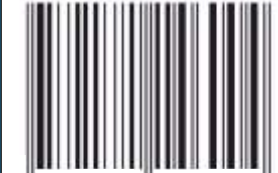




E-ISSN 2181-1466



9 772181 146004

ISSN 2181-6875



9 772181 687004

CONTENTS

EXACT AND NATURAL SCIENCES		
Durdiyev D.Q. Suyarov T.R.	The inverse problem for systems of incompressible viscoelastic polymeric fluid at rest	2
Jumaev J.J.	Two dimensional inverse problem of determining the kernel of the integro-differential heat equation	26
Merajova Sh.B.	Numerical study of the first and second boundary value problems posed for a model equation of mixed type in space, and the calculation of the approximation of a stable finite-difference scheme of these problems	46
Orujova Sh.M.	Landscape studies of the Nakhchivan Autonomous Republic: semi-desert and mountain xerophytic landscapes	63
LINGUISTICS		
Valiyeva K.S.	Analysis of lexico-semantic characteristics of pathological terms in azerbaijani and arabic	75
Avezov S.S.	Corpus linguistics: historical and linguodidactic aspects	86
Djuraeva Z.R.	The problem of studying the concepts “gender” and “sex” in world linguistics	95
Tursunov M.M.	Proverb as an essential lexical unit representing national culture and identity	105
LITERARY CRITICISM		
Turaeva U.R.	Literary critic Begali Kasimov is a researcher of modern literature	114
Kakharova M.Y.	Classification of various types of characters in literary works	122
“NAVOIY GULSHANI”		
Sayliyeva Z.R.	Chu jilva ayladi ul husn istabon oshiq...	132
PHILOSOPHY, LAW AND POLITICAL SCIENCES		
Kadyrova A.B.	Al-Hakim at-Termizi “Hidden issues” (“Al-masoil al-maknuna”) brochure. A statement of the issue of “The soul is the master of the organs of the body”	138
Jumayev R.G.	Ahmad Donish’s views on social justice and political equality	146

CORPUS LINGUISTICS: HISTORICAL AND LINGUODIDACTIC ASPECTS

Avezov Sukhrob Sobirovich

Lecturer at the Department of Russian Language and Literature Bukhara State University

1990senigama@gmail.com

Abstract.

Introduction. *The article considers the main stages in the development of corpus linguistics. First, the prerequisites for its development are described, it is indicated which branches of linguistics laid the foundation for corpus linguistics. This article details the history of the creation of the first foreign and domestic corpora, and also highlights the main controversial issues about the status of corpus linguistics as an independent scientific discipline. The last issue addressed by the author in the article is the current state of corpus linguistics and its importance in teaching a foreign language.*

Materials and methods. *The development of this approach was the beginning of the creation of corpus linguistics. Many of the technologies that are now used in building cases were invented long before the advent of computers and electronic resources.*

Results. *Based on the experience of foreign countries, we see that NC covers all layers of the language, and for the corpus to work effectively, it is required that each of the levels be equipped with the required information.*

Discussions. *Let's go back a few decades, to the moment when zealous disputes began about its status as an independent discipline. I would like to remind you that in the 60-80s of the XX century, when the first corpora began to be created in the USA and Western Europe, the rational (Chomskian) approach dominated the scientific community.*

Conclusions. *. Thus, the linguistic corpus is a means for solving not only scientific, but also educational and methodological problems. The benefits of its application in various fields are beyond doubt, although the theoretical basis has not yet been fully developed*

Key words: *corpus linguistics, national corpus, Brown Corpus, British National Corpus, National Corpus of the Russian language, learning with the help of corpora.*

Introduction. The development of the science of translation in general and various areas of corpus linguistics in particular is largely associated with the use of modern linguistic resources and methods, including large volumes of texts, scientific, socio-political, legal, economic and their translations into foreign languages, collected in parallel boxes. The possibility of using extensive parallel corpora of translations makes real changes in the ways of accessing bilingual linguistic

information, opens up new possibilities for its analysis and generalization, followed by the application of the acquired knowledge for the purposes of manual and machine translation, compiling dictionaries, and teaching foreign languages.

The most important role for the use of these resources is played by annotation of source text materials collected from a variety of sources, primarily from the Internet, starting from a simple comparison of sections, chapters, paragraphs of the source text with similar structures of the target text - to the alignment of sentences, phrases and words inside sentences, with morphological, syntactic and semantic parsing and comparison of parse trees, either in the model of immediate constituents, or in the model of semantic-syntactic dependencies. The size and degree of annotation of the corpus is determined by the tasks of its researchers. Large corpora, balanced in terms of subject matter, chronology and genres, serve to study the general properties of the language, while when creating specialized parallel corpora, they try mainly to solve the problems of analyzing specific aspects of translation, or specific types of texts, determining authorship, etc.

The development of corpus linguistics, as well as the construction of corpora, is one of the urgent problems of modern linguistics. At the moment, the use of corpora plays a leading role in most linguistic research. Although in the first half of the twentieth century it was possible to build buildings only by hand. It took a lot of time, effort and expense. Therefore, the appeal to the corpora of texts was minimal and only in the case when a large number of people were interested in this. But thanks to the development of corpus linguistics, the use of corpora has become possible everywhere. Structuring and synchronizing material has become much easier, and costs have dropped dramatically.

The importance of the development of corpus linguistics is beyond doubt. It combines many positive features of the humanities and technical sciences. But having arisen relatively recently, it did not have time to fully develop, and its status as an independent science has not yet been established. Thus, many scientists consider corpus linguistics as a sub-field of traditional linguistics. They argue that it has only applied application, and there are no theoretical justifications. In their opinion, corpus linguistics is an improved method of collecting and processing material, a new information resource. On the other hand, if there is applied corpus linguistics, why can't there be theoretical, i.e. discipline that has its own subject, method and theory. The lack of a theoretical basis at the moment does not affect the possibility of its development in the future.

Materials and methods. In the middle of the XIX century. the scientific community was dominated by a rational approach based "on linguistic intuition, distinguishing between correct and incorrect constructions" [1, 14]. In contrast to this approach, an empirical approach appears, which proposes to consider the language as "a resource that provides a set of opportunities for communication" [1, 14]. The development of this approach was the beginning of the creation of corpus linguistics. Many of the technologies that are now used in building cases were invented long before the advent of computers and electronic resources. Some of them were used as early as the 18th-19th centuries, when linguistics began to be considered an independent and independent scientific discipline. The famous Rosetta stone,

created in the 2nd century BC, can be considered the beginning of the parallel building . and containing a parallel text in 2 variants of ancient Egyptian writing and in ancient Greek [2]. Back in the 60s , R.G. Piotrovsky argued that only from a large array of texts can one obtain reliable data on the semantic structure of language and speech [3,144].

In his textbook "Corpus Linguistics" Zakharov V.P. names the technologies that influenced the creation of cases. He identifies three main areas of linguistic research that formed the basis of corpus linguistics, although he notes that there were many more [4, 25].

The first such area he singles out is comparative-historical linguistics. Scientists working in this field have always turned to a huge number of different texts. The use of technologies for the reconstruction of proto-languages can also be found in modern linguistics. The second area, which Zakharov V.P. preference is the compilation of grammars and dictionaries and language teaching. Indeed, any grammatical rule needs to be illustrated. And in this case, examples from the text can perfectly help with this. Corpora as sources of empirical data play an important role in teaching a foreign language. The last area to influence the development of corpora is sociolinguistics. As early as the 19th century, scientists began to develop dialect maps and compile collections of dialect relations. At the same time, it was necessary to take into account various criteria when compiling manuals on dialects. All these factors served as the beginning of corpus linguistics.

At the moment, corpus linguistics is successfully developing the technologies and methods that led to its birth. Also, do not forget about the technical side of the issue. There has been a sharp leap in the development of computer and information technologies. These possibilities began to be successfully used in linguistics and linguistics. Thanks to the development and popularization of the global Internet, a huge number of users from different countries could use the data from the corpus. In addition, the problem of completeness and expansion of the buildings no longer arose. Most of the research in the field of corpus linguistics has been carried out on the material of English texts. There are two main reasons for this: firstly, there is an active development of computer technology in North America and Western Europe, and secondly, there is a favorable climate for the development of British linguistics in the 60-80s. 20th century

Brown corpus is considered to be the first linguistic corpus . corpus). It was developed in 1963 by N. Francis and G. Kuchera at Brown University. The volume of the first corpus is 1 million word usages, i.e. it includes 500 fragments of 2000 word usages each from texts of different genres: literary texts of famous writers and poets, articles from newspapers and magazines, examples of written business speech and texts on religious topics. There are several reasons for the creation of the Brown Corpus. First of all, this is the provision of a systematic study of English texts belonging to various genres. Secondly, providing a sufficient amount of material to compare these data. Thirdly, attracting the interest of many scientists to the emergence of a new scientific discipline. This was a breakthrough in applied linguistics, which caused a lot of controversy and discussion. As for the compilation of the corpus itself, it was built, on the one hand, on the basis of statistical data, and,

on the other hand, on the experience and intuition of scientists. To achieve objectivity, it was necessary to use simple and transparent texts.

Later, in 1971-78, following the example of their American colleagues, European scholars began compiling another body of texts. It is called the Lancaster - Oslo- Bergen Corpus of British Modern English (The Lancaster - Oslo / Bergen corpus of british English , LOB). The compilation of this dictionary was carried out mainly by British and Norwegian scientists. They were guided by the same principles as scientists from Brown University. For the period of creation, it consisted of 500 texts belonging to various genres, with a volume of 2000 word usages.

The most popular at the moment is the British National Corpus (British National Corpus , BNC). It was created in 1991-1994. researchers from Oxford University and Lancaster University. Its volume is 100 million words and it is much larger than its predecessors. It consists of 90% written texts and 10% oral texts. The texts belong to the end of the 20th century and represent various genres. In it you can find newspaper articles, non-fiction literature, examples of business correspondence, texts on religious topics, transcribed recordings of informal speeches, radio shows, government speeches, etc.

Results. It is worth noting that it was the British Corpus that received the status of "national" first. Plungyan V.A. writes that “for the British, the word “national” meant primarily “characterizing the British national version of the English language”. After all, there are also American and Australian English. But after a while, this corpus became the standard of all corpus and the meaning of the word “national” changed somewhat. The corpus, which is the largest and most representative and which characterizes the language of a given country as a whole, began to be called national. Plungyan V.A. notes that the national corpus should not only be large in volume, but also contain texts of various genres in all their diversity in a given historical period, and at the same time contain them in the correct proportion [5, p. 7].

Take for example the British National Corpus (BNC from English. *British National Corpus* is a 100-million-word corpus of written and spoken British English from a wide range of sources. The corpus covers British English of the late 20th century, represented by a wide variety of genres, and is intended to be representative of the typical spoken and written British English of the time. [12]



Many countries, following the example of the British National Corpus, decided to create their own national Corpus. Russia is no exception. For seven years from 2003 to 2010, scientists worked on creating a single text base. The Association "National Corpus of the Russian Language" offered cooperation to the "Yandex" company and with their technical support a website was developed, which hosts the current version of the National Corpus of the Russian Language. It contains more than 163 million word usages for the period from the middle of the 18th to the beginning of the 21st century. Thanks to the presence of texts from various eras, it is possible to find examples of both modern and historical texts. At the same time, the texts belong to the literary, colloquial, official business and scientific styles.

Based on the experience of foreign countries, we see that NC covers all layers of the language, and for the corpus to work effectively, it is required that each of the levels be equipped with the required information.

As we can see from Table 1, the BNC structure consists of 90% written and 10% oral corpus, which includes all styles and genres of the language. BNC is monolingual, that is, it contains only British English words. For example, let's turn to the National Corpus of the Russian Language (NCRL) [13]

Национальный корпус русского языка —
представительная коллекция текстов на русском языке,
оснащенная лингвистической разметкой и
инструментами поиска

ИРЯ Иллиран Яндекс

Подробнее о Корпусе

Введите слово или фразу Искать

Основной (336 млн)	Параллельный (151 млн)	Поэтический (13 млн)	Мультимедийный (5 млн)
Газетный (790 млн)	Обучающий (664 тыс)	Устный (13 млн)	Мультипарк (229 тыс)
Синтаксический (1 млн)	Диалектный (485 тыс)	Акцентологический (133 млн)	Исторический (13 млн)

The difference between NCRL and BNC is that there is no division into written and oral corpus. It consists of 12 corpora covering the entire language structure.

On the official website of the National Corpus of the Russian Language there is a reference article about the meaning of the term "corpus", about the rules of use, and the purpose of its creation is also clearly defined. It also describes the structure of the National Corpus. It currently consists of the following sections:

- Main corpus (prose written texts of the 18th - early 21st centuries);
- Syntactic corpus (in which a complete morphological and syntactic structure is built for each sentence);

- Newspaper corpus (articles from the media of the 1990-2000s);
- Parallel corpora (in which you can find all translations for a certain word or phrase into Russian or from Russian);
- Corpus of dialect texts;
- Corpus of poetic texts;
- Educational corpus of the Russian language (corpus with removed homonymy, the markup of which is focused on the school curriculum of the Russian language);
- Corpus of oral speech;
- Multimedia corpus (includes fragments of films from the 1930s-2000s provided with video and audio sequences);
- Corpus of the history of Russian stress (texts containing information about the history of Russian stress).

Russian National Corpus covers almost all facets of the language.

It should be noted that the national corpus of the Russian language is not the only corpus of the Russian language. Grudeva E.V. in his work "Corpus Linguistics" he singles out several more corpora [8, 35]. The first of these is the Uppsala Corpus of the Russian Language (The uppsala Russian corpus). It is the first corpus of the Russian language. Scientists from the University of Uppsala in Sweden worked on the creation of the Uppsala Corpus in the late 1980s and early 1990s. It includes 600 fragments of artistic and informative texts with a volume of about 1 million word usages.

It is also worth paying attention to another corpus of the Russian language. It was developed in 1999-2004. in Germany, at the University of Tübingen . It got its name according to the place where it was created (Russian Tübingen Corpus). It was developed on the basis of the Uppsala corpus, but the number of word usages has grown to 25 million word usages.

Discussions. Let's go back a few decades, to the moment when zealous disputes began about its status as an independent discipline. I would like to remind you that in the 60-80s of the XX century, when the first corpora began to be created in the USA and Western Europe, the rational (Chomskian) approach dominated the scientific community. Noam Chomsky and his supporters shared the opinion that the construction of correct and incorrect language constructions can be carried out only on the basis of the intuition of native speakers. The American linguist was an ardent opponent of the corpus approach, and there is a lot of evidence for this. N. Chomsky's answer to the interviewer's question about Chomsky's attitude to corpus linguistics is well-known: "There is no such thing" [8, 195]. And in one of the Corpora-List teleconferences, supporters of N. Chomsky entered into a discussion about the irrelevance of corpora [9, 334]. A supporter of N. Chomsky, Professor Robert Leese, in 1962 at a conference at Brown University, said that the creation of the corpus "is a waste of your time and government money. You are a native English speaker; within ten minutes you are able to present more examples of any phenomenon of English grammar than you can find in many millions of words of random texts" [9, 26]. There is undoubtedly some truth in their statements, however, corpus linguistics has found many supporters, both among scientists and in the government. More and more funds

began to be allocated for the creation and modernization of buildings. “Linguistics has now largely rid itself of early generative illusions, in particular, of the belief that linguistic mechanisms as such can be known using a very limited set of examples (usually composed by the linguist himself). These rather naive ideas are being replaced by an understanding of the need to build a study of even the “smallest” fragment of the language system using a representative set of texts of the corresponding language” [10, 94]. As a set of texts, of course, there is a linguistic corpus. Although the criteria for the representativeness of such a corpus are not yet clear enough, the task that is set for the corpus is already well defined. “The corpus must have quantitative and qualitative parameters necessary and sufficient to build on its basis an adequate dictionary and grammar of the corresponding language” [11, 93].

When creating the National Corpus of the Russian Language, scientists, of course, asked the obvious question: for whom will the corpus be useful? In what areas of life can it be applied? Plungyan V.A. answered this question quite fully . in his work "Why do we need the National Corpus of the Russian Language?". First of all, it is an invaluable tool for professional linguists: “One way or another, they deal with the facts of the language, which means they must collect and systematize these facts” [5, 8]. Secondly, it is very useful for programmers. Not surprisingly, they immediately supported the idea of creating it. Of course, the linguistic corpus is also useful for people who have connected their lives with the word, for example, writers, editors of newspapers and magazines. Any difficult moment can be checked by referring to the body.

And, finally, the building will become an excellent assistant for teachers and teachers, both at school and at the university. In the world's leading universities, it is becoming a daily practice to use corpus data as material for lecture courses, student assignments, and independent projects.

All over the world, in higher education institutions, data from linguistic corpora are used in the preparation of various lecture courses and assignments for students. Many students themselves use corpus data when working on projects and homework. It can be assumed that students who are encouraged to independently study the language, its features and traits, master language competencies faster and more efficiently than those who are hammered into the head with the rules.

At the moment, one of the most popular areas of the corpus approach in teaching a foreign language is learning with the help of corpora, or data - driven . learning (DDL). The essence of this training is that students use "raw" language data directly from the corpus. This direction is based on the assumption that students can learn a language much more effectively when the learning process is encouraged to use the observe-guess-experiment model, i.e. when they have the opportunity to draw their own conclusions about the meanings of words, phrases, grammatical rules based on linguistic material. The process is not necessarily limited to a computer terminal. The results of corpus searches (concordances) in printed form can be easily incorporated into handouts, teaching aids, etc.

The corpus approach in teaching a foreign language changes the nature of the student's learning activity and puts him at the center of the learning process. The role

of the teacher will be to organize and control the research activities of students. At the same time, the formation of students' ability to autonomously master linguistic knowledge and skills will require the teacher to make certain efforts to create methodological materials. As N.L. Baidikova, “one of the functions of a teacher in organizing training in effective methods of independent work is the methodological support of this process. The development of manuals seems to be the most important task, otherwise the organization of independent work of students turns into a laborious, cumbersome and uneconomical process” [9, p. 108]. You should be prepared that the corpus will not always give out data corresponding to the explanatory dictionary. Many of the examples can be difficult for students to understand.

At the current stage of the cultural and intellectual development of our country, one of the priority tasks is the formation of intellectual potential in the field of science, modern personnel, and high technologies. Cultural and information exchange presupposes active interaction. This, in turn, serves the intensive development of corpus linguistics in all its fields. The large-scale reforms in the language policy carried out in the republic led not only to the beginning of a new era in the development of the Uzbek language, but also to the transition to a new stage of research work in linguistics, increased attention to applied research that studies the language taking into account the human factor. Along with a number of tasks set for scientists, new goals were defined, which are to create the National Corpus of the Uzbek Language, software for the corpus, create linguistic foundations and linguistic models, as well as raise theoretical, practical and applied research to the level of the world standard.

Conclusion. Thus, the linguistic corpus is a means for solving not only scientific, but also educational and methodological problems. The benefits of its application in various fields are beyond doubt, although the theoretical basis has not yet been fully developed. That is why scientists still cannot answer the question: “What is corpus linguistics: a new scientific discipline or just an information resource?” We hope that soon the answer to this question will be found and corpus linguistics will become an independent scientific discipline.

REFERENCES:

1. Sharov S.A. *Representative corpus of the Russian language in the context of world experience* / S.A. Sharov // *Scientific and technical information*. - Ser. 2. - No. 6. - P. 12–16.
2. http://lab314.brsu.by/kmp-lite/kmp2/JOB/Corpus/CorpusLL/00-multilingual_corpora.pdf
3. Piotrovsky R.G. *The study of the vocabulary and grammar of the text with the help of an electronic computer* // *Problems of synchronous study of the grammatical structure of the language: Proceedings of the scientific. conf.* - M.: MGPIIA, 1965. - p.144-146.
4. Zakharov V.P. *Corpus linguistics: textbook.-method. Allowance* / V.P. Zakharov. - St. Petersburg, 2005. - 48 p.
5. Plungyan V.A. *Why do we need the National Corpus of the Russian Language? Informal introduction* / V.A. Plungyan // *National Corpus of the Russian Language: 2003 - 2005*. - M.: Indrik. - S. 6 - 20.
6. Ventsov A.V., Grudeva E.V., Kasevich V.B., Yagunova E.V. *National corpus of the Russian literary language: some results, applications and tasks* / Ventsov A.V., Grudeva E.V., Kasevich V.B., Yagunova E.V. // *Scientific and technical information*. - Ser. 2. - No. 6. - P. 35–36.

7. Ventsov A.V., Grudeva E.V. *On the corpus of the Russian literary language* / A.V. Ventsov, E.V. Grudev // *Russian Linguistics* . - - No. 2. - S. 195 - 209.
8. Grudeva E.V. *Corpus linguistics: textbook. allowance* / E.V. Grudev . - 2nd ed., erased. – M.: FLINTA, 2012. – 165 p.
9. Baidikova N.L. *Formation of educational competence among students of language faculties of universities in the process of teaching theoretical disciplines* / N.L. Baydikova // *Actual problems of international cooperation in the field of science and education: Proceedings of the III International Scientific and Practical Conference (in absentia)*. - Tambov: TSU im. G. R. Derzhavin. - S. 105 - 109.
10. *Brown Corpus* : [electronic resource]: <http://clu.uni.no/icame/brown/bcm.html#bc3> (Accessed: 07/05/2022).
11. *The Lancaster-Oslo/Bergen Corpus of British English, LOB*: [electronic resource]: <http://clu.uni.no/icame/manuals/LOB/INDEX.HTM> (date circulation: 07/05/2022).
12. *British National Corpus, BNC*: [electronic resource]: <http://www.natcorp.ox.ac.uk/> (date of access: 07/05/2022).
13. *National corpus of the Russian language*: [electronic resource]: <http://www.ruscorpora.ru/index.html> (date of access: 07/05/2022).
14. *The uppsala Russian Corpus* : [electronic resource]: <http://www.slaviska.uu.se/korpus.htm> (date of access: 07/05/2022).
15. *Tübingen corpus of the Russian language*: [electronic resource]: <http://www.sfb441.uni-tuebingen.de/b1/-en/korpora.html> (date of access: 07/05/2022).
16. Juraeva M. M., Khamidova M. K. *Linguoculture and Linguocultureme* // *Scientific reports of Bukhara State University*. – 2020. – T. 3. – №. 4. – P. 111-114.
17. Khamidova M. K. *The structure and means of expression of metaphorical units with deopoetonymy in french and uzbek languages* // *Scientific reports of Bukhara State University*. – 2021. – T. 5. – №. 2. – P. 67-83.
18. Khamidova M. K., Qahhorova G. S. *Features separate linguistic geopoetika in french and uzbek languages* // *Scientific reports of Bukhara State University*. – 2020. – T. 4. – №. 2. – P. 104-108.
19. Khamidova M. K. *Artistic application of depoethonyms in french and uzbek poetry* // *Scientific reports of Bukhara State University*. – 2020. – T. 4. – №. 3. – P. 210-213.
20. Kuldashova N. B. et al. *The importance of terminology in linguistics* // *Scientific reports of Bukhara State University*. – 2020. – T. 3. – №. 3. – P. 106-109.
21. *Deopoetonyms U. Teacher of department of German and French languages, BSU* // *Scientific reports of bukhara state university* – P. 57.