

Complex Corpus Of Turkisms Of The Russian Language*

Gulnaz I. MARDANOVA, Guzel N. KARIMULLINA, Rezeda N.

KARIMULLINA, Tatyana E. KARPENKO

Kazan State Federal University

Ablai Khan University of International Relations and World Languages, Almaty

Abstract. *The complex (multiple) corpus of Turkisms that is described in this article is the result of a long-standing study of the Turkic layer of Russian vocabulary, which has been formed in a multicultural language environment. The data obtained during analyzing the Turkic lexical elements from ten academic explanatory dictionaries of the Russian language, published within the period from the 18th century till the 21st century, have been involved. A large number of lexicographic parameters in dictionaries of this type make it possible to create an adequate picture of the functioning of lexical units in the Russian literary language. The study uses descriptive, comparative-contrastive and statistical methods.*

The article fixes various shifts and changes associated with the functioning of the Turkic vocabulary layer during the period under consideration, in particular phonetic-graphic, grammatical, semantic, functional-stylistic changes. In addition, it analyzes the methods of lexicography (dictionary fixation) of loan words, for example, the features of the historical and etymological description of the Turkisms, which, in turn, enables to reveal various inaccuracies and inconsistencies relating to the dictionary description of Turkic words. In addition, the article discloses the advantages of a complex fund of Turkic borrowings of the Russian language, outlines the prospects and possibilities for its use. The information resource being created allows to fix and generalize the changes, characteristic of this circle of loan words of the Russian language.

Keywords: borrowing, dictionary, Turkism, lexicography, complex corpus, the Russian language.

Introduction

The problem of language contacts has always been in the field of view of domestic and foreign linguistics. This problem takes on especial significance in the context of a multinational state, that is, Russia.¹ Many contemporary works note the importance of a thorough study of the Turkic elements existing in the Russian language.²

The study of the Turkic elements in the Russian vocabulary has a long tradition. The analysis was based on various materials (literary monuments, fictional literature, dialects, regions, folklore, the language of mass media, etc.), from various viewpoints (etymological, chronological, phonetic, spelling, semantic, grammatical,

* The work is performed according to the Russian Government Program of Competitive Growth of Kazan Federal University.

¹ G. R. Galiullina, A. Sh. Yusupova A, "European linguistic borrowings in the Tatar sprachraum (by written sources of the XIX - XX centuries)", in *Life Science Journal*, XI (2014), Issue 7, p. 697-700; G. Kh. Gilazetdinova, I. Zh. Edikhanov, A. A. Aminova, "Problems of ethnocultural identity and cross-language communication," in *Journal of Language and Literature II* (2014), p. 39-42.

² M. V. Oreshkina, (2013). "The Dictionary of borrowings into Russian from the languages of the peoples of Russia and the Near Abroad Countries: principles and concepts," in *IV International conference «Language, Culture and Society in Russian - English studies» (22 - 23 July, 2013) London, University of London; School of Advanced Study; Institute of English Studies; Russian Academy of Sciences; Russian Academy of Linguistics; Moscow Institute of Foreign Languages, London, Senate House, University of London, 2013, p. 170 – 173.*

functional, comparative-contrastive, etc.). The objects of scientific research were the Turkic borrowing words of certain thematic groups (the names of plants, animals, clothing, buildings and construction, household utensils and appliances, the names of persons, military vocabulary, vocabulary of food, etc.).

Nowadays, it is urgent to generalize, thoroughly comprehend the rich material on borrowings in modern linguistics, including a multidimensional description of the Turkic layer in the Russian lexicon.

The lexical variety of language at a certain stage of its development, as is generally known, is reflected in lexicographical sources. Explanatory academic dictionaries present this information most completely. The chief task of these publications is to explain the meanings of the units in them. However, the scope of information contained in these dictionaries is much broader. They contain the information about pronunciation, stress, writing, grammatical characteristics, the peculiarities of the use of a word, in a number of cases, they give set expressions, etymological and historical comments.

Materials and methods

A large number of lexicographic parameters in explanatory dictionaries make it possible to create a fairly complete picture of functioning of the Turkic units in the Russian literary language.

The explanatory dictionaries of the Russian language of the academic type compiled in the 18th-21st centuries were used as the sources: 1) "Словарь Академии Российской" (1789-1794) ("The Dictionary of the Russian Academy"), 2) "Словарь Академии Российской, по азбучному порядку расположенный" (1806-1822) ("The Alphabetical Dictionary of the Russian Academy"), 3) "Словарь церковно-славянского и русского языка" (1847) ("The Dictionary of the Church Slavonic and Russian Language"), 4) "Толковый словарь русского языка" под ред. Д.Н.Ушакова (1935-1940) ("D. Ushakov's Explanatory Dictionary of the Russian Language" (hereinafter UD), 5) "Словарь современного русского литературного языка" (1948-1965) ("The Dictionary of the Modern Russian Literary Language") The Dictionary of the Russian Language: in 4 volumes (1981-1984)/ Edited by A.P. Evgenieva.- 2nd edition, revised and enlarged.- M.: Rus, Yaz,³ (hereinafter GAD), 6) "Словарь русского языка" под ред. А.П.Евгеньевой (1981-1984) "A.P. Evgenieva's Dictionary of the Russian Language" (further SAD), 7) "Толковый словарь русского языка" С.И.Ожегова и Н.Ю.Шведовой (1997) ("Explanatory Dictionary of the Russian Language" by S.I. Ozhegov and N. Y. Shvedova, 8) "Большой толковый словарь русского языка" под ред. С.А.Кузнецова (1998,) ("Large Dictionary of the Russian Language" edited by S. A. Kuznetsov) (hereinafter LED), 9) "Новый словарь русского языка. Толково-словообразовательный" Т.Ф.Ефремовой (2000) ("New Dictionary of the Russian Language. Explanatory and Word-Formative" by T. Efremova) (hereinafter ED); 10) "Толковый словарь русского языка с включением

³ A.P. Evgenieva (ed.), *The Dictionary of the Russian Language: in 4 volumes (1981-1984)*, 2nd edition, Moscow, RUs, Yaz, 2015.

сведений о происхождении слов”, изданный коллективом авторов (2007) “Explanatory Dictionary of the Russian language with the Inclusion of Information about the Origin of Words”, published by the team of authors.

In addition to the descriptive method (in analyzing the materials of explanatory dictionaries, their content), the study also uses linguistic statistics method (for quantitative characteristics of Turkisms) and comparative-contrastive analysis.

Discussion

In contactological lexicography, the complex corpus of Turkisms of the Russian language (based on the material of explanatory dictionaries), according to its concept and structure, \ the number of characteristics reflected in it, is primarily new development and has no analogues.

The task of informatization of linguistic data is urgent at the present stage of the development of linguistics, which is directly connected with the development and use of computer machinery and computer technologies. At the same time, of particular importance are the reference systems, which allow of generalizing, structuring, supplementing various information⁴

Creation of complex corpora, including the materials of dictionaries, is one of the promising areas at present.

The complex corpus of Turkisms of the Russian language is an information resource that reflects the information about the Turkic lexical elements, fixed in the Russian dictionaries of the period. It comprises a description of the Turkic loan words according to the following lexicographic parameters:

- orthography of borrowing (reflecting phonetic and graphic variants);
- phonetic (orthoepic) characteristic;
- grammatical characteristic (belonging to parts of speech, peculiarities of word changing);
- semantic characteristic (explanation);
- stylistic characteristic;
- etymology.

The data of the complex corpus, thus, allow to note the changes associated with the functioning of the layer of Turkic vocabulary within the period from the 18th century till the 21st century, to reveal the features characteristic of this scope of Russian language borrowings.

Results

⁴ A. R. Gizatullina, C. R. Galiullin, "Dictionaries and texts of ancient manuscripts in Kazan region during XVI-XVII centuries: Development prospects," in *Journal of Language and Literature*, VI (2015), Issue 1, p. 267-272; M. V. Oreshkina, "Language and culture: lexicographical description," in *The Third International Conference 'Language, culture, and society in Russian - English studies' Senate House, University of London, 9 - 10 July, / University of London; School of Advanced Studies; Institute of English Studies; Russian Academy of Sciences; Russian Academy of Linguistics*, Moscow, Institute of Foreign Languages, 2012, p. 1-11.

According to our calculations, 16% of Turkic words (about 150 units) have variations of spelling or pronunciation, i.e. they are presented with phonetic-graphic variants in the analyzed dictionaries (this number does not include the accentological variants): *бергамо'т – баргамо'т, калу'ффер – кану'ффер, сайда'к – саада'к* and so on. None of pairs to occur in all lexicographic sources or even dictionaries of the twentieth century has been fixed. Orthographic variants, as a rule, result from the representation in writing of the variation of the vowels *a* and *o* in the pretonic syllables (*карачу'н – корочу'н, карга' – корга', магары'ч – могоары'ч, ата'ра – ата'ра* ad others). In a number of occasions (for 30%), one can say about profound mastering of words, choice or confirmation of one of the variants existing earlier; see, for example (the variant to have been ousted was highlighted in bold): *салта'н – султа'н, сергу'ч – сургу'ч, чевя'ки – чувя'ки* and others. During the analysis, there were revealed the accentological variants, ousted from the Russian language within the analyzed period of time: *бели'берда – белиберда', домра' – до'мра, жемчу'г – же'мчуг* and others. To compare, the gender variants of the loan words under study function without visible changes in the language – the variants, fixed in the dictionaries of the beginning of the 20th century, occur in the sources of the end of the 20th century: *каптан – каптана, мамон – мамона, тут – тута, хабар – хабара, чинар – чинара* and others.

Some Turkic words or their meanings are not reflected in lexicographic publications as time goes by. The reasons for this, in our opinion, are:

1) occurrence of Turkic borrowing words or their individual meanings in the layer of obsolete vocabulary as a result of vanishing of the designated realia or notion – about 20 units (*ронтаь* 'non-Orthodox prayer-house', *селямлик* '1. Parade of the sultan to the mosque. 2. Admission to sultan's Bairam. 3. Man's half of the house typical of the Turkish people', *ерлаш* '<...> 3. *Obsolete*. Blend of different sorts of dry jam, sweets' and others);

2) a restricted scope of the use of certain lexemes (as well as some of their meanings), inclusion or non-inclusion of which into the vocabulary of the dictionary depends on the principle of choosing the material by the compilers (*ергак* 'тулуп из жеребьячьих шкур' / sheepskin coat, *пирог* 'пшеничный хлеб' / wheat bread, *талагай* 'невежа, неуч, boor, ignoramus, тумак ' меховая шапка' / fur cap and others).

The portion of the Turkic words (more than 200 units) acquire new meanings with time. Most of them are recorded for the first time in UD (about 100) and GAD (more than 50): *каланча* 'tall person', *камьш* 'overgrowth with such grass', *лапша* 'spineless, weak-willed man', *табор* 'a large group of people', etc. A considerable part of new lexical-semantic variants of Turkic words arose as a result of metaphorical transfer; the most productive are the models "the name of animal" → the name of the person (*ишак* 'donkey → a person resignedly fulfilling the hardest work'), 'the name of the subject → the name of the person' (*фитиль* 'ribbon in lighting fixtures → thin and very tall man'), "the name of the object → something similar to it" (*шалаш* 'an easy building → something that resembles such construction'), etc. Most of the semantic derivatives resulted from the metonymic transfer are formed by the models "the name of the material → an article made

from it (*собака* 'fur, the skin of such animal → the furs of such animal'), the plant → the product from its fruit (*баклажан* 'plant → the food from the fruits of this plant'), the name of the animal (fish) → the product from it" (*тюлька* 'small fish → canned food from such fish'), etc.

A number of Turkic loan words of the Russian language are gradually included in the layer of common vocabulary (as a consequence of the spread of the corresponding realia), and therefore their lexicographic fixation (*аджика, азу, йогурт*, etc.) becomes possible. It is obvious at the same time that the ousting of the lexical-semantic variant from the language, judging from the data of the dictionaries, is much slower than the appearance of a new meaning.

The involvement of the materials of the corpus also provides the information for identifying and correcting inaccuracies contained in the dictionaries and studies, adjusting and supplementing information and provisions concerning various aspects of the characteristics of the Turkic lexical elements of Russian vocabulary. Thus, the etymological glosses of explanatory dictionaries convey the Tatar prototypical words not always correctly: АЙДА' interjection <...> (Tatar *айда*) in (The Dictionary of the Russian Language: in 4 volumes, 1981) instead of *әйдә* (here and after the Tatar words are given from ⁵; БИРЮ'К, а, м. <...> Tatar *бире*, Turkic *бөрү* – wolf in ⁶ instead of *буре*; ТЮФЯ'К, -а', м. <...> (From Tatar *тууэк*) in (The Dictionary of the Russian Language: in 4 volumes, 1981), ТЮФЯ'К, а, м. (from Tatar *тууэк*) in (Great Explanatory Dictionary of the Russian Dictionary, 1998) instead of *тууэк* and others.

In a number of cases the dictionaries present the meanings of words belonging to one thematic group non-uniformly, they can be given individually or in the context of lexico-semantic variant. For example: ПАША' м. 1. The title of the highest military and civil dignitaries in the former Sultan Turkey <...>. 2. A person who has such title; and КАГА'Н м. 1. The title of the head of state of the ancient Turkic peoples. // A person who had such title. 2. <...>; БАДЬЯ' ф. 1. A wooden or metal wide bucket, narrowed downwards. 2. The amount of smth., being contained in such vessel; and СТАКА'Н м. 1. A glass vessel of cylindrical shape, without handle, serving for drinking. // The amount of substance that can be contained in it <...>.⁷

Conclusion

Thus, the complex corpus Turkisms of the Russian language, compiled on the basis of the information of academic explanatory dictionaries, allows not only of reflecting the specific functioning of this layer of vocabulary, but demonstrates the importance of a structured, systematized dictionary representation of Turkic borrowing words.

⁵ Татар теленең аңлатмалы сүзлеге, Баш ред. Ф.Ә. Ганиев, Казан: Матбугат йорты, 2005.

⁶ S. A. Kuznetsov (ed.), *Great Explanatory Dictionary of the Russian Dictionary*, Sankt Petersburg, Norint, 1998.

⁷ T. F. Efremova (ed.), *New Dictionary of the Russian Language: Explanatory- Word-Formative: more than 136000 dictionary entries: in 2 volumes*, Moscow, Rus. Yaz, 2000.

Complex Corpus Of Turkisms Of The Russian Language, Astra Salvensis, V (2017), no. 10,
p. 145-150

The inclusion of the complex corpus of Turkisms in the international computer network will enable to fulfil operative information exchange between scientific collectives; will provide a reliable factual basis for various studies in this field. In general, it will serve for enrichment and informatization of the theory of language contacts in a multinational state of Turkic-Russian language contacts.