UDC 81'33
DOI: 10.17223/19986645/72/5

Istvan Lenart[1], Irina Yu. Markovina[1], Orsolya Endrody[2]

[1] I.M. Sechenov First Moscow State Medical University (Moscow, Russian Federation)
[2] Eötvös Loránd University (Budapest, Hungary)
E-mail: istvan.lenart@1msmu.ru, markovina_i_yu@staff.sechenov.mail.ru, endrody.orsolya@ppk.elte.hu

Preschool Children's Verbal Image of the World: A Cross-Cultural Russian-Hungarian Comparison Based on Word Associations[1]

This psycholinguistic-pedagogical interdisciplinary research investigates 4- to 5-year-old Russian and Hungarian preschool children's linguistic image of the world. Applying the word association method, kindergarteners (N=100 in both countries) were asked to freely associate from 10 word-stimuli, then the results were contrasted. The research shed light on a similar perception of a family; of a friend; on the effects of globalization (Lego, Trudi, tablet); and on lacunas (devil and angel).

Keywords: *association experiment, linguistic consciousness, early childhood, psycholinguistics, cross-cultural communication*

1. Introduction

This study investigated preschool children's verbal image of the world from a Russian-Hungarian cross-cultural perspective. The approach merged perspectives of psycholinguistics, linguistics, and pedagogy. The methodology is based on the psycholinguistic theory of linguistic consciousness deeply rooted in Vygotsky's cultural-historical psychology [1, 2], A.N. Leontiev's psychological theory of activity, and the theory of speech activity by A.A. Leontiev [3–6]. The key research method to collect the linguistic data was an association test also referred to as an association experiment [7, 8]. The authors modified the method to better adjust it to the target group of respondents. The shoulder-to-shoulder method [9] was used while collecting the data, and the results were analysed using the psycholinguistic theory of linguistic consciousness and methods of corpus linguistics [10].

The major contribution of the pedagogical perspective was to characterise the age group chosen and select the stimulus words for the association test. When compiling the list of word-stimuli, the authors relied on describing and understanding childhood from the children's perspective. Early childhood experts emphasise that children must be regarded as 'actors' determining their own life [11]. In order to correctly analyse the data obtained, it was important to remember that

[1] The reported study was funded by the Russian Foundation for Basic Research, project number 18-512-23004.

learning begins with birth, and 4- to 5-year-old children have a vast knowledge of the world around and of their own selves.

The research objective was to reveal and analyse similarities and differences in 4- to 5-year-old children's linguistic consciousness across cultures, in the Russian-Hungarian context. Apart from the cross-cultural approach, the research is a part of a longer series of studies in which the 10- to 12-year-old age group will be investigated with the same methodology, allowing the researchers to diachronically contrast the 4- to 5-year-old age group, with the 10- to 12-year-old age group and the adult language users; the latter by using existing associative dictionaries.

In the course of the research, the authors aimed at gaining an insight into the *linguistic consciousness* (the psycholinguistic equivalent for *image of the world*) of Russian and Hungarian preschool children. The aim of cross-cultural research typically is to promote better understanding between potential dialogue partners by demonstrating common and culture-specific characteristics of the image of the world of people speaking different languages. In order to achieve this, the authors started investigating 4- to 5-year-old children as respondents. Thus, two possible Russian and Hungarian childhood narratives were compared based on the fact that childhood as an abstract concept varies within different societies and cultural groups [12, 13, 14].

In order to map and compare the linguistic consciousness of Russian and Hungarian preschool children, the association experiment was selected as the most appropriate and easy-to-use research method. Each participating respondent group included 100 children aged 4 to 5. Ten concepts were investigated: *friend, child, family, water, black, toy/game, devil, home/house, foreigner,* and *angel.* Two lists of stimulus words in both languages were compiled based on the concepts selected. In this article, key findings of the research are presented, and the selection method is discussed in detail in Section 3.

2. Psycholinguistic approach to linguistic consciousness

2.1.1. Definitions of linguistic consciousness

As already mentioned, the present research is based on the psycholinguistic theory of linguistic consciousness. Even though the phenomenon of linguistic consciousness has been widely discussed and investigated by Russian psychologists, linguists and psycholinguists [1–7, 15–18], there is still no agreed definition.

There have been a number of approaches. For instance, Shcherba applied the term to refer to *individual psychophysiological verbal organization* [15], where linguistic consciousness referred to an individual language system as opposed to the language of the community.

Ushakova emphasised the difficulties in interpreting the term *linguistic consciousness*, focusing on the two underlying phenomena: the mental and the material. The mental phenomenon is of a non-material nature, it cannot be measured, seen, or heard; whereas the material phenomenon implies speech that can be produced or

recorded and the physiological process of building verbal links [16]. Even though the author admits the polysemy of the term, she believes that the key idea is that it refers to consciousness expressed verbally. In other words, the term illustrates the transition from a mental phenomenon into a material one [16].

Ufimtseva stresses that the concept of *linguistic consciousness* used for studying or modelling the linguistic picture of the world is synonymous with the psychological concept of "the image of the world" [6].

Based on Alexey N. Leontiev's psychological theory of activity, Tarasov describes linguistic consciousness as images of mental consciousness that constitute a person's perceptual and conceptual knowledge about the objects of the world. These images of consciousness require externalizations that can be observed, which may be represented as objects, actions, or words [5]. Thus, linguistic consciousness is viewed as a set of linguistically externalised mental images of consciousness developed by members of a certain culture. This contains concepts of man and their activities as well as concepts of objects and phenomena [5]. Tarasov also stresses that an important function of externalization is to communicate mental images across generations, for example, a mental image of the Russian *house/home* can be developed if one perceives it from its both inside and outside and lives in a Russian family setting. This shows that externalization reveals cultural and mental peculiarities that can be studied [5]. Ufimtseva states that people communicate by using special signs (mostly linguistic) drawing on the knowledge accumulated within their native culture [19].

Leading Russian psycholinguists agree on the following characteristics of *linguistic consciousness*:

1. It manifests images of consciousness indirectly and materially, and can therefore be studied;

2. It can be described as mental images having linguistic equivalents, which are characteristic of a specific community and can be extrapolated;

3. It reflects cultural and mental peculiarities and moral standards of a community. This makes linguistic consciousness a tool not only for linguistic and psychological analysis but also for the analysis of cultural perspectives of a community.

The reason for applying an approach that is relatively unknown in the academic community outside the Russian Federation lies in the fact that the methodology of Russian psycholinguistics enables researchers to easily and effectively investigate the individual language user's mental inventory and organization of lexemes (mental lexicon). Furthermore, associations as a research technique have been used since the late 19th century when Galton's first association experiments [20] shaped the psycholinguistic research of the 20th century [21, 22]. In recent years, the technique has experienced a new renaissance due to new informatic analytical tools able to collect word association data [23, 24] that are able to shed new insights into the structure of the mental lexicon based on word association data [23, 25].

In Hungary, one of the most notable applications of the association technique was the Agykapocs (Connectyourmind) experiment conducted to demonstrate how

the networked structure of the lexicon can be captured and how these structures can contribute to a better understanding of the mental lexicon. The association data were collected online with different methods of network science being applied to analyse the data. It was shown that words that have a globally central position (can be seen as central elements of the networks) are at the same time partly and locally central (they have a central role in local, small structures as well) [26]. Furthermore, it was revealed that the central elements were in parallel in the analysed Hungarian network and in the Florida Word Association Database [27] (see more about the advantages of the association method in Section 2.2.).

2.1.2. How linguistic consciousness can be studied

The notion of *linguistic consciousness* is a result of the cooperation between linguists and psychologists, and being at the interface between these two sciences, it enables researchers to study speech, language, mental consciousness, and culture. As Ushakova states, developing the term in a specific study is supposed to provide an opportunity to enrich our knowledge not only of the speech and language phenomena, but also of consciousness as a mental phenomenon [16].

At the same time, the investigation of linguistic consciousness is associated with a number of problems, one of which is the inability to study consciousness directly and objectively. Tarasov states that "*the distorting character of images of consciousness in their externalization*" is a challenge for researchers [5].

To meet this challenge, an association test involving a large number of subjects can be conducted as it can been seen as a way of averaging. This implies that statistically processed association test data show the various fragments of linguistic consciousness of a typical language speaker. Conducting an association experiment results in constructing associative fields of stimulus words. Ufimtseva asserts that an associative field is not only a fragment of verbal memory but also a fragment of the world image of a particular language speaker reflected in the consciousness of an average cultural representative, his or her judgments and motivations, i.e. cultural stereotypes [28].

Distortion of the images of consciousness can be caused by mental mechanisms that externalise images in words as well as by many other factors that influence the validity of the association experiment. For instance, the limitation may be that some respondents tend to react in a biased way, i.e. in the form of a favourable approach to a certain topic or word-stimulus.

The absence of an agreed definition of the term *linguistic consciousness* gives rise to one more challenge, i.e. the conclusions drawn ultimately depend on the individual researcher's interpretation [16]. Linguistic consciousness lies at the interface between both mental and material phenomena. A researcher may tend to present a transition from one phenomenon to another as a simple and direct one [16]. The transition is possible dependent on whether we know how the image of consciousness is transformed into an externalised sign or object (in our case, a word). Without understanding this process, we cannot scientifically explain how mental and material phenomena are related to each other.

However, with the advent of the so-called *association dictionaries* and a*ssociative thesauri* [29, 30] the focus of interest moved to the problem of interpreting the results of association tests.

2.1.3. The role of linguistic consciousness in communication

As mentioned above consciousness is a mental phenomenon that cannot be studied directly. To gain access to it, it has to be externalised as linguistic signs that people select based on their knowledge acquired in their cultural environment. It could be suggested that we may succeed in communicating our consciousness images to other people via linguistic signs if we share with them common knowledge as well as cultural stereotypes. Within a community, mutual understanding and interaction are possible because communicators use definitions and interpretations shared by all members of this community [18].

This explains why representatives of one culture freely understand each other and why there are failures in intercultural communication (even if there is no language barrier). The latter is determined by the fact that people of different cultural backgrounds lack commonality of consciousness [5].

It can be concluded that in order to improve intercultural communication it is important to reveal features of linguistic consciousness in people from different cultural backgrounds. Intercultural communication has a pivotal role in the modern world of globalization and growing international cooperation, consequently, studying problems of intercultural miscommunication is essential. Studying common and distinct features of linguistic consciousness in people of different cultural backgrounds will contribute to problem solving.

2.1.4. Techniques to study the content of consciousness images

Ufimtseva points out that peculiarities of images of consciousness (cultural stereotypes) can be revealed as a result of either conscious introspection or organised experimental research [31]. Association test technique and associative dictionaries and thesauri developed based on the data obtained are now widely used to gain access to the content of mental images and, through this, reconstruct the verbal image of the world of particular language speakers. The benefit of this approach is that an association dictionary entry may be regarded as a model of linguistic consciousness of a language speaker representing a certain culture. It demonstrates the world image, features of culture-specific mindset, ethnic character, and communicative potential of native language speakers [31].

2.2. Association experiment as a research tool to reconstruct the content of the consciousness image externalised by a word

2.2.1. Association experiment from the psycholinguistic perspective

As Ufimtseva states, linguistic consciousness can be studied only as a result of the former activity in its "converted" form and alienated from an individual's

consciousness forms [28]. In other words, to be able to analyse it, we need an externalised, materialised form of linguistic consciousness.

There are several types of an association experiment: free, directed, and linking associative experiments. The technique most commonly used to externalise linguistic consciousness is the free association experiment, wherein an individual is given a list of words and is instructed to respond with the first word that comes to mind. In most cases, respondents are university students. The test is commonly carried out in the native language of the subject.

A major advantage of an association experiment is its relative simplicity in conducting: no special equipment or setting is needed, and the instructions are easy to understand and follow. Another benefit of this approach is that it saves time. The test may be conducted in a room with a large number of respondents, which saves time in data collection. Another advantage is that a large number of responses may be collected simultaneously, enabling the extrapolation of results to all representatives of a particular culture. The large number of respondents secures a relative stability of association fields. Ufimtseva [7] summarises the advantages of applying the association experiment as follows: 1) it reflects the experiences of native speakers of a certain language; 2) it displays the relative importance of lexemes, thus refers to their hierarchy in language (with this method it is also possible to identify the core elements of the linguistic lexicon); 3) it can be applied to any language; and 4) it is not artificially constructed by a linguist (as in the majority of traditional dictionaries).

In the context of the current research, the authors underline the additional benefits of the application of the association experiment including: 5) the experiment can be conducted as a kind of game, which can effectively be applied with young respondents; 6) the natural situation created during the survey enables respondents to provide spontaneous and less artificial responses; 7) the experiment is easily administered with no special technical equipment needed; 8) the experimental process can easily be explained to different age groups, including children.

However, there are certain limitations of this experimental technique. Many factors may influence the results of the test. One of these is that there are some factors independent of the experiment itself (neither the researcher nor the respondent can influence it). Among the factors is the active vocabulary of the respondent, age, sex, profession, and geographical conditions. A high frequency of the response *man* to the stimulus word *woman* is explained by a large number of female respondents [5]. In the experiment performed by Alexey A. Leontev, respondents reacted differently to the stimulus word *brush* depending on their profession and geographical conditions. Research by Russian psychologists and linguists showed that respondents with technical background give more paradigmatic responses whereas subjects studying the humanities tend to give more syntagmatic reactions [32, 33].

Factors such as time and conditions of a test as well as a respondent's physical and emotional state may have an impact during the experiment. Thus, results may be distorted because of the respondent's fatigue caused by the length of the

experiment (e.g. phonetic associations that are not typical of adult respondents *врач — грач* [*doctor — rook*]). The place where the experiment is conducted can also have an effect on the results (e.g. *institute — here*). All these limitations should be well understood and noted in each association research.

In the Hungarian literature, application of the association experiment as a linguistic research method is less widespread than in Russia. Several authors attempted though to map Hungarian language users' mental lexicon by applying the association experiment, including the compilation of the *Encyclopedia of Hungarian Norms of Associations* [34]. Kovacs strongly relies on the association method when he investigates concept systems and lexical networks in the Hungarian mental lexicon [29]. He discusses the role of associative networks both on the level of language users in general and on the language acquisition process in early childhood.

2.2.2. Association experiment from the pedagogical perspective

Since the research is related to the Early Childhood Development and Conception of Childhood theory, it is important to mention the basics of the attitude toward such young learners. The authors believe that all knowledge is based on human interactions and is constructed based on social context [35]. Childhood conceptions are based on the idea that children must be seen as social actors [11] and be capable of determining their own lives [36]. According to the related research, 4- to 5-year-old children look for differences, patterns and change, and are able to ask questions about why things happen and how things work. They also show interest in and curiosity about understanding social structures such as family or institutional groups.

2.2.3. Association field

The association field is a set of reactions to a stimulus word. An association field consists of a nucleus with the most frequent reactions and a periphery. In terms of its content, it reflects both a person's verbal memory and a fragment of the verbal image of the world, judgments, attitudes, and motivations of the respondents as representatives of a certain culture (see: *Russian Association Dictionary 2002*). Furthermore, it needs to be emphasised again that, although there are some individual variations in responses, the fact that the respondents represent one culture secures association fields with relative stability.

The so-called *semantic gestalt of an association field* developed by Karaulov is one of the methods to reconstruct the knowledge about the surrounding world in the linguistic consciousness of native speakers [17]. The method is based on dividing the reactions of one association field into several semantic zones by uniting similar features of an object or a phenomenon. To make it convenient, the semantic zones are marked by pronouns that reflect general ideas, e.g. who (persons), what (objects), which (attributes), this (structures in which the pronoun 'this' acts as a hypothetical link), to do/to make (actions), where (places), and

when (time). At the same time, it is possible to identify additional semantic zones [17].

2.2.4. Associative dictionaries

Conducting large-scale association experiments resulted in creating the so-called association dictionaries, or thesauri of word associations. Today, there are two major Russian dictionaries: *Russian Association Dictionary* (*RAD*) developed by Juri Karaulov, Galina Cherkasova, Natalia Ufimtseva, Yuri Sorokin, and Evgeny Tarasov, and the *Slavic Association Dictionary* (*SAD*) developed by Natalia Ufimtseva, Juri Karaulov, Galina Cherkasova, and Evgeny Tarasov. The RAD contains around 1300 stimulus words and about 13,000 different reactions. The fullest English association dictionary is *The Associative Thesaurus of English, The Edinburgh Associative Thesaurus* developed by G.R. Kiss, C. Armstrong, and J. Piper. *The English Associative Thesaurus* contains more than 23,000 words.

The reactions in a dictionary entry are given in a decreasing order of frequency. Each reaction has its own index, the number of respondents who reacted to a stimulus word with this word.

An association dictionary is truly a unique reference source because it gives information about the most frequent word links that characterise cultural features, and no other dictionary contains such information. Thus, an association dictionary holds a large amount of data that enables researchers to study culture, linguistic consciousness, and 'text potential', i.e. the cultural and linguistic background of the respondents for text perception and understanding. Association dictionaries allow the researcher to identify and study the systemic character of the world image of representatives of different cultures. To do this, researchers reveal the nucleus of language consciousness, i.e. the units of a semantic network that have the largest number of links with other units of this semantic network represented in the thesaurus (dictionary) [37].

Thus, researchers in different fields will find their own field-specific data in an association dictionary. A philosopher will reveal the constituent parts of the image of the world; a culture studies expert will discover the system of axiological patterns and attitudes of a certain culture; a psychologist will see the proportion of linguistic and extra-linguistic knowledge contained in the image of the world of representatives of a certain culture [37].

The present cross-cultural research has been conducted in line with the theoretical and experimental approaches described. The authors attempted to make a step forward by developing and modifying the techniques already existing in this research field. We hope some innovative approaches presented and results obtained will form a convincing argument in favour of an interdisciplinary study conducted by an intercultural team of researchers.

3. Methods

The underlying methodological background of the research is the association experiment [6], which is based on the free associations of respondents to a given

stimulus-word. Researchers apply this method for conducting smaller-scale research on well-defined research questions, as well as for compiling large- scale associative dictionaries including the *Russian Associative Dictionary* [37]; the *Slavic Associative Dictionary* of the Russian, Belarusian, Bulgarian and Ukrainian languages [38]; and the *Russian Regional Associative Dictionary "EURAS"* [39] to mention a few.

In this research, 10 stimulus words *(friend, child, family, water, black, toy/game, devil, home, foreigner, angel)* were selected from the following sources: five stimuli from the core lexicon of Russian linguistic consciousness [40] with their ranks in parentheses *дом* (*home,* #2), *друг* (*friend*, #10), *вода* (*water*, #18), *ребенок* (*child,* #19,5), *черный* (*black*, #49,5); one stimulus word – *игрушка* (*toy*) – from the 18 initial stimuli of the *Russian Children's Associative Thesaurus* [41]. The authors' essential endeavour when selecting the above mentioned six stimulus words was to guarantee that 1) the selected words were within the core linguistic lexicon (most frequently used words of the language); 2) the words were semantically closely connected to children's everyday life; and 3) the words were simple, easy-to-understand lexemes, appropriate for conducting research with 4- to 5-year-old respondents.

Finally, four stimuli were selected as the authors' subjective choice in order to, on the one hand, better map the children's self-perception and their concept of a family (*семья*/*family*), and, on the other hand, add some atypical lexemes that were less frequently used by this age group, including *angel* (*ангел*), *foreigner* (*иностранец*), and *devil* (*чёрт*). In line with ethical standards of academic research, informed consent was obtained.

The selection of the ten stimulus words and the more detailed analysis of the five items most closely connected with childhood *(friend, child, family, toy/game, home/house)* as well as their investigation with a combination of the association experiment, corpus linguistic methods and a pedagogical approach enabled the authors to map and compare Hungarian and Russian preschoolers' perception of the outside world.

Regarding the pedagogical perspective of the dichotomy of *angel* and *devil*, it must be stated that, in the studies of history of childhood, one of the thought-provoking questions is about the polarised approaches whether children are divine or of diabolical nature [14]. According to the Christian doctrine of the original sin, based on St. Augustine's theories, it is necessary to be baptised in order to wash away the diabolical taint. On the other hand, Luther believes that human beings are innocent in the first 5 to 6 years of their lives.

The 10 selected Russian stimulus words were translated into Hungarian, with the following remarks on the translation problems of equivalence. The stimulus noun *ребенок (child)* can be translated into both *gyermek* (more formal) and *gyerek* (less formal), and, in line with the respondents' age group, the latter, less formal word form was prioritised. The Hungarian word *játék* has a wider denotation in the Hungarian language when compared to Russian and means both *игрушка* (*toy*) and *игра* (*game*). *Дом* denotes both a *house* (*ház*) and a *home* (*otthon*) in the Russian language; in line with the focus of the research *otthon*

(home) was selected as its Hungarian equivalent and was included in the list of initial stimuli. A peculiar element of the Russian lexicon *чёрт* was singled out. This noun is one of the possible translations of the English word *devil*, besides *дьявол*, *бес*, and *сатана:* the lexeme in addition means in colloquial language usage: *heck*.

Two series of kindergarten surveys were accomplished with 100–100 respondents: one in Russia (in two kindergartens in Moscow), and the second one in Hungary (97 interviews in urban areas, in 7 cities including Budapest, Szeged, Tata, Zalaegerszeg, Gödöllő, Kistarcsa, and Veresegyház as well as three respondents from a village, Kesztölc). Participants of the research were 4- to 5-year-old preschool children born in 2013–2014. During a 15–20-minute interview, the children were requested to say out aloud whatever came to their mind when hearing the stimulus words. The questionnaire consisted of 27 questions on the 10 word-stimuli (2 to 4 sub-questions for each stimulus word). In contrast with a typical association experiment, several questions were applied for each stimulus word in order to gain a more significant amount of linguistic information that could be transformed into a relatively large corpus. This enabled the authors to analyse the results not only through the classical frequency lists of the association experiment, but also with corpus linguistic tools.

The interviews were conducted with the application of the shoulder-to-shoulder method [9], an approach that originates from pedagogy [42], which presumes that children from this age group are typically reluctant to speak openly and express their feelings in a classical interview situation. The shoulder-to-shoulder approach goes back to paired or partner reading; a teaching strategy that enhances reading fluency by two students sitting next (shoulder to shoulder) to each other, sharing a book. The method itself enhances children's elaboration of a certain theme, which is primarily important in this age group in order to reduce their inner tensions and stress created during everyday life activities. In this research, the method was applied so that interviewers were asked to create an informal situation, allowing the children to free-play, to walk, draw and move during the interview. This contributed to a better atmosphere, where children were more willing to communicate. Moreover, if the 15–20-minute conversation could not be completed at once, then the data collectors stopped and continued later when the child felt ready for the conversation. According to Ginnis, everyone needs to feel emotionally secure and psychologically safe [43].

Answers to all 27 questions were registered either in written form or by voice recorder, and later transcribed to text. All replies were entered into Excel files, and then the authors created 22 separate searchable corpora: 10–10 corpora were based on the 10 stimulus words from the Hungarian and the Russian results, another two corpora were created containing all responses from the Hungarian and from the Russian respondents, respectively. The Hungarian corpus consisted of 19,967 tokens and 15,319 words, while the Russian corpus incorporated 16,268 tokens and 12,646 words.

The linguistic data were analysed with the Sketch Engine corpus linguistic tool [10] firstly on the separate (stimulus-based), then on the cumulated (coun-

try/language-based) level. Frequency results were compiled and contrasted with two comparable gigaword corpora from the Araneum corpus family [44]: the Araneum Russicum Russicum Maius corpus (1,200,000,258 tokens, 859,319,823 words) and the Araneum Hungaricum Maius corpus (1,200,001,609 tokens, 792,549,686 words), while keyword analysis was executed with Russian Web 2011 (ruTenTen11; 18,280,486,876 tokens, 14,553,856,113 words) and Hungarian Web 2012 (huTenTen12; 3,161,920,362 tokens, 2,572,620,694 words) [10].

The results of the association experiment which served as the core method of this research were compared with the huge stock of linguistic data of the Araneum corpora, utilising an in-build function of the Sketch Engine tool; the thesauri of the investigated stimulus words [10]. Thesauri, as defined by Sketch Engine, are semantically similar or synonymous items of a corpus that can be identified by analysing the typical collocations of the given stimulus words.

4. Results

4.1. Individual stimulus words

Both the Hungarian and the Russian responses (up to 10 respective stimulus words each) were collected and analysed individually.

In this section, we discuss in detail five of the ten stimulus words; each of them being easy to understand, and the most closely connected to early childhood: *friend, child, family, toy/game,* and *home/house*. The results are grouped in accordance with the association method; according to the frequency of each response word, a list is created beginning with the most frequent word. The top four responses are displayed in the tables below (Tables 1–5). The responses are collected and displayed in three groups depending on their word classes: nouns, adjectives and verbs were taken into account (other word classes were excluded).

4.1.1. Friend

Друг (*friend*) is the 10th most typical item in the Russian linguistic consciousness nucleus [40]. *Barát*, a noun that, although obviously not semantically identical to its Russian counterpart, is without doubt the Hungarian equivalent. However, there were no translation problems.

The results showed a similarity (Table 1), with the top responses nearly coinciding in the category of adjectives (*kind/dear*, *good* and *cute* among the Hungarian and *good*, *kind* and *nice* among the Russian responses). The top two verbs associated with the concept of *friend* also proved to be identical (*play* and *love*), while the third and fourth most frequent words differed from each other although referring to a common activity with someone. The most remarkable difference could be observed between the noun responses, where the Russian results reflected a personality focus (*boy*, *girl*, *person*, *female friend*). Meanwhile, in the Hungarian results, the inanimate entities objects prevailed (*toy*, *nursery*, *court*).

Table 1

Associations of the stimulus words *BARÁT (FRIEND)* and *ДРУГ (FRIEND)*, results of the research

barát *(friend)*			***друг*** *(friend)*		
NOUN	ADJ	VERB	NOUN	ADJ	VERB
játék ***(game / toy) 14***; *anyuka* ***(mother) 13***; *ovi* ***(nursery) 10***; *udvar* ***(court) 6***	*kedves* ***(kind/dear) 44***; *jó* ***(good) 33***; *aranyos* ***(cute) 5***	*játszik* ***(play) 130***; *szeret* ***(like/love) 42***; *épít* ***(build) 7***	*мальчик* ***(boy) 14***; *девочка* ***(girl) 12***; *человек / машинка* ***(person/car) 10***; *подруга* ***(female (friend) 6***	*хороший* ***(good) 35***; *добрый* ***(kind) 16***; *красивый* ***(nice) 7***	*играть* ***(play) 85***; *любить / дружить* ***(love / be friends) 42***; *помогать* ***(help) 30***; *гулять* ***(walk) 16***
	ad ***(give) 6***				

Table 2

Thesauri of the stimulus words *BARÁT (FRIEND)* and *ДРУГ (FRIEND)* from the Araneum comparative corpora

barát *(friend)*			***друг*** *(friend)*		
NOUN	ADJ	VERB	NOUN	ADJ	VERB
lány ***(girl) 316,289***; *fiú* ***(boy) 162,514;*** *apa* ***(father) 125,193;*** *család* ***(family) 350,449***	*kedves****(kind/dear) 214,709***	*szeret* ***(like/love) 983,599***; *gondol* ***(think) 568,644***	*человек* ***(person) 2,021,397***; *родитель* ***(parent) 220,563***; *ребенок* ***(child) 510,992***; *жизнь* ***(life) 908,095***	*другой* ***(other) 1,716,427***; *такой* ***(such) 2,544,728***; *каждый* ***(each) 1,108,658***; *самый* ***(most) 1,354,155***	*стать* ***(become) 1,074,175***
	hisz ***(believe) 470,096***; *mond* ***(say) 1,324,910***				

The thesauri results of the words *barát (friend)* and *друг (friend)* automatically identified similar words or synonyms extracted from respective corpora. In this paper, the comparative corpus family Araneum and its Hungarian and Russian gigaword corpora were used to compare the preschool children's associations with adults' language use based on a massive amount of linguistic data (1,2 billion tokens in each of the two corpora).

Similarities between Hungarian adults' and children's language use (Table 2) include the key role of the adjective *kind* and the verb *love* when analysing the stimulus word *friend*, as well as a curious difference in associating *mother* to a friend among the children's results, and *father* to the same stimulus word in the adults' corpus.

A visualization of the aforementioned results is displayed in Figure 1, where the most typical collocations of the word *barát (friend)* and *друг (friend)* are grouped depending on their grammatical relations and their role in the sentence (verbs with; modifiers of; subject of; object of etc.).

Fig. 1. Visualization of the associations of stimulus words *BARÁT (FRIEND)* and *ДРУГ*[1]

In the Russian chart, *настоящий (real)* appears besides *хороший (good)* as opposed to the Hungarian adjective *jó (good)* as the main characteristic of a friend. In the Hungarian linguistic consciousness, the chart suggests that a friend usually *helps*, *plays with the pony*, moreover, friends *stick together*, while in the Russian example, the diagram shows *playing*, *helping* and *being friends*. It must be noted that Russian morphology allows the use of the verb *дружить (be friends)*, originating from the noun *друг (friend*). The reaction word *дружить* is an example from the Russian language of a stimulus-reaction connection partly based on similar sound.

4.1.2. Child

The image of a child also proved to be similar in the Hungarian and Russian linguistic consciousness. All four of the most frequent adjectives (*small*, *big*, *good*, *kind*) coincided in the lists of associations, while two of the four most typical verbs (*grow up*, *love*) were also identical. *Cry* (*плакать)* as a verb connected to a child's image only appeared in the Russian top 4 but not in the Hungarian.

Based on the visualization charts (Figure 2), the *small–big* dichotomy appears amongst the associations of a child in both groups of respondents (*big* being present in the Hungarian and *small* in the Russian sample). Significant differences are visible in the chart in terms of activities associated with a child: in Hungarian, collocations of a child describing their activities include *néz* (*look*), *csinál* (*do*), *lesz* (*will be*), *van* (*is*), meanwhile, the Russian sample displays *воспитывать (raise), обижают* (*hurt*), and *бывать* (*be*).

[1] Labels of the Hungarian chart: barát (friend); jó (good); közös (common); Ágika (Ágika); kedves (kind); anyuka (mother); Nati (Nati); segít (help); van (be); lesz (will be); összetart (stick together); pónizik (plays with the pony. Russian labels: друг (friend); настоящий (real); хороший (good); играть (play); со (with); дружить (be friends); помочь (help); быть (be); значит (mean).

Table 3

Associations of the stimulus words *GYEREK (CHILD)* and *РЕБЁНОК (CHILD)*, results of the research

gyerek *(child)*			***ребёнок*** *(child)*		
NOUN	ADJ	VERB	NOUN	ADJ	VERB
kisbaba ***(small baby) 10***; *baba* ***(baby) 9***; *játék* ***(toy/game) 7***; *óvoda / anyuka / fiú/lány* ***(nursery / mother / boy / girl) 6***	*kicsi* ***(small) 54***; *nagy* ***(big) 22***; *jó* ***(good) 20***; *kedves* ***(kind) 13***	*felnő* ***(grow up) 35***; *játszik* ***(play) 28***; *válaszol* ***(reply) 13***; *szeret* ***(like/love) 11***	*малыш* ***(kiddy) 46***; *человек* ***(person) 18***; *мама* ***(mom) 17***; *мальчик* ***(boy) 8***	*маленький* ***(little) 44***; *хороший* ***(good) 10***; *добрый* ***(kind) 6***; *большой* ***(big) 6***	*расти / плакать* ***(grow up / cry) 6***; *любить / родить* ***(love/be born) 5***

Table 4

Thesauri of the stimulus words *GYEREK (CHILD)* and *РЕБЁНОК (CHILD)* from the Araneum comparative corpora

gyerek *(child)*			***ребёнок*** *(child)*		
NOUN	ADJ	VERB	NOUN	ADJ	VERB
gyermek ***(child) 376,113***; *ember* ***(man/person) 1,248,639***; *szülő* ***(parent) 180,688***; *lány* ***(girl) 316,289***	*fiatal* ***(young) 318,105***; *kicsi* ***(small) 369,164***; *jó* ***(good) 2,041,516***; *kis* ***(little) 644,146***	*szeret* ***(love) 983,599***; *akar* ***(want) 799,720***; *lát* ***(see) 1,050,286***	*человек* ***(person) 2,021,397***; *малыш* ***(kiddy) 136,028***; *женщина* ***(woman) 377,996***; *родитель* ***(parent) 220,563***	*детский* ***(children's) 298,287***; *данный* ***(given) 731,108***; *молодой* ***(young) 223,831***	*стать* ***(become) 1,074,175***

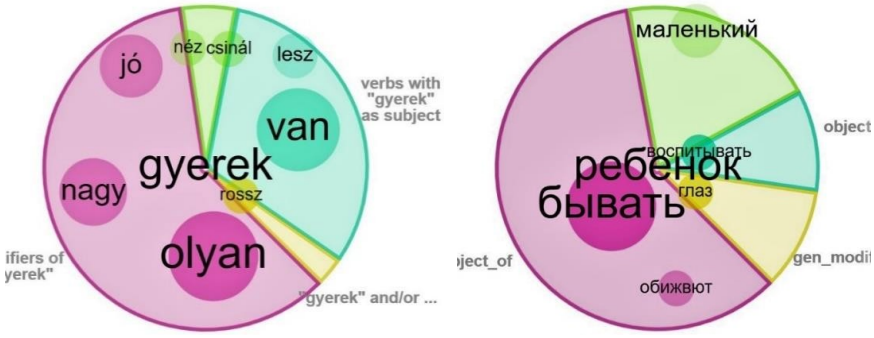

Figure 2. Visualization of the associations of stimulus words *GYEREK (CHILD)* and *РЕБЁНОК (CHILD)*[1]

[1] Labels of the Hungarian chart: gyerek (child); nagy (big); jó (good); néz (look), csinál (do), lesz (will be), van (is), rossz (naughty); olyan (such). Russian labels: ребёнок (child); воспитывать (raise); обижают (hurt); бывать (be) глаз (eye); маленкий (small/little).

Results of the Araneum corpus displayed several corresponding results in the Hungarian language: a child is *small* and *good*, and *love* as a verb plays a central role in the semantics of the word. Some top noun results of the Russian data coincided including the nouns, *малыш (kiddy) and человек (person)*.

4.1.3. Family

The word *family* does not belong to the core lexicon of the Russian linguistic consciousness, nor is it included in the list of 18 initial stimuli of the Russian children's associative thesaurus [41]. Nevertheless, this lexeme was included in this study as the concept of a family constitutes an integral part of a child's linguistic consciousness. Furthermore, it helps in comparing the Hungarian and the Russian concepts of a family.

Table 5

Associations of the stimulus words *CSALÁD (FAMILY)* and *СЕМЬЯ (FAMILY)*, results of the research

család *(family)*			***семья*** *(family)*		
NOUN	ADJ	VERB	NOUN	ADJ	VERB
anya/anyuka ***(mother) 140***; *apa/apuka* ***(father) 121***; *gyerek* ***(child) 35***; *mama* ***(mom/grandma) 23***	*kedves* ***(kind) 27***; *jó* ***(good) 23***; *nagy* ***(big) 17***; *igazi* ***(real) 10***	*szeret* ***(love) 26***; *játszik* ***(play) 12***; *mond* ***(say) 11***; *megy* ***(go) 10***	*мама* ***(mom) 146***; *папа* ***(daddy) 134***; *бабушка* ***(grandma) 50***; *дедушка* ***(grandpa) 40***	*хорошая* ***(good) 40***; *добрая* ***(kind) 20***; *большая* ***(big) 19***; *дружная* ***(friendly) 17***	*любить* ***(love) 26***; *жить* ***(live) 13***; *ходить* ***(go) 7***; *дружить* ***(be friends) 5***

A general feature of the frequency list of both respondent groups reveals that the strongest association from the stimulus word *family* is *mother*, followed by *father*. In the Russian context *grandmother* and *grandfather* follow. The Hungarian children mentioned *child* and *mama* (which may mean either *mother* or *grandmother* depending on the context) after their parents. The comparison of activities connected with the term *family* displays *love* in the first place and *to go* within the top 4 associations (presumably referring to common activities of the family that include travelling or going somewhere together perceived as a common activity). Adjectives used in describing a family almost fully coincide: *kind*, *good*, *big*, and *real* are identified in the Hungarian, while *good*, *kind*, *big*, and *friendly* occur in the Russian group.

Significant differences can be pinpointed in the perception of a family when comparing the children's and adults' results. Children's perspective reflects a self-centred approach when the family is characterised by the *mother*, the *father*, and grandparents, meanwhile key elements of the concept of a family are the *child,* and the verb *live* as the data of the Araneum corpus suggests.

Table 6

Thesauri of the stimulus words *CSALÁD (FAMILY)* and *СЕМЬЯ (FAMILY)* from the Araneum comparative corpora

család *(family)*			***семья*** *(family)*		
NOUN	ADJ	VERB	NOUN	ADJ	VERB
gyermek ***(child) 376,113***; *gyerek* ***(child) 412,046***; *ember* ***(man / person) 1,248,639;*** *élet* ***(life) 901,881***	*fiatal* ***(young) 318,105***	*él* ***(live) 675,558***; *szeret* ***(love) 983,599***	*родитель* ***(parent) 220,563***; *дитя* ***(child) 617,06***; *жизнь* ***(life) 908,095***; *ребёнок* ***(child) 510,992***	*молодой* ***(young) 223,831***; *русский* ***(Russian) 364,077***; *данный* ***(given) 731,108***	*стать* ***(become) 1,074,175***; *хотеть* ***(want) 600,660***; *жить* ***(live) 310,494***; *смочь* ***(be able to) 454,904***

The visualization charts complement the above results with further information, including the fact that in the Hungarian context the closest term to *family* is *real*, while in Russia it is *love* and *nice*. A diminutive form of family appears in the Russian chart: *семьячка*, similarly to other reaction words in the study e.g. *братик* (*bro*), *сестрёнка* (*sis*).

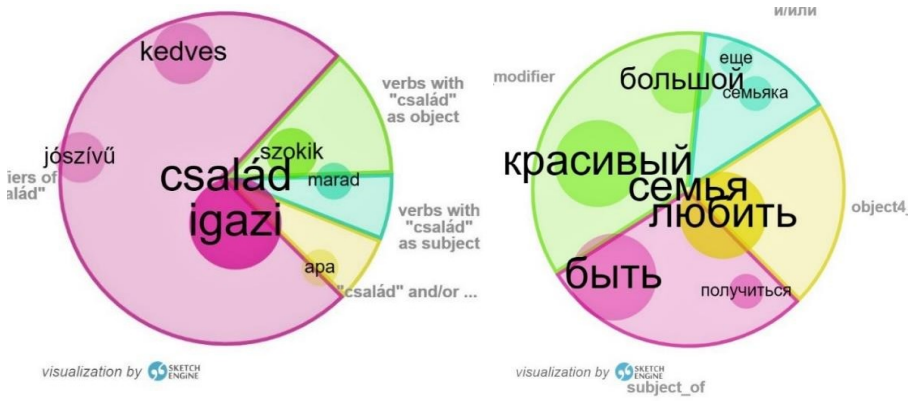

Figure 3. Visualization of the associations of stimulus words *CSALÁD (FAMILY)* and *СЕМЬЯ (FAMILY)*[1]

[1] Labels of the Hungarian chart: család (family); igazi (real); jószívű (warm-hearted); kedves (kind); szokik (used to); marad (stay); apa (father). Russian labels: семья (family); красивый (nice); любить (love); быть (be); получиться (form); большой (big); ещё (and); семьячка (family [diminutive]).

4.1.4. Toy/game

In the Hungarian language, there are no separate lexemes for *toy* and *game*, the two denotations are expressed by the same word: *játék*. In Russian however *toy* is translated as *игрушка* and *game* as *игра*. In this research, *игрушка (toy)* was selected from the stimuli of the Russian children's associative thesaurus [41] and *játék (toy/game)* was chosen as the Hungarian equivalent, keeping in mind that the Hungarian word has a narrower meaning.

Notwithstanding the semantic differences, both groups of respondents associated *toy* with the same two most frequent verbs (activities): *play* and *like/love*. The top noun associations also displayed remarkable similarities depicting the objects preschoolers play with on a daily basis including a *doll* and a *car*. For the above reasons, *игра (game)* appeared exclusively among the Russian associations, while the Hungarian group frequently mentioned *Lego* and *train*.

Adjectives from the collection of the Hungarian associations diverged from the Russian results: Russians described a *toy* as *favourite*, *small*, *real*, and *soft*, meanwhile *good* and *with dolls* proved to be the most typical description of *toy* for the Hungarian participants.

Table 7

Associations of the stimulus words *JÁTÉK (TOY/GAME)* and *ИГРУШКА (TOY)*, results of the research

játék *(toy/game)*			***игрушка*** *(toy)*		
NOUN	ADJ	VERB	NOUN	ADJ	VERB
legó ***(Lego) 15***; *baba* ***(doll) 13***; *autó* ***(car) 11***; *vonat* ***(train) 9***	*jó* ***(good) 6***; *babás* ***(with dolls) 5***	*játszik* ***(play) 75***; *szeret* ***(like) 31***; *rajzol* ***(draw) 10***; *legózik* ***(play Lego) 7***	*кукла* ***(doll) 36***; *машинка* ***(small car) 30***; *игра* ***(game) 15***; *машина* ***(car) 13***	*любимый* ***(favourite) 10***; *маленький* ***(small) 9***; *настоящий* ***(real) 6***; *мягкий* ***(soft) 5***	*играть* ***(play) 70***; *любить* ***(like/love) 14***

The charts below (Figure 4) shed light on a grammatical feature of the Russian language, namely the presence of prepositions as the most significant collocations of *игрушка (toy)*. Both prepositions *в* and *во* mean *in*, and are typical co-occurrences of the word *играть* in phrasal verbs such as *играть в игрушки (play with toys)*, *играть в самолётики (play with small helicopters)*, or *играть в ракеты (play with rockets)*. Typical collocations of the Russian word *игрушка (toy)* include *шарик (ball)* and *бусинка (bead)*; according to the collected Hungarian data, *játék (toy/game)* implies *homework* and *paw*. Typical Hungarian verb associations of *játék (toy/game)* are *szeret (love/like)* and *beszél (talk)*.

Both in the children's associations and in the adults' corpora, the word *toy* is connected with the verb *love/like*. The Hungarian data coincide in the adjective *jó (good)*, while the Russian results confirm the relevance of the adjectives *маленький (small)* and *любимый (favourite)*.

Table 8

Thesauri of the stimulus words *JÁTÉK (TOY/GAME)* and *ИГРУШКА (TOY)* from the Araneum comparative corpora

játék *(toy/game)*			***игрушка*** *(toy)*		
NOUN	ADJ	VERB	NOUN	ADJ	VERB
dolog ***(thing) 689,192***; *kép* ***(picture) 457,672***; *élet* ***(life) 901,881***; *film* ***(film) 240,233***	*jó* ***(good) 2,041,516***	*szeret* ***(like) 983,599***; *lát* ***(see) 1,050,286***; *ad* ***(give) 1,130,654***; *akar* ***(want) 799,720***	*вещь* ***(thing) 204,821***; *кукла* ***(doll) 25,945***; *подарок* ***(gift) 149,121***; *украшение* ***(clothing) 70,343***	*маленький* ***(small) 233,154***; *любимый* ***(favourite) 131,875***; *детский* ***(children's) 298,287***; *новогодний* ***(New Year's) 52,824***	*любить* ***(love) 226,488***

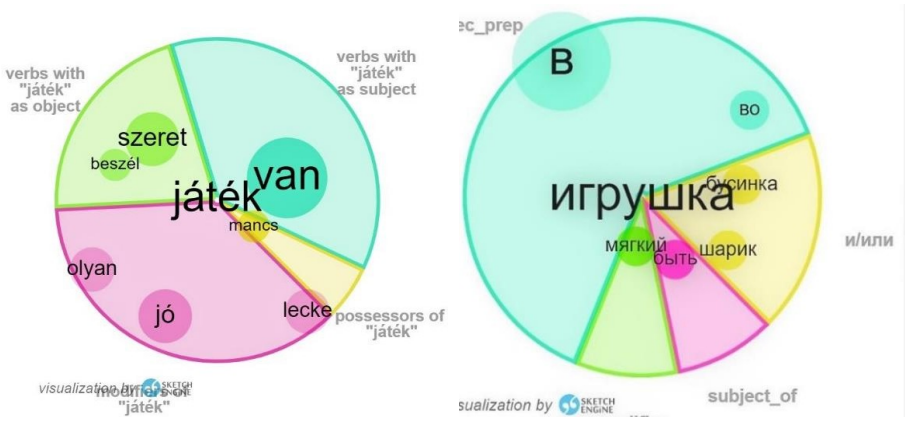

Figure 4. Visualization of the associations of stimulus words *JÁTÉK (TOY/GAME)* and *ИГРУШКА (TOY)*[1]

4.1.5. Home/house

The problem of equivalence arose in the case of the stimulus words *дом* and *otthon*. *Дом* means *house* and *home* at the same time, while in the Hungarian language *otthon* means *home* and *ház* denotes *house*. This is reflected in the associations of the Russian stimulus *дом* as it includes lexemes typical for the notion of a building: *построить (build)*, *кирпич (brick)*, *этаж (floor)*, and *крыша (roof)*. The Hungarian results reflected a meaning that is emotionally closer to the re-

[1] Labels of the Hungarian chart: játék (toy/game); jó (good); olyan (such); lecke (homework); szeret (love/like); beszél (talk); van (be); mancs (paw). Russian labels: игрушка (toy); в (in); во (in); мягкий (soft); быть (be) шарик (ball); бусинка (bead).

spondents, among the identified associations objects and concepts from the preschoolers' daily lives prevailed including, *szeret (like/love), játszik (play), alszik (sleep), játék (toy), ágy (bed),* and *szoba (room)*. One colour term appeared in the top Russian results, *белый (white)* and another one in the Hungarian data set: *sárga (yellow)*.

Table 9

Associations of the stimulus words *OTTHON (HOME)* and *ДОМ (HOUSE/HOME)*, results of the research

otthon *(home)*			***дом*** *(house/home)*		
NOUN	ADJ	VERB	NOUN	ADJ	VERB
ház ***(house) 41***; *játék* ***(toy/game) 24***; *ágy* ***(bed) 17***; *szoba* ***(room) 16***	*jó* ***(good) 34***; *nagy* ***(big) 28***; *emeletes* ***(storeyed) 12***; *sárga* ***(yellow) 9***	*lakik* ***(live) 32***; *szeret* ***(love) 19***; *játszik* ***(play) 17***; *alszik* ***(sleep) 9***	*кирпич* ***(brick) 33***;*этаж* ***(floor) 32***; *человек* ***(person) 26***; *крыша* ***(roof) 16***	*большой* ***(big) 33***; *белый* ***(white) 22***; *высокий* ***(high) 19***; *красивый* ***(nice) 11***	*жить* ***(live) 77***; *построить* ***(build) 10***; *сделать* ***(did) 7***

The aforementioned results are presented below (Table 10). Similar to the associations of the noun *family*, the adjective *настоящий (real)* is one of the most typical adjectives of the analysed word *дом (house/home)*.

Table 10

Thesauri of the stimulus words *OTTHON (HOME)* and *ДОМ (HOUSE/HOME)* from the Araneum comparative corpora

T10					
otthon *(home)*			***дом*** *(house/home)*		
NOUN	ADJ	VERB	NOUN	ADJ	VERB
ház ***(house) 325,627***; *hely* ***(place) 845,533***; *gyerek* ***(child) 412,046***; *lakás* ***(flat) 142,856***	*fiatal* ***(young) 318,105***; *szép* ***(nice) 390,404***	*szeret* ***(love) 983,599***; *lát* ***(see) 1,050,286***; *akar* ***(want) 799,720***; *vesz* ***(buy) 921,675***	*квартира* ***(flat) 376,505***; *здание* ***(building) 257,693***; *помещение* ***(premises) 257,693***; *место* ***(place) 257,693***	*новый* ***(new) 1,400,083***; *жилой* ***(residential) 134,344***; *большой* ***(big) 911,214***; *данный* ***(given) 731,108***	*стать* ***(become) 1,074,175***; *иметь* ***(have) 1,243,296***; *являться* ***(appear) 1,282,940***

Otthon (home) and *ház (house)* are closely connected in the Hungarian language based on both the results of this survey and the large-scale data of the Araneum corpus. The Russian results suggest that the main characteristic of *дом (house/home)* is *большой (big)*.

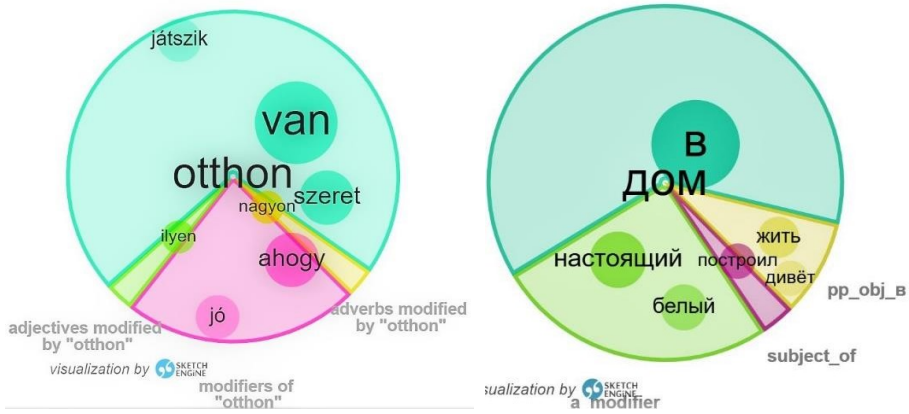

Figure 5. Visualization of the associations of stimulus words *OTTHON (HOME)* AND *ДОМ (HOUSE/HOME)*[1]

4.2. Overall data

Based on the respondents' utterances, two corpora were created from the Hungarian and the Russian data set (see Table 11). We excluded articles, conjunction words, prepositions, and modal verbs from this summary to be able to focus on the key elements. The analysis of the overall inputs of the Hungarian (HU100) and the Russian (RU100) research phases demonstrates the following phenomena observable in the linguistic consciousness of the Russian and Hungarian preschool children: 1) the most frequent noun in both cultures was *mother (мама/anya)*, followed by *father (папа/apa)*; 2) the nouns *child* or *kid (малыш/gyerek)*; *house* or *home* (*дом/ház)*; and *friend (друг/barát)* play a central role in the linguistic consciousness of preschoolers in both countries, ranking in the top 10 most frequent items; 3) *playing* is a key concept in both corpora: in the Hungarian corpus, it appears in the lexical item *játék (game/toy)*, while in the Russian data set in the form of a verb *(играть/play)*; 4) the Russian results are more diverse in terms of word classes. Top 15 Russian responses include personal pronouns (3), nouns (6), adjectives (4) and verbs (2), while the Hungarian top 15 is limited to personal pronouns (4) and nouns (11).

Keywords are useful and illustrative indicators of a corpus: they signal those lexemes that are significantly more frequently present in the investigated corpus than in the reference corpus. In this research, two giant corpora were used as references to the overall results of the Hungarian (HU100) and Russian (RU100) research phases. The two reference corpora are web-based collections of texts: Russian Web 2011 (14,553,856,113 words) and Hungarian Web 2012 (2,572,620,694 words) [10].

[1] Labels of the Hungarian chart: otthon (home); játszik (play); van (be); szeret (love), ilyen (such); jó (good), ahogy (how); nagyon (a lot). Russian labels: дом (home); в (in); настоящий (real); белый (white) построил (built); жить (live).

Table 11

Top 15 most frequent words (excluding articles, conjunction words, prepositions, and modal verbs)

№	**Hungarian** (HU100)		**Russian** (RU100)	
	word	frequency	word	frequency
1	*én* ***(I)***	228	*он* ***(he)***	161
2	*ő* ***(he/she)***	194	*мама* ***(mother)***	158
3	*anya* ***(mother)***	145	*я* ***(I)***	144
4	*gyerek* ***(child)***	129	*папа* ***(father)***	136
5	*apa* ***(father)***	118	*маленький* ***(small)***	117
6	*játék* ***(game/toy)***	111	*человек* ***(person)***	103
7	*barát* ***(friend)***	77	*дом* ***(house/home)***	75
8	*mi* ***(we)***	76	*играть* ***(play)***	69
9	*család* ***(family)***	58	*малыш* ***(kid)***	54
10	*ház* ***(house)***	58	*мы* ***(we)***	53
11	*ők* ***(they)***	54	*большой* ***(big)***	50
12	*anyuka* ***(mother)***	53	*друг* ***(friend)***	47
13	*ruha* ***(clothes)***	52	*люблю* ***(love)***	47
14	*víz* ***(water)***	48	*белый* ***(white)***	47
15	*szárny* ***(wing)***	46	*хорошая* ***(good)***	45

Keywords were selected in two groups: single-words (containing one lexeme) and multi-words (containing two lexemes). Results of the single-word keywords search display a clear reference to the general activities of preschool children (Table 12). In the Hungarian context, they tend to play *Lego*, play with the *dollhouse*, *play hide and seek*, *play with dolls*, *play board games*, play with *dinos*, *play tag*, and play with *pony*. In the Russian context, children mentioned the following activities according to the list of the top 15 key single words: play *Lego*, *play hide and seek*, *play tag*, and play with *cars*. Playing Lego is without doubt the most popular activity in both groups, taking first and third place in the Hungarian sample, and second in Russian.

Further activities that appeared in the key single word analysis of the Hungarian data included *misbehaving (rosszalkodik)*, while in the Russian context they were *to be friends* (*дружить*) and *to draw (чертить)*, as well as *pencil (карандаш)* referring presumably to the activity of drawing as well.

The Hungarian results reflected preschool children's linguistic consciousness centred around their daily activities, toys, and playing games (12 key single words of 15). However, this was less typical in the Russian sample where only six lexical items of 15 referred to such activities or to toys. Furthermore, family members appeared in a significantly higher proportion in the Russian sample where *bro* (*братик)* took first place, and *little sister (сестрёнка)* and *dad (папа)* also ranked in the top 15. The Hungarian sample only included one reference to family members, ranking no. 14 of 15: *kistesó (small brother/sister)*.

The appearance of *Jézuska* (little Jesus) and *szarv* (horn) can only be observed in the Hungarian results. In the Russian sample *крылышко (little wing)* and *крыльев (wings)* are over-represented when compared to the reference corpora. Both items may have religious/cultural connotations.

Table 12

Top 15 key single-words (without proper names, double occurrences)

№	**Hungarian** (HU100)		**Russian** (RU100)	
	word	frequency	word	frequency
1	*legózik* (***play Lego***)	13	*братик* (***bro***)	31
2	*babakonyha* (***dollhouse***)	11	*лего* (***Lego***)	15
3	*lego* (***Lego***)	8	*прятки* (***hide and seek***)	16
4	*bújócska* (***hide and seek***)	9	*догонялки* (***play tag***)	7
5	*babázik* (***play with dolls***)	10	*ребёнок* (***child***)	5
6	*szarv* (***horn***)	28	*дружить* (***be friends***)	49
7	*Jézuska* (***little Jesus***)	14	*сестрёнка* (***little sister***)	5
8	*társasozik* (***play board game***)	6	*домик* (***small house***)	4
9	*plüss* (***plush***)	12	*крылышко* (***little wing***)	10
10	*bújócskázik* (***play hide and seek***)	6	*крыльев* (***wings***)	3
11	*rosszalkodik* (***misbehave***)	8	*чертить* (***draw***)	3
12	*dínósat* (***play with dinos***)	4	*карандаш* (***pencil***)	3
13	*fogócskázik* (***play tag***)	5	*чёрный* (***black***)	3
14	*kistesó* (***small brother/sister***)	8	*папа* (***dad***)	146
15	*póni* (***pony***)	10	*машинка* (***car*** *[diminutive]*)	46

Several examples of lacunas could be identified as a result of the keyword analysis. Lacunas are linguistic gaps, or lexemes with no equivalent in the other, contrasted culture/language. These lexical items can be explained in the light of the so-called lacuna theory [4, 18, 46]. The Russian word *дружить (to be friends)* is an example of a cross-cultural lacuna as the expression has no Hungarian equivalent. Its meaning is an activity when the participants stay together and spend time or play together as friends. Similarly, *társasozik (play board games)* is a widespread term often used by Hungarian children that does not have a clear equivalent in the Russian language.

A second type of keywords are key multi-words consisting of two lexemes, such as *igazi család (real family)*. These words frequently appear together in the observed language; and, similar to key single-words, appear more frequently in the analysed corpus than in the reference corpus. Table 13 displays these strong word connections that are over-represented in the preschool children's corpus, with a minimum of three occurrences.

Similarly to the key single-word results (Table 12) the Russian sample contains more items connected to the family *(лучшая семья/best family; красивая семья/nice family)* as well as reference to family members *(младшая сестра/younger sister)*. The key role of family in the Hungarian preschoolers' linguistic consciousness is reflected by the strongest keyword combination, *igazi család (real family)*. However, there is no further example of the notion of family or of family members in the Hungarian list of most typical key multi-words.

Black and *white* are seemingly the most typical colours preschoolers actively use in both languages. They appeared six times of 18 in the Hungarian sample and three times of 17 in the Russian results. Two more colours were identified in the top results of the Hungarian sample including *red (piros szarv/red horn)* and *blond (szőke haj/blond hair)*.

Table 13

Top key multi-words (with at least 3 occurrences)

№	**Hungarian** (HU100)		**Russian** (RU100)	
	multi-word	frequency	multi-word	frequency
1	*igazi család* ***(real family)***	8	*наименьший малыш* ***(tiny*** *[lit.:**smallest]** **baby)***	10
2	*fehér ruha* ***(white dress/clothes)***	8	*наименьший человек* ***(tiny*** *[lit.:**smallest]** **person)***	7
3	*külföldi ember* ***(foreign person)***	6	*настоящий дом* ***(real house/home)***	6
4	*nagy gyerek* ***(big kid)***	5	*плохое слово* ***(bad word)***	5
5	*másik ország* ***(other country)***	5	*чёрный цвет* ***(black colour)***	5
6	*fekete szarv* ***(black horn)***	5	*лучший друг* ***(best friend)***	4
7	*rossz gyerek* ***(naughty kid)***	4	*наибольший кирпич* ***(huge*** *[lit.:**largest]** **brick)***	4
8	*szőke haj* ***(blond hair)***	4	*лучшая семья* ***(best family)***	4
9	*nagy ház* ***(big house)***	4	*худший человек* ***(worst person)***	4
10	*fekete ruha* ***(black dress/clothes)***	4	*белый дом* ***(white house)***	3
11	*fehér szárny* ***(white wing)***	4	*наибольший дом* ***(huge*** *[lit.:**largest]** **house)***	3
12	*fekete ceruza* ***(black pencil)***	3	*игре прятки* ***(hide and seek game)***	3
13	*plüss cica* ***(plush kitten)***	3	*большая комната* ***(big room)***	3
14	*rossz dolog* ***(bad thing)***	3	*красивая семья* ***(nice family)***	3
15	*emeletes ház* ***(storeyed house)***	3	*младшая сестра* ***(younger sister)***	3
16	*fehér a ruha* ***(the dress/clothes is white)***	3	*наибольший человек* ***(huge*** *[lit.:**largest]** **person)***	3
17	*piros szarv* ***(red horn)***	3	*чёрный человек* ***(black person)***	3
18	*Zsuzsi vonat* ***(Zsuzsi train)***	3		

Conclusions

Any interdisciplinary research such as the current one inevitably needs to take several perspectives into account. Thus, linguistic, psychological and pedagogical aspects were emphasised in the design, execution and interpretation of the study. Relying on the psychological and psycholinguistic theories [activity theory: Leontiev 1978; speech activity theory: Leontiev 1993; lacuna theory: Markovina 2006], as well as the methods [association experiment: Ufimtseva 2014a; shoulder-to-shoulder method: Griffin et al. 2014] and blending those with corpus linguistics and pedagogical theories [childhood as a social construction: James&Prout, 1997], the research aimed at shedding light on the linguistic consciousness of 4- to 5-year-old preschool children coming from the target cultures.

Applying perspectives stemming from multiple disciplinary fields and blending methods from both linguistic and pedagogical areas, contributed to gaining

an overall picture of Russian and Hungarian preschoolers' linguistic consciousness. In this cross-cultural research, not only an interdisciplinary approach was successfully applied, but the linguistic methods were also merged into an amalgam of the classical Russian psycholinguistic perspective and of the state-of-the-art corpus linguistic frame of reference.

The authors identified a significant proportion of universalistic features in terms of the Russian and Hungarian preschoolers' linguistic consciousness. A *family* is seen in both the Russian and the Hungarian sample as represented by the *mother* in first place, followed by the *father* (top 2 noun associations in both groups), the main characteristics of a family are *kind*ness and being *good* (top 2 adjective associations in both samples) and the most representative emotion (verb) in both countries proved to be *love*. Similarly, in both groups, friend is described as a *kind*, *good* and *cute/nice* person, and is associated with the activities, to *love* and to *play* with. Furthermore, a child is characterised as *small*, *good* and *kind* in both samples, and main noun and verbal associations of a *child* include the response-words *baby/kiddy*, *mother*, to *grow up* and to *love*. The above results reconfirm that the linguistic consciousness in this age group bears a significant proportion of similar, universalistic features, with key notions of childhood (*friend, child, family, toy, home*) being similar across Hungarian and Russian cultures.

Nevertheless, a remarkable amount of culture- and language-bound items were identified in the research such as the strong association between *друг–дружить* or *маленький–малыш* (*friend–to be friends; small–kiddy*): etymologically and morphologically similar words); and *чёрт–чёрточка–чёрный* (*small devil–line–black*: three lexical items in the Russian language of similar sounding but different etymology). Moreover, several phenomena of lacunarity [18, 46] were pinpointed in the research. The phenomenon of the Lacuna Paradox [46] between Hungarian and Russian stimulus words was identified, including the words *дом* (that denotes *house* and *home* at the same time in the Hungarian); the Hungarian word *játék* that denotes both *toy (игрушка)* and *game (игра)* in Russian; or *ördög (devil)* that can be equally translated into *дьявол (devil), чёрт (small devil), бес (demon)* or *сатана (Satan)*.

Examples abound of the effects of globalization revealed on identifying objects and notions of modern technology, including *tablets, cartoons, brand names (Lego, Duplo, Trudi)*. *Lego* and to *play with Lego* proved to be significantly over-represented in the children's responses (compared to the respective adults' corpora). Besides playing with *Lego*, preschoolers' linguistic consciousness precisely described some of the central activities of children in both countries. The key single-word analysis proved to be an effective tool for gaining insight into preschoolers' activities: *hide and seek, tag, play with dinos, with pony, play with dolls* or *draw*.

All in all, the selection of the 10 stimulus words and the more detailed analysis of the five items most closely connected with childhood *(friend, child, family, toy/game, home/house)*, as well as their investigation with a combination of the association experiment, corpus linguistic methods and a pedagogical approach

assisted us in identifying and comparing Hungarian and Russian preschoolers' perception of the outside world in a more successful way.

References

1. Vygotsky, L.S. (1962) *Thought and language.* Cambridge MA: MIT Press. DOI: 10.1037/11193-000

2. Vygotsky, L.S. (1978) *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

3. Leontiev, A.A. (1993) Yazykovoe soznanie i obraz mira [Language consciousness and the image of the world]. In: *Yazyk i soznanie: Paradoksal;naya ratsionalnost'* [Language and consciousness: Paradoxical rationality]. Moscow: Institute of Linguistics, RAS. pp. 16–21.

4. Sorokin, J.A. (1993) Die Lakunen-Theorie. Zur Optimierung interkultureller Komminikation. In: Ertelt-Vieth, A. (ed.) *Sprache, Kultur, Identität. Selbst-und Fremdwahrnehmungen in Ost-und Westeuropa.* Frankfurt am Main; Berlin: Peter Lang. pp. 163–173.

5. Tarasov, Y.F. (1996) Yazykovoe soznanie – perespektivy issledovaniya [Language consciousness – research perspectives]. In: *Etnokulturnaya spetsifika yazykovogo soznaniya* [The ethnocultural specificity of language consciousness], Moscow: Institute of Linguistics, RAS. pp. 7–22.

6. Ufimtseva, N.V. (2014a) Russian Psycholinguistics: Contribution to the Theory of Intercultural Communication. *International Communication Studies*. XXIII:1. pp. 1–13.

7. Ufimtseva, N.V. (2014b) The Associative Dictionary as a Model of the Linguistic Picture of the World. *Procedia – Social and Behavioral Sciences*. 154. pp. 36–43. DOI: 10.1016/j.sbspro.2014.10.108

8. Lenart, I. (2017) Associations and verbal consciousness: an analysis based on four English and one Hungarian translation of Bulgakov's novel: The Master and Margarita. *Neohelicon.* 44. pp. 487–504. DOI: 10.1007/s11059-017-0386-9

9. Griffin, M., Lahman, M. & Opitz, M. (2014) Shoulder-to-shoulder research with children: Methodological and ethical considerations. *Journal of Early Childhood Research.* 14 (1). pp. 18–27. DOI: 10.1177/1476718X14523747

10. Kilgarriff, A. et al. (2014) The Sketch Engine: ten years on. *Lexicography.* 1. pp. 7–36. DOI: 10.1007/s40607-014-0009-9

11. Hendrick, H. (2000) The child as a social actor in historical sources. In: Christiansen, P. & James, A. (eds) *Researck with children.* London: Falmer Press.

12. Heywood, C. (2001) *A History of Childhood.* Blackwell Publishers Ltd.

13. Endrody-Nagy, O. (2016) Paintings and Illuminated Manuscripts as Sources of the History of Childhood: Conceptions of Childhood in the Renaissance, In: Benedek, A. & Veszelszki, Á. (eds) *In the Beginning was the Image: The Omnipresence of Pictures: Time, Truth, Tradition*. Frankfurt am Main: Peter Lang GmbH, Internationaler Verlag der Wissenschaften. pp. 91–100.

14. Endrody-Nagy, O. (2018) A brief history of childhood – as seen on visuals. *Revista de Didácticas Específicas.* 19. pp. 100–107. DOI: 10.15366/didacticas2018.19

15. Scherba, L.V. (1974) *Language system and speech activity*. Leningrad: Nauka. (In Russian).

16. Ushakova, T.N. (2003) Yazykovoye soznaniye i prinzypy ego issledovaniya [Verbal consciousness and the principles of its research]. In: *Yazykovoye soznanie i tekst: teoreticheskiye i prikladnye aspekty* [Verbal consciousness and text: theoretical and practical aspects]. Barnaul. pp. 6–17.

17. Karaulov, Yu.N. (2000) Pokazatyeli natsionalnogo mentaliteta v associativno-verblanoy setii [National mentality indicators in an associative-verbal network]. In: Ufimtseva, N.V. (ed.) *Yazikovoe soznanie i obraz mira* [Language consciousness and image of the world]. RAN Institute of Lingusitics. Moscow. pp. 191–206.

18. Markovina, I. & Sorokin, Y. (2006): The lacuna phenomenon and the problem of foreign culture comprehension: An experimental study of lacuna elimination strategies. In: Panasiuk, I. & Schröder, H. (Hrsg.) *Lakunen-Theorie: Ethnopsycholinguistische Aspekte der Sprach- und Kulturforschung. Semiotik der Kultur*. Band 5. Münster-Berlin-Hamburg-London-Wien: LIT Verlag. pp. 154–160.

19. Ufimtseva, N.V. (2006) Etnopsikholingvistika: vchera i segodnya [Ethnopsycholinguistics: yesterday and today]. *Voprosy psikholingvistiki – Journal of Psycolinguistics*. 4. pp. 92–100.

20. Galton, F. (1879) Psychometric experiments. *Brain*. 2. pp. 149–162.

21. Cramer, P. (1968) *Word Association*. London: Academic Press.

22. Nelson, D.L., McEvoy, C.L. & Schreiber, T.A. (1998) *The University of South Florida word association, rhyme, and word fragment norms*. [Online] Available from: http://www.usf.edu/FreeAssociation/.

23. De Deyne, S. & Storms, G. (2008) Word Associations: Network and semantic properties. *Behavior Research Methods*. 40/1. pp. 213–231.

24. Kovács, L., Orosz, K. & Pollner, P. (2012) Magyar szóasszociációk hálózata [The network of Hungarian word associations]. *Magyar Tudomány*. 6. pp. 699–705.

25. Gravino, P. et. al. (2012) Complex structures and semantics in free word association. *Advances in Complex Systems*. 15/3-4. p. 1250054-1.

26. Bóta, A. & Kovács, L. (2015) The community structure of word association graphs. *The Proceedings of the 9th International Conference on Applied Informatics*. Eger. pp. 113–120.

27. Kovács, L. et al. 2020) Központi szavak és azok közösségei a mentális lexikon hálózataiban. In: Balázs, G., Imrényi, A. & Simon, G. (eds.) *Hálózatkutatás – Hálózatok a nyelvben*. Budapest: Magyar Szemiotikai Társaság. pp. 207–220.

28. Ufimtseva, N.V. (2009) Image of the world of Russians: the systemic characteristics and the content. *Yazyk i kul'tura – Language and Culture*. 4 (8). pp. 98–111.

29. Kovacs, L. (2013) *Fogalmi rendszerek és lexikai hálozatok a mentális lexikonban* [Concept systems and lexical networks in the mental lexicon]. Budapest: Tinta Konyvkiadó.

30. Kovacs, L. (2014) *Agykapocs.hu / Connectyourmind.com online database.*

31. Ufimtseva, N.V. (2011) *Yazykovoe soznanie: dinamika i variativnost'* [Language consciousness: dynamics and variability]. Moscow: Institute of Linguistics, RAS.

32. Vasilevich, A.P. (1987) *Issledovaniye leksiki v psikholingvisticheskom eksperimente* [Investigation of lexis in a psychological experiment]. Moscow: Nauka.

33. Ilyasov, I. (1974) Psikhologocheskiye I fisiologicheskiye metody psikholongvisticheskogo issledovaniya [Psychological and physiological methods of psycholinguistic research]. In: Leont'ev, A.A. (ed.) *Osnovy teorii rechevoy deyatel'nosti* [Fundamentals of the theory of speech activity]. Moscow: Nauka.

34. Lengyel, Zs. (2008) *Encyclopedia of Hungarian Norms of Associations.* Budapest: Tinta Kiado.

35. Crotty, M. (ed.) (2008) *The foundations of social research: meaning and perspective in the research process.* Thousand Oaks, CA; Sage.

36. James, A. & Prout, A. (eds) (1997) Constructing and Reconstructing Childhood. In: *Contemporary Issues in the Sociological Study of Childhood.* London; Washington, D.C.: Falmers Press, Taylor & Francis Group. DOI: 10.4324/9781315745008

37. Karaulov, Yu.N. et al. (2002) *Russkiy assotsiativnyy slovar'* [Russian Associative Dictionary]. Vols 1, 2. Moscow: AST-Astrel'.

38. Ufimtseva, N. et al. (eds) (2004) *Slavyanskiy Assotsiativnyy Slovar': Russkiy, Belorusskiy, Bolgarskiy, Ukrainskiy* [Slavic Associative Dictionary: Russian, Belorussian, Bulgarian, Ukrainian]. Moscow: Institute of Linguistics, RAS.

39. Cherkasova, G.A. & Ufimtseva, N.V. (2014) *Russian Regional Associative Dictionary "EURAS"*. Moscow.

40. Sergieva, N.S. (2006) Semanticheskiy geshtal't i yadro yazykovogo soznaniya russkikh [Semantic gestalt and the core of Russians' verbal consciousness]. 2006. *SUSU Bulletin. Social Sciences and Humanities*. 2 (57). pp.160–165.

41. Sokolova, T.V. (1998) *Assotsiativnyy tezaurus rebyenka 3–6 let* [The associative thesaurus of children aged 3 to 6]. Philology Dr. Diss. Astrakhan.

42. Meisinger, E.B. et al. (2004) Interaction quality during partner reading. *Journal of Literacy Research,* 36. pp. 111–140. DOI: 10.1207/s15548430jlr3602_1

43. Ginnis, P. (2007) *The Teacher's Toolkit.* Crown House, UK.

44. Benko, V. (2014) Compatible Sketch Grammars for Comparable Corpora. *Proceedings of the XVI EURALEX International Congress: The User In Focus*. 15–19 July 2014. Bolzano/Bozen: Eurac Research. pp. 417–430.

45. Leontiev, A.N. (1978) *Deyatel'nost'. Soznanie. Lichnost'* [Activity, Consciousness, and Personality]. Moscow: Politizdat.

46. Lenart, I. (2016) *Intercultural lacunae in Hungarian-Vietnamese communication, with emphasis on entrepreneurial interactions*. Doctoral dissertation. Budapes: Eotvos Lorand University.

Языковой образ мира детей дошкольного возраста: кросскультурное русско-венгерское сопоставительное исследование на основе словесных ассоциаций
И. Ленарт, И.Ю. Марковина, О. Эндроди

Ключевые слова: ассоциативный эксперимент, языковое сознание, раннее детство, психолингвистика, межкультурная коммуникация

Данное российско-венгерское кросскультурное междисциплинарное исследование посвящено изучению *языкового образа мира* или *языкового сознания* детей дошкольного возраста от 4 до 5 лет. С целью выявить сходства и различия в восприятии мира детьми авторы применили метод словесных ассоциаций. Были отобраны 10 слов-стимулов, на которые в ходе свободного ассоциативного эксперимента были получены реакции от респондентов исследуемой группы (N = 100 в каждой стране). Слова-стимулы отобраны таким образом, чтобы они обозначали понятия из ближайшего социального, материального и духовного окружения ребенка дошкольного возраста. Результаты, сгруппированные в частотные списки, сопоставили как фрагменты языкового сознания носителей венгерского и русского языков. Полученные данные позволяют установить, прежде всего, универсальные характеристики русского и венгерского языкового сознания в раннем детстве: например, схожим оказалось восприятие *семьи*, *друга*, *ребенка*. Исследование показало, что как российские, так и венгерские дети ассоциируют *семью* в первую очередь с *матерью*, а затем с *отцом*. И в том и в другом сознании *семья* характеризуется ассоциатами *добрый*, *хороший, любовь*. *Друг* описывается прилагательными *добрый*, *хороший,* а основные действия, ассоциируемые с другом, – *любить* и *играть*. В обеих группах *ребенок* описывается как *маленькое*, *хорошее* и *доброе* существо, а основными словесными ассоциациями на данный стимул являются: *малыш*, *мама*, *расти* и *любить*. В восприятии 10 исследованных понятий выявлены и расхождения, отчасти как результат различий лингвокультурного фона респондентов (например, ассоциативные поля слов-стимулов: *дьявол* и *ангел*), а отчасти как следствие языковой лакунарности, например *játék* (*игра* / *игрушка*). На материале венгерских и русских параллельных слов-стимулов авторы также описывают феномен *лакунарного парадокса*. Например, слово *дом* (которое в венгерском языке означает *здание* и *жилище, квартира* одновременно); венгерское слово *játék*, которое в русском языке имеет значения «*игрушка*» и «*игра*»; или *ördög* (*дьявол*), что может переводиться как *дьявол*, *чёрт*, *бес* или *сатана*. Результаты продемонстрировали эффекты глобализации в обеих культурах, о чем свидетельствует присутствие в ассоциативных полях названий международных торговых марок (*Lego*) и объектов современных технологий (*планшеты*).

Кроме того, исследование позволило выявить типичные занятия дошкольников, отраженные в их языковом образе мира: например, *прятки*, *метки*, *игры с динозаврами*, *пони*, *игры с куклами* и *рисование*. В заключение следует отметить, что сочетание психолингвистических и педагогических подходов оказалось эффективным способом изучения языкового сознания дошкольников, в частности представлений о социальном, материальном и духовном окружении, имеющем наиболее важное значение для русских и венгерских детей исследуемой возрастной группы.