

Types of Synchronous Corpus and Their Significance

Mahmudova Dildora Murodilloyevna

*PhD student of Bukhara State University, ESP teacher at Bukhara State Medical Institute named
after Abu Ali ibn Sino*

Abstract. *A synchronous corpus is a collection of texts or conversations that have been recorded or transcribed in real-time. There are several types of synchronous corpora, each serving a specific purpose in linguistic research and language study. Some common types of synchronous corpora include spoken language corpora, chat corpora, and text messaging corpora. Synchronous corpora play a crucial role in linguistic research by providing researchers with real-life examples of language use in various contexts. They offer a rich source of data for studying language variation, language change, and the impact of technology on language use. By analyzing synchronous corpora, researchers can gain a deeper understanding of how language is used in everyday communication and how it evolves over time.*

Key words: *corpus, literature, corpora, stylistician, literary texts.*

Introduction. A text corpus is a very large collection of texts (often many billion words) produced by actual users of a language and used to analyze words, phrases, and how language is used in general. It is used by linguists, lexicographers, social scientists, humanities, natural language processing specialists, and many others. The corpus is also used to build various language databases used in software development, such as predictive keyboards, spell checking, grammar correction, text/speech understanding systems, text-to-speech modules, machine translation systems, etc.

The corpus cannot be easily classified into a specific category. Instead, a corpus can have features or properties that can be used to group them. The same body can have one or more of these properties.

Monolingual corpus. A monolingual corpus is the most common type of corpus. It contains texts in only one language. Corpora are usually labeled for parts of speech and are used by a wide range of users for a variety of situations from very practical tasks, such as: checking the correct usage of a word or searching for the most natural word combinations, to scientific uses, such as: in language identifying frequent patterns or emerging trends. Sketch Engine includes hundreds of monolingual corpora in dozens of languages.

Parallel corpus, multilingual corpus. A parallel corpus consists of two or more monolingual corpora. Corpora is a translation of each other. For example, a novel and its translation or the translation memory of a CAT tool can be used to build a parallel corpus. Both languages must be aligned, that is, matching segments, usually sentences or paragraphs, must match. The user can then search for all instances of a word or phrase in one language, and the results will be displayed along with corresponding sentences in the other language. Then the user can observe how the searched word or phrase is translated. Parallel corpora are central to translation studies and contrastive linguistics. Many of the parallel corpora are accessible through convenient concordancers that greatly facilitate the study of cross-linguistic phenomena. Such corpora are also a rich source of materials for language teaching. In addition, parallel corpora serve as training data for statistical machine translation systems.

Comparative corpus. A comparative corpus is one corpus in a collection of two or more monolingual corpora, usually each in a different language, built according to the same principles. Therefore, the content is similar and the results can be compared between the corpora, even though they are not translations of each other (so they are not the same). When users search these corpora, they can take advantage of the fact that the corpora also have the same metadata. Examples of comparable corpora in Sketch Engine are the CHILDES corpora or various corpora from Wikipedia. Arcanum cases are also comparable.

Diachronic corpus. A diachronic corpus is a corpus containing texts from different periods and is used to study the development or change of a language. Sketch Engine allows you to search the corpus as a whole or only containing selected time intervals. In addition, there is a special diachronic feature called Trends that identifies the words whose usage has changed the most during a selected time period.

Synchronous case. The opposite is a synchronic corpus, whose texts originate from the same time. This is a snapshot of the language. The TenTen family is such a picture, because their content is collected over several months.

A static body (also called a reference body (although it refers to something else in Sketch Engine)) is a body whose development is complete. The content of the body does not change. Most bodies are static bodies. Benefits of an immutable body: analysis consistency of results is important in many scenarios.

Monitor corpus. A monitor case is used to monitor language changes. This is a regularly (or even permanently) updated corpus, with new texts being added as they are produced. Search results change because the corpus is always growing. The Trends case in Sketch Engine is an example of a monitor case, also known as a timestamp case.

Student corpus. A learner's corpus is a corpus of texts created by language learners. The corpus is used to study students' errors and problems in learning a foreign language. Sketch Engine allows a corpus of readers to annotate the error type and provides a custom interface to search for the error itself, the error fix, the error type, or a combination of the three.

Annotated corpus with error. This corpus contains texts created by language learners or translators. Errors are annotated and can be used to study the types of errors made by different groups of learners or translators.

Specialized corpus. A specialized corpus contains texts limited to one or more topics, domains, topics, etc. Such a corpus is used to learn how to use a specialized language. A user can create a specialized subbody from a generic body in Sketch Engine.

Multimedia corpus. A multimedia corpus contains text enhanced with audio or visual material or other types of multimedia content. For example, the conversational part of the British National Corpus in Sketch Engine contains links to relevant records that can be played from the Sketch Engine interface.

The significance of these types of synchronous corpora lies in their ability to facilitate various kinds of linguistic and cross-linguistic research, including translation studies, contrastive analysis, language teaching, natural language processing, cross-cultural communication, and more. They provide valuable insights into language use, variation, and translation processes across different languages and cultures. Additionally, they serve as essential resources for developing and evaluating language technologies and tools that operate across multiple languages.

In summary, synchronous corpora are indispensable resources for advancing our understanding of language, facilitating cross-linguistic research, improving language technologies, and promoting effective communication across linguistic boundaries. Their importance extends to various fields, including linguistics, translation studies, language education, NLP, and cross-cultural communication.

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. Mahmudova, D. (2023). CORPUS LINGUISTICS. B ACADEMIC RESEARCH IN MODERN SCIENCE (Т. 2, Выпуск 23, сс. 104–106). Zenodo. <https://doi.org/10.5281/zenodo.10025001>
7. Mahmudova, D. (2023). CORPORA AND LITERATURE. Current approaches and new research in modern sciences, 2(10), 63-64. Zenodo. <https://doi.org/10.5281/zenodo.10013233>
8. Karimov Rustam Abdurasulovich, Mengliev Bakhtiyor Rajabovich (2019). The Role of the Parallel Corpus in Linguistics, the Importance and the Possibilities of Interpretation. *International Journal of Engineering and Advanced Technology (IJEAT)* ISSN: 2249 – 8958, Volume-8, Issue-5S3 July 2019.
9. Davlatova M.H. Relation of lexical-semantic structure of verbs in the linguistic essence.// *IEJRD-International Multidisciplinary Journal*, № 5, 2021. – P. 58-62.
10. Davlatova M.H. Semantic implementation of resultative structures.// *JournalNX - A Multidisciplinary Peer Reviewed Journal*, № 6, 2021. – P. 291–294.
11. Davlatova M.H. Different aspects of resultative structures according to their linguistic essence.// *Academia Globe: Inderscience Research*, 2 (05), 2021. – P. 475–479.
12. Davlatova M.H. Aspectual and lexical-semantic classification of verbs. *JournalNX-A Multidisciplinary Peer Reviewed Journal*, 8(2). 2022. – P. 116-121.
13. Davlatova M.H. On the interrelationship of resultative and causative meanings.// *World Bulletin of Public Health* № 9. 2022. – P. 212-215.
14. 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>
15. Nematova Zebo Tursunboevna. (2023). The usage of authentic materials in the language classroom. *International Journal of Literature And Languages*, 3(02), 68–73. <https://doi.org/10.37547/ijll/Volume03Issue02-13>
16. Zebo Nematova Tursunboevna. (2023). AUTHENTIC MATERIALS AS AN ESSENTIAL PART OF FORMATION OF COMMUNICATIVE COMPETENCE. *Academia Repository*, 4(10), 217–224. Retrieved from <https://academiarepo.org/index.php/1/article/view/64>
17. Nematova, Z. (2022). Advantages of using videos in English lessons. *International Bulletin of Applied Science and Technology*, 2(9), 8-12.
18. 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.

19. 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>
20. 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>
21. Tursunboevna, N. Z. (2022). Various types of assessment in language teaching and learning. *Eurasian journal of social sciences, philosophy and culture*, 2(3), 140-145.
22. 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.
23. Habibova, M. N. (2021). The theme feminism in the epistolary novels in modern times. *ISJ Theoretical & Applied Science*, 11(103), 1101-1105.
24. 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.
25. Habibova, M. N. (2022). EVALUATIVE OBSERVATION ON DH LAWRENCE'S EPISTOLARY ACHIEVEMENT. *EURASIAN JOURNAL OF ACADEMIC RESEARCH*, 2 (4), 489–494.
26. 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.
27. 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.
28. 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>
29. 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>
30. 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>
31. 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>
32. 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>
33. 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>

34. 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>
35. 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.
36. 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>
37. 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>
38. 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>
39. 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–Retrieved from <https://academiarepo.org/index.php/1/article/view/96>
40. 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>
41. Habibova, M. N. (2023). Women and the Epistolary Novel. *American Journal of Science on Integration and Human Development (2993-2750)*, 1(9), 119-124.
42. Manzila Nuriddinovna Habibova. (2024). Pragmatics of Epistolary Technique Used in Biographical Works. *American Journal of Language, Literacy and Learning in STEM Education (2993-2769)*, 2(2), 135–139. Retrieved from <http://grnjournal.us/index.php/STEM/article/view/3061>