

УДК 811.111'373.4

DOI <https://doi.org/10.32782/2410-0927-2023-18-2>

Ірина ДІЛАЙ

кандидат філологічних наук, доцент, доцент кафедри англійської філології Львівського національного університету імені Івана Франка, вул. Університетська, 1, м. Львів, Україна, 79000

ORCID: 0000-0001-9626-290X

Бібліографічний опис статті: Ділай, І. (2023). Ключовість у художньому тексті (на матеріалі роману М. Етвуд «Оповідь служниці»). *Актуальні питання іноземної філології*, 18, 12–18, doi: <https://doi.org/10.32782/2410-0927-2023-18-2>

КЛЮЧОВІСТЬ У ХУДОЖНЬОМУ ТЕКСТІ (НА МАТЕРІАЛІ РОМАНУ М. ЕТВУД «ОПОВІДЬ СЛУЖНИЦІ»)

Стаття присвячена автоматизованому виявленню з метою подальшої інтерпретації ключових лексичних одиниць у романі «Оповідь служниці» відомої канадської письменниці-постмодерністки Маргарет Етвуд. Мета дослідження – апробувати оптимальну методологію корпусокерованого виокремлення ключових слів та словосполук з художнього тексту.

Методологія. Аналіз ключовості є важливою складовою контент-аналізу художніх творів, оскільки це дає уявлення про основні теми, мотиви, ідеї, які намагається донести автор, а також про особливості ідіостилу автора. Матеріалом аналізу є текст роману М. Етвуд, який, як і інші твори письменниці, раніше не досліджувався на предмет ключовості.

Проведене дослідження складалося із кількох етапів відповідно до поставлених завдань: 1) створено анотований електронний корпус тексту роману; 2) підібрану порівнюваний референтний корпус для виявлення ключових лексичних одиниць; 3) автоматизовано встановлено ключові лексичні одиниці у тексті роману та здійснено інтерпретацію даних. Для цього застосовано корпусний інструментарій Sketch Engine, в основі якого закладена проста статистична міра для встановлення ключовості.

В ході здійсненого аналізу визначено такі ключові лексичні одиниці: авторські неологізми, створені М. Етвуд для опису суспільного ладу в дистопії під назвою Джилід (Gilead, Aunt, Commander, Handmaid, Econowife, Unwom-en, Mayday, Prayvaganza, Salvagings, Birthmobile тощо), опису інтер'єру (sitting room, white curtain, sitting room door, kitchen table, white cloth, birthing stool, window seat, double door, back door, folding wooden chair), опису побуту та обов'язків служниць (nightgown, rustle (of a dress), bedsheet, dishtowel, hand lotion, paper napkin; to kneel, to outspread, to unbutton), кольори (red, white). За частиномовною належністю найбільший коефіцієнт ключовості спостерігаємо у займенників (2,7), за якими слідує прислівники (1,5) і дієслова (1,3).

Методологія автоматизованого визначення ключовості апробована в цьому дослідженні може бути застосована до аналізу текстів різного типу. Результати здійсненого дослідження встановлені на основі об'єктивних статистичних критеріїв, слугують попереднім етапом контент-аналізу та дискурс-аналізу, тому ще потребують більш детальної інтерпретації.

Ключові слова: ключовість, художній текст, ключові лексичні одиниці, цільовий корпус, референтний корпус.

Ірина ДІЛАЙ

Candidate of Philological Sciences, Associate Professor at the Department of English Philology, Ivan Franko National University of Lviv, 1 Universytetska str., Lviv, Ukraine, 79000

ORCID: 0000-0001-9626-290X

To cite this article: Dilai, I. (2023). Klyuchovist' v khudozhnyomu teksti (na materiali romanu M. Etwud «Opovid' sluzhnytsi») [Keyness in literary text (based on the novel «The Handmaid's Tale» by M. Atwood)]. *Current Issues of Foreign Philology*, 18, 12–18, doi: <https://doi.org/10.32782/2410-0927-2023-18-2>

KEYNESS IN LITERARY TEXT (BASED ON THE NOVEL «THE HANDMAID'S TALE» BY M. ATWOOD)

The article deals with the automatic detection and further analysis of the key lexical items in the novel «The Handmaid's Tale» by the prominent Canadian postmodern author Margaret Atwood. The aim of the study is to elaborate an efficient methodology of a corpus-driven extraction of key words and phrases from literary text.

Keyness analysis serves as a significant component of the content analysis of the works of art providing clues about prevailing themes, motifs and messages conveyed by the author as well as illuminating insights into the author's narrative style. The keyness in the novels by M. Atwood has not been subject to scrutiny before.

The conducted study consisted of several stages according to the tasks set: 1) the corpus of the text of the novel under analysis has been created and annotated; 2) the comparable reference corpus to detect the key items has been selected and applied; 3) the automatic keyness analysis has been conducted, followed by the further interpretation of the findings. The Sketch Engine corpus tool based on simple math for keywords has been used in order to process the data.

In the course of the analysis, the following key lexical items have been identified: author's neologisms created by M. Atwood to describe the social order in the dystopia called Gilead (Gilead, Aunt, Commander, Handmaid, Econowife, Unwomen, Mayday, Prayvaganza, Salvagings, Birthmobile, etc.), to describe the interior (sitting room, white curtain, sitting room door; kitchen table, white cloth, birthing stool, window seat, double door, back door, folding wooden chair), to describe the routine and duties of Handmaids (nightgown, rustle (of a dress), bedsheet, dishtowel, hand lotion, paper napkin; to kneel, to outspread, to unbutton), colours (red, white), etc. According to the part of speech, the highest keyness score is observed in pronouns (2.7), followed by adverbs (1.5), and verbs (1.3).

The implications of the conducted analysis go beyond the study of the chosen sample of text in view of the fact that the procedure applied to automatically detect key lexical items can be implemented both for literary and non-literary texts. However, the interpretation of the keyness in literary texts requires more subtle content and discourse analyses.

Key words: keyness, literary text, key lexical items, target corpus, reference corpus.

Актуальність проблеми. Ключові слова – це слова, які зустрічаються статистично значно частіше в тексті або корпусі, ніж у порівнюваному, більшому референтному тексті або корпусі (Scott, 1999). Їхня частотність відображає значущість змісту або структури, яка аналізується. Тому, вивчення вживання ключових слів може слугувати відправним пунктом художньої інтерпретації даних (Fischer-Starche, 2010, 65), адже вони вказують на теми або структурні ознаки, які переважають. Актуальним у сучасному прикладному мовознавстві є автоматизоване встановлення ключових слів у різних типах текстів.

Аналіз останніх досліджень і публікацій. Визначення ключових слів має довгу лінгвістичну традицію. Якщо спочатку ключові слова встановлювали інтуїтивно, зараз це здійснюється автоматизовано. Найперші спроби виокремлення ключових слів знаходимо у О. Ладендорфа (1906) «Historisches Schlagwörterbuch (Словнику історичних ключових слів)» та Ф. Леппа (1908) у його праці про ключові слова періоду Реформації. Обидва дослідники інтуїтивно встановлювали ключові слова, які були найбільш важливими для періодів, які вони вивчали. Такий підхід знайшов продовження у німецькомовних країнах: Г. Штраус, У. Гас і Г. Гаррас (1989), а також Д. Герберг (1997), В. Тойберт (1989) аналізували значущі політичні терміни. У Франції Е. Бенвеніст (1954) дослідив конотації та значення французького терміна «цивілізація». У Великобританії Дж. Фірт (1935, 1957) та Р. Вільямс (1983) перші інтуїтивно визначали суспільно вагомі ключові слова (Fischer-Starche, 2010, 67–68).

Сучасний корпусний інструментарій дозволяє автоматизовано визначати ключові слова у різних типах текстів. Детальний опис методик зустрічаємо зокрема у К. Габріелятоса, А. Кілгарріффа, С. Гріса та ін. (Gabrielatos, 2018; Kilgarriff, 2009; Gries, 2010). Встановлення ключових слів у художніх текстах за допомогою корпусного інструментарію є доволі поширеним у сучасній корпусній стилістиці. Зокрема, ключові слова у текстах Дж. Остен досліджувала Б. Фішер-Старке (Fischer-Starche, 2010). Інші дослідження у цьому напрямі здійснюються колективом під керівництвом М. Мальбенг і П. Стоквела в рамках т. з. *CLIC Dickens project* (Mahlberg et al, 2020). Нами раніше було досліджено ключовість у політичних промовах Дж. МакЕйна на основі корпусно-керованого підходу до досліджуваного матеріалу із застосуванням корпусного менеджера UAM (Dilai, 2020).

Мета цього дослідження – апробувати методику виявлення ключових слів у художньому тексті за допомогою корпусного інструментарію. Матеріалом дослідження слугував текст роману-дистопії відомої сучасної канадської письменниці Маргарет Етвуд «Оповідь служниці» (1984), який, як і інші твори письменниці, раніше не досліджувався на предмет ключовості. Мета передбачає виконання таких завдань: 1) здійснити морфолого-синтаксичну анотацію корпусу роману «Оповідь служниці» М. Етвуд; 2) підібрати а) сумісний референтний корпус для здійснення порівняння лексичної канви та виокремлення ключових слів у цільовому корпусі, б) підібрати ефективний корпусний інструментарій, який опирається на

найбільш оптимальні статистичні міри обчислення ключовості у тексті; 3) автоматизовано виокремити ключові слова у тексті роману «Оповідь служниці» М. Етвуд з метою їх наступної інтерпретації.

Виклад основного матеріалу дослідження. Автоматизоване виявлення ключових слів покликане сприяти повній та об'єктивній швидкій інтерпретації тексту, дозволяє здійснити автоматизований контент-аналіз тексту. Не можна стверджувати, що автори навмисно підбирають ключові слова для передачі певних ідей та створення образів. Водночас, у певних випадках, не можна відкидати телеологічної сутності ключових слів та словосполук у тестах, зокрема художніх. Корпусний аналіз покликаний виявити ключові слова незалежно від їхнього свідомого чи несвідомого вживання. Інтуїтивне виявлення ключових слів у тексті є завжди суб'єктивним та намагається встановити випадки свідомого використання автором певних слів в якості стилістичних прийомів, носіїв вагомого змістового навантаження тощо.

Відповідно до різних підходів до ключовості в тексті дослідники розрізняють ключові слова (*keywords*) та ключові слова (*key words*). Якщо перші встановлюють на основі об'єктивних емпіричних корпусних даних, то останні – внаслідок стилістичної інтерпретації тексту, здійсненої фахівцем переважно без використання підрахунків. Обидва способи можуть ефективно та науково переконливо виявити як однакові, так і відмінні ключові слова в одному і тому ж тексті. У нашому дослідженні ключові слова (одиниці) визначаємо за допомогою статистичних прийомів та корпусного інструментарію, спеціально розробленого та апробованого для вивчення ключовості. Застосований підхід видається найбільш оптимальним, оскільки враховує останні напрацювання в галузі обробки природної мови та може бути застосованим до різних текстів.

Особливістю автоматизованого виявлення ключовості у художніх текстах є потреба виходу за межі одного слова, встановлення ключовості словосполук. У термінології корпусної лінгвістики це означає виявлення ключових колокацій (*lexical bundles* за Д. Байбером, 1999), кластерів (Scott, 1999) або *n-grams*. Образність у художньому тексті автоматично експлікується за допомогою таких ключових словосполук. Проте, слід зауважити, що корпусний (корпусо-керований) підхід до ключовості може оминути слова та словосполуки, які є вагомими для розуміння та сприйняття тексту, але не мають високої частоти представленості, щоб їх кваліфікувати як ключові. Проблема охоплення салієнтної інформації корпусним інструментарієм – одна з найбільш актуальних та досі не вирішених.

Отож, беручи за основу частотність лексичного наповнення у художньому тексті, за допомогою корпусного менеджера було встановлено ключовість окремих лексичних одиниць у романі М. Етвуд «Оповідь служниці». Дослідження складалося із кількох етапів, які відповідають викладеним вище завданням. Для здійснення анотації електронного тексту роману, а також для встановлення ключових слів на основі статистичних підрахунків було застосовано *Sketch Engine*, який після апробації інших аналогів обробки лексичної інформації в тексті виявився найоптимальнішим.

Загальну статистичну інформацію щодо лексичної наповненості роману, встановлену на основі *Sketch Engine*, подано у Таблиці 1. Зокрема, наведено частотність вживання унікальних слів (типів) та їх загальну частотність (токенів), що дає можливість обчислити *type/token ratio*. Наведено також частотні показники повнозначних частин мови, які потенційно можуть бути ключовими словами. Дослідження ключовості, як правило, залишає поза увагу так звані *stop words* – функціональні слова.

Таблиця 1

Статистика лексичної наповненості роману М. Етвуд «Оповідь служниці»

Лексичні одиниці	унікальна частотність (<i>types</i>)	загальна частотність (<i>tokens</i>)
Іменники	3898	18969
Дієслова	1547	20650
Прикметники	1563	5836
Прислівники	451	7416
Лексичні одиниці (загал.)	97 314	118 710

Type-token ratio – статистична величина, яка вказує на лексичну насиченість тексту. Для твору, що досліджується, вона складає $0,82 * 100 = 82\%$. Такий показник є доволі високим порівняно з багатьма іншими авторами, дослідженими раніше (Ali, Husein, 2014), та свідчить про значне лексичне розмаїття твору.

Ключові слова хибно вважати найчастотнішими словами у корпусі. З метою виявлення ключових слів застосовують порівняння вживання слів у кількох текстах. Слова, які суттєво відрізняються за статистичними показниками у цих текстах, претендують на ключовість. Саме статистичні критерії ключовості, базовані на відносній частотності, а не абсолютній частотності, застосовуються для її об'єктивного встановлення. Обов'язковою складовою автоматизованого виявлення ключовості є підбір порівнюваного референтного корпусу. Низка дослідників схиляються до думки про те, що референтний корпус повинен бути більшим за розміром, достатньо репрезентативним та збалансованим. Проте для встановлення ключовості в окремому художньому тексті можна підібрати також схожі за стилем, тематикою тексти, референтним корпусом може слугувати корпус усіх творів автора або якийсь інший твір того ж автора. Зрозуміло, що залежно від обраного референтного корпусу результати дослідження можуть суттєво відрізнитися. Тому, проблема відбору оптимального референтного корпусу задля встановлення ключовості є однією із основних у корпусній лінгвістиці.

Після кількох спроб підбору оптимального референтного корпусу для цільового корпусу – тексту роману М. Етвуд «Оповідь служниці» – було прийнято рішення використати доступний у *Sketch Engine* корпус *English Web 2020 (eTenTen)*. Цей корпус великий за розміром (36 млрд слів), охоплює тексти різних жанрів та стилів, доступних у мережі. Обраний референтний корпус анований за тією ж схемою, що й цільовий корпус з використанням *English Penn Treebank tagset*.

Sketch Engine застосовує прості математичні обчислення для встановлення ключовості, які ґрунтуються на відносній (нормалізованій на млн слововживань) частотності (Kilgarriff, 2009), а саме:

$$\frac{fpm_{rmfocus} + N}{fpm_{rmref} + N}$$

де $fpm_{rmfocus}$ – нормалізована (на млн) частотність слова у цільовому корпусі,

fpm_{rmref} – нормалізована (на млн) частотність слова у референтному корпусі, ($N = 1$).

Отож, в результаті зіставлення цільового та референтного корпусів було виявлено ключові слова та ключові словосполучення. При цьому, *Sketch Engine* дозволяє окремо дослідити унікальні ключові слова та звичайні слова, які набувають статусу ключових у творі. Найвищий коефіцієнт унікальної ключовості природно спостерігаємо у власних назвах – іменах персонажів та місцях, зображених у творі: *Moira* (595,3), *Ofglen* (585,5), *Janine* (452,4), *Lydia* (229,4), *Cora* (191,8), *Serena* (172,3), *Gilead* (141) та ін. Таблиця 2 відображає інші унікальні ключові слова у корпусі «Оповідь служниці», на які покладено певне змістове та ідейне навантаження.

Як видно з Таблиці 2, перші позиції серед унікальних ключових слів займають авторські

Таблиця 2

Ключові слова у романі М. Етвуд «Оповідь служниці»

Ключове слово	Коефіцієнт	Ключове слово	Коефіцієнт	Ключове слово	Коефіцієнт
1. Aunt	140	11.Salvagings	43	21.rustle	34
2. Handmaid	130	12.pre-Gilead	43	22.dishtowel	33
3. Gileadian	94	13.unwomen	43	23.salvager	33
4. Mayday	74	14.kneel	42	24.oblong	32
5. nightgown	71	15.Scrabble	38	25.tulip	32
6. washroom	62	16.bedsheet	36	26.outspread	32
7. searchlight	48	17.Prayvaganza	35	27.banister	32
8. shatterproof	47	18.Econowives	35	28.unbutton	30
9. Commander	46	19.cupboard	34	29.ashtray	28
10.Birthmobile	43	20.whispering	34	30.footstool	27

неологізми, створені М. Етвуд для опису суспільного ладу в дистопії під назвою Джилід (*Gil-ead*), серед яких виділяємо суспільні ролі (*Aunt, Commander, Handmaid, Econowife, Unwomen* тощо), суспільні події (*Mayday, Prayvaganza, Salvagings* тощо), реалії (*Birthmobile*). Інші ключові слова вказують на правила поведінки та атмосферу, яка панувала у цьому вигаданому тоталітарному суспільстві, яке нагадувало в'язницю для жінок: *searchlight, shatterproof (windows), whispering*. Жінка у цьому суспільстві була нижчою істотою, на що вказує авторський неологізм, який описує особливо нещасних жінок – *Unwomen*, які мешкали ізольовано від інших і виконували надважку працю в особливо шкідливих умовах. Служниці (*Handmaids*) були сексуальним об'єктом та сурогатними матерями для *Commanders*, позбавленими будь-яких прав і свобод, тому й не випадковим є потрапляння до ключових слів елементів їхнього побуту та їхніх обов'язків: *nightgown, rustle (of a dress), bedsheet, dishtowel; to kneel, to outspread, to unbutton* тощо. У такому суспільстві з'являється бажання займатися давно забутими і забороненими речами, наприклад, пограти у відому настільну гру *Scrabble* потай від інших.

До загальноновживаних ключових слів у романі М. Етвуд потрапляють у порядку спадання ключовості: займенники (*she, I, her, me, it, they, he, we*), дієслова (*say, know, would, think, do, look, could, go, sit, must, want, tell, remember, see, turn, stand, walk, hear, smile, smell, come, suppose, wait, be*), іменники (*woman, hand, eye, room, door, thing, face, head, bed, wife, way, man, dress, mother, voice, hair, chair, foot, joy*), прикметники (*red, white*), прислівники (*there, too, now, maybe, here, then*). З'ясовано, що за части-

номовною належністю найбільший коефіцієнт ключовості є у займенників (2,7), за якими слідує прислівники (1,5) і дієслова (1,3). Ключовість займенників *she* та *I* відображає різні типи оповіді у творі, при цьому переважає відсторонений наратор-спостерігач, який доволі скупо описує дії головних героїв (*walk, sit, stand, go, wait* тощо), практично не вживаючи прислівники на позначення способу дії.

Окрім ключових слів виявлено ключові словосполучення у тексті роману «Оповідь служниці», найважливіші з яких наведено у Таблиці 3. До них також потрапили власні назви, які не включаємо у цю таблицю.

Ключові лексичні одиниці можна об'єднати за семантичною та тематичною схожістю (Stubbs, 2001). Детальніший аналіз ключових словосполучень засвідчує, що значна частина їх належить до тематичної групи опису інтер'єру (*sitting room, white curtain, sitting room door, kitchen table, white cloth, birthing stool, window seat, double door, back door, folding wooden chair* та побутових речей (*hand lotion* – лосьйон для рук, що був розкішною, *paper napkin*). Та особливу увагу привертає кольорова гамма твору: червоний (*red shoe, Red Center, red dress, red brick, red smile*, а також *red skirt, red cloth, red-gloved hand, red habit, red robe, red veil, red light, red berries, red Handmaids, red face, red umbrella, red car, red women, red vinyl, red bodies, red velvet cushions, red stripes, red smoke, red stockings, red arm bands, red spreads, red events, red Birthmobile, red cloak, red pairs, red radiation, red shroud, red writing, red spangles, red balloons*) та білий кольори (*white wing, white cotton, white curtain, white cloth*, а також *white bag, white enamel, white headdress, white cotton socks, white gauze mask, white shift, white night-*

Таблиця 3

Ключові словосполучення у романі М. Етвуд «Оповідь служниці»

Ключові кластери	Коефіцієнт	Ключові кластери	Коефіцієнт	Ключові кластери	Коефіцієнт
1. <i>sitting room</i>	98	11. <i>kitchen table</i>	44	21. <i>dried flower</i>	31
2. <i>white wing</i>	92	12. <i>white cloth</i>	43	22. <i>red brick</i>	29
3. <i>hand lotion</i>	73	13. <i>sitting room door</i>	43	23. <i>back door</i>	27
4. <i>soul scroll</i>	68	14. <i>red dress</i>	41	24. <i>Gileadean regime</i>	27
5. <i>night chapter</i>	57	15. <i>birthing stool</i>	34	25. <i>pig ball</i>	26
6. <i>red shoe</i>	51	16. <i>window seat</i>	34	26. <i>plaster eye</i>	26
7. <i>Red Center</i>	50	17. <i>black van</i>	33	27. <i>underground femaleroad</i>	26
8. <i>white cotton</i>	49	18. <i>double door</i>	33	28. <i>folding wooden chair</i>	26
9. <i>white curtain</i>	49	19. <i>paper napkin</i>	32	29. <i>red smile</i>	26
10. <i>former time</i>	47	20. <i>shirt sleeve</i>	31	30. <i>shopping partner</i>	26

gown, white shape, white steps, white dress, white hair, white bathroom тощо). Служниці носили довгі червоні сукні, червоне взуття та білі каптури. Ключовість цих кольорів вказує на їх символізм та художню образність. Зокрема, червоний колір символізує кров, а також плодючість та непристойність служниць.

Наступним етапом аналізу ключових слів може бути встановлення їхньої дисперсії у тексті та детальніший аналіз високочастотних кластерів, що дає загальне уявлення про структурування тесту (Thornbury, 2010, 280).

Висновки і перспективи подальших досліджень. Виявленні автоматизованим способом ключові лексичні одиниці у романі М. Етвуд «Оповідь служниці» підтверджують тематичну спрямованість роману-дистопії: сто-

суються опису тоталітарного режиму дистопії, гендерної тематики. Окрім цього, ключові слова та словосполучення вказують на своєрідний авторський художній стиль, стосуються символів та образів, створених авторкою, унікального нараторського стилю, які вимагають подальшої інтерпретації та залучення інших методів дослідження. Таким чином, ключовість художнього твору є об'єктивним підґрунтям для проведення контент-аналізу та критичного аналізу дискурсу.

Перспективи подальших досліджень вбачаємо саме у якісній інтерпретації результатів статистичного виявлення ключовості, а також в апробації методології автоматизованого виявлення ключових слів на інших художніх творах.

ЛІТЕРАТУРА:

1. Ali S. M., Husein K. S. The comparative power of type/token and hapax legomena/type ratios: a corpus-based study of authorial differentiation. *Advances in Language and Literary Studies*. 2014, Vol. 5, No 3. P. 112–119.
2. Atwood M. *The Handmaid's Tale*. London: Vintage, 1996. 320 p.
3. Dilai I., Dilai M. Automatic extraction of keywords in political speeches. *Proceedings of IEEE CSIT 2020*. 23–24 September 2020, Zbarazh – Lviv. P. 291–294.
4. Fischer-Starcke B. *Corpus Linguistics in Literary Analysis. Jane Austen and her Contemporaries: Studies in Corpus and Discourse*. W. Teubert, M. Mahlberg [Eds.]. London, New York: Continuum International Publishing Group, 2010. 226 p.
5. Gabrielatos C. Keyness analysis: Nature, metrics and techniques. Taylor, C., Marchi, A. (eds.) *Corpus Approaches to Discourse: A critical review*. London: Routledge, 2018. P. 225–258.
6. Gries S. Th. Useful statistics for corpus linguistics. *A Mosaic of Corpus Linguistics: Selected Approaches*. Sánchez, A. & Almela, M. eds. Frankfurt am Main: Peter Lang, 2010. P. 269–291.
7. Kilgariff A. Simple maths for keywords. Proceedings of Corpus Linguistics Conference CL2009. Mahlberg, M., González-Díaz, V. & Smith, C. (eds.), University of Liverpool, UK, July 2009. URL: <https://www.sketchengine.eu/wp-content/uploads/2015/04/2009-Simple-maths-for-keywords.pdf>
8. Mahlberg M., Stockwell P., Wiegand V., Lentin J. *CLiC 2.1. Corpus Linguistics in Context*, 2020. URL: <https://www.clic.bham.ac.uk>
9. Scott M. *WordSmith Tools. 3.0*. Oxford: OUP, 1999. 149 p.
10. Stubbs M. *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford: Blackwell, 2001. 267 p.
11. Sketch Engine. URL: <https://www.sketchengine.eu/>
12. Thornbury S. What can a corpus tell us about discourse? *The Routledge Handbook of Corpus Linguistics*. A. O'Keefe, M. McCarthy (eds.) London: Routledge, 2010. P. 270–287.

REFERENCES:

1. Ali, S. M. & K. S., Husein. (2014). The comparative power of type/token and hapax legomena/type ratios: a corpus-based study of authorial differentiation. *Advances in Language and Literary Studies*, 5(3), 112–119.
2. Atwood, M. (1996). *The Handmaid's Tale*. London : Vintage.
3. Dilai, I. & Dilai, M. (2020). Automatic extraction of keywords in political speeches. *Proceedings of IEEE CSIT 2020*. 23–24 September 2020, Zbarazh – Lviv. 291–294.
4. Fischer-Starcke, B. *Corpus Linguistics in Literary Analysis. Jane Austen and her Contemporaries: Studies in Corpus and Discourse*, W. Teubert, M. Mahlberg (Eds.). London, New York : Continuum International Publishing Group, 2010.
5. Gabrielatos, C. (2018). Keyness analysis: Nature, metrics and techniques. Taylor, C., Marchi, A. (eds.) *Corpus Approaches to Discourse: A critical review*. London : Routledge. 225–258.

6. Gries, S. Th. (2010). Useful statistics for corpus linguistics. *A Mosaic of Corpus Linguistics: Selected Approaches*, Sánchez, A. & Almela, M. (Eds.), (pp. 269–291). Frankfurt am Main : Peter Lang.
7. Kilgarriff, A. (2009). Simple maths for keywords. Proceedings of Corpus Linguistics Conference CL2009, Mahlberg, M., González-Díaz, V. & Smith, C. (Eds.), University of Liverpool, UK. Retrieved from: <https://www.sketch-engine.eu/wp-content/uploads/2015/04/2009-Simple-maths-for-keywords.pdf>
8. Mahlberg, M., Stockwell, P., Wiegand, V. & J. Lentin. (2020). *CLiC 2.1. Corpus Linguistics in Context*. Retrieved from: <https://www.clic.bham.ac.uk>
9. Scott, M. (1999). *WordSmith Tools. 3.0*. Oxford : OUP.
10. Stubbs, M. (2001). *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford : Blackwell.
11. *Sketch Engine*. Retrieved from: <https://www.sketchengine.eu/>
12. Thornbury, S. (2010). What can a corpus tell us about discourse? *The Routledge Handbook of Corpus Linguistics*, A. O'Keeffe, M. McCarthy (Eds.), (pp. 270–287). London : Routledge.