

PRINCIPLES OF CREATING COMPARATIVE BILINGUAL SYNCHRONIC CORPUS

Mahmudova Dildora Murodilloyevna
BSMI, the teacher of English language department

Abstract: *This paper aims to explore the challenges of comparison in cross-linguistic corpus-assisted discourse studies, raising important questions such as how to compare different languages in corpora, how to determine relevant language units for comparison, and how to ensure comparable data. Through case studies, the paper first examines lexical comparison and then explores methods for analyzing beyond the lexical level, including semantic fields, discourse frames, and rhetorical features. By discussing issues and potential solutions for comparing multilingual corpora, the paper seeks to spark a constructive dialogue to enhance the field's understanding and resources.*

Key words: *cross-linguistic, corpus-assisted discourse studies, lexical level, lexical comparison, multilingual corpora, methodology.*

INTRODUCTION

The integration of discourse analysis and corpus linguistics, as discussed in Baker (2006), Partington, Duguid, et al. (2013), and Mautner (2015), combines different methodological approaches and theoretical perspectives on language analysis. This methodology remains flexible and adept at balancing qualitative and quantitative research demands. Corpus linguistics offers the advantage of analyzing a large volume of texts, providing a unique vantage point on the data. Fairclough (1989: 54) highlighted the cumulative impact of media power through repeated patterns in language use, such as causality, agency, and reader positioning. Corpus linguistics allows for tracking this cumulative effect, especially in examining word associations like collocations. Understanding which words typically co-occur can provide insights into the nuanced meanings, including evaluative connotations, of the specific lexical item under scrutiny. The critical discourse studies aspect, complementing corpus and discourse analysis, contributes both theoretically and methodologically to the framework. The theoretical dimension involves assumptions about the functioning of language in society and the interactive nature of this relationship, underscoring the significance of discourse analysis. Methodologically, a range of analytical tools enables the recognition and labeling of language usage patterns.

The resulting fusion, commonly known as corpus-assisted discourse studies (CADS), has been widely utilized in various interdisciplinary and multidisciplinary settings. Ancarno (2018) explains that these interactions can occur in three main ways: 1) where CADS research informs or is informed by other disciplines, 2) where CADS research contributes to non-academic contexts, and 3) where CADS research collaborates with other disciplines. In the context of this special issue, our primary focus lies on the third scenario. Our goal is to delve deeper into the area of cross-linguistic CADS and address the

challenges associated with studying discourse across languages, aiming to initiate a dialogue with Translation Studies.

As evident from the earlier discussion on the various types of studies analyzing discourse across languages through corpus linguistics, most studies do not utilize multilingual corpora. In fact, this field is more traditional than it may initially seem. The majority of research relies on multiple monolingual sub-corpora. For example, Del Fante (2018) examined the prominence of country of origin in representations of immigrants in the British and Italian press pre and post the 2016 Brexit referendum. They compiled two sets of monolingual corpora: two Italian newspapers and two English-language newspapers. These corpora were comparable as they were sourced from similar outlets (newspapers) and used similar search terms to identify articles on migration. This type of synchronic comparable monolingual corpora is the most commonly used approach in current CL-CADS studies. Additionally, Nardone (2018), who focuses on German-Italian comparisons, notes that English is often included as one of the language sub-corpora (depending on the publication language).

Another frequently employed type of corpus in cross-linguistic research is the parallel corpus, although discourse analysis has not been as prominently featured in this context. In a parallel corpus comparison, researchers typically work with multiple monolingual corpora, with one serving as the source text and one or more as the translated texts.

One of the primary challenges in working with corpora is determining whether similar text genres exist across different languages and cultures. For example, when analyzing press discourse, researchers often encounter differences in categorizations such as tabloid versus broadsheet in the UK, which may not be applicable in other countries where distinctions could be based on factors like frequency of publication or geographic scope. Understanding the purpose and target audience of each text type is crucial for evaluating their comparability in cross-linguistic studies.

Creating a discourse-complete corpus involves compiling all texts within a specific domain and time period. For example, the SiBol corpora in Sketch Engine include articles from various newspapers published in 1993, 2005, 2010, and 2013. Designing comparable corpora becomes simpler when comparable newspapers are identified. On the other hand, in a search-term corpus, which is more commonly used due to practical reasons, the primary challenge lies in determining relevant topics.

Another essential aspect of corpus work involves examining collocations, which are words that frequently appear together and have a significant textual connection. Unlike keyness analysis, collocation analysis begins with selecting specific words as nodes for comparison. This process requires ensuring that the chosen terms are functionally equivalent, similar to the initial step of identifying search terms for constructing a corpus. Therefore, before delving into collocation studies, a comparative analysis is necessary to establish the appropriate terms for analysis.

At its core, a corpus-based discourse analysis can involve comparing lexical elements by examining the lexical characteristics of two or more equivalent terms in different languages. This comparison can serve as the primary focus of the analysis or act as a

preliminary step to determine search terms for corpus compilation and identify comparable terms for collocation analysis. This approach is often considered the most conventional starting point and aligns closely with the historical roots of corpus linguistics in lexicography.

REFERENCES:

1. Olimova, D. Z., & Mahmudova, M. D. (2022). POLITICAL DISCOURSE AND TRANSLATION. RESEARCH AND EDUCATION, 1(3), 176–179.
2. Azimovich: Anticipation strategy in simultaneous... - Академия Google
https://scholar.google.com/scholar?hl=ru&as_sdt=0,5&cluster=11648826108631162856&authuser=1
3. Mahmudova, D. (2023). CORPUS LINGUISTICS. В ACADEMIC RESEARCH IN MODERN SCIENCE (Т. 2, Выпуск 23, сс. 104–106).
4. Mahmudova, D. (2023). CORPORA AND LITERATURE. В CURRENT APPROACHES AND NEW RESEARCH IN MODERN SCIENCES (Т. 2, Выпуск 10, сс. 63–64).
5. Mahmudova Dildora Murodilloyevna. (2023). EXPLORING THE POWER OF CORPORA IN LITERATURE: UNVEILING INSIGHTS, ENHANCING INTERPRETATIONS. Academia Repository, 4(11), 167–174.
6. Mahmudova Dildora Murodilloyevna. (2023). CORPORA AND LITERATURE. Academia Repository, 4(11), 159–166.
7. Davlatova M.H. Relation of lexical-semantic structure of verbs in the linguistic essence.// IEJRD-International Multidisciplinary Journal, № 5, 2021. – P. 58-62.
8. Davlatova M.H. Semantic implementation of resultative structures.// JournalNX - A Multidisciplinary Peer Reviewed Journal, № 6, 2021. – P. 291–294.
9. Davlatova M.H. Different aspects of resultative structures according to their linguistic essence.// Academicia Globe: Inderscience Research, 2 (05), 2021. – P. 475–479.
10. Davlatova M.H. Aspectual and lexical-semantic classification of verbs. JournalNX-A Multidisciplinary Peer Reviewed Journal, 8(2). 2022. – P. 116-121.
11. Davlatova M.H. On the interrelationship of resultive and causative meanings.// World Bulletin of Public Health № 9. 2022. – P. 212-215.
12. Qizi, N. G. Z. (2022). Psychoanalytical Approach in English Literature. Central Asian Journal of Literature, Philosophy and Culture, 3(12), 34-38.
13. 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>
14. 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>

15. 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>
16. 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.
17. 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.
18. Habibova, M. N. (2021). The theme feminism in the epistolary novels in modern times. ISJ Theoretical & Applied Science, 11(103), 1101–1105.
19. 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.
20. Habibova, M. N. (2022). EVALUATIVE OBSERVATION ON DH LAWRENCE'S EPISTOLARY ACHIEVEMENT. EURASIAN JOURNAL OF ACADEMIC RESEARCH, 2 (4), 489–494.
21. 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.
22. 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.
23. 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>
24. 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>
25. 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>
26. 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>
27. 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>
28. 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>
29. Хабибова, М. Н. (2022). ЭПИСТОЛЯРНЫЙ РОМАН КАК НАУЧНАЯ ПРОБЛЕМА. Results of National Scientific Research International Journal, 1(9), 394–402. Retrieved from <http://academics.uz/index.php/rnsr/article/view/1379>
30. 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>
31. 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>
32. 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>
33. 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>
34. Habibova, M. (2023). PRAGMATICS AND COGNITION. Current approaches and new research in modern sciences, 2(10), 90-96.
35. Manzila Nuriddinovna Habibova. (2023). THE CONCEPT OF THE WHITE MAN'S SUPERIORITY AND THE PORTRAYAL OF THE ARABIC CHARACTERS IN "THE LAWRENCE OF ARABIA" AND "THE SEVEN PILLARS OF WISDOM". Academia Repository, 4(10), 24–31. Retrieved from <https://academiarepo.org/index.php/1/article/view/96>
36. Habibova, M. N. (2023). Letter Fiction and Diary Fiction. American Journal of Public Diplomacy and International Studies (2993-2157), 1(9), 153–158. Retrieved from <http://grnjournal.us/index.php/AJPDIS/article/view/1386>
37. Habibova, M. N. (2023). Women and the Epistolary Novel. American Journal of Science on Integration and Human Development (2993-2750), 1(9), 119-124.
38. Habibova, M. (2023). TE LAWRENCE AND ARAB NATIONALISTS IN THE OTTOMAN EMPIRE. THEORY AND ANALYTICAL ASPECTS OF RECENT RESEARCH 2 (20), 183-193
39. Kholova Madina Boboqulovna. (2023). POLITENESS AND CULTURE. Academia Repository, 4(10), 4–9. Retrieved from <https://academiarepo.org/index.php/1/article/view/139>

40. Kholova, M. (2023). PRAGMATICS. В МЕЖДУРОДНАЯ КОНФЕРЕНЦИЯ АКАДЕМИЧЕСКИХ НАУК (Т. 2, Выпуск 10, сс. 68–74). Zenodo. <https://doi.org/10.5281/zenodo.10054322>
41. Kholova Madina Boboqulovna. (2023). Cognitive Linguistics. American Journal of Language, Literacy and Learning in STEM Education (2993-2769), 1(10), 357–363. Retrieved from <https://grnjournal.us/index.php/STEM/article/view/2018>
42. Kholova Madina Boboqulovna. (2023). COGNITION AND LANGUAGE. Scholastic: Journal of Natural and Medical Education, 2(12), 64–71. Retrieved from <https://univerpubl.com/index.php/scholastic/article/view/3014>