Potential Of Using Abai's Poetic Corpara For Research, Educational And Methodical Purposes

Pirmanova K.K.[1*], Serikbayeva A.D.[2], Nurlybayev N.M.[3], Sabirova D.A.[4]

[1*]PhD postdoctoral student of Al-Farabi Kazakh National University, Almaty, Kazakhstan
e-mail: kunsulu.pirmanova@mail.ru, https://orcid.org/0000-0003-3783-3199
[2]PhD postdoctoral student of Al-Farabi Kazakh National University, Almaty, Kazakhstan. e-mail: sd_aizhan@mail.ru
https://orcid.org/0000-0002-2585-4892
[3]Candidate of philological Sciences, Shymkent University, Shymkent, Kazakhstan. e-mail: nuraraika@mail.ru
https://orcid.org/0009-0000-3617-6049
[4]PhD of Abai Kazakh National Pedagogical University. e-mail: d.sabirova@abaiuniversity.edu.kz
[4]https://orcid.org/0000-0002-6001-5314

Abstract
The article describes the goals and objectives of creating national language corpora and their educational potential. The creation of Abai's poetic corpora, the potential of the poetic corpora in solving research, educational and methodological problems is described as part of the research carried out by the authors.
The article contains a demonstration of the capabilities of national language corpora, including poetic sub-corpora. The latter, from the authors' point of view, can become a tool for studying both the theory and history of literature, the author's style, the system of artistic images, means of expression and poetic syntax, rhythm, meter, rhyme, etc.
As a potential for using the national corpora in general and the poetic corpus in particular, the authors name a conceptual analysis of the author's poetic text, determining the "attraction" or "repulsion" of a lexeme to a metrical structure, identifying moments of intertextuality, searching and analyzing moments of homonymy, synonymy, antonymy, polysemy and etc.
The article highlights different points of view on the possibilities of using the resources of national language corpora, including polar ones. The authors offer specific tasks for practicing various skills in teaching language and literature.
The authors conclude that the poetic corpus of Abai as part of the National Corpora of the Kazakh language in the future opens up wide opportunities for linguistic and literary studies, preservation of the cultural and linguistic fund, as well as opportunities for use in the process of teaching language and literature by both native speakers and foreigners.

Keywords: Abai Kunanbaev, research potential, lexeme, methodology, meter, national language corpora, poetic corpus, educational potential, poetry, methodological capabilities, tasks

Introduction

The development of national language corpora is currently considered one of the priority areas of modern linguistics in recent decades. Valuable markups – syntactic, morphological, semantic, frame (for example, in FrameNet) and others – provide ample opportunities for researching linguistic material with high objectivity. Currently, national corpora of American English, British, Russian, Czech, Croatian, Hungarian, Italian, Persian, Bashkir, Tatar and other languages have been developed.

The development and creation of national language corpora is a significant and integral area of modern linguistics. They are a valuable resource for the study of language, its structure, semantics and use in various contexts. National corpora typically include extensive textual data collected from a variety of sources that represent the language in its diversity and evolution. A corpus is understood as "a unified, structured and labeled array of language (speech) data in electronic form, intended for certain philological and, more broadly, humanitarian research" [1]. The corpora provides opportunities to search for individual word forms by lemmas, groups of word forms, determine text types, determine context, and obtain lexical and grammatical data. The types of marks introduced into the corpora include metatextual, morphological, accentual, syntactic, sematic, etc.

In addition to general language material, national corpora may also contain specialized sub-corpora dedicated to specific language genres or styles, such as poetry. Poetry sub-corpora offer researchers and literature lovers a unique opportunity to study and analyze poems by different authors, their stylistic features, lexical richness and rhythmic structure.

In addition to research value, national corpora, including poetic sub-corpora, have high educational potential. They can be used for educational purposes, both for students studying language or literature and for teachers who want to provide recipients of philological knowledge with access to original texts and teach them the analysis and interpretation of poetry. In

addition, national corpora can be used to develop automatic text analysis tools, machine translation applications, grammar checking, and other language technology-related tasks.

National corpora, including poetic sub-corpora, can be used for educational purposes in both language and literature teaching. Their use for educational purposes allows students and teachers to familiarize themselves with original poetic texts, expand their vocabulary and improve their skills in analyzing and interpreting poetry. Positive aspects include a high concentration in one resource centre of texts of all styles, genres, and authors with hyperlinks containing linguistic, cultural, biographical and other types of information.

For students studying languages or literature, National Corpora offer the opportunity to immerse themselves in the culture and history of a language. By reading and studying original poetic texts, students can improve their understanding of language, style and poetic techniques. This can help to develop their skills in all types of linguistic activities as well as their creative thinking.

This research is in the field of the development of the Kazakh national corpus, especially the poetic one. The authors develop the poetic corpus of Abai.

Abai Kunanbayev (1845-1904) is the first classicist and reformer of Kazakh language, culture and literature. The legacy of the outstanding Kazakh poet consists of 170 poems and 56 translations, three poems — "Iskander", "Masgut", "The Tale of Azim" and "Words of Edification" ("Kara Sozder"). Currently, the translations into world languages form an extensive fund and include texts in Russian, English, German, Turkish, Tajik, Arabic and other languages.

Material and methods

As a resource material for the article were used the National Corpora of the Russian languages, scientific articles on the problems of using materials from corpora of national languages, as well as the poetic heritage of Abai Kunanbayev, which is subject to digital accumulation in the space of the poetic sub-corpora of the Kazakh national language corpus.

In the framework of this study, the hypothetical-inductive method, analysis of scientific research in the field of corpus linguistics, and synthesis of scientific ideas in the context of the idea of using the resources of national languages in research and education were used.

To determine the research and educational potential of the national corps, the following actions were carried out:
- a review of sources containing information on the purpose of national language corpora was conducted;
- the features of language corpora that are of research, educational and methodological value are identified.

Results

The development of national language corpora is an important area of modern linguistics, and this fact is undisputed. Moreover, national corpora, especially poetic sub-corpora, offer a wealth of possibilities for literary studies and also hold great educational potential. Let us now turn to the studies that demonstrate these points.

I. Vorobey, considering the application of the concordance method in the study of R.M. Rilke's poetic texts, points out the conceptual approach to the study of the text, which helps to create a "linguistic picture of the world" of a particular author. Considering the characteristics, more precisely the nuances of meaning of the terms "term", "content", "meaning", "concept", "sense", the researcher emphasizes that the main difference between a term as a linguistic phenomenon is its emotional component, certainty through the cultural context, "translingual character" "" [2] and acts as "a conditional mental entity that does not exist outside of thought and is materialized with the help of language, through the word" [2]. Referring to the division of concepts into universal, ethnic, group and individual concepts, including the author's, the researcher goes into detail about the phenomenon of crystallization of meaning, i.e. the formation of a word with a definite meaning. Using the method of conceptual analysis, I. Vorobey divides these into concordance and interpretation.

As is well known, a concordance is "a list of all uses of a particular linguistic expression (e.g. a word) in context, possibly with references to the source" [3]. The concordance allows us to consider the poetic word not as a unit of vocabulary, but as a particular, culturally and emotionally colored and semantically charged word form within the poetic world. I. Vorobey, who analyzes Rilke's poetry on the basis of the concordance, identifies the poetic word, defines the crystallization and considers the moment of deviation from the general linguistic concept.

From this it can be concluded that the indexing of Abai's poetic corpora could, in the future, enable a conceptual analysis to be carried out and the author's concepts to be identified through the use of concordance and interpretation methods. Abai's poetic word allows us to create his poetic world through the poetic word. A poetic word is understood as a particular word – an artistic image that has a meaning in the poetic world that is lost outside the poetic text.

The poetic corpus of Abai will enable us in particular to determine the severity of the lexical units in relation to the poetic meter. B.V. Orekhov writes about this [4]. For this purpose, the already traditional structural analysis of A. Stefanovich and S. Gris can be used, which is based on measuring the degree of "attraction" or "repulsion" of a structure with words. To determine these two phenomena, the Fisher method is used, in which the frequency of occurrence of a lexeme in a construction is calculated throughout the corpus. If the frequency exceeds the expected value, one can speak of attraction, and conversely, if the expected value decreases, one can speak of repulsion of the word by the construction. For this purpose, a lexeme analysis should be carried out, which includes the following steps: recording the number of occurrences of each lexeme in the corpus and, separately, statistics of occurrences in the construction. In this way, the required ratio and determine the moments of "attraction" or "repulsion" can be calculated. This applies in particular to the counting of lexemes

and their occurrence in subcorpora of meters. In this way, you can determine the attraction of one or another word form to one or another metrical construction or, according to M.L. Gasparov, the "semantic halo of meter" [5]. The application of this method will allow simulating the genesis of a poetic text and helping the modern reader to master this model by superimposing it on the perceptual model. Such methods help in the use of corpora to solve various lexical problems, e.g. in the description of accent variants of lexemes or in literary studies in determining the genesis of a poetic text and other phenomena. The national poetic corpora of Abai will, in our opinion, help to test Gaspar's theory about the historical relationship between word and meter, even if it is a hypothetical idea. In any case, the corpus offers the opportunity to conduct research at the intersection of linguistics and literary criticism in order to determine the "attraction" and "repulsion" of a lexeme by meter and to draw conclusions accordingly. Moreover, a researcher interested in semantics must limit himself to analyzing lexemes and not word forms and changes, for example prosodic ones. But in general, it is possible to determine the author's individual style preferences.

The next theoretical, as it can be said, perspective of using the poetic corpus is, according to B.V. Orekhov, with the search for intertextuality. The author gives an example of an analysis that allows determining formulaic and original passages in the texts of F.M. Tyutchev, in particular, on the basis of the material of the lexeme "catastrophe". If B.V. Orekhov establishes connections between the chosen lexeme and its predecessors, the Russian poets, then the future researcher of Abai's poetry will undoubtedly have the opportunity to trace intertextual connections between the lexemes of Khakim Abai and Oriental texts, since the Kazakh thinker, as is known, knew ancient Oriental poetry well and undoubtedly drew from it a source of inspiration and technical methods for the realization of poetic inspiration. At the same time, we know Abai as the first translator from Russian into Kazakh. We assume that the national corpus of the Kazakh language will enable researchers of Russian-Kazakh literary connections to identify moments of intertextuality at the level of Kazakh and Russian poetry.
Regarding the use of the corpus' potential to preserve and disseminate the poetic heritage of Abai, we can say that the resource can be used to analyze the grammatical features of the Abai language and the evolution of poetic language in the different stages of creativity based on a linguistic database.

In addition to the research interest, a corpus of national languages can become a practical tool if the developers create opportunities for a variety of research inquiries that do not limit the use of the website. For example, poetic tagging aims to highlight the structural elements of poems such as stanzas, lines, rhymes and rhythm, which will help in creating specialized analysis tools that will help researchers understand Abai's poetic manner, his use of metaphors and stylistic characteristics.
The purpose of the study is to create and structure an easy-to-use language corpus that can serve as the basis for linguistic or literary studies, analysis of language trends and a deep understanding of the poetic heritage of Abai Kunanbaev. Prospects for the research include expanding the corpus, and creating educational resources based on it, which will ultimately serve to preserve and promote the Kazakh language and literature through the prism of the creativity of the creator of the Kazakh literary language.
Modern methods of teaching language and literature are aimed at instilling the ability to use what has been created by man for personal or social purposes, that is, functional literacy. Let's consider the prospects for using the corpus for educational purposes.
I.V. Deryabina [6] notes such features of language corpora as location on machine media, standardized representation, finite size and representativeness. "The corpora must present with maximum objectivity the diversity of the phenomenon being studied and at the same time give an objective picture of the application of this phenomenon in the speech practice of a native speaker," notes V.V. Rykov. [7, p. 89-96]. These findings highlight the methodologically significant capabilities of national language corpora.

N.R. Dobrushina points out the capabilities of the Russian language corpora in composition tasks: "... the presence of the National Corpora greatly facilitates the process of composing exercises of various kinds" [8]. S. Anokhina [9] draws attention to the possibilities of using the National Corpora of the Russian language in the professional training of primary school teachers. The author proposes to study the categories of polysemy and homonymy using examples from the corpora of the sub-corpora "children's literature". According to the author of the article, students go through several stages of working with the text material: getting to know texts from the National Corpora of the Russian language sub-corpora, analyzing the meanings of words in different texts, determining the primary-secondary nature of lexical- semantics variants, and studying homonymy, synonymy and antonymy by searching for them in texts of children's literature of one or different historical periods. In this case, the analysis is based on the search for the author's motifs in the selection of synonyms.

E.A. Grishina in her article highlights interesting approaches to the analysis of markers of Russian colloquial speech based on film transcripts in the corresponding sub-corpora and finds the following signs of oral speech: the introduction of special interphrasal particles, ellipsis of conjunctions, explication of the performative, introduction of metatext, explication of roles, intensification of phatic, apocope, rematization and emphatization of the text, implementation of annotation excess, discrepancy between the transcript and subtitles, etc. [10]

Let's consider what types of background a poetic corpora can provide during teaching the Kazakh language and literature at various levels of education:
1. Search for a word and its meaning in the modern Kazakh language and in Abai's poetry (comparative analysis).
2. Comparison of the original and the translation from different positions (for foreign speakers).
3. Compilation of an explanatory dictionary of a certain poem (cycle), for example, poems about the seasons.

4. Analysis of possible changes in the semantics of the word in Abai's text and in modern poetic texts.
5. Search for homonymy and clarification of the semantics of homonyms.
6. Selection of words with the same root using the example of Abai's poetry.
7. Synthesizing different forms of one verb.
8. Search for the author's punctuation marks.
9. Search for original neologisms.
10. Analysis of borrowings in Abai's poetry.
11. Etymological analysis of a certain poem, cycle, or poems of a certain period.
12. Creating a chronological picture of the creation of poems.
13. Search for thematically similar poems and writing
14. Search for examples of words used in a metaphorical meaning.
15. Comparison of the use of words in Abai's poetry and modern prose/poetic texts in the Kazakh language corpora (if available) or in other modern sources.
16. Comparison of contextual meanings of words in Abai's poetic corpora and modern explanatory dictionaries of the Kazakh language.
17. Search for original, unusual tropes in Abai's poems.
18. Comparison of means of artistic expression transformed in Abai's translations.
19. Analysis of syntactic structures characteristic of Abai's poetic style.
20. Identification of the diachronic aspect of the formation of linguistic phenomena of the Kazakh language through the prism of Abai's poetry.
21. Determination of the equivalence of translations of Abai's texts into European languages and Turkic languages.
22. Analysis of the semiotic structure of Abai's poetic corpora (in genre, thematic and other types of concentration).

One of the important points reflecting the capabilities of the national language corpora of specifically Kazakh, precisely poetry, is the widespread use of the capabilities of the language corpora in the professional training of future literature teachers. All possible ways of using corpora vocabulary outlined above can also be successfully used in a course on methods of teaching language and literature. In particular, write tasks based on the resources of the national corps while developing methodological competence in the context of the ability.

The principles that students – future philologists must follow when compiling assignments based on the corpora include the following:
1. Taking into account the age characteristics of students who are the target audience of the assignments. It is necessary to make it clear to future teachers that the material contained in the corpora does not have age markings and therefore when choosing texts, a responsible selective approach to vocabulary is required;
2. The principle of developmental education;
3. A differentiated approach or a personalized position in the assignment policy. If educational material in paper textbooks, as a rule, is designed for the average student, then the use of extensive material from national corpora provides ample opportunities for an individual approach;
4. Optimization of the volume of material used for assignments, taking into account the educational time planned for completion;
5. Diversity of material used, including, for example, multimedia material;
6. Taking into account the real speech context while considering a linguistic phenomenon. This principle means a situational approach to the selection of materials and the creation of communicatively oriented dialogical and monological problematic moments.
7. Taking into account the relevance of the selected language material. This means relevance for the target audience "here and now".
8. A critical attitude towards the selected material, the use of additional sources of linguistic and literary knowledge.
9. Using standardized texts, and also texts that violate norms.
10. Demonstration of active processes in language that is depiction of language as a developing phenomenon.

Analysis of articles on the topic allows us to conclude that compiling tasks based on national language corpora can be implemented in two ways: with an already developed task, when only suitable material is needed, and for correcting ready-made tasks.
It should also be noted that using the capabilities of national language corpora for educational and methodological purposes has several advantages, which include saving time, increasing motivation, synchronizing modern educational practice with the demands of our time, etc.

Discussion

The overview and research position presented in the materials at hand may trigger a scholarly controversy in the sense that the linguistic community is divided into supporters and opponents of corpora development and its use in teaching and research. For example, Apresyan believes that an ill-considered use of corpora resources can lead to falsification due to misuse of quantitative data. This point of view certainly deserves consideration, but does not rule out the use of corpora materials. Widdowson, H. G. [11] also argues against careless use of housing materials. We agree with the authors of these ideas that there can be quantitative, qualitative and other biases in the texts of corpora, but nevertheless feel it necessary to

defend the position of the value of corpora for research and methodological purposes. A critical approach to the use of resources is important.

The position of S.V. Gusarenko and M.K. Gusarenko [12] deserves attention, they consider the possibility of creating corpora of texts with violations of language and speech norms in order to use them for various purposes, including methodological ones. At the same time, according to linguists, the accumulation of texts with errors, for example in the speech of bilinguals, can become valuable material and a kind of "counter-model" for anyone familiarizing themselves with standardized oral and written language.

Being able to accept different points of view, we take the liberty of expressing our point of view that the creation of the poetic corpora of Abai is still beyond the risks mentioned, since it must represent examples of classical language and texts from practically canonized sources.

Conclusion

A corpora is digitized content that contains all texts in a particular language in various formats and linguistic markups. Several conclusions can be drawn from the considerations on the use of national language corpora in research and teaching.

The use of national language corpora in the educational process plays an important role in improving the quality of education through the introduction of information technology. National language corpora not only serve as collections of texts, but also contain a number of tools that extend the possibilities of traditional educational materials.

An important feature of national corpora is its structuring into sub-corpora, which provides teachers and students with valuable material for the learning process. Homonymy is important.

National corpora offer the following possibilities:

1. Automated search

The presence of search engines in corpora makes it possible to quickly and efficiently find specific words, phrases or structures in a large number of texts. This greatly simplifies the process of learning language norms and rules and also helps to study the features of the language in depth.

The presence of search engines in text corpora is an important aid in the study of language norms and rules as well as in the study of language features. Search engines make it possible to search quickly and efficiently for specific words, phrases or structures in a large amount of text.

The advantage of the corpora approach is that it not only makes it possible to find the necessary linguistic and literary information, but also to analyze it in context, i.e. it provides a picture of how exactly this word form is used in different contexts or in different ways of speaking.

Search engines in text corpora reveal features of the grammatical structure and contextual meanings of words, e.g. by searching for collocations for a particular word or phrase.

2. Linguistic markers

Linguistic markers in corpora provide information about cases, tenses, parts of speech and other grammatical aspects that can be used to automate morphological information. In addition, linguistic markers can indicate the syntactic role of a word in a sentence, e.g. subject, predicate, object, etc. This undoubtedly raises the problem of losing the ability to determine grammatical categories independently. But this problem is solved individually in each methodological case.

Corpora information makes it possible to better understand the structure of the language and to analyze texts more precisely. With the help of these markers, grammatical rules and norms can be learned and understood, and specific information can be extracted from text. For example, corpora offer the possibility of studying cases, declensions and word agreement when studying native and foreign languages.

Furthermore, linguistic markers in language corpora facilitate the work of researchers and developers of linguistic systems and tools. They make it possible to automate text processing, e.g. to create automatic morphological analyzers or syntactic parsers.

Linguistic markers in language corpora are an important tool for language learning and linguistic research. They help to understand the grammatical properties of words and their syntactic functions, which contribute to a deeper understanding and application of language norms and rules. Accordingly, poetic markers advance literary research.

3. Regulatory data

The national language corpora include dictionaries and reference works with the current language norms. In this way, the correct spelling, pronunciation and use of words can be checked quickly and easily and their interpretations determined.

In addition, national language corpora can also contain text samples of different genres and styles, making it possible to examine the characteristics of the language in the context of its use. They are a valuable resource not only for linguists and philologists, but for anyone interested in language and its proper use. National language corpora are actively used for educational purposes, helping to study and analyze language based on real data. This tool promotes the development of language analysis and understanding of linguistic phenomena.

Thanks to the constant updating and supplementation of information, the national language corpora remain current and relevant for research in the fields of linguistics, literary criticism and language teaching. They are an important tool for the preservation and standardization of language in modern society and contribute to its development.

4. Authenticity of the texts

A special feature of language corpora is the presence of texts from real language use, which in their composition come as close as possible to the vocabulary used in everyday life. In contrast to traditional teaching texts, which can be artificially composed or adapted, the texts in corpora reflect the natural language use of native speakers.

This makes language corpora more realistic and useful for language learning as they provide students with real-life examples of words and expressions used in different contexts. In this way, students can better understand how to use language in real-life situations and develop their communication skills.

In addition, language corpora provide the opportunity to study different aspects of a language, such as vocabulary, grammar, stylistics and semantics, using large amounts of data. This helps linguists and researchers to gain a deeper understanding of language and its use in different contexts.

5. Corpora analysis

The ability to carry out corpora analysis makes it possible to examine language patterns, determine the frequency of use of words and expressions, trace semantic connections and explore the historical development of language. This deepens the understanding of language processes and promotes the development of research skills.

An important part of corpora analysis is the ability to examine the context in which words and expressions are used. This enables a better understanding of their meanings and the differences in their use in different situations. In addition, analysing the semantic relationships between words enables the identification of subtle nuances of meaning, which is useful both for learning a language and for its translation and interpretation.

Corpora analysis also opens up possibilities for studying change in diachronic and synchronic aspects. Analysis of the frequency of use of words and expressions in different periods allows observing the evolution of a language and identify trends in its development. This is especially important for linguistic research and the creation of dictionaries and grammars that reflect the current state of the language.

Overall, corpora analysis is a powerful tool for studying language at a deeper level and understanding its functioning in real-life communication settings. It helps not only linguists and translators but also ordinary language users to better master its structure and features.

6. Linguistic and literary research tools

Language corpora provide access to various linguistic and literary research tools, such as identifying collocations, analyzing semantic fields, and searching for synonyms and antonyms. These tools help students develop critical analysis and language interpretation skills.

7. Basis for the development of educational materials

Language corpora also serve as a source for the development of modern teaching materials. Authentic texts, normative data and linguistic corpora labels can be integrated into higher education syllabuses, providing deeper and more comprehensive language learning. The use of language corpora becomes an important moment of immersion in the language environment and makes it possible to develop tasks of both a research and pragmatically oriented active nature to develop oral and written speech.

Language corpora are a comprehensive resource for linguistic research, literary studies, and language teaching. They provide access to a variety of tools, such as collocation detection, which helps to identify common word combinations in a text. Analysis of semantic fields allows understanding the semantic relationships between words and identifying their meanings in context. Finding synonyms and antonyms helps to expand your vocabulary and better understand the differences in word meanings. These tools promote the development of critical analysis and interpretation of language data, a key skill for student researchers in language and linguistics.

We would like to note another relevant point related to the educational context of the use of language corpora and the poetic corpora of Abai, in particular. This aspect is associated with the formation of functional literacy of students. It is about the semantic reading of literary texts. Working with the meanings of words contained in texts, searching for semantics and etymology, searching for historicisms and archaisms, determining connections with modern meanings of words and other types of work are associated with a conscious perception of texts, penetration into their structural and content features.

Therefore, we can conclude that, despite all possible data inaccuracies, incomplete information or any other negative aspects, national corpora are a valuable source of linguistic material and can be used for both research and methodological purposes. At the same time, of course, professional users and developers have a great responsibility to use corpora data within reasonable limits and on the basis of a critical professional approach.

The research on which this article is based is in the development phase, so many of the ideas presented here are hypothetical. The prospects for the research are the development of Abai Kunanbaev's poetic corpora and the search for ways to use its resources for research, educational and methodological purposes, which should be tested during the experimental work.

In the modern educational situation, a language corpora is already recognized as one of the convenient tools for working with text. The positive aspects of using Abai's poetic corpora include the possibility of linguistic and literary analysis, the development of oral and written speech, the possibility of checking knowledge, the help of various types of markers that allow modeling and searching for various specified parameters, depending on the educational situation.

Reference

1. Zakharov, V. P. (2003). Internet search engines as a tool for linguistic research [Poiskovye sistemy Interneta kak instrument lingvisticheskikh issledovaniy]. Russian Language on the Internet. Kazan, pp. 48-59. (in Russian)

2. Vorobey, I. A. Concordance as part of the method of poetic concept analysis [Konkordans v sostave metoda analiza poeticheskogo kontsepta]. Retrieved from https://www.academia.edu/ (in Russian)

3. Concordance [Konkordance]. (n.d.). Lomonosov Fund Encyclopedia. Retrieved from http://www.lomonosov-fund.ru/enc/ru/encyclopedia:0127200#google_vignette (in Russian)

4. Orekhov, B. V. (n.d.). Once again about the research potential of poetic corpus: meter, vocabulary, formula [Eshche raz ob issledovatel'skom potentsiale poeticheskogo korpusa: metr, leksika, formula]. Retrieved from https://publications.hse.ru/articles/178496831(in Russian)

5. Gasparov, M. L. (2012). Meter and meaning: On one of the mechanisms of cultural memory [Metr i smysl. Ob odnom iz mekhanizmov kul'turnoy pamyati]. Moscow: Fortuna EL, pp. 3-16. (in Russian)

6. Deryabina, I. V. (2012). Characteristics of pedagogical work with the English National Corpora [Kharakteristika pedagogicheskoy raboty s angliyskim natsional'nym korpusom]. Humanities. Pedagogy and Psychology, 9(113), pp. 156-160. (in Russian)

7. Rykov, V. V. (1999). Pragmatically oriented corpus of texts [Pragmaticheski orientirovannyy korpus tekstov]. Tverskoy Linguistic Meridian, 3, pp. 89-96. (in Russian)

8. Dobrushina, R.M. (n.d.). How to use the National Corpora of the Russian Language in education [Kak ispol'zovat' natsional'nyy korpus russkogo yazyka v obrazovanii]. Retrieved from
https://ruscorpora.ru/old/sbornik2005/18dobrushina.pdf(in Russian)

9. Anokhina, S. A. (2007). Using materials of the National Corpora of the Russian Language in the study of lexical categories [Ispol'zovanie materialov natsional'nogo korpusa russkogo yazyka pri izuchenii leksicheskikh kategoriy]. In N. R. Dobrushina (Ed.), National Corpora of the Russian Language and problems of humanitarian education. Moscow, pp. 150-159. (in Russian)

10. Grishina, E. A. (n.d.). On markers of colloquial speech: Preliminary research of the cinema subcorpus in the National Corpora of the Russian Language [O markerakh razgovornoy rechi: predvaritel'noe issledovanie podkorpusa kino v natsional'nom korpuse russkogo yazyka]. Retrieved from www.ruscorpora.ru(in Russian)

11. Widdowson, H. G. (2000). On the limitations of applied linguistics. Applied Linguistics, 21, 3-25. https://doi.org/10.1093/applin/21.1.3(in English)

12. Gusarenko, S. V., & Gusarenko, M. K. (2022). About the corpus of speech samples with errors in the use of Russian as a foreign language: Ways of data representation and parameters of deep annotation [O korpuse rechevykh obraztsov s oshibkami v upotreblenii russkogo yazyka kak inostrannogo: sposoby predstavleniya dannykh i parametry glubinnoy razmetki]. Humanities and Legal Studies, 9 (4). (in Russian)