

ЧАСТОТНИЙ ПРОФІЛЬ СПОНУКАЛЬНИХ КОНСТРУКЦІЙ У СУЧАСНОМУ АНГЛІЙСЬКОМОВНОМУ ТА УКРАЇНСЬКОМОВНОМУ МЕДІАДИСКУРСІ: ЛІНГВОСТАТИСТИКА ТА АЛГОРИТМ АВТОМАТИЗОВАНОГО ВІДБОРУ

Олена Рупа

студентка магістерського рівня кафедри прикладної лінгвістики,

Національний університет «Львівська політехніка»

ORCID ID 0009-0002-3266-3506

olena.rupa.mflpl.2025@lpnu.ua

Ірина Карамішева

кандидат філологічних наук,

Національний університет «Львівська політехніка»

ORCID ID 0000-0001-8683-2040

Scopus Author ID 56979297100

iryna.d.karamysheva@lpnu.ua

Ключові слова:

*спонукальні конструкції,
дієслово-предикат,
ужитковобазований
підхід, медіадискурс,
лінгвостатистика, обробка
природної мови, синтаксична
компресія.*

У статті розглянуто функціональні та лінгвостатистичні особливості спонукальних конструкцій у сучасному англійськомовному та українськомовному медіадискурсі. Актуальність дослідження зумовлена необхідністю об'єктивного вивчення синтаксичних структур у великих текстових масивах із застосуванням інструментарію автоматизованої обробки природної мови (NLP). Методологічну основу становить ужитковобазований підхід. Спираючись на нього, ми трактуємо частотність ужитку як ключовий чинник формування граматичних моделей. Метою дослідження є виявлення частотного профілю та прагматики цільових дієслів-предикатів на основі вибірки обсягом 1000 речень. Для агрегації та обробки макротекстового масиву (понад 43 тис. вебпублікацій) ми розробили та застосували інноваційний алгоритм мовою Python (із залученням бібліотеки spaCy та модуля DependencyMatcher), який уможливив точний автоматизований синтаксичний аналіз із подальшою експертною валідацією. Результати аналізу засвідчують виразну міжмовну асиметрію: відносна частотність дієслів-предикатів, що виражають переконання/прохання та примус/наказ, в англійськомовному медіадискурсі у 3–4 рази перевищує показники українськомовних відповідників. Встановлено, що в англійській мові нефінітна комплементація функціонує як високопродуктивний засіб синтаксичної компресії. Натомість в українській мові інфінітивний комплекс слугує маркером інституційної директивності у формально-діловому регістрі новин. Крім того, виявлено показову тенденцію до використання конструкцій із неживими та абстрактними ініціаторами дії в обох мовах. У висновках зазначено, що вибір спонукальних конструкцій є дієвим інструментом прагматичного позиціонування, а розроблений алгоритм створює підґрунтя для автоматизованого аналізу дискурсивних моделей із залученням великих текстових масивів.

FREQUENCY PROFILE OF CAUSATIVE CONSTRUCTIONS IN CONTEMPORARY ENGLISH AND UKRAINIAN MEDIA DISCOURSE: LINGUOSTATISTICS AND AUTOMATED SELECTION ALGORITHM

Olena Rupa

*Master's Student of Applied Linguistics Department,
Lviv Polytechnic National University*

Iryna Karamysheva

*Candidate of Philological Sciences,
Lviv Polytechnic National University*

Key words: *causative constructions, complement-taking predicate, usage-based approach, media discourse, linguostatistics, natural language processing, syntactic compression.*

The article deals with a functional and linguostatistical analysis of causative constructions in contemporary English and Ukrainian media discourse. The relevance of the study lies in the need for an objective investigation of syntactic structures across macro-textual arrays using natural language processing (NLP) tools. The methodological framework integrates the usage-based approach. Drawing on this, we interpret usage frequency as a key determinant in the formation of grammatical patterns. The purpose of the study is to identify the frequency profile and pragmatics of target predicate verbs based on a 1,000-sentence sample. To aggregate and process a macro-textual array (over 43,000 web publications), we developed and applied an innovative Python-based algorithm (utilizing the spaCy library and DependencyMatcher), enabling precise automated syntactic parsing followed by expert validation. The analysis reveals a distinct cross-linguistic asymmetry: the relative frequency of predicate verbs expressing persuasion/requests and coercion/commands in English media discourse is three to four times higher than that of their Ukrainian counterparts. The research establishes that non-finite complementation serves as a highly productive mechanism of syntactic compression in English. Conversely, in Ukrainian, the infinitive complex functions as a marker of institutional directiveness in the formal news register. Furthermore, an indicative trend toward employing constructions with inanimate and abstract causers is identified in both languages. The conclusions indicate that the choice of causative constructions serves as an effective tool for pragmatic positioning, while the developed algorithm lays the groundwork for automated analysis of discourse patterns in large textual arrays.

Вступ. Дослідження синтаксичних структур у великих текстових масивах вимагає інтеграції традиційних методів лінгвістичного аналізу з інструментарієм автоматизованої обробки природної мови (NLP). У межах ужитковобазованого підходу постулюється (usage-based approach)¹, що

на формування та засвоєння граматичних конструкцій значною мірою впливає частота їхнього вживання, відтак “корпусні дані слугують незамінним інструментом для виявлення та аналізу синтаксичних моделей” (Horsch, 2023b: 18; Karamysheva, 2025a: 136). Автоматизований збір вебданих (web scraping) та алгоритми обробки природної мови (NLP) забезпечують оперування репрезентативними вибірками, усуваючи суб’єктивізм ручного відбору фактологічної бази.

Одним із ключових об’єктів корпусноорієнтованого аналізу є категорія предикації, з типами якої пов’язані фундаментальні характеристики речення – його здатність виражати відносно закінчену думку та слугувати основним засобом репрезентації повідомлення (Кульбабська, 2011: 12).

¹ Джоан Байбі, одна із засновниць ужитковобазованої лінгвістики (Usage-Based Linguistics), стверджує, що “Визнання ролі частотності вживання у процесах засвоєння, створення та зміни мовної структури часто вважають визначальною характеристикою теоретичної моделі, орієнтованої на вживання (usage-based approach)” (J. Bybee, 2023, с. 18). На важливості частоти вживання (frequency effects) також наголошують провідні представники конструкційної граматики (Usage-Based Construction Grammar) Адель Е. Голдберг (A. E. Goldberg 2006, 2019), Томас Гоффманн (Th. Hoffmann et al, 2019), Анатоль Стефанович та Штефан Т. Гріс (S. T. Gries, & A. Stefanowitsch, 2004; S. T. Gries, 2023).

Спираючись на концепцію В. Крофта, конструкцію трактуємо як усталене поєднання форми та функції, що слугує для вираження певної комбінації семантичного змісту та інформаційного структурування (information packaging) (Croft, 2022: 17; Karamysheva, 2025a: 134). Водночас оформлення семантичного змісту безпосередньо реалізується через предикативні відношення, оскільки саме предикація є механізмом його синтаксичного структурування (Karamysheva, 2025b: 32).

У цьому річизці особливий дослідницький інтерес становлять синтаксичні конструкції, що виражають маніпулятивні події. У типології М. Нунана розрізняють предикати, які дозволяють подіям функціонувати як аргументи, – complement-taking predicates або CTPs (Noonan, 2007: 53; Karamysheva, 2025b: 32). Предикати, що репрезентують маніпулятивні події, поділяються на каузативні та пермісивні підтипи (Karamysheva, 2024: 3). Маніпулятивні предикати експлікують відношення між агенсом, що функціонує як причина, афективом (affectee) та результативною ситуацією, неодмінним учасником якої є сам афектив (Karamysheva, 2025b: 33). Відповідно, каузативні конструкції за своєю синтаксичною природою належать до структур вторинної предикації (Karamysheva, 2025a: 135). У межах нашого дослідження терміни “спонукальні” та “каузативні” конструкції вживаються як синоніми для позначення єдиного синтаксичного феномену, однак надаємо перевагу українському відповіднику “спонукальні конструкції”.

Попри ґрунтовну теоретичну базу, лінгвостатистичні параметри спонукальних конструкцій у певних видах дискурсу залишаються маловивченими. Зокрема, детального з'ясування потребує частотний профіль цих одиниць у сучасному медіадискурсі (editorial news), де вони слугують дієвим синтаксичним інструментом впливу на реципієнта. Для забезпечення релевантності контрастивного аналізу англійськомовних та українськомовних текстів виникає потреба розрізняти абсолютну (raw) частотність, нормалізовану частотність, частотність токенів (token frequency), що безпосередньо корелює з укоріненням (entrenchment)² у ментальній структурі мовця та відображається через кількість мовних структур у корпусі, та частотність типів (type frequency), яка визначає продуктивність граматичної моделі (Karamysheva, 2024: 3).

² Термін “entrenchment” (дотримуємось українського відповідника “укорінення”) був уведений у когнітивну лінгвістику Рональдом Ленекером (R. Langacker, 1987) як теоретичний конструкт для пояснення того, як мовні структури створюються та формуються в процесі вживання, припускаючи, що збільшення частотності вживання поглиблює закріплення тої чи іншої мовної структури у пам'яті/свідомості мовця.

З огляду на це, **актуальність** статті полягає в необхідності комплексного лінгвостатистичного аналізу спонукальних конструкцій із застосуванням інноваційної методології автоматизованого збору фактологічної бази.

Мета дослідження – виявити функціональні та лінгвостатистичні особливості спонукальних конструкцій у сучасному англійськомовному та українськомовному медіадискурсі на основі вибірки обсягом 1000 речень (500 англійськомовних та 500 українськомовних текстів новин), згенерованої шляхом автоматизованого синтаксичного аналізу із застосуванням розробленого програмного алгоритму мовою Python. Для досягнення мети передбачено розв'язання таких завдань: описати алгоритм NLP-фільтрування з використанням бібліотеки spaCy та алгоритму DependencyMatcher; проаналізувати абсолютну та відносну частотність цільових дієслів-предикатів (із детальним поясненням високої частотності дієслова *have* в загальному контексті); визначити специфіку функціонування та синтаксичні критерії відбору виокремлених конструкцій із нескінченною комплементацією в текстах новин.

Матеріал і методи дослідження. Традиційна методологія ручного відбору текстів виявляється недостатньо ефективною для опрацювання макротекстових масивів. Опрацювання масштабних масивів даних вимагає мінімізації суб'єктивного чинника. Це є фундаментальною умовою для отримання об'єктивних лінгвостатистичних показників. Дослідники конструкційної граматики наголошують на критичній залежності результатів від обсягу аналізованого матеріалу: “[...] кількість прикладів, необхідних для аналізу коваріантних колексем (covarying collexemes), є надзвичайно великою: навіть корпус BNC зі своїми 100 мільйонами слів не забезпечив достатньої кількості збігів для їхнього аналізу; обсяг корпусу довелося збільшити вчетверо, перш ніж з'явилися значущі результати. Отже, деякі конструкції можна аналізувати лише на базі газетних корпусів або ж доведеться чекати на появу збалансованих гігакорпусів” (Gries & Stefanowitsch, 2004: 235). З огляду на це, концептуальною основою нашого дослідження стала розробка оптимізованого алгоритму веб-скрейпінгу (scraper_EN_thesis, scraper_UK_thesis) мовою Python. Використання відкритого алгоритму автоматизованого пошуку, на відміну від готових корпусних менеджерів, продиктоване необхідністю створення гнучкого інструмента. Він потрібен для багаторівневого автоматизованого синтаксичного аналізу складних синтаксичних дерев (Dependency Trees) у неструктурованих інформаційних масивах.

Формування джерельної бази розпочалося з укладання реєстру предикатів, що виражають каузативну семантику. Цей перелік ґрунтується на лексико-семантичних класифікаціях, верифікованих у попередніх дослідженнях. Загальний список містить 27 англійських дієслів. Вони диференціюються на дві підгрупи: конструкції примусу/наказу до дії (*coercive constructions*, 11 дієслів-предикатів), та конструкції переконання/прохання (*persuasive constructions*, 16 дієслів-предикатів). Відповідний масив українських маніпулятивних предикатів налічує 23 базові каузативні дієслова, що вводять спонукальні конструкції, та 6 додаткових лексико-граматичних варіацій. Відібрані лексеми репрезентують широкий спектр інтенсивності впливу: від потужної маніпулятивної сили (*примусити*, *змусити*, *спонукати*, *наказувати*) до м'якого переконання (*попросити*, *порадити*, *порекомендувати*) (Karamysheva, 2026: 103). Водночас периферійна лексика залучалася з урахуванням частотних обмежень. Дієслова з виразним діалектним або архаїчним забарвленням (наприклад, *веліти* чи *раяти*, див. Karamysheva, 2026: 103) виключалися з основного фокуса, аби не спотворювати лінгвостатистичний профіль сучасного медіадискурсу.

Масиви новинних текстів було отримано шляхом автоматизованого збору даних за період 2021–2026 років. На етапі агрегації генеральна сукупність проаналізованих вебджерел склала 43 274 одиниці (20 892 англійськомовні та 22 382 українськомовні публікації). Для уникнення джерельного зміщення застосовано алгоритм ітеративного рівномірного відбору вебсторінок. Водночас автоматизований алгоритм було налаштовано на пріоритетне опрацювання розгорнутих переконливих текстів (*editorial*), де спонукальні конструкції функціонують найефективніше. Алгоритм вилучав винятково автентичні контексти, що мінімізувало ризик джерельного зміщення. Це дозволило інтерпретувати відсутність певних рідкісних дієслів як об'єктивний лінгвістичний факт нульової частотності.

Ядром методології стала автоматизована ідентифікація спонукальних конструкцій, реалізована через формалізацію лінгвістичної моделі за допомогою бібліотеки обробки природної мови *sraCy* та алгоритму *DependencyMatcher*. Пошукова архітектура ґрунтувалася на розпізнаванні чотирьох обов'язкових компонентів: ініціатора дії (S1), каузативного предиката (P1), виконавця другої дії / афектива (S2) та нефінітного компонента (P2). Оптимізований Python-скрипт здійснював автоматизований аналіз дерева синтаксичних залежностей, де каузативне дієслово ідентифікувалося за тегами частин мови (POS-тегами) як стрижневий вузол (ROOT). За допомогою

алгоритму *DependencyMatcher* інші компоненти виокремлювалися за синтаксичними зв'язками (DEP-тегами): ініціатор дії (S1) як номінативний суб'єкт (*nsubj*), афектив (S2) як додаток (*dobj* або *obj*), а нефінітний комплемент (P2) як відкрите дієслівне доповнення (*xcomp*). Його безпосередня наявність у межах однієї поліпредикативної клаузи формувала базовий синтаксичний шаблон (шаблон "Pattern A"). Для обробки синтаксично ускладнених явищ (наприклад, в умовах еліпса номінативного суб'єкта) було розроблено розширений алгоритм (шаблон "Pattern B"), який відновлював логічного ініціатора з макроконтексту шляхом аналізу кореферентності. Конструкції з пасивним станом головного дієслова автоматично відхилялися.

Розроблений автоматизований алгоритм передбачав вилучення щонайменше 1500 речень для кожної мови, з яких переважна більшість мала відповідати жорсткому шаблону (шаблон "Pattern A"). Оскільки сучасні інструменти автоматизованого синтаксичного аналізу не здатні безпомилково розпізнавати глибинні метафоричні зміщення або контекстуально зумовлену десемантизацію предикатів, до методології було включено етап інтерактивної експертної валідації. Здійснювався ручний семантичний аналіз попередньо зібраного масиву з вилученням нерелевантних результатів. Цей підхід дозволив сформувати еталонну збалансовану вибірку (по 500 англійськомовних та українськомовних речень).

Результати та обговорення. Реалізація розробленого алгоритму видобування даних дозволила сформувати репрезентативний масив для лінгвостатистичного та прагматичного аналізу. Загальний обсяг електронних адрес (URL-посилань) склав 43 274 одиниці, з яких 20 892 припадало на англійськомовні, а 22 382 – на українськомовні джерела. Алгоритм цілеспрямовано відібрав визначену квоту – по 1500 речень для кожної мови, серед яких щонайменше 1000 одиниць відповідали базовому синтаксичному шаблону (Pattern A).

У результаті автоматизованого синтаксичного аналізу та подальшої експертної валідації було сформовано лінгвостатистичний профіль цільових предикатів у медіадискурсі. Загальна кількість опрацьованих токенів для англійської мови склала 2 248 493 лексеми. Найвищу відносну частотність продемонстрували предикати, що виражають переконання/прохання, *ask* (0,7249 на 10 тис. слів) та *urge* (0,6493). Висока частотність дієслів *have* (0,6760) та *make* (0,2224) як дієслів-предикатів, що виражають примус/наказ, підтверджує їхню частотність у новинних текстах. Точне виокремлення конструкцій із цими дієсловами стало можливим виключно завдяки автоматизованій лематизації та POS-тегуванню: без морфологічної розмітки розрізнення *have* як

каузативного предиката та *have* як допоміжного дієслова чи дієслова володіння було б неможливим. Кількісні параметри англійської мови наведено в Таблиці 1.

Для українськомовного масиву загальна кількість проаналізованих текстових одиниць виявилася вдвічі більшою та склала 4 555 761 лексему. Попри розширену емпіричну базу, абсолютна та відносна частотність спонукальних конструкцій із нефінітним комплементом в українській мові є суттєво нижчою порівняно з англійською. Лідером серед українських маніпулятивних предикатів виявилось дієслово *просити* (відносна частотність 0,2173), за яким ідуть *змусити* (0,1602) та *запропонувати* (0,1449). Зведені дані представлено в Таблиці 2.

Зіставний аналіз кількісних параметрів виявляє виразну міжмовну асиметрію. Відносна частотність провідних англійських предикатів у 3–4 рази перевищує показники найуживаніших українських відповідників. Ця диспропорція зумовлена не лише структурно-типологічними властивостями досліджуваних мов, а й глибинними прагматичними стратегіями конструювання медіадискурсу. В англійськомовному новинному

тексті, що функціонує в умовах жорстких вимог до економії простору та динаміки подачі інформації, нефінітна комплементация виступає як високопродуктивний механізм синтаксичної компресії. Ця модель дозволяє безперешкодно інтегрувати дві пропозиції (акт каузації та результативну подію) в межах однієї клаузи, утворюючи нерозривний причинно-наслідковий ланцюг.

Перехід від констатації кількісних метрик до якісного лінгвопрагматичного розбору став можливим завдяки здатності алгоритму *DependencyMatcher* автоматично відсіювати синтаксичний шум і вилучати контексти з десемантизованими дієсловами. Отриманий у такий спосіб еталонний масив, який пройшов фінальну стадію ручної експертної валідації, дозволив здійснити точну верифікацію статистичних показників і дослідити глибинні механізми синтаксичного оформлення (*syntactic packaging*).

Англійськомовні приклади ілюструють високу частотність прототипних персуазивних моделей, де ініціатором виступає інституційний агент: (1) *“Initially these consultations were required to be bulk billed, but the AMA has now persuaded the government to remove this requirement”* (The

Таблиця 1

Частотний профіль англійських спонукальних предикатів

№	Мова	Лексема	Абсолютна частотність	Загальна кількість слів	Відносна частотність на 10 тис. слів
1	EN	ask	163	2248493	0,7249
2	EN	have	152	2248493	0,6760
3	EN	urge	146	2248493	0,6493
4	EN	encourage	96	2248493	0,4270
5	EN	force	60	2248493	0,2668
6	EN	lead	50	2248493	0,2224
7	EN	make	50	2248493	0,2224
8	EN	prompt	50	2248493	0,2224
9	EN	advise	24	2248493	0,1067
10	EN	persuade	12	2248493	0,0534

Таблиця 2

Частотний профіль українських спонукальних предикатів

№	Мова	Лексема	Абсолютна частотність	Загальна кількість слів	Відносна частотність на 10 тис. слів
1	UK	просити	99	4555761	0,2173
2	UK	змусити	73	4555761	0,1602
3	UK	запропонувати	66	4555761	0,1449
4	UK	пропонувати	58	4555761	0,1273
5	UK	попросити	54	4555761	0,1185
6	UK	рекомендувати	48	4555761	0,1054
7	UK	спонукати	21	4555761	0,0461
8	UK	радити	20	4555761	0,0439
9	UK	заохочувати	15	4555761	0,0329
10	UK	переконувати	8	4555761	0,0176

Conversation, 2021). (2) *“Jersey Police are urging football fans to enjoy watching the England-Norway World Cup match ‘responsibly’”* (BBC News, 2026b). У наведених прикладах каузативні предикати (*persuaded, are urging*) експлікують чіткий вектор впливу на афективів (*the government, football fans*). Використання інфінітива безпосередньо маркує очікувану подію як логічний і безальтернативний наслідок маніпулятивного акта. Висока синтаксична щільність характерна і для конструкцій із заперечним інфінітивом, які ефективно кодують каузацію заборони (чи наполегливу рекомендацію утриматися від дії): (3) *“It [Cambridge Water] urged people not to call 999 if they see someone using a hosepipe”* (BBC News, 2026a).

Критично важливим показником валідності розробленого програмного комплексу стала його здатність ідентифікувати ускладнені синтаксичні моделі (шаблон “Pattern B”), у яких ініціатором маніпулятивного впливу виступає абстрактне поняття, подія чи неістота (*inanimate agent*), а логічні зв’язки реалізуються через анафоричні елементи або відносні займенники: (4) *“Aggressive coverage of the Trump administration often led the former president to attack it...”* (AP News, 2021a). (5) *“It’s not uncommon to hear shootings in the courthouse area that force government employees to go home early...”* (AP News, 2021b). (6) *“In 1936, Lambert published his findings in a book... This prompted Sir Cecil Levita to seek Lambert’s dismissal...”* (The Conversation, 2022). У прикладі (4) абстрактний концепт *coverage* (висвітлення подій) метафорично концептуалізується як активний агент, що спричиняє реакцію об’єкта. У реченні (5) алгоритм коректно розпізнав відносний займенник *that* як номінативний суб’єкт, автоматично відновивши його кореферентність із лексемою *shootings* (стрілянина), яка й виступає маніпулятивною силою. Нарешті, у прикладі (6) каузатором є вказівний займенник *This*, що конденсує в собі усю попередню макропропозицію. Здатність алгоритму фіксувати такі складні топологічні відношення доводить, що англійська мова системно використовує вторинну предикацію для кодування комплексних подій, перетворюючи їх на дієві агенти впливу.

Водночас якісний аналіз корпусу виявив абсолютну необхідність інтерактивної експертної валідації (*human-in-the-loop*) після автоматизованого синтаксичного аналізу. Алгоритм [*S1 + V + S2 + Vinf*] очікувано фіксував хибнопозитивні результати (*false positives*), у яких дієслово *make* функціонувало не як каузативний предикат, а як компонент десемантизованої конструкції (*light verb construction*): (7) *“... they’re making rational choices based on their*

economic circumstances...” (The Guardian, 2026). (8) *“...England were well on course to make it 11 from 12 in Wellington.”* (BBC News, 2026c). Вилучення подібних контекстів на фінальному етапі очищення даних убезпечило результати від статистичних викривлень та підтвердило тезу про те, що суто формального автоматизованого аналізу без семантико-прагматичного фільтрування недостатньо для точного аналізу граматики конструкцій.

Переходячи до аналізу українськомовного масиву, варто зазначити, що суттєво нижча частотність інфінітивної комплементції свідчить про гостру конкуренцію цієї моделі зі складнопідрядними з’ясувальними реченнями. Перевага, яку українські медіа віддають підрядним конструкціям (*“змусив когось, щоб він зробив щось”*), створює ефект прагматичного дистанціювання, дозволяючи журналістам розмежувати зони відповідальності та пом’якшити категоричність тверджень. Відповідно, використання українськими медіа жорсткої моделі нефінітної комплементції є завжди стилістично й прагматично маркованим кроком.

Так, модель інфінітивного компонента безальтернативно обирається для опису каузації, що виражає примус/наказ, коли необхідно підкреслити неминучість дії або акт насильства. Відсутність підрядного зв’язку тут навмисно усуває будь-яку синтаксичну дистанцію: (9) *“На півдні Франції масштабна лісова пожежа змусила владу евакуювати тисячі мешканців...”* (Ukrainska Pravda, 2026b). (10) *“У відповідь на це Перевозкін наказував тюремникам бити її...”* (Suspilne, 2026a).

Проте аналіз розширеної вибірки дозволив виявити ще одну показову тенденцію – інституційну формалізацію предикатів м’якого переконання. У формально-діловому, дипломатичному та юридичному реєстрах новинних текстів дієслова типу *просити* або *запропонувати* втрачають своє первинне значення необов’язкової поради і набувають статусу безапеляційного інституційного запиту чи прихованої директиви: (11) *“Компанія Apple просить суд призупинити виконання рішення про зміни правил...”* (Hromadske, 2021b). (12) *“Через це вони [прокурори] просять суд заарештувати депутата з можливістю вийти під заставу...”* (Suspilne, 2026d). (13) *“Водночас Міністерство оборони та посольство Південної Кореї попросили владу Ірану підтвердити безпеку екіпажу та якомога швидше звільнити судно.”* (Hromadske, 2021a). У цих прикладах вторинна предикація (*призупинити, заарештувати, підтвердити, звільнити*) нерозривно прив’язує очікувану дію до інституційного афектива (*суд, влада Ірану*). Використання у цьому випадку підрядного речення знизило б статус

офіційності та перетворило б правову вимогу чи дипломатичний демарш на побутове прохання.

Ця ж трансформація характерна для предиката *запропонувати*, який у політичному та економічному дискурсі кодує спонування до стратегічних кроків: (14) “*Україна запропонувала Хорватії укласти спеціальну угоду і, продавши яхту Медведчука, розділити кошти між країнами.*” (Suspilne, 2026c). (15) “*...очільник ОП запропонував Кличку самому подати у відставку.*” (Tyzhden, 2021). Якщо приклад (14) репрезентує офіційну міжнародну пропозицію, то контекст (15) демонструє використання спонукальної конструкції для передачі політичного ультиматуму, формально замаскованого під “пропозицію”. Така прагматична багатовимірність підтверджує, що інфінітивний комплекс в українській мові є точним інструментом медіаманіпуляції. Високий рівень синтаксичної деталізації простежується і в конструкціях із запереченням, де предикат ініціює стримування від хибної оцінки: (16) “*Він радив Штілерману не недооцінювати детективів бюро.*” (Suspilne, 2026b).

Важливо підкреслити, що розроблений алгоритм (шаблон “Pattern B”) успішно ідентифікував ускладнені каузативні моделі з неживими та абстрактними ініціаторами дії і в українськомовному корпусі. Це спростовує уявлення про те, що такі структури є винятковою прерогативою аналітичних мов: (17) “*...феноменально позитивне відчуття спільноти й усвідомлення власної співучасті у творенні мистецтва спонукає публіку раз у раз голосувати за театр*” (Tyzhden, 2024). (18) “*Але зараз маємо не просто мовчання: ЄС фінансово стимулює Київ робити антиреформи...*” (Ukrainska Pravda, 2026a). У прикладі (17) абстрактні ментальні концепти (*відчуття, усвідомлення*) алгоритм безпомилково розпізнав як суб’єкти спонування. У реченні (18) інституційний актор (ЄС) виступає каузатором дії через предикат “*стимулювати*”, що формує виразну оцінно-маніпулятивну рамку. Українська мова, як й англійська, ефективно застосовує вторинну предикацію для кодування когнітивних і макрополітичних процесів, наділяючи абстрактні феномени дієвою силою.

Підсумовуючи, лінгвопрагматичний аналіз корпусу повністю підтверджує валідність і евристичний потенціал розробленого NLP-алгоритму. Інструментарій *DependencyMatcher* дозволяє не лише акумулювати масштабні частотні показники, що відповідають засадам ужитковобазованого підходу, а й виявляти найтонші нюанси синтаксичного структурування інформації. Зібрані дані доводять, що конфігурація граматичних конструкцій безпосередньо зумовлюється комунікативними стратегіями: від максимальної

синтаксичної компресії в англійській мові до інституційної директивності та тонкого прагматичного маневрування в українській. Розподіл каузативних моделей та їхніх комплементів не є випадковим, а слугує дієвим лінгвопрагматичним інструментом, за допомогою якого медіа конструюють причинно-наслідкові зв’язки та маніпулюють сприйняттям реципієнта.

Висновки. Розроблений спеціалізований алгоритм на базі мови Python та NLP-бібліотеки spaCy дав змогу розв’язати методологічну проблему автоматизованої ідентифікації низькочастотних спонукальних конструкцій. Одержані емпіричні дані безпосередньо підтверджують тезу про вичерпаність евристичного потенціалу традиційних лінгвістичних баз даних обмеженого обсягу, а також доводять неефективність нерозмічених масивів для глибинного синтаксичного пошуку. Завдяки розробці та застосуванню алгоритмів автоматизованого сегментування й ітеративного обслуговування джерел удалося нейтралізувати діахронічне та джерельне зміщення, що дозволило сформувати високорепрезентативний інформаційний корпус, який охоплює понад 43 тисячі медіапублікацій.

Застосування інструментів автоматизованого синтаксичного аналізу в поєднанні з етапом експертної валідації довело свою методологічну перевагу над стандартним лексичним пошуком, уможлививши формування збалансованої вибірки з високою точністю.

Лінгвостатистичний аналіз цієї вибірки зафіксував виразну міжмовну асиметрію синтаксичного структурування. Відносна частотність провідних англійських маніпулятивних предикатів суттєво перевищує показники їхніх найуживаніших українських відповідників, що ілюструє високу продуктивність нефінітної комплементції як засобу економного структурування семантичного змісту в англійськомовному медіадискурсі. Водночас показники автоматизованого синтаксичного аналізу підтверджують, що українські спонукальні конструкції є цілком функціональними синтаксичними одиницями, які активно використовуються в новинних текстах для передачі ієрархічних і причинно-наслідкових відношень. Зафіксована частотна дистрибуція предикатів природно узгоджується з постулатами ужитковобазованого підходу, підтверджуючи, що конфігурація граматики безпосередньо зумовлюється регулярністю вжитку в реальному комунікативному середовищі.

Крім того, виокремлені предикати формують суворо детерміновані семантичні кластери в межах каузативної моделі (фізичний примус, вербальне переконання), що віддзеркалюють лексико-синтаксичну солідарність і регламентують

вибір дієслова залежно від вектора маніпулятивного впливу як в англійській, так і в українській мовах.

Отже, розроблена NLP-методологія не лише підтвердила свою ефективність в опрацюванні макротекстів, а й дозволила довести, що вибір спонукальних граматичних конструкцій є ключовим інструментом медіаманіпуляції та прагматичного позиціонування в сучасному інформаційному просторі. Перспективи подальших наукових розвідок полягають в екстраполяції розробленого NLP-алгоритму на збалансовані гігакорпуси, що є безумовною методологічною вимогою для детального аналізу коваріантних колексем.

ЛІТЕРАТУРА

1. Кульбабська, О. В. (2011). *Вторинна предикація у простому реченні: монографія*. Чернівецький нац. ун-т.
2. Кульбабська, О. В. (2025). Кореферентність засобів вторинної предикації в інтелектуально-комунікативній діяльності мовців. *Вчені записки*, 39, 39–44. <https://doi.org/10.32782/2710-4656/2025.3.1/07>
3. Bybee, J. (2023). What is usage-based linguistics? In M. Díaz-Campos & S. Balasch (Eds.), *The handbook of usage-based linguistics* (pp. 9–29). Wiley-Blackwell. <https://doi.org/10.1002/9781119839859.ch1>
4. Croft, W. (2022). *Morphosyntax: Constructions of the world's languages*. Cambridge University Press. <https://doi.org/10.1017/9781316145289>
5. Davies, M. (2012). Expanding horizons in historical linguistics with the 400-million word Corpus of Historical American English. *Corpora*, 7(2), 121–157. <https://doi.org/10.3366/cor.2012.0024>
6. Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press.
7. Goldberg, A. E. (2019). *Explain me this: Creativity, competition and the partial productivity of constructions*. Princeton University Press.
8. Gries, S. T. (2023). New technologies and advances in statistical analysis in recent decades. In M. Díaz-Campos & S. Balasch (Eds.), *The handbook of usage-based linguistics* (pp. 561–579). Wiley-Blackwell. <https://doi.org/10.1002/9781119839859.ch31>
9. Gries, S. T., & Stefanowitsch, A. (2004). Co-varying collexemes in the into-causative. In M. Achard & S. Kemmer (Eds.), *Language, culture, and mind* (pp. 225–236). CSLI.
10. Hoffmann, T., Horsch, J., & Brunner, T. (2019). The more data, the better: A usage-based account of the English comparative correlative construction. *Cognitive Linguistics*, 30(1), 1–36. <https://doi.org/10.1515/cog-2018-0036>
11. Horsch, J. (2023a). From corpus data to constructional networks: Analyzing language with the Usage-based Construction Grammar framework. *Journal of Linguistics*, 74(3), 701–740. <https://doi.org/10.2478/jazcas-2024-0019>
12. Horsch, J. (2023b). *The comparative correlative construction in World Englishes: A usage-based construction grammar approach* [Doctoral dissertation, Catholic University of Eichstätt-Ingolstadt]. <https://doi.org/10.17904/ku.opus-833>
13. Karamysheva, I. (2024). Challenges of corpus-based contrastive studies: A case study of the English “make”-construction and its Ukrainian counterpart. *2024 IEEE 19th International Conference on Computer Science and Information Technologies (CSIT)*, 1–4. <https://doi.org/10.1109/CSIT65290.2024.10982618>
14. Karamysheva, I. (2025a). The choice of corpora for contrastive analysis: A case study of the English causative construction with “encourage” and its Ukrainian counterpart. *Inozemna Filologia*, 138, 133–147. <https://doi.org/10.30970/fpl.2025.138.5074>
15. Karamysheva, I. (2025b). A corpus-based approach to the study of permissive constructions in English and Ukrainian. *Journal of Linguistics/ Jazykovedný časopis*, 76(1), 31–40. <https://doi.org/10.2478/jazcas-2025-0003>
16. Karamysheva, I. (2026). Corpus-based, usage-based and construction-based approaches towards teaching of Ukrainian nonfinite complement constructions in contrast with their English counterparts. *Teorija i Praktyka Vykladannja Ukrajins'koji Movy jak Inozemnoji*, 21, 96–110. <https://doi.org/10.30970/uf.2026.21.5147>
17. Langacker, R. W. (1987). *Foundations of cognitive grammar, volume I, theoretical prerequisites*. Stanford University Press.
18. Noonan, M. (2007). Complement-taking predicates. In T. Shopen (Ed.), *Language typology and syntactic description* (pp. 52–150). Cambridge University Press. <https://doi.org/10.1017/CBO9780511619434>

ДЖЕРЕЛА ІЛЮСТРАТИВНОГО МАТЕРІАЛУ

1. AP News. (2021a, 26 січня). Washington Post's Marty Baron says he's retiring next month. <https://apnews.com/article/liev-schreiber-newspapers-ed9a919e05e4d80b00d6e72aa168cfc7>
2. AP News. (2021b, 28 січня). Strikes, violence overwhelm Haiti's crumbling judiciary. <https://apnews.com/article/port-au-prince-prisons-judiciary-latin-america-united-nations-86043172b27e2fa42cd105a097ae3380>

3. BBC News. (2026a, 10 липня). What does a hosepipe ban mean for the UK's driest region. <https://www.bbc.co.uk/news/articles/c1wyxgl72pro>
4. BBC News. (2026b, 11 липня). Fans asked to enjoy World Cup match 'responsibly'. <https://www.bbc.com/news/articles/c4gy1l6mp7do>
5. BBC News. (2026c, 12 липня). The highs and lows of Bazball. <https://www.bbc.com/sport/cricket/articles/c4grpj0lz7ro>
6. Hromadske. (2021a, 4 січня). Іран захопив танкер під південнокорейським прапором у своїх водах. Сеул вимагає звільнити судно. <https://hromadske.ua/posts/iran-zahopiv-tanker-pid-korejskim-praporom-v-ormuzkij-protoci-seul-vimagaye-zvilniti-sudno>
7. Hromadske. (2021b, 9 жовтня). Apple хоче призупинити рішення у справі проти Epic Games, поки його оскаржують. Це може відтягнути зміни в Apple Store. <https://hromadske.ua/posts/apple-hoche-prizupiniti-rishennya-u-spravi-proti-epic-games-poki-jogo-oskarzhuyut-ce-mozhe-vidtyagnuti-zmini-v-apple-store>
8. Suspilne. (2026a, 6 червня). Слідство.Інфо: у російському СІЗО, де загинула Вікторія Рощина, почали краще годувати полонених. <https://suspilne.media/1324413-slidstvoinfo-u-rosijskomu-sizo-de-zaginula-viktoria-rošina-pocali-krase-goduvati-polonenih/>
9. Suspilne. (2026b, 19 червня). Як Fire Point стала незамінним постачальником Сил оборони – і фігурантом розслідувань. Історія компанії. <https://suspilne.media/1328696-ak-fire-point-stala-nezaminnim-postacalnikom-sil-oboroni-i-figurantom-rozsliduvan-istoria-kompanii/>
10. Suspilne. (2026c, 29 червня). Кошти від продажу яхти Медведчука Україна планує ділити з Хорватією. <https://suspilne.media/1342752-kosti-vid-prodazu-ahti-medvedcuka-ukraina-planue-diliti-z-horvatieu/>
11. Suspilne. (2026d, 4 липня). “Дайте мені константу, і я буду інструментом”. Як і за що суд призначив нардепу Тищенко 10 мільйонів гривень застави. <https://suspilne.media/1347198-dajte-meni-konstantu-i-a-budu-instrumentom-ak-i-za-so-sud-priznaciv-nardeputisenku-zastavu-v-10-miljoniv/>
12. The Conversation. (2021, 13 серпня). Media and politicians often defer to the AMA on COVID policies. But what role should the doctors' group have in the pandemic? <https://theconversation.com/media-and-politicians-often-defer-to-the-ama-on-covid-policies-but-what-role-should-the-doctors-group-have-in-the-pandemic-165074>
13. The Conversation. (2022, 6 квітня). 100 years of the BBC: the rebels who reshaped broadcasting – and paid the price. <https://theconversation.com/100-years-of-the-bbc-the-rebels-who-reshaped-broadcasting-and-paid-the-price-173784>
14. The Guardian. (2026, 12 липня). Dining across the divide: 'She's fine with billionaires – I would call them hoarders'. <https://www.theguardian.com/lifeandstyle/2026/jul/12/dining-across-the-divide-emma-aly-billionaires-defence-welfare-tax>
15. Tyzhden. (2021, 25 листопада). Битва за Київ 2.0. <https://tyzhden.ua/bytva-za-kyiv-2-0/>
16. Tyzhden. (2024, 2 січня). 2023 рік: культурні підсумки. <https://tyzhden.ua/2023-rik-kulturni-pidsumky/>
17. Ukrainska Pravda. (2026a, 6 липня). Два варіанти покарання України. Як ЄС опинився перед складним вибором через судову «реформу». <https://www.eurointegration.com.ua/articles/2026/07/6/7241001/>
18. Ukrainska Pravda. (2026b, 6 липня). У Франції через масштабні лісові пожежі евакуювали понад 10 тисяч людей. <https://www.eurointegration.com.ua/news/2026/07/6/7241058/>

Дата першого надходження статті до видання: 21.07.2026

Дата прийняття статті до друку після рецензування: 12.08.2026

Дата публікації (оприлюднення) статті: 01.10.2026



Стаття поширюється на умовах ліцензії відкритого доступу (CC BY 4.0)