DEVELOPMENTAL CHANGES OF EARLY READING SKILLS IN SIX-YEAR-OLD POLISH CHILDREN AND GRAPHOGAME AS A COMPUTER-BASED INTERVENTION TO SUPPORT THEM

JOANNA KAMYKOWSKA[*], EWA HAMAN[**], JUHA-MATTI LATVALA[***], ULLA RICHARDSON[****] & HEIKKI LYYTINEN[****]

**Independent researcher,*
***Faculty of Psychology, University of Warsaw, Poland,*
****Niilo Mäki Institute, Finland *****
Agora Centre, University of Jyväskylä, Finland

Paper submitted: June 8, 2013
Paper resubmitted: November 20, 2013

Abstract
The aim of this randomised crossover study was to investigate the development of reading skills in children with low letter knowledge in the first year of their formal reading instruction and to assess the effectiveness of GraphoGame Polish (PL) – intensive computer game-based training in recognition of grapheme-phoneme associations.
The results show that even though children with initial poor letter knowledge finally achieved the level of the reference group in terms of letter naming, there was still a significant gap between them in reading speed of words and pseudowords.
The experiment investigating the training effectiveness did not show significant differences between the GraphoGame playing group and the group playing a maths game in respect to changes in letter knowledge. The reading skills of both groups of players increased between the assessments with a similar pace.

Kamykowska, J., Haman, E., Latvala J.-M., Richardson, U., Lyytinen, H. (2013). Developmental changes of early reading skills in six-year-old Polish children and GraphoGame as a computer-based intervention to support them. Contribution to a double special issue on Early Literacy Research in Poland, edited by Elżbieta Awramiuk and Grażyna Krasowicz-Kupis. L1-Educational Studies in Language and Literature, 13, p. 1-18. http://dx.doi.org/10.17239/L1ESLL-2013.01.05
Corresponding author: Joanna Kamykowska, email: j.piotrowska2013@gmail.com

The analysis of levels' difficulty showed that differentiating between phonologically similar letters was more difficult than between visually similar letters for learners with reading acquisition difficulties.
The findings of this study in comparison to other similar studies conducted for other languages are discussed together with its limitations and possible modifications which could improve GraphoGame efficacy for Polish poor readers.

Keywords: early literacy, letter knowledge, reading intervention, computer game-based intervention

1. INTRODUCTION

Poor letter knowledge and phonemic awareness have been shown to be sensitive indicators of future reading difficulties in children (Lyytinen et al., 2007; Hulme et al., 2012). Accuracy and stability in the perception of letters, i.e., the abstract grapheme-phoneme correspondences, is a prerequisite for achieving automatic retrieval and fluent decoding. Fast and effortless decoding of single words together with good linguistic comprehension leads to fully developed reading skills (Gough & Tunmer, 1986; Hoover & Gough, 1990).

As the ability to learn letter-sound correspondence is a critical bottleneck in reading acquisition, it is crucial to identify, support and carefully follow children who experience difficulties in this early stage of literacy skills development (Lyytinen et al., 2007).[1]

The growing need for evidence-based tools in resource-squeezed schools leads educators to search for effective remediation methods that could free teachers' resources (Davies, 1999). Computer-assisted interventions seem to be a good answer to that need because they can offer intensive individualised training with frequent repetitions (Nowicka, 2007a).

Indeed the effectiveness of computer-assisted remedial reading interventions is scientifically proven, even though the overall effect sizes are usually small, e.g. Soe and collaborators (2000) reported the effect sizes from d = 0.045 to even d = 0.762; Blok and collaborators (2002) reported the corrected overall effect size d = 0.19 (± 0.06); Slavin and collaborators (2008) reported d = 0.10 in studies assessing intervention effectiveness in adolescents.

Although the idea of using computer-based interventions for early reading acquisition was developed as early as 1970 (Atkinson, 1970), there are still not enough research-based computer applications for young children available in Poland. A review of the computer game-based programs for Polish speaking children with reading difficulties included no studies researching the actual effectiveness of these interventions (Nowicka, 2007). The question arises whether Polish speaking children could take advantage of the research-based computer program called

[1] *This is particularly important during the process of implementing changes in the curriculum. This study was conducted in the first year of implementing the new curriculum in primary schools in Poland. According to the new curriculum (Dz. U. 2012, poz. 176) children should be able to read the required reading after one year of formal reading instruction starting at the age of six (instead of being able to do it after two years of formal instruction).*

GraphoGame (Finnish name: *Ekapeli*) that was developed at the University of Jyväskylä and Niilo Mäki Institute, and subsequently spread to other countries (Lyytinen et al., 2007, 2009; Lovio et al., 2012; Hintikka et al., 2005, 2008; Huemer et al., 2008, 2010; Brem et al., 2010; Saine et al., 2010, 2011; Bach et al., 2013; Kyle et al., 2013, Nakeva von Mentzer et al., 2013). This computer game is designed to provide intensive training in rapid recognition of grapheme-phoneme associations and further reading skills, aimed at helping children to overcome 'the storage/retrieval bottleneck' that prevents reading skills development (Lyytinen, 2007).

The design of GraphoGame training in early reading skills was proven to be efficient in a number of published studies. Researchers in Finland reported that playing GraphoGame for as little as 1-2 hours improved blending skills of non-reading 6 to 7-year old kindergarteners (in comparison with a math-game playing group), and playing for 1-3 hours improved reading accuracy in children in the first grade (Lyytinen et al., 2007). Another similarly designed study revealed that GraphoGame players' gains in phonological processing, letter recognition and writing were higher than the gains of a math-game playing group. Furthermore, the intervention was proven to be beneficial for the neural basis of phonetic discrimination (Lovio et al., 2012).

Also a randomised controlled trial (RCT) longitudinal study in Finland reported gains in letter knowledge, reading accuracy, fluency, and in spelling after playing GraphoGame for 28 weeks in the first grade (compared with traditional reading remediation, the long-term gains were still evident in the third grade, Saine et al., 2010, Saine et al., 2011).

Positive results were also reported for the German language. Kindergarten GraphoGame players improved their letter knowledge, compared to number-knowledge game players but they did not improve their reading skills (Brem et al., 2010). Furthermore, a study of English readers revealed that 12 weeks of training improved reading skills, spelling, and phonological skills of GraphoGame players in comparison with the untreated control group (Kyle et al., 2013). Potentially promising results were also published for deaf and hard of hearing children in Sweden (though the study lacks a control group; Nakeva von Mentzer et al., 2013).

In light of the above findings, the Polish version of GraphoGame was designed especially for this study with the aim of supporting children with low letter knowledge during the first year of their formal reading instruction. Its structure was directly based on *Ekapeli*, however the specificity of Polish letter – sound correspondences was carefully addressed in preparing the trials for the game. Detailed information on GraphoGame PL design is presented in the Method section (2.4.1).

The present study aims to investigate the development of early reading skills (i.e. single letters identification, single word decoding speed and single pseudoword decoding speed) in children in the first year of their formal reading instruction. In particular it contrasts the reading skills of children with low letter knowledge with those of children with typical letter knowledge to assess to what extent poor readers can make up their skills during the first months of formal in-

struction. Apart from that, our study aims to determine whether GraphoGame could be an effective method to support these children in letter naming acquisition and in general reading skills.

Going in depth into the nature of children's problems with learning grapheme-phoneme associations, the study also examines the difficulty of distinguishing visually similar letter shapes vs. phonologically similar letter sounds. As visual processes play a role in the emergent literacy skills and problems in visual processing have long been believed to be partly responsible for reading problems (Krasowicz-Kupis, 2009), it is worth considering whether visual traits of similar letters (similarity in shape) make learning them difficult at the early stage of reading acquisition.

On the other hand, the ability to identify phonemes is a good predictor of children's literacy skills development (Krasowicz-Kupis, 2009) and interventions aimed at overcoming the difficulties in phonemic awareness are highly effective (the effect size for reading outcomes, d = 0.53, NRP, 2000). Thus we investigated whether poor performance in letter learning is related to problems in discriminating visual and/or phonological traits of letters.

2. METHOD

2.1 *Participants*

The participants (n = 62) were six-year-old first grade pupils (6;3 – 7;4 at the beginning of GraphoGame PL training) in the school year 2009/2010. They were sampled from three public primary schools in Warsaw, mainly the recently reformed classes[2] dedicated to six-year-old children only (the remaining 11.3% were children from mixed-classes for 6- and 7-year old pupils). This group might not be fully representative for all six-year-old children as assigning a child the reformed first grade (instead of the traditional 'zero' grade) was their parents' decision. There were fewer boys than girls in the group: 61.0% were girls (n = 38). Written permissions for all children were obtained from their parents. Children agreed orally to take part in the research before they took their screening test. All children were native Polish speakers with no mental, physical or sensory difficulties (according to their teachers' knowledge).

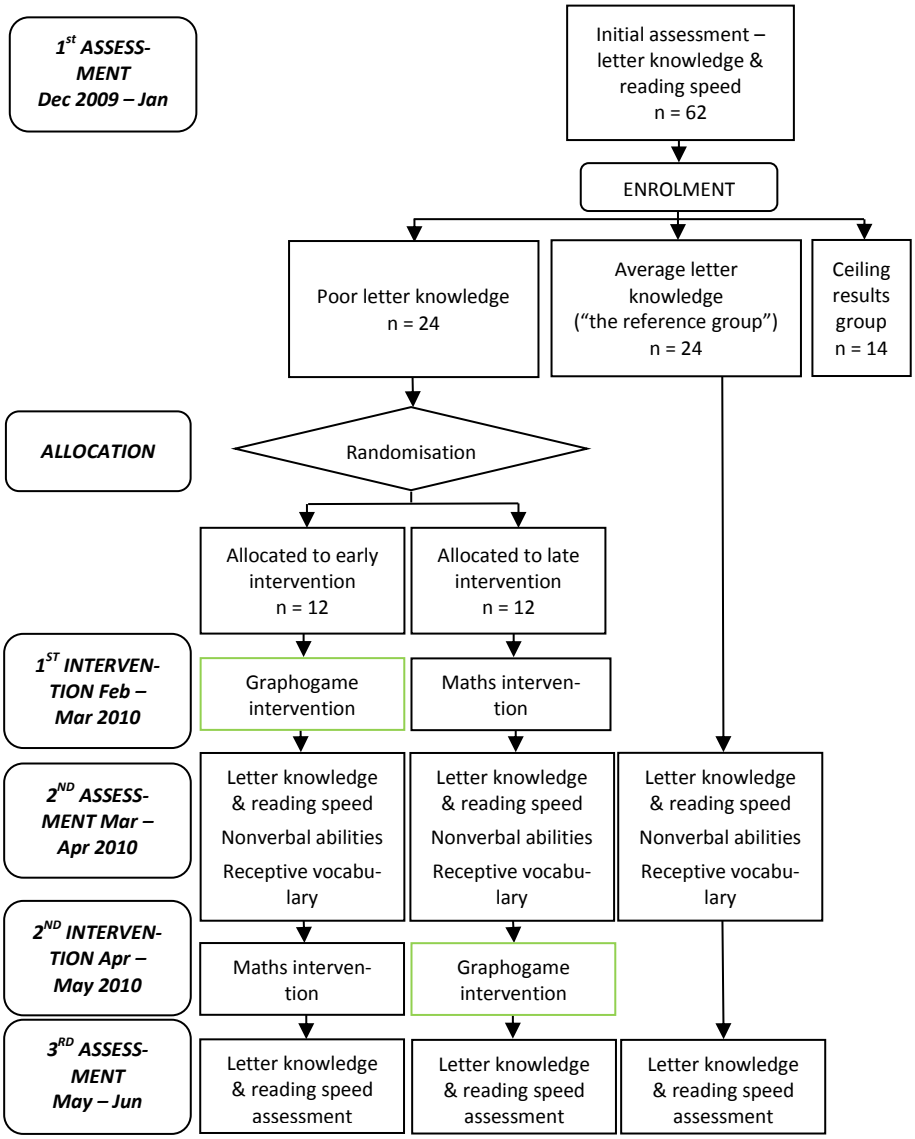
2.2 *Study design*

The training crossover design was used to assess the development of children's reading skills and the effectiveness of the GraphoGame intervention (see Figure 1). Participants were divided into groups based on their results in the screening assessment of the letter knowledge (a task from *Przesiewowy Test Dekodowania*,

[2] *According to the recent educational policy, in years 2009/2010 – 2013/2014 parents can choose for their children to start school either at the age of six or seven. From 2015/2016 all six-year-old children are going to start compulsory school education (Dz. U. 2012, poz. 176).*

Szczerbiński, Pelc-Pękala, 2008 – detailed description in 2.3 Assessment section). Children with the lowest results (n = 24, M = 36.17, SD = 10.92) formed two randomised crossover intervention groups at-risk of reading difficulties. The next 24 participants with results higher than the intervention groups (M = 55.71, SD = 3.69) formed the reference group. Children with the highest, ceiling results (n = 14, M = 62.5, SD = 0.86) were excluded from further analyses.

Figure 1 Crossover intervention study design.



The experiment was conducted between the fourth and ninth month of children's formal literacy instruction at school. There were two training phases and three assessment points. Each training phase lasted 32 days (23 working days for 'Training I' and 22 working days for 'Training II'). Because of limitations due to school timetables, holidays etc., it was impossible to balance the intervals between the assessment phases. Hence, the length of the assessment periods and breaks between them were unequal (105 days between the beginning of Test I and beginning of Test II, 51 days between the beginnings of Test II and Test III).

2.3 *Assessment*

At all three assessment points, children were asked to participate in three tasks from *Przesiewowy Test Dekodowania* [Screening Test of Decoding] (Szczerbiński, Pelc-Pękala, 2008): letter knowledge task, reading words and reading pseudowords.

Letter knowledge; measured by the number of letters named in the test; consists of 32 lower-case and 32 upper-case Polish letters (hence *min* = 0, *max* = 64 points). A child is asked to name the letters. One point is given for each correct answer. The result from the first testing session (Test I) served also as a screening task.

Word reading speed; measured by the number of words read per minute (*min* = 0, *max* = 150 points); consists of a short warm-up subtask and two 30 seconds subtasks. A child is asked to read lists of maximum 75 items as quickly as possible. Items become progressively more difficult, e.g. easy item *my* (eng. "we"), difficult item *prześcieradło* (eng. "sheet").

Pseudoword reading speed; measured by the number of pseudowords read per minute (*min* = 0, *max* = 138 points); the procedure is identical to Word reading speed task (e.g easy item *du*; difficult item *dziąbajszy).*

At the 2nd assessment point, children were also asked to take part in two additional tests*:*

Nonverbal abilities; a score achieved in *Test Matryc Ravena w wersji Kolorowej* [Raven's Coloured Progressive Matrices – Polish version] (Jaworowska, Szustrowa, 2003);

Receptive vocabulary; the number of correct answers in the experimental/unpublished version of a Picture Vocabulary Test: Comprehension (Haman

& Fronczyk, 2009)[3]. It is an individually administered test that consists of 129 colourful picture items. Each item is a page with four pictures: correct answer and semantic, phonetic or thematic distracters. The child is asked to point to the picture that represents the meaning of the word presented orally by the examiner.

2.4 *Interventions*

Two randomised crossover groups received two types of interventions: the training in letter-sound knowledge with GraphoGame PL, and the training of basic mathematical skills with GraphoGame Maths. The math-game was used as parallel nonlinguistic training during the non-GraphoGame period of the study in order to ensure similar conditions of training for both groups. This was to control for potentially important external variables influencing the reading skills development (time spent in a computer lab, teacher's or parent's beliefs about children's training, etc., as in Lyytinen et al., 2007, Brem et al., 2010). The reference group received no training at all, but as the other two groups, was influenced by the typical school reading instruction.

The training programs were made available online for the duration of this study through the GraphoGame server in Finland. Participants' playing time and progress was saved on the game server.

2.4.1 *GraphoGame PL*

GraphoGame is based on a simple concept of catching items embedded within 'balls' falling downward on the screen. Children are presented via headphones with a single sound. Their task is to click on the correct answer out of a maximum of five concurrently presented 'letter balls' (up to four letter-distracters included). Correct answers are awarded with stickers in a virtual stamp album. If the player clicks the wrong 'ball' or does not identify the target before it hits the ground, the player is still forced to choose the right answer right away, and the same target is repeated in the next trial in a colour-highlighted ball. Progression to subsequent levels is based on the achievement of 80% successfully identified phoneme-grapheme relations.

The Polish version of the game consists of 21 levels, twenty trials in each level. Its aim is to teach children 36 Polish phonemes and 39 graphemes. Some levels are designed especially to teach players to differentiate between visually similar letter shapes (3 levels: 10, 12, 14), while others are supposed to train differentiation be-

[3] *The test was under development at the time of assessment. Currently the normed version is published as Obrazkowy Test Słownikowy – Rozumienie (OTSR) [Picture Vocabulary Test – Comprehension] (Haman & Fronczyk, 2012). The normed and published version comprises 88 items in each of the two parallel forms.*

tween similar sounding phonemes (3 levels 11, 13, 15). Levels 16 to 18 include letters that are both visually and phonologically similar (examples of levels in Table 3).

2.4.2 *GraphoGame Maths*

The Polish version of GraphoGame Maths is based on similar rules. The child's task is to click on the correct item out of several concurrently presented items in squares. The provided verbal stimulus is always a number, and the visual stimuli are either pictures (e.g. flowers, dots) or numbers to be matched with the number name heard.

2.4.3 *Training procedure*

During the training periods, the researcher and a group of six volunteers interchangeably supervised players[4] on training sessions that lasted 15 minutes a day. Once a day, during a break between lessons, a group of 6 – 12 participants was taken by the adult from their classroom to the computer lab (GraphoGame playing group together with Maths group). All players were sitting in the same computer lab with the researcher, who was ready to help children and motivate them by commenting on their training progress. On the program's desktop, children could choose to play the game or watch the sticker album with his or her rewards in the form of stickers. After completing each level, the player saw the results in a form of stickers and heard the automatic approval.

3. RESULTS

3.1 *Group characteristics*

As mentioned above, the intervention groups and the reference group were formed on the basis of the initial naming letters test results (Table 1). Randomized intervention groups did not differ in any of the initially measured characteristics (age, gender, letter knowledge, reading speed). As intended, the reference group had much better reading skills (as a result of selection criteria). Interestingly, there were far more girls (n = 27) than boys (n = 11) in this group, $\chi^2(1, N = 24) = 6.00$, $p < 0.05$.

In order to assess potential group differences in cognitive and language skills (nonverbal abilities and receptive vocabulary) the results were subjected to one-way ANOVA. The groups differed in nonverbal abilities (assessed by Raven Coloured Matrices), $F(2, 44) = 3.28$, $p = 0.047$. The only *post hoc* test that showed any

[4] *A group of volunteers was recruited from the Faculty of Psychology (University of Warsaw) to provide additional care for players. The volunteers were instructed about the rules of the experiment's procedure and about their tasks. These were mainly preparing the computers and headphones, taking care of children's needs (e.g. taking a child to the toilet).*

differences was LSD. The *post hoc* test indicated that the reference group achieved higher results (*M* = 27.92, *SD* = 3.89) than both early (*M* = 24.55; *SD* = 6.44) and late intervention group (*M* = 24.50; *SD* = 3.63). The difference in receptive vocabulary was also significant, $F(2, 44) = 6.63$, $p = 0.007$. Due to inequality of variances in Levene statistic, Tamhane *post hoc* test was implemented to show that the reference group (*M* = 108.26, *SD* = 5.28) achieved slightly higher results than the late intervention group (*M* = 101.42, *SD* = 7.87). No other differences were noticed.

Table 1. Descriptions of the four groups of children formed on the basis of pretest

	Early intervention (*n* = 12)	Late intervention (*n* = 12)	Reference group (*n* = 24)	Ceiling results (*n* = 14)
	Participants' characteristics [*M* (*SD*)]			
Age (years; months)	6;7 (2.90)	6;6 (4.02)	6;5 (3.04)	6;9 (4.48)
Gender: % of boys	58.3[a]	50.0[b]	25.0[a, b]	35.7
Nonverbal abilities	24.55 (6.44)[c]	24.50 (3.63)[d]	27.92 (3.89)[c, d]	*X*
Receptive vocabulary	99.83 (11.44)	101.42 (7.87)[e]	108.26 (5.28)[e]	*X*
	Initial assessment results [*M* (*SD*)]			
Letter knowledge (*max* = 64)	34.25 (13.75)[f, h]	38.08 (7.22)[g,i]	55.71 (3.69)[f,g,j]	62.5 (0.86)[h,i,j]
Word reading speed (*max* = 150)	4.08 (4.30)[k]	5.58 (4.54)[l]	17.46 (9.42)[k, l]	*X*
Pseudoword reading speed (*max* = 138)	3.00 (3.67)[m]	4.83 (4.47)[n]	15.13 (5.94)[m, n]	*X*

Note: the pairs with same letters differ significantly at the level $p < 0.05$.

In conclusion, the intervention groups were similar in terms of mean age, gender, cognitive skills measured and reading skills in pretest. However, slight but statistically significant differences in cognitive skills (nonverbal abilities and receptive vocabulary) were observed for some comparisons between intervention and reference groups.

3.2 *Letter naming*

The influence of training on children's performance in the letter naming task and reading tasks was subjected to univariate repeated measures ANOVA with (3) groups (early intervention, late intervention, reference group) as the between-subjects factor x (3) assessments (Test I, II and III) as the within-subjects factor. Results of three children (a boy from late intervention group and two children from

the reference group) were not included, due to the lack of one of their test results (child's absence at school at one of the assessment points).

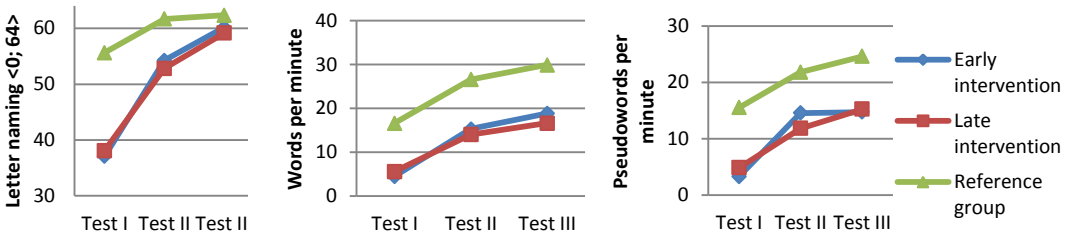


Figure 2. The results of the assessment tasks.
Intervention groups do not differ significantly at any stage.

Table 1. Pairwise comparisons of the mean differences in the assessment tasks results

	Letter knowledge			Word reading speed			Pseudo-word speed		Reading
	Test I	Test II	Test III	Test I	Test II	Test III	Test I	Test II	Test III
Early intervention	37.09[a]	54.27[b]	60.18[c, d]	4.46[f]	15.27[g]	18.81[h]	3.27[i]	14.55[j,l]	14.72[k,l]
Late intervention	38.08[a]	52.83[b]	59.17[c]	5.58[f]	14.00[g]	16.58[h]	4.83[i]	11.83[j]	15.25[k]
Reference group	55.64	61.68[e]	62.36[d, e]	16.59	26.55	29.9	15.52	21.81	24.62

Note: the pairs with same letters did not differ significantly at the level $p < 0.05$ (with Bonferroni adjustment for multiple comparisons).

Significant main effects were found for all examined factors. The result in letter knowledge depended on the group factor, $F(2, 42) = 31.63$, $p < 0.001$; $\eta^2 = 0.60$. Also the main effect for the assessment point was found to be significant (with Greenhouse-Geisser epsilon 0.81), $F(2, 84) = 176.89$, $p < 0.001$, $\eta^2 = 0.81$; same for their interaction, $F(4, 84) = 19.08$, $p < 0.001$, $\eta^2 = 0.48$.

In order to learn if the groups that played GraphoGame achieved higher results than those who played GraphoGame Maths, pairwise comparisons were implemented (Table 2). They showed no differences between the experimental groups in each assessment point. The reference group scored much better than the experimental groups in all comparisons except one. In Test III the late intervention group ($M_{III} = 59.17$, $SD_{III} = 0.83$) had only one point less than the early intervention group ($M_{III} = 60.18$, $SD_{III} = 0.87$). The latter achieved the level of the reference group ($M_{III} =$

62.36, SD_3 = 0.60). All groups were very close to the maximum result of the test (*max* = 64).

3.3 *Reading*

3.3.1 *Reading words*

The results were quite similar in terms of reading. Significant main effects were found both for groups, $F(2, 42) = 16.16$, $p < 0.001$, $\eta^2 = 0.44$, and for assessment points, $F(2, 84) = 89.90$, $p < 0.001$, $\eta^2 = 0.68$ (with Greenhouse-Geisser epsilon 0.62). Their interaction was insignificant, $F(4, 84) = 0.45$, $p > 0.05$, $\eta^2 = 0.02$.

Pairwise comparisons (Table 2) showed that each group made progress in reading between assessments. Still, the reference group outperformed both experimental groups in each assessment – the reference group read on average 1.8 times as many words as the late intervention group. There were no statistically significant differences between the intervention groups.

3.3.2 *Reading pseudowords*

In order to examine the influence of playing the game on children's performance in the reading pseudowords task, the model of univariate repeated measures ANOVA (described above) was used. The reference group has one less case than in other tests (n = 21), due to the lack of one child's result. Similarly to the word reading results, the significant main effects were found both for groups, $F(2, 41) = 15.88$, $p < 0.001$, $\eta^2 = 0.44$, and for assessment points, $F(2, 82) = 62.37$, $p < 0.001$, $\eta^2 = 0.60$ (with Greenhouse-Geisser epsilon 0.69). Their interaction was insignificant, $F(4, 82) = 1.30$, $p > 0.05$, $\eta^2 = 0.06$.

Pairwise comparisons (Table 2) showed that groups made progress in reading between tests, with the exception of the late intervention group that made no progress during GraphoGame Maths playing phase (M_{II} = 14.55, SD_{II} = 2.27, M_{III} = 14.42, SD_{III} = 2.28). Still, the reference group outperformed both experimental groups in each assessment, with no statistically significant differences between the experimental groups. The mean results of the intervention groups were much lower than the results of the reference group. The late intervention group result in Test III (M_{III} = 15.25, SD_{III} = 2.18) was still below the reference group's Test I result (M_I = 15.52, SD_I = 1.13).

3.4 *The role of cognitive abilities*

3.4.1 *Nonverbal abilities*

Nonverbal abilities were tested for their correlation with the pretest results. Nonverbal test results (n = 47) appeared to correlate significantly with the pretest re-

sult in the letter knowledge (r = 0.37, p = 0.011), word reading speed (r = 0.38, p = 0.009) and pseudoword reading speed (r = 0.32, p = 0.029).

3.4.2 *Receptive vocabulary*

Receptive vocabulary, measured by *Obrazkowy Test Słownikowy - Rozumienie* (Haman & Fronczyk, 2009), was also tested for its correlations with different pretest results. Moderate correlations were found between receptive vocabulary and all pretest results: letter naming task (r = 0.40, p = 0.005), reading words (r = 0.45, p = 0.002) and reading pseudowords (r = 0.42, p = 0.004).

3.5 *Game results*

According to data from the *graphogame.com* server, both groups spent statistically equal time on training grapheme-phoneme correspondence with GraphoGame (the exposure time for early intervention group during Training I phase, M = 57.36 min; SD = 15.75; late intervention group during Training II period, M = 48.90 min; SD = 8.34), U = 43.00, $p > 0.05$[5]. There were no gender differences in time spent on training (for boys M = 51.82 min, SD = 13.47, for girls M = 54.45 min, SD = 13.09), $t(22) = -0.47$, $p > 0.05$.

The mean number of levels completed by players was M = 17.54, SD = 4.34. Almost all participants (n = 22) reached the part of training with visually similar or phonologically similar items (levels 10 to 15), 18 of them completed this part of training. Half of the players (n = 12) managed to complete the training by playing all 21 levels, and ten of those decided to continue and started the game from the beginning again.

In order to check if certain traits of letters influenced learning progress, analysis of progress through levels was conducted. Table 3 presents the list of the most difficult game levels.

The indicator for comparisons of the training difficulty was the mean number of attempts needed to master and complete the levels containing either visually (levels 10, 12, 14) or phonologically (levels 11, 13, 15) similar items (examples in Table 3). Wilcoxon signed rank test was implemented to show that players (n = 18) needed significantly more attempts to complete the 'phonological levels' (M = 2.86, SD = 2.24) than the 'visual levels' (M = 1.73, SD = 0.85), Z = -1.96, p = 0.0498.

The game also contained three subsequent levels (16, 17, 18 – see Table 3) with such Polish items that were similar phonologically and visually at the same time. Pearson correlations of those levels with 'phonologically' and 'visually' difficult levels were investigated. It appeared that there was a strong correlation with phono-

[5] *The exposure to trained items time is much shorter than the actual training session time, as some time is usually consumed by logging in and out, and by activities related to game rewards (e.g. choosing a sticker).*

logically difficult levels (r = 0.59, p = 0.013) but no correlation with visually difficult levels (r = -0.02, p > 0.05).

Table 3. The list of game levels that were the most difficult to complete for players

Level type	Level content (Polish script | IPA)		Nr of player's attempts to complete the level [*Mdn*; *M* (*SD*)]
Basic (7)	h, ć, ź, dż, ę	/x/ /t͡ɕ/ /ʑ/ /d͡ʑ/ /ɛ̃/	3; 3.50 (3.09)
Phonological (13)	dż, cz, k, g, sz	/d͡ʒ/ /t͡ʂ/ /k/ /g/ /ʂ/	2; 4.06 (5.46)
Phonological (15)	ż, sz, dź, ć, cz	/ʒ/ /ʂ/ /d͡ʑ/ /t͡ɕ/ /t͡ʂ/	2; 3.24 (3.05)
Phonol.-visual (16)	h, n, m, w, ń	/x/ /n/ /m/ /v/ /ɲ/	2; 2.86 (3.17)
Visual (14)	g, d, p, a, h	/g/ /d/ /p/ /a/ /x/	2; 2.33 (1.68)
Phonol.-visual (18)	p, b, d, i ,j	/p/ /b/ /d/ /i/ /j/	1.5; 2.07 (1.27)

4. DISCUSSION

The present study aimed to assess the development of reading skills in six-year-old children sampled from the recently reformed first grade that follows the new curriculum. The focus of the study was on children with initial low letter knowledge. The second aim was to evaluate the effectiveness of the computer-based intervention GraphoGame to support these children (Lyytinen 2007, 2009). The game effectiveness was assessed on the basis of players' basic literacy skills development. Other questions that were addressed included assessing the role of receptive vocabulary and nonverbal abilities in reading acquisition and searching for the traits of letters that may affect establishing the mental representations of grapheme-phoneme relationships.

The study showed that the gap between children with initial poor letter knowledge and the reference group was reduced by the end of the first year of formal reading instruction. Even though both groups achieved ceiling results in the naming letters task in their ninth month of education, there was still a significant difference in their reading speed. The reference group read words and pseudowords more than 1.5 times faster than the group with the initial poor letter knowledge. The gap between these two groups remained stable despite the fact that parents and teachers knew which children were identified as poor readers (so they could offer them the additional support), and that children were supported by GraphoGame intervention.

The experiment which assessed the effectiveness of GraphoGame did not show any differences between children who played GraphoGame and those who played GraphoGame Maths in naming letters and reading performances at any assessment

point. This indicates that playing GraphoGame PL had no more influence on literacy skills than playing a game which trained skills not related to letter – sound correspondence. The results of both groups increased in each assessment, though. The only exception was the late intervention group's pseudoword reading result that did not increase after training with the Maths game.

In the present study, the assessment of training effectiveness was strengthened by the analysis of the difficulty of differentiating phonologically and visually similar letters or letter sounds. A conclusion from this analysis is that the phonological traits of letters are significantly more difficult to differentiate than the visual traits for poor readers in this study. This result seems to be in accordance with the general view of different studies describing reading acquisition difficulties in relation to phonological skills (e.g. studies related to the phonological deficits hypothesis, Snowling, 2000).

The present results diverge from the majority of results from Lyytinen and collaborators (e.g. 2007, 2009; also Saine et al., 2011; Lovio et al., 2012; Kyle et al., 2013) that showed significant improvements in reading-related skills in GraphoGame players in languages such as Finnish, German and English. The attempt to explain the results of the present study, in particular the lack of expected effect of the intervention involves analysing its limitations.

The effect of "learning letters despite the type of training" might be explained by the role of school literacy instruction. Polish children are taught in the analytic-synthetic phonics approach (Jurek, 2012) that is similar to the GraphoGame approach. The main rule of both is presenting a child with a letter and its sound and linking them together. It seems that the training provided at school was sufficient to build and strengthen mental representations of single grapheme-phoneme connections (even in children with initial letter learning difficulties) but it did not help to teach children to read fluently. In a similar way GraphoGame PL provided training in letter knowledge but it did not provide training in rapid retrieval needed for fluent decoding. This might suggest that letter knowledge training is needed at the very beginning of a school year instead and is a prelude to helping children acquire reading fluency by providing more complex training with syllables and words. Such training seems to be needed as the gap between the average readers and poor readers was still significant at the end of the first grade. The intervention groups only reached the level of initial reading speed (Test I) of the reference group.

Apart from that, the reason for the lack of specific training effects (improvement in letter knowledge and in reading speed) might be connected to different aspects of the research design. Firstly, the training was short in duration. Players spent on average only 50 minutes being exposed to training items. Secondly, the training was not strictly regular due to school timetable limitations. There were some breaks (1-3 days) between the training sessions, which were against the initial plan of the study design. Short training time and its irregularities might have prevented building the stable representations of the trained grapheme-phoneme connections. Thirdly, the game design itself might not have been motivating

enough to focus children's attention on trained items for the proper amount of time. According to the researcher's observations, there were sessions, when players were much more interested in their rewards than the training itself. Moreover, the game construction might have not motivated children to respond quickly. The lack of attention and slow responding might have hindered the crucial process of learning the rapid retrieval of grapheme-phoneme connections needed for future fluent reading. According to Hintikka and collaborators (2008), changes in the game construction may produce better results in players' reading fluency through increasing their motivation to respond quickly. Those might include the gradual increasing of the speed of items' presentation, training one-to-one with a tutor, who snatches the point when the player's response is too slow, or giving additional points for very quick response.

Changes in the game content might include adding more levels to train phonological differentiations of letter sounds (instead of visual differentiations). Also enriching the game with training of syllables and words is highly recommended. Finally, the training should be very intensive and GraphoGame PL should be introduced at the very beginning of formal education to fulfil its aim of stimulating the initial process of building stable mental representations of grapheme-phoneme correspondence.

5. CONCLUSIONS

First, we conclude that the first 9 months of formal reading instruction based on the new curriculum was enough to learn the letters even for children with initial low letter knowledge. On the other hand, children with low initial letter knowledge did not manage to close the gap between them and the reference group in terms of reading speed; and therefore the intensive training in reading speed should be provided for these children in the second grade.

Secondly, the RCT experiment showed that GraphoGame training designed to support for children with poor letter knowledge was no more effective than playing non-linguistic game. As this result differs from the other results of GraphoGame interventions (Lyytinen et al., 2007, 2009; also Saine et al., 2011; Lovio et al., 2012; Kyle et al., 2013), it is possible that the mode of delivery for this particular training was not sufficient.

Thirdly, the support offered to children in the present study was limited to letter-phoneme correspondence identification and thus was not enough to close the gap between them and the reference group in terms of reading speed. Hence it is important to carefully plan the extra support for children at-risk of reading difficulties in the first grade directed at specific deficits they encounter. Unfortunately, the very strict school timetables in the Polish school curriculum left little time for extra support in the form of regular training in the computer lab.

ACKNOWLEDGEMENTS

We thank the Graphogame development team at the University of Jyväskylä and all Polish volunteers involved in preparing the game and the study, especially Joanna Osesik, Krzysztof Osesik, Marian Kamykowski, Radio Józef Station, doc. dr Maria Przybysz-Piwko and students from Students' Association of Logopaedics at the University of Warsaw. We tender our thanks to Laura Lee for polishing the language. We also thank the students, families, and schools that participated in the study. The project presented here was partially supported by STSM grant awarded to the first author within the COST Action A33 in 2009[6] and by BST grant awarded to the second author by Faculty of Psychology, University of Warsaw. Results presented here were previously described in the first author's (nee: Piotrowska) unpublished MA thesis, prepared under the supervision of the second author at the Faculty of Psychology, University of Warsaw (2010) and presented in Polish as a poster during *XIX Ogólnopolska Konferencja Psychologii Rozwojowej* (XIX National Conference on Developmental Psychology) in Łódź, June 2010.

REFERENCE LIST

Atkinson, R.C. (1970). Instruction in initial reading under computer control: *The Stanford Project* (Report No. 158). Washington, D. C: National Science Foundation.

Bach, S., Richardson, U., Brandeis, D., Martin, E., Brem, S. (2013). Print-specific multimodal brain activation in kindergarten improves prediction of reading skills in second grade. *NeuroImage, 82*, 605-615. http://dx.doi.org/10.1016/j.neuroimage.2013.05.062

Blok, H., Oostdam, R., Otter, M. E., & Overmaat, M. (2002). Computer-assisted instruction in support of beginning reading instruction: A review. *Review of Educational Research, 72 (1),* 101–130. http://dx.doi.org/10.3102/00346543072001101

Brem, S., Bach, S., Kucian, K., Guttorm, T. K., Martin, E., Lyytinen, H., Brandeis, D. & Richardson, U. (2010). Brain sensitivity to print emerges when children learn letter-speech sound correspondences. *PNAS, 107 (17)*, 7939-7944. http://dx.doi.org/10.1073/pnas.0904402107

Davies, P. (1999). What is Evidence-based Education? *British Journal of Educational Studies*, *47 (2)*, 108-121. http://dx.doi.org/10.1111/1467-8527.00106

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7,* 6–10. http://dx.doi.org/10.1177/074193258600700104

Haman, E. & Fronczyk, K. (2009). *Obrazkowy Test Słownikowy - Rozumienie (OTSR)* [Picture Vocabulary Test - Comprehension], experimental version 129 items. Unpublished material. Faculty of Psychology University of Warsaw.

Haman, E. & Fronczyk, K. (2012). *Obrazkowy Test Słownikowy - Rozumienie (OTSR)* [Picture Vocabulary Test - Comprehension]. Gdańsk: Pracownia Testów Psychologicznych i Pedagogicznych.

Hintikka, S., Aro, M., & Lyytinen, H. (2005). Computerized training of the correspondences between phonological and orthographic units. *Written Language & Literacy, 8,* 155 – 178. http://dx.doi.org/10.1075/wll.8.2.07hin

Hintikka, S., Landerl, K., Aro, M. & Lyytinen, H. (2008). Training reading fluency: is it important to practice reading aloud and is generalization possible? *Annals of Dyslexia, 58*, 59–79. http://dx.doi.org/10.1007/s11881-008-0012-7

Hoover, W. A. and Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal, 2,* 127-160. http://dx.doi.org/10.1007/BF00401799

[6] *http://www.cost.eu/domains_actions/isch/Actions/A33*

Huemer, S., Landerl, K., Aro, M., & Lyytinen, H. (2008). Training reading fluency among poor readers of German: many ways to the goal. *Annals of Dyslexia, 58,* 115 – 137. http://dx.doi.org/10.1007/s11881-008-0017-2

Huemer, S., Aro, M., Landerl, K. & Lyytinen, H. (2010). Repeated Reading of Syllables Among Finnish-Speaking Children With Poor Reading Skills. *Scientific Studies of Reading, 14 (4)*, 317-340. http://dx.doi.org/10.1080/10888430903150659

Hulme, C., Bowyer-Crane, C., Carroll, J., Duff, F. & Snowling, M. (2012). The Causal Role of Phoneme Awareness and Letter-Sound Knowledge in Learning to Read. Combining Intervention Studies With Mediation Analyses. *Psychological Science, 23 (6),* 572-577. http://dx.doi.org/10.1177/0956797611435921

Jaworowska A. & Szustrowa T., (2003). *Podręcznik do Testu Matryc Ravena w wersji Kolorowej* [Raven's Coloured Progressive Matrices – Polish version]. Warszawa: Pracownia Testów Psychologicznych PTP.

Jurek, A. (2012). *Metody nauki czytania i pisania z perspektywy trudności uczniów* [Methods for teaching reading and writing from students' difficulties perspective]. Gdańsk: Harmonia Universalis.

Krasowicz-Kupis, G. (2009). *Psychologia dysleksji* [Psychology of dyslexia]. Warszawa: Wydawnictwo Naukowe PWN.

Kyle, F., Kujala, J., Richardon, U. & Lyytinen, H. (2013). Assessing the effectiveness of two theoretically motivated computer-assisted reading interventions in the United Kingdom: GG Rime and GG Phoneme. *Reading Research Quarterly, 48 (1)*, 61-76.

Lovio, R., Halttunen, A., Lyytinen, H., Näätänen, R. & Kujala, T. (2012). Reading skill and neural processing accuracy improvement after a 3-hour intervention in pre-schoolers with difficulties in reading-related skills. *Brain research, 1448,* 42-55. http://dx.doi.org/10.1016/j.brainres.2012.01.071

Lyytinen, H., Erskine, J., Kujala, J., Ojanen, E. & Richardson, U. (2009). In search of a science-based application: A learning tool for reading acquisition. *Scandinavian Journal of Psychology, 50*, 668–675. http://dx.doi.org/10.1111/j.1467-9450.2009.00791.x

Lyytinen, H., Ronimus, M., Alanko, A., Poikkeus, A.-M. & Taanila, M. (2007). Early identification of dyslexia and the use of computer game-based practice to support reading acquisition. *Nordic Psychology*, *59,* 109-126. http://dx.doi.org/10.1027/1901-2276.59.2.109

Nakeva von Mentzer, C., Lyxell, B., Sahlen, B., Wass, M., Lindgren, M., Ors, M., Kallioinen, P., Uhlen, I. (2013). Computer-assisted training of phoneme–grapheme correspondence for children who are deaf and hard of hearing: Effects on phonological processing skills. *International Journal of Pediatric Otorhinolaryngology. http://dx.doi.org/10.1016/j.ijporl.2013.10.007*

National Reading Panel (2000). *Report of the National Reading Panel. Teaching children to read.* Washington, D.C.: U.S. Government Printing Office.

Nowicka, E. (2007a). Zastosowanie technologii informacyjnych w terapii pedagogicznej [Using information technology in remedial classes]. In: M. Kostka-Szymańska & G. Krasowicz-Kupis (red.), *Dysleksja – problem znany czy nieznany?* (p. 309-315). Lublin: Wydawnictwo UMCS. [Dyslexia – is problem well known or not?].

Nowicka, E. (2007b). Zawartość edukacyjnych programów komputerowych wspierających zajęcia korekcyjno-kompensacyjne [The content of educational computer programs used in remedial classes]. In M. Kostka-Szymańska & G. Krasowicz-Kupis (Eds.), *Dysleksja – problem znany czy nieznany?* (pp. 309-315). Lublin: Wydawnictwo UMCS.

Saine, N., Lerkkanen, M.-K., Ahonen, T., Tolvanen, A. & Lyytinen, H. (2010). Predicting word-level reading fluency outcomes in three contrastive groups: Remedial and computer assisted-remedial reading intervention, and mainstream instruction. *Learning and Individual Differences, 20*, 402-414. http://dx.doi.org/10.1016/j.lindif.2010.06.004

Saine, N., Lerkkanen, M.-K., Ahonen, T., Tolvanen, A. & Lyytinen, H. (2011). Computer-assisted remedial reading intervention for school beginners at risk for reading disability. *Child Development, 82 (3)*, 1013–1028. http://dx.doi.org/10.1111/j.1467-8624.2011.01580.x

Slavin, R. E., Cheung, A., Groff, C., & Lake, C. (2008). Effective reading programs for middle and high schools: a best–evidence synthesis. *Reading Research Quarterly, 43*, 290–322. http://dx.doi.org/10.1598/RRQ.43.3.4

Snowling, M. J. (2000). *Dyslexia*. Oxford: Blackwell Publishers.

Soe, K., Koki, S., & Chang, J. M. (2000). *Effect of computer assisted instruction (CAI) on reading achievement: A metaanalysis.* Honolulu, HI: Pacific Resources for Education and Learning.

Szczerbiński, M. & Pelc-Pękala, O. (2008). *Przesiewowy Test Dekodowania* [Screening Test of Decoding]. Kraków: Autorzy.