Comparative Analysis of Concreteness / Abstractness of Russian Words

Irina Zhuravkina
Belarusian State Pedagogical University
named after M. Tank
Minsk, Belarus
irishazhu@mail.ru

Alexander Lobanov
Belarusian State Pedagogical University
named after M. Tank
Minsk, Belarus
kaf-vipp@bspu.by

Valery Soloviev
Kazan Federal University
Kazan, Russia
maki.solovyev@mail.ru

Andrey Danilov
Kazan Federal University
Kazan, Russia
tukai@yandex.ru

Abstract – The article presents the results of studying the perception of Russian words by parameters: concreteness / abstractness. The purpose of the research: creation of a special dictionary where the degree of its abstractness is specified for each word. The research used semantic differentiation method, where respondents evaluated the degree of abstractness / concreteness of the word on a scale from 1 to 5 in online mode. The sample of 154 respondents in the research was used. A 500-word analysis was made. The survey was conducted in Russia. The analysis of words allowed classifying them into three categories: abstract / concreteness / undifferentiated.

I. INTRODUCTION

Modern man is in the global information space, which is important to comprehend and structure. The person in the process of obtaining, processing and reproducing information has become the subject of study in a number of sciences. As V. V. Krasnykh notes, “there is a shift in scientific research to the study of a real person in real conditions of real life and real communication” [1, p. 173]. A personality is viewed through the prism of its communication, and, consequently, of language and speech. The ways of perception and assimilation of information flows by a person are increasingly determined by language constructs and linguistic properties of words. Research on the impact on the quality of information assimilation of the properties of speech utterances, such as the number of syllables in a word, regularity or irregularity of spelling of words, imagery / abstractness / concreteness of words, frequency of their use in a print publication, is becoming relevant.

Research at the intersection of sciences, including the work in the field of mathematical linguistics, the use of computerized psycholinguistic databases that contribute to understanding the language and creating a literate text in every way, is becoming popular.

The mediation (interface) function of a word as a means of access to a single information base of a person also becomes a focus of attention. Based on interface word theory A. A. Zalevskaya [2] defines the function of a word as a means of access to a single information base of an individual, which is a multifunctional product of a perceptual, cognitive and emotional-evaluative processing of a person’s knowledge and communication experience. The author defines the word as a unit of the mental lexicon of a person and a means of access to the individual consciousness of a person. In the process of awareness and verbalization the word is identified and the text is understood. Cognitive traits can be more or less specific in terms of abstractness, in terms of generalization, and figurative, informational, and interpretative in content.

Z. D. Popova [3] notes that a cognitive picture of the world is formed in the consciousness of a native speaker, which includes the following structural components of the concept: image, information content, and interpretation field. Thus, the sensory image combines perceptual, cognitive and figurative signs; information content of concepts, reflects artifacts, scientific concepts, conceptualizing abstract entities. In addition, the interpretation field of the concept includes cognitive characteristics of the carrier of verbal utterances.

The concept of abstraction is one of the most controversial and difficult to understand. The problem of determining the abstract arises in the process of determining the boundary of the abstract and the concrete. Consideration of this issue requires access to data from disciplines such as philosophy, logic, psychology. Logic considers the concept of abstraction in connection with various kinds of concepts. In philosophy the concepts of “concreteness / abstract” are among the fundamental categories. In psychology the study of the above phenomena takes place in the context of the development of cognitive processes, which presupposes the flow of knowledge from various representation modules, which involve taking into account the data of the psychology of perception, memory and thinking.

In the humanitarian encyclopedia by V. S. Shvyrev and G. I. Ruzavin [4] the abstract and the concrete are both interconnected and opposite in meaning of the concepts of philosophical, scientific and everyday discourses. In their
relationship they express the manifestation of unity between abstract and concrete knowledge. The concept "abstract" is defined as a conceptualized mental image obtained by distracting (abstracting) from certain non-essential properties or relations of an object in order to highlight its essential features.

A theoretical analysis of the stated problem is quite fully presented in the article [5]. Therefore, adhering to the theoretical and methodological foundations mentioned above, we restrict ourselves to placing emphasis on individual issues.

The purpose of our research is to create a special dictionary where the degree of its abstractness is specified for each word. The research objectives: firstly, conducting a comparative analysis of words on the concepts of concreteness / abstractness in Russian and English to find out if there is a difference in the perception of these categories by native speakers of Russian and English; secondly, the definition of the boundaries of concreteness / abstractness / non-differentiation of words.

II. RELATED WORK

The research of abstractness / concreteness (C / A) goes back to the works of A. Paivio, J. Yuill and S. Madigan, which, using subjective scaling and an associative experiment, resulted in figurative coefficients, concreteness - abstractness for 925 English nouns [6]. Figurativeness, according to A. Paivio, J. Yuill and S. Madigan, the ability of subjects to present the meaning of the word as an image of any modality (for example, the image of the words sun, silence or heaviness), and specificity - abstractness - it is the availability of denotation (substantive attribution) of the meaning of the word to sensory perceptions. According to the results of the researches, the words with a strong emotional connotation, the imagery coefficients were highly rated, while the specificity coefficients were low, and many scientific terms had a high index on the specificity scale and a low figurative level.

Based on the theory of double coding, A. Paivio [7] notes that the verbal system is responsible for the abstract, logical, and figurative - for the specificity, analog way of thinking. In other words, verbal and figurative systems differ in the way their units are organized into higher-order structures.

In the research of V. F. Petrenko [8], carried out according to the scheme of A. Paivio on the material of Russian vocabulary, similar results were obtained. It has been established that the concepts of "concreteness", "abstractness" and "figurative" conceal psychic realities different for the subject. Emotionally colored concepts, as noted by V. F. Petrenko, associated with emotional conflicts in the emotional experience of an individual, have sensory fabric, which is experienced as a figurative concept.

J. Richardson [9] concludes that the subjects frequently reported the use of images in the associative learning of pairs consisting of highly shaped words, and in this situation we should expect a high level of reproduction of the material. At the same time, figurative instructions improve the results of reproduction at the expense of the larger organization of connections between the memory elements.

The most promising direction "artificial intelligence", according to D. A. Pospelov (see [8]), makes the development of cognitive graphics, where the solution of problem situations is realized in the framework of visual thinking. Cognitive graphics methods can be used when it is necessary to transform the textual description of tasks in their figurative representations, and in the generation of textual descriptions of the paintings that occur in the input and output blocks of intelligent systems, as well as in human-machine systems, designed for solving complex, poorly formalized tasks.

Experiment on the analysis of semantic distance and verification of semantic relations by L. Rips, E. Shoben and E. Smith allowed us to conclude that the concept of "animals" (more abstract from a logical point of view, than the concept "mammals") are closer in semantic distance to a concreteness animal names than the concept of "mammals" [10].

The representatives of the set-theoretic memory model such as: G. Bower, D. Meyer, B. Schaffer, R. Wallace are noted that the semantic distance between the concepts and, accordingly, the time of their sequential retrieval from memory is proportional to the number of common functional characteristics of their subject denotations [11].

V. F. Petrenko [8] noted that the level of development of values (abstract concepts) can be described through the forms of relationships which it enters with other meanings. Thus, the research of the functioning of values in human consciousness requires the consideration of actual psychological processes, which exist as individual values and personal meanings.

W. Kintsch [12] suggested that concrete nouns are grammatically simple and basic, whereas abstract nouns are derived from concreteness and thus are grammatically complex. J. R. Anderson and G. H. Bower have indicated that concrete words have fewer distinct values in the dictionary, but more semantic features than abstract. G. V. Jones suggested that the concreteness words may be considered as potential properties or "predicates" rather than abstract words [12].

As noted by F. A. Bleasdale [13], the degree of concreteness / abstractness of verbal information is important to consider in the research of human nature of representation and method of information coding. A. P. Lobanov [14] in their research noted that words acquire a concreteness or abstract nature in the consciousness of the subject and are grouped by him in an associative and / or conceptual way in accordance with the principle of typification or classification. The choice of grouping is determined by the mental experience of the subject, and the nature of the grouping demonstrates his cognitive development.

In a research of A. P. Lobanov [15] as a result of cluster analysis of indicators of intellectual development of the subjects, diagnosed using tests by R. Amthauer, D. Wechsler and G. Raven, they managed to localize verbal intelligence and structurally to present it as a vertical continuum of two factors: specific (SAs) and the abstract (SP) verbal intelligence and to offer definitions of two of the above factors in the paradigm of intelligence as an individual mental experience. Abstract intelligence formed a common subgroup with verbal scales test by D. Wechsler, and then together with concreteness...
intelligence (as a set associative abilities, the functioning of which is based on thematic representations and the mechanism of typification) were included in the cluster of higher order.

In linguistics different contents are put into the notion "abstractness" and " concreteness ". So, V. G. Huck [16] notes that such words that denote properties, relations, abstracted from material things, the state of things (kindness, reflection, causation) are called abstract. On the semantic level, the relation between abstract and concrete (e.g., human – human) words is reduced to interchanges of overseme and differential semes, that is metonymical transfer. Possibly, the use of the abstractness / concreteness of words connected with language and its connection with social conditions of existence. The abstract nature of the vocabulary noted by many linguists as a distinctive feature of the French language. V. G. Huck, who compared the French language with other languages, also notes its characteristic tendency to use the words of general value.

When perceiving and processing information by virtue of mental-linguistic unity, as D.M. Mironova notes in her research [17], the semantics of the representatives carries traces of visual, auditory and tactile sensory processes. The results obtained by the researcher confirm that the visual figurative component is leading.

In the study by K. Wiemer-Hastings and X. Xu [18] of the differences between concreteness and abstractness, the following results were obtained: 1) abstract concepts have less internal properties of objects and more properties expressing subjective experiences associated with the social aspects of situations; 2) they include more subjective signs that express the emotional and evaluative attitude of the speaker. At the same time, the concepts of concrete objects comprise the knowledge of the “internal” inherent characteristics of the object itself (properties, structure, material). Abstract concepts, according to L. W. Barsalou [19], refer to entities that are neither purely physical nor spatially limited.

L. O. Cherneyko [20] points out that different mental actions are necessary for a word to highlight a fragment of reality (and to realize the content of the word itself); in one case, attention (and consciousness) is directed to an external object (if it is a concrete name), in another case - attention is directed to the internal aspects of consciousness.

Conducting an analysis of abstract nouns on the material of the English language E. V. Pupynina [21] concludes that understanding the word depends on other words in a certain semantic space. Therefore, to determine the meaning of a single word it is important to understand its position in a particular semantic field. According to E. V. Pupynina abstract words verbalize the conceptual content associated with the spiritual life of a person, with life in society. In this regard, such words indicate the properties and phenomena for which one or another assessment is fixed in society.

In today's dynamically changing world, dictionaries and encyclopedias occupy an important place among books and electronic products. On the one hand, knowledge of the real world in lexicographic (vocabulary) form becomes the most popular for transferring information from generation to generation, on the other hand, the dictionary represents the most concise form, depth and breadth of the description of knowledge. Dictionaries are most in demand in social and professional groups of people for self-expression, communication and training [22].

The concepts and terms, with which we operate, as noted in psycholinguistics, are organized in our minds on the basis of the thesaurus. A thesaurus is a special kind of vocabulary of general or special vocabulary, in which semantic relationships (synonyms, antonyms, paronyms, hyponyms, hyperonyms, etc.) between lexical units are indicated. Thanks to the thesaurus, it is possible to reveal meaning not so much by definition, but by correlating words with other concepts and their groups.

The first thesaurus was created by British lexicographer Peter Mark Roger [23] and received the name "Thesaurus of English Words and Phrases", published in 1852. In thesaurus dictionaries, vocabulary is organized according to the thematic principle. P. Roger's thesaurus includes six main semantic categories: 1) abstract relations; 2) space; 3) matter; 4) reason; 5) will; 6) sensual and moral strength.

Studying the history of the thesaurus of P. M. Roger W. Hullen draws attention to the fact that most of the vocabulary is abstract. It is used to describe the names of categories: classes (class), subclasses (division), groups within subclasses, headwords. W. Hullen also notes that the enumeration of synonyms in the first paragraph goes from more abstract to more concrete words, and in the subsequent paragraphs, as a rule, mostly concrete vocabulary is given [24].

That is why, the creation of a dictionary of concreteness / abstract words is quite relevant in connection with the complex worldview of people and various information processing systems.

Thus, the study of abstraction / concreteness doesn’t lose their relevance, while preserving its interdisciplinary status. The obtained results enable to have a different look at intercultural communication, and hence the specific structure of the educational process.

III. MATERIALS AND METHODS

This study is one of the successive steps in collecting information on creating a special dictionary in which the degree of its abstractness will be indicated for each word. This study was carried out in the framework of the joint project of Kazan Federal University (KFU) and the Belarusian state pedagogical university named after Maxim Tank (BSPU). In fact, it continues the experiment of a group of scientists of Kazan Federal University [25] at the Belarusian Russian-speaking sample. The purpose of the study is comparing and matching C / A ratings of nouns in the Russian language.

For carrying out the survey software toolsets were used in which online respondents were required to rate on a scale from 1 to 5 the degree of abstractness / concreteness of each word from the list. The 1st position (on the left) means " the highest degree of concreteness", the 2nd – "a high degree of concreteness", the 3d – "the presence of equal levels of concreteness and abstraction", the 4th – "a high degree of
abstraction" and the 5th means "the highest degree of abstraction". From 40 to 60 student responses were received for each word, assuming the statistical validity of the results. In the process of the study 10 profiles of 50 words in each questionnaire were processed. As a result of a student survey 500 words were analyzed.

The words are taken from the frequency dictionary of the Russian language [22]. In accordance with the traditions of Russian linguistics the concept of concreteness / abstractness is considered only in relation to the noun. From the dictionary [22] we highlighted the nouns, and then for the experiments conducted in our university (BSPU), we selected the nouns with numbers from 501 to 1000. The first 500 nouns were used in the experiments in the KFU. Full data are given on the website of the project (https://kpfu.ru/tehnologiya-sozdaniya-semanticheskix-elektronnyh.html).

The study was conducted on the basis of BSPU (autumn, 2019). The respondents were 2-4 year full-time students (N = 154), students at the faculty of social-pedagogical technologies who are native speakers of the Russian language. The age of the subjects ranged from 18 to 25 years.

The following instruction was presented to the students: it is necessary to choose where the position of the word is, in the interval between the two poles, concrete (1) and abstract (5) Fig. 1. Upon presentation of the instruction, students were not given any hints, directions, only an example was given: “table” is a concrete concept, and “trust” is abstract. There were no time restrictions during the survey.

Then for each word the average index of abstraction was calculated, according to the results of which it could be stated that the higher the word index is, the more abstract it is.

Of the two experiments for English in the first one [26] used a 7-point scale was used, where 7 is the concreteness word, and 1 is the most abstract, and in the second one also a 5-point scale. In this regard, for comparison of ratings of Russian and English words two scales had to be combined. The Russian scale had to be turned on the formula Si = 6 - x, where x is the initial value in the Russian questionnaire. Then the scale was being stretched according to the formula: 100 * (1.5 * (Si - 1)) +1).

Thus identical scales were obtained to assess the degree of concreteness / abstractness: from 700 to 100 with the highest C / A, equal to 700, and the lowest C / A – 100. The most abstract words are located at the border of about 100 points and the most concreteness words are closer to 700 points.

IV. RESULTS

When transferring data in standard (adopted for English, see above) system the following figures (Fig. 2, 3) of the rating of Russian words were obtained. According to the results of the conducted research the following data were obtained. The most abstract were the following words: memory (Russian: 3,83 / English: 275), opening (3,83 / 275), efficiency (3,83 / 275), creativity (3,85 / 272), demand (3,85 / 272), study (3,88 / 269), basic (3,9 / 264), dream (3,91 / 263), beauty (4,00 / 247), adoption (4,02 / 247) – Fig. 2.

The results of our study are consistent with the assessment of specificity in the Russian semantic dictionary by N.Yu. Shvedova [27]. We compared the obtained abstract list of words with the data from the dictionary. The following words confirmed the status of abstract words: beauty, dream, creativity, need, discovery, memory. However, such words with abstract meaning as adoption, main, study, efficiency are absent in the dictionary by N.Yu. Shvedova. The results of our experiment can be used to replenish the semantic dictionary by N.Yu. Shvedova.

The most concreteness of the cases were: hospital (Russian: 1,39 / English: 642), vodka (1,41 / 639), dress (1,41 / 638), bed (1,42 / 638), photography (1,43 / 636), travel (1,45 / 633), bird (1,45 / 633), glass (1,46 / 631), neighbor (1,47 / 630), island (1,47 / 630) – Fig. 3.

With the most specific words obtained a comparative analysis was also performed with words from the Russian semantic dictionary [27]. So, the specific words obtained in our study confirmed their meaning in the specificity section of the dictionary: hospital, vodka, dress, bed, photograph, travel, bird, island.

The calculation of the arithmetic mean, standard deviation and median was done. The arithmetic mean for the sample was equal to 2.59; standard deviation was equal to 0.67; the index of the midpoint - 2.51, which enables to note the downward trend in results.

---

Fig. 1. A fragment of the questionnaire

Fig. 2. C / A rating indicators of the most abstract words of the Russian language
On the basis of the rule of two sigmas the concreteness words include those which have an index of 1.92 (104 words), the abstract words include the ones with an index greater than or equal to about 3.26 (109 words). Note that most of the words are in the interval of 1.92 and 3.26, which enabled to assign them to the category of poorly differentiated (287 words).

The analysis of Figures 2, 3 show that the C / A rating ranges from 247 for the word acceptance to 642 for the word hospital.

To analyze the results we also used a multidimensional statistical method - cluster analysis. The results of multidimensional scaling made it possible to note that there are certain latent signs by which subjects combined these words into certain groups, but due to the large number of words it is not possible to identify these signs Fig. 4.

The results obtained allow us to note the following points. Most of the words in Russian and American have the same situation with the categorization of concreteness. However, in the category abstractness the following words have obtained great differences: beauty (Russian 247, 513 English sample) and dream (263 and 485, respectively), efficiency (275 and 139), opening (275 and 391). You can say that for the Russian sample the concepts such as beauty and dream have a more abstract image for real people.

Let’s consider the most concreteness Russian words with English equivalents Fig. 6.

For most native English respondents beauty / dream is perceived as a more concrete concept. This can be explained as follows. The Russian word mechta is translated as dream. However, the dream in English also has the meaning of a sleep, which is absent in the Russian word mechta. The meaning of the word sleep is much more concrete, which is probably reflected in the evaluation of this word in the English experiment. Also the English word beauty, in addition to beauty, means a beautiful female, which is more concreteness.

In rating of concreteness the greatest differences were given to the following words: neighbor (630 Russian, 548 English samples) and island (630 and 520, respectively).
The most recent discussion of linguistic aspects of concreteness / abstractness can be found in [32]. Establishing precise criteria and boundaries of concreteness / abstractness is extremely difficult. To a greater extent, words are best characterized as more or less abstract / concrete in a number of other words. In different languages there are certain nuances in the definition of concreteness / abstract concepts. Perhaps, for a more detailed study of concreteness / abstractness it is necessary to increase the number of words considered and expand the age range of subjects.

An important issue not addressed in this article is the polysemy of words and the ambiguity in translation.

VI. CONCLUSION

Thus, awareness of the meaning of the word, perception and understanding of texts has become an important area of research in various fields of science. There are questions about how to convey information to the teacher, how to promote advertising products, creating readable texts, how to build communication in business, family, so that the terms you use are understandable to the other side and understood with the accuracy of your message. These and many other issues still require more detailed consideration. Our research is the beginning of the journey. It does not pretend to be a global solution to the problem because of the uniformity of the sample: the student audience. However, in conjunction with other studies of this problem, it will contribute to the accumulation of data for the development of dictionaries (indicating the concreteness / abstractness of words), and is also a definite contribution to the study of natural language and conversational artificial intelligence.

In a joint study conducted at the Belarusian State Pedagogical University and Kazan Federal University a dictionary of the Russian language was first created with the rating of abstract / concreteness words. This article is the first to compare the ratings of abstractness / concreteness of words in Russian and English.

The results of the research confirm that in most cases the representation of abstractness / concreteness of words have an identical performance both in the Russian and English sample.
The ascertained differences are few and can be explained by the specificity of translation and perception of certain concepts. Nevertheless, one should pay attention to the noticeable difference in the evaluation of a number of abstract words with a very exact coincidence of the ratings of concrete words. Further research is required in this direction. Words can be classified in terms of abstraction / concreteness / weak differentiation that allow you to use them in the communication process, the educational process and promotional activities.

Further research should provide estimates of more words and reach a wider audience of respondents, not just students, as in the current study.

ACKNOWLEDGMENT

This research was financed by Grant 19-07-00807 of Russian Foundation for Basic Research.

REFERENCES


