Representation of Dictionaries in the Russian Collocations Database*

Maria Khokhlova
m.khokhlova@spbu.ru

Saint-Petersburg State University
Saint-Petersburg, Russian Federation

Abstract

The study of collocability is an important task and is still highly relevant in linguistics. The present paper discusses the issue of collocability and collocations in a number of Russian dictionaries (the Dictionary of the Russian Language, Dictionary of Set Verb-Noun Phrases, the Dictionary of Russian Idiomatics and the Dictionary of Collocations). These dictionaries were analyzed and used to create a Russian collocations database which includes both verified or dictionary collocations, as well as data from text corpora. At present, the database includes about 18,500 collocations.

Keywords: *collocations, lexicographic sources, lexical database, Russian language, dictionaries, corpora*

1   Introduction

Within the last decades we have witnessed a movement from printed dictionaries to electronic ones accessible via web, desktop or mobile applications. Nowadays the latter are not digital copies of their printed "ancestors" as it was true at the beginning but self-dependent reference books (if we use this word in a new sense). Having a solid theoretical background lexicography quickly mastered various tools that could improve the process of dictionary compiling. The lexicography of the 21st century can be fairly called smart as its methodology now successfully combines both theoretical and computational approaches.

Russian dictionaries have a long tradition but when it comes to computational lexicography it is long overdue. It can be partially explained by the historical events of the Soviet and post-Soviet era that had so much influence on the selection of lexis and its definitions and also by the approach of Russian lexicographers to describing language that differs from the one of their Western colleagues. Russian language does not belong to less resourced ones but there is still a need in up-to-date tools and hence every effort in this direction should be welcomed.

The paper is organized as follows. The Introduction explains the motivation of the project. Section 2 describes the notion of collocation in Russian linguistics and how it is shown in the database. The next two sections give an overview of the printed and electronic dictionaries used as a source for the database. Section 5 discusses the processing of the dictionaries for the database. The last section concludes the paper and proposes plans for future work.

2   Collocations in Russian

Our project deals with the process of building a database that will represent information on collocability in Russian from dictionaries and corpora [Khokhlova, 2018]. When building a database we had to answer the following question: what kind of collocation should be considered as an appropriate item for the database. And here we need to analyse existing lexicographic sources.

Although the term "collocation" appeared in the Russian language long ago [Akhmanova, 1966], it is not generally recognized by Russian linguists and even is absent in the Russian Linguistic Dictionary [Yartseva, 1990]. For a long time there was no agreement among scholars how to call such lexical units; cf. "set verbal-noun expressions" [Deribas, 1983], "analytic lexical phrases" [Teliya, 1996], etc. It can be seen as a reason why collocations dictionaries are not so well-known compared to English ones. The majority of modern authors understand under a collocation a statistically set phrase. Collocations can be put between free phrases and idioms on a scale of phrases. The monograph [Borisova, 1995a] has proved to be the first work in Russian linguistics that officially "introduced" this term in Russian studies. According to the author one of the key properties of a collocation is "the impossibility of prediction of such combinations on the basis of meanings of their components" [Borisova, 1995a: 13]. In the "Meaning–Text" model [Mel'chuk, 1974] collocations are considered as a subclass of more extensive class of set phrases, or phrasemes.

Following Testelets's approach [Testelets, 2001] we interpret collocability of a lexical unit extremely widely and understand it as the ability to connect with other lexical units. Thus, we will consider phrases of different degrees of stability (from collocations to idioms and phraseological units with non-compositionality). For example, according to Telia's [Telia, 1996] classification: idioms (*rabochaja loshad'* 'working horse'), phraseological units (*teljachij vostorg* 'foolish enthusiasm'), fixed expressions (*vsegokhoroshego* 'all the best'), cliche (*minutu vnimanija* 'minute of attention'). Also here we add terms (*kontrol'naja palata* 'control chamber') and set phrases (*podvergnut'sja deformatsii* 'undergo deformation').

We came to the conclusion that it could be a certain problem to find an appropriate dictionary that describes collocations and is large enough in order to both list high frequency word combinations and give a sufficient number of examples for them. The dictionaries differ in their representation of information on collocability and in coverage of the examples. Data on collocability can be found in various dictionaries and sources and in different parts of the dictionary entries, e.g. quotes, speeches or specialized zones. The latter can list set expressions representing a certain lexical unit. Dictionary entries should be structured in order to make possible further collocation extraction. And here lies a certain problem as the majority of Russian dictionaries (if can ever be found in digital form) represent scanned copies of printed versions without OCR. Their recognition is then followed by further division of entries into a structured format that takes into account possible grammatical information, examples, quotations etc.

3 Printed Russian Dictionaries

For our project we made a survey of various dictionaries that can be used as sources of collocations. In the present paper we will dwell on those of them that have been already processed during the project.

The explanatory dictionaries implement various ways to represent the information on combinatorial restrictions but they share much in common. They use different fonts and labels in their entries. Set phrases, multiword expressions and collocations can be described not only in special sections of the entries but also in the examples, sayings and quotations.

The "Dictionary of the Russian Language" [DRL, 1981–1984] followed the "Dictionary of Contemporary Literary Russian Language" [DCLRL, 1948–1965] (one of the most important projects in Soviet lexicography started before the Second World War but its first volume appeared only in 1950) and comprises more than 80,000 lexical items resulting in 4 volumes. Its first edition was published in 1957–1961, the revised version came out in 1981–1984 and soon became popular among linguists and other scholars.

Compared to [DCLRL, 1948-1965] the given dictionary is more concise and we can see it on the excerpt from its online version (Fig. 1). Here collocations do not have a special mark-up (they are given as examples in italics) but the phraseological units do.

Traditionally dictionaries that describe collocability are intended at language learners and teachers. Below we will discuss Russian dictionaries that represent collocability and can be to a certain degree thus called collocations dictionaries. Also we can name several glossaries that were compiled for language learners, some of them printed in a small number of copies in the republics of the Soviet Union and hence are now rare books (for example, [Anisimova et al., 1975]).

The "Dictionary of Set Verb-Noun Phrases in Russian" [Deribas, 1983] was intended for students of Russian and comprises 5,197 collocations for 744 verbs and 1,345 nounsin total. The authors put collocations between free phrases and phraseological units: *opravdyvat' (opravdat') ozhidaniya* 'to confirm expectations'; *pitat' uvazheniye* 'to respect'; *chitat' lektsiyu* 'to hold a lecture' etc. The majority of phrases consist of bigrams including verbs and nouns as direct or indirect objects. The authors emphasize the dictionary is focused on language learners and does not provide any definitions or explanations merely listing the collocations of literary

РЕШЕ́НИЕ, -я, ср.

1. Действие по знач. глаг. реши́ть[1]—реша́ть (в 1, 2 и 3 знач.). *Решение вопроса. Решение дела. Участвовать в решении чьей-л. судьбы.* □ *Мне нужно было не только самому решать подобные задачи, — но и объяснять, как решения производятся в уме.* М. Павлов, Воспоминания металлурга. *Самое главное — это в понять глубоко и безошибочно, — не надо торопиться со своими решениями.* Воронин, Чужое письмо. *Мы не нашли еще современного ключа к сценическому решению комедии Аристофана и Плавта.* Акимов, О театре.

2. То, что является результатом обдумывания каких-л. действий, обдуманное намерение сделать что-л. *[Талиенок:] Я не изменил решения. Я никуда не еду.* Леонов, Нашествие. *За всех вот этих хороших людей, одетых в солдатскую форму, был он в ответе. От принятого им решения во многом зависела их судьба.* М. Алексеев, Солдаты. ‖ Заключение, вывод, мнение, разрешающее что-л. *Интересное решение проблемы.* □ *Я вкратце объяснил командирам создавшееся положение и сообщил им свое решение: 45-й и 62-й дивизиям во что бы то ни стало прорвать вражеское кольцо и форсировать Десну.* Москаленко, На юго-западном направлении.

3. То, что принято в результате обсуждения; постановление. *Решение суда. Решение съезда.* □ *[Петр] продиктовал Ибрагиму несколько указов и решений.* Пушкин, Арап Петра Великого. *Бригада созвал совещание, которое вынесло решение: вторую линию [нефтепровода] строить!* Ажаев, Далеко от Москвы.

4. Способ воплощения чего-л. *Я не вижу пока таких же удачных решений внутри павильона, какие Жолтовский нашел в понял в его строительном чертеже.* Шагинян, Брюссельская Всемирная выставка 1958 г. *Он еще лист. Ему, эрудиту и знатоку моторов, в одну минуту стало ясно: у него в руках решение, которое он почти отчаялся найти.* Бек, Талант.

5. Способ нахождения ответа к задаче, кроссворду и т. п., а также ответ к задаче, к кроссворду и т. п. *Легкое решение. Правильное решение. Переписать решение задачи.* □ *Екатерина Сергеевна стала объяснять второму классу решение какой-то арифметической задачи.* Исаковский, На Ельнинской земле.

Figure 1: Headword *reshenije* 'decision' in the "Dictionary of the Russian Language"

language. There are two lists sorted alphabetically according to verbs and nouns. The verbal entry (Fig. 2) includes the infinitive of main imperfective form and if it has the corresponding perfective form gives it in bold italics in parentheses. The collocates are used in their word forms followed by the interrogative pronouns (with or without prepositions) and cases that indicate possible distributions of collocations (a kind of a valency frame).

The nominal entry (Fig. 3) lists all the collocations that include the noun. Verbs with possible close or opposite meanings are marked as synonyms and antonyms and are given in parentheses. As we can see the nominal part of the dictionary was not so elaborated as the verbal one.

The monograph [Borisova, 1995a] initially published in Austria proved to be the first work that was devoted to Russian collocations. The preface mentions students and teachers of Russian as a foreign language as the main audience of the reference book. In continuation in the same year the author compiled also the "Dictionary of Russian Collocations with English-Russian Dictionary of Keywords" [Borisova, 1995b]. In the dictionary collocations are structured according to their semantics with numbers that correspond to lexical functions and are represented graphically in capital letters (Fig. 4), e.g. *tverdoje reshenije* 'firm decision'.

Nouns are the most frequent headwords in the dictionary followed by verbs, adjectives and phrases. For example, the word list of the dictionary [Borisova, 1995b] has only 512 items. The dictionary also has an English-Russian list of keywords aimed at foreign users.

Figure 2: Headword *prinimat'* 'take, accept' in [Debibas, 1983]

Figure 3: Headwords *retsenzija* 'review', *retsept* 'prescription', *rech* 'speech', *reshenije* 'decision' in [Deribas, 1983]

4 Electronic Russian Dictionaries and Databases

At the moment, there is no such a system for the Russian language that would allow researchers to obtain not only corpus data, but also "reference" (verified) information on the vocabulary and behavior of lexical units, combined due to different characteristics. At the same time, there is a number of unique and valuable lexicographic projects that describe collocability of units, although in different ways.

There are online resources (for example, slovari.ru or gramota.ru), which provide an opportunity to learn the meaning of a word and view dictionary entries containing information about collocability. But at the same time, users may have difficulty reading the entries, since collocability can be presented both in different parts of the entries and directly in the quotes themselves. When it comes to phrases, a complete dictionary entry is not always necessary, but there is a need in examples of real data, properly described and accompanied by an assessment of the correctness or frequency of use in speech.

The Russian National Corpus [RNC, 2003-2019] has a number of tools (n-grams search with statistical evaluation, lists of set phrases and fixed expressions, and lexical graphs) and dictionaries based on it. Here we can name the Dictionary of Russian Abstract Nouns' Verbal Collocability [Biryuk, Gusev and Kalinina, 2008]. It presents information for over 10,000 phrases of the following structures: 1) noun+verb; 2) verb+noun; 3) verb+adjective+noun. The dictionary uses lexical functions for describing and classifying collocations and their senses. The collocation candidates were initially extracted from the parsed subcorpus of RNC. The nouns occupied one of the following syntactic positions: 1) direct object of a transitive verb; 2) indirect object of a transitive verb; 3) subject of an intransitive verb. The Dictionary of Russian Idiomatic Expressions [Kustova, 2008] presents word combinations with the meaning of a high intensity. The author describes lexical items with restricted collocability, i.e. idiomatic expressions that should be learned by native and non-native speakers. The initial word list was based on RNC and printed dictionaries. One can find phraseological units (*kruglyj sirota* 'orphan'), collocations (*plakat' navzryd* 'to sob violently'), idiomatic expressions (*glubokoo gorchen* 'deeply
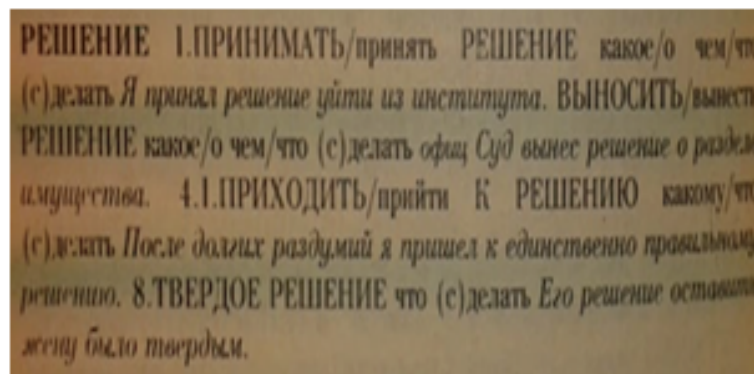
Figure 4: Headword *reshenije* 'decision' in [Borisova, 1995b]

disappointed') and semantically motivated free phrases (*chrezvychajno ogorchen* 'extremely disappointed').

5   Database

The dictionaries mentioned in the previous section were used as a source for the database. Below we present the pipeline for data processing and discuss the principles of the database.

The analysis of the dictionaries suggests that there is a need in a unified format that can be used for describing data. An entry has the following characteristics:

- a collocation;

- a syntactic model;

- a dictionary index (if applicable);

- references to the dictionaries (if applicable);

- frequencies in corpora (in ipm);

- references to the dictionaries;

- values of the association measures;

- visualization.

The database includes two kinds of collocations, i.e. dictionary and statistical ones. The former present in various lexicographic sources whereas the latter can be extracted automatically from corpora. In order to obtain data on co-occurrences in the Russian language, we process AraneumRussicum Maximum corpus (about 15 billion words), which was created automatically and is based on web texts of different genres being one of the largest collection of Russian texts [Benko andZakharov, 2016]. We use a statistical approach for automatic extraction of word combinations from corpora that implied several association measures (t-score, MI, log-likelihood).

At present the data from the described dictionaries are available in the database. We analyzed dictionary entries and extracted collocation candidates either from phraseological sections or as separate items written with a special type. Altogether the list comprises about 18,500 collocations. The initial merged lists were longer but a certain part of the dictionaries word lists overlapped. These collocations received the corresponding index, i.e. the number of the dictionaries they were presented in. The index indicates the given items are highly reproducible in speech and can be used by language learners.

Preprocessing involved also morphological analysis in order to present collocations in their canonical form but grammatical information was also preserved.

5.1 The "Dictionary of the Russian Language"

Altogether we extracted 11,210 phraseological units that were marked with a special diamond '◇' symbol. The total amount of the headwords that had phraseological units in their entries was 5,955 (about 7.4% from the entire grams (*makovoj rosinki v rot ne brat'* 'starving'). In the database we focused on the following models from the dictionary :

- adjective + noun.

Examples: *bolotnyj gaz* 'marsh gas', *morskay amilja* 'nautical mile', *gomericheskij khokhot* 'Homeric laughter', *zolotaj asvad'ba* 'golden wedding', *zhivoj um* 'nimble mind'.

- verb + noun / verb + preposition + noun.

Examples: *bit' kartu* 'to cover a card', *boltat' jazykom* 'to jabber away', *nesti otvetstvennost'* 'to be responsible', *oblechd overijem* 'to trust', *stavit' tochku* 'to end'.

Among other models found in the candidates list we can name the following ones: 1) noun + noun (*kniga pocheta* 'book of honorable guests'); 2) preposition + noun (*bez umolku* 'nonstop'); 3) pronoun + noun (*nechego skazat'* 'nothing to say').

The phraseological units found in the dictionary often enumerate certain semantic groups lexical items of which collocate with a keyword. Examples: *v pylu (srazhen'ja, bitvy, sporaetc)* 'heat of the (fight, battle, debate)', *v rassrochku (kupit', prodat')* 'in instalments (to buy, to sell)', *v storonu (skazat', proiznesti)* 'aside (to say, to utter)'. We considered each unit with an appropriate word as a separate phrase. Such an approach enabled us to enrich the data and also in future we plan to use special semantic tags to make the information clear for language learners.

Collocations can be also embedded into longer phraseological units and thus were treated separately. The dictionary entries do not represent the synonymic expressions separately but by commas or conjunctions. E.g. *v khristianskij vid privesti* or *khristianskij vid pridat* 'ironic, to put in order'. In the database thus we have two different records with verbs and one record with noun phrase (*khristianskij vid*).

5.2 The "Dictionary of Russian Collocations"

The initial preprocessing involved OCR procedure and extraction of the collocations written in capital letters. This resulted in a total sum of 3,058 collocation candidates. We also extracted collocations from quotations that were not highlighted in the entries and marked them with the asterisk '*' symbol. These phrases vary in their fixedness but they can be an important source of information. Hence additional 232 collocations were added to the database. Examples: *ozhivlennajabeseda* 'lively talk', *vypusknojekzamen* 'final examination', *proizvoditel'nost' truda* 'workforce productivity'. Polysemic headwords marked in the dictionary with digits were preserved in the database as separate entries. Examples: *otvlekat' vnimanije 1* 'to distract attention' (meaning: 'showing interest') vs *okazyvat' vnimanije 2* 'to give attention' (meaning 'care').

In the dictionary 2,044 verb pairs (imperfective and perfective aspects) are given via slash (e.g. *vnosit' /vnesti jasnost'* 'to clarify something'). They were separated in order to make two records resulting in total 4,088 verb phrases.

The dictionary lists a large number of collocations that slightly differs in their meaning and can be called synonyms to a certain degree. They vary either in wordforms (case or number) or in prepositions. Examples: *nakhodit'sja vo vlasti* vs *nakhodit'sja pod vlast'ju* 'to be in smb's power', *ostavit' pamjat' o sebevs ostavit' pamjat' po sebe* "to leave memories'.

5.3 The "Dictionary of Russian Idiomatic Expressions"

The preprocessing of the data extracted from the given electronic dictionary was not so elaborated as it was the case with other dictionaries. The following models are present in the dictionary:

- adjective + noun.

Examples: *glubokaja blagodarnost'* 'deep gratitude', *krupnyj uspekh* 'enormous success', *nadezhnaja opora* 'firm support', *ogromnyje perspektivy* 'exciting prospects', *jaryj storonnik* 'stickler'.

- adverb + verb.

Examples: *aktivno obsuzhdat'* 'to discuss actively', *bezgranichno verit'* 'to trust implicitly', *vovsju moch nestis'* 'to rush headlong'.

- adverb + adverb.

Examples: *bezumno interesno* 'extremely interesting', *nesterpimo dolgo* 'unbearably long', *sovsem nedavno* 'just recently'.

- adverb + predicative.

Examples: *absoljutno ochevidno* 'absolutely obvious', *zhutko prijatno* 'extremely pleasant', *iskljuchitel'no vazhno* 'exceptionally important'.

- adverb + adjective.

Examples: *gluboko porjadochnyj* 'totally honest', *zhutko bogatyj* 'awfully rich', *otchajanno smelyj* 'extraordinarily courageous', *khorosho znakomyj* 'well-known', *jarkovyrazhennyj* 'strongly marked'.

- particle + noun.

Example: *prjamochudo* 'really badly'.

We confined ourselves to noun phrases and lemmatized them. The total amount was about 7,000 collocations.

5.4   The "Dictionary of Set Verb-Noun Phrases in Russian"

The dictionary comprises more than 3,770 collocations with perfective and imperfective verb pairs and 383 with only imperfective verb forms. After processing the dictionary data, excluding prepositional phrases and representing phrases with perfective and imperfective verbal forms as different pairs we had a list of 7,923 items.

5.5   Dictionaries: Results

The volume of the verified (dictionary) collocations depends on the volume of the dictionaries that are used. As it was mentioned above dictionaries' volume is not sufficient enough to describe vaster groups of lexis and hence to give a broader coverage that could be comparable to a word list of an explanatory dictionary (it often counts several thousand).

Table 1 demonstrates the number of collocation candidates (noun and verb phrases) from four dictionaries. It can be noticed that an electronic dictionary based on corpus data has larger volume compared to a printed dictionary.

Table 1: Statistics of collocation in the dictionaries

Dictionary	Number of collocations	
	adjective + noun	verb + noun
Dictionary of the Russian Language	3,243	5,23
Dictionary of Russian Collocations	613	4,177
Dictionary of Russian Idiomatics	6,962	0
Dictionary of Set Verb-Noun Phrases in Russian	0	7,923

The Dictionary of the Russian Language [1981-1984] and [Kustova, 2008] show the lowest overlap. The merged list had 10,205 positions and only 59 of them were described in both dictionaries (less than 1% of the whole list). It can be explained by the fact that [Kustova, 2008] describes collocations with denoting high intensity whereas the Dictionary of the Russian Language [1981-1984] aims at a comprehensive representation of the lexicon.

7

6   Conclusion

We made an overview of the database that includes data on collocability from various lexicographic sources. It is already available online (http://collocations.spbu.ru).

Even though Russian cannot be counted as a less resourced language there is a certain need in online tools that can be used both in research and language learning. The low overlap between the dictionaries (about 15%) suggests that they describe different lexical units. The Dictionary of the Russian Language [DRL, 1981-1984] puts emphasis on phraseological units whereas other lexicographic sources pay attention to collocations. There are also other syntactic models that should be taken into account as well as examples from quotations. There is a need in further processing of other dictionaries.

It is planned to open free access to the system for a possible assessment by the users of the degree of stability of the structures found in corpora. This will allow getting an interpretation of the data by the speakers of the language and can be used in further improving the resource and other tasks, for example, when creating a specialized dictionary or in systems using machine learning. Also, the amount of illustrative material and information on collocability presented in lexicographic sources (and, accordingly, in the "gold standard") may indicate the frequency of use of the unit and correlate with it. This may be required when developing teaching and learning materials for students of the Russian language.

Acknowledgements

References

[Khokhlova, 2018] Khokhlova M. (2018) Building a Gold Standard for a Russian Collocations Database. In Proceedings of the XVIII EURALEX International Congress: Lexicography in Global Contexts. Ljubljana. Pp. 863–869.

[Akhmanova, 1966] Akhmanova O.S. (ed.). (1966) Dictionary of Linguistic Terms. Moscow.(In Rus.) = Akhmanova O.S.Slovar' lingvisticheskikhterminov. Moskva, 1966.

[Yartseva, 1990] Yartseva V.N., (ed.). (1990) Linguistic Encyclopedic Dictionary. Moscow.(In Rus.) = Yartseva V.N., (red.). Lingvisticheskijenciklopedicheskijslovar'. Moskva, 1990.

[Deribas, 1983] Deribas V.M., (1983) Verb-Noun Collocations in Russian. Moscow: Russian language.(In Rus.) = Deribas V.M. Ustojchivyjeglagol'no-imennyjeslovosochetanijarusskogojazyka. Moskva: Russkijjazyk, 1983.

[Teliya, 1996] Teliya V.N. (1996) Russian Phraseology: Semantic, Pragmatic and Cultural Aspects. Moscow.(In Rus.) = Teliya V.N. Russkajafrazeologija: semanticheskij, pragmaticheskijilingvokul'torologicheskijaspekty. Moskva, 1996.

[Borisova, 1995a] Borisova E.G. (1995a) Collocations. What is it and how to study them. Klagenfurt.(In Rus.) =Borisova E.G. Kollokatsii. Chtoetotakojeikakikhizuchat'. Klagenfurt, 1995a.

[Mel'chuk, 1974] Mel'chuk I.A. (1974) The experience of the theory of linguistic models "Meaning-Text". Moscow.(In Rus.) =Mel'chuk I.A.Opytteoriilingvisticheskikhmodelej" Smysl-Tekst". Moskva, 1974.

[Testelets, 2001] Testelets Y. (2001) Introduction to the General Syntax. Moscow: RGGU.(In Rus.) =Testelets Y. Vvedeniye v obshchiysintaksis. Moskva: RGGU, 2001.

[DRL, 1981-1984] Dictionary of the Russian Language (1981–1984) Yevgen'yeva A. P. (ed.-in-chief). Vol. 1–4, 2nd edition, revised and supplemented. Moscow: Russkijjazyk.(In Rus.) =Slovar' russkogojazyka v 4 tomakh. T. 1-4, 2-e izdanije. Moskva: Russkijjazyk, 1981-1984.

[DCLRL, 1948-1965] Dictionary of Contemporary Literary Russian Language (1948–1965)Chernyshev V.I. (ed.). Moscow-Leningrad: Izd-voAkademiinauk SSSR.(In Rus.) =Slovar' sovremennogorusskogoliteraturnogoyazyka v 17 tomakh. Moskva-Leningrad: Izd-voAkademiinauk SSSR, 1948-1965.

[Anisimova et al., 1975] Anisimova T. I., Ivanova Z. Je. and Ul'janko R. V. (1975) A Textbook on Lexical Collocability in Russian. A Dictionary..Pleshhenko, T. P. and Sakovec, L. F. (eds.). Minsk.(In Rus.) =Anisimova T. I., Ivanova Z. Je. and Ul'janko R. V. Posobie po leksicheskojsochetaemostislovrusskogojazyka. Slovar'-spravochnik. Pleshhenko, T. P. and Sakovec, L. F. (red.). Minsk,1975.

[Borisova, 1995b] Borisova E. G. (1995b) A Word in a Text. A Dictionary of Russian Collocations with English-Russian Dictionary of Keywords. Moscow.(In Rus.) = Borisova E. G. Slovo v tekste. Slovar' kollokatsij (ustojchivykhsochetanij) russkogojazyka s anglo-russkimslovaremkljuchevykhslov. Moskva, 1995.

[RNC, 2003-2019] Russian National Corpus, (2003-2019).(In Rus.) = Natsional'nyjkorpusrusskogojazyka. Available athttp://ruscorpora.ru.

[Biriuk, Gusev and Kalinina 2008] Biriuk O. L., Gusev V. Ju., Kalinina E. Ju. (2008) Dictionary of Russian Abstract Nouns' Verbal Collocability. A Dictionary based on the Russian National Corpus.(In Rus.) = Slovar' Glagol'nojSochetaemostiNepredmetnykhImenRusskogoJazyka. Slovar' naosnoveNatsional'nogoKorpusaRusskogoJazyka. Available at http://dict.ruslang.ru

[Kustova, 2008] Kustova G.I. (2008). Dictionary of Russian Idiomatic Expressions.(In Rus.) = Slovar' russkojidiomatiki. Sochetanijaslov so znachenijemvysokojstepeni. Available at http://dict.ruslang.ru

[Benko and Zakharov, 2016] Benko V., Zakharov V. (2016) Very large Russian corpora: New opportunities and new challenges.InComputational Linguistics and Intellectual Technologies 15:22.Pp. 79–93. Moscow: Izd-vo RGGU.