



The new life path of world languages or national language corporations

Mahammadiyeva Diyora

SamSU 2nd year master's degree

ABSTRACT

This article highlights the importance of using the language corpus for representatives of various fields if we create a national corpus of the Uzbek language. It has been proven how important the corpus is for lexicographers, literary critics, teachers, programmers, journalists, pragmatologists, forensic linguists.

Keywords:

Digitization process, language corpus, electronic library, educational corpus, national corpus, corpus usage, corpus importance

In our country, the focus on language has risen to the level of one of the priorities of spirituality. Therefore, along with the preservation, enrichment of our native language, increasing the effectiveness of its practical use, achieving widespread use of the Uzbek language in the modern information and communication system has become an urgent task. Because achieving the emergence of our native language in the world is one of the main ways to improve and enhance the national spirituality. As digital technologies advance, their use is becoming a daily necessity. In the new Uzbekistan today, every industry is in the process of digitization. Concepts like digital economy, e-government, e-application have become a part of our lives. Linguistics is no exception. Language is a key tool for preserving and enriching our spiritual heritage. Preserving languages means preserving national spirituality. After all, the original spirituality exists only in the national form. Modern information technology has opened the door to a wide range of conveniences in the use of the functional capabilities of language. Computer translation, editing, analysis, electronic dictionaries and thesaurus (language treasures) are the confirmation of our opinion. In

particular, the creation of a culture of creating and using modern electronic dictionaries is important in expanding the possibilities of language. In our country, great attention is paid to the effective use of information technology in all areas, projects based on software. The changes taking place in our new Uzbekistan require our linguists to study in depth the lexical system of the Uzbek language, the narrowing and expansion of meaning in words, ambiguity, new concepts and terms. puts. With the advent of the twentieth century, a number of countries have created their own national corps. Of course, in the XXI century, the Internet, the spread of information technology is determined by the fact that the language is recognized and promoted on a global scale, associated with modern communications. In order for the language, in particular, the Uzbek language to be studied and recognized by the world community, first of all, its basic lexical and grammatical bases must be formed and software must be created. So, what are the advantages of digitizing the Uzbek language, creating a national corpus of the Uzbek language. The national corps means the national language treasure. It is widely used by linguists, lexicographers, computer linguists,

programmers, editors, translators, journalists, publishers, scientists, teachers, learners and any other field specialist. Language corpus is an undeniable tool for language research and practical problem solving. It is different from a normal electronic library. The purpose of the e-library is a comprehensive coverage of works of art and journalism that reflect the socio-political, spiritual, economic life of the people. E-library texts are inconvenient for research because they are not processed from a linguistic point of view. This is because the e-library will not be created to prepare a database of scientific research material, but will be aimed at consolidating the national spiritual heritage. The language corpus, in contrast to the electronic library, involves the collection of necessary, useful and interesting texts for the study and research of a language. Corpus in various scientific literature different described. "Corps - known purpose collected texts set e _ _ til units sum " "Corpus - separately of the tongue write obtained verbally and written texts based on Linguistic data set", "corpus" - an electron in any natural language in the form of stored written or verbally computerized search to the system The body is a set of language units that make up a set ¹of texts for a specific purpose, written or spoken in electronic form in a natural language, placed on a computerized search engine based on software on-line or offline -line is ²like a set of texts running on a system. In OTIL, the word corpus is interpreted as follows: Corpus [lot.corpus - body, body, substance; a whole thing] 1. The part that forms the base or shell of a mechanism, apparatus, etc. Hour corpus. 2. One of the buildings in a common plot, part of a complex, a separate building. 3. a large military unit consisting of several divisions or brigades. 4. All diplomats, consuls in a country. The legislature consists of all deputies of individual legislative bodies or chambers belonging to the legislature ³. Apparently, OTIL does not provide any information about the language corpus that is the object of our study. Given that OTIL was

published in 2006, it is clear that corpus linguistics is a new branch of the Uzbek language that is just beginning. However, world linguistics has made great strides in this regard. In world linguistics, a lot of work has been done in the field of corpus linguistics and the creation of language corpora. In many countries, such corps began to be formed in the 80s of the twentieth century. They serve a variety of purposes and tasks. The Bank of English and the British National Corpus (BNC) have developed projects in the UK, and the Russian National Corpus in Russia. For example, the volume of the National Corpus of the Russian language is now 149 million words. In recent years, the development of the Internet has led to the emergence of a corpus of virtual texts. That is, search engines on the Internet, electronic libraries, virtual encyclopedias serve as a corps. The genre and thematic diversity of the case depends on the interests of the internet user. Corpus users can be representatives of different industries, professions. The range of corpus users is much wider than we thought, and we list them below. Just as dictionaries and grammars are used not only by linguists, the corps is an essential base for the humanities researcher, literary critic, and historian. The corps for the teacher is an incomparable treasure. For example, he wants to use only fiction (or a non-fiction text: sports, an autobiography, or just a limited work of art) to prepare an assignment. At this point, there is a need to distinguish between the educational and the national corps. The main difference of the educational corps from the national corps is that the units (texts) of the educational corps have a didactic character. Since the main goal in the national corpus is to show live language, texts of different styles and genres are used as a corpus unit, while in the national corpus, representativeness (genre diversity of text) prevails. The texts in the educational corps will consist of texts aimed at language education, spiritual and moral education of students. The educational corps includes more genres of

¹ Nurillayeva O. Principles of creating the national corpus of the Uzbek language. Graduation qualification work. Tashkent - 2018

²Xamroyeva Sh. A concise glossary of corpus linguistics terms. Tashkent - 2018

³ OTIL. Volume 2

proverbs, parables, fairy tales, and narrations. The main goal is to teach the mother tongue (Uzbek), as well as the educational goal. The National Language Corps is a national content of scientific, artistic, journalistic, colloquial and official texts in the Uzbek language, as well as contemporary, historical, poetic, prose, dramatic and dialectal texts. If the national language corps is created, the educational corps will be further improved. The National Language Corps is of a research and analytical nature. The educational building is an electronic resource with a practical purpose⁴. The corpus is able to meet any requirement as it covers texts of different genres and styles. For example, in a textbook written for journalism students, journalistic texts are much outdated. Taking new examples from the corpus, which is regularly enriched every day, presenting them to the student or giving him the same task, of course, brings education closer to life. In a textbook that has been reprinted many times, the examples remain the same. For a specialist who is becoming a philologist or journalist today to feel all the subtleties of our language, it is not enough to have an old sample of text 20-30 years ago, he must work with works in all fields, styles and genres written today. Only then is education considered to have fulfilled the social order. The language corps can be used more effectively by teachers, students and schoolchildren. Because it is only through the corpus that one can easily find, use, and spell (spell) a word, phrase, and compound that is rarely used. It should be noted that in the language corpus, language is not as described by the author of grammar and textbooks, but as it lives and works in society. This is the most productive tool in the study of the student's environment - the vernacular and literary language. The specialist who needs the language corpus the most is a programmer who works with automatic word processing (for example, a

translation program) and various search engines. Because he works with natural language and must "understand" the structure of all texts written in this language (in natural and living language, not relying on examples from scientific grammar and textbooks!). Another of the professionals who feel the need for a language corps are those who frequently appeal to the charm of written and oral speech in their daily work: newspaper and magazine editors, journalists, radio and television staff. Because these experts need more than a grammar scholar to know the status, level of use of a particular word, phrase or sentence, who first used it, when, and for what style. No information bank other than the corps can answer such questions immediately. Just as we have relied on textbooks, scientific grammars and dictionaries so far, so many representatives of the modern field feel the need for the corps and use it⁵. Lingvodidactics is a field that is in great demand in the corps, which is equally important in the teaching of the native language and in the teaching of foreign languages. The corpus is very useful in showing the array of examples to show the magnitude of the vocabulary in language teaching, to explain the possibility of using the word through this or that grammatical construction. The fact that the example, which is important for language education, is constantly updated, and the ability and opportunity to reflect it is unique to the corpus. The teacher can find new, convincing, endless and varied examples here, it is not difficult to define assignments and exercises, in a few minutes he can prepare an exercise consisting of new examples on the topic.⁶ As noted by President Sh.M. Mirziyoyev, "the presentation of our native language in an attractive and interesting way requires the creation of computer programs, online textbooks, electronic dictionaries of the Uzbek

⁴Raupova L., Zaripboyeva N. Educational building: as an applied, practical electronic resource. Theoretical and practical issues of creating Uzbek national and educational corps. Vol.1.2021

⁵ Nurillayeva O. Principles of creating the national corpus of the Uzbek language. Graduation qualification work. Tashkent - 2018

⁶Bakhriddinova B., Bahodirova G. National language renaissance and lexical development in Uzbekistan. Theoretical and practical issues of creating Uzbek national and educational corps. Vol.1.2021

language . " ⁷Even psycholinguists, pragmalinguists, and forensic linguists working in relatively new fields of linguistics can use the corpus effectively, achieving the expected result in seconds. To do this, first of all, it is necessary to create an open national corpus program, and devotees of the field are required to constantly replenish its content with oral and written texts of various genres. In particular, with the help of a pragmalinguist corpus, one can find hundreds of different texts and sort them out in a short time in determining the pragmatic value of certain linguistic units in the communication process. If we take a deeper approach to the topic, the semantics of some words, the pragmatic purpose of speech, are interpreted differently by people of different ages, genders, social classes and groups. For example, the qualities of "bad", "horror" and "crazy" in the speech of teenagers can be a means of positive assessment today. But for the older generation or parents, the opposite may be true . Nowadays, in world linguistics, forensic linguists often use language corpus. Because perfectly crafted language corporations cover everything from suicide letters to cell phone messages , police statements, transcripts of police conversations, to witness statements. The corpus has the following meanings for linguistic expertise:

- Phonetic, lexical, grammatical, features of speech of different regions;
- Identify the exact and similar aspects of plagiarism ;
- insults, curses, and defamation;
- Murojaat. Determining the national ethnic identity of extremist materials in the form of summonses using corpus bases ⁸. The creation of a national corps in the Uzbek language is still being postponed. Many dictionaries have been

⁷ Mirziyoyev Sh. The Uzbek language is a symbol of our national identity and independent statehood, invaluable spiritual wealth and great value for our people. Speech by President Sh. Mirziyoyev at the solemn ceremony dedicated to the thirtieth anniversary of the official status of the Uzbek language. October 21, 2019. <http://uza.uz/oz/politics/>

⁸Musulmonova K. The role and prospects of the creation of the National Corps in the development of criminolinguistics.

created in Uzbek, but no serious research has been done on the development of the language corpus and its theoretical foundations. Shahlo Khamrayeva, an independent researcher from Bukhara, conducted a study of the author's corps today. His research has a worthy place in the list of important sources created in the Uzbek language for the criteria of creating a national author's corps . Also, today our researchers, scientists and programmers, such as G. Begmatova ⁹, A. Eshmuminov ¹⁰, Khidirov O. ¹¹, Juraev J., Rabbimov I., have done a lot of work on corpus linguistics and its software. We are also in our future work The essence of corpus linguistics, based on the practical experience of the author's corpus of this field abroad, we have set as our goal to create a linguistic supply of Osman Azim's dramas in the language corpus as well.

In order to create a model for the national author's corps of the Uzbek language on the basis of Osman Azim's dramas, Usmon Azim's "Life of a writer", "Nights without days", "Abdullah Qahhor", "Return of Alpomish", "Hangomalari of the past", "Bor ekan-u" dramas such as "No Step", "One Step Way", "Market" were selected.

In short, the scope and results of previous research on language need to be studied in a comprehensive manner, and the theoretical knowledge of corpus linguistics needs to be supplemented, researched and practiced. One of the tasks before us is to create author's corps, considering that one of the main directions of modern corpus linguistics is based on author's corpora to conduct new and large-scale linguistic research . It is very important to place materials in various fields on the basis of the national corps to be addressed . The creation of various network corporations will serve as a

Theoretical and practical issues of creating Uzbek national and educational corps. Vol.1.2021

⁹ Begmatova G. Creating a database of idioms in the Uzbek National Corps. PhD dissertation. 2021

¹⁰ Eshmuminov A. Synonym database of the Uzbek language national corps. PhD dissertation. 2019.

¹¹ Khidirov O. Linguistic bases of creating parsing program for the National Corps. .PhD dissertation abstract. 2021

factor for the development of science and the development of a particular field.

13. Khidirov O. Linguistic bases of creating parsing program for the National Corps. .PhD dissertation abstract. 2021.

Books:

1. Xamroyeva Sh. A concise glossary of corpus linguistics terms. Tashkent - 2018
2. OTIL. Volume 2
3. Raupova L., Zaripboyeva N. Educational building: as an applied, practical electronic resource. Theoretical and practical issues of creating Uzbek national and educational corps. Vol.1.2021
4. Nurillayeva O. Principles of creating the national corpus of the Uzbek language. Graduation qualification work. Tashkent - 2018
5. Bakhriddinova B. , Bahodirova G. National language renaissance and lexical development in Uzbekistan . Theoretical and practical issues of creating Uzbek national and educational corps. Vol.1.2021
6. Mirziyoyev Sh. The Uzbek language is a symbol of our national identity and independent statehood, invaluable spiritual wealth and great value for our people. President Mirziyoyev has given Uzbek the status of the state language
7. speech at the anniversary ceremony. October 21, 2019. <http://uza.uz/oz/politics/>
8. Musulmonova K. The role and prospects of the creation of the National Corps in the development of criminolinguistics. Theoretical and practical issues of creating Uzbek national and educational corps. Vol.1.2021
9. Mengliyev B et al. National corpus of Uzbek language // Ma'rifat, - Tashkent, 2018.
10. Xamroyeva Sh. Linguistic bases of creation of the Uzbek language author's corps. PhD dissertation abstract. 2018.
11. Begmatova G. Creating a database of idioms in the Uzbek National Corps. PhD dissertation. 2021.
12. Eshmuminov A. Synonym database of the Uzbek language national corps. PhD dissertation. 2019.