

What Factors Help and Hinder Children's Progress in Reading?

[Inna Antipkina](#), [Marina Kuznetsova](#), [Elena Kardanova](#)

Received in
January 2017

Inna Antipkina

Junior Researcher at the Center for Monitoring the Quality in Education, Institute of Education, National Research University Higher School of Economics. E-mail: iantipkina@hse.ru

Marina Kuznetsova

Candidate of Sciences in Pedagogy, Research Fellow at the Center for Monitoring the Quality in Education, Institute of Education, National Research University Higher School of Economics. E-mail: mikuznetsova@hse.ru

Elena Kardanova

Candidate of Sciences in Mathematics and Physics, Associate Professor, Head of Center for Monitoring the Quality in Education, Institute of Education, National Research University Higher School of Economics. E-mail: ekardanova@hse.ru

Address: 20 Myasnitskaya St, 101000 Moscow, Russian Federation.

Abstract. This study looked at the effects of phonological preparedness and vocabulary size in children, who just started primary school, on their progress in reading at the end of the first grade while controlling for other factors that can be related to increasing or decreasing reading achievements (such as SES, parenting activities and non-cognitive development of children). The study was conducted using data from the IPIPS project which assesses the preparedness of children for school-

ing and their progress at the end of the first school year. The sample consisted of 2740 first-graders living in two large Russian cities (Krasnoyarsk and Kazan) whose reading skills were assessed twice, at the beginning and at the end of the 2014–2015 school year. The results demonstrated that low levels of phonological ability and vocabulary are related to lower results not only for those who just started learning to read (as is suggested by the theoretical framework of reading skills acquisitions) but also for children who already have basic reading skills or read well. To compensate for this, special teaching approaches might be needed. Among family factors the main predictors of reading results were the level of the father's education and language at home. Parenting activities related to reading were divided into informal (reading a book, discussing a book, reading street signs out loud during walks etc.) and formal (deliberate teaching of letters and writing letters or words), with informal activities being a significant predictor of reading outcomes at the end of the first year. Conclusions and limitations of the study are discussed.

Keywords: primary school, progress in reading, parenting activities, noncognitive development, phonological ability, vocabulary.

DOI: 10.17323/1814-9545-2017-2-206-233

The study was conducted with the support of the Russian Foundation for the Humanities (project No. 16-18-10401 "Predicting Teaching Effectiveness in Elementary Education").

Why study reading development?

The importance of conscious reading skills for successful learning not in question. The 2009 Federal State Elementary Education Standard describes conscious reading of texts of different styles and genres as a critical meta-skill. Back in the 1930s, researchers had already justified the need for reading acquisition as follows: "Children learn to read during the first three years of school and read to learn the rest of the time." [Hilliard, Wilson 1936: 226] Although alternative methods of obtaining information outside of school have become available since then, this skill has remained the most frequent predictor of academic performance in pedagogical psychology research. For instance, "Matthew effects" in reading development have been the subject of ample research over the last three decades. This research provides strong evidence that individual differences in reading development widen the achievement gap between low- and high-performers [Hattie, Dörfler, Artelt 2014: 203]. Relevant studies have also revealed some correlations between the level of reading development and academic performance and proved the peculiar importance of such reading skill components as reading speed and reading comprehension (i. e. constructing meaning from decoded phonemes) in general academic progress. Therefore, not only does research in reading development help enhance the reading skills of individual students but it also solves more generalized methodological problems related to reducing the inequality of children's access to education.

The question about the most effective ways of preparing preschool children for reading development remains relevant. The deep-seated notion in parents' minds is that studying is easier for first-graders who learned to read at preschool age. As a consequence, there has long been a social demand for preschool reading development. The problem is that in this case the necessary sequence of stages—first preparing kids for reading development, then teaching them to read—is inverted. Full-value reading acquisition is impossible without providing a good foundation for the skill. It is important to analyze again the reading performance of first-graders as determined by cognitive processes underlying the skill as well as preschool home literacy practices. The data obtained will help update the beliefs about effective techniques for getting children ready for reading development as well as improve the reading instruction methods used in elementary school.

Modern reading development models

Russian reading development methodology universally accepts the model proposed by DaniilElkonin. Having determined that "at the initial stage of development, reading is reproducing the phonetic form of a word based on its graphic (letter) representation" [Elkonin 1976: 20], he demonstrated that it was the phonological system of language—phonemes and their sequences—that underlay the reading process. Reading mechanisms rely directly on the phonetics and writing system of a specific language. In order to read in Russian, it is essential to dis-

criminate between vowels and consonants, hard and soft consonants, back and front vowels and their effects on the preceding consonant. In light of this, Elkonin identified three stages in initial reading development: (i) the preparatory stage, aimed at developing phonemic analysis skills and general phonological awareness; (ii) learning the system of vowels, their letter representations, and being guided by vowel letters when reading; and (iii) learning the system of consonants, their letter representations, and developing the basic reading mechanism [Ibid.: 64]. Obviously, phonological awareness plays a crucial role in this model. Meanwhile, it is worth pointing out that, despite the explicit focus on reproducing the phonetic forms, the meaning of words is also a priority from the very first days of teaching, since reading itself is senseless unless the reader is focused on understanding what they are reading. As initial reading skills develop—meaning that a child learns to build letter-sound correspondences, assign stress correctly, observe pausing and intonation rules—the technical aspect requires less and less effort from a child, who begins to focus more and more on extracting the meaning from texts.

Despite all the differences in reading instruction methods across various countries, we found it relevant to analyze foreign reading development models as well. Most present-day English-language models include three successive stages:

1. Logographic stage, where children perceive whole words as pictures. Visual representation of a word is perceived as a whole, indivisible symbol, associated with the word's meaning. At this stage, children can even "read" individual words because they remember their graphic features;
2. Phonological decoding, where children read by extracting necessary sounds (phonemes) from letters (graphemes);
3. Orthographic stage, where readers accumulate a set of familiar word elements (e. g. syllables, letter combinations, morphemes) and decode those elements or whole words instead of individual letters or phonemes when reading [Chiappe, Siegel 2006:135].

This English-language model, the most cited one, emphasizes the high importance of phonological awareness for reading acquisition.

Imaging findings show that phonological awareness may play even a greater role than has been believed up to now. According to neurobiologists, at the earliest stage of reading development neural brain structures involved in speech production are the most active [Goswami 2010: 318]. It is only as reading experience is accumulated that the neural mechanisms in the region involved in visual perception, dubbed "visual word-form area" (VWFA), grow more active. Although this part of the cortex serves to process visual forms, there is no reason to believe that it is also responsible for extracting the meaning from graphic representation of whole words, as the same region is activated when

reading non-words. Researchers assume that this area “stores” familiar word elements that are processed immediately. Therefore, drawing on the imaging data, we can assume that early reading strategies are based on phonological recoding of letters into phonemes, not on perceiving words as images (i.e. reading development does not begin with the logographic stage, as implied by the model described above). However, as compared to the findings of traditional studies on pedagogical psychology, the available imaging data is too small to draw firm conclusions yet. Which is why the logographic stage is still a fundamental element of most reading development models in English-speaking countries.

Russian researchers of the psychology of reading have determined that the key factors of successful reading acquisition include specific levels of spatial representation, visual perception and speech development. All speech development components are important for reading acquisition: pronunciation skills, general phonological awareness, identification of phonemes within words based on their distinctive features, in order to be able to reproduce the phonetic forms of words, and a high level of lexico-grammatical analysis and synthesis skills, in order to fully understand what has been read [Ananyev 1960: 456; Gvozdev 1961: 140; Yegorov 1953: 30; Zhinkin 1966: 14–15; Zhurova, Elkonin 1963: 225; Tsvetkova 1988: 189–190; Shvachkin 1948: 106; Elkonin 1958: 101; Elkonin 1962: 16].

Phonological awareness and phonological perception develop as children naturally learn to speak their native language. For the brain to identify each word as a unique and distinct set of sounds associated with a specific meaning, children should be able to discriminate among phonemes similar in their acoustic and articulatory characteristics. However, the specific phonological word analysis skill only develops when children start learning to read. Johannes Ziegler and Usha Goswami believe that the effect of literacy on spoken language processes is the most intriguing aspect of learning to read and write. They also compare, taking their cue from Uta Frith, the alphabetic code to an infectious virus: whole word sounds are automatically broken up into sound constituents, and “language is never the same again.” [Ziegler, Goswami 2005: 14; Frith 1998: 1051].

Vocabulary size used to be considered a more important reading development tool than even phonological awareness, right up until the early 1980s. For example, Irene Athey in her article of 1983 writes that “vocabulary development may be the single most important preparatory step for reading, but it must consist of true development of the ideas surrounding a concept and not just the dictionary definitions of more and more new words.” [Athey 1983: 198]. In order to provide this, new words should be learned in the context of their application. Russian researchers and speech-language pathologists associate the limited vocabularies of children with underdeveloped phonemic and phonological awareness: without understanding the meaning of words,

Figure 1. **Reading acquisition skills arranged in accordance with the extent to which teaching them is complex and time-consuming**

<i>Skills that develop comparatively quickly and easily</i>			<i>Skills that require a lot of time and repeated effort to develop</i>			
Ability to write one's name	Knowing the alphabet	Using high-frequency words	Ability to pronounce sounds properly	Phonological processes (phonological awareness)	Oral reading fluency	Vocabulary and comprehension

children pronounce them wrong, omitting sounds, changing their sequence, or substituting other sounds for them [Filicheva, Cheveleva, Chirkina 1993: 6, 153].

Contemporary foreign researchers divide reading development skills into constrained and unconstrained. The term “unconstrained” was coined to concisely describe the teaching of such abilities as “complex and time-consuming”. Figure 1 shows the vital reading development skills depending on their “constrainedness” [Dougherty Stahl 2011: 53].

Vocabulary development requires more time than development of phonological awareness. Unfortunately, school research and reading readiness tests focus most often on constrained skills, which also dominate the school curriculum, while truly important unconstrained skills are largely overlooked [Dougherty Stahl, 2011: 55].

How non-cognitive skills and environment contribute to reading acquisition

The level of reading development has to do with other factors as well, apart from the well-formedness of phonological, semantic and grammatical language components. First of all, these other factors include the non-cognitive development of children and the environment that they grow up in.

A number of studies have proved the correlation between non-cognitive—social, emotional and personal—development of children and their academic achievement. In particular, the development of self-regulation skills correlates with that of phonological awareness. The level of social skills predicts letter recognition, sound-letter correspondence and non-word repetition skills (according to teachers) [Ritchey 2004: 375]. Inattention affects reading development more than any other non-cognitive factors, which has been demonstrated by middle school longitudinal studies that did not involve direct assessment of reading skills [Dittman 2016: 660.].

A meta-analysis of 41 studies on the relation of parental involvement to urban elementary school student academic achievement [Jeynes 2005] revealed that all the analyzed studies but one report-

ed a high coefficient (0.7 on average) of positive correlation between parental involvement and the academic performance of school students. Regression coefficients were higher in the studies which used few control variables, so the author believes that parental involvement is an indicator of differentiating factors, such as socioeconomic status, ethnicity, etc.

Canadian researchers demonstrated in a five-year reading development longitudinal project [Sénéchal, LeFevre 2002] that reading fluency development trajectories vary across children depending on home literacy practices selected by their parents: informal shared book reading or teaching the alphabet and reading in a formal style. Meanwhile, reading fluency development trajectories were found to correlate with the results of reading fluency assessment in the 3rd grade. Shared book reading is a predictor of vocabulary size and listening comprehension, which are related directly to reading skills in the 3rd grade. Formal-style teaching of the alphabet and reading skills to children (where parents deliberately make children focus on letter recognition and writing, e. g. by reading an ABC book together) correlates with early development of literacy, which is related indirectly (via word reading at the end of the 1st grade) with the third-grade level of reading development. In addition, the authors establish that vocabulary, listening comprehension and the meta-linguistic skill of phonological awareness form a single factor in factor analysis, which appears to be important for understanding the relationship between vocabulary and phonological development.

Hypothesizing on the factors associated with reading progress in children

An analysis of studies produced a number of factors essential for reading development. These include, first of all, phonological awareness, vocabulary size, non-cognitive skills, and home literacy environment. The goal of this study was to assess the effects of these factors on reading progress in children at different stages of reading development.

The following hypotheses were tested during the course of research:

- (1) Children with larger vocabularies and better phonological awareness will have better reading progress during the first year in school, provided that the starting level of reading development is the same for all.
- (2) Small vocabularies and low phonological awareness inhibit reading progress even in children who learned to read at preschool age.
- (3) Reading progress in children is affected by the level of their social and emotional development.
- (4) Family factors (cultural capital, parental involvement) are significant predictors of progress at different stages of reading development.

Sampling and the assessment tool

The sample included 2,741 first-graders from Krasnoyarsk and Kazan who were involved in the IPIPS¹ (International Performance Indicators in Primary School) study in 2014–2015. Children’s skills were assessed twice: in autumn, soon after placement in the 1st grade, and in spring, shortly before the end of the academic year, which allows for measuring reading progress over the first year of school. The IPIPS tool provides a comprehensive assessment of children’s characteristics—their cognitive skills as well as the level of socioemotional development—and allows for collecting various contextual information (via parental and teacher surveys). This study focused on assessing reading abilities. Reading assessment tests were arranged in the IPIPS tool by increasing complexity and used the following assessment model:

- Basic perceptions of text structure: children are asked to show the beginning and the end of a sentence, the period, and the capital letter.
- Letter recognition: children are asked to name a few letters written on a sheet of paper.
- Word recognition: children are asked to read a few words.
- Ability to read a short text correctly (without deep understanding): the only focus of the tasks in this module was on the ability to read properly; no comprehension questions were asked.
- Reading and text analysis (close reading): this skill was assessed using passages with several multiple-choice tasks. Children were asked to choose the right words to put in a sentence from three available options in the process of reading, on the basis of what they understood. This is a task example: *Yulya bistro (zabralas', sobralas', probralas') i vyshla (u, ot, iz) doma.* (“Yulya (climbed, dressed, sneaked in) quickly and went out (by, from, of) the house.”)

In addition to reading assessment tests, children also performed vocabulary size tests (vocabulary knowledge tasks using words of varying frequency) and phonological awareness tests (real-word and non-word repetition as well as multiple-choice rhyme-matching tasks).

The IPIPS tool assessed children’s skills using the adaptive testing principle: if a child committed a specified number of mistakes, assessment in the current module was stopped and more complex modules were not offered. The adaptive testing algorithm allows for reducing the assessment time, encouraging children’s motivation, and preventing child fatigue.

Non-cognitive (personal, social and emotional) development of children was assessed through a teacher questionnaire that involved 11 aspects:

¹ IPIPS website: <http://ioe.hse.ru/ipips>

- comfort (adaptation to school conditions);
- independence/autonomy;
- self-confidence;
- concentration in teacher-led activities;
- concentration in independent activities;
- prudence / impulsivity;
- relationship with peers;
- relationship with adults;
- observance of rules;
- awareness of cultural differences;
- communication.

Teachers rated all the characteristics for every child on a five-point scale. Each parameter was accompanied by a detailed description of typical behavior corresponding to specific points on the scale, so that teachers could choose the best match based on their observations. Data on the eleven indicators of personal, social and emotional development was divided into two subscales: "behavior in the classroom" and "communicative skills". "Behavior in the classroom" embraces the skills that help children remain concentrated and observe school regulations and schedules as well as promote awareness of cultural differences (i. e. understanding that other people can have a different lifestyle and their customs should be respected). The "communicative skills" scale describes the level of children's independence and autonomy as well as their social skills, i. e. the ability to maintain relationships with other people: peers and adults, both in school and in broader social contexts.

The parent questionnaire data was used to build an index of parental involvement and collect information on students' families. The questionnaire consisted of 17 questions on home literacy practices, asking how often parents engaged in various preschool home activities, such as book reading, counting, table games, jigsaw puzzles, playing with construction toys, drawing, poem reciting, singing, etc. The unified index of parental involvement was built for all 17 variables by applying the scale method used in the TIMSS study [Martin, Preuschof 2008: 282] to parents' answers. In accordance with the theoretical framework borrowed from the publications of Monique Sénéchal and her coauthors [Sénéchal, LeFevre 2002], two separate parental involvement indexes were also built, one of formal parent involvement in reading development (which included the variables "learned how to write letters or words together" and "played letter games, e. g. letter blocks") and one of the informal reading acquisition practices ("read books to the child", "discussed what has been read in books", "played word games", "read aloud street billboards, signs, words displayed in shop windows, etc."), in order to find out which types of home literacy activities predicted reading performance better.

All the results (in reading, vocabulary, phonological awareness, social and emotional development) in the IPIPS project were transformed from raw scores to z-scores, with the mean of 50 and standard deviation of 10. The same scores were used to measure reading progress.

The data was analyzed using standard statistical methods (dispersion analysis, chi-squared test, factor analysis, regression analysis) and SPSS and HLM software.

Data description Tasks of different complexity allow for identifying categories of first-graders at different stages of reading development.

Six categories, described in Table 1, were identified in compliance with the theoretical model of reading development stages. For convenience, the table presents z-scores, which show by how many standard deviations each category is on average higher or lower than the sample mean. Fourteen percent of first-graders cannot read at the beginning of the 1st grade; of them, a little over 7% are unable to recognize any letters or show the beginning/end of a text, and another 7% recognize letters well but cannot read a single word. Seventeen percent of the children were able to read individual words but not sentences. Almost half of the first-graders (47.8%, category 4 in Table 1) were placed in elementary school with some reading skills, yet with a low level of comprehension: they were able to read a short story but failed to perform the reading tasks that involved text analysis. Finally, one fifth of first-graders showed a good or excellent level of reading literacy as they entered elementary school. The differences in reading development at the end of the year are not significant between categories 2 and 3, and those in reading progress are insignificant between categories 4 and 5. All the other differences are statistically significant (*t*-test, $p=0.001$)

Figures 2, 3, 4 and 5 show the percentage distribution of children from different categories of reading literacy on the basis of the following contextual variables: number of books at home, parental education, and attending preschool classes. All of the differences between categories 1 and 2 are statistically significant except those in the number of books at home; categories 2 and 3 differ statistically significantly only in the level of mother's education; all the differences are statistically significant between categories 3 and 4; categories 4 and 5 differ statistically significantly only in the level of education of both parents; and, finally, the only statistically significant difference between categories 5 and 6 is in the number of books at home (*t*-test, $p=0.05$). Otherwise speaking, it means that children's reading ability correlates with the educational resources available to them.

Children's vocabulary was measured using a series of tasks assessing the knowledge of words of various frequency. Children were asked to match the word with one of five pictures; the test includ-

Table 1. **Categories of first-graders by the level of reading development**

Category No.	Criterion: performing a task of a specific complexity level	Description of reading abilities	N	%	Average reading score, z-scores, autumn	Reading-progress, z-scores
1	Failed the letter recognition task as well as the task that required understanding the text structure	No knowledge of letters or understanding of what a text is, i. e. where it begins or ends	199	7.3	-1.99	1.16
2	Obtained a few scores in the letter recognition task but were unable to read words	Letter recognition; no reading ability	203	7.4	-0.92	0.54
3	Solved the tasks that required understanding of what reading is; recognized letters, read words, but were unable to read sentences, either because they never reached the short story level due to word reading mistakes or because they committed too many mistakes when reading the story	Ability to read individual words; inability to read sentences	477	17.4	-0.61	0.19
4	Succeeded in reading the short story but failed to read the text that required knowledge of letter-sound correspondences and deep comprehension	Ability to read words and sentences; however, building sound-letter correspondences is so effort-consuming that it inhibits simultaneous semantic processing	1,310	47.8	0.14	-0.15
5	Succeeded in reading the first text, which required simultaneous reading and comprehension, but failed to demonstrate the same level in reading the second (more complex) one	Ability to read and analyze texts at the same time	243	8.9	0.75	-0.15
6	Read both texts requiring simultaneous reading and comprehension and made it to the third (the most complex) one	Very good reading and comprehension skills	309	11.3	1.3	-0.66
Total			2,741	100		

ed 16 sets of pictures. The complexity of the tasks was gradually increased, from more frequent words to less frequent ones. One score was awarded for each correct answer. Depending on the number of scores obtained, the sample was divided into four categories (Table 2).

Phonological awareness was assessed using two tests. The first asked the children to repeat eight real words and non-words (e. g. *ta-*

Figure 2. **Number of books at home**

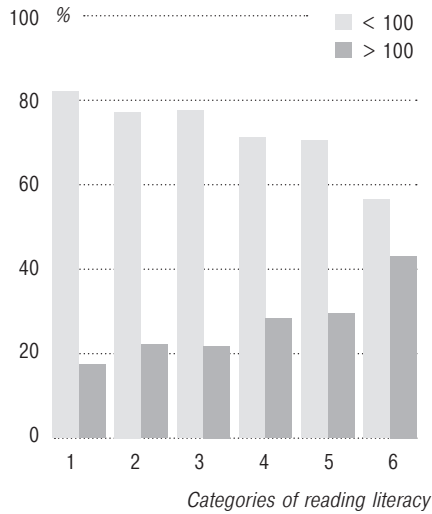


Figure 3. **Preschool classes**

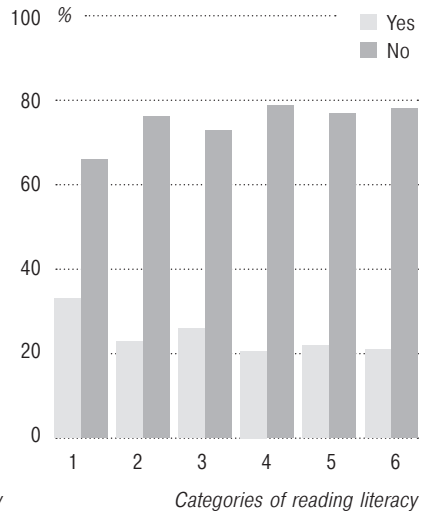


Figure 4. **Mother's education**

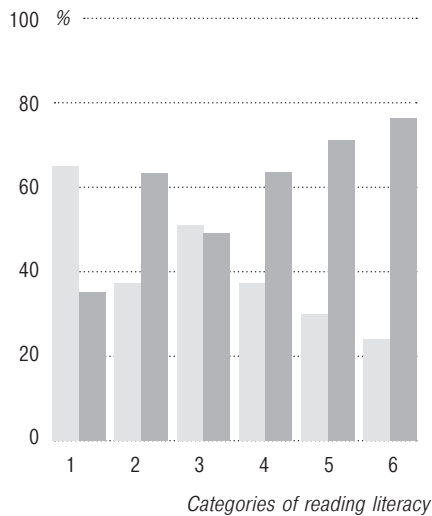
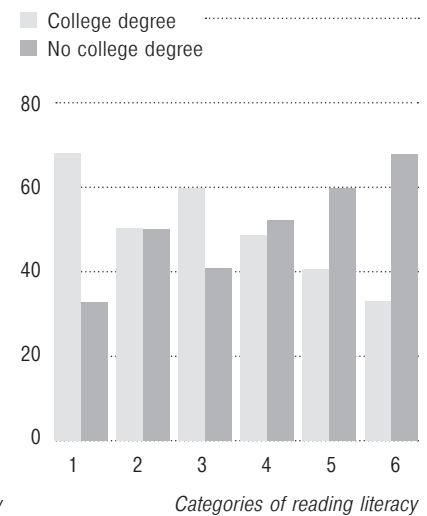


Figure 5. **Father's education**



nets(~“pance”), which had varying syllable structures. The children were supposed to concentrate on the perception of words and repeat them in exactly the same way, reproducing their sound and syllable structure. The second test asked the children to choose the word from a set of options that rhymed with the given one, following a demonstration on how it worked. The children were supposed to find rhymes for five words. Based on performance in each of the two tests (word repetition and rhyme matching), the students were divided into three

Table 2. **Student categories depending on vocabulary size**

Category No.	Criterion: rawscores	Vocabulary description	N	%	Mean vocabulary size score expressed in z-scores
1	0–3	Knowledge of the most frequent words only, such as <i>stado</i> ("herd")	330	12	-1.7
2	4–8	Knowledge of common words, such as <i>kompas</i> ("compass")	1,202	43.9	-0.4
3	9–12	Large vocabulary, knowledge of words like <i>saksofon</i> ("saxophone")	770	28.1	-0.51
4	13–16	Rich vocabulary, knowledge of words like <i>siluet</i> ("silhouette")	439	16.0	1.5
Total			2,741	100	

Table 3. **Categories of students identified based on their ability to repeat words with different syllable structures**

Category No.	Criterion: rawscores	Description of word repetition skills	N	%	Mean score for the whole phonological awareness module (z-score)
1	0–2	Ability to repeat very simple words only, such as <i>stop</i> ("stop")	197	7.2	-1.3
2	3–6	Ability to repeat words with syllable structures of medium difficulty	1,458	53.2	-0.45
3	7–8	Ability to repeat words with complex sound and syllable structures, such as <i>predpriimchivy</i> ("entrepreneurial")	1,086	39.6	0.85
Total:			2,741	100	

Table 4. **Categories of students identified based on their ability to match rhyming words**

Category No.	Criterion: rawscores	Description of rhyme matching skills	N	%	Mean score for the whole phonological awareness module (z-score)
1	0–1	Inability to find rhymes even in the easiest tasks, such as <i>dver'</i> — <i>zver'</i>	592	21.6	-0.44
2	2–3	Ability to find simple rhymes only	860	31.4	-0.15
3	4–5	Ability to find rhymes for any of the given words	1,289	47.0	0.30
Total:			2,741	100.0	

categories by the number of raw scores (Tables 3 and 4). In addition, unified scores for the whole phonological awareness module were estimated based on these two tests. The mean values of such scores are

Figure 6. **Vocabulary of children from categories with different levels of reading development**

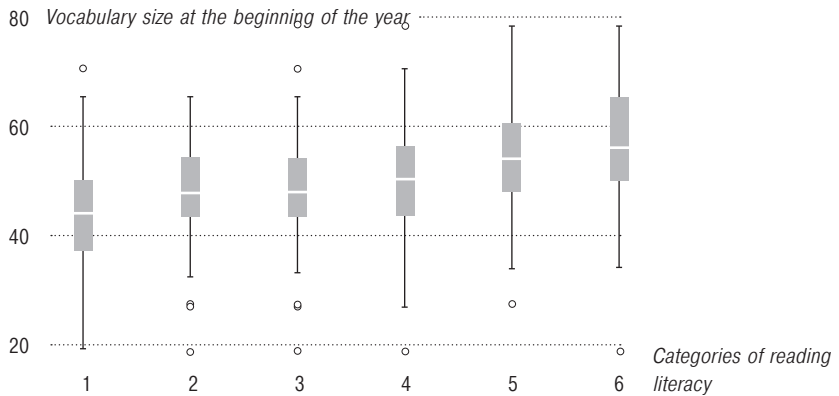
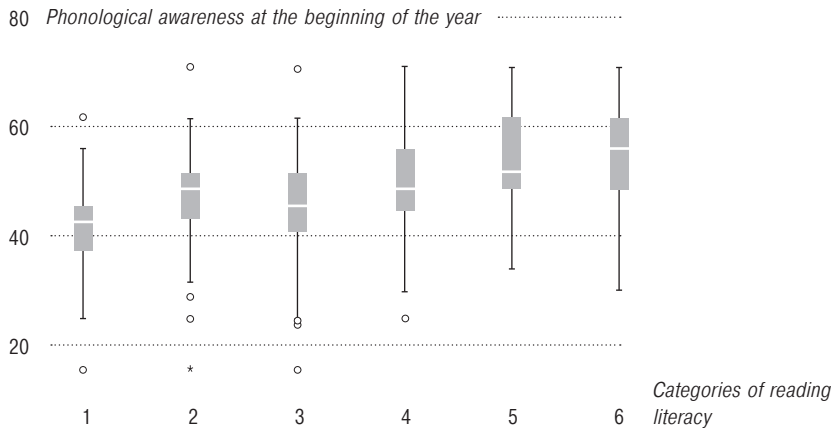


Figure 7. **Differences in phonological awareness between categories with different levels of reading development**



given as z-scores for each of the identified student categories in the far right columns of Tables 3 and 4.

The vocabulary size and phonological awareness of children correlate with their reading skills at the 0.4 level (Pearson correlation, $p=0.01$). Figures 6 and 7 contain diagrams showing the vocabulary and phonological awareness of children from categories with varying levels of reading development. In Figure 6, all of the differences are statistically significant (t -test, $p=0.01$) except those in vocabulary size between categories 2 and 3. In Figure 7, all of the differences are statistically significant (t -test, $p=0.000$). On the whole, these findings con-

Table 5. Percentage distribution of the sample into subcategories based on vocabulary size and phonological awareness

	N	Percentage of the category (%)	Percentage of the sample (%)
1A	44	22	1.6
1B	155	78	5.7
2A	90	44	3.3
2B	113	56	4.1
3A	183	38	6.7
3B	294	62	10.7
4A	715	55	26.1
4B	595	45	21.7
5A	166	68	6.1
5B	77	32	2.8
6A	249	81	9.1
6B	60	19	2.2
Total:	2,741		100

firm the hypothesis on the relation between vocabulary, phonemic development and reading skills. The children in category 2 (who know the letters but cannot read) are a curious case: while having socioeconomic characteristics (parental education, number of books at home), vocabulary size and phonological awareness similar to those of category 4 children (who can read but find comprehension challenging), these students, however, did not learn to read at preschool age. Meanwhile, they show good reading progress (see Table 1). Specific features of this student category are discussed in the final section of this article.

In order to test hypotheses 1 and 2 (limited vocabulary and low phonological awareness can inhibit reading development and progress), each of the six categories of reading literacy was divided into two parts. The "A" subcategory students had rich vocabularies and high levels of phonological awareness, while the "B" subcategory included first-graders with poor vocabularies and low levels of phonological awareness. The "B" subcategory students had been classified into category 1 or 2 by their vocabulary size and phonological awareness (word repetition and/or rhyme matching). The resulting subcategories are shown in Table 5.

Table 6 contains the main characteristics of the resulting "A" and "B" subcategories within each category. Statistically significant differences between "A" and "B" subcategories within each category are given in bold (*t*-test was used to measure the differences). Catego-

Table 6. **Characteristics of analyzed subcategories** (all indicators are given in z-scores)

Subcategory	Reading, autumn	Reading, spring	Reading progress	Communication skills, autumn	Behavior in the classroom, autumn	Parental involvement index	Socioeconomic status
1A	-1.9	-0.79	1.47	-0.29	-0.21	-0.54	-0.18
1B	-2.01	-1.31	1.07	-0.78	-0.67	-0.48	-0.54
2A	-0.91	-0.44	0.65	0.2	-0.09	-0.03	0.25
2B	-0.92	-0.7	0.38	-0.1	-0.3	-0.17	-0.15
3A	-0.48	-0.36	0.21	-0.2	-0.14	-0.04	-0.12
3B	-0.69	-0.6	0.15	-0.38	-0.26	-0.25	-0.41
4A	0.2	0.21	-0.02	0.15	0.11	0.1	0.19
4B	0.07	-0.19	-0.32	-0.08	-0.07	-0.02	-0.17
5A	0.75	0.81	-0.05	0.36	0.32	0.2	0.26
5B	0.73	0.5	-0.37	-0.01	-0.05	0.21	0.14

ry 6 was not included in the assessment, as many children from this category showed maximum possible results at the beginning of the year and had no opportunity to show their full progress in the spring assessment.

As can be seen from Table 6, low phonological awareness and small vocabulary correlate with less reading progress and lower reading assessment results at the end of the year in almost every categories. All of the children who could be defined as nonreaders at the elementary placement stage (categories 1, 2 and 3) showed similar reading progress. Category 4, the largest one (48% of the sample) originally consisted of children who were unable to read the small text in the placement test. By the end of the academic year, those with lower phonological awareness and smaller vocabularies in this subcategory had made less reading progress than others—such children account for a little more than one fifth of the sample. This subcategory also differs significantly from other children in category 4 by the level of social and emotional development, socioeconomic status, and parental involvement.

Regression analysis

Comparison of means was used in the previous section to demonstrate the significance of phonological awareness and vocabulary development in children for their reading progress in school. Whilst comparison of means reveals relations at the level of categories, the contribution of different variables at the level of individual students was assessed using regression analysis, while controlling for the variables that described contextual information on children's life (home literacy practices, cultural and social capital, language spoken at home, attendance of preschool classes, and non-cognitive development).

Table 7. **Description of auxiliary variables used in the model**

Variable	Description
1-2-3A	Nonreaders and children able to read individual words only (from categories 1, 2 and 3), who have a large vocabulary or a high level of phonological awareness
1-2-3B	Nonreaders and children able to read individual words only (from categories 1, 2 and 3), who also have a small vocabulary or low phonological awareness
4A	Readers with poor reading comprehension (who succeeded in reading the short text but were unable to perform in-depth reading comprehension tasks), yet with large vocabularies or high phonological awareness
4B	Readers with poor reading comprehension (who succeeded in reading the short text but were unable to perform in-depth reading comprehension tasks) and small vocabularies or low phonological awareness

Since the sample had a natural hierarchy (children were grouped into classes and schools), it was necessary to consider its cluster structure. Two-level regression analysis was performed using HLM software, with individual students at level one and classes at level two.

Further analysis did not include the group of 309 children who performed virtually all the comprehension tests in autumn (category 6), as the same tests were offered in spring and the children had no opportunity to demonstrate their progress on more complex tasks. With a view to better illustrate regression analysis results, two dummy variables were created; they are described in Table 7.

Children included in any of these subcategories were encoded as 1 and the others as 0, which resulted in four dummy variables. Children who reached the second reading comprehension task but did not make it to the third one because of mistakes (reading development category 5) were chosen to be the reference group. These children are pretty good readers, and even though smaller vocabularies or lower phonological awareness in this category also correlated with lower reading assessment results at the end of the year (which was verified in a separate analysis), the difference was found to be rather small (about two scores) and less statistically significant (at the level of $p=0.05$ only). For this reason, the whole of category 5 was assigned the status of the reference group without dividing it on the basis of vocabulary size and phonological awareness. Two-level regression analysis produced the following model.

$$\begin{aligned} \text{End-of-year reading assessment results} = & \gamma_{00} + \gamma_{10} \times \text{Informal home literacy practices} + \\ & + \gamma_{20} \times \text{Father's education} + \gamma_{30} \times \text{Language spoken at home} + \\ & + \gamma_{40} \times \text{Behavior in the classroom} + \gamma_{50} \times \text{Subcategory 1-2-3A} + \\ & + \gamma_{60} \times \text{Subcategory 1-2-3B} + \gamma_{70} \times \text{Subcategory 4A} + \gamma_{80} \times \text{Subcategory 4B} + u_0 + r, \end{aligned}$$

where γ_{00} is the intercept, $\gamma_{10}-\gamma_{80}$ denote coefficients of relevant variables (showing the number of end-of-year reading assessment scores

that is added in case the variable is incremented by 1), u_0 is second-level measurement error, and r is first-level measurement error.

Table 8 presents the results of a few preliminary regression models and the final one. As dummy variables 1–2–3A, 1–2–3B, 4A and 4B, which take into account both baseline reading assessment results and vocabulary and phonological awareness development, are introduced, expected results of children from categories 1–2–3 and 4 (i. e. nonreaders and “mechanical” readers) are significantly lower than those of good readers, despite the great progress of all. The differences between “A” and “B” subcategories hover around 3 scores for nonreaders as well as those who could only read simple sentences. When controlling for the personal and socioemotional development of children (adding the “behavior in the classroom” scale to model 2), this gap between the “A” and “B” subcategories is reduced to 2.5 scores for nonreaders and 2 scores for “mechanical” readers, yet the factor of “behavior in the classroom” as such contributes very little to differences in reading development, its coefficient being as low as 0.17 scores. The second factor of socioemotional development (“communication skills”) revealed no significant correlations and was removed from the final model.

Mother’s education was found to be an insignificant factor, while father’s education, on the contrary, had a high level of significance. After controlling for family characteristics (father’s education, home literacy practices) and the language spoken at home, the differences between the “A” and “B” subcategories were reduced a little more, yet persisted, now being 2 scores.

Two “sets” of parental involvement indicators were tested one by one in the model described above: the common index for various activities and two indexes assessing home literacy environment, i. e. those of formal and informal reading practices. The formal reading practices indicator was found to be insignificant and was removed from the final model. The models built using the common parental involvement index and the index of informal reading practices were virtually identical (equal index coefficients, coefficient differences of no more than 0.1–0.2 between the other variables, equal proportions of explained variation at both levels in the alternative models), so the final model included the index of informal reading practices, as it used fewer variables and its values were easier to interpret.

The variable encoding the type of school (regular or advanced, such as gymnasium or lyceum) was tested at the second level of the regression, but no significant difference was revealed, so this variable was not included in the final model either.

The intraclass correlation coefficient (ICC), showing the proportion of variation in children’s performance at the second level explained by their distribution among classes, was pretty high in autumn, when measured based only on the baseline reading assessment results (ICC = 0.13, not shown in the table), which means that children at different reading development stages were distributed unevenly among

Table 8. **Regression analysis results**

	Model 0	Model1	Model 2	Model 3
Fixed effects				
Intercept	57.91** (0.34)	65.4** (0.42)	64.8** (0.47)	62,0** (0,73)
Subcategory 1, 2, 3A		-10.4 ** (0.59)	-9.48** (0.61)	-9,11** (0,64)
Subcategory 1, 2, 3B		-13.29** (0.56)	-11.5** (0.59)	-10,83** (0,62)
Subcategory 4A		-4.34** (0.49)	-4.04** (0.51)	-3,9** (0,53)
Subcategory 4B		-7.39** (0.51)	-6.57** (0.54)	-6,14** (0,58)
Behavior in the classroom			0.13** (0.02)	0,17** (0,02)
Home literacy practices				0,32* (0,14)
Father’s education				1,24** (0,34)
Language spoken at home				2,09** (0,63)
Random effects				
u_0 (school level)	10.56	7.75	8.91	7,14
E_{ij} (student level)	55.64	40.00	38.17	37,24
ICC	0.16			
R^2 (class level)		0.27	0.16	0.32
R^2 (student level)		0.28	0.31	0.33

** significance level $p=0.01$; * significance level $p=0.05$.

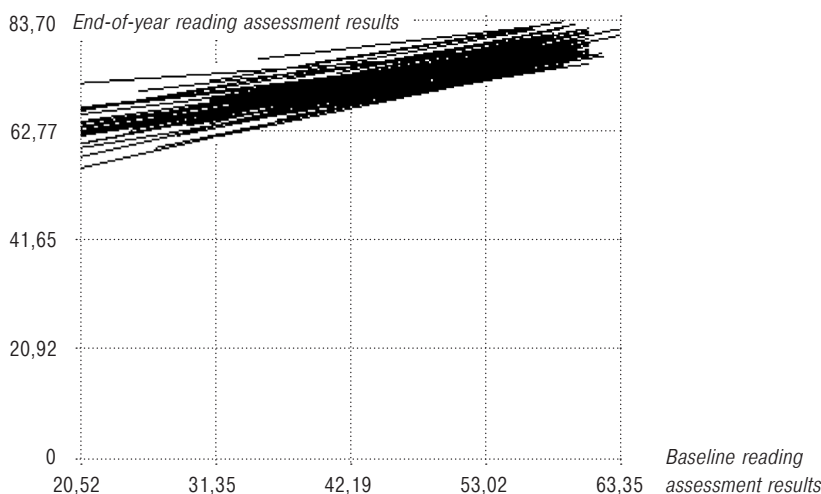
classes. The ICC at the end of the year (based on the end-of-year reading assessment) was 0.16, i. e. it increased a little more, probably because children in different classes developed at different speeds, thus widening the gap measured in autumn.

Two-level regression modeling allows for assessing differences in the regression curves across different classes. In order to evaluate the predictive power of placement reading test results, we built an additional model with a similar set of variables—except that dummy grouping variables, measuring children’s reading, vocabulary and phonological awareness test results, were replaced with individual children’s scores in baseline “reading results”, “vocabulary size” and “phonological awareness”.

$$\text{End-of-year reading assessment results} = \gamma_{00} + \gamma_{10} \times \text{Informal home literacy practices} + \gamma_{20} \times \text{Father's education} + \gamma_{30} \times \text{Language spoken at home} + \gamma_{40} \times \text{Behavior in the classroom} + \gamma_{50} \times \text{Subcategory 1–2–3A} + \gamma_{50} \times \text{Baseline reading assessment results} + \gamma_{60} \times \text{Vocabulary size} + \gamma_{70} \times \text{Phonological awareness} + u_0 + r,$$

This model was used to construct a graph, which is shown in Figure 8. There is a statistically significant difference in the location and an-

Figure 8. **Regression models for reading development in different classes**



gles of regression curves across individual classes, significance level $p=0.01$. The figure demonstrates that the curves have a slight, yet statistically significant tendency toward converging. This may be interpreted as an indicator of reading development “leveling” as classes with worse-reading children catch up to the mean performance level.

Conclusion This study aimed at evaluating the role of phonological awareness and vocabulary size in reading development at various stages. The analysis confirmed hypotheses 1 and 2 (stating that phonological awareness and vocabulary size correlate with reading progress in children with different levels of preschool reading development). Hypotheses 3 and 4 (on the role of non-cognitive skills and family factors) were confirmed partially. The contribution of socioemotional development was found to be significant only on the “behavior in the classroom” scale (which brings together the indicators of self-regulation skills), the coefficient being less than 1. Among the family factors, language spoken at home and father’s education turned out to be the most significant predictors of end-of-year reading assessment results (baseline reading results, phonological awareness, vocabulary size and non-cognitive characteristics being controlled for). Preschool home literacy practices correlate slightly yet significantly positively with end-of-year reading assessment results, informal practices such as shared book reading and discussion being more important predictors of reading development than formal (“school-style”) ones, such as learning the alphabet. Classes formed by the schools differ significantly in the level of their students’ abilities and reading progress.

Practical implications and avenues of further research

As phonological awareness and vocabulary size are critical factors of reading development not only at the beginning but at all stages of school education, teachers and parents should not be satisfied with children's basic reading skills as they enter the first grade. Possession of phonological decoding skills at the beginning of the 1st grade alone does not guarantee further success in reading development. Reading abilities at the end of the 1st grade are better predicted by a complex of three factors: reading skills as such, phonological awareness, and vocabulary size. Because the development of vocabulary and phonological awareness requires a lot of time and effort, these skills cannot be left aside when getting children ready for school.

Non-cognitive skills of children show a very weak yet statistically significant correlation with reading progress. Family characteristics—language spoken at home, father's education, and parental involvement, especially informal home literacy practices—are more reliable predictors. The low coefficient of parental involvement effect can be explained by social desirability bias (when filling out questionnaires, some parents may overestimate their involvement in their child's education), which cannot be controlled for in the framework of this study.

Quite surprisingly, the level of mother's education was found to be an insignificant factor, unlike father's education. In the course of preliminary data analysis (not included in the article due to the length restrictions), we observed that the mother's education variable was significant when it was used, together with baseline reading assessment results, to predict end-of-year reading assessment results. However, it would lose its significance as soon as the "phonological awareness" or "vocabulary size" variable was introduced. These findings may have a theoretical justification: the development of phonological awareness and vocabulary should correlate with mother's cultural capital, which means that the differences described by the level of mother's education have already been considered in the variation of other variables. Besides, allowance should be made for the composition of the sample, which consisted of megalopolis dwellers (Krasnoyarsk and Kazan), where 54% of mothers reported having college degrees. Such a high percentage of families with college-educated mothers could cause bias. Nevertheless, the high significance and great role of such a predictor as father's education became a curious finding. Perhaps, a father's college degree is a more significant indicator of both socioeconomic status and family composition. Only about half of the parents in the surveyed sample answered the question about father's education, while the others omitted the item for whatever reason. Further research is needed so as to establish whether father's education is an indicator of socioeconomic status or family composition—both hypotheses offer prospects for research.

Category 2 of reading literacy already stood out at the descriptive statistics stage of our research: it did not fit into the set of gradually increasing "steps" in Figures 2–7. These children cannot read

despite knowing the letters, yet they have better phonological awareness, larger vocabularies, a lot of books at home and college-educated parents, as compared to other nonreading first-graders. Identifying this category in a larger sample to find out why these children did not learn to read at the preschool age would be a prospective avenue of further research. It may be that their parents made a deliberate decision to take no part in preschool reading development of their children, while providing a high level of their cognitive development. It may also be that such children have some psychophysiological disorders that prevented them from learning to read despite the overall high level of school readiness. Inadequate reading instruction methods could also be a reason why children with high reading readiness were unable to develop reading skills.

Further research should include analysis of reading development dynamics at subsequent elementary school stages (e. g. upon completing the first two years) as well as of classroom practices and teacher characteristics, since a high level of variation across classes was revealed in the course of this study.

References

- Ananyev B. (1960) Analiz trudnostey v protsesse ovladeniya detmi chteniyem i pismom [Analysis of Challenges Encountered by Children Learning to Read and Write]. Ananyev B. *Psikhologiya chuvstvennogo poznaniya* [Psychology of Sensuous Cognition], Moscow: Academy of Pedagogical Sciences of the Russian Soviet Federative Socialist Republic, pp. 399–459.
- Athey I. (1983) Language Development Factors Related to Reading Development. *The Journal of Educational Research*, vol. 76, no 4, pp. 197–203.
- Bojczyk K. E., Davis A. E., Rana V. (2016) Mother–Child Interaction Quality in Shared Book Reading: Relation to Child Vocabulary and Readiness to Read. *Early Childhood Research Quarterly*, vol. 36, no 3, pp. 404–414.
- Burns M. S., Kidd J. K. (2010) Learning to Read. *International Encyclopedia of Education* (eds P. Peterson, E. Baker, B. McGaw), Oxford: Elsevier, pp. 394–400.
- Chiappe P., Siegel L. S. (2006) A Longitudinal Study of Reading Development of Canadian Children from Diverse Linguistic Backgrounds. *The Elementary School Journal*, vol. 107, no 2, pp. 135–152.
- Dittman C. K. (2016) The Impact of Early Classroom Inattention on Phonological Processing and Word-Reading Development. *Journal of Attention Disorders*, vol. 20, no 8, pp. 653–664.
- Dougherty Stahl K. A. (2011) Applying New Visions of Reading Development in Today's Classrooms. *The Reading Teacher*, vol. 65, no 1, pp. 52–56.
- Egorov T. (1953) *Ocherki psikhologii obucheniya detey chteniyu* [Sketches on the Psychology of Reading Development], Moscow: Uchpedgiz.
- Elkonin D. (1958) *Razvitiye rechi v doshkolnom vozraste* [Early Childhood Speech Development], Moscow: Academy of Pedagogical Sciences of the Russian Soviet Federative Socialist Republic.
- Elkonin D. (1962) Eksperimentalny analiz nachalnogo etapa obucheniya chteniyu [Initial Reading Development Stage Experimental Analysis]. *Voprosy psikhologii uchebnoy deyatel'nosti mladshikh shkolnikov* [Issues of the Psychology of Learning in Elementary School] (eds D. Elkonin, V. Davydov), Moscow: Academy of Pedagogical Sciences of the Russian Soviet Federative Socialist Republic, pp. 7–50.

- Elkonin D. (1976) *Kak uchit detey chitat* [How to Teach Children Read], Moscow: Znanie.
- Filicheva T., Cheveleva N., Chirkina G. (1993) *Narusheniya rechi u detey. Posobie dlya vospitateley doshkolnykh uchrezhdeniy* [Child Speech Disorders. Study Guide for Preschool Teachers], Moscow: Professionalnoe obrazovanie.
- Frith U. (1998) Editorial: Literally Changing the Brain. *Brain*, vol. 121, no 6, pp. 1051–1052.
- Goswami U. (2010) The Neuroscience of Reading. *International Encyclopedia of Education* (eds P. Peterson, E. Baker, B. McGaw), Oxford: Elsevier, pp. 394–400.
- Gvozdev A. (1961) *Voprosy izucheniya detskoy rechi* [Child Speech Research Issues], Moscow: Academy of Pedagogical Sciences of the Russian Soviet Federative Socialist Republic.
- Hattie J., Dörfler T., Artelt C. (2014) Individual Differences in Reading Development: A Review of 25 Years of Empirical Research on Matthew Effects in Reading. *Review of Educational Research*, vol. 84, no 2, pp. 203–244.
- Hilliard G. H., Wilson M. C. (1936) A Transitional Period in Reading Development. *Peabody Journal of Education*, vol. 13, no 5, pp. 226–232.
- Jeynes W. (2005) A Meta-Analysis of the Relation of Parental Involvement to Urban Elementary School Student Academic Achievement. *Urban Education*, vol. 40, no 3, pp. 237–269.
- Martin M. O., Preuschof C. (2008) Creating the TIMSS2007 Background Indices. *TIMSS2007 Technical Report* (eds J. F. Olson, M. O. Martin, I.V.S. Mullis), Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College, pp. 281–338. Available at: http://timss.bc.edu/timss2007/PDF/T07_TR_Chapter12.pdf (accessed 10 April 2017).
- Ritchey K. (2004) From Letter Names to Word Reading: The Development of Reading in Kindergarten. *Reading Research Quarterly*, vol. 39, no 4, pp. 374–376.
- Segers E., Damhuis C. M.P., Van de Sande E., Verhoeven L. (2016) Role of Executive Functioning and Home Environment in Early Reading Development. *Learning and Individual Differences*, no 49, pp. 251–259.
- Sen Y., Kuleli M. (2015) The Effect of Vocabulary Size and Vocabulary Depth on Reading in EFL Context. *Procedia—Social and Behavioral Sciences*, vol. 199, pp. 555–562.
- Senechal M., LeFevre J. (2002) Parental Involvement in the Development of Children's Reading Skill: A Five-Year Longitudinal Study. *Child Development*, vol. 73, no 2, pp. 445–460.
- Shvachkin N. (1948) Razvitiye fonemacheskogo vospriyatiya rechi v rannem vozraste [Early Development of Phonemic Awareness]. *Izvestiya APN RSFSR. Otdelenie psikhologii*, no 13, pp. 101–132.
- Tsvetkova L. (1988) *Afaziya i vosstanovitelnoe obuchenie* [Aphasia and Rehabilitation Training], Moscow: Prosveshchenie.
- Zhinkin N. (1966) *Psikhologicheskie osnovy razvitiya rechi* [Psychological Foundations of Speech Development]. *V zashchitu zhivogo slova* [In defense of the Living Word], Moscow: Prosveshchenie, pp. 5–26.
- Zhurava L., Elkonin D. (1963) K voprosu o formirovaniy fonemacheskogo vospriyatiya u detey doshkolnogo vozrasta [Towards the Issue of Phonemic Awareness Development in Preschool Children]. *Sensornoe vospitanie doshkol'nikov* [Sensory Education of Preschool Children] (eds A. Zaporozhets, A. Usova), Moscow: Academy of Pedagogical Sciences of the Russian Soviet Federative Socialist Republic, pp. 213–227.
- Ziegler J. C., Goswami U. (2005) Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory. *Psychological Bulletin*, vol. 131, no 1, pp. 3–29.