

REPUBLIC OF AZERBAIJAN

On the rights of manuscript

ABSTRACT

of the dissertation for the degree of Doctor of Philosophy

**LITERARY TEXT AS AN OBJECT OF
LINGUOSTATISTICAL STUDY**

Speciality: 5704.01 – Theory of Language

Field of science: Philology

Applicant: **Hamida Azizagha Khalilova**

BAKU – 2025

The dissertation work was performed at the Department of Artificial Intelligence and Computational Linguistics of the Institute of Linguistics named after Nasimi ANAS

Scientific supervisor: Doctor of Philological Sciences, Professor
Masud Ahmad Mahmudov

Official opponents: Doctor of Philological Sciences, Professor
Ilham Mikail Tahirov

Doctor of Philosophy Sciences in Philology,
Associate professor
Javanshir Khankishi Muradov

Doctor of Philosophy in Philology, Associate
professor
Sabina Ali Suleymanova

Dissertation Council ED 1.06 of the Supreme Attestation Commission under the President of the Republic of Azerbaijan operating at the Institute of Linguistics named after Nasimi of ANAS

Chairman of the Dissertation council: Doctor of Philological Sciences,
Professor
Nadir Balaoghlan Mammadli

Scientific secretary of the Dissertation council: Doctor of Philosophy in Philology,
Associate professor
Sevinj Yusif Mammadova

Chairman of the scientific seminar: Doctor of Philological Sciences,
Professor
Bahar Jumay Jafarova



Handwritten signature in blue ink, likely belonging to Bahar Jumay Jafarova, the Chairman of the scientific seminar.

INTRODUCTION

The research is devoted to the linguostatistical and corpus-based study of the literary texts of Nizami Ganjavi, M.A. Sabir, M. Fuzuli, and A.S. Pushkin. The research encompasses translations of Nizami Ganjavi's poems from the Khamsa into Azerbaijani and, in part, into Russian; M. Fuzuli's Divan and other poems written in Azerbaijani; the works of A.S. Pushkin (including The Captain's Daughter and others); as well as satires, ghazals, letters, short stories, feuilletons, and other writings by M.A. Sabir, as presented in the two-volume edition Hop-Hopname.

Relevance and the Degree of Scholarly Development. At present, corpus linguistics and linguostatistics represents one of the most rapidly developing fields of science. Although these disciplines emerged relatively recently, they have been developing at an accelerated pace. In the era of information technologies, their progress is irreversible, which has led to their recognition as independent scientific domains. Thanks to the methods of corpus linguistics and linguostatistics, it is possible to compile frequency dictionaries and concordances, create text corpora, and conduct comparative studies of text collections. Moreover, researchers can compare the language of an original text with its translation, identify the characteristic features of literary language, and more. By applying methods of linguostatistics and corpus linguistics, it also becomes possible to calculate the coefficient of lexical diversity (CLD), determine the number of tokens, and so forth.

The relevance of this study is also determined by the importance of creating Azerbaijani-language corpora, as no such corpora are currently in existence. The choice of material further emphasizes the topicality of this research – namely, the literary texts of M. Fuzuli, M.A. Sabir, and N. Ganjavi, which have not previously been the object of corpus-based and linguostatistical research. The choice of these authors is not accidental. Scholarly interest in the life and creative legacy of N. Ganjavi, M. Fuzuli, and M.A. Sabir remains

strong. N. Ganjavi's *Khamsa*, the work of a renowned poet and philosopher, has been extensively studied both in Azerbaijan and abroad. The significance of studying Nizami's legacy is substantiated by the agreement signed between the Nizami Ganjavi Center and Oxford University, with the participation and support of academician Nargiz Pashayeva, Vice President of the Azerbaijan National Academy of Sciences (ANAS) and Rector of the Moscow State University Branch in Baku.

The satires of the outstanding poet and founder of Azerbaijani satirical poetry, M.A. Sabir, remain relevant to this day. Sabir was a progressive thinker: through the complexity and urgency of the topics and issues raised in his satires, the great satirist was far ahead of his time. The language and style in which M.A. Sabir wrote are no less significant for our study. He wrote in a clear and accessible manner, combining colloquial speech with literary language, thereby creating a new poetic style. Thanks to this particular feature, Sabir's poetic language was easily understood by the people, which contributed to the wide dissemination of his works among the masses.

Another central figure of Azerbaijani literature is M. Fuzuli – an outstanding poet, philosopher, thinker, and translator. Beyond this, M. Fuzuli is regarded as the founder of the literary Azerbaijani language, which makes his works particularly important for this study.

A.S. Pushkin, in turn, is considered the founder of the modern Russian literary language. He created a national literary language by blending bookish language with spoken Russian, synthesizing Church Slavonicisms, Europeanisms, and elements of living Russian speech. The importance of studying Pushkin's creative legacy is noted by many scholars and cultural figures. The national leader Heydar Aliyev also spoke highly of Pushkin's works, which once again confirms the enduring relevance of Pushkin's personality and literary heritage. It should also be emphasized that the Corpus of Pushkin's Texts has been included in this research for comparative purposes. Studying the Pushkin Corpus clearly demonstrates how corpus research facilitates the work of the scholar and reveals its unlimited potential for various

types of research. It may also serve as a model, helping to identify the strengths and weaknesses of the Azerbaijani text corpora created in this study.

Significant contributions to corpus and linguostatistical research have been made by V.P. Zakharov, A.A. Polikarpov, O.V. Kukushkina, A.A. Akhundov, J.A. Rahmanov, M.A. Mahmudov, K. Valiyeva, R. Mahmudov, A.I. Izotov, A.V. Zubov, M.A. Kajarova, E.A. Kazak, A.V. Kasheyeva, M.V. Kopotev, and many others.

The ideas of A.A. Polikarpov, regarded as one of the pioneers of Russian corpus linguistics, underlie the development of corpus managers and approaches to corpus building at the Laboratory of General and Computer Lexicology and Lexicography at the Faculty of Philology of Moscow State University. Thanks to the efforts of A.A. Polikarpov and O.V. Kukushkina, author corpora of A.S. Pushkin, A.P. Chekhov, and others were created at Laboratory.

A.A. Akhundov's work significantly contributed to the development of linguostatistics in Azerbaijan. His article *From the Experience of Statistical Analysis of Vowels in the Azerbaijani Language* is considered one of the earliest works applying statistical and computer methods. Later, Akhundov expanded this line of research and authored *Mathematical Linguistics*. Furthermore, the development of linguostatistical research in Azerbaijan is associated with the names of M.A. Mahmudov, K. Valiyeva, M. Mammadova, among others. For instance, M.A. Mahmudov was one of the first in the country to conduct research in automatic language analysis, machine translation, and statistical lexicography. He systematically studied the problem of formal text analysis based on Azerbaijani material, developed algorithms for the automatic processing of formal signs at various language levels, and implemented them in computing machines.

In the field of corpus linguistics, the foundational principles were reflected in the works of J.A. Rahmanov, M.A. Mahmudov, K. Valiyeva, and others. The earliest corpus studies belong to J.A.

Rahmanov, whose research focused on problems of machine translation.

Thus, the works of Azerbaijani scholars have significantly advanced the study of corpus linguistics and linguostatistics in Azerbaijan, paving the way for the creation of text corpora. As a result, to date, an electronic dictionary corpus of the Azerbaijani language has been created. The above-mentioned studies are valuable and important primarily because they raised fundamental questions regarding emerging scientific fields. However, they do not address the significance of creating corpora of individual authors' texts or author-specific lexical dictionaries. Moreover, those works are largely theoretical, whereas the present research attempts to create the first Azerbaijani text corpora, analyze the concordances produced, and examine the distinctive features of authorial language.

Object and subject of the research. The object of the study is the literary texts of Nizami Ganjavi, M. Fuzuli, M.A. Sabir, and A.S. Pushkin.

The subject of the research is the application of corpus linguistics and linguostatistics methods to the study of the language of these authors, including the compilation of text corpora, the generation of concordances, the creation of frequency dictionaries, and the statistical analysis of lexical and grammatical units.

The purpose and tasks of the research. The purpose of the dissertation is to compile text corpora of Azerbaijani authors (poets and writers), as well as concordances of units contained in the studied literary texts and frequency lists (dictionaries) of these units. To achieve this purpose, the following tasks have been set in the course of the research:

- To describe and analyze various interpretations of the concept of “text”;
- To describe the categories and components of a text;
- To describe and identify the difference between the concepts of “text” and “discourse”;

- To describe and identify the difference between the concepts of “text” and “literary text”;
- To describe the origins and advantages of new interdisciplinary fields – computational linguistics, linguostatistics (quantitative linguistics), and corpus linguistics;
- To identify the scope of application of corpus linguistics and linguostatistical methods;
- To describe the text corpora of A.P. Chekhov and A.S. Pushkin, which can serve as models and help to determine the strengths and weaknesses of the corpora being created;
- To involve the “A.S. Pushkin Text Corpus” in this research in order to demonstrate the potential of text corpora for conducting diverse and large-scale studies;
- To analyze A.S. Pushkin’s novella *The Captain’s Daughter* and its translation into Azerbaijani: to compare the use of certain lexical units of “situational modality” in the original language and the translation in order to reveal differences;
- To create text corpora of Azerbaijani poets and writers;
- To compile frequency lists and concordances;
- To determine the number of tokens and word forms;
- To identify usage patterns;
- To reveal, through the linguostatistical study of a given author’s works, the role of low-frequency and high-frequency vocabulary.

Research methods. The study employed the descriptive method, the componential analysis method, corpus linguistics methods, lexicographic analysis, statistical methods for text processing, and others.

The main provisions submitted for defense:

- As a result of the analysis of authorial literary corpora, specific structural features of literary texts were identified, expressed in the presence of interstylistic inclusions and a high density of stylistically marked vocabulary, which allows for

clarifying their status in the typology of texts as a multi-level linguistic system.

- Based on the corpus analysis of texts of various genres, distinctive features of literary texts were revealed, including the predominance of expressive vocabulary, extensive use of stylistic figures, as well as the tendency to include elements of colloquial and journalistic vocabulary, confirming their functional hybridity.

- For the first time, on the basis of corpus analysis of literary texts in Russian and Azerbaijani, the distinction between “text” and “discourse” was empirically confirmed, where the text is recorded as the result of speech activity, and discourse is understood as a process within a communicative situation; it was justified that only text is subject to statistical and frequency-based analysis.

- Frequency dictionaries and lists (of nouns, verbs, adjectives) based on corpus data, including translations into Azerbaijani, were developed and implemented in the practice of teaching Russian as a foreign language, which enhanced the effectiveness of vocabulary acquisition among Azerbaijani students.

- Quantitative patterns of the use of different parts of speech and certain syntactic constructions in literary texts were identified based on the analysis of the corpora of Pushkin, Sabir, and Fuzuli, allowing for the formalization of authors’ stylistic features and contributing to more accurate automatic text annotation. The corpora provide a comprehensive picture of the usage of linguistic units and allow for determining their degree of frequency.

- For the first time, the necessity of creating literary text corpora and including them as subcorpora in the national corpus of Azerbaijan was substantiated. These subcorpora, based on literary works and oral folk creativity, contain precedent phenomena, culturally specific realities, and authorial stylistic

models, which expands the analytical potential of the corpus and enables its interdisciplinary application.

- The A.S. Pushkin text corpus was used for comparative purposes as a model of a mature authorial corpus, confirming the high efficiency of corpus linguistics methods in linguistic research. The application of these methods to the Pushkin corpus demonstrated significant simplification of analytical work, the identification of frequency, stylistic, and structural features of texts. Furthermore, this corpus served as a reference in the creation and testing of authorial corpora of Azerbaijani writers and poets, allowing the identification of both strengths and problematic aspects of the developed databases.

- The created authorial corpora of literary texts made it possible to identify the features of individual authors' linguistic personality, including lexico-semantic preferences and individual stylistic techniques, confirming their value as a tool for linguo-personological analysis. Since an author's language is fully realized as an integrated system within their texts, such corpora are particularly valuable for the formation of a national corpus linguistics school, especially during the early stages of Azerbaijani corpus linguistics, where systematic description of linguistic personality carries both scientific and practical significance.

- Based on the created authorial corpora, frequency dictionaries of individual authors' language (M.A. Sabir, M. Fuzuli, N. Ganjavi) were compiled, reflecting characteristic thematic fields and predominant parts of speech, ensuring a deeper level of lexico-stylistic analysis.

- Analysis of frequency lists and concordances of originals and translations of classical poetry and prose (based on Nizami *Ganjavi's* "Leyli and Majnun" and *A.S. Pushkin's* "The Captain's Daughter") revealed significant differences in lexical content. It was established that translation involves substitutions affecting semantic completeness, stylistic coloring, and the

frequency of certain linguistic units, reflecting features of the text adaptation process and highlighting characteristic traits of the translator's style.

- Distinctive features of the language of A.S. Pushkin, M.A. Sabir, and M. Fuzuli were identified based on the comparison of their authorial corpora with data from contemporary Russian and Azerbaijani literary language. Lexico-grammatical features and frequency deviations were revealed, reflecting the historical-cultural context and individual style of the authors.

- For the first time, Azerbaijani language text corpora were created and described (literary texts): the corpus of M.A. Sabir, the translations of N. Ganjavi's "Khamsa" into Azerbaijani, as well as the poetic corpus of M. Fuzuli. Frequency dictionaries and concordances were developed based on these corpora, representing a significant contribution to the development of Azerbaijani corpus linguistics and expanding opportunities for lexicographic and linguistic-stylistic research, as well as serving as a foundation for new scientific investigations in the fields of language, translation, and beyond.

- Corpus and linguostatistical analysis of authorial corpora revealed the role of low-frequency vocabulary as an indicator of individual style. It was established that units with minimal frequency in the corpus generally represent individual-author usage, including unique nominations, neologisms, or idiolectal expressions.

Scientific novelty of the research. The scientific novelty of the dissertation is substantiated, firstly, by the attempt to define the special status of the literary text and its significance in conducting corpus and linguostatistical studies. In addition, the scientific novelty of this dissertation lies in the application of corpus linguistics methods and linguostatistical methods within a specific study, as well as in applying these methods to literary texts presented in Azerbaijani and partially in Russian. This approach allows for identifying the language

features of individual authors, the specifics of the literary text in the original language and in translation, and the characteristic features of the author's language. The primary significance of the scientific novelty is justified by the creation of text corpora. It should be particularly noted that such studies are innovative, as research in Azerbaijan in the fields of corpus linguistics and linguostatistics is still limited. More specifically, the Azerbaijani language is not yet represented by text corpora.

Thus, this research contributes to the formation of corpora of texts by Azerbaijani authors (writers and poets) and of the Azerbaijani language, which will facilitate the creation of the National Corpus of the Azerbaijani language. That is, the creation of corpora of individual Azerbaijani authors is equally important for the development of the National Corpus. National corpora include all styles and genres. For instance, the National Corpus also contains literary works, since it is literary texts and oral folk creativity that carry cultural significance and include precedent phenomena.

The compilation of frequency dictionaries is valuable because their application has exceptional significance in automated data processing, machine translation, summarization, publishing, forensic linguistics, and the identification of authors of anonymous texts and documents.

It should also be noted that the creation of author-specific corpora has an innovative character. For instance, in Russia, corpora have been created for the texts of Pushkin and Chekhov, while in England, a corpus has been developed for Shakespeare's texts. Authorial corpora have been particularly actively developed in recent times, as noted by V.P. Zakharov in his work. Such corpora allow researchers to investigate the characteristics of an individual linguistic personality. Specifically, within the texts of a single author, language is realized in its entirety as a holistic system. Author-specific corpora can be supplemented with dictionaries, which enable researchers to locate the necessary contexts. In addition to author-specific corpora, depending on the goals and tasks set, a researcher can compile new

corpora that meet specific requirements, thus independently becoming the compiler of a particular corpus. For example, a corpus processing scientific texts has been initiated. Therefore, this work represents the first creation of author-specific corpora in Azerbaijani scholarship.

Theoretical and practical significance of the research. The theoretical significance of this dissertation lies in clarifying a number of concepts, expanding scientific understanding of the text, and defining the special status of the literary text. Furthermore, the theoretical significance includes substantiating the development of corpus linguistics, highlighting the advantages of applying corpus and linguostatistical methods, and determining the importance of creating corpora of individual authors. The work also contributes to the field of theoretical and practical lexicography, frequency dictionaries, and corpora.

The practical significance of the work consists in the fact that the obtained results can be used in courses on lexicology, text linguistics, and other practical classes. In addition, the created frequency dictionaries and concordances can be applied for educational and lexicographic purposes. The created text corpora themselves can also be used for teaching and lexicographic work. Moreover, this research will contribute to the development of corpus linguistics and linguostatistics in Azerbaijan and to the creation of text corpora.

Approbation and application of the research. The main points of the work were presented at international and national conferences and symposia. Articles, theses, and conference materials have been published on the topic of this dissertation.

The organization where the dissertation was carried out. The research was carried out at the Department of Artificial Intelligence and Computational Linguistics of the Institute of Linguistics named after Nasimi, Azerbaijan National Academy of Sciences.

Structure and volume of the dissertation, in characters, with the volume of each section specified separately. The

dissertation consists of an introduction, table of contents, three chapters, conclusions, a list of references (bibliography), appendices, and a list of abbreviations. The introduction comprises 11 pages, 20,174 characters; Chapter I – 39 pages, 68,537 characters; Chapter II – 30 pages, 53,897 characters; Chapter III – 61 pages, 93, 492 characters; Conclusions – 2 pages, 3, 129 characters. The total number of characters in this dissertation is 239,229.

MAIN CONTENT OF THE DISSERTATION

The introduction substantiates the relevance and degree of development of the research topic, presents the subject and object of the study, defines the aim and objectives, describes the research methods, highlights the main findings submitted for defense, outlines the scientific novelty of the research, the scientific-theoretical and scientific-practical significance of the work, provides information on the research materials, and indicates the testing and application of the main provisions and results of the study.

Chapter I. “The Concepts of ‘Text’ and ‘Literary Text’. The Special Status of the Literary Text” consists of five subsections. In the subsection **“Text as an Object of Philology. Different Approaches to the Problematic”**, the origins of the emergence and study of the concept of “text” in Russian, Azerbaijani, and international scholarship are examined.

Text is one of the most important and fundamental terms in linguistics. At the same time, it represents the primary subject of philology. As Yu. M. Lotman notes, text is *“undoubtedly one of the most widely used terms in the humanities.”*¹ *“Text (from Latin Textus – fabric, weave, connection) is a sequence of sign units united by semantic coherence, the main properties of which are cohesion and integrity”*², meaning that the sign units within a text must share a

¹ Лотман, Ю.М. Семиосфера / Ю.М. Лотман. – СПб.: Искусство-СПб, – 2000. – 704 с.

² Лингвистический энциклопедический словарь: [Электронный ресурс] / URL: <http://tapemark.narod.ru/les/index.html>

common semantic component on the basis of which they are combined to form a coherent text. Therefore, not every set of sign units constitutes a text; cohesion and integrity are required. The etymology of the word “text” reflects one of its fundamental principles – connection, i.e., its semantic coherence.

Text is a system, and due to its multifaceted nature and complexity, different approaches to its interpretation exist in the scientific literature.

L.G. Babenko notes that *“text is a verbal speech product in which all linguistic units are realized.”*³

According to Z.Ya. Turayeva, text is *“an ordered set of sentences connected by various types of lexical, logical, and grammatical links, capable of conveying information organized and directed in a specific manner. The text is a complex whole functioning as a structural-semantic unity.”*⁴ The researcher also considers text to be a complex and multifaceted system and emphasizes the structural and semantic unity of all units that constitute the text.

Q. Kazimov emphasizes the importance of theme–rheme relations in text construction. According to him, new information, which constitutes the communicative basis of a sentence, leads to the deepening and development of thought, as a result of which utterances, when combined, determine thematic development, i.e., the creation of a text. The researcher further summarizes his idea: a text is a unity of sentences, interconnected on the basis of a common purpose and forming a structural–semantic integrity.⁵

³ Бабенко, Л.Г. Лингвистический анализ художественного текста [Текст] / Л.Г. Бабенко, И.Е. Васильев, Ю.В. Казаров. – Екатеринбург: Изд-во Урал, унта, – 2000. – 534 с.

⁴ Тураева, З.Я. Лингвистика текста (Текст: Структура и семантика): [Электронный ресурс] / URL: https://www.phantastike.com/linguistics/lingvist_teksta/djvu/view/

⁵ Kazimov, Q.Ş. Müasir Azərbaycan dili. Sintaksis / Q.S. Kazimov. – Bakı: Ünsiyət, – 2000. – 496 s.

According to T.M. Dridze, however, a text is a holistic communicative unit that represents a complex sign. At the same time, the elements within this system share a unified structure and are subordinated to a certain hierarchy.⁶

N.B. Mammadli points out that text represents the unity of structure, composition, grammatical-communicative completeness, and content. According to other definitions, text is considered a “multilayered system” and a complex whole, serving to create structural-semantic unity.⁷

N.F. Alefirenko notes that text is “the totality of what is written and spoken in a given language.” Thus, a text is formed by various linguistic units and represents a multi-level and complex system.⁸

Therefore, more precise definitions are attributed to Z.Ya. Turayeva, Q. Kazimov, and N.B. Mammadli. N.F. Alefirenko’s work is also significant, as he distinguishes between written and oral forms of text. In this chapter, approaches to the study of text were examined. It was determined that the main aspect is the linguistic or linguo-centric aspect.

In the subsection “*Components and Categories of Text*”, the components, elements, and categories of text are analyzed. It is established that the sentence constitutes the fundamental textual unit. Depending on their approaches, researchers define the sentence in different ways: as a “*minimal communicative unit*”⁹ (Groshenkova), a

⁶ Дридзе, Т.М. Текстовая деятельность в структуре социальной коммуникации: проблемы семиосоциопсихологии [Текст] / Т.М. Дридзе. – М.: Наука, – 1984. – 268 с.

⁷ Mətn dilçiliyi problemləri / N.B. Məmmədli. – Bakı: Zərdabi Nəşr MMC, – 2023. – 260 s.

⁸ Алефиренко, Н.Ф.: [Электронный ресурс] / URL: <https://core.ac.uk/download/pdf/151213879.pdf>

⁹ Groshenkova, V.A. Обучение чтению в специальной (коррекционной) школе: [Электронный ресурс] / URL: <https://cito-web.ysspu.org/link1/metod/met143/node3.html>

“small predicative link in the chain of thought”¹⁰ (N.B. Mamedli), among others. The primary categories of text are cohesion and integrity.

The subsection “**Text and Discourse**” of Chapter I addresses the distinction between the concepts of “text” and “discourse,” a differentiation that determines the object of research. This subsection describes the origins of the concept of “discourse” and presents different scholarly perspectives on the matter.

Discourse is defined as “*a coherent text in conjunction with extralinguistic factors – pragmatic, socio-cultural, psychological, and others; a text considered in the aspect of events; speech viewed as a purposeful social action, as a component participating in human interaction and cognitive mechanisms.*”¹¹ Discourse is not merely text, but speech “*immersed in life*”¹², i.e., it acts as a complex and coherent unity of speech and situation.

In the subsection “**Literary Text. The Linguistic Nature of the Literary Text**”, various approaches to the interpretation of the literary text are discussed, as well as similarities and differences between the concepts of “text” and “literary text.”

A literary text is an independent, original product of literary speech, created in a particular language and representing a complex, integrated system. It reflects the characteristics of the people and the language of the author and holds a special status due to the use of the literary-artistic style – the richest of functional styles, encompassing elements of all others. This feature makes the literary text a valuable object for linguostatistical, corpus-based, and other types of research.

¹⁰ Mətn dilçiliyi problemləri / N.B. Məmmədli. – Bakı: Zərdabi Nəşr MMC, – 2023. – 260 s.

¹¹ Клокова, А.Г. Понятие дискурса в отечественном и зарубежном языкознании: [Электронный ресурс] / URL: <https://elib.bsu.by/handle/123456789/199391>

¹² Лингвистический энциклопедический словарь: [Электронный ресурс] / URL: <http://tapemark.narod.ru/les/index.html>

Therefore, “text” is a generic concept, and the literary text is one of its varieties.

The subsection “*Genres of the Literary Text*” in Chapter I examines literary types and genres. Since the object of the research is literary texts and the subject is their structural units, concordances, and corpora, the study of genres is essential for understanding their structural unity and the specific features of the use of linguistic units.

Chapter II. “Corpus Linguistics and Linguostatistics: The Study of Literary Texts and Language through Text Corpora” also consists of five subsections. In the subsection “**Computational Linguistics. In the subsection “Computational Linguistics. The Formation of the Discipline”**, a new interdisciplinary field, computational linguistics, is examined.

With the advancement of technologies, new scientific fields began to emerge, one of which is computer linguistics. The concept of “computer linguistics” is understood broadly. Computer linguistics can be considered a new direction that emerged at the intersection of other scientific disciplines. All issues related to the application and use of computers in linguistics are considered objects of study within computer linguistics.

The emergence and development of computer linguistics as such are directly connected with the advent of electronic computing machines. In the 1950s, a kind of “revolution in scientific thinking” occurred – a fundamental transformation in the field of scientific thought. The appearance of electronic computing machines predetermined the development of science, generated new terms and fields, and demonstrated that interdisciplinary research yields more precise results, which became the impetus for the emergence of computer linguistics.

It should be noted that there is no unanimous opinion in the scientific community regarding the definitive name of this field. This field has become known under various names, such as “mathematical linguistics,” “structural linguistics,” “cybernetic linguistics,” “engineering linguistics,” “corpus linguistics,” and others.

K.K. Boyarsky also notes that the emerging new science was referred to by various names in its early stages, and eventually the designation “computational linguistics” became firmly established.¹³

According to K. Valieva, computational linguistics is a branch of applied linguistics that studies the linguistic foundations of informatics through computer programs, language, and thought. It also examines linguistic modeling, thereby creating and refining a common language, and sets tasks that are considered among the main problems of artificial intelligence, namely: addressing the interface problem of natural languages and the modeling of information systems.¹⁴

In our view, the term “computer linguistics,” used by M.A. Mahmudov, K.K. Boyarsky, and K. Valieva, encompasses various aspects of this new scientific field and meets the demands of researchers from different disciplines. Some branches may evolve into independent sciences, as is the case with corpus linguistics, which emerged from computer linguistics.

The second subsection of Chapter II, “*Corpus Linguistics*”, describes the origins and fundamental concepts of corpus linguistics, defines and characterizes its terminological framework, and discusses the methods of corpus linguistics that are essential for text analysis.

Corpus linguistics is a branch of linguistics that deals with the creation of text corpora, employing methods of computer linguistics and linguostatistics.

The terminological apparatus of corpus linguistics includes: text corpus, concordance, corpus manager, and others.

A corpus (linguistic or language corpus) is an electronic-information system comprising a collection of texts compiled according to specific rules. “*A linguistic or language text corpus is understood as a large, electronically represented, unified, structured,*

¹³ Боярский, К.К. Введение в компьютерную лингвистику / К.К. Боярский. Учебное пособие. – СПб: НИУ ИТМО, – 2013. – 72 с.

¹⁴ Vəliyeva, K. Kompüter dilçiliyinə giriş / K. Vəliyeva. – Bakı: Elm və Təhsil, – 2018, – 384 s.

annotated, and philologically competent array of linguistic data intended for solving specific linguistic tasks."¹⁵

The methods of corpus linguistics date back to the 13th century: "*pre-electronic corpora began to form from the 12th century up to the 1960s*"¹⁶, and the very term "corpus" entered linguistic usage only by the end of the pre-electronic era.

The main function of corpora is to simplify research and language study, identify rules and tendencies, investigate the features of different types of texts based on quantitative data, and obtain a comprehensive picture of unit usage.

A text corpus is also the object that corpus linguistics strives to create. In fact, a text corpus functions, "*on the one hand, as its object, i.e., as the primary linguistic material for corpus linguistics and for other linguistic disciplines*"¹⁷; on the other hand, a text corpus is the result of the activity of corpus linguistics. This duality emphasizes the following: corpus linguistics is aimed at creating a corpus, but it also involves the use of text corpora to solve specific tasks set before the researcher.

For corpus linguistics, terms such as annotation, meta-annotation, lemmatization, and tokenization are important. Annotation is additional linguistic and metalinguistic information about texts and word usages.¹⁸ "*Meta-annotation provides significant extralinguistic information about the text (genre, author, domain of use, subject matter, etc.)*"¹⁹ From the definition of meta-annotation, it follows that it supplies important extralinguistic information about a text, such as

¹⁵ Захаров, В.П. Корпусная лингвистика: Учебно-метод. пособие. – СПб., 2005. – 48 с.

¹⁶ Солнышкина, М.И. История развития корпусной лингвистики (на примере англоязычных корпусов), URL: <https://cyberleninka.ru/article/n/istoriya-razvitiya-korpusnoy-lingvistiki-na-primere-angloyazychnyh-korpusov>

¹⁷ Бойкова, С.Н. Введение в корпусную лингвистику: [Электронный ресурс] / URL: <https://core.ac.uk/download/pdf/75999967.pdf>

¹⁸ Лозовая, А.А.: [Электронный ресурс] / URL: http://elib.rshu.ru/files_books/pdf/rid_647a25806a7e42d5b628f117f8c5f6e1.pdf

¹⁹ Там же.

its genre, author, topic, and other aspects that help to understand the context and features of the text beyond its linguistic structure. “*Lemmatization is the process of forming the base (dictionary) form of a word from its other word forms*”²⁰, while tokenization involves dividing the text into individual units, without which further automatic processing is impossible²¹”; that is, tokenization is the segmentation of the text into tokens and word forms.²²

It is also worth noting that research in corpus linguistics is underrepresented in Azerbaijan. The Azerbaijani language lacks text corpora, although a corpus of electronic dictionaries²³ has been created. The absence of corpora limits opportunities for deeper language analysis based on real texts, slowing the development of linguistic research and applications such as automatic text analysis, machine translation, and other technologies requiring a high-quality corpus. The creation of full-fledged text corpora in Azerbaijani will open new perspectives for linguistic research and contribute to the further development of corpus linguistics in the country.

A special type of corpora is authorial corpora. They have been actively created in recent years, as they allow the study of the features of an individual linguistic personality. In the texts of a single individual, language is fully realized as an integrated system.

The advantages of corpus linguistics methods should be emphasized: they allow the resolution of linguistic, sociological, and historical tasks, facilitate comparative analysis of literary works by different authors and their translations, reveal features of authorial style, study the frequency of units, and perform search and selection

²⁰Токенизация: [Электронный ресурс] / URL: https://ozlib.com/1010207/literatura/osnovnye_protsedury_obrabotki_estestvennog_o_yazyka_tokenizatsiya_lemmatizatsiya_stemming_parsing#google_vignette

²¹ Там же

²²Гречачин, В.А. Понятие токенизации в корпусной лингвистике: [Электронный ресурс] / URL: https://www.elibrary.ru/download/elibrary_25307213_89263284.pdf

²³ Azərbaycan dilinin elektron lüğətlər korpusu: [Электронный ресурс] / URL: <https://korpus.azerbaycandili.az/>

of words, word forms, grammatical categories, and collocations considering morphological characteristics and contextual parameters.

The third subsection, “***National Corpora***”, addresses the concept of the national corpus, its significance, and provides descriptions of several of the most representative corpora. Of particular importance is the discussion of the relevance of creating a corpus of texts by Azerbaijani writers and poets, as well as a National Corpus of the Azerbaijani Language itself. The creation of the National Corpus of the Azerbaijani Language is regarded as a strategically important issue, and at present this matter is especially urgent.

The subsection “***Information on the Laboratory of General and Computational Lexicology and Lexicography (LGCLL) of the Faculty of Philology, Lomonosov Moscow State University***” presents an overview of the Laboratory of General and Computational Lexicology and Lexicography (LGCLL) of the Faculty of Philology at Lomonosov Moscow State University, where research is conducted on theoretical developments and the study of natural language. The subsection provides information about the research carried out at LGCLL, as well as on corpus managers. Since the materials of this study include A.S. Pushkin’s *The Captain’s Daughter* and the *Corpus of A.S. Pushkin’s Texts*, their consideration was both essential and justified.

The subsection “***Linguostatistics***” of Chapter II explores a new domain – linguostatistics, or quantitative linguistics.

Linguostatistics (linguistic statistics) is defined as: “1) in a broad sense: the application of statistical methods in linguistics; 2) in a narrow sense: the study of certain mathematical problems related to linguistic material, primarily concerning the types of statistical distributions of linguistic units in a text.”²⁴

²⁴ Лингвистический энциклопедический словарь [Электронный ресурс] / URL: <http://tapemark.narod.ru/les/index.html>

Linguostatistical methods are particularly productive in the study of literary texts. They allow researchers to evaluate the nature of language use, determine the diversity of speech, the frequency of words and word forms, count tokens, and identify linguistic features of the author and specifics of the literary text in the original and translation. This facilitates the creation of corpora for specific authors' languages.

Chapter III. “The Scope of Application of Corpus Linguistics Methods and Technologies. Creation of Corpora of Literary Texts by Azerbaijani Authors (Writers and Poets)” consists of eight subsections. The first subsection, **“Application of Corpus Linguistics and Linguostatistics Methods in the Study of Language on the Basis of Literary Texts”**, outlines the scope of application of corpus linguistics methods and technologies. It is noted that the methods and technologies of corpus linguistics and linguostatistics are employed not only by researchers, but also by educators and students in the process of learning a new language.

Thus, thanks to the application of corpus linguistic methods and technologies, it has been possible to create, in a short period, a dictionary of the most frequent words of the Russian language with translations into Azerbaijani, as well as lists of the most frequent nouns, verbs, adjectives, and other parts of speech with their Azerbaijani translations. This approach and method significantly increase the efficiency of Russian as a Foreign Language (RFL) acquisition and contribute to a higher level of language proficiency. It should be noted that the contexts extracted from the National Corpus of the Russian Language (NCRL) pertain to literary texts; that is, the contexts (materials) were drawn from literary texts included in the corpus, which is also significant for the present study.

The second subsection, **“Linguostatistical Study Based on A.S. Pushkin’s The Captain’s Daughter”**, is devoted to a comparative corpus-based analysis of lexical means of situational modality that express the subject’s attitude toward an action (e.g., “can,” “must,” “want”) in A.S. Pushkin’s novella *The Captain’s*

Daughter and its translation into Azerbaijani, with the aim of identifying differences between the original and the translation.

The subsection **“Results of the Analysis of Certain Nouns and Core Predicates Denoting the Microfields of the Field ‘Situational Modality’ Based on the A.S. Pushkin Text Corpus”** in Chapter III presents the outcomes of research conducted with the *A.S. Pushkin Text Corpus* developed at LGCLL. It should be emphasized that the *A.S. Pushkin Text Corpus* is employed in this study as a model: research carried out on its material will in the future allow for the evaluation of the strengths and weaknesses of newly created corpora and will demonstrate their research potential.

Significant findings of the subsection **“Comparative Analysis of Contemporary and Pushkin’s Usage of Certain Markers of Possibility Based on Corpus Data”** include the identification of specific features of Pushkin’s language, as well as certain divergences from contemporary Standard Russian.

The subsection **“Creation of Corpora of Azerbaijani Writers’ and Poets’ Texts”** in Chapter III emphasizes the importance of creating corpora of individual authors’ texts, highlighting the significance of developing corpora of Azerbaijani authors, which is equally important for the formation of the National Corpus of the Azerbaijani Language.

It should be noted that in the scholarly world, corpora of individual authors already exist. For example, in Russian studies, the *A.S. Pushkin Text Corpus* and *A.P. Chekhov’s Literary Texts*; in England, the *Shakespeare Text Corpus*, among others. Depending on research objectives, a scholar may also create a specialized corpus, as demonstrated in the article by E. Callegari²⁵, which describes the process of building a corpus for the automatic analysis of scientific articles.

²⁵ Callegari, Elena Xhura, Desara A corpus for Automatic Article Analysis: [Электронный ресурс] / URL: <https://aclanthology.org/2022.clib-1.2.pdf>

The creation of corpora of Azerbaijani authors' texts is also important for the formation of the National Corpus, which encompasses all styles and genres, including literary works and folklore, as carriers of cultural significance and precedent phenomena.

It should also be noted that compiling frequency dictionaries is valuable because their application is crucial in automatic data processing, machine translation, summarization, printing, forensic linguistics, and in identifying the authors of anonymous texts and records.

The subsection “**Creation of the Corpus of Nizami Ganjavi’s Khamsa**” describes the process of constructing the *Corpus of Nizami Ganjavi’s Khamsa* in its Azerbaijani translation. As a result, corpus-based and linguostatistical studies were conducted, concordances were compiled both for individual works and for the entire corpus, frequency dictionaries were prepared, as well as a dictionary of ethnospecific vocabulary based on the translation of the poem *Leyli and Majnun*. At present, this dictionary of ethnospecific vocabulary contains 588 entries.²⁶

Quantitative indicators for the *Khamsa* corpus have also been determined: 290,476 tokens and 45,453 wordforms, with the most frequent lexical items identified. The ten most frequent units in the corpus are: *bir* (one; an emphatic particle *also*) – 8,193 occurrences, *bu* (this) – 4,942, *o* (he, she, it; that) – 4,473, *hər* (every, each) – 2,617, *ki* (emphatic particle; conjunction *that*) – 2,079, *nə* (*neither...nor*; particle; interrogative pronoun *what*) – 1,830, *də* (conjunction *and*; emphatic particle) – 1,802, *da* (conjunction *and*; emphatic particle) – 1,489, *sən* (you) – 1,488, *mən* (I) – 1,341.

²⁶ Халилова, Г.А. «Развитие азербайджанской корпусной лингвистики. Создание корпуса текстов азербайджанских авторов (поэтов и писателей) и национального корпуса азербайджанского языка». Вестник Дагестанского государственного университета. Серия2: Гуманитарные науки 2024 / 2. [Электронный ресурс] / URL: <https://vestnik.dgu.ru/pol.aspx?razdel=2&artId=4670>

The functioning of Arabic borrowings was examined, and discrepancies in the frequency of words such as *Allah*, *yaradan*, *tanrı*, etc., were identified. In addition, a comparative analysis was carried out of the Azerbaijani and Russian translations of the poem *Leyli and Majnun*, including the frequency of personal pronouns, predicates, religious vocabulary, and ethnospecific lexicon. The functioning of Arabic borrowings was examined, and differences were identified in the frequency of words such as *Allah*, *yaradan* (creator), *tanrı* (god), among others.

The subsection “**Creation of the Corpus of M.A. Sabir’s Texts**” in Chapter III describes the compilation of the *M.A. Sabir Text Corpus*, which currently includes: 274 satires, 12 letters, 4 stories and feuilletons, 5 translations, 22 children’s poems and stories, 11 miscellaneous poems, 11 ghazals, 38 articles and correspondences, and 37 tahmises. Frequency dictionaries and concordances have been compiled, and the most frequent lexical items have been identified both across the entire corpus and in individual works. The total number of word tokens (occurrences) and word types (forms) in Sabir’s works is: 67,138 tokens and 21,361 types. The most frequent among them are: *bir* (one; an emphatic particle *also*) – 1,371 occurrences, *bu* (this) – 1,037, *ki* (emphatic particle; conjunction *that*) – 861, *nə* – 632, *və* (and) – 504, *ilə* (with) – 476, *də* – 475, *hər* (every, each) – 459, *da* – 304, *var* (there is/are) – 304, *mən* (I) – 223. Certain linguistic peculiarities of Sabir’s works have also been identified. For instance, the word *школа* (school) is used by the great satirist in the following forms: *uşqol* – 1, *uşqola* – 3, *uşqolada* – 1, *uşqoladan* – 1, *uşqolalar* – 1, *uşqolaların* – 1, *uşqolanın* – 2, *uşqolasına* – 1, and many others.

The final subsection, “**Creation of the Poetic Corpus of M. Fuzuli’s Texts**” in Chapter III, discusses the creation of a poetic corpus of M. Fuzuli’s works, which currently includes 305 ghazals and the poems *Layli and Majnun*, *Opium and Wine*, *The Dispute of Fruits*, and *Complaint*. Frequency dictionaries and concordances have been compiled, the specific features of Fuzuli’s language have been identified, and quantitative correlations of tokens and word forms have

been determined both for individual works and for the corpus as a whole. The principal outcome of this work is the creation of the *Poetic Corpus of M. Fuzuli's Ghazals*, which records 31,372 tokens and 9,289 wordforms. The study also revealed notable features of Fuzuli's poetic language: extensive use of the auxiliary verbs *etmək* (to do), *eyləmək* (to do), and *olmaq* (to be); the use of pronouns in connective functions; the formation of complex words by combining certain Arabisms with Azerbaijani verbs; and many others. For example, the word form *ol* occurs 324 times in the ghazals, making it the fourth most frequent word. The word form *olmaq* occurs 16 times, *olsam* – 12, *olsun* – 13, *olmuş* – 45, and others. The form *et* appears 90 times in the ghazals of the *Divan*, *etdi* – 55, *etdim* – 44, *etdin* – 29, *etmək* – 17; *eylər* – 90, *eylə* – 60, *eyləmək* – 17, *eylədi* – 15, *eylədin* – 13 times, etc.

The **Conclusion** presents the main findings, results, and generalizations derived from the dissertation.

This study has been carried out with the purpose of compiling corpora of Azerbaijani authors' texts (both poets and writers), concordances of the units contained in the literary texts under investigation, as well as frequency lists and dictionaries of these units.

The main content and provisions of the dissertation were reflected in the following theses and articles of the author:

1. The concept of text has been clarified as a multi-level system; the key categories of text (cohesion, coherence) have been identified, and the notions of “text” and “discourse” have been distinguished.
2. The special status of the literary text has been established, characterized by its rich stylistic potential, which enables large-scale linguostatistical and corpus-based studies.
3. The development of corpus linguistics and linguostatistics as independent scientific fields has been traced; it has been demonstrated that text corpora facilitate language research and reveal patterns and tendencies of unit usage.

4. The role of linguostatistics and corpus linguistics in the study of language has been substantiated; the necessity of creating a National Corpus of the Azerbaijani Language as a strategic project has also been justified.

5. The effectiveness of corpus methods has been confirmed on the material of A. S. Pushkin's text corpus, which opens up prospects for their application in Azerbaijani philology.

6. The applied significance of corpus technologies has been determined, including the compilation of frequency dictionaries and concordances, and their use in automatic text processing, machine translation, authorship attribution, and foreign language teaching.

7. Within the framework of applying corpus linguistic methods in foreign language teaching, a dictionary of the most frequently used Russian words with Azerbaijani translations has been compiled. Frequency lists of nouns, verbs, adjectives, and other parts of speech with corresponding translations have also been prepared, which ensures their use in the educational process and in the development of teaching materials.

8. Corpora of texts by Nizami Ganjavi, M. A. Sabir, and M. Fuzuli have been created; frequency lists and concordances have been compiled, and their frequency as well as lexico-grammatical features have been identified, including ethnospecific vocabulary, Arabic-Persian borrowings, and individual authorial usages.

9. It has been revealed that low-frequency vocabulary serves as an indicator of individual style; it has been established that rarely occurring units include unique words, neologisms, and idiolectisms.

10. A linguostatistical analysis has been conducted of A. S. Pushkin's novella *The Captain's Daughter* and its translation into Azerbaijani, as well as of the translations of Nizami Ganjavi's *Layli and Majnun* into Russian and Azerbaijani; frequency indices of the main predicates of situational modality and their differences in the original and in translation have been identified.

11. The established characteristic linguistic features of M. A. Sabir and M. Fuzuli confirm the impact of their works on the development of the Azerbaijani literary language.

12. This study is pioneering in the field of corpus linguistics in Azerbaijan: for the first time, corpora and frequency dictionaries of texts by Azerbaijani authors have been systematically compiled, laying the foundation for the creation of the National Corpus of the Azerbaijani Language.

List of published scientific works on the topic of the dissertation:

1. Ситуативная модальность и лексические средства ее выражения как объект корпусных исследований: [Электронный ресурс] / URL: <https://www.stephanos.ru/izd/2019/2019-36.pdf>

2. Компьютерная лингвистика и применение лингвостатистических методов при исследовании художественных текстов // – Ankara: Bilimsel arařtırmalar, – 2021. – S. 33-35.

3. Особенности этноспецифической лексики на материале поэмы «Лейли и Меджнун» Низами Гянджеви: [Электронный ресурс] / – 2023. URL: <https://bsu-university.edu.az/storage/konfranslar/azerbaycansunasligin-aktual-meseleleri-2023-2-cild-2.pdf>

4. Bədii mətn linqvostatistik tədqiqat obyektı kimi // – Bakı: II International congress of modern studies in social sciences and humanities «In Memory of Victory Day and Martyrs» konfrans materialları, – 2021. Vol. 1, – s. 248-255;

5. Azərbaycan yazıçılarının bədii mətn korpuslarının yaradılmasında süni intellektin metod və texnologiyaları: [Электронный ресурс] / – Bakı: “Azərbaycan və dünya ədəbiyyatşünaslığı: problemlər və perspektivlər” mövzusunda Respublika Konfransının materialları. URL: <https://baau.edu.az/az/announcement/azerbaycan-ve-dunya->

edebiyatunasligi-problemler-ve-perspektivler-movzusunda-respublika-konfrans-84

6. Создание корпуса художественных текстов азербайджанского языка: [Электронный ресурс] / – М.: Материалы Международной научной конференции студентов, аспирантов и молодых учёных «Ломоносов-2025» – в Азербайджане. URL: https://lomonosov-msu.ru/archive/Lomonosov_2025/data/9614/183797_uid145271_report.pdf

7. «Хамсе» Низами Гянджеви как объект корпусного и лингвостатистического исследования // – М.: Сборник статей LVII международной научно-практической конференции, – 2023. – с. 170-171.

8. Корпусное исследование на материале азербайджанского и русского языков. Создание корпуса текстов Низами Гянджеви // – М.: Текст. Исследования молодых учёных, – 2023. № 1–2 (8–9), с. 121-128.

9. Лингвостатистический анализ основных предикатов «ситуативной модальности на материале повести А.С. Пушкина «Капитанская дочка» и ее перевода на азербайджанский язык»: [Электронный ресурс] / – М.: Материалы Международного молодежного научного форума «Ломоносов-2022», – 2022. URL: https://lomonosov-msu.ru/archive/Lomonosov_2022/data/25755/uid145271_5e4c9960aa04707225038b26eb9137a62141a505.doc

10. Лингвостатистический анализ на материале перевода «Лейли и Меджнун» Низами Гянджеви»: [Электронный ресурс] / – М.: Материалы Международного молодежного научного форума «Ломоносов-2022». URL: https://confhub.ru/archive/Lomonosov_2022/data/26116/146919_uid145271_report.pdf

11. Понятия «текст» и «художественный текст» в современной лингвистике // – Bakı: Filologiya məsələləri, – 2023. № 12, – s. 196-203.

12. Проблема перевода. Трудности перевода теологической литературы на азербайджанский язык: [Электронный ресурс] / – Баку: “Müasir dünyada din təhsili” mövzusunda respublika elmi konfransının materialları. URL: <https://ait.edu.az/az/news/all/muasir-dunyada-din-t-hsili-movzusunda-respublika-elmi-konfransinin-materiallari-cap-olunub/>

13. Употребление некоторых показателей возможности в корпусе текстов А.С. Пушкина: [Электронный ресурс] / – М.: Материалы Международного молодежного научного форума «Ломоносов-2020». URL: https://lomonosov-msu.ru/archive/Lomonosov_2020_2/data/section_36_19498.htm

14. Корпусное исследование сатир Мирза Алекпера Сабиря. Создание корпуса текстов Сабиря // Filologiya məsələləri, – Bakı: Elm və təhsil, – 2023. №14, – s. 425-432.

15. Лингвостатистика, или квантитивная лингвистика: закон Ципфа, закон Мензерата – Альтманна преимущества лингвостатистических исследований // – Gəncə: Elmi xəbərlər, – 2023. № 4, s. 226-229.

16. Лингвостатистический анализ перевода «Лейли и Меджнун» Н. Гянджеви //– Bakı: II International congress of modern studies in social sciences and humanities «Year of Shusha-2022» konfrans materialları, – 2022. Vol. 3, – s. 207-211.

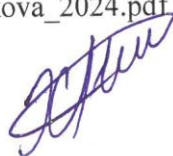
17. Создание поэтического корпуса газелей М. Физули // – М.: Балтийский гуманитарный журнал, – 2024. Т. 13. № 1(46), – с. 113-116.

18. Развитие азербайджанской корпусной лингвистики. Создание корпуса текстов азербайджанских авторов (поэтов и писателей) и национального корпуса азербайджанского языка: [Электронный ресурс] / – Махачкала: Вестник Дагестанского государственного университета. Серия 2, Гуманитарные науки, – 2024. № 2. URL: <https://cyberleninka.ru/article/n/razvitie-azerbaydzhanskoy-korpusnoy-lingvistiki-sozdanie-korpusa-tekstov-azerbaydzhanskih-avtorov-poetov-i-pisateley-i>

19. Корпусная лингвистика: истоки зарождения, терминологический аппарат. Применение методов корпусной лингвистики при анализе текста // – Bakı: Elmi əsərlər, – 2023. T. 17, № 4 (56), – с. 149-153.

20. Сопоставление морфологических категорий имени существительного в славянских и тюркских языках (на примере русского и азербайджанского языков): [Электронный ресурс] / URL:

https://www.philol.msu.ru/~slavphil/books/shirokova_2024.pdf



The defense will be held on 28 octyabr 2025 at the meeting of the Dissertation council ED 1.06 of Supreme Attestation Commission under the President of the Republic of Azerbaijan operating at the Institute of Linguistics named after Nasimi of the Azerbaijan National Academy of Sciences.

11:00

Address: AZ 1143, Baku, H.Javid avenue 115, V floor, Institute of Linguistics named after Nasimi of ANAS.

Dissertation is accessible at the Institute of Linguistics named after Nasimi of the Azerbaijan National Academy of Sciences library.

Electronic version of the abstract is available on the official website of the Institute of Linguistics named after Nasimi of the Azerbaijan National Academy of Sciences.

Abstract was sent to the required addresses on 27 sentyabr 2025

Signed for print: 25.09.2025

Paper format: 60x84 1/16

Volume: 43 825

Number of hard copies: 20