CORPORA AND LITERATURE

Mahmudova Dildora Murodilloyevna
ESP Teacher, Bukhara State Medical Institute named
after Abu Ali ibn Sino Bukhara, Uzbekistan

Abstract

Literary corpora comprise poetry and fictional prose texts, such as novels, short stories and plays. They bring together the collected works of a single author or representative texts from a specific literary period. Since literary corpora are often available through powerful concordances, they are especially well suited for a quantitative and qualitative approach to comparative literary analysis, within or across different genres and historical periods. This paper attempts to analyze the relationship of corpus and literature.

Keywords: corpus, literature, corpora

КОРПОРА И ЛИТЕРАТУРА

Аннотация

Литературные корпуса включают стихи и художественные прозаические тексты, такие как романы, рассказы и пьесы. Они объединяют собрания сочинений одного автора или репрезентативные тексты определенного литературного периода. Поскольку литературные корпуса часто доступны через мощные симфонии, они особенно хорошо подходят для количественного и качественного подхода к сравнительному литературному анализу внутри или между различными жанрами и историческими периодами. В данной статье предпринята попытка проанализировать взаимосвязь корпуса и литературы.

Ключевые слова: корпус, литература, корпусы.

Inroduction

The analysis of literary texts using corpora is a common methodological approach in the field of stylistics. Stylistics, which is the study of style in language, is a sub-discipline of linguistics within applied linguistics. While stylisticians typically focus on analyzing literary texts, there is no reason why stylistics should be limited to this type of text. In fact, some stylisticians have applied the same analytical techniques to non-literary texts. However, the principles of stylistics are rooted in the study of

literary texts, and corpora are valuable tools for testing hypotheses and theories about texts and stylistic effects. In recent years, the use of corpora has provided statistical evidence that supports foregrounding theory, a fundamental aspect of stylistics.

Foregrounding theory posits that an author's intentional linguistic choices play a crucial role in shaping their style. These choices give rise to specific stylistic effects, which can be observed at the overall authorial level, within a particular genre, or even in a specific text. According to this theory, the most significant stylistic elements are those that deviate from the established norms of language, whether those norms are external or internal to the text itself. Psychological studies have lent support to this theory by demonstrating that readers generally agree on what is considered foregrounded in a text.

However, determining the extent of foregrounding has historically been challenging due to limited access to frequency information about linguistic norms. The advent of large-scale electronic corpora has greatly facilitated this process, as it provides linguists with the necessary tools to analyze linguistic patterns and establish statistical norms. Some linguists have questioned the utility of relying on frequency information, even if it were readily available. Nevertheless, these objections are flawed because understanding the statistical norms of language is crucial in determining whether a particular usage is indeed unusual and therefore foregrounded.

Modern corpus linguistics has emerged as an invaluable resource for stylisticians, enabling them to uncover the statistical norms of language and assess the foregrounded elements within a text accurately. This development has debunked views opposing the use of corpora, such as Freeman's, and highlighted the immense value of corpora in stylistic analysis. Although corpus-related work in stylistics has not progressed as rapidly as in mainstream linguistics, it is steadily growing, and the term "corpus stylistics" is increasingly employed to describe this approach.

The term "collocates" refers to words that are often found near a specific word in a text. In the field of corpus stylistics, analyzing the collocates of a word in a reference corpus can provide insights into patterns and associations that can aid in understanding its meaning or stylistic effects. For instance, by examining the collocates of the word "priest" in the British National Corpus, we can see that the most frequently occurring collocate is "lecherous," which suggests a potential alternative interpretation or meaning of the poet's description. Approaches to corpus stylistics can be categorized into three types: corpus-assisted analysis, corpus-based analysis, and corpus-driven analysis. Although these distinctions may have their flaws, they can be helpful as a heuristic tool for identifying common practices in corpus research. Corpus-assisted analysis involves using large reference corpora to support the

analysis of a single text. This type of analysis can be employed to confirm or challenge a stylistician's intuition regarding the stylistic effects of specific linguistic structures. For example, in Roger McGough's poem "Vinegar," the poet depicts waiting alone in a queue at a fish and chip shop. While contemplating how it would be nicer to buy supper for two, the poet compares himself to a priest. Drawing on our knowledge of priests, particularly the typical assumption that they are unmarried, we might infer that the speaker is reflecting on feeling lonely. However, conducting a search for the collocates of the word "priest" in a reference corpus reveals an alternative potential meaning. The collocates of a particular word are other words that commonly appear nearby. By examining the British National Corpus, which consists of 100 million words of late twentieth century British English, we discover that the most frequent collocate of "priest" is "lecherous."

This suggests that an alternative interpretation of the poet's meaning in 'Vinegar' is that he is expressing a desire for a sexual rather than an emotional relationship. It is important to consider the range of texts in which a collocate is present in a corpus for stylistic analysis. However, even if there was only one occurrence of "lecherous" as a collocate of "priest," it would validate the response of a reader who interpreted this connotation in the poem. Collocation does not indicate the "correct" interpretation of a text, but it does indicate the range of valid interpretations. A notable example of corpus-assisted analysis is O'Halloran's work (2007), and it is perhaps this type of corpus-related research that will be most appealing to traditional literary critics. It does not require the construction of specialized corpora, as there are commercially available corpora that can be used for corpus-assisted work. Additionally, it does not necessitate expertise in computing and corpus analytical techniques. Instead, corpus-assisted literary analysis utilizes readily accessible information to provide support for more conventional qualitative analyses of literary texts.

There are three types of approaches to corpus stylistics: corpus-assisted analysis, corpus-based analysis, and corpus-driven analysis. While these categorizations may have their limitations, they can serve as a useful heuristic tool for identifying common practices in corpus research. Corpus-assisted analysis involves utilizing large reference corpora to support the analysis of a single text. This type of analysis can be employed to validate or challenge a stylistician's intuition regarding the stylistic effects of specific linguistic structures.

The techniques described in this article require the use of corpus analysis software. There are numerous commercially available packages for this purpose, with WordSmith Tools (Scott, 2004) being perhaps the most well-known. However, other packages such as AntConc (described in Anthony, 2004) are equally effective and, in

some cases, free. The web-based tool Wmatrix (Rayson, 2008) is also highly recommended. In addition to the conventional corpus tools featured in WordSmith and AntConc, Wmatrix automatically annotates uploaded texts for part-of-speech and semantic categories, enabling the analysis of key parts-of-speech and key semantic domains. All of these software packages can be found online. Furthermore, a range of electronic texts and corpora, many of which are freely available, are stored at the Oxford Text Archive (http://ota.ahds.ac.uk/). As literary scholars become increasingly aware of the advantages offered by corpora, the use of corpora in literary analysis is likely to increase. This approach to literary studies significantly improves upon more subjective methods by allowing for a more objective and falsifiable approach to text analysis. As Sinclair (2004) emphasizes, in all approaches to linguistic analysis, it is crucial that we "trust the text." Utilizing corpora enables us to do just that.

In conclusion, understanding the statistical norms of language is essential for identifying fore grounded elements within a text. The rise of modern corpus linguistics has provided stylisticians with the necessary means to analyze linguistic patterns and establish these norms accurately. Despite some initial skepticism, corpora have proven to be invaluable tools in stylistic analysis, and their usage in the field is steadily increasing.

References

1. Louw, B. (1997). The role of corpora in critical literary appreciation. In A. Wichman, S. Fligelstone, T. McEnery & G. Knowles (Eds.), Teaching and Language Corpora, (pp. 240-251). Harlow: Longman.

2. D. Z. Olimova, & M. D. Mahmudova. (2022). POLITICAL DISCOURSE AND TRANSLATION. RESEARCH AND EDUCATION, 1(3), 176–179. 2022

3. Saidov Akmal Azimovich, Mahmudova Dildora Murodilloyevna. Anticipation strategy in simultaneous interpretation of political discourse. Spanish journal of Innovation and integrity Volume:12,110-116. November-2022.

4. Mahlberg, M. (2007). A corpus stylistic perspective on Dickens' Great Expectations. In M. Lambrou and P. Stockwell (Eds.), Contemporary Stylistics, (pp. 19-31). London: Continuum.

5. O'Halloran, K. A. (2007). The subconscious in James Joyce's 'Eveline': a corpus stylistic analysis which chews on the 'Fish hook'. Language and Literature, 16(3), 227-244.

6. Davlatova M.H. Relation of lexical-semantic structure of verbs in the linguistic essence.// IEJRD-International Multidisciplinary Journal, № 5, 2021. – P. 58-62.

7. Davlatova M.H. Semantic implementation of resultutive structures.// JournalNX - A Multidisciplinary Peer Reviewed Journal, № 6, 2021. – P. 291–294.

8. Davlatova M.H. Different aspects of resultative structures according to their linguistic essence.// Academicia Globe: Inderscience Research, 2 (05), 2021. – P. 475–479.

9. Davlatova M.H. Aspectual and lexical-semantic classification of verbs. JournalNX-A Multidisciplinary Peer Reviewed Journal, 8(2). 2022. – P. 116-121.

10. Davlatova M.H. On the interrelationship of resultive and causative meanings.// World Bulletin of Public Health № 9. 2022. – P. 212-215.

11. Habibova, M. N. (2021). Jorjina Houellning "Queen of the desert" biografik asarida gertruda Bell timsoli tasviri. *Academic research in educational sciences*, *2*(2), 770-778.

12. Habibova, M. N. (2021). The theme feminism in the epistolary novels in modern times. *ISJ Theoretical & Applied Science*, *11*(103), 1101-1105.

13. Habibova, M. N. (2022). THE HISTORY OF THE RUSSIAN EPISTOLARY NOVEL IN EIGHTEENTH CENTURY. EURASIAN JOURNAL OF SOCIAL SCIENCES, PHILOSOPHY AND CULTURE, 2 (3), 135–139.

14. Habibova, M. N. (2022). EVALUATIVE OBSERVATION ON DH LAWRENCE'S EPISTOLARY ACHIEVEMENT. EURASIAN JOURNAL OF ACADEMIC RESEARCH, 2 (4), 489–494.

15. Habibova, M. (2022). THEORIES OF INTERTEXTUALITY AND THE BASIC FRAMEWORK OF KRISTEVA'S FORMULATION OF HER THEORY OF INTERTEXTUALITY. *Eurasian Journal of Social Sciences, Philosophy and Culture*, *2*(5), 301-307.

16. Habibova, Manzila Nuriddinovna (2022). A BIOGRAPHY IS A SPECIALIZED FORM OF HISTORY AND BASIC TYPES OF BIOGRAPHIES. Oriental renaissance: Innovative, educational, natural and social sciences, 2 (5), 495-503.

17. Habibova, M. N. (2022). Epistolary Novel as a Scientific Problem. American Journal of Social and Humanitarian Research, 3(10), 211–214. Retrieved from https://www.grnjournals.us/index.php/ajshr/article/view/1575

18. Habibova, M. N. (2022). The Epistolary Form and Art in Modernist Literature. American Journal of Social and Humanitarian Research, 3(10), 206–210. Retrieved from https://www.grnjournals.us/index.php/ajshr/article/view/1574

19. Habibova, M. (2022). WRITING THE NOVEL OF LETTERS AND THE ANALYSIS OF LETTER FICTION AND DIARY FICTION. International Bulletin of Applied Science and Technology, 2(11), 260–264. Retrieved from https://researchcitations.com/index.php/ibast/article/view/317.

20. Habibova, M. (2022). EPISTOLARY NARRATOLOGY AND CITATIONS FROM AN EPISTOLARY SOURCE: SAMUEL RICHARDSON'S PAMELA AND CLARISSA AND FRANCES BURNEY'S EVELINA. *Евразийский журнал академических исследований*, *2*(12), 1131–1135. извлечено от https://in-academy.uz/index.php/ejar/article/view/6655

21. Manzila Nuriddinovna Khabibova. (2022). JOYCE'S DEVELOPMENT AS AN AUTHOR AND HIS EXPERIMENTS WITH THE EPISTOLARY FORM. American Journal of Philological Sciences, 2(11), 73–81. https://doi.org/10.37547/ajps/Volume02Issue11-10

22. Manzila Nuriddinovna Habibova. (2022). LETTER WRITING AS ALLEGORY FOR THE WOMAN WRITER. American Journal Of Philological Sciences, 2(11), 88–92. https://doi.org/10.37547/ajps/Volume02Issue11-12

23. Хабибова, М. Н. (2022). ЭПИСТОЛЯРНЫЙ РОМАН КАК НАУЧНАЯ ПРОБЛЕМА. Results of National Scientific Research International Journal, 1(9), 394–402. Retrieved from https://academicsresearch.com/index.php/rnsr/article/view/1379

24. Habibova, M. N. (2023). Origins of the Epistolary Novel and the Principal Epistolary Works. *INTERNATIONAL JOURNAL OF LANGUAGE LEARNING AND APPLIED LINGUISTICS*, *2*(2), 70–74. Retrieved from https://inter-publishing.com/index.php/IJLLAL/article/view/1159

25. Khabibova, M. N. . (2023). The Historical Influence of the Epistolary Novel. *INTERNATIONAL JOURNAL OF LANGUAGE LEARNING AND APPLIED LINGUISTICS*, *2*(2), 75–79. Retrieved from https://inter-publishing.com/index.php/IJLLAL/article/view/1161

26. Habibova, M. (2023). SPECIFIC POSSIBILITIES WRITING TECHNIQUES USED IN EPISTOLARY FORM AND THEIR EFFECTS ON BOTH THE CHARACTERS AND READERS. *Eurasian Journal of Academic Research*, *3*(5),172–177.Retrieved from https://in-academy.uz/index.php/ejar/article/view/14085

27. Habibova, M. (2023). EPISTOLARY COMPARISONS AND TRENDS. *Евразийский журнал академических исследований*, *3*(5 Part 4), 330–338. извлечено от https://in-academy.uz/index.php/ejar/article/view/16170

28. Nasriyeva Guzal Zulfiddin Qizi. (2023). DEPICTION OF SOCIAL LIFE IN THE NOVEL "MEHRODDAN CHAYON" BY ABDULLA KADIRI. International Journal Of Literature And Languages, 3(04), 88–94. https://doi.org/10.37547/ijll/Volume03Issue04-16.

29. Nasriyeva Guzal Zulfiddin Qizi. (2023). THE CONCEPT OF TOLERANCE IN ENGLISH MEDIA AND PRESS. International Journal of Literature and Languages, 3(02), 74–81. https://doi.org/10.37547/ijll/Volume03Issue02-14

30. Qizi, N. G. Z. (2022). Psychoanalytical Approach in English Literature. Central Asian Journal of Literature, Philosophy and Culture, 3(12), 34-38. Retrieved from https://www.cajlpc.centralasianstudies.org/index.php/CAJLPC/article/view/632

31. Qizi, N. G. Z. (2022). Literature as Discourse. American Journal of Social and Humanitarian Research, 3(10), 266–271. Retrieved from https://www.grnjournals.us/index.php/ajshr/article/view/1600

32. qizi, N. G. Z. . (2022). Hermeneutic Method in Literature. American Journal of Social and Humanitarian Research, 3(9), 197–202. Retrieved from https://www.grnjournals.us/index.php/ajshr/article/view/1510

33. Davlatova M.H. The process of transformation of philosophy understanding as factor of information culture of the period of the English renaissance (XVI c.).// Oriental renaissance: Innovative, educational, natural and social sciences, № 5. 2021. – P. 949-957.

34. Davlatova M.H. The expression of resultative and depictive constructions in English and Uzbek languages.// So'z san'ati xalqaro jurnali. – Toshkent, 2021. – P. 102-107.

35. Давлатова М.Х. Результативликнинг аспектуал тадқиқотлар доирасида ўрганилиши.// Scientific progress, № 2. – Тошкент, 2021. – Б. 1678-1683.

36. Давлатова М.Х. Инглиз ва ўзбек тилларида результативликни ифодаловчи лексик-грамматик воситалар.// Тил адабиёт та`лим илмий-услубий журнал. – Самарқанд, 2021. - Б. 70-77.

37. Давлатова М.Х. Результатив мазмундаги воқеа - ҳодисалар таркибланишининг лисоний ва когнитив хусусиятлари.// Хоразм маъмун академияси ахборотномаси. № 11. – Хоразм, 2021. – Б. 361-363.

38. Tursunbaevna, N. Z. (2022). Types of Interactive Methods of Teaching Foreign Languages in Higher Education Institutions. American Journal of Social and Humanitarian Research, 3(10), 151-155.

39. Nematova , Z. . (2022). THE TYPES OF ERROR IN SECOND LANGUAGE LEARNING. Евразийский журнал академических исследований, 2(12), 310–314. извлечено от https://www.in-academy.uz/index.php/ejar/article/view/7227

40. Nematova , Z. . (2022). VARIOUS EXAMPLES OF DIFFERENTIATED INSTRUCTION IN THE CLASSROOM, ADVANTAGES AND DISADVANTAGES OF DIFFERENTIATED INSTRUCTION. Евразийский журнал академических

исследований, 2(12), 315–318. извлечено от https://in-academy.uz/index.php/ejar/article/view/7228

41. Nematova, Z. (2023). THE METHODS OF DEVELOPING LOGICAL THINKING USING AUTHENTIC MATERIALS IN THE CLASSROOM. International Bulletin of Applied Science and Technology, 3(3), 183–188. Retrieved from https://www.researchcitations.com/index.php/ibast/article/view/749.