

Comparing Similarity of Words Based on Psychosemantic Experiment and RuWordNet

Valery Solovyev

Kazan Federal University,
Kazan, Russia,

maki.solovyev@mail.ru

Natalia Loukachevitch

Lomonosov Moscow State University,
Moscow, Russia,

louk_nat@mail.ru

Abstract

In the paper we compare the structure of the Russian language thesaurus RuWordNet with the data of a psychosemantic experiment to identify semantically close words. The aim of the study is to find out to what extent the structure of RuWordNet corresponds to the intuitive ideas of native speakers about the semantic proximity of words. The respondents were asked to list synonyms to a given word. As a result of the experiment, we found that the respondents mainly mentioned not only synonyms but words that are in paradigmatic relations with the stimuli. The words of the mental sphere were chosen for the experiment. In 95% of cases, the words characterized in the experiment as semantically close were also close according to the thesaurus. In other cases, additions to the thesaurus were proposed.

1 Introduction

Semantic proximity of words is an important parameter required in various tasks of natural language processing. It can be estimated in different ways: by corpus, using distributional methods (Mikolov 2013, Bojanowski et al., 2017), expert assessments; from psychosemantic experiments; using thesauri such as WordNet (Fellbaum, 1998).

The general concept of *semantic similarity* can be subdivided to paradigmatic (taxonomical) similarity and semantic associations (Agirre et al., 2009; Hill et al., 2015; Kliegr and Zamazal, 2018; Majewska et al., 2020). Paradigmatic similarity can be defined in terms of shared superordinate category or shared semantic features. Se-

mantic associations correspond to co-occurrence (syntagmatic relations) in texts.

To study automatic methods of word similarity calculation, specialized datasets are created. Some researchers try to create datasets distinguishing different subtypes of semantic similarity of words, which requires additional efforts and guidelines. Agirre et al. (2009) subdivided the existing semantic word dataset WordSim353 (Finkelstein et al., 2002) into two subsets: WordSim353-similarity and WordSim353-relatedness datasets. SimLex-999 guidelines (Hill et al., 2015) aim to distinguish word pairs in taxonomical semantic similarity relation (synonyms, hypernyms, hyponyms) from remaining types of relations (antonymy, co-hyponyms). The authors of WIN353 dataset (Kliegr and Zamazal, 2018) ask respondents about word similarity based on word interchangeability in sentences.

We can also try to use human scores of word semantic similarity to assess the quality of descriptions in electronic lexical-semantic resources (thesauri). Such resources are built on the basis of synsets – sets of synonyms – linked by semantic relations such as hyponymy, hypernymy, antonymy, and some others. The automatic use of thesauri requires high quality descriptions of semantic senses and semantic relations between them.

In this paper, we compare the results of a survey of respondents and the similarity of words according to the RuWordNet thesaurus (Loukachevitch et al., 2018) for the Russian language. Currently, the published RuWordNet version comprises about 110 thousand Russian words and expressions. A new version of RuWordNet is being prepared and RuWordNet data are tested from different points of view.

In the psychosemantic experiment the respondents were asked to list synonyms for stimu-

li words without any guidelines. We found that their answers mainly contain paradigmatically similar words, practically without words related via any other similarity relationships, which makes it possible to check the taxonomic structure of the thesaurus.

The paper is structured as follows. Section 2 provides information on related work. Section 3 describes a psychosemantic experiment to determine the semantic proximity of words. Section 4 analyzes the data obtained. Section 5 discusses the results of the experiment.

2 Related work

The paper concerns two directions of studies: revision and updating existing lexical-semantic resources for natural language processing (thesauri) and studies on relation types exploited by native speakers in word association experiments.

2.1 Revision of Existing Lexical Semantic Resources

Procedures for revising and verifying resources are important for the developers of WordNet-like resources. Some ontological tools have been proposed to check the consistency of relationships in WordNet (Guarino and Welty, 2004; Alvez et al., 2018). Rambousek et al. (2018) considered a crowdsourcing tool allowing a user of the Czech wordnet to report errors. Users may propose an update of any data value. These suggestions can be approved or rejected by editors. Visualization tools can also help to find problems in wordnets (Piasecki et al. 2013; Johannsen et al. 2011). Cristea et al. (2004) and Rudnicka et al. (2012) reported on the revision of mistakes and inconsistencies in their wordnets in the process of linking the wordnet and the English WordNet.

McCrae et al. (2019) discussed a new project: Open-Source WordNet for English, which is based on the Princeton WordNet. This project has already fixed errors found in the current version of WordNet, including spelling mistakes in definitions and examples. Some problematic issues were reported (for example, synset duplicates, missed or incorrect relationships) for further revision.

Recently, verification and enrichment methods have been systematically developed for the RuWordNet thesaurus. In (Loukachevitch, 2019), the following method for enriching the RuWordNet thesaurus was proposed. For a large

text corpus, words are searched for which 20 words closest in the corpus (based on the standard method for evaluating the semantic similarity of words) are located far from each other in the thesaurus. The distance between words in the thesaurus is the length of the shortest path between them in the graph of semantic relations. For found words with such properties, the reasons for such discrepancy are analyzed. The analysis of the data presented in (Loukachevitch, 2019) was continued in (Bayrasheva, 2019).

In work (Soloviev et al., 2020), RuWordNet synsets were compared with synonymous sets according to published 10 dictionaries of Russian synonyms. The work (Erofeeva et al., 2020) presents the results of an experiment in which the respondents were asked to list synonyms for a given word. The results are compared with the RuWordNet synsets. Usmanova et al. (2020) analyzed pairs of quasi-synonyms and the distance between them in RuWordNet. It was expected that quasi-synonyms, as semantically close words, should be located at a short distance in the thesaurus.

The general result of above-mentioned studies of RuWordNet is as follows. RuWordNet data, including the composition of synsets and the structure of semantic relations, correlate well with all the other considered sources of information about semantically close words. At the same time, a number of gaps in RuWordNet were identified, taking into account of which allows improving descriptions in the thesaurus. This article continues research in this direction.

2.2 Lexical relations in associations experiments

One of the most known associative experiments for Russian was organized by Karaulov in 1986-1997 (Karaulov et al., 2002). In experiments such demographic information such as age, gender, specialization, and location was also considered and recorded. Currently, these data are considered as outdated.

Many researchers classify word associations into syntagmatic and paradigmatic relations (Fitzpatrick, 2006). The researchers study the structure of associations for language learners (Fitzpatrick, 2006), patients (Arias-Trejo et al., 2018), children (Wojcik and Kandhadai, 2019), and other social groups.

Vylomova et al. (2018) study types of relations in associative responses of Russian native speakers in dependence on socio-demographic characteristics. They organized associative ex-

periments in various Russian regions, including Siberia and the Urals. The age of participants ranged from 16 to 26, most of them were university students of approximately 50 specialties. In their analysis, Vylomova et al. (2018) classified the lexical relations in associations to syntagmatic (they calculated word co-occurrences in a text corpus) and paradigmatic (according to RuWordNet thesaurus). The authors found that men more frequently list paradigmatic associations whereas women are more likely to produce syntagmatic associations. It was also revealed that most students of technical specializations and natural sciences demonstrate high scores for paradigmatic association types.

Sinopalnikova (2004) studies approaches to extract useful lexical relations from existing word association thesauri to assist in developing new wordnets.

3 Experiment Setting

In the current study we present the results of a psychosemantic experiment, carried out in accordance with the methodology described in (Petrenko, 2010). The experiment reveals semantically close words (synonyms) as seen by native speakers. In (Erofeeva et al., 2020) only synonyms from the RuWordNet synsets were considered, in this work all semantic relations are involved.

The experiment is as follows. The respondents (Russian native speakers) receive a number of words, and they have to list synonyms for these words in a limited time. The respondents are students (18-23 years old, 200 people) of Kazan Federal University (Kazan, Russia). About half of the students are philologists, the second half are non-philological students. The definition of a synonym is not explained to the respondents; we rely on intuitive understanding of synonyms by native speakers. For the experiment, words related to the mental sphere are selected. This semantic area is the most difficult for clear differentiation of synonym sets and their semantic relations.

The results of philologists and non-philologists differ insignificantly. However, it is worth noting that synonyms for word *мечта* (*mechta* – dream as imaginative thoughts) listed by philologists and non-philologists have interesting distinctions. So, for philologists, the word *фантазия* (*fantasia* – fantasy) is in 3rd place, and for non-philologists, the word *стремление* (*stremleniye* – aspiration) is in the 3rd position. Conversely, for philologists, *stremlenie* (aspira-

tion) is listed in the 5th place, and for non-philologists, *fantasia* (fantasy) is in the 4th position. It seems that the figurative thinking of philologists, the reading and study of fiction, which form their linguistic personalities, are reflected in the results of the experiment: for them, the word *mechta* (dream) is associated with fantasy and dreams, that is, with something unreal, ephemeral. Non-philologists are more pragmatic: the third position in their lists is occupied by the word *stremlenie* (aspiration), in the semantics of which the presentation of concrete results is conveyed (Erofeeva et al., 2020).

Further we will write Cyrillic Russian words in Latin transcription.

Since the respondents, naturally, did not use the criteria of synonymy, such as interchangeability in different contexts and did not have much time to complete the task, they suggested words that have something semantically in common with the given word, but not necessarily synonyms in the strict sense of the term. For example, for the word *mechta* (dream as imaginative thoughts), the following words were listed as synonyms in RuWordNet: *gresa*, *mechtaniye*, *fantasia* (fantasy). The respondents most often indicated the following words: *zhelaniye* (desire), *tsel'* (goal), *fantasia* (fantasy), *gresa*, *stremleniye* (aspiration), *nadezhda* (hope). Only two of them are synonymous. The rest of the words – *zhelaniye* (desire), *tsel'* (goal), *stremleniye* (aspiration), *nadezhda* (hope) – at first glance may seem like associations with the given word dream. However, this assumption is not true.

In the Karaulov's dictionary of Russian associations (Karaulov et al., 2002), the word *mechta* (dream) has the following most frequent associations: *goluboy* (blue), *zhizn'* (life), *moya* (mine), *sbylas'* (come true), *idiota* (idiot), *nesbytochnaya* (unrealizable), *rozovaya* (pink). The words *zhelaniye* (desire), *stremleniye* (aspiration), *nadezhda* (hope) are not mentioned as associations at all, and the word *tsel'* (goal) is mentioned only once in 101 responses. We can see that in fact words having syntagmatic relations with the original one are also mentioned as associations by respondents. In the current experiment, the respondents indicated words that were not in syntagmatic but in paradigmatic relations with the stimulus. Rather, they can be characterized as belonging to the semantic field of the original word or as its analogues.

It is worth noting that in the dictionary (Apresian, 2004) the words *namereniye* (intention) and *mysl'* (thought) are considered as ana-

logues (near-synonyms) of the word *mechta* (dream) (its synonyms are not given in the dictionary). For the verb *mechtat'* (to dream), the synonyms, according to (Apresyan, 2004), are *khotet'* (to want), *zhelat'* (to desire), and the analogue is the word *nadeyat'sya* (to hope). Thus, the words indicated by the respondents are close in meaning to the word *mechta* (dream). Our experiment can be characterized as aimed at identifying paradigmatic associations, while the Karaulov's dictionary (Karaulov et al., 2002) in fact mixes paradigmatic and syntagmatic associations.

4 Analysis of Results

In this work, the associations for the words *obida* (offense, as a feeling caused being offended), *radost'* (joy), *talant* (talent), *strast'* (passion), *lyubov'* (love), *mysl'* (thought), *vostorg* (delight) are considered. For each stimulus word, six most frequently mentioned responses are studied.

Obida (offense feeling). The informants most often indicated the words: *ogorcheniye* (grief), *dosada* (annoyance), *bol'* (pain), *grust'* (sadness), *razocharovaniye* (disappointment), *zlost'* (anger). The first of them is interpreted in RuWordNet as a hypernym for *obida*. The word *grust'* (sadness) in RuWordNet also has a direct connection with *obida* – it is a hypernym-hypernym for *obida*. *Dosada* (annoyance) is a co-hyponym for *obida*, having the common hypernym *nedovol'stvo* (discontent).

There is also a short path between the words *obida* and *razocharovaniye* (disappointment): *obida* (offense) – *nedovol'stvo* (discontent) – *dushevnoye perezivaniye* (emotional experience) – *razocharovaniye* (disappointment). There is a similar path between the words *obida* (offense) and *razocharovaniye* (disappointment): offense – discontent – emotional experience – disappointment. Finally, the path between the words *bol'* (pain) and *obida* is only slightly longer: pain – suffering – emotional experience – discontent – offense. Semantic distances of 4 steps or less are treated in (Loukachevitch, 2019) as short. All semantic relations are hypohypernymic.

Radost' (joy). For this word, respondents indicate the following word associations: *schast'ye* (happiness), *vostorg* (delight), *vesel'ye* (fun), *ulybka* (smile), *likovaniye* (exultation), *udovol'stviye* (pleasure).

The words *veseliye* (fun), *likovaniye* (exultation), *udovol'stviye* (pleasure) are hyponyms in relation to *radost'* (joy). The words *vostorg* (delight) and *schast'ye* (happiness) are co-hyponyms with *radost'* (joy) with a common hypernym – *dushevnoye perezivaniye* (emotional experience). But between the words *radost'* (joy) and *ulybka* (smile) there is only a very long way: *radost'* (joy) – *dushevnoye perezivaniye* (emotional experience) – *mental'nyy ob'yekt* (mental object) – *abstraktnaya sushchnost'* (abstract entity) – *kachestvo* (quality) – *vneshnost'* (appearance) – *vyrazheniye litsa* (facial expression) – *ulybka* (smile). Such a long path reflects the fact that in RuWordNet the word *ulybka* (smile) is interpreted only as a facial expression and, accordingly, *radost'* (joy) and *ulybka* (smile) in RuWordNet refer to different spheres – the mental world and the physical.

Princeton WordNet presents the point of view that a person smiles to communicate something to others about his condition (to change one's facial expression by spreading the lips, often to signal pleasure¹) and thus it is classified as communication. Still, it should be noted that a person can smile at own thoughts, pleasant memories while alone with yourself, i.e. a smile is also possible outside the communication situation. As we can see, the situation here is very difficult. According to the Russian explanatory dictionary (Ozhegov and Shvedova, 1997), *ulybka* (smile) has the following definition: “mimic movement of the face, lips, eyes, showing disposition to laughter, expressing pleasure or ridicule and other feelings (translation from Russian)”. This definition takes into account both facial expressions and communicative intentions.

It is possible to take into account the intuition of native speakers and the dual nature of a smile by making certain changes to the thesaurus. It can be described with the entailment relationship between concepts *ulybat'sya* (to smile) and *radovat'sya* (to joy). If a person smiles, then usually this person is really happy, or at least seeks to show happiness to others. Conversely, if a person is really happy about something, then this manifests itself in a smile.

Talant (talent). For this word, the respondents indicate the following synonymous words: *sposobnost'* (ability), *dar* (gift), *umeniye* (skill), *darovanie*, *odarenost'* (giftedness), *talent*, *geniy* (genius). In RuWordNet *darovaniye* (giftedness),

¹ <http://wordnetweb.princeton.edu/perl/webwn>

dap (gift), are listed as synonyms to the word *talant* (talent). *Sposobnost'* (ability) is a hypernym for *prirodnaya sposobnost'* (natural ability), which is a hypernym for *talant* (talent). *Umeniyе* (skill) is a co-hyponym with *talant* (talent) through the general hypernym *prirodnaya sposobnost'* (natural ability). *Odarennost'* (giftedness) is a hyponym in relation to *sposobnost'* (ability), i.e. is at distance 3 from the word *talant* (talent). The word *geniy* (genius) is a hyponym to *odarennost'* (giftedness), at a distance 4 from the word *talant* (talent).

Strast' (passion). For the word *strast'* (passion), the respondents indicate the synonymous words: *zhelaniye* (desire), *vlecheniye* (attraction), *любовь* (love), *увлечение* (infatuation), *pokhot'* (lust), *интерес* (interest). In RuWordNet, *увлечение* (infatuation) is a hypernym for *strast'* (passion). The word *interes* (interest) is a hypernym of the hypernym for the word *strast'* (passion). The words *vlecheniye* (attraction), *zhelanie* (desire), *lyubov'* (love) are co-hyponyms with the word *увлечение* (infatuation) with a common hypernym, *dushevnoye perezkhivaniye* (emotional experience), i.e. are at a distance of 3 from *strast'* (passion). The word *pokhot'* (lust) is a hyponym in relation to *vlecheniye* (attraction), i.e. is at a distance of 4 from the initial word *strast'* (passion).

The scheme of semantic relations in this group of words can be represented in Fig. 1. This is a typical scheme for the student answers in the experiment. The arrows show links from hypernyms to hyponyms.

Lyubov' (love). For this word, the respondents indicated such words as *privyazannost'* (attachment), *vlyublennost'* (falling in love), *sympatia*

(sympathy), *nezhnost'* (tenderness), *vlecheniye* (attraction).

We saw above that *lyubov'* (love) is at a distance of 3 from *strast'* (passion) and 2 from *vlecheniye* (attraction). *Privyazannost'* (attachment) is a hypernym for *lyubov'* (love). The words *lyubov'* (love) and *vlyublennost'* (falling in love) are co-hyponyms with a common hypernym *dushevnoye perezkhivaniye* (emotional experience). The word *sympatia* (sympathy) is a co-hyponym with word *vlyublennost'* (falling in love) through a hypernym *lichnostnyye otnosheniya* (personal relationships). Thus, between the words *sympatia* (sympathy) and *lyubov'* (love) there is a distance of length 4. But *nezhnost'* (tenderness) is interpreted in RuWordNet only as a character trait (two other senses: *nezhnost'* 1 (soft, gentle to the touch), *nezhnost'* 3 (fragile, too weak) are not here discussed as irrelevant), and not as mental experience and there is no close way between them.

In fact, in RuWordNet one of the senses of the word *nezhnost'* (tenderness) is missing. In the dictionary (Ozhegov and Shvedova, 1997), *nezhnost'* (tenderness) refers to the word *nezhnyi* (tender), which is interpreted (in this sense) as "affectionate, full of love: tender feelings". According to the dictionary (Apresian, 2004) "*nezhnyi* (tender) – showing a feeling of love or affection in communication with a person."

Thus, in the interpretation of this word, the word *lyubov'* (love) invariably appears, indicating the correctness of the students' assessment. Therefore, it is recommended to add a new sense of the word *nezhnost'* (tenderness) in RuWordNet, in accordance with the above-mentioned dictionary definitions.

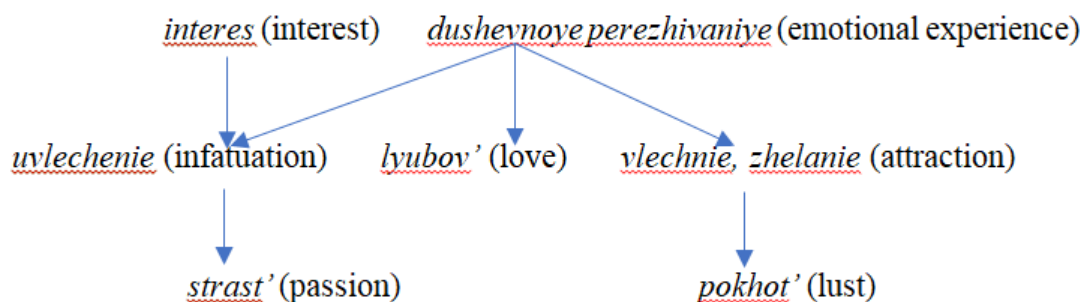


Fig1. Scheme of semantic relations of words-reactions to the stimulus *strast'* (passion)

Vostorg (delight). Respondents indicate the following words: *radost'* (joy), *voskhishcheniye* (admiration), *udivleniye* (surprise), *schast'ye* (happiness), *likovaniye* (jubilation), *voodushevleniye* (inspiration).

In RuWorNet, *vostorg* (delight) and *voskhishcheniye* (admiration) are synonyms, *schast'ye* (happiness), *udivleniye* (surprise) and *radost'* (joy) are co-hyponyms with *vostorg* (delight) with the common hypernym *dushevnoye perezhivaniye* (emotional experience). Word *likovaniye* (jubilation), as noted above, is a hyponym in relation to *radost'* (joy), i.e. is at a distance of 3 from *vostorg* (delight). *Voodushevleniye* (inspiration) is a co-hypernym with the word *radost'* (joy) with the common hyponym *euphoria*, i.e. is at a distance of 4 from *vostorg* (delight).

Mysl' (thought). Respondents indicate the following words: *ideya* (idea), *duma* (thought), *mneniye* (opinion), *dogadka* (guess), *soobrazheniye* (consideration), *suzhdeniye* (judgment).

Words *soobrazheniye* (consideration) and *duma* (thought) are synonymous with *mysl'* (thought). *Ideya* (idea) is a hyponym for *mysl'* (thought), *suzhdeniye* (judgment) is a hypernym for *mysl'* (thought). *Mneniye* (opinion) is a co-hyponym with *mysl'* (thought) via common hypernym *suzhdeniye* (judgment).

Between the words *mysl'* (thought) and *dogadka* (guess) there is a path of length 4: *mysl'* (thought) – *suzhdeniye* (judgment) – *mneniye* (opinion) – *dopushcheniye* (assumption) – *dogadka* (guess).

5 Discussion

We analyzed 40 word pairs (out of a total of $7 \times 6 = 42$ pairs, two pairs were repeated). In 38 cases (95%), word pairs listed by the respondents as synonyms are also close according to the thesaurus descriptions: 13 pairs are at a distance of 1; 12 pairs are at a distance of 2; 7 pairs are at a distance of 3; 6 pairs are at a distance of 4. The number of mentioned words located at a certain path distance in the thesaurus decreases monotonically with increasing distance. In all these cases, it turned out to be sufficient to consider only hypo-hypernymic relations. In two cases, it

is necessary to make certain changes in the thesaurus to obtain smaller distance for semantically close words. These pairs of words are as follows: *lyubov'* (love) – *nezhnost'* (tenderness) and *radost'* (joy) – *ulybka* (smile). In the first case, it is proposed to add a new sense of the word *nezhnost'* (tenderness) to the thesaurus and to establish the necessary additional relation of hyponymy, in the second case we suggest to add the relation of entailment.

Thus, most words frequently mentioned by respondents are located close to the stimulus word in RuWordNet, which indicates good consistency of the thesaurus with the intuition of native speakers. At the same time, taking into account the data of a psychosemantic experiment makes it possible to identify some problem areas in the thesaurus. Let us consider whether there is a correlation between the frequency with which the response word is chosen by the respondents and the distance in the thesaurus from the stimulus word to the response word. We sort words-reactions according to the frequency of their mention.

Table 1 summarizes the data, sorted by the frequency of the words in the respondents' answers. The asterisk indicates the distances that will take place after the implementation of the above-mentioned suggestions for improving the thesaurus structure. We can see that the words mentioned more often are at a shorter distance from the stimulus word in the thesaurus, which is also a good confirmation of the correct structure of the thesaurus and the adequacy of the experiment.

6 Conclusion

Thesauri are created by professionals who rely on both the theory of language and their ideas about the semantics of linguistic units. However, semantics are not described in the literature in as much detail as required by the thesaurus developers. Taking this into account, it is of natural interest to compare thesaurus data with the linguistic intuition of native speakers, manifested in psychosemantic experiments.

Stimulus word	Obida (offense)	Radost' (joy)	Talant (talent)	Strast' (passion)	Lyubov' (love)	Mysl' (thought)	Vostorg (delight)	Average
Words in the 1 st places of the respondents' associations								
Frequency (%)	24	49	55.5	43.5	26	61.5	48.5	40.3
Relation dist.	1	2	2	3	1	1	2	1.7
Words in the 2 nd places of the respondents' associations								
Frequency (%)	19	34.5	44.5	28.5	24.5	28.5	39.5	31.3
Relation dist.	3	2	1	3	4	1	1	2.1
Words in the 3 ^d places of the respondents' associations								
Frequency (%)	16	32	30	17.5	22	12.5	29.5	22.7
Relation dist.	2	1	2	3	2	2	2	2.0
Words in the 4 th places of the respondents' associations								
Frequency (%)	14	12	14	12.5	20.5	11	13.5	13.9
Relation dist.	4	1	3	1	3	1	2	2.1
Words in the 5 th places of the respondents' associations								
Frequency (%)	13.5	10.5	13	11.5	18	10.5	12.5	12.7
Relation dist.	3	3*	4	2	2	1	3	2.6
Words in the 6 th places of the respondents' associations								
Frequency (%)	13	7	11.5	9.5	14.5	9	8.5	10.4
Relation dist.	2	1	1	4	2*	4	4	2.6

Table 1. Positions, frequencies (percentage of answers) and RuWordNet distances of word associations. The sign *) means distances after the suggested corrections.

Most words frequently mentioned by respondents or synonyms with the stimulus word or are located close to it in RuWordNet, which indicates good consistency of the thesaurus with the intuition of native speakers. This confirms the high quality of the RuWordNet thesaurus. The experimental results also support the choice of distance 4 as a measure of the semantic proximity of words in the thesaurus. At the same time, taking into account the experimental data made it possible to identify some problem areas in the thesaurus.

Acknowledgments

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

References

Agirre, E., Alfonseca, E., Hall, K., Kravalova, J., Pasca, M., and Soroa, A. 2009. A study on similarity

and relatedness using distributional and WordNet-based approaches. In *Proceedings of NAACL-HLT: 19–27*.

Alvez, J., Gonzalez-Dios, I., and Rigau, G. 2018. Cross-checking WordNet and SUMO using meronymy. 2018. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC-2018)*.

Apresian, I.D. *New explanatory dictionary of Russian synonyms*. Moscow, Vena: Iazyki slavianskoi kul'tury, 2004. (in Russian)

Arias-Trejo, N., Minto-García, A., Luna-Umanzor, D. I., Ríos-Ponce, A. E., Mariana, B. P., and Bel-Enguix, G. 2018. Word-word relations in Dementia and typical aging. In *Proceedings of the first international workshop on language cognition and computational models:85-93*.

Bayrasheva, V. 2019. Corpus-based vs thesaurus-based word similarities: expert verification. *The XXth International Scientific Conference "Cognitive Modeling in Linguistics"*. Proceedings:56-63.

- Bojanowski, P., Grave, E., Joulin, A., and Mikolov, T. 2017. Enriching word vectors with subword information. *Transactions of the ACL*, 5:135–146.
- Cristea, D., Mihaila, C., Forascu, Trandabat, D., Hursariuc, M., Haja, G., and Postolache, O. 2004. Mapping Princeton WordNet synsets onto Romanian WordNet synsets. *Romanian Journal of Information Science and Technology*, 7(1-2):125–145.
- Erofeeva, I., Solovyev, V., and Bayrasheva, V. 2020. Psychosemantic experiment as a tool for objectifying data on the ways of representing synonymy in modern Russian language. *Bulletin of the Volgograd State University. Series 2: Linguistics*. 2020. - T. 19, No. 1:178–194 (in Russian).
- Fellbaum, Ch. 1998. *WordNet: An electronic lexical database*. MIT Press. 1998.
- Finkelstein, L., Gabrilovich, E., Matias, Y., Rivlin, E., Solan, Z., Wolfman, G., and Ruppin, E. 2002. Placing search in context: The concept revisited. *ACM Transactions on Information Systems*, 20(1):116–131.
- Fitzpatrick, Tess. 2006. "Habits and rabbits: Word associations and the L2 lexicon." *EuroSLA Yearbook 6.1* (2006):121-145.
- Guarino, N., and Welty, Ch. 2004. An overview of OntoClean. *Handbook on ontologies*. Springer, Berlin, Heidelberg:151-171.
- Hill, F., Reichart, R., and Korhonen, A. 2015. SimLex-999: Evaluating semantic models with (genuine) similarity estimation. *Computational Linguistics*, 41(4):665–695.
- Johannsen, A. and Pedersen, B. 2011. Andre ord?—a wordnet browser for the Danish wordnet, DanNet." *Proceedings of the 18th Nordic Conference of Computational Linguistics (NODALIDA 2011)*.
- Karaulov, Yu.N., G.A. Cherkasova, N.V. Ufimtseva, Yu.A. Sorokin, E.F., and Tarasov. 2002 *Russian associative dictionary*. In 2 volumes. M.: AST-Astrel, 2002. 784 p. (in Russian)
- Kliegr T., and Zamazal O. 2018. Antonyms are similar: Towards paradigmatic association approach to rating similarity in SimLex-999 and WordSim-353. *Data & Knowledge Engineering*. V. 115:174-193.
- Loukachevitch N., Lashevich G., and Dobrov B. 2018. Comparing Two Thesaurus Representations for Russian. *Proceedings of Global WordNet Conference GWC-2018*:35-44.
- Loukachevitch, N. 2019. Corpus-based Check-up for Thesaurus. *Proc. of the 57th Annual Meeting of the Association for Computational Linguistics ACL-2019*: 5773-5779.
- Majewska, O., McCarthy, D., van den Bosch, J., Kriegeskorte, N., Vulić, I., and Korhonen, A. 2020. Spatial Multi-Arrangement for Clustering and Multi-way Similarity Dataset Construction. In *Proceedings of The 12th Language Resources and Evaluation Conference*: 5749-5758.
- McCrae, J., Rademaker, A., Bond, F., Rudnicka, E., and Fellbaum, Ch. 2019. English WordNet. An Open Source WordNet for English. In *Proceedings of Global WordNet Conference GWC-2019*.
- Mikolov, T., Chen, K., Corrado, G. S., and Dean, J. (2013). Efficient estimation of word representations in vectorspace. *CoRR*, abs/1301.3781.
- Ozhegov, S.I. and Shvedova, N.Yu. 1997. *Explanatory Dictionary of the Russian Language*. Russian Academy of Sciences. Institute of the Russian Language named after V.V. Vinogradov. - 4th ed., Add. - M.: Azbukovnik,. (in Russian).
- Petrenko, V.F. 2010. *Fundamentals of psychosemantics*. M.: Eksmo. (in Russian)
- Piasecki, M., Marcińczuk, M., and Maziarz, M. 2013. WordNetLoom: a WordNet development system integrating form-based and graph-based perspectives. *International Journal of Data Mining, Modelling and Management*, 5(3) :210-232.
- Rudnicka, E., Maziarz, M., Piasecki, M., and Szpakowicz, S. 2012. A strategy of Mapping Polish WordNet onto Princeton WordNet. In *Proceedings of COLING 2012* :1039-1048.
- Sinopalnikova, A. 2004. Word association thesaurus as a resource for building wordnet. *Proceedings of the 2nd International WordNet Conference*: 199-205.
- Soloviev, V., Gimaletdinova, G., Khalitova, L., and Usmanova, L. 2020. Expert Assessment of Synonymic Rows in RuWordNet. *Proc. of Analysis of Images, Social Networks and Texts AIST-2020 conference*. CCIS, vol. 1086:174-183.
- Usmanova, L., Erofeeva, I., Solovyev, V., and V. Bochkarev. 2020. Analysis of the Semantic Distance of Words in the RuWordNet Thesaurus. 2020. *Proceedings of the DAMDID/RCDL'2020 Conference*.
- Vylomova, E., Shcherbakov, A., Philippovich, Y., and Cherkasova, G. 2018. Men Are from Mars, Women Are from Venus: Evaluation and Modelling of Verbal Associations. In: *Analysis of Images, Social Networks and Texts. AIST 2017*. Lecture Notes in Computer Science, vol 10716. Springer, Cham:106-115.
- Wojcik, E. H., and P. Kandhadai. 2019. Paradigmatic associations and individual variability in early lexical-semantic networks: Evidence from a free association task. *Developmental Psychology*. 56(1):53-69.