

Assessment of Comparative Abstractness: Quantitative Approach

Mariia Andreeva^{1,2}[0000-0002-5760-093], Marina Solnyshkina²[0000-0003-1885-3039],
Artem Zaikin³[0000-0002-5596-3176], Olga Bukach²[0000-0001-8675-2243],
Radif Zamaletdinov⁴[0000-0002-2692-1698]

¹ Kazan State Medical University, 420012, Butlerova, 49, Kazan, Russia

² Kazan Federal University, Department of theory and practice of language teaching, Research laboratory 'Intellectual technologies for text management', 420008, Kremlyovskaya, 18, Kazan, Russia

³ Kazan Federal University, Laboratory of modern geoinformation and geophysical technologies, Department of mathematical statistics and information technologies, 420008, Kremlyovskaya, 18, Kazan, Russia

⁴ Kazan Federal University, 420008, Kremlyovskaya, 18, Kazan, Russia
mesoln@yandex.ru

Abstract. Estimation of abstractness ratings of Russian words have remained almost entirely unexplored, though by contrast, a vast literature has addressed questions relating to the same notion in English. In this article we present the results of the experiment aimed at quantitative assessment of abstractness / concreteness of the top 1000 Russian words and contrasting their ratings with those of American English equivalents. We also performed a quantitative verification of abstractness ratings for polisemous words in Russian. The latter was conducted based on the survey results in which respondents were exposed to at least two collocations explicating separate senses of the word. Overall, our study reveals consistent similar patterns in respondents' ratings of concrete and abstract words in Russian and English. The findings of empirical dissimilarities in A/C ratings for different senses of one word validate the suggested method of presenting not single words but collocations. The results provide additional more fine-grained data on A/C ratings which could be employed as useful indicators of text complexity and features in a multi-factor text analysis.

Keywords: Quantitative Assessment, Abstract and Concrete Nouns, Experimental Study, Polysemous Words.

1 Introduction

The quantitative representation of concreteness-abstractness continuum is viewed as a primary task in semantic networks analysis; therefore its theoretical importance is widely acknowledged [1]. In the modern scientific paradigm, abstract / concrete words discrimination is based on the assumption that concrete words denote referents

Copyright © 2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

which can be experienced through sense, i.e. available to the senses, whereas referents nominated with abstract words lack the attribute and refer to ideas or concepts [2].

The notion of abstractness / concreteness is nowadays a focus of numerous studies [3] and for a few decades the problem of discriminating concrete and abstract words has been viewed as relevant by researchers in a number of areas: linguistics, psychology, pedagogy, medicine etc. At present, ratings of abstract/concrete words or A/C ratings are used in studies of statistical models of word distribution [4], Text Leveling Systems for ranging texts in difficulty thus profiling them for different categories of readers as well as in literacy education where ratings are implemented to help students with learning difficulties [5]. The spectrum of modern research in the area varies from the problems of mental performance and processing [6] to psychological disorders [7] and global aphasia [8].

However, the study of differences in perceiving bilingual equivalents in two unrelated languages, namely Russian and English, to the best of our knowledge, has never been pursued though it has a big potential to contribute both to the general theory of abstractness/concreteness and intercultural studies.

The Research Questions we aimed at in this study are as follows:

RQ 1: How similar or different are A/C ratings polysemous words presented to respondents separately and as parts of collocations?

RQ 2: How similar or different are A/C ratings of the top 1000 Russian words and their English equivalents?

2 Related work

2.1 Cognitive processing of abstract and concrete words: speed, emotions, associations

The modern paradigm of research in the area designed and developed a number of methods and techniques to rate the perception of abstract and concrete words. The data obtained by Kroll [8] in the study of lexical judgments on abstract and concrete words suggest that abstract words take more time to be comprehended and as such they are processed by human brain significantly longer as compared to concrete words. The works by Kiehl [9] and Noppeney [10] introduced a notion of ‘the concreteness effect’ confirming the idea that concrete words are processed more efficiently and faster than abstract ones. Another relative finding received after a series of associative experiments provides evidence that in the majority of cases while processing abstract words people experience emotions [11]. Moffat et al [12] added to this notion arguing that if people experience emotional involvement while conducting verbal semantic categorization, they process abstract words faster. More affective associations with abstract words received from informants led scholars to the introduction of the so-called ‘imageability variable’ that measures how easily a mental image can be formed for a concept. The authors proved that concrete words are more imageable than abstract. It has also been proved that, speakers tend to develop more associations with words bearing a higher degree of abstractness than concrete words [13].

2.2 Questionnaire survey as a computational method to compile A/C dictionaries

One of the first methods of compiling an A/C dictionary was developed in psycholinguistics where the degree (or level) of A/C of words was estimated by native speakers of the language on a limited scale from ‘the most abstract’ to ‘the most concrete’. E.g. MRC Psycholinguistic database was also compiled based on the survey followed by the Semantic Differential Measurement technique in which respondents were asked to assess the degree of A/C of words on a bipolar scale ranging from the most abstract (100) to the most concrete (700) [14].

Since 1981, MRC Psycholinguistic database has been extensively used to generate different resources including A/C dictionaries. One of the first English Dictionaries of 4.000 abstract / concrete words released in 1981 [15] still serves as a reference list in various research [16–19]. The latest edition of Abstract / Concrete Words Dictionary comprises A / C ratings of 37058 English words and 2896 two-word collocations (such as zebra crossing and zoom in), obtained from over 4.000 participants by means of a norming study based on the data collected by crowd-sourcing [2].

2.3 Corpora used to compile A/C dictionary: vector-based analysis

Numerous studies in different languages corpora prove that ‘concrete words co-occur with other concrete words, whereas abstract words co-occur with abstract words’ [19–22].

The analysis conducted by B. Sneffjella et al. [13] on the Corpus of Historical American English (COHA) revealed an increasing tendency to use concrete words more frequently than abstract words in English. Based on the historical data analysis, B. Sneffjella et al. also argue that the average concreteness in English texts tends to grow over years [13]. The corpus-based approach also enabled researchers to explore possibilities of automated compilation of dictionaries of abstract/concrete words. E.g. the authors of [13] suggested and applied a method of designing a COHA-based dictionary of abstract/concrete words. They started with eliciting a core list of obviously abstract and concrete words from COHA. For each word in the core list, with the help of word embeddings method, they constructed a vector characterizing the word’s joint occurrences with other words in the corpus. Further, they measured the distance between each word and vectors of other words in the core list. That allowed assessing the degree of closeness of a given word to a concrete or an abstract extreme. The dictionary constructed with this method was compared with the dictionary based on MRC database [2]. Spearman’s correlation coefficient of the two dictionaries estimated at 0.70 is statistically significant thus establishing a high level of reliability of the data registered in the COHA-based dictionary of abstract/concrete words.

In education, A/C ratings are used to assess text readability / complexity [23, 24]. E.g., in works of D. McNamara, lists of abstract words and online A / C ratings are used as resources for the automated tools, such as Coh-Metrix, TAACO, SiNLP developed to profile texts and teach effective comprehension [23].

Two dictionaries of 64000 and 88000 Russian nouns tagged with numerical estimates of abstractness/concreteness were compiled by a group of Russian researchers [19]. Both dictionaries are based on the Google Ngram Books Corpus and “the core list of obviously abstract and concrete words” from the Dictionary of the Russian language [25]. Based on the extracted from the Google Books Ngram bigrams [1] generated two sets of words (Nabs, Ncon) and their collocations contexts thus providing a solid foundation for measuring degrees of A/C of the selected words.

3 Materials and methods

The material used in the current study comprises two sets of data: (1) The List of 1000 most Frequent Russian Words [26]; (2) the list of their English equivalents registered in MRC collected at a preliminary stage of the research.

The initial analysis of the top 1000 Most Frequent Russian nouns indicated that they are primarily polysemous. For this reason we had to reconsider the Semantic differential measurement technique and design an additional stage in the experiment in which the words in question were presented to respondents not as separate words but in collocations in which different senses of the word were explicated in different collocations. E.g. *mesto* (Eng. place.): (1) *mnogo mesta* (Eng. a lot of space), (2) **rabochee mesto** (Eng. working place); **vzglyad** (Engl. gaze): (1) *obvesti vzglyadom* (Engl. look around); (2) *politicheskie vzglyady* (Eng. political views).

The research was conducted in three stages:

Stage 1. Defining the ratings of abstractness of the top 1000 most frequent Russian nouns with the help of the Semantic differential measurement technique and thus forming the List of Most Frequent Russian Words with A/C ratings (FRAC 1000).

Stage 2. Contrasting the levels of abstractness of the top 1000 Russian and their English (American) equivalents.

Stage 3. Defining A/C ratings of separate senses of polysemous words in FRAC 1000.

Defining the ratings of abstractness of the top 1000 Russian most frequent nouns as abstract / concrete, **Stage 1**, was performed with the help of an online semantic differential measurement technique involving experts' assessment.

We designed 20 Google online survey forms to measure the rating of abstractness / concreteness of 1000 most frequent Russian nouns with the help of Semantic differential measurement technique. Top 1000 Russian nouns were obtained from the frequency dictionary of Sharov and Lyashevskaya (2009).

Semantic differential measurement technique presents a questionnaire in which people are asked to rate a word within a five-point abstractness / concreteness scale. The 1st position on the left corresponds to ‘the highest level of concreteness’, the 2nd corresponds to ‘a high level of concreteness’, the 3rd position corresponds to ‘bearing equal levels of concreteness and abstractness’, the 4th is ‘a high level of abstractness and the 5th position corresponds to ‘the highest level of abstractness’ (see Fig. 1) [27]. The respondents evaluated the level of abstractness of each word by choosing a number on the scale based on the perceptions the word produces in their mind. Participation in the Survey was completely anonymous and voluntary.

The respondents (n = 800) are full-time University students, native speakers of Russian aged 17 – 25. All the respondents signed the consent form approved by the Local Ethic Committee to participate in the study [28].

mesto (place)

	1	2	3	4	5	
concrete	<input type="radio"/>	abstract				

krug (circle)

	1	2	3	4	5	
concrete	<input type="radio"/>	abstract				

Fig. 1. A bipolar scale to rate level of abstractness/concreteness

In each of the 20 Google online survey forms respondents rated 50 words. The A/C ratings of each Russian word were computed as an average of all the assessments received in the range from 1 to 5 (see Fig. 1). The number of respondents varied from 40 to 52. The ratings obtained from the respondents per word were registered and the total and average A/C ratings for initial scale value were calculated (see Fig. 2). E.g. the rating of the word *polozhenie* (Eng. position) was assessed as 1 by three respondents, as 2 by five respondents, 3 by 19 respondents, rating 4 eleven times, and rating 5 fifteen times. So the average (see the bottom line in Table 1) was estimated as 3.6 (Table 1).

Table 1. Initial scale calculating average A / C ratings in FRAC 1000

Respondents	mesto (place)	tsel' (goal)	polozhenie (position)	krug (circle)	tehnika (technique)
	initial scale value				
1	4	4	5	3	1
2	1	5	3	1	1
3	4	1	3	1	3
4	5	5	4	2	2
5	5	4	5	1	3
...
38	1	5	4	2	3
39	5	3	5	5	5
40	3	2	4	1	3
Total	122	124	189	100	139
A/C ratings	3,0	2,8	3,6	2,3	2,6

As a result of the study we computed FRAC 1000 available at [29] under the heading Frequency Russian Abstractness-Concreteness 1000 [27].

Stage 2. Contrasting the levels of abstractness of the top 1000 Russian words and their American English equivalents.

On the second stage of the research all the words in FRAC 1000 were translated into English and the A/C ratings of Russian words and their English equivalents were contrasted. The A/C ratings of the English equivalents were elicited from MRC Psycholinguistic Database [14]. In similar experiments aimed at compiling an MRC Psycholinguistic database [14] researchers applied a 7-point inverted scale with 7 being the most concrete word and 1 – the most abstract score. Thus, to contrast the A/C ratings of American English and Russian words we had to unify the two scales. For this purpose we inverted the Russian scale using the formula:

$$S_i = 6 - x, \quad (1)$$

where x is the initial value in the Russian questionnaire.

Afterwards, the scale was stretched based on the formula:

$$100 (1.5 (S_i - 1)) + 1 \quad (2)$$

In this way, we got identical scales for evaluating the level of abstractness: from 700 to 100 with the highest A/C to be 700 and the lowest A/C – 100 [27] (see Table 2 for the inverted scale value). E.g. the word *krug* (Eng. circle) received the A/C rating of 501 while the word *polozhenie* (Eng. position) was assessed as a word with the A/C of 315.

Table 2. The Inverted scale values of A/C ratings in FRAC 1000

Respondents	mesto (place)	tsel' (goal)	polozhenie (position)	krug (circle)	tehnika (technique)
	inverted scale value				
1	250	250	100	400	700
2	700	100	400	700	700
3	250	700	400	700	400
4	100	100	250	550	550
5	100	250	100	700	400
...
38	700	100	250	550	400
39	100	400	100	100	100
40	400	550	250	700	400
Total	16550	18800	16700	21550	24200
A/C ratings	408	427	315	501	457

A small subset of the obtained results presented in Fig.2 indicates that A/C ratings fluctuate in the range of 270 in *mysl'* (Eng. thought) to 638 in *krovat'* (Eng. bed).

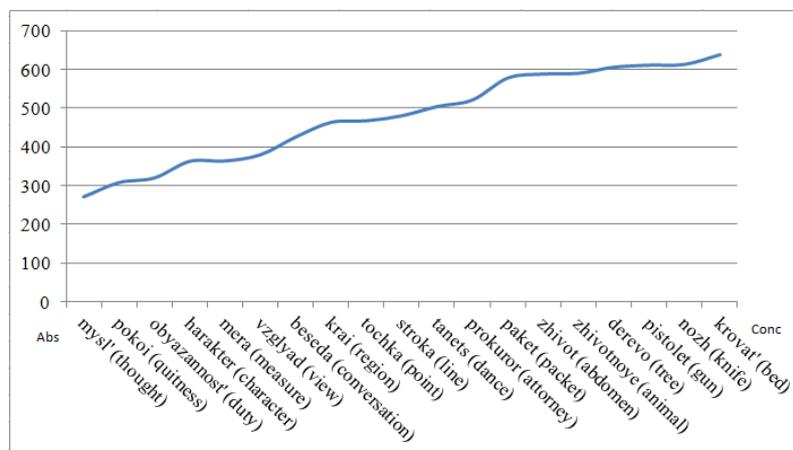


Fig. 2. A/C ratings of Russian nouns (FRAC 1000)

As MRC dataset does not register 227 American English equivalents of the Russian words in the List of 1000 Most Frequent Words, the finalized list contains 773 Russian words and their English equivalents tagged with A/C ratings. The subset from the complete Table of Russian / English equivalents in Table 3 demonstrates the words with the lowest and highest A/C difference.

Table 3. Russian /English equivalents with A/C ratings

#	Russian word	A/C metric (Rus)	A/C metric (Eng)	A/C Rus and A/C Eng difference	English word
1	sil'a	340	339	1	strength
2	derevo	606	604	2	tree
3	effekt	288	295	7	effect
4	tsvetok	566	584	18	flower
5	mozg	606	556	50	brain
...
770	strana	565	465	100	country
771	remont	543	394	149	repair
772	kontsert	492	252	240	concert
773	administratsiya	599	331	268	administration

Based on the differences in their A/C ratings we classified 773 Russian/English equivalents into three groups:

Group 1 comprises 611 words with similar A/C ratings in Russian and English (ratings difference < 33%). E.g. sil'a (340) / strength (339), effect (288) / effect (295), mozg (606) / brain (556).

In Group 2 there are 78 nouns with A/C ratings difference in the range of 34- 66 %. E.g. remont (543) / repair (394), reshenie (430) / decision (297), chislo (543) / number (395).

Group 3 is formed by 46 nouns perceived with drastically different in Russian and English, i.e. with ratings difference above 67 %. E.g. *kontsert* (492) / *concert* (252), *administratsiya* (599) / *administration* (331), *mesyats* (567) / *month* (345).

As we can see, 611 nouns which make the prevailing number of the assessed nouns show little or no difference in A/C ratings in Russian and English. The revealed differences may be explained either by cross-cultural differences or and homonymy. E.g. the word ‘surprise’ is rated with little difference in both languages – *surprise* (326) / *udivlenie* (357). The contexts, obtained from the British National Corpora ‘The news came as a complete surprise to workers at the Oxfordshire base’ [30] and Russian National Corpus ‘The project caused us a great surprise, and a pleasant one at the same time’ [31]. The word *scene* / *stena* is also perceived and rated with little difference – *scene* (408) / *stena* (496). E.g. ‘The village centre is once again the scene of chaos as the roads are being dug up, filled in and tarmaced over’ [30] ‘quite a beautiful scene of a trip through the waking city’ [31].

Stage 3. Defining A / C ratings of separate senses of polysemous words in FRAC 1000.

At Stage 3 we tested the hypothesis that separate senses of polysemous words or homonyms bear different A/C ratings.

The effect of polysemous words on text comprehension has been studied by Mason, 1979, Williams, 1992, Paul, 1988. Devorah E. Klein and Gregory L. Murphy in [32] conclude that “polysemous words have separate representations for each sense” in the brain and as such in studies on abstractness their senses are supposed to be viewed separately. Willims in [33] the study aimed at examining whether the various meanings of polysemous adjectives (e.g., *firm* as in *solid* or *firm* as in *strict*) are functionally independent in language comprehension” reports that only “central” meanings of polysemous words become active in comprehension even if they are irrelevant in the context [33].

Lexicographic analysis demonstrated that only 206 words from FRAC 1000 [29] bear more than one sense. To discriminate A/C ratings of two separate senses of each of the 206 words we retrieved a collocation exemplifying it either from [34] or Russian National Corpus [31] which are viewed as the most reliable sources. E.g., [MAS] registers five senses of the word **sfera** (Engl. *sphere*): 1. *A ball or its inner surface*; 2. *Mat. A closed surface, all points of which are equally distant from the center; surface of a ball*; 3. *The space within the range of smth. As well as the scope of smth.;. an area of activity, interest, or expertise; a section of society or an aspect of life distinguished and unified by a particular characteristic*; 4. *public environment, environment, setting*; 5. *(spheres, spheres) with a definition. The circle of persons united by a common social status or occupation. Of the five above we selected (1) and (3) functioning in the Russian discourse in the following collocations: (1) **zemnaya sfera** (lit. sphere of the Earth, i.e. biosphere, hydrosphere and lithosphere) and (3) **sfera zdra-vookhraneniya** (lit. a sphere of healthcare).*

Later all the selected collocations were grouped in seven online Google forms with no more than 30 words (i.e. 60 collocations) per form (see Fig. 3).

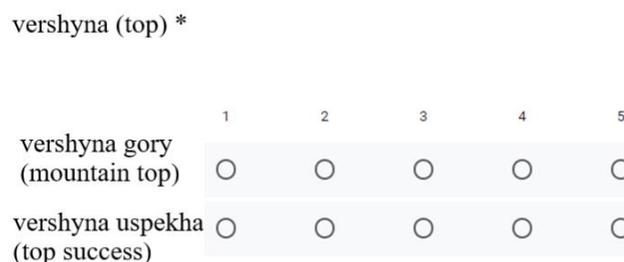


Fig. 3. A/C scales of the collocations with VERSHINA (lit.top)

Next, similarly to the procedure described above, respondents (n=280), native Russians aged 18 – 60, were requested to rate A/C of two senses of polysemous words on a five-point scale.

The A/C ratings were further inverted to be compatible with the results, achieved at the previous stages.

Further, the A/C ratings of separate senses were contrasted twice: (1) with each other and (2) with A/C ratings of the word assessed earlier as a semantic whole. E.g. for the word **doroga** (Eng. *road, way*), we contrasted (1) A/C ratings of two senses realized in **prosyolohnaya doroga** (Eng. *countryside road*) (192) and **sobirat'sya v dorogu** (Eng. *set off for a trip / road*) (475); (2) each sense rating with the A/C rating of the word assessed in the previous experiment, i.e. 199. As we can see in this particular case two ratings, i.e. *countryside road* and *road*, are similar (192 vs 199) while those of *set off for a trip / road* and *road* differ considerably. The former may indicate that while comprehending the word **doroga** (Eng. *road, way*) Russians tend to visualize **prosyolohnaya doroga** (Eng. *countryside road*) and this particular sense is the brightest of all the senses registered in [34]. Table 4 presents differences in A/C ratings of polysemous words.

Table 4. The difference of the A/C ratings revealed for polysemous words and their word combinations

#	word A/C	abstract sense A/C	concrete sense A/C	A/C differences in collocations (in descending order)
1	Duty (Rus. dolg) 470	A duty to the country (Rus. dolg pered otechestvom) 522	Money debt (Rus. denezhny dolg) 151	371
2	Wave (Rus. volna) 547	Wave of protests (Rus. volna protestov) 496	A sea wave (Rus. morskaya volna) 148	348
3	Place (Rus. mesto) 407	A lot of space (Rus. mnogo mesta) 451	Working place (Rus. rabochee mesto) 189	262
4	Face (Rus. litso) 505	Face of a project (Rus. litso proekta) 455	Features of a face (Rus. cherty litsa) 199	256
5	End (Rus. konets) 332	End of year (Rus. konets goda) 335	End of a table (konets stola) 175	160
6	Party (Rus. partiya) 490	To win a chess party (Rus. vyigrat' partiyu v shakhmaty) 291	A party of goods (Rus. partiya tovarov) 235	56

Fig.4 visualizes the range of differences discovered in A/C ratings when they are estimated in single words and in collocations.

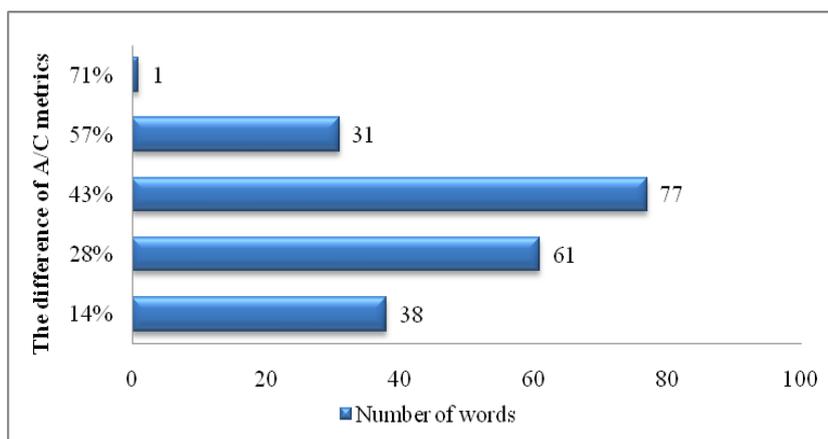


Fig. 4. A/C ratings difference in collocations

The maximum difference in A/C ratings (71%, over 400 points) is revealed in the word *povorot* (Eng. turn) defined as 1. ‘the place where a road turns, deviates to the side’; 2. ‘complete change in the development of something’ [34]. The perception of combinations of *povorot nalevo ot doma* (the left turn from home) (129) and *povorot sud’by* (a twist of fate) (540) indicates the distinction between concrete and abstract ratings rated by respondents.

57% difference (in the range of 301 – 400 points) in A/C ratings is estimated for 31 words. E.g. *golova* (Eng. head) defined as (1) ‘the upper part of the human body, the upper or anterior part of the animal body containing the brain’ [34] is more concrete than in (2) ‘mind, consciousness; reason’ [34]. The metrics obtained for (1) in *golova bolit* (sb’s head aches) (158) and (2) *dumat’ na svezhuyu golovu* (Eng. lit. think on a clear head) (465) differ significantly.

The phrases *massa tela* (Eng. a body mass) (213) and *massa vpechatleniy* (lots of impressions) (525) correspond to two meanings of the word *massa* (Eng. a mass, lot of) defined as 1. ‘one of the main physical characteristics of matter, which is a measure of its inertial and gravitational properties’; 2. ‘a large number, a lot of things’ [34].

43% difference of the A/C ratings, i.e. within the range of 201 – 300-points, is determined for 77 nouns. In particular, the word *znak* (Eng. sign) when used in the collocation *dorozhny znak* (Eng. road sign) explicates a more concrete sense, i.e. ‘an image with a certain conventional meaning’ [34], rated at 174. The collocation *znak soglasiya* (Eng. sign of consent) realizes another sense, i.e. ‘an external detection, manifestation of something, evidence, sign of something’ [34] which was rated at 465. A relatively high difference of 291 units between the two senses in two collocations demonstrates a clear distinction between concrete and abstract meanings of the word.

The temporal semantic element in the meaning of the word *obed* (Eng. lunchtime) [34] has a low A/C rating of 426, when used in the collocation *zakryt' kabinet na obed* (Eng. close an office for lunchtime). When comprehended as 'food, dishes' [34] in the phrase *obed iz tryokh blyud* (Eng. three-course lunch), the word *obed* (Eng. lunch) is more concrete with A/C at 178.

Senses in 61 words exhibit 28%, i.e. 100–200 points, difference in A/C ratings. E.g., the sense 'something published or is being published (about books, magazines, etc.)' [34] in the collocation *periodicheskoe izdanie* (Eng. periodical) is estimated at 267, while in the collocation *izdanie ukaza* (Eng. release of a decree) the rating is 386 exemplifying a more abstract meaning of the word *izdanie* (Eng. issue, production) [34]. Similarly, the sense 'color, coloration, and also a shade of some color that differs in the degree of brightness, saturation' [MAS] in the collocation *zelyonye tona* (shades of green) was assessed by respondents at 241. While *ne govori takim tonom* (Eng. do not speak in this manner) eliciting the sense 'the character, tone of the sound of an instrument or voice' [34] received A/C rating of 429.

The minimal difference of A/C ratings 14 %, lower than 99 points, is found in 38 words. E.g., the A/C rating revealed for the sense 'a certain, usually significant, quantity of some items, goods, etc' [34] in the collocation *partiya tovarov* (Eng. a set of goods) 235. The meaning 'a game (chess, cards, etc.) from beginning to end' [34] of the word *partiya* (game) presented to respondents in the collocation *vyigrat' partiyu v shakmaty* (Eng. to win a game of chess) was estimated as 291.

The sense of the word *vzglyad* (Eng. gaze) is more concrete when defined as 'direction of the eyes, view of someone, something' [34] in the collocation *obvesti vzglyadom* (Eng. look around): it's A/C rating is 356. *Vzglyad* (Eng. opinion) in *politicheskie vzglyady* (Eng. political views) received 424. As we can see the difference in the respondents' perception of these two senses in collocations is insignificant. Obviously, both collocations with the word *partiya* (Eng. a set, game) are perceived as more concrete (the rating does not exceed in its range 200), while both ratings of the word *vzglyad* (Eng. gaze, opinion) are perceived as more abstract with the corresponding ratings of about 400.

4 Conclusion

The empirical analysis of abstractness/ concreteness nature of the top 1000 Russian words confirmed a striking similarity of the majority of these words with their American English equivalents. A quantitative analysis of separate senses of each word under study revealed differences in their perception of native speakers. The findings indicate that ratings of separate senses of the words are to be assessed in collocations exemplifying one sense only. In summary, the research made available abstractness/concreteness ratings of 1000 Russian words, which expands future theoretical research of abstractness/ concreteness effect on text readability. FRAC 1000 compiled as a result of the study may be viewed as a valuable resource for numerous comparative and contrastive studies involving rating the words or texts A/C levels. Another area of the research results implementation is a multi-factor automated text analysis.

It takes into account an important conceptual feature of abstract/concrete discrimination when conducting cross-language research such as bilingualism and translation profiling texts for certain categories of readers. The algorithm to compute A/C ratings of Russian nouns introduced and implemented in the research may be used for other parts of speech in different languages.

5 Acknowledgements

This research was financially supported by grant RFBR 19-07-00807.

References

1. Solovyev, V, Ivanov, V.: Automated Compilation of a Corpus-Based Dictionary and Computing Concreteness Ratings of Russian. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12335, 554–561 (2020).
2. Brysbaert, M. Warriner, A. B., Kuperman, V.: Concreteness ratings for 40 thousand generally known English word lemmas. In: Behavior research methods, 46(3), 904–911 (2014).
3. Reuter, K., Werning, M., Kuchinke, L., Cosentino, E.: Reading words hurts: the impact of pain sensitivity on people’s ratings of pain-related words. In: Language and Cognition, 9(3), 553–567 (2017).
4. Paivio, A., Allan, J. Clark, M.: Dual coding theory and education. Draft chapter presented at the conference on Pathways to Literacy Achievement for High Poverty Children at The University of Michigan School of Education (2006).
5. Hines, D.: Recognition of verbs, abstract nouns and concrete nouns from the left and right visual half-fields. *Neuropsychologia*, 14(2), 211–216 (1976).
6. Binney, R.J., Zuckerman, B., Reilly, J.: A Neuropsychological Perspective on Abstract Word Representation: From Theory to Treatment of Acquired Language Disorders. *Current neurology and neuroscience reports*, 1–26 (2018).
7. Crutch, S.J., Warrington, E.K.: The differential dependence of abstract and concrete words upon associative and similarity-based information: Complementary semantic interference and facilitation effects. *Cognitive neuropsychology*, 27(1), 46–71 (2010).
8. Kroll, J.F., Merves, J.S.: Lexical access for concrete and abstract words. In: *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12(1), 92-107 (1986).
9. Kiehl, K.A., Liddle, P.F., Smith, A.M., Mendrek, A., Forster, B.B., Hare, R.D.: Neural pathways involved in the processing of concrete and abstract words. *Human brain mapping*, 7(4), 225–233 (1999).
10. Noppeney, U., Price, C.J.: Retrieval of abstract semantics. *Neuroimage*, 22(1), 164–170 (2004).
11. Ponari, M., Norbury, C.F., Vigliocco, G.: Acquisition of abstract concepts is influenced by emotional valence. *Dev. Sci*, 2(2) (2018).
12. Moffat, M., Siakaluk, P.D., Sidhu, D. M., Pexman, P.M.: Situated conceptualization and semantic processing: effects of emotional experience and context availability in semantic categorization and naming tasks. *Psychonomic bulletin & review*, 22(2), 408–419 (2015).
13. Snefjella, B., Génereux, M., Kuperman, V.: Historical evolution of concrete and abstract language revisited. *Behavior research methods*, 51(4), 1693–1705 (2018).

14. Coltheart, M.: The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, 33(4), 497–505 (1981).
15. Warschauer, M.: Motivational aspects of using computers for writing and communication. *Telecollaboration in foreign language learning*, 29–46 (1996).
16. Deschamps, I., Baum, S.R., Gracco, V.: Phonological processing in speech perception: What do sonority differences tell us? *Brain and language*, 149, 77–83 (2015).
17. Alexeeva, S.V., Slioussar, N.A., Chernova, D.A.: *Stimulstat: a database for linguistic and psychological studies on Russian language* (2016).
18. Hubbard, J., Molapour, T., Morsella, E.: The Subjective Consequences of Experiencing Random Events. *International Journal of Psychological Studies*, 8(2) (2016).
19. Solovyev, V.D.; Ivanov, V.V.; Akhtiamov, R.B.: Dictionary of Abstract and Concrete Words of the Russian Language: A Methodology for Creation and Application. *Journal of Research in Applied Linguistics*, 10, 215–227 (2019).
20. Frassinelli, D., Naumann, D., Utt, J., im Walde, S.S.: Contextual characteristics of concrete and abstract words. In: *IWCS 2017–12th International Conference on Computational Semantics–Short papers* (2017).
21. Naumann, D., Frassinelli, D., im Walde, S.S.: Quantitative semantic variation in the contexts of concrete and abstract words. In: *Proceedings of the Seventh Joint Conference on Lexical and Computational Semantics*, 76–85 (2018).
22. Bhaskar, S.A., Köper, M., im Walde, S.S., Frassinelli, D.: Exploring multi-modal text+image models to distinguish between abstract and concrete nouns. In: *Proceedings of the IWCS workshop on Foundations of Situated and Multimodal Communication* (2017).
23. Crossley, S.A., Skalicky, S., Dascalu, M., McNamara, D.S., Kyle, K.: Predicting text comprehension, processing, and familiarity in adult readers: New approaches to readability formulas. *Discourse Processes*, 54(5-6), 340–359 (2017).
24. Dashtestani, R.: EFL teachers' and students' perspectives on the use of electronic dictionaries for learning English. *CALL-EJ*, 14(2), 51–65 (2013).
25. Ozhegov, S.I., Shvedova, N.Yu.: *Tolkovyj slovar' russkogo yazyka* (1992).
26. New frequency dictionary, <http://dict.ruslang.ru/freq.php>, last accessed 2020/10/30
27. Solovyev, V., Andreeva, M., Solnyshkina, M., Zamaletdinov, R., Danilov, A., Gaynutdinova, D.: Computing Concreteness Ratings of Russian and English Most Frequent Words: Contrastive Approach. In: *2019 12th International Conference on Developments in eSystems Engineering (DeSE)*, Kazan, Russia, 403–408 (2019).
28. Zhuravkina, I., Soloviev, V., Lobanov, A.: Comparative Analysis of Concreteness/Abstractness of Russian Words. In: *Conference of Open Innovation Association, FRUCT*, 464–470 (2020).
29. Technologies of semantic dictionaries compilation, <https://kpfu.ru/tehnologiya-sozdaniya-semanticeskikh-elektronnyh.html> FRAC 1000, last accessed 2020/10/30.
30. British National Corpus, <https://www.english-corpora.org/bnc/>, last accessed 2020/10/30
31. Russian National Corpus, <https://ruscorpora.ru/new/>, last accessed 2020/10/30.
32. Klein, D.E., Murphy, G.L.: The representation of polysemous words. *Journal of Memory and Language*, 45(2), 259–282 (2001).
33. Williams, J.N. Processing polysemous words in context: Evidence for interrelated meanings. *Journal of Psycholinguistic Research*, 21(3), 193–218 (1992).
34. Concise academic dictionary, <https://gufo.me/dict/mas>, last accessed, 2020/10/30.