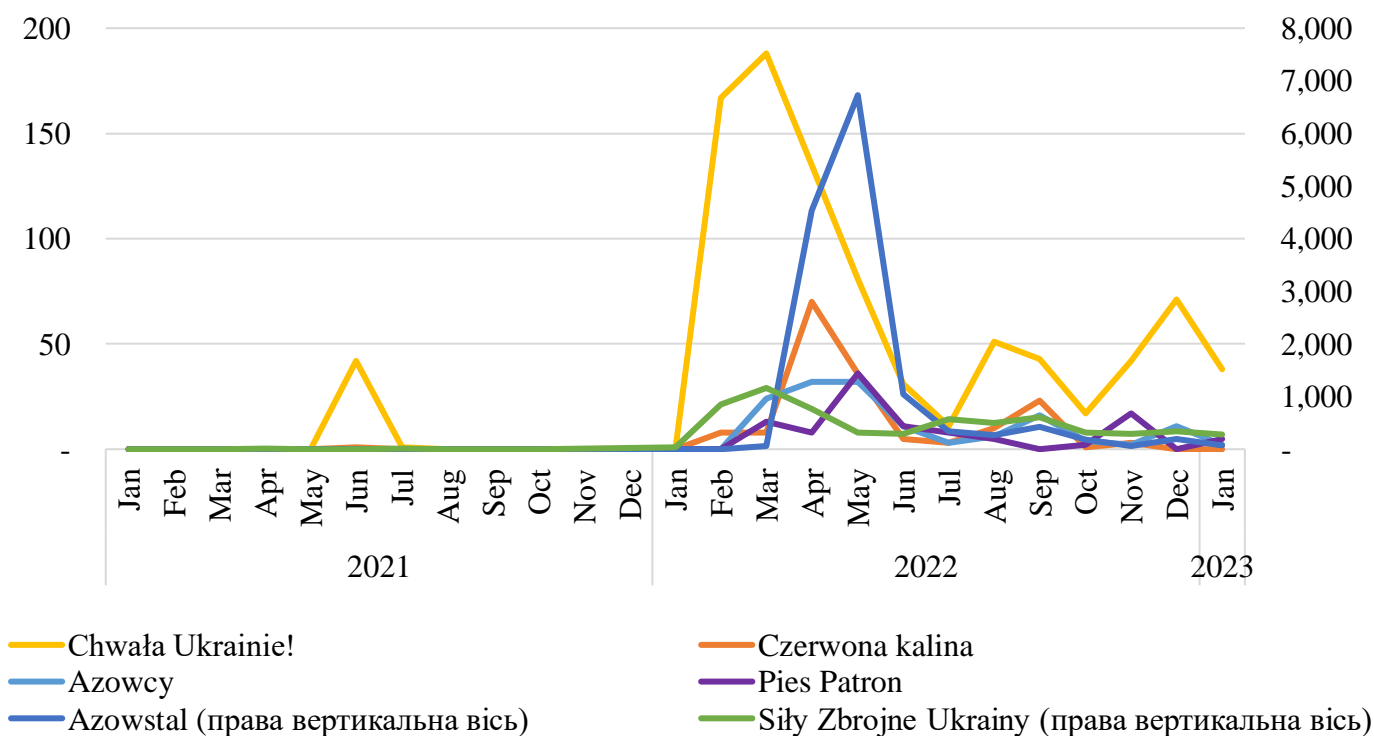


Рис. 3.2. Частотність вживання українських виразів з групи «Символи героїзму та національної ідентичності», що найбільше згадувалися у польських новинах протягом лютого 2022 р. – січня 2023 р.

Джерело: побудовано на основі даних на платформі MoncoPL (<http://monco.frazeo.pl/>).

3.2.2. Назви і види зброї

Загалом очікуваним є спостереження, що серед потенційних запозичень найчисленішою групою (32 позиції) є лексичні одиниці, безпосередньо пов'язані з видами й назвами зброї, як представлено у табл. 3.4. Слова саме з цієї групи найчастіше та найрегулярніше з'являлися у текстах польських новин: з-посеред 21 лексичної одиниці з фінальної вибірки, що згадувалися понад 1 000 разів протягом лютого 2022 р. – січня 2023 р., 11 відносяться до групи «Назви й види зброї».

Таблиця 3.4.

Назви і види зброї, що були запозичені до суспільного польського медіа-дискурсу за посередництвом української мови в умовах воєнного часу

Вираз	Частотність вживання			Частка згадувань, %
	січень 2021 р. – січень 2022 р.	лютий 2022 р. – січень 2023 р.	Всього	лютий 2022 р. – січень 2023 р.
<i>Obrona przeciwlotnicza, przeciwrakietowa</i>	1242	12477	13719	91
<i>HIMARS</i>	128	7648	7776	98
<i>Leopard</i>	174	6975	7149	98
<i>Patriot</i>	240	5048	5288	95
<i>Bajraktar</i>	288	3596	3884	93
<i>S-300, S-200, S-400</i>	473	3019	3492	86
<i>Iskandery</i>	419	2710	3129	87
<i>Krażownik Moskwa</i>	8	2788	2796	100
<i>Dron kamikadze</i>	12	2216	2228	99
<i>Javeliny</i>	345	1172	1517	77
<i>NASAMS</i>	14	1246	1260	99
<i>Kalibry</i>	18	703	721	98
<i>Gepard</i>	1	684	685	100
<i>ATACMS</i>	9	621	630	99

Продовження табл. 3.4.

<i>Precyzyjna broń</i>	16	609	625	97
<i>Neptun</i>	3	566	569	99
<i>NLAW</i>	54	502	556	90
<i>Lend-Lease</i>	32	518	550	94
<i>Bradley</i>	9	525	534	98
<i>Grad</i>	5	353	358	99
<i>Bomby fosforowe</i>	9	263	272	97
<i>Stinger</i>	27	210	237	89
<i>Stuhna</i>	3	166	169	98
<i>Iris-T</i>	0	166	166	100
<i>Solncepioki</i>	2	132	134	99
<i>Panzerfuast</i>	18	87	105	83
<i>Szahed</i>	0	72	72	100
<i>Mohadzer</i>	0	63	63	100
<i>SAMP/T</i>	0	54	54	100
<i>GIS Arta</i>	0	48	48	100
<i>Armata</i>	5	30	35	86
<i>Kindżal</i>	0	6	6	100

Джерело: побудовано на основі даних на платформі MoncoPL (<http://monco.frazeo.pl/>).

Найбільш уживаним у цій тематичній групі був вираз *Obrona przeciwlotnicza/przeciwrakietowa*, що спричинено активною роботою засобів протиповітряної та протиракетної оборони України через регулярні атаки російської армії протягом періоду аналізу, які широко висвітлювалися у засобах масової інформації.

Серед інших часто згадуваних слів можна виділити окрему підкатегорію – назви зброї країн-членів НАТО, що надається Україні союзниками: *HIMARS, Leopard, Patriot, Bajraktar, Javeliny, NASAMS, Gepard, ATACMS* тощо. Хоча більшість з цих назв існувала й до повномасштабного вторгнення, вони згадувалися значно рідше та/або у інших

контекстах у польському медіа-просторі до лютого 2022 р. Частотному згадуванню цих слів серед широких верств населення сприяли передусім такі два чинники:

1. Досить тривалий період погодження та політичних перемовин щодо постачання зброї, опис її основних характеристик та напрямів застосування.
2. Згадування зброї при висвітленні новин, пов'язаних з бойовими діями.

Іншою підкатегорією цієї тематичної групи є назви радянської і російської зброї, що застосовується проти України, зокрема: *S-300, Iskandery, Kraqzownnik Moskwa, dron kamikadze, Kalibry* тощо.

Окрім цього, досить часто згадувалися військові технології та зброя українського виробництва внаслідок її успішного застосування у відомих військових операціях: *Neptun, Stuhna, GIS Arta*.

3.2.3. Новітні реалії, породжені війною

Протягом 2022 р. в українському та світовому суспільному дискурсі активно формувалися нові слова та поняття під впливом трагічних подій війни, які представлені у табл. 3.5. Деякі слова, відомі ще з часів Другої світової війни або попередніх військових конфліктів, увійшли до мови польських мас-медів із новим значенням. До таких запозичень можна віднести, наприклад, *prywatne firme wojskowe, filtracje, brudna bomba, gauleiter*.

Таблиця 3.5.

Новітні реалії, породжені війною, які були запозичені з українського суспільного дискурсу до текстів польських новин в умовах воєнного часу

Вираз	Частотність вживання			Частка згадувань, % лютий 2022 р. – січень 2023 р.
	січень 2021 р. – січень 2022 р.	лютий 2022 р. – січень 2023 р.	Всього	
<i>Prywatne firme wojskowe</i>	96	1200	1296	93
<i>Filtracja</i>	5	1286	1291	100

Продовження табл. 3.5.

<i>Brudna bomba</i>	7	1134	1141	99
<i>Ruski trol</i>	149	788	937	84
<i>Denazyfikacja</i>	0	814	814	100
<i>Raszystowski, raszyzm, raszysty</i>	0	633	633	100
<i>Orkowie</i>	0	215	215	100
<i>Russkij wojennyj korabl</i>	0	77	77	100
<i>Zombirowani</i>	0	12	12	100
<i>Mordor</i>	1	10	11	91
<i>Gauleiter</i>	0	6	6	100

Джерело: побудовано на основі даних на платформі MoncoPL (<http://monco.frazeo.pl/>).

Цікавим з боку лінгвістики було формування неологічної лексики шляхом семантичної і стилістичної модифікації понять і конструкцій, що існували раніше, але наразі набули нового специфічного емоційного й смислового навантаження, формуючи нові образи у світовій суспільній думці. Наприклад, такі слова як *Orkowie* та *Mordor* були запозичені до українського медіа-дискурсу з творчості одного із засновників високого фентезі Дж. Р.Р. Толкіна. Не лише в українських, але й у текстах польських новин ці слова почали використовуватися на позначення країни-агресорки і російської армії: *Orki ... znowu porywają kierowników władz lokalnych* (wnp.pl).

3.2.4. Іронічне зображення риторики російської пропаганди

Досить незвичною та дещо парадоксальною виявилася доля деяких виразів з російського політичного дискурсу, які спершу були частиною риторики пропаганди. Внаслідок частого уживання у текстах новин та через переосмислення в українському суспільно-політичному дискурсі такі слова увійшли до польського медіа-простору, але

не в первинному значенні, а з іронічною конотацією. Перелік таких слів представлено у табл. 3.6.

Іронічний відтінок у вживанні даних виразів можна прослідкувати й у текстах польських ЗМІ: *Najwyraźniej tak wiele jest takich "planowanych przegrupowań", że ... musiał zmienić nie tylko język rosyjski, ale i geografię, by je ukryć* (wiadomosci.onet.pl); *Po bojowych nietoperzach czas na zarażone komary* (tvp.info); *...zestrzelili własny śmigłowiec. "Kolejny gest dobrej woli"* (i.pl); *...ponownie mówił o "amerykańskich biolaboratoriach założonych w Ukrainie"* (wiadomosci.wp.pl).

Таблиця 3.6.

**Слова, що іронічно відображають риторику російської пропаганди,
що були запозичені з українського суспільного дискурсу до текстів
польських новин в умовах воєнного часу**

Вираз	Частотність вживання			Частка згадувань, %
	січень 2021 р. – січень 2022 р.	лютий 2022 р. – січень 2023 р.	Всього	лютий 2022 р. – січень 2023 р.
<i>Specjalna operacja wojskowa</i>	3	4056	4059	100
<i>Częściowa mobilizacja</i>	2	3711	3713	100
<i>Druga armia świata</i>	0	464	464	100
<i>Bojowe komary, nietoperze,</i>	0	69	69	100
<i>Mobilizacja</i>	0	51	51	100
<i>Gest dobrej woli</i>	0	43	43	100
<i>Biolaboratoria</i>	0	37	37	100
<i>Planowe przegrupowanie</i>	2	12	14	86
<i>Nie mający analogów</i>	1	5	6	83
<i>Czymbiki</i>	0	5	5	100

Джерело: побудовано на основі даних на платформі MoncoPL (<http://monco.frazeo.pl/>).

Хоча більшість слів з цієї групи не вживалися в польських медіа до лютого 2022 р., протягом 2022 р. – початку 2023 р. вони почали досить активно згадуватися у текстах новин в контексті війни та пов'язаних з нею подій. Як видно з рис. 3.3., більшість слів даної групи висвітлювалися у польських новинах з досить регулярною періодичністю, не зникаючи з публічного дискурсу на тривалий час.

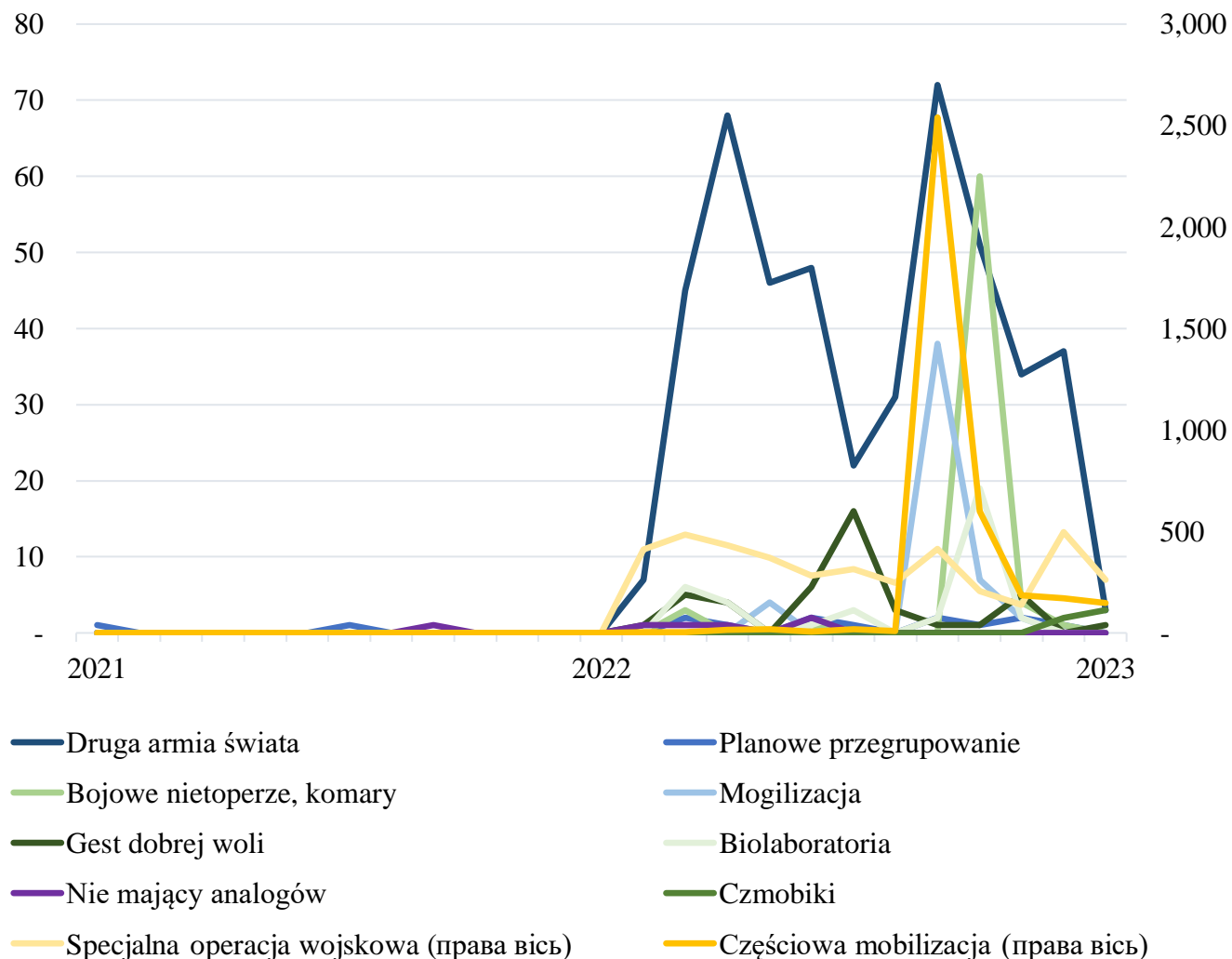


Рис. 3.3. Частотність вживання українських виразів з групи «Іронічне зображення риторики пропаганди», що найбільше згадувалися у польських новинах протягом лютого 2022 р. – січня 2023 р.

Джерело: побудовано на основі даних на платформі MoncoPL (<http://monco.frazeo.pl/>).

Вагому частину серед новітніх запозичень, що увійшли до польської мови протягом 2022 р. – початку 2023 р., складає група українських топонімів, які набули символічного значення з початку повномасштабних бойових дій (всього виявлено 23 слова).

Цікавим також є аналіз особливостей вживання українських культурем, які з'явилися у польській мові кілька століть тому, але досі функціонують у засобах масової інформації в різноманітних контекстах. Детальний перелік топонімів і культурем, а також особливостей їх уживання протягом періоду аналізу, представлено далі у тексті роботи.

У табл. 3.7. наведено перелік інших запозичень-українізмів, які не було класифіковано до жодної з тематичних груп, адже вони різняться за сферою уживання, емоційним забарвленням або семантичним навантаженням.

Таблиця 3.7.

Інші запозичення з українського суспільного дискурсу до текстів польських новин в умовах воєнного часу

Вираз	Частотність вживання			Частка згадувань, %
	січень 2021 р. – січень 2022 р.	лютий 2022 р. – січень 2023 р.	Всього	лютий 2022 р. – січень 2023 р.
<i>przesiedleńcy/ uchodźcy z Ukrainy</i>	167	56061	56228	100
<i>Wojna/inwazja na pełną skalę</i>	35	1127	1162	97
<i>Alarm powietrzny</i>	5	1023	1028	100
<i>Umowa zbożowa</i>	0	512	512	100
<i>Wojsko informacyjne</i>	0	23	23	100

Джерело: побудовано на основі даних на платформі MoncoPL (<http://monco.frazeo.pl/>).

Перераховані у табл. 3.7. лексичні одиниці відображають реалії повсякденного життя багатьох мешканців України за умов воєнного часу. Хоча окремі елементи цих

виразів вживалися у текстах польських новин і до початку 2022 р., вони набули нові варіанти лексичних та/або конотаційних значень протягом останнього часу.

Вираз *przesiedleńcy/uchodźcy z Ukrainy* (найчастіше уживаний вираз в цій тематичній групі) у польських та українських новинах ілюструє солідарність польського народу з вимушеними українськими переселенцями, а також демонструє тісні культурні, економічні та соціальні зв'язки між двома країнами.

На основі аналізу десяти найбільш частотних колокатів виразу «*przesiedleńcy/uchodźcy z Ukrainy*», що представлені на рис. 3.4., можна зробити висновок, що більшість контекстів уживання цих слів у польському медіа-дискурсі стосується передусім гуманітарної допомоги тимчасовим мігрантам з України, їх поселенню, забезпеченню необхідними засобами для комфортного проживання: серед найбільш пов'язаних колокатів близько 73% випадків уживання належать до лексико-семантичної групи «допомога».

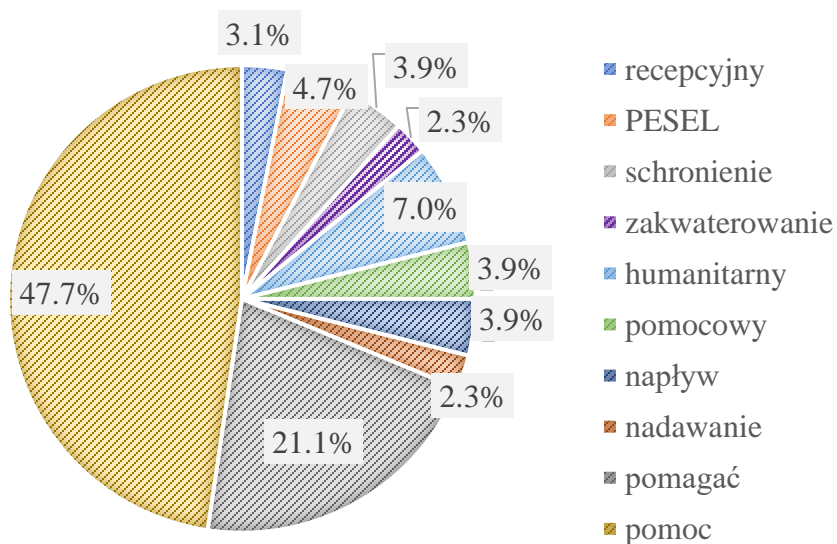


Рис. 3.4. Топ-10 найчастотніших колокатів виразу «*przesiedleńcy/uchodźcy z Ukrainy*» у корпусі польських новин за період з лютого 2022 р. по січень 2023 р.

Джерело: побудовано на основі даних на платформі MoncoPL (<http://monco.frazeo.pl/>).

Це спостереження підтверджується й вибіркоким аналізом конкордансу, як представлено у наступних прикладах: *Każda osoba, która jest **uchodźcą** wojennym z*

Ukrainy, może skorzystać ze wsparcia, którego udziela Ośrodek Pomocy... (radioolsztyn.pl); Przebywającym w Polsce uchodźcom z Ukrainy przysługuje jednorazowe świadczenie pieniężne w wysokości 300 zł, przeznaczone na najpilniejsze potrzeby (wnp.pl); ...by zapewnić finansowanie wszelkich działań i projektów niezbędnych do pomocy (i integracji) uchodźców ukraińskich (wgospodarce.pl); Rada Ministrów planuje przyjąć do końca marca projekt ustawy, która m.in. zwolni z opłat za dostęp do infrastruktury kolejowej tych przewoźników, którzy uruchomią pociągi przeznaczone do przewozu uchodźców z Ukrainy (wnp.pl).

Отже, корпусний аналіз польських новин з січня 2021 р. до січня 2023 р. дозволив визначити, які зі слів, що з'явилися або почали використовуватися з новою конотацією в українському медіа-дискурсі, були запозичені до польського суспільного мовлення з початком повномасштабних воєнних дій. Вибірку таких слів було розподілено за кількома тематичними групами: символи героїзму та національної ідентичності; назви й види зброї; новітні реалії, породжені війною; іронічне зображення риторики російської пропаганди; культуреми, топоніми та інші.

3.3. Новий погляд на українські культуреми

Внаслідок тісних і тривалих контактів між двома народами спостерігається не лише запозичення елементів однієї мови до іншої, а й відбувається взаємне проникнення й збагачення культур цих народів, що має безпосередній вплив на розвиток мовних систем, адже культура й мова є нерозривними чинниками формування одна одної.

Особливе місце серед запозичень посідають *культуреми*, які відображають колективну пам'ять іншого народу, його самоідентифікацію та формують його національну мовну картину світу.

Існує декілька підходів до визначення поняття «культурема». Відповідно до А. Нагурко, М. Лажінського та Х. Бурхарда (2004), М. Рака (2015), це ключові слова, які є важливими для самовираження певного суспільства, характеризують його картину

світу, успадковані цінності, традиції та поточне світосприйняття. О. Палатовська (2023) зазначає, що в сучасних дослідженнях культурами трактують як символ колективної пам'яті та самоідентифікації певної етнічної спільноти, які можуть формуватися у відповідному культурно-історичному контексті (Палатовська, 2023, с.275). В етнолінгвістичному розумінні культурами визначають як гіпероніми терміну «ключові слова», які часто вживаються у певній культурі та здатні її характеризувати (Rak, 2015, с.308-309).

Л. Надаць (2009) визначила три основні функції, які виконують культурами: естетична, когнітивно-герменевтична та аргументативну. Відповідно, культурами характеризуються:

- 1) стійкістю та життєздатністю, функціонуючи у сучасній мові;
- 2) фразеологічною пластичністю, утворюючи фразеологічні вузли;
- 3) домінуванням конотації над денотатом у семантичній структурі слова (Nadal, 2009, с. 104-107).

М. Рак (2015) зазначає, що через відсутність усталених об'єктивних процедур ідентифікації культурем, цей процес ґрунтується на інтуїтивному виборі та залежить від суб'єктивного світосприйняття дослідника. Через це існує ризик надінтерпретації, а також відсутності єдиної типології слів-культуремів: слова, що сприймаються деякими науковцями як культурами, можуть не мати такого статусу для інших (Rak, 2015, с.312).

До деяких польських культурем можна віднести наступні слова у кожній категорії:

- 1) національні символи: *ojczyzna, szlachta, ziemiaństwo*;
- 2) топографічні назви: *Częstochowa, Wawel, Westerplatte, Wisła*;
- 3) національні риси та цінності: *honor, pracowitość, wolność*;
- 4) символи культурної самобутності: *bigos, chleb, kotlet schabowy* (J. Bartmiński, 2006; M. Rak, 2015; A. Nagórko, 2004).

Вивчення культурем досі залишається актуальним дослідницьким питанням у польській та українській лінгвістиці. В рамках польсько-української митецької виставки «Словник українізмів. Інтерпретація польських та українських ілюстраторів», яка

відбувалася влітку 2022 року, було сформовано набір з дванадцяти слів-українізмів та відповідних ілюстрацій до кожного з них (*Słownik ukrainizmów*). Матеріали даного словника слугували основною для корпусного дослідження українських культуремів, що були запозичені польською мовою, як представлено далі у тексті роботи. Хоча більшість із розглянутих українізмів увійшли до польської мови протягом попередніх століть, в умовах воєнного часу вони актуалізують своє значення та конотативне забарвлення, відображаючи риси та культурний код українського народу.

Щоб перевірити, які з даних культурем функціонують у сучасному польському медіа-дискурсі, було проведено оглядове корпусне дослідження частотності з використанням інструментарію MoncoPL.

Аналіз частотності уживання культурем проводився для періоду з січня 2021 р. до січня 2023 р. відповідно до класифікації, запропонованої у праці О. Палатовської (Палатовська, 2023):

1. Слова-символи України (*Majdan* / Майдан, *step* / степ, *bandura* / бандура).
2. Українізми-етнографізми (*mereżka* / мережка, *kureń* / курінь, *koromysło* / коромисло, *horak* / гопак, *szarawary* / шаровари).
3. Українсько-польські міжмовні омоніми (*bogomaz* / богомаз, *wataha* / ватага, *chmara* / хмара).

Слова-символи України

Загальна динаміка частотності згадування слів *Majdan*, *step*, *bandura* у польських медіа наведено на рис. 3.5. Слова *Majdan* та *step* регулярно й часто згадувалися протягом періоду аналізу у різноманітних контекстах, зокрема й у ролі українських культурем. Як продемонстровано на рис. 3.5., пік згадування слова *Majdan* припадає на початок повномасштабних воєнних дій проти України, що підтверджує міцну асоціацію цієї лексеми з волелюбством і жагою до свободи народу України.

Хоча слово *majdan* є загальноживаним у польській мові і позначає передусім площу у населеному пункті, двір, місце торгівлі й зберігання товарів або навіть великий багаж, воно набуло додаткового значення протягом останнього десятиліття у зв'язку з

суспільно-політичними процесами в Україні. Як ілюструє наступний приклад з польських новин від липня 2021 р., це слово вже досить тривалий час існує у суспільній свідомості поляків та набуло специфічного конотаційного забарвлення, асоціюючись з народним волевиявленням, свободою, незламністю чи протестом: «*Wzywał m.in. do „polskiego Majdanu”*» (dziennikwschodni.pl). У іншому прикладі вираз «часи Майдану» вживається у символічному значенні: «*Od ośmiu lat, od czasów Majdanu, robią to Ukraińcy*» (wiadomosci.wp.pl).

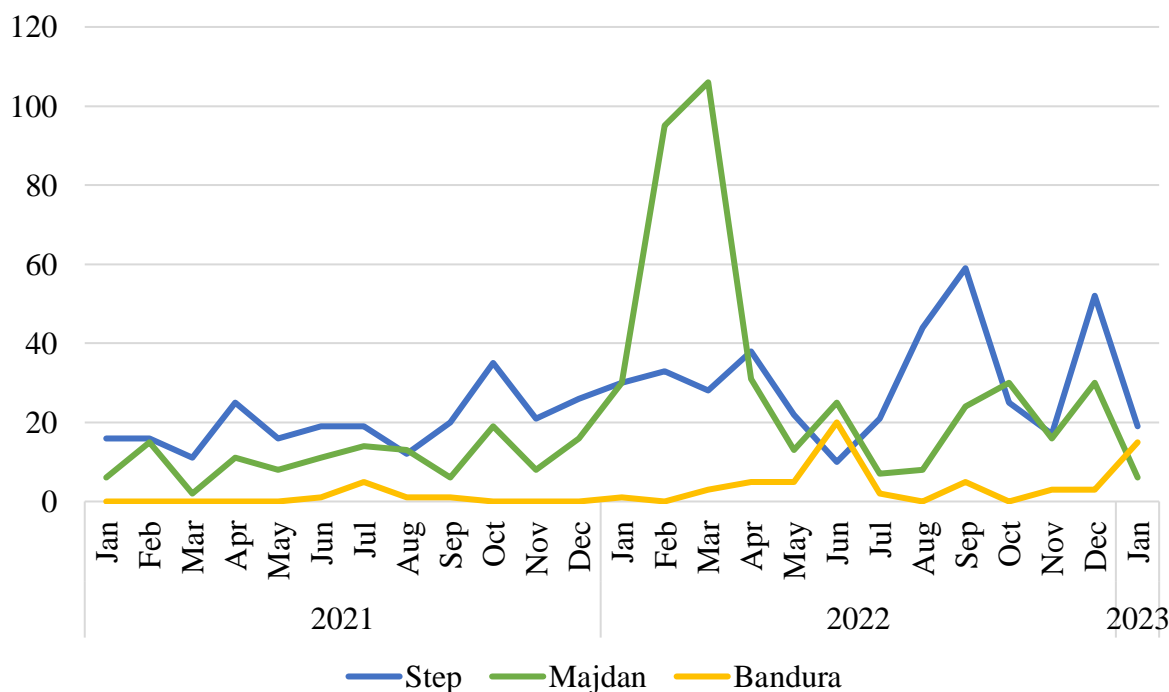


Рис. 3.4. Частотність вживання українських культуремів *step*, *majdan*, *bandura* у польських новинах протягом лютого 2022 р. – січня 2023 р.

Джерело: побудовано на основі даних на платформі MoncoPL (<http://monco.frazeo.pl/>).

Слово *step* вживалося у польській мові не лише в загальному нейтральному значенні трав'янистої природної смуги помірного поясу, а й у контексті символів України – її жаги до свободи та нескореності. Якщо раніше слово *step* вживалося передусім в історико-культурному аспекті, то з лютого 2022 р. воно набуває сучасного символізму, адже в українському степі відбувається боротьба за свободу, як це було протягом минулих століть: «...*na bezkresnych stepach Ukrainy, na której odbędzie się*

wcześniej lub później walka o demokrację XXI wieku» (dorzeczy.pl). Окрім цього, ця українська культура вживається й у традиційному художньо-символічному значенні: «...którego akcja toczy się w znacznej części na bezkresnych **stepach** Ukrainy» (gk24.pl); «Obok tradycyjnych utworów wywodzących się z ukraińskich **stepów** i chat, muzycy zagrali kilka kompozycji pochodzących z Polski» (i.pl).

Слово-символ української музичної культури *bandura* також згадувалося в польських новинах протягом досліджуваного періоду, однак зі значно меншою частотою. Вибірковий аналіз текстів новин демонструє, що цей музичний інструмент дуже часто вживався на позначення української культури: «...najpiękniejsze ukraińskie kolędy przy towarzyszeniu **bandur** - tradycyjnych instrumentów strunowych» (czluchow.naszemiasto.pl); «...bohatera ukraińskich podań ludowych, przedstawianego z tradycyjnym ukraińskim instrumentem **bandurą**» (polsatnews.pl).

Українізми-етнографізми

Друга група культурем охоплює обмежені етнографізми, які знайомлять поляків з культурою України. Хоча вони увійшли до польської мови вже досить давно, вони досі використовуються у медіа-дискурсі. Контекст згадування цих слів є досить вузьким та пов'язаний передусім з українською культурою, народними традиціями та мистецтвом, як проілюстровано нижче:

1. **Mereżka:** «Uczestniczki projektu już wcześniej dowiedziały się, że **mereżka** jest najpopularniejszym wykończeniem haftów w naszym regionie» (chojnice24.pl); «Mają do tego fartuszki z mereżką, to jest taki haft» (radio.bialystok.pl).
2. **Kureń:** «W Kaliszu wydawano następujące pisma: „Zaliczny Strileć”, miesięcznik literacki „Weselka”, satyryczne „Oko”, „Sicz” „Za drotom”, „Dżereło”, „Wowa internowanyj”, „Za Ukrainu” i „Wisnyk internowanego **kurenia**”» (ostrow.naszemiasto.pl).
3. **Koromysło:** «... dobrze znany kołowrotek i samowar elektryczny, ale też mniej rozpoznawalne **koromysło** czyli drewniany element do zawieszenia i przenoszenia wiader z wodą» (hajnowka.naszemiasto.pl).

4. **Хорак:** «*Bojowy **hopak** to słowiańska, kozacka sztuka walki*» (poranny.pl); «*Chodzi też o łamanie stereotypu o kulturze ukraińskiej. Chcemy pokazać, że tradycje te, to o wiele więcej niż tylko pierogi albo **hopak**, czyli taniec ludowy*» (warszawa.naszemiasto.pl).
5. **Шаравары:** «*Żuawi nosili charakterystyczne **szarawary**, wyposażeni byli zarówno w karabin, jak i w różne rodzaje broni białej*» (i.pl).

Українсько-польські міжмовні омоніми

Особливістю слів з даної групи є те, що вони зазнали суттєвих семантичних трансформацій у польській мові, змінивши своє оригінальне значення, що було запозичено з української мови.

Наприклад, слово *bohomasz*, яке в українській мові означає «іконописець» (Палатовська, 2023, с.279), у польській мові набуло значення «неякісний малюнок», що можна проілюструвати уривком з польських новин: «*Co można już zaliczyć do sztuki z gatunku "street art", a co jest jedynie wysmarowanym **bohomaszem**, bez żadnej wartości, ba! aktem prymitywizmu, wyrazem barbarzyństwa, zwykłą dewastacją zabytkowych i nowych dynków?*».

Слово *wataha* звузило своє значення з «великої групи людей» в українській мові до «небезпечної групи людей» «стада звірів чи вовків» у польській мові: «***Wataha** wilków zaatakowała stado krów*» (fakt.pl); «*Zwykle wataha liczy od 4 do 6 osobników - jest to zwykle para dorosłych wilków i ich wilcze potomstwo z ostatniego roku-dwóch*» (rmf24.pl).

Слово *chmara* навпаки розширило значення зі «скупчення водяної пари в тропосфері» в українській мові до «великої кількості когось або чогось» у польській мові: «*Pomocni urzędnicy trzepoczą wokół człowieka jak **chmara** motyli*» (liberte.pl); «*od najmłodszych lat przyzwyczajona jest do wszędobylskich **chmar** paparazzi, które śledzą jej każdy krok, gdy przemierza ulice Nowego Jorku*» (pudelek.pl).

Отже, українські культурери, що увійшли до польської мови протягом багатьох попередніх століть, на сьогодні набули в ній значного поширення і дещо нового

семантичного наповнення, про що свідчить високий рівень частотності та регулярності їх уживання (окрім українізмів-етнографізмів).

3.4. Неологічні топоніми в текстах польських ЗМІ

Після початку повномасштабного вторгнення та введення воєного стану військово-політичні події в Україні почали активно висвітлюватися у засобах масової інформації по всьому світу, зокрема й у Польщі, що передбачало широке використання топонімічної лексики, передусім ойконімів – назв населених пунктів.

Оскільки більшість з цих географічних назв України стали активно згадуватися у текстах ЗМІ і, відповідно, стали широко відомими для більшості носіїв польської мови лише з лютого 2022 р., то існують підстави вважати їх неологічними.

Основні положення та результати дослідження неологічних українізмів-ойконімів, представлених далі, були висвітлені у статті О. Палатовської та Є. Сиви «Неологічні українізми в мові польських новин: січень 2021 – січень 2023 (корпусне дослідження)» (2023).

На першому етапі дослідження було сформовано первинну вибірку ойконімів, які є неологічними для польського медіа-дискурсу, що охоплювала майже 150 назв населених пунктів. При дослідженні враховувалися також деривативи ойконімів у різних відмінкових формах (наприклад, *Wolnowacha*, *Wolnowachi*, *Wolnowachq*, *Wolnowaski*). Первинна вибірка формувалася на основі аналізу текстів українських і польських новин за період з січня 2021 р. по січень 2023 р., зокрема з використанням комп'ютеризованого корпусу MoncoPL (<http://monco.frazeo.pl/>).

Для перевірки ойконімів застосовувалися такі критерії:

- 1) уживання назви населеного пункту до повномасштабних воєнних дій (січень 2021 р. – січень 2022 р.) не перевищувало 30 згадувань;
- 2) згадування назви населеного пункту після початку введення воєнного стану перевищувало 1 000 разів (лютий 2022 р. – лютий 2023 р.).

За результатами емпіричного дослідження з використанням корпусу MoncoPL було виявлено 23 населені пункти, що одночасно відповідали зазначеним вище критеріям, що представлені у табл. 3.8.

Таблиця 3.8.

**Неологічні українській ойконіми,
що почали активно згадуватися у текстах польських ЗМІ
після початку повномасштабних воєнних дій в Україні**

Населений пункт	Частотність вживання			Частка згадувань, %
	січень 2021 р. – січень 2022 р.	лютий 2022 р. – січень 2023 р.	Всього	лютий 2022 р. – січень 2023 р.
<i>Bucz</i>	8	18,775	18,783	100
<i>Bachmut</i>	11	12,406	12,417	100
<i>Siewierodoneck</i>	1	11,924	11,925	100
<i>Izium</i>	5	9,013	9,018	100
<i>Irpiń</i>	4	6,781	6,785	100
<i>Lisiczańsk</i>	15	5,795	5,810	100
<i>Melitopol</i>	2	5,344	5,346	100
<i>Hostomel</i>	6	4,665	4,671	100
<i>Kramatorsk</i>	7	4,319	4,326	100
<i>Słowiańsk</i>	7	4,086	4,093	100
<i>Enerhodar</i>	2	3,699	3,701	100
<i>Borodzianka</i>	-	3,648	3,648	100
<i>Kachowka</i>	1	3,060	3,061	100
<i>Popasna</i>	3	2,764	2,767	100
<i>Rubiżne</i>	-	2,570	2,570	100
<i>Berdiańsk</i>	6	2,295	2,301	100
<i>Awdijiwka</i>	16	2,214	2,230	99
<i>Ołeniwka</i>	-	2,019	2,019	100
<i>Kreminna</i>	-	1,853	1,853	100
<i>Wołnowacha</i>	2	1,452	1,454	100
<i>Ochtyrka</i>	1	1,087	1,088	100

<i>Balaklija</i>	-	1,083	1,083	100
<i>Swatowe</i>	2	1,061	1,063	100

Джерело: побудовано на основі даних на платформі MoncoPL (<http://monco.frazeo.pl/>).

Хоча при дослідженні корпусу новинних текстів перевірялися різні варіанти правопису назв українських населених пунктів польською мовою, а нерелевантні спільнокореневі слова виключалися, все ж показники частотності можуть містити незначні похибки, що є загалом притаманним для досліджень подібного спрямування й не має вирішального впливу на одержані результати.

Для вивчення особливостей контексту вживання визначені неологічні ойконіми було відображено на карті України на рис. 3.5. (розмір позначень пропорційний до частоти вживання відповідного ойконіму).

На рис. 3.5. відображено, що найчастіше в текстах польських ЗМІ згадувалися ті населені пункти, де відбувалися трагічні події у період окупації, що набули широкого розголосу у світових медіа (*Buczka, Irpiń, Hostomel, Borodziańska*), а також міста й містечка, які були розташовані поруч із лінією бойових дій (*Bachmut, Siewierodoneck, Popasna, Rubiżne*), або деокуповані території (*Izium, Balaklija*).

Динаміку частотності згадування п'яти найуживаніших українізмів-ойконімів протягом лютого 2022 р. – січня 2023 р. представлено на рис.3.6. Можна спостерігати, що пік висвітлення виявлених неологічних ойконімів у засобах масової інформації кореляє з періодом, коли відбувалися події, що безпосередньо пов'язані з цими населеними пунктами. Наприклад:

1. *Buczka, Irpiń, Hostomel, Borodziańska* (березень-квітень 2022 р.) – період трагічних подій окупації.
2. *Siewierodoneck* і *Lisiczańsk* (червень 2022 р.) – період захоплення міст.
3. *Izium* (квітень 2022 р. та вересень 2022 р.) – події, що пов'язані з окупацією та звільненням міста, відповідно.
4. *Bachmut* (липень 2022 р., грудень 2022 р., січень 2023 р.) – періоди активних бойових дій та загострення ситуації на полі бою.

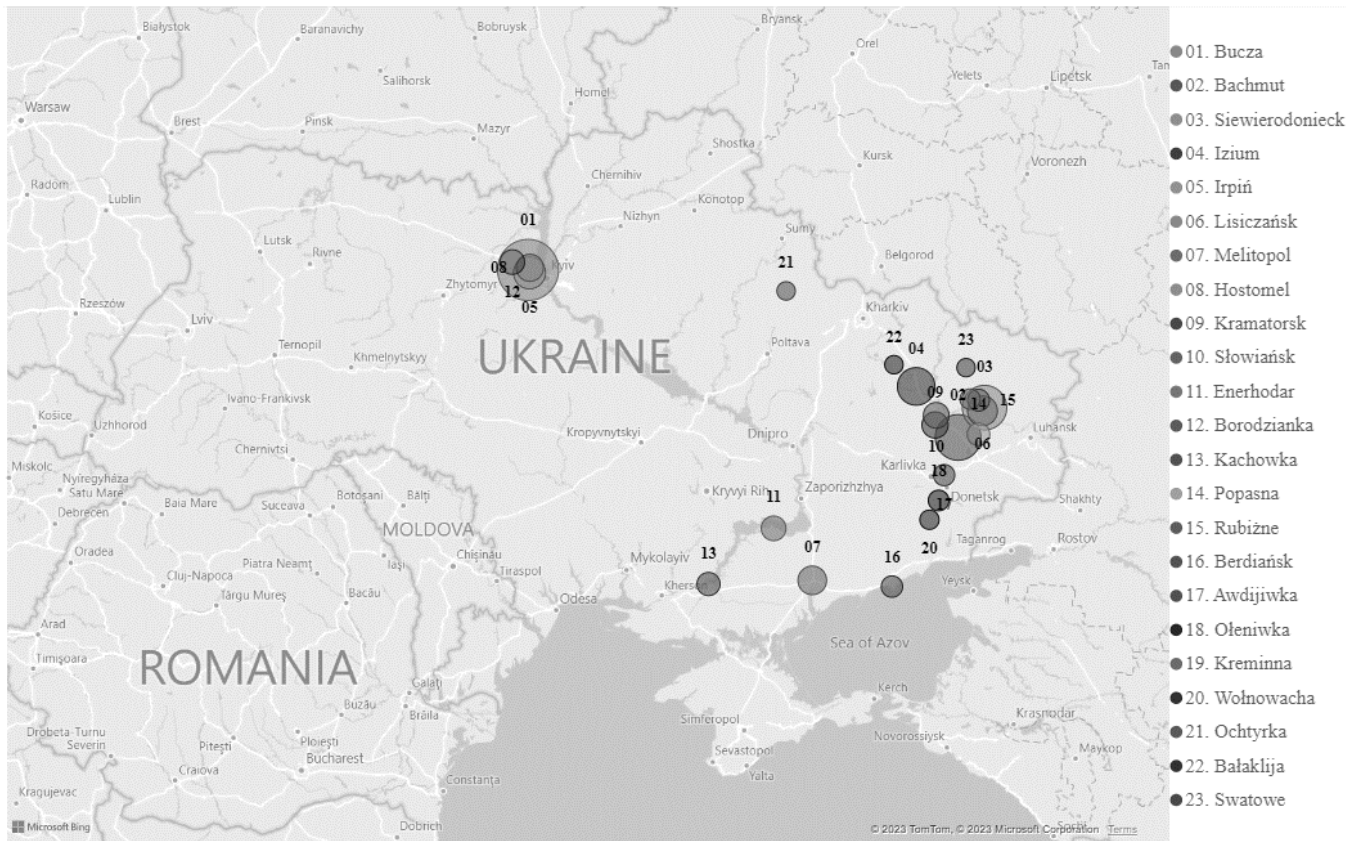


Рис. 3.5. Розподіл визначених неологічних ойконімів на карті України з урахуванням частотності їх вживання у текстах польських новин протягом лютого 2022 р. – січня 2023 р.

Джерело: побудовано на основі даних на платформі MoncoPL (<http://monco.frazeo.pl/>)

На думку дослідників, частотність вживання лексичних одиниць тісно пов'язана з явищем *салієнтності*, тобто поступової активації у свідомості та короткочасній пам'яті людини набору слів, що часто використовуються у публічних й приватних комунікативних актах (Палатовська, & Сова, 2023, с.274-275). Феномен салієнтності проявляється внаслідок регулярної та активної когнітивної обробки мовних одиниць, що широко висвітлюються у медійному дискурсі, у свідомості носіїв мови, внаслідок чого формуються щільні асоціативні зв'язки з цими словами.

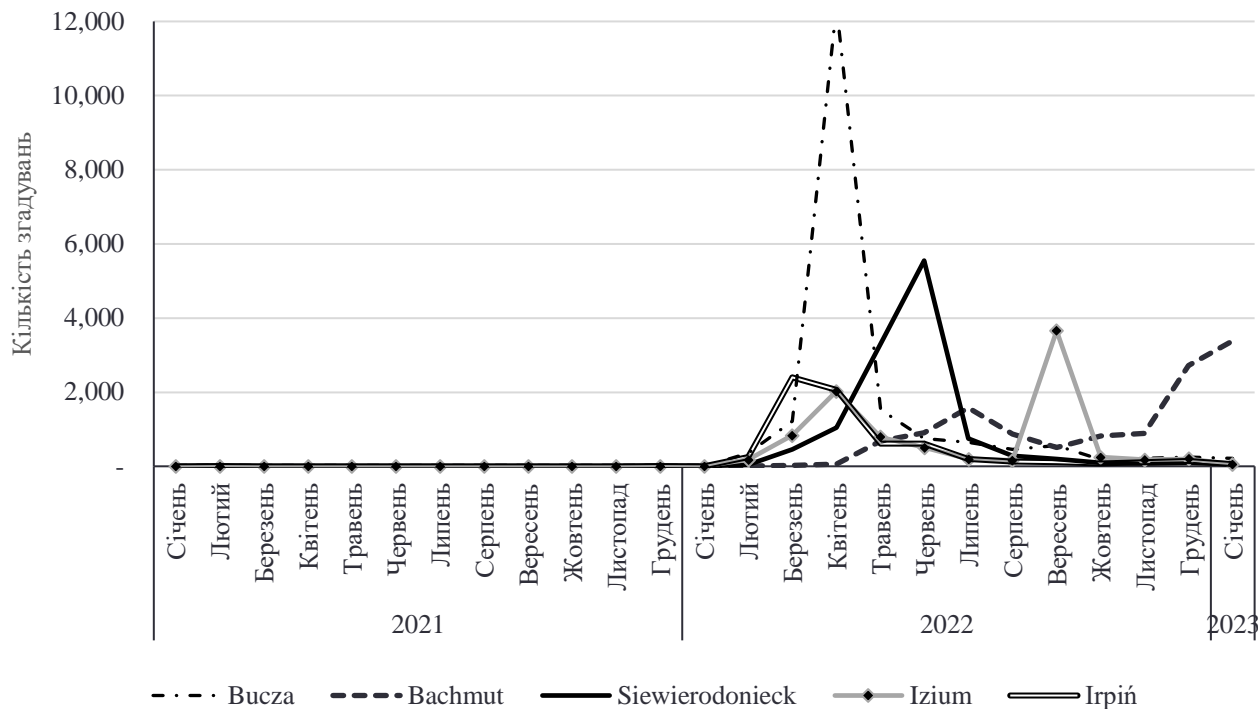


Рис. 3.6. П'ять найуживаніших неологічних українізмів-ойконімів упродовж лютого 2022 р. – січня 2023 р. у текстах польських новин

Джерело: побудовано на основі даних на платформі MoncoPL (<http://monco.frazeo.pl/>).

Зважаючи на те, що механізми салієнтності відіграють ключову роль у процесі запозичення іншомовних слів, можна припустити, що найбільш частотні неологічні українізми-ойконіми мають високий потенціал входження до польської мовної картини світу у середньостроковій перспективі. Водночас варто враховувати, що салієнтність лексичних одиниць не є постійною, а може змінювати свій ступінь залежно від низки лінгвальних та екстралінгвальних чинників. Окрім цього, існує ризик, що неологічні запозичення можуть лише тимчасово увійти до активного лексикону польської мовної спільноти, поступово зникнувши з суспільного мовлення внаслідок зниження інтенсивності вживання у текстах польських ЗМІ.

Висновки до розділу 3

Перші запозичення з української мови потрапили до польської ще у XIV ст., а найбільш активне імпортування слів припало на XVI – XIX ст. Більшість українізмів у польській мові пов'язані з людиною, її побутом та культурою, а також природою.

Основними чинниками, що сприяли активному поширенню українізмів у польській мові, була активна творча діяльність письменників-білінгвів, а також тривалі й тісні суспільно-політичні, культурно-історичні та економічні контакти між українським і польським народами.

З початком повномасштабних військових дій в Україні з лютого 2022 р. в українській мові активізувався процес формування неологізмів, що безпосередньо пов'язані з військовими реаліями, символами героїзму та національної ідентичності тощо. За посередництвом медіа та завдяки тісним контактам між українським і польським народами частина неологічних українізмів поступово запозичується польською мовою.

У процесі емпіричного дослідження було ідентифіковано близько 300 потенційних неологічних українізмів, зокрема слів, що з'явилися в українській мові після початку повномасштабного вторгнення, а також виразів, які вже існували, проте суттєво змінили своє значення та/або конотаційне забарвлення. За результатами перевірки частотності вживання кожного з виразів на основі моніторингового корпусу польських новин MonsoPL було сформовано фінальну вибірку, що містить понад 100 слів, які були запозичені до польського медіа-дискурсу протягом лютого 2022 р. – січня 2023 р. за посередництвом української мови. Кожен із виразів було класифіковано до однієї з таких тематичних груп: символи героїзму та національної ідентичності українців; назви й види зброї; новітні реалії, породжені війною; іронічне зображення риторики росвійської пропаганди; українські культурами та топоніми; а також інші запозичення. Проведений аналіз частотності, колокацій та конкордансу для найбільш згадуваних

українізмів дозволив визначити контекст їх уживання, а також потенціал до запозичення польською мовою.

Корпусно-медійний аналіз українських культурем, запозичених до польської мови у попередніх століттях, дозволив з'ясувати, що більшість з них не втратила актуальності, а навпаки, на сьогодні підвищила активність уживання у сучасній польській мові.

За результатами поглибленого дослідження неологічних українізмів-топонімів, було визначено 23 ойконіми, що регулярно та широко висвітлювалися у польських засобах масової інформації протягом лютого 2022 р. – січня 2023 р.

З огляду на високу кореляцію між частотністю уживання та салієнтністю мовних одиниць, існують достатні підстави припускати, що значна частина виявлених українізмів, які активно згадувалися у польському медіа-дискурсі протягом досліджуваного періоду, може бути запозичена до активного лексикону носіїв польської мови у середньостроковій перспективі.

ВИСНОВКИ

У роботі здійснено ґрунтовне теоретико-методологічне обґрунтування сутності запозичень іншомовних слів і чинників їх входження до мови-реципієнта, а також проведено емпіричне корпусне дослідження новітніх українізмів, що були запозичені польськими засобами масової інформації після початку повномасштабних воєнних дій в Україні. У процесі дослідження було виявлено близько 100 українізмів, запозичених до активного складу польської мови за посередництвом засобів масової інформації.

За результатами проведеного дослідження було вирішено усі поставлені завдання та сформовано такі висновки:

1. У сучасному трактуванні поняття «запозичення» відображає не лише результат, а й процес входження та адаптації іншомовних слів до системи мови-реципієнта. Більшість запозичень охоплюють переважно лексико-семантичний рівень мови. У лінгвістиці існує багато підходів до класифікації іншомовних запозичень: залежно від ступеня морфологічної субституції (лексичні запозичення без субституції, напівкальки з частковою субституцією, кальки), залежно від мови походження (англіцизми, галліцизми, іспанізми, українізми, японізми тощо), залежно від близькості до мови-джерела (прямі, непрямі, інтернаціоналізми) та ін.

2. Основними передумовами до запозичення іншомовних слів є існування тісних й тривалих контактів між мовою-донором та мовою-реципієнтом, активна мовленнєва діяльність білінгвів, глобальні інтеграційні процеси у різних сферах суспільного, культурного, політичного й економічного життя, а також інші позалінгвальні чинники. Ступінь та швидкість адаптації іншомовних слів залежить від внутрішньої стійкості мовної системи до зовнішніх змін, а також низки соціолінгвістичних і психологічних чинників.

3. Внаслідок тісних і тривалих контактів між українським і польським народами перші підтвержені випадки запозичень українізмів до польської мови відбувалися ще у XIV-XV ст. (напр., *ataman*, *rohatyňa*, *proskura*), а найбільших

масштабів досягли на протязі XVI – XVII ст. (напр., *bohater, chata, hreczka, kobza, wataga, oczeret, bandura, czereda, poroh*) і XVIII – XIX ст. (напр., *hajdamaka, sicz, step, baraban, czumak, hopak, kutia*). Більшість із них безпосередньо пов'язані з культурою і побутом людини, а також історичними подіями відповідної доби.

4. Для ідентифікації, вивчення частотності та контексту уживання сучасних українізмів, що були запозичені до польської мови протягом 2022 р. – початку 2023 р., доцільним є використання методології корпусної лінгвістики, яка є одним із найперспективніших наукових напрямів на сучасному етапі розвитку лінгвістичної науки. Цей напрям характеризується міждисциплінарним підходом, поєднуючи інструментарій загальної лінгвістичної теорії, когнітивістики, інформаційних та комп'ютерних технологій, статистики й математичних наук. Однією з переваг цього підходу є застосування об'єктивних методів вивчення мовного матеріалу у первинному вигляді, що відрізняє її від багатьох інших напрямів лінгвістики, які покладаються на суб'єктивний підхід. Водночас корпусна лінгвістика дозволяє визначити лише факт існування певного мовного елементу або явища, однак підтверджує коректність та не пояснює причину його функціонування в такому вигляді. Основними методами сучасного корпусного дослідження є аналіз частотності; аналіз колокацій та конкордансу; лексико-семантичний аналіз тексту тощо.

5. Дослідження особливостей вживання неологічних українізмів у сучасному польському медіа-дискурсі було проведено на основі моніторингового корпусу польських новин, який постійно наповнюється новим мовним матеріалом. У процесі дослідження було розроблено набір критеріїв, згідно з якими до потенційних неологічних українізмів було відібрано ті лексичні одиниці, що почали активно вживатися у польських ЗМІ переважно з лютого 2022 р. до січня 2023 р., але водночас не використовувалися чи вживалися в інших вузьких контекстах до цього періоду.

6. У процесі емпіричного корпусно-медійного аналізу було сформовано первинну вибірку з майже 300 потенційних неологічних українізмів, що з'явилися в українській мові або суттєво змінили своє значення та/або конотацію після початку

повномасштабного вторгнення. Використовуючи інструментарій корпусного методу на порталі MonsoPL, було перевірено кожну лексичну одиницю, що увійшла до первинної вибірки.

7. За результатами проведеного дослідження було встановлено понад 100 українізмів, які активно використовувалися у польському медійному дискурсі переважно з лютого 2022 р. Ідентифіковані в текстах польських новин українізми було розподілено за кількома тематичними групами: символи героїзму та національної ідентичності українців (усього 11 лексичних одиниць, зокрема: *czzerwona kalina, Dobry wieczór, jesteśmy z Ukrainy, pies Patron, Siły Zbrojne Ukrainy (ZSU), palanysa, oborona terytorialna*); назви й види зброї (всього 32 лексеми, зокрема: *Bajraktar, Kindzał, HIMARS, Stuhna, NASAMS, Patriot, szahed, dron kamikadze, Neptun*); новітні реалії, породжені війною (всього 11 лексем, зокрема: *brudna bomba, filtracja, denazyfikacja, Mordor, orkowie, zombiowani*); іронічне зображення риторики російської пропаганди (всього 10 лексичних одиниць, зокрема: *biolaboratoria, bojowe komary, gest dobrej woli, nie mający analogów, czmobiaki, druga armia świata*); українські культури (усього 11 лексем, зокрема: *mereżka, chmara, step, hopak, wataha, koromysło, bandura, kureń*); топоніми (всього 23 лексеми, зокрема: *Awdjiwka, Bachmut, Balaklija, Borodzianka, Bucza, Hostomel, Kachowka, Melitopol, Swatowe*); інші потенційні неологізми (всього 6 виразів, зокрема: *alarm powietrzny, umowa zbożowa, wojna na pełną skalę, przemieszczający/uchodzący z Ukrainy, wojsko informacyjne*).

Одержані висновки і результати мають практичне значення та характеризуються високим ступенем новизни й актуальності, створюючи підґрунтя до подальших корпусно-медійних та контрастивних досліджень української і польської мов. Вважаємо, що у подальшому перспективним напрямом дослідження було б поглиблене вивчення частотності, активності й контексту уживання лексичних одиниць, що увійшли до фінальної вибірки неологічних українізмів, особливо після завершення бойових дій на території України. Це дозволило б перевірити ступінь адаптації потенційних неологізмів

до польської мови, їх салієнтості у контексті використання в активному лексиконі носіями сучасної польської мови.

PE3IOME

Aktualność tematu badawczego wynika przede wszystkim z praktycznej potrzeby zbadania najnowszych zmian w języku polskim, będących rezultatem wpływu neologicznych ukrainizmów. W kontekście realiów wojskowych, te ukrainizmy aktywnie przenikają do codziennego słownictwa, przemówień polityków oraz tekstów medialnych, nadając im nowe znaczenia i konotacje.

Cele badawcze. Głównym celem badania jest identyfikacja najnowszych ukrainizmów, które zostały zapożyczone do polskich tekstów medialnych w związku z realiami wojskowymi w Ukrainie. Aby osiągnąć ten cel, w trakcie pisania pracy sformułowano i rozwiązano następujące kluczowe zadania:

- wyjaśnienie istoty pojęcia i rodzajów zapożyczeń językowych;
- generalizowanie czynników przenikania i charakterystyki adaptacji zapożyczeń;
- zbadanie społeczno-historycznych i kulturowych uwarunkowań przenikania ukraińskich zapożyczeń do języka polskiego;
- usystematyzowanie i opisanie metodologii wykorzystania korpusów językowych w badaniach językoznawczych;
- opracowanie metodologii badania ukraińskich zapożyczeń w języku polskim w oparciu o badania korpusowe i metody ilościowe w językoznawstwie;
- identyfikacja grup tematycznych ukrainizmów, które zostały zapożyczone przez język polski w czasie wojny;
- weryfikacja, przy użyciu metod badań korpusowych, które neologiczne ukrainizmy zostały zapożyczone do języka polskiego w ostatnim okresie od lutego 2022 r. do stycznia 2023 r.

Obiektem badania są zapożyczenia językowe. **Przedmiotem** badania w ramach zdefiniowanego obiektu są ukrainizmy zapożyczone przez polskie środki masowego przekazu.

Materiał badawczy. Baza informacyjna badania opierała się na polskim korpusie informacyjnym MoncoPL, który jest stale aktualizowany o nowe materiały. Na początku

września 2023 r. korpus ten zawierał około 447 milionów zdań, co odpowiada prawie 8,4 miliarda słów. W ciągu ostatnich dziesięciu lat średni miesięczny przyrost wynosił około 2,7 miliona zdań. Przed przeprowadzeniem badania, utworzono pierwotną próbkę, która obejmowała około 300 słów o różnym ukierunkowaniu tematycznym. Na podstawie wyników badania korpusowego, przy wykorzystaniu narzędzi portalu MoncoPL, utworzono ostateczną próbkę zawierającą ponad 100 ukrainizmów, głównie neologizmów, które zostały zapożyczone z polskiego dyskursu medialnego w okresie od lutego 2022 do stycznia 2023.

W *pierwszym* rozdziale przedstawiono przegląd badań z zakresu zapożyczeń w językoznawstwie, a także omówiono czynniki przenikania i charakterystykę adaptacji pożyczek. W *drugim* rozdziale opisano metodologię badań korpusowych, podejście do ilościowych metod analizy w lingwistyce oraz zdefiniowano etapy badania ukraińskich zapożyczeń. *Trzeci* rozdział zawiera opis społeczno-kulturowego i historycznego tła przenikania ukrainizmów do języka polskiego. W tej części pracy przedstawiono wyniki przeprowadzonego badania empirycznego neologicznych ukrainizmów z różnych grup tematycznych, które zostały przyjęte w tekstach polskich mediów.

Wyniki i wnioski. Praca zawiera dogłębne teoretyczne i metodologiczne uzasadnienie istoty zapożyczeń obcojęzycznych oraz czynników ich przenikania do języka odbiorcy. Ponadto, przeprowadzono empiryczne badanie korpusowe najnowszych ukrainizmów, które zostały zapożyczone przez polskie media po wybuchu działań wojennych na pełną skalę w Ukrainie. W badaniu zidentyfikowano około 100 ukrainizmów, które zostały zapożyczone do języka polskiego za pośrednictwem środków masowego przekazu. Na podstawie wyników badania wszystkie cele zostały osiągnięte i sformułowano następujące wnioski:

- Większość zapożyczeń obejmuje głównie leksykalny i semantyczny poziom języka. Głównymi warunkami zapożyczania obcych słów są bliskie i długotrwałe kontakty między językami dawcy i biorcy, aktywność językowa osób dwujęzycznych, globalne procesy integracyjne w różnych sferach życia społecznego, kulturalnego, politycznego i gospodarczego, a także inne czynniki pozajęzykowe. Ze względu na bliskie i długotrwałe kontakty między narodami ukraińskim i polskim, pierwsze potwierdzone przypadki

zapożyczeń miały miejsce w XIV i XV wieku (np. *ataman, rohatyna, proskura*), a największej skali osiągnęły w XVI-XVII w. (np. *bohater, chata, hreczka, kobza, wataga, oczeret, bandura, czereda, poroh*) i XVIII-XIX w. (np. *hajdamaka, sicz, step, baraban, czumak, hopak, kutia*).

- Językoznawstwo korpusowe jest jednym z najbardziej obiecujących obszarów naukowych na obecnym etapie rozwoju nauk lingwistycznych. Dziedzina ta charakteryzuje się interdyscyplinarnym podejściem, łączącym narzędzia ogólnej teorii lingwistycznej, kognitywistyki, IT, statystyki i matematyki. Do głównych metod współczesnych badań korpusowych należą: analiza częstości występowania wyrazów, analiza kolokacji i konkordancji, analiza leksykalna i semantyczna tekstu itp.
- W badaniu zidentyfikowano ponad 100 ukrainizmów, które aktywnie przenikają do polskiego dyskursu medialnego głównie od lutego 2022 r. Ukrainizmy zidentyfikowane w polskich tekstach informacyjnych zostały podzielone na kilka grup tematycznych: symbole bohaterstwa i tożsamości (łącznie 11 wyrażen, w tym: *czzerwona kalina, Dobry wieczór, jesteśmy z Ukrainy, pies Patron, Siły Zbrojne Ukrainy (ZSU), palanyca*); nazwy i rodzaje broni (łącznie 32 wyrażenia, w tym: *Bajraktar, Kindżał, HIMARS, Stuhna, NASAMS, Patriot, szahed, dron kamikadze, Neptun*); nowa rzeczywistość stworzona przez wojnę (łącznie 11 wyrażen, w tym: *brudna bomba, filtracja, denazyfikacja, zombirowani*); ironiczne ujęcie retoryki propagandowej (łącznie 10 wyrażen, w tym: *biolaboratoria, bojowe komary, gest dobrej woli, nie mający analogów*); kulturemy (łącznie 11 wyrażen, w tym: *mereżka, chmara, step, hopak, wataha, koromysło, bandura, kureń*); nazwy miejscowości (łącznie 23 wyrażenia, w tym: *Awdijiwka, Bachmut, Bałaklija, Borodzianka, Bucza, Hostomel, Kachowka, Melitopol, Swatowe*) oraz inne potencjalne neologizmy.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Ажнюк, Б. М. (2008). Англізми в сучасній українській, російській і чеській мовах. *К.: Мовознавство, (2-3).*
2. Бобкова, Т. В. (2014). Корпус текстів: основні аспекти визначення. *Науковий вісник кафедри ЮНЕСКО Київського національного лінгвістичного університету. Філологія, педагогіка, психологія, (29), 11-20.*
3. Глибовець, А., & Точицький, В. (2017). Алгоритм токенізації та стемінгу для текстів українською мовою.
4. Гудима, Н. В. (2009). Класифікація основних типів запозичень у сучасній українській літературній мові. *Наукові праці Кам'янець-Подільського національного університету імені Івана Огієнка. Філологічні науки, (20), 147-151.*
5. Дарчук, Н. (2019). Лінгвістичні засади автоматичного сентимент-аналізу українськомовного тексту. *Science and education a new dimension, 189, 10-13.*
6. Ділай, І. П., & Ділай, М. П. (2022). Роль частотності у лінгвістичних дослідженнях. *Закарпатські філологічні студії, 21 (1), 146–151.*
7. Жмаєва, Н. С. (2011). Лінгвістичне тлумачення запозичення. *Науковий вісник Південноукраїнського національного педуніверситету ім. К. Д. Ушинського. Лінгвістичні науки, 12, 48-60.*
8. Жуковська, В. В. (2020). Лінгвістичний корпус як новітній інформаційно-дослідницький інструментарій сучасного мовознавства. *Вчені записки ТНУ ім. ВІ Вернадського. Серія: Філологія. Соціальні комунікації, 113-119.*
9. Загнітко, А. (2020). *Сучасний лінгвістичний словник.* Вінниця : ТВОРИ.
10. Кононенко, І. (2012). *Українська та польська мови: контрастивне дослідження / Język ukraiński i polski: studium kontrastywne.* Warszawa: Wydawnictwo Uniwersytetu Warszawskiego. 808 с.

11. Кононенко, І. (2015-2016). Семантико-стилістичне трансформування українізмів у польській мові. *Вісник Прикарпатського університету ім. В. Стефаника. Філологія*, (44-45), 24-30.
12. Кочерган, М. П. (2008). *Загальне мовознавство*. Київ: Академія.
13. Лук'янець, Г. Г. (2012). Основні напрямки сучасних корпусних досліджень мови та перспективи їх подальшого розвитку.
14. Палатовська, О.В. (2023). Українські культурами: полімодальна рецепція польських і українських митців. *Матеріали III Міжнародної славістичної конференції, присвяченої пам'яті святих Кирила і Мефодія: «Національні наукові традиції в слов'янському світі»*. Харків – Шумен : ХНПУ; ХІФТ, 272-278.
15. Палатовська, О. В., & Сова, Є. С. (2023). Неологічні українізми в мові польських новин: січень 2021 – січень 2023 (корпусне дослідження). *Лінгвістичні дослідження*, (58), 268-279.
16. Пампура, С. Ю. (2013). Р. Раск і проблеми етимології індоєвропейських мов. *Лінгвістика*, (2), 27-32.
17. Совтис, Н. М. (2005). *Українські лексичні запозичення в польській літературній мові* (Doctoral dissertation, –Луцьк, 2005.–233 с).
18. Соссюр, Ф. (1998). *Курс загальної лінгвістики*. Київ: Основи.
19. Стишов, О. А. (2017). Нові іншомовні слова в українськомовних ЗМІ початку ХХІ ст. *Studia philologica*, (2).
20. Філон, М. І., & Хомік, О. Є. (2010). *Сучасна українська мова. Лексикологія: навчальний посібник для студентів філологічних спеціальностей вищих навчальних закладів. Ч.1*. Харків.
21. Bennett, G. R. (2010). Using corpora in the language learning classroom: Corpus linguistics for teachers. (*No Title*).
22. Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.

23. Bloomfield, L. (1933). *Language*. New York, Chicago, San Francisco, Toronto. *Holt Rhinehart and Winston*.
24. Chen, H., Wang, X., Pan, S., & Xiong, F. (2019). Identify topic relations in scientific literature using topic modeling. *IEEE Transactions on Engineering Management*, 68(5), 1232-1244.
25. Dubisz, S. (2003). Słownictwo pochodzenia ukraińskiego we współczesnej polszczyźnie. *Warszawskie Zeszyty Ukrainoznawcze*, (15-16), 409-417.
26. Fałowski, A. (2023). Ile jest słów pochodzenia ukraińskiego we współczesnej polszczyźnie ogólnej? *LingVaria*, 18(1 (35)), 63-72.
27. Greavu, A., & Blaga, L. (2014). A classification of borrowings: Observations from Romanian/English contact. *Diversité et identité culturelle en Europe*, 10(2), 95-104.
28. Gries, S. T. (2009). *Quantitative corpus linguistics with R: A practical introduction*. London, New York: Routledge.
29. Haspelmath, M. (2009). Lexical borrowing: Concepts and issues. *Loanwords in the world's languages: A comparative handbook*, 35, 54.
30. Haugen, E. (1950). The analysis of linguistic borrowing. *Language*, 26(2), 210-231.
31. Heller, M. (1988). Strategic ambiguity: Codeswitching in the management of conflict. *Codeswitching: Anthropological and sociolinguistic perspectives*, 77-96.
32. Hoffer, B. L. (2002). Language borrowing and language diffusion: An overview. *Intercultural communication studies*, 11(4), 1-37.
33. Hunston, S. (2022). *Corpora in applied linguistics*. Cambridge University Press.
34. Jespersen, O. (1922). *Language: Its nature, development and origin* (Vol. 68). Allen & Unwin.
35. Kennedy, G. (1991). Between and through: The company they keep and the functions they serve. *English corpus linguistics*, 95-110.
36. Klapper, M. (2016). M. Bańko, D. Svobodova, J. Rączaszek-Leonardi, M. Tatjewski: Nie całkiem obce. Zapożyczenia wyrazowe w języku polskim i czeskim, Wydawnictwo

- Uniwersytetu Warszawskiego, Warszawa 2016, ss. 188. *Socjolingwistyka*, (30), 289-293.
37. Krawczuk, A. (2011). *Leksykologia i kultura języka polskiego*, t. 1–2. Kijów: [nakład własny autora].
38. Labov, W. (1966). *The Social Stratification of English in New York City* (Washington DC: Center for Applied Linguistics).
39. Myers-Scotton, C. (2006). *Multiple voices: An introduction to bilingualism*. (No Title).
40. Nagórko, A., Łaziński, M., & Burkhardt, H. (2004). *Dystynktywny słownik synonimów*. TAIWPN Universitas.
41. Olejniczak, J. (2018). Using corpora to aid qualitative text analysis. *The Journal of Education, Culture, and Society*, 9(2), 154-164.
42. Peźnik, P. (2020). Budowa i zastosowania korpusu monitorującego MoncoPL. In *Forum Lingwistyczne* (No. 7, pp. 133-150). Wydawnictwo Uniwersytetu Śląskiego.
43. Poplack, S. (1988). Contrasting patterns of code-switching in two communities. *Codeswitching: Anthropological and sociolinguistic perspectives*, 48, 215-244.
44. Rak, M. (2015). Co to jest kulturem?. *LingVaria*, (20), 305-316.
45. Sapir, E. (1921). *Language, an Introduction to the Study of Speech*, by Edward Sapir. New York:: Harcourt, Brace.
46. Scotton, C. M. (1988). Codeswitching as indexical of social negotiations. *Codeswitching: Anthropological and sociolinguistic perspectives*, 151, 86.
47. Sorbet, P. (2012). Propozycja nowej kategorii zapożyczeń w językoznawstwie diachronicznym. *Linguistica Copernicana*, 7(1), 243-253.
48. Stefanowitsch, A. (2020). *Corpus linguistics: A guide to the methodology*. Language Science Press.
49. Van Coetsem, F. (1988): Loan phonology and the two transfer types in language contact. *Publications in Language Sciences*, 27.

50. Weinreich, U. (1953). *Languages in contact*. New York: Publications of the Linguistic Circle of NY.
51. Witalisz, A. (2016). *Przewodnik po anglicyzmach w języku polskim*. Towarzystwo Miłośników Języka Polskiego, Kraków.
52. Zabawa, M. (2012). *English lexical and semantic loans in informal spoken Polish*. Wydawnictwo Uniwersytetu Śląskiego.

СПИСОК ДЖЕРЕЛ ІЛЮСТРАТИВНОГО МАТЕРІАЛУ

1. Шведова, М., Яригін, С., Рисін, А., Старко, В., & Возняк, М. (2017). Генеральний регіонально анотований корпус української мови (ГРАК)/М. Шведова, Р. фон Вальденфельс, С. Яригін, М. Крук, А. Рисін, В. Старко, М. Возняк.–Київ, Осло, Єна, 2019. URL: <http://uacorpus.org/Kyiv/ua/poshuk-u-grak/grafiki-cql>
2. Korpus monitorujący polszczyzny MoncoPL. URL: <http://monco.frazeo.pl/>
3. Słownik języka polskiego PWN. URL: <https://sjp.pwn.pl/>
4. Słownik ukrainizmów. Interpretują polscy i ukraińscy ilustratorzy / Словник українізмів. Інтерпретація польських та українських ілюстраторів. URL: <https://warsztatykultury.pl/slownik-ukrainizmow-interpretuja-polscy-i-ukrainscy-ilustratorzy-wystawa/>.