Distribution of Quasi-Synonyms in Thesaurus for Natural Language Processing

Valery Solovyev[a], Natalia Loukachevitch[b] and Vladimir Bochkarev[a]

[a] *Kazan Federal University, Kremlevskaya Street 18, 420008 Kazan, Russia*
[b] *Lomonosov Moscow State University, Leninskie Gory 1, 119234, Moscow Russia*

Abstract

The paper examines the relative positions of quasi-synonyms in the RuWordNet thesaurus. Quasi-synonyms (analogs) are taken from the dictionary of synonyms by Yu.D. Apresyan. The aim of the research is to test the RuWordNet thesaurus. It is shown that in more than 90% of cases quasi-synonyms are located in the thesaurus at a small semantic distance from each other. This is a confirmation of the good organization of RuWordNet. The analysis of the remaining 10% cases made it possible to formulate corrections for improving the thesaurus.

Keywords

thesaurus, natural language processing, quasi-synonym

1. Introduction

Modern electronic thesauri such as Princeton WordNet [1] and similar resources in other languages, including the Russian language thesaurus RuWordNet [2], are one of the most demanded computer resources in the tasks of natural language processing (NLP). Such resources are built on the basis of synsets – sets of synonyms linked by semantic relations of hypo-hypernymy, antonymy and some others [3]. A detailed description of the RuWordNet thesaurus can be found in [4]. For practical application, thesauri must include tens of thousands of words and expressions, which requires significant efforts for their development and maintenance, including testing the quality of their formalized descriptions. In this paper, we propose an approach to testing existing thesauri based on the analysis of quasi-synonyms.

The general idea is as follows. Almost all existing NLP thesauri are based on the concept of synonymy and do not have means to represent quasi-synonymy. Meanwhile, this concept also reflects the semantic similarity of words and therefore it should be closely related to the structure of thesauri. In this article, we analyze this relationship on the basis of the RuWordNet thesaurus. The consideration of the semantic distance in the thesaurus between quasi-synonyms makes it possible to assess the quality of the thesaurus and identify possible gaps in its structure.

2. Related Work

Traditionally, quasi-synonyms are words that are close in meaning, but not interchangeable in all contexts. In approximately the same sense, Yu.D. Apresyan uses the term *analogs* [5]: analogs are words of the same part of speech as a source word, the senses of which substantially intersect with the general sense of a given set of synonyms, although they do not reach the degree of synonymic closeness. In the English language literature, the terms near synonyms [6], plesionyms [7] and almost synonyms [8] are also used. We do not distinguish among all of these terms, assuming that they mean

Proceedings of the Linguistic Forum 2020: Language and Artificial Intelligence, November 12-14, 2020, Moscow, Russia
EMAIL: maki.solovyev@mail.ru (1); louk_nat@mail.ru (2); vbochkarev@mail.ru (3)
ORCID: 0000-0003-4692-2564 (1); 0000-0002-1883-4121 (2); 0000-0001-8792-1491 (3)

the same phenomenon. Quasi-synonyms often include hyponyms and hypernyms of a given word, as well as cohyponyms - words that have a common hypernym. However, quasi-synonymy is still a broader concept [5]. Quasi-synonymous relations are quite important in the tasks of natural language processing. In [9] the authors discuss the task of extracting multiword quasi-synonyms for multiword terms. In [10] the importance of this concept for the expansion of search queries is noted. However, quasi-synonyms have not been studied in as much detail as synonyms. In particular, there are no quasi-synonym dictionaries for the Russian language. Quasi-synonyms are not reflected in the Princeton WordNet and RuWordNet.

It seems natural that quasi-synonyms in thesauri, including RuWordNet, should be located at a short distance from each other. Distance means, as normal, the length of the path in the semantic relationship graph of a thesaurus. It is shown in [11] that the average distance between words in RuWordNet differs for different parts of speech. For adjectives, it is equal to 5.3, for nouns – 5.5, for verbs – 8.7.

We use quasi-synonyms in this article to test RuWordNet[1]. As far as we know, such use of quasi-synonyms for correcting the thesauri structure has not been described earlier in the literature.

The following two main approaches are used in lexical semantics, when studying synonyms and other semantically related words: comparison of dictionary definitions, including component analysis [12], and interchangeability in various contexts. The second approach is the main one when defining synonymy, although, for example, in the dictionary of synonyms [13], the main attention is paid to definitions and highlighting very subtle differential semes.

The interchangeability approach appears not as effective when studying quasi-synonyms. For example, quasi-synonyms are often referred to hypernyms (*chair – furniture*) and co-hyponyms (*chair – bookcase*). They will not be interchangeable in most contexts*: an upholstered furniture store,* but not *an upholstered chair store,* or *he leaned back in an armchair,* but not *he leaned back in a bookcase*. Thus, we rely in this study on the interpretation of words in dictionaries and the component analysis.

Among other approaches to verification and extension of thesauri, the corpus approach proposed in [14] should be noted as well. It consists in identifying words which are close in any large corpus (according to one of the standard approaches to determining the semantic similarity of words), but distant in the thesaurus. The work [15] has been carried out in the framework of this approach. The comparison of the RuWordNet thesaurus with the existing dictionaries of synonyms is carried out in [16].

3. Methods and Data

Although, as stated above, there is no complete dictionary of quasi-synonyms, however, the dictionary [13] lists all analogues for each of the 354 synonym sets described in it. In accordance with this description, we have compiled a list of pairs <word, its analogue>. The list is available at https://kpfu.ru/kompleksnyj-analiz-struktury-i-soderzhaniya-366287.html .

The dictionary [13] covers only a small part of the Russian lexicon. Therefore, the use of the compiled list of analogues is not sufficient for an exhaustive study of the RuWordNet structure. Our goal is rather to identify and classify possible types of errors or problem areas in the structure of the thesaurus. In this paper, we will limit ourselves to considering two parts of speech - adjectives and verbs. The analysis of analogous nouns was started in [17]. Our list contains 1601 verb pairs and 685 adjective pairs. Some of the words contained in the pairs are not included in the thesaurus. After removing them, 1410 pairs of verbs and 558 pairs of adjectives still remain.

We calculated distances between words in RuWordNet, as well as generated all possible paths between the words in the pairs. The list of paths is available at https://kpfu.ru/kompleksnyj-analiz-struktury-i-soderzhaniya-366287.html. The shortest paths were found through the well-known breadth-first search algorithm. An adjacency matrix was constructed at the first stage based on the

[1] RuWordNet: ruwordnet.ru

data of semantic relationships of the types of our interest extracted from RuWordNet. Neighborhoods of increasing radius for a word A were sequentially constructed in order to find the length of the shortest path from a word A to a word B. The calculations stopped when, at a certain radius, the vertex B fell into the resulting neighborhood. Figure 1 shows the scheme for calculating the distance between words A and B. The circles symbolically indicate the neighborhood of the word A, corresponding to a distance 1, 2 and 3, respectively.

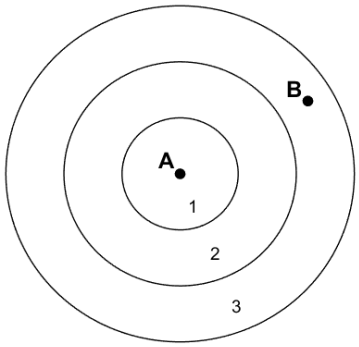

Figure 1. Scheme for calculating the distance between words A and B

The average distance between quasi-synonymous adjectives is 2.43, and between quasi-synonymous verbs – 2.27, which, as expected, is significantly less than the average distance between words of the corresponding parts of speech. Table 1 shows the distribution of pairs by distances – the number of quasi-synonymous pairs for each distance.

Table 1
Number of analogous pairs depending on the distance

Distance	Number of adjective pairs	Percent	Number of verb pairs	Percent
1	171	30.65	392	27.80
2	150	26.88	447	31.70
3	122	21.86	386	27.38
4	63	11.29	170	12.06
5	45	8.07	13	0.92
6	7	1.25	2	0.14

On the basis of some empirical studies, distances between words in the thesaurus, which are less than or equal to 4, are interpreted in [14] as short. Apparently, most of quasi-synonymous pairs are separated by short distances in the thesaurus. For pairs with a distance of 5 and 6, you can expect to find certain gaps in the structure of the thesaurus. Although it is not always necessary, it is still possible that the formalism adopted in the construction of the thesaurus does not capture some of the nuances in semantics; and within the framework of the given formalism it is impossible to obtain short paths by natural additions. The program we developed for further analysis generates all possible shortest path length between pairs of quasi-synonyms. 138505 paths have been generated for verb pairs, and 10837 – for adjectives.

Several examples of detected paths are shown in Tables 2 and 3. Table 2 presents examples of pairs with short path lengths between them, and Table 3 shows examples of word pairs with long paths.

Table 2
Examples of quasi-synonymic pairs with the distance of not exceeding 4 in RuWordNet

Word	Step 1	Step 2	Step 3	Step 4
безлюдный ‘uninhabited’	пустой ‘desolate’			
безлюдный ‘uninhabited’	малолюдный ‘poorly populated’	уединенный ‘lonely’		
храбрый ‘brave’	отчаянный ‘desperate’	удалой ‘bold’	бравый ‘ dashing’	
беспокоить ‘disturb’	надоедать ‘bother’	раздражить ‘irritate’	нервировать ‘annoy’	
хитрый ‘crafty’	непростой ‘intricate’	непонятный ‘unclear’	неопределенный ‘obscure’	уклончивый ‘elusive’

All semantic relations in Table 2 are those of hypo-hypernymy relations and appear to be natural. In contrast to them, Table 3 shows examples of the shortest path lengths between pairs of quasi-synonyms, which are too long and do not reflect the degree of their semantic proximity. The problematic quasi-synonyms connected with long paths will be considered in the next section.

Table 3
Examples of quasi-synonymic pairs with the distance of more than 4 in RuWordNet

Word	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6
коварный ‘insidious’	ролевой ‘thematic’	антиинфляци-онный anti-inflation	качественный qualitative	неясный obscure	путаный confused	каверзный tricky
ладить ‘agree’	знавать ‘used to know’	восторгаться admire	делаться become	измениться alter	подстраивать-ся tune up	подладиться adjust oneself to
каяться ‘repent’	виниться ‘plead guilty’	сделать признание ‘confess of’’	открываться ‘declare’	прийти ‘come’	успокаиваться ‘calm down’	облегчить душу ‘relieve soul’
отчетливый ‘distinct’	ясность ‘clarity’	оценка ‘evaluation’	отношение ‘relation’	несовпадение ‘discordancy’	контраст цвета’ contract color’	контрастный ‘contrastive’
гостеприим-ный ‘hospitable’	радушный' ‘cordial’	дружелюбный ‘amiable’	доброжелатель-ный ‘considerate’	межличностный ‘interpersonal’	контактный ‘sociable’	компанейский ‘companion-able’
громкий ‘loud’	качествен-ный ‘qualitative’	обычный ‘usual’	общераспростра ненный ‘widespread’	известнейший ‘acclaimed’	знаменитый ‘famous’	

4. Linguistic analysis of the data

Let us consider the examples from Table 3.

1. Коварный – каверзный (‘insidious’ – ‘tricky’)

RuWordNet contains 50 paths of the length 6 between these words. As an example, one of the paths is shown in Table 3. It is noteworthy that the path goes through words that have little in common with the original ones. *Qualitativ*e – is a category of adjectives that includes a very large number of words. All other 49 paths also pass through words with a very general meaning or through words, which semantically are very far from them. At the same time, the words *коварный* (‘*insidious’)* and *каверзный (‘tricky’)* have clearly a lot in common. At the very least, they have a common seme of negativity, which has been lost along the way. The connection of these adjectives with human actions is also lost. It seems that here are two reasons for this – common and specific. A common reason is that there are no satisfactory classifications of adjectives. In [18], a conclusion is made about the plurality of classification options, as well as the complexity of a universal classification development on any basis. Adjectives in [19] are divided into classes associated with the type of the object being defined. Probably, this or that transition from adjectives to nouns is a good way of classifying them and, accordingly, establishing short paths.

Consider the words *коварство* (‘*insidiousness’)* and *каверза (‘trick’)*. The first of these words is listed in RuWordNet as a cognate word and also as a part of speech synonym for *коварный (‘insidious’)*. The second of these words – *каверза (‘trick’)* – is not associated with *каверзный (‘tricky’)* by relationships of this type. In *каверзный* (‘*tricky’)*, only *каверзность (‘trickery’)* is indicated as a cognate word and a part of speech synonym. It is proposed to add the word *каверза (‘trick’)* to it.

There is the following further path between *коварство* (‘*insidiousness’)* and *каверза (‘trick’)*: *каверза – баловство – поступок человека – дурной поступок – коварство (‘trick’ – ‘mischief’ – ‘human deed’ – ‘bad deed’ – ‘insidiousness’)*. It has a distance of 4 and, together with two transitions: ‘*insidious’* – ‘*insidiousness* ‘and ‘*tricky’* – ‘*trick’*, it is possible to get a path with length 6. However, this path is already much better than the one considered above. All words in it are directly related to the given words. The seme ‘human deed’ is preserved in all words (collocations) along this path. Although, the seme of negativity is lost on the way. *Баловство* (‘*mischief’)* in RuWordNet has a hypernym *развлечение (‘entertainment’)* and belongs to the domain of “отдых” (‘*recreation’)*. Let us consider the senses of the word *каверза (trick)* in the dictionaries. The Wiktionary gives the following senses: *злая интрига, происки, затеваемые с целью запутать что-либо, повредить кому-либо; коварная, злая шалость, проделка* (‘*evil intrigue, machinations undertaken in order to confuse something, to harm someone; insidious, evil mischief, trickery’*). The similar senses are in D.N. Ushakov Dictionary [20].

Evidently, the word *каверза (‘trick’)* has a definitely negative sense (and not just *баловство* (*mischief)*) and clearly approaches the word *коварство* (‘*insidiousness’)*. It is proposed to add the second sense of the word *каверза (‘trick’)*: *каверза* 2 ‘злая интрига*’ (‘evil intrigue’*) with the hypernym *дурной поступок* (‘*bad deed’)*. The sense of the ‘evil intrigue’ is manifested, for example, in the following sentence: *Он чувствовал, что готовится какая-то новая каверза со стороны союзников* (‘He felt that some new trick was being prepared by the allies. A.N. Tolstoy. The Adventures of Nevzorov, or Ibicus’). In this case, the length of the path between the original words is reduced to 4: *коварный – коварство – дурной поступок – каверза 2 – каверзный (‘insidious’ – ‘insidiousness’ – ‘bad deed’ –‘ trick’ 2 – ‘tricky’).*

2. *Коварный – злой (‘Insidious’ – ‘evil’)*

All paths for this pair, similarly, have a length 6 and pass through common words such as *ролевой (‘thematic’)*. Obviously, the words *коварный* (‘*insidious’)* and *злой* (‘*evil’)* are semantically closer to each other than to the word *ролевой* (‘*thematic’)*. In the Wiktionary, коварство (‘*insidiousness’)* is defined as “a property of a person's character: the ability and tendency to hide evil intentions behind ostentatious friendliness”. Thus, the seme ‘evil’ is inherent in both of these words. There are several possibilities to build shorter paths between them. First, adding between the words *злодейство* (‘*insidiousness’)* and the *злой (‘evil’)* the relation of the ‘cognate words’ and / or ‘part of speech synonymy’ we get a path length 4: *коварный – коварство – плохой поступок –- злодейство –- злой* (‘*insidious ‘*– ‘*insidiousness’* – ‘*bad deed ‘– ‘evil deed’ – ‘evil’)*. Another possibility is to introduce a new collocation *коварный умысел* (‘*insidious intent’)* into RuWordNet, adding it as a synonym for *злой умысел* (‘*evil intent’)*. As a result, the path of a length 3: *коварный – коварный*

умысел – злой умысел – злой (‘insidious’ – ‘*insidious intent’ – ‘evil intent’ – ‘evil*)’ is created. Both of these possibilities do not contradict each other and can be implemented together. The presence of several paths of the same distance between words, as we can see, is quite frequent.

3. *Ладить – подладиться (‘to get along with’ – ‘to adjust oneself to’)*

All discovered paths between these words pass through common words, such as *делаться, быть (‘to become’,’ to be’)*. As noted above, the presence of words with general senses in the path is a direct indication that some important connections in the thesaurus are missing. Meanwhile, *ладить* (‘*get along with’*) and *подладиться (‘adjust oneself to’)* have a direct relationship: *ладит (‘get along with’)* is a state achieved as a result of the action *подладиться* (‘*to adjust oneself to’)*. In terms of the Theory of lexical functions by Melchuk [21], this can be expressed as *подладиться (‘get along with’)* = Caus (*ладить (‘to adjust oneself to’*)). The RuWorNet structure includes the following semantic relationship for verbs: ‘entailment’ (logical consequence). This relation in particular connects in the thesaurus the words: *призадумываться – думать* (‘*become thoughtful’ – ‘to think’)*; collocations: *наладить отношения – состоять в отношениях (‘to improve relationships’ – ‘to be in a relationship’),* etc. It seems natural to establish the same relationship between the words *подладиться – ладить (‘to adjust oneself to’ – ‘to get along with’)*. Then the distance between them will be equal to 1. In general, it is necessary to systematically consider the advisability of introducing this relation in pairs, such as *пробежаться – бежать (‘take a run’ – ‘to run’*), etc., as well as the widespread use of lexical functions. Very often the use of a lexical function does not lead beyond the synsets. For example, for the function ‘начинать’ (start) – Incep, we have the relation *побежать* (‘*take a run’*) = Incep (*бежать (‘run’)*).

The meaning of the derived word does not go beyond the inflectional nest (considering the transformation of the imperfect form into the perfect form to be the inflection, which is natural for thesauri and is used in RuWordNet). However, there are also more complex cases. Thus, in [13], the words *бороться* (‘*fight’*) и *устранить (‘eliminate’)* are considered to be analogues. The second word is the result of the (successful) execution of the first word.

4. *Каяться – облегчить душу (‘to repent’ – ‘to relieve (one's soul)’)*

An example of the shortest length path 6 is shown in Table 3. All such paths go through common words such as *прийти* (*come)*. Meanwhile, words *каяться – облегчить душу (‘to repent’ – ‘to relief’)* are directly related semantically. The dictionary [22] treats them as synonyms. Repentance usually leads to a relief of the soul and is carried out for this purpose, for example:

> *После этой знаменательной беседы я решила искать утешения в церкви и облегчить душу покаянием на святой исповеди.* (‘After this significant conversation, I decided to seek consolation in the church and to relieve my soul by repentance in holy confession’. Hoffman E. Elixirs of Satan. 2020).

> *Но по опыту прошлых причастий не уверена, что найдется человек, который мне поможет покаяться во всех грехах и действительно облегчит душу.* (‘But from the experience of the past sacraments, I am not sure that there will be a person who will help me repent of all my sins and really relieve my soul’ (https://omolenko.com/928.html)).

Thus, it makes sense to establish the ‘entailment’ relationship between them. Although it is possible to relieve the soul in other ways, this does not eliminate the indicated semantic relationship.

5. Отчетливый – контрастный (‘distinct’ – ‘contrastive’)

This example is most complex and controversial among those considered. We propose to obtain a short path between them through rather serious changes in the thesaurus. Perhaps this also indicates that these words should not be interpreted as quasi-synonyms. The example of the shortest path length of 6, is shown in Table 3. All such paths go through common words, such as *оценка, отношение*

(‘evaluation’, ‘relation’), which do not reflect the specific semantics of the words in question. To get a short distance, it is necessary to add new synsets. We propose to introduce a synset based on the collocation *отчетливо различимые детали (‘clearly distinguishable details’)*. This synset will be associated with the synset of the word *отчетливый (‘clear’)* through the relation 'Component synsets', connecting the synset of a phrase with synsets of the component words. Next, we introduce a synset based on the collocation *контрастное изображение (‘contrast image’)*. According to the dictionary [23], the phrase *контрастное изображение (‘contrast image’)* is defined as “An image with clearly distinguishable individual details”. This synset is associated with the synset of clearly distinguishable details through the 'whole-to-part' relationship. Finally, the word *контрастный (‘contrastive’)* is associated with the collocation *контрастное изображение (‘contrast image’)* through the ‘Component Synsets’ relation. As a result, a path length of 3 can be found.

6. *Гостеприимный – компанейский (‘hospitable’ – ‘companionable’)*

Not all paths of the distance 6 are necessarily inadequately long.

The shortest path between the words *гостеприимный (‘hospitable’)* and *компанейский (‘companionable’)*, has a length of 6 (Table 3), but all semantic relationships in it are logical, and it is not clear how this path can be shortened in a proper way.

7. *Громкий – знаменитый (‘loud’ – ‘famous’)*

All paths between *громкий (‘loud’)* and *знаменитый (‘famous’)* go through the word *качественный (‘qualitative’)*. This is due to the fact that the word *громкий (‘loud’)* in RuWordNet has two senses ‘громкий по звуку’ (‘*loud*’) and ‘высокопарный’ (‘*pompous*’). At the same time, according to the Ozhegov Dictionary (as well as the Wiktionary), this word has also the sense *получивший широкую известность (‘widely known’)*. If we add this sense to the word *громкий* (*loud)*, then it will become a co-hyponym with *знаменитый* (‘*famous’)* through the hypernym *широкоизвестный (‘widely known’)*, thus establishing paths of the length 2.

5. Discussion

Analysis of the examples enabled to reveal the following typical types of gaps in the thesaurus.

1. Missing words, which are in relation to part-of-speech synonymy and cognate words (example 1);
2. Missing senses of words from the standard dictionaries of the Russian language (examples 1, 7)
3. Missing semantic relation of entailment for verbs (examples 3, 4)
4. Missing synsets (example 5).

The present study has shown that the proposed method of studying the shortest paths between analogous words is highly effective and makes it possible to systematically identify gaps in the thesaurus. The solutions proposed to remedy the situation are subjective where appropriate, as is the compilation of dictionaries and thesauri in general. In some cases, the proposed solutions are suggestive, while in others they may be controversial. The most doubtful is the decision to add new collocations, as in the example 5, where it is recommended to add the collocation *отчетливо различимые детали (‘clearly distinguishable details’)* to the thesaurus. It is obvious that in general the question whether a collocation is subject to the inclusion in the thesaurus or not is non-trivial by itself. This issue is discussed in [24], proposing the criteria for the inclusion of collocations in the thesaurus.

It is clear that the inclusion of such collocations in the thesaurus as *печально известный (‘sadly remembered’)* or *как облупленный* (‘*inside out’)* and many others is beyond doubt. In the example considered, the inclusion of the collocation *отчетливо различимые детали (‘clearly distinguishable details’)* is more controversial. However, regardless of whether it meets the criteria of [24] or not, its addition improves the structure of the thesaurus by introducing new short natural paths. A methodological question arises as to whether collocations should be introduced specifically for the

purpose of improving the structure of the thesaurus. It seems that the approaches to verification of thesauri put forward and discussed in the present article can facilitate solving this issue.

6. Conclusion

The research work proposes a new method for analyzing the structure of thesauri. It consists in considering semantically close words, which are not as close as synonyms. These are, first of all, quasi-synonyms (analogues). However, this approach can be applied to the words obtained by using lexical functions. We proceed from the fact that semantically close words should be close including in terms of the distance between them according to the thesaurus [14].

Using the Explanatory Dictionary of Synonyms [13], we have arranged many pairs of analogous words and built the shortest paths between them in RuWordNet. As a result of the analysis of long paths (length 6), we revealed typical situations, in which semantically close words turn out to be distant according to the thesaurus. In addition to the length of the path, the presence of words having too general meaning in the path shows that the semantic similarity of words is not captured by the thesaurus in its current state. The search for shorter paths is based on a careful analysis of the analogous words interpretations.

7. Acknowledgements

This research was financially supported by RFBR, grants № 18-00-01238 and № 18-00-01226 as parts of the project № 18-00-01240 (K).

8. References

[1] Ch. Fellbaum (Ed.), WordNet: An electronic lexical database, MIT Press, 1998.

[2] N. Loukachevitch, G. Lashevich, B.Dobrov, Comparing Two Thesaurus Representations for Russian, /Proceedings of Global WordNet Conference GWC-2018, 2018, pp. 35–44.

[3] N. Lukashevich, Thesauri in information retrieval tasks, Moscow State University, 2011. (in Russian)

[4] N.V.Lukashevich, G. Lashevich, A. Gerasimova, V. Ivanov, B.V. Dobrov, Generation of the thesaurus of the WordNet type for the Russian language, Proceedings of the conference on artificial intelligence KII-2016, V.2, 2016, pp, 89–97. (in Russian)

[5] Yu. Apresyan, Selected Works, Volume II. Lexical Semantics. School “Languages of Russian Culture”. M.: Oriental Literature, 1995. (in Russian)

[6] D. Divjak, Structuring the lexicon: A clustered model for near-synonymy, De Gruyter Mouton, Berlin, 2010.

[7] D.A. Cruse, Lexical semantics. Cambridge University Press, Cambridge, 1986.

[8] Ph. Edmonds, G. Hirst, Near synonymy and lexical choice, Computational Linguistics, 2002 28, pp. 105–144.

[9] A. Hazem, B. Daille, Word embedding approach for synonym extraction of multi-word terms, Proceedings of the Eleventh International Conference on Language Resources and Evaluation, LREC-2018, 2018.

[10] Sh. Ali, R. Crawford, Query Expansion Behavior Within a Thesaurus-Enhanced Search Environment: A User-Centered Evaluation, Journal of the American Society for Information Science and Technology, 57(4), 2006, pp. 462–478.

[11] V.V. Bochkarev, V.D. Solovyev, Properties of the network of semantic relations in the Russian language based on the RuWordNet data, Journal of Physics: Conference Series, Vol. 1391, paper 012052.

[12] I.M. Kobozeva, Component analysis of lexical meaning, Linguistic semantics: Textbook. - 4th ed, M.: Book House "LIBROKOM", 2009. pp. 109–122. (in Russian)

[13] Y.Apresyan (ed.), A New Explanatory Dictionary of Synonyms of the Russian Language. 2nd edn. Moscow, Yazyki russkoy kultury, 2000. (in Russian)

[14] N. Loukachevitch, Corpus-based Check-up for Thesaurus. Proc. of the 57th Annual Meeting of the Association for Computational Linguistics, 2019, pp. 5773–5779.

[15] V. Bayrasheva, Corpus-based vs thesaurus-based word similarities: expert verification, The XXth International Scientific Conference “Cognitive Modeling in Linguistics”, 2019, pp. 56-63.

[16] V. Soloviev, G. Gimaletdinova, L. Khalitova, L. Usmanova, Expert Assessment of Synonymic Rows in RuWordNet, Analysis of Images, Social Networks and Texts (AIST'2019). CCIS, vol. 1086. 2020, pp. 174-183.

[17] L. Usmanova, I. Erofeeva, V. Solovyev, V. Bochkarev, Analysis of the Semantic Distance of Words in the RuWordNet Thesaurus, Proceedings of the DAMDID/RCDL-2020, 2020.

[18] Z.A. Haritonchik, An adjective: problems of classification [Imja prilagatelnoe: problemy klassifikacii] Theory of grammar: lexico-grammatical classes and ranks of words [Teorija grammatiki: leksiko-grammaticheskie klassy i razrjady slov] Berezina F. M., Kubrjakovoj E. S. (eds.) INION RAN: Moscow, 1990, pp. 94–119.

[19] A.N. Shramm, Essays on semantics of quality adjectives: on the data of modern Russian language. [Ocherki po semantike kachestvennyh prilagatelnyh: na materiale sovremennogo russkogo jazyka], Publishing house of the Leningrad State University: Leningrad, 1979.

[20] D.N. Ushakov (Ed.), Explanatory dictionary of the Russian language: In 4 volumes, M .: Sov. encycl. : OGIZ, 1935–1940. (in Russian)

[21] I. A. Melchuk, A. K. Zholkovsky and others, Explanatory Combinatorial Dictionary of the Modern Russian Language. Experiments in the semantic and syntactic description of Russian vocabulary, Vienna: Wiener Slavistischer Almanach, 1984.

[22] A.S. Gavrilova, Dictionary of synonyms and antonyms of the modern Russian language, M. 2014, (in Russian)

[23] L.M. Nevdyaev, Telecommunication technologies. English-Russian explanatory dictionary-reference book, M., 2002

[24] N. Loukachevitch, G. Lashevich, Multiword expressions in Russian thesauri RuThes and RuWordnet, IEEE Artificial Intelligence and Natural Language Conference, AINL-2016, IEEE, 2016, pp. 1–6.