

Relations Between Reading, Vocabulary and Phonological Awareness in low-Income Children¹

Heloisa Helena Motta Bandini²
*Universidade Estadual de Ciências da
Saúde de Alagoas, Maceió-AL, Brazil*

Carmen Silvia Motta Bandini
*Universidade Estadual de Ciências da
Saúde de Alagoas, Maceió-AL, Brazil*

Adhemar Ranciaro Neto
*Universidade Federal de Alagoas,
Maceió-AL, Brazil*

Abstract: Phonological awareness, intelligence, vocabulary and socioeconomic status (SES) have been considered important variables in the acquisition of reading. Nevertheless, attention deficit is associated to low reading performance. The purpose of this work was to investigate correlations among phonological awareness, vocabulary, intelligence, focused attention and reading skills in children with low SES exposed to a non-effective teaching environment. This study included 111 children belonging to socioeconomic classes D and E who were enrolled at a state school. Results pointed to a positive correlation among vocabulary, intelligence, phonological awareness and reading scores. There was no significant correlation between attention measures and reading ability. The relations among the variables continued, even in a population with low SES, although bad teaching environment acts as a limiting factor for the students' development of their reading ability.

Keywords: reading, vocabulary, phonological awareness, attention, socioeconomic status

Relações entre Leitura, Consciência Fonológica e Vocabulário em Crianças de Baixa Renda

Resumo: Consciência fonológica, inteligência, vocabulário e nível socioeconômico vêm sendo consideradas variáveis importantes para a aquisição da leitura. Ao mesmo tempo, déficits de atenção vêm sendo relacionados ao mau desempenho em leitura. O objetivo foi investigar as correlações entre as habilidades de consciência fonológica, vocabulário, inteligência, atenção concentrada e leitura em crianças de baixo nível socioeconômico expostas a ambiente precário de ensino. Participaram 111 crianças de uma escola pública, pertencentes às classes socioeconômicas D e E. Os resultados indicaram correlações positivas entre as habilidades de vocabulário, inteligência, consciência fonológica e leitura. A atenção não contribuiu para a explicação da competência em leitura. Verificou-se que as relações entre as variáveis em questão são mantidas, mesmo em populações de baixo nível socioeconômico, entretanto, o ambiente precário de ensino pode ser um fator limitante para a expansão da capacidade de leitura dos estudantes.

Palavras-chave: leitura, vocabulário, consciência fonológica, atenção, nível socioeconômico

Relaciones entre lectura, conciencia fonológica y vocabulario en niños de bajos ingresos

Resumen: Conciencia fonológica, inteligencia, vocabulario y estatus socioeconómico se han considerado variables importantes para el aprendizaje de la lectura. Al mismo tiempo, el déficit de atención se ha relacionado con los malos resultados de lectura. El objetivo de este estudio fue investigar la correlación entre conciencia fonológica, vocabulario, inteligencia, atención enfocada y lectura en niños de bajo nivel socioeconómico expuestos a malas condiciones de enseñanza. En este estudio participaron 111 niños que pertenecen a las clases socioeconómicas D y E y fueron inscritos en una escuela pública. Los resultados indican correlaciones positivas entre las puntuaciones de vocabulario, inteligencia, conciencia fonológica y lectura. No fue encontrada correlación significativa entre la medida de atención y la aptitud para la lectura. Se encontró que la relación entre las variables en cuestión se mantiene incluso en poblaciones de bajos ingresos, sin embargo, la pobreza del entorno de enseñanza actúa como factor limitante para el desarrollo expansión de la capacidad de lectura de los estudiantes.

Palabras clave: lectura, vocabulario, conciencia fonológica, atención, nivel socioeconómico

Learning to read is a gradual process, for which a series of prerequisites have been described in the literature. Skills like phonological awareness (Justi & Roazzi, 2012; Lundberg, Frost, & Petersen, 1988), intelligence (T.P. Alloway & R.G. Alloway, 2010) and vocabulary (T.P.

Alloway & R.G. Alloway, 2010; Bandini, Santos, & de Souza, 2013; Berendes, Schnitzler, Willmes, & Huber, 2010; Engel, Santos, & Gathercole, 2008; Santos & Befi-Lopes, 2012) have been highlighted as fundamental prerequisites to acquire reading proficiency.

Despite the importance of these skills for reading acquisition, Justi and Roazzi (2012) affirm that few studies have considered all of these simultaneously. The authors investigated the relations among phonological working memory, phonological awareness, rapid serial naming (vocabulary) and morphological awareness and the

¹Support: São Paulo Research Foundation (FAPESP, Grant # No. 2008/57705-8) and National Council of Scientific and Technological Development (CNPq, Grant # No. 573972/2008-7).

²Correspondence address: Heloisa Helena Motta Bandini, Universidade Estadual de Ciências da Saúde de Alagoas, Centro Especializado em Reabilitação, Rua Jorge de Lima, 113 Trapiche da Barra, CEP – 57010-300 Maceió-AL, Brazil. heloisabandini@gmail.com

development of reading and writing. The study included 94 children in the third and fifth year of elementary education of elementary school at private schools. The results of the hierarchical regression analyses indicated that phonological processing and vocabulary contributed to the reading skills.

In view of the lack of studies that take into account phonological reading skills, vocabulary, intelligence and reading simultaneously, the objective in this study was to investigate the correlations among phonological awareness, vocabulary, intelligence, focused attention and reading in children with low SES exposed to a non-effective teaching environment. Four distinctive aspects can be considered in this study in relation to the literature in the area: the type of population that participated, the expansion of the phonological awareness tests, the inclusion of the focused attention variable and the simultaneous use of the variables in question.

What the research sample is concerned, in this study, participants from a low-income school were selected where teaching has shown non-effective outcomes, as appointed by the results on the official Portuguese language test (*Prova Brasil para Língua Portuguesa*), provided by the National Institute for Educational Studies and Research Anísio Teixeira [INEP] (2014). In this assessment, the school scored 147 points, which ranked it below the state average for Brazilian public schools (INEP, 2014). Studying these skills in a population with those characteristics can be relevant as the socioeconomic level has been appointed as an explanatory or control variable for the analysis of language acquisition (Engel et al., 2008). That is so because the variation in the socioeconomic level seems to be widely related to the individuals' possible exposure to teaching/training situations, which result in the acquisition and development of language, in line with Engel et al. (2008) and Macedo, Andreucci and Montelli (2004).

A second difference is the expanded use of the phonological awareness variable. Defined as the skill to segment and manipulate speech sounds to compose new words (A.G.S. Capovilla & F.C. Capovilla, 1998), phonological awareness can be subdivided in different skills, which involve competencies related to the discrimination, segmentation and manipulation of syllables and phonemes. Nevertheless, many studies assess only some of these competencies, such as (Justi and Roazzi, 2012), who assessed only two phonological awareness skills, i.e. phonemic subtraction and spoonerism. In this study, we chose to assess the ten skills tested by A.G.S. Capovilla and F.C. Capovilla (1998).

Finally, in this study, the focused attention variable was included through the use of the D2 Test (Brickenkamp, 2000). Attention deficits have been related with bad reading performance, as appointed by Cain and Bignell (2014), Franceschini, Gori, Ruffino, Pedrolli and Facoetti (2013) and Kendeou, van den Broek, Helder and Karlsson (2014). Those authors affirm that, when a child presents an attention deficit, his/her performance in the reading of regular and irregular words is significantly inferior to the performance of children without these difficulties.

Method

Participants

The study participants were 63 children from the third-year of elementary education, between seven years and eight months and ten year and two months of age (mean age eight year and six months), being 30 girls and 33 boys, and 48 children from the fifth year of elementary education, between nine years and nine months and 16 years of age (mean age 12 years and two months), being 25 girls and 23 boys. All children were registered in the same public school on the outskirts of Maceió-AL. The children belonged to socioeconomic classes D and E, according to the Brazilian Association of Research Companies [ABEP] (2012), with an assumed family income between R\$477.00 (four hundred and seventy-seven *Reais*) and R\$714.00 (seven hundred and fourteen *Reais*).

Instruments

In this study, the following instruments were used:

Raven's Progressive Matrices, adapted to Portuguese by Angelini, Alves, Custódio, W.F. Duarte and J.L.M. Duarte (1999). It is intended to assess children's intellectual development. The instrument consists of three series of 12 items, arranged in increasing order of difficulty. In each item, a colored matrix is presented with one part missing. Under this matrix, six alternative matrices are displayed to complement the missing piece in the model matrix, only one of which completes it correctly. Then, the child is asked to choose one option for each item and the option chosen is marked on a specific protocol. The detailed description of the tasks, the instructions for their execution and the psychometric properties of this test are available in Angelini et al. (1999).

Phonological Awareness Test by Oral Production. Developed by A.G.S. Capovilla and F.C. Capovilla (1998), it is intended to verify children's skill at identifying and manipulating speech sounds. The test consists of ten subtests, each of which contains six items: two training and four assessment items. The subtests assess the suprasegmental skills (alliteration and rhyme) and synthesis, segmentation, manipulation and transposition skills at the syllabic and phonemic levels. For each correct answer, one point is attributed, up to the maximum of 40 points. The detailed description of the tasks, test instructions and psychometric properties of this test are available in A.G.S. Capovilla and F.C. Capovilla (1998).

Test of Silent Word Reading Fluency (TSWRF). Developed by F.C. Capovilla and A.G.S. Capovilla (2010), it assesses the reading development level. The test consists of sheets with figures, accompanied by a written element (which can be a word written correctly, a word written incorrectly, a pseudoword or a word spelled correctly but which does not correspond to the figure it is associated with). The student's task is to judge whether the word printed below the figure correctly represents its. The training items are assessed with the help of the evaluator, while the student performs the test items without any help. The detailed description of the tasks and the psychometric properties of this task are available in

F.C. Capovilla and A.G.S. Capovilla (2010).

Vocabulary Test by Figures (TVFUSP). Developed by F.C. Capovilla (2011), it is available in two versions. In this study, the 92-item version was used, which is aimed at assessing children's receptive vocabulary. To apply the test, the student receives a notebook with 92 figures, grouped in lines of four and is instructed to mark one figure on each line, corresponding to the word the experimenter dictated. The test result is calculated by the amount of figures marked correctly. The instructions for its execution and the psychometric properties of this test are available in F.C. Capovilla (2011).

D2 Test. Adapted to Portuguese by Brickenkamp (2000), it is intended to assess the focused attention of children over nine years of age. In the test, the participant is instructed to mark the letter 'd' with a line only when accompanied by two lines (above, below, or one on top and the other below). The student completes the training sheet together with the instructor, and the instructor corrects the answers in the participant's presence. The test sheet consists of 14 lines and the participant has 20 seconds for each line. Thus, the participant starts the activity by line 1. When the time is reached, the instructor informs the participant and indicates that he needs to move on to the next line, even if line 1 has not been fully completed yet. The participant who is able to find the 'd' with two lines in the shortest time possible with the lowest error rate scores best. The instructions for its execution and the psychometric properties of this test are available in Brickenkamp (2000).

Procedure

Data collection. The individual tests were applied in classrooms booked for this purpose. The children were invited to participate in the activities as if they would engage in play or a game and sat in front of or beside the experimenter. The group tests (TSWRF and TVFUSP) were applied in the classroom, during sessions of approximately 40 minutes. A team previously trained for this purpose applied all tests. Psychology professionals, regularly registered in the Alagoas Regional Psychology Council, applied the Raven and D2 tests, exclusive to psychologists. The data collected took about 90 days, between September and November 2012.

Data analysis. The following statistical procedures were employed: Kolmogorov-Smirnov and Shapiro-Wilk normality tests; transformation of variables; Pearson's correlation; Mann-Whitney test for intergroup comparison of the research variables; principal component analysis to reduce the number of variables in a measure; quantile regression and hierarchical regression. Calculations were executed in STATA - Data Analysis and Statistical Software v. 12.

Ethical Considerations

The research procedures received approval from the Ethics Committee for Research involving Human Beings at *Universidade Estadual de Ciências da Saúde de Alagoas* (protocol #1771) and due care was taken in their execution to guarantee the participants' wellbeing and avoid their exposure to risks of any kind.

Results

The descriptive statistics of the study sample (third and fifth-year students) were displayed in Table 1. In both groups, the frequencies of some variables differ from the Gaussian distribution. Hair, Black, Babin and Anderson (2010) consider that the population can present normal distribution in a specific variables, which is the case when Asymmetry/Standard-Error and Kurtosis/Standard-Error ratios between -1.96 and 1.96 are observed in the sample. The same authors appoint that hypothesis tests like the modified Kolmogorov-Smirnov and the Shapiro-Wilk are good alternatives to verify the normal distribution and recommend a 5% significance level as a criterion to reject the null hypothesis.

For all hypothesis tests undertaken in this study, significance was set at 5%, except for those situations in which the tolerance of type I errors was properly expressed. For the second series, strong evidence exists for a non-Gaussian curve in the variables representing the application of Raven's Progressive Matrices (RAVEN) (Kolmogorov-Smirnov, $p = .007$ and Shapiro-Wilk $p = .002$), representing the results of the Phonological Awareness Test (PA) (Kolmogorov-Smirnov and Shapiro-Wilk p -value inferior to .01).

What the fifth year is concerned, the variables that showed strong evidence (5% significance) that the population distribution differed from the Gaussian type were age (AGE) (Kolmogorov-Smirnov, $p = .109$ and Shapiro-Wilk, $p = .016$), standardized coefficient obtained by applying the Test of Silent Word Reading Fluency (TSWRF) (Kolmogorov-Smirnov, $p = .100$ and Shapiro-Wilk, $p = .000$), gross score on the same test (TSWRF_b) (Kolmogorov-Smirnov, $p = .001$ and Shapiro-Wilk, $p = .001$), score on the Vocabulary Test by Figures USP (VOC) (Kolmogorov-Smirnov $p = .006$ and Shapiro-Wilk $p = .002$), Gross Result of D2 (GR) (Kolmogorov-Smirnov $p = .021$ and Shapiro-Wilk $p = .000$), Percentage error on the same test (E%) (Kolmogorov-Smirnov $p = .098$ and Shapiro-Wilk $p = .021$) and Oscillation Range, also on the D2 test (OR) (Kolmogorov-Smirnov $p = .093$ e Shapiro-Wilk $p = .023$).

When comparing the samples of both groups, it was observed that the mean TSWRF_b for the third-year students (42.09) was lower than that of the fifth-year students (52.31). The same fact repeated itself for the variables VOC (mean third year = 56.33; mean fifth year = 66.60) and PA (mean third year = 11.59; mean fifth year = 18.54). The RAVEN variables and the D2 components cannot be compared as the D2 test can only be applied to children over nine years of age.

To check for differences between the means in the two groups of students, Mann-Whitney's non-parametrical statistical procedure was applied to the variables VOC and PA, due to their non-normal distribution. This procedure tests the null hypothesis that the groups belong to similar populations, against the alternative hypothesis that they were extracted from different populations. The test results point towards the rejection of the null hypothesis (5% significance) for the variables VOC ($z = -4.920$; $p < .01$) and PA ($z = -4.586$; $p < .01$).

The instrument employed to analyze the occurrence of

intergroup differences in the TSWRF variables was quantile regression with standard errors robust to heteroscedasticity. The mathematical model involved using the gross scale of the TSWRF (TSWRF_b) as the dependent variables and PA, AGE and VOC as independent variables, besides a dichotomous variable (independent) called YEAR, described by 0 to

represent the third year and 1 for the fifth year. This statistical procedure serves to estimate the variation in the median value of the dependent variable in function of the explanatory variables. The application of this estimator is justified by the non-Gaussian distribution of the residues obtained by the classical linear regression model.

Table 1
Descriptive statistics of the participants' variables analyzed

Variable	Min.	Max.	Third-year (n = 63)		Asymmetry		Kurtosis	
			Est.	Standard Deviation Standard Error	Est.	Standard Error	Est.	Standard Error
Third-year (n = 63)								
AGE	7.67	10.17	8.56	0.57	0.53	0.30	-0.03	0.59
TSWRF	42.48	110.66	78.40	15.85	-0.02	0.30	-0.37	0.59
TSWRF_b	20	62	42.09	9.74	-0.01	0.30	2.57	0.59
VOC	16.00	77.00	56.33	11.47	-0.62	0.30	1.15	0.59
RAVEN	1.00	90.00	29.86	18.53	0.82	0.30	1.04	0.59
PA	5.00	30.00	11.59	5.31	1.68	0.30	2.74	0.59
sqrt_RAVEN	1.00	9.49	5.17	1.78	-0.12	0.30	-0.20	0.59
inv_sqrt_PA	0.18	0.45	0.31	0.06	0.02	0.30	0.31	0.59
Fifth-year (n = 48)								
AGE	9.75	16.00	12.13	1.47	0.72	0.34	0.01	0.67
TSWRF	13.43	99.53	74.00	17.93	-1.39	0.34	2.10	0.67
TSWRF_b	27	63	52.31	7.49	-1.34	0.34	4.74	0.67
VOC	40.00	80.00	66.60	8.67	-1.13	0.34	1.41	0.67
GR	201.00	530.00	305.48	82.50	1.04	0.34	0.46	0.67
RL	149.00	331.00	240.48	44.61	-0.01	0.34	-0.33	0.67
E%	2.48	45.47	19.27	11.44	0.48	0.34	-0.81	0.67
OR	6.00	40.00	20.31	9.72	0.53	0.34	-0.78	0.67
PA	1.00	34.00	18.54	8.06	-0.09	0.34	-0.95	0.67
ln_AGE	2.28	2.77	2.49	0.12	0.49	0.34	-0.43	0.67
TSWRF_cube	2422.30	9.8x10 ⁵	4.6x10 ⁵	2.4x10 ⁵	-0.06	0.34	-0.32	0.67
VOC_square	1600.00	6400.00	4509.77	1078.63	-0.73	0.34	0.32	0.67
ln_GR	5.30	6.27	5.69	0.25	0.59	0.34	-0.45	0.67
sqrt_E%	1.57	6.74	4.18	1.35	-0.01	0.34	-0.93	0.67
ln_OR	1.79	3.69	2.89	0.50	-0.17	0.34	-0.93	0.67
PA_sqrt_PA	1.00	198.25	85.60	50.97	0.26	0.34	-0.97	0.67

Note. AGE= participant's age; TSWRF = standardized test score; TSWRF_b = gross test score; VOC = test score; RAVEN = test score; PA = test score; sqrt_RAVEN = square root of RAVEN variable. The variable inv_sqrt_PA = inverse of square root of variable PA; GR, RL, E%, OR are scores obtained based on D2 test; ln_AGE = natural logarithm of AGE variable; TSWRF_cube = cubed TSWRF; VOC_square = squared VOC; ln_GR = natural logarithm of GR; sqrt_E% = square root of E%; ln_OR = natural logarithm of OR; PA_sqrt_PA = PA multiplied by square root of PA.

In combination, the four variables presented coefficients different from zero ($F = 23.15$, $df1 = 4$, $df2 = 106$, $p < .01$) and, separately, the variables that presented coefficients statistically different from zero were PA (coef.= 0.53, $t = 4.20$, $p < .01$) and VOC (coef. = 0.22, $t = 2.72$, $p < .01$). On the opposite, YEAR

(coef. = 2.38, $t = 0.73$, $p = .46$) and AGE (coef. = 0.54, $t = 0.80$, $p = .42$) did not present statistical significance. Therefore, it cannot be affirmed based on the data that the difference between the school years, extracting the effect of VOC, PA and AGE, may be related to variations in the median TSWRF_b. The same

is true for AGE. An estimator robust to heteroscedasticity was chosen out of prudence after the results of the Machado–Santos Silva test ($chi\text{-}sq. = 5.66$, $df. = 2$, $p = .06$) had been observed. The proportions found in the group samples for the levels of the net TSWRF score were: third year (very low = 23.8%, low = 44.4% and medium = 31.8%) and fifth year (very low = 23%, low = 43.7% and medium = 33.3%).

Pearson’s correlation coefficients were calculated for the variables associated with both groups separately. Due to the need to extract the variables from a population with a Gaussian frequency distribution for statistical calculations, transformations were needed to comply with that requirement. For the third-year group, the transformations were made in the variables RAVEN and PA to $\sqrt{\text{RAVEN}}$ and $\text{inv}\sqrt{\text{PA}}$, respectively. The descriptive statistics and meaning of the changed variables were described in Table 1.

Table 2 evidences the results obtained based on the calculated correlations. The level of association of the variables complies with the proposal by Cohen (1988). Hence, only the inference between $\text{inv}\sqrt{\text{PA}}$ and TSWRF can be considered strong. The negative signal shows that, as the TSWRF score increases, the same happens with the PA score. Moderate correlations were observed between VOC and AGE, VOC and TSWRF, $\sqrt{\text{RAVEN}}$ and TSWRF, TSWRF and $\sqrt{\text{RAVEN}}$, VOC and $\sqrt{\text{RAVEN}}$, VOC and $\text{inv}\sqrt{\text{PA}}$ and between $\text{inv}\sqrt{\text{PA}}$ and $\sqrt{\text{RAVEN}}$. In the first four pairs, variations in the same sense tend to be confirmed while, in the last two cases, the negative signal has the same meaning as the signal presented in the strong correlation. A weak and statistically non-significant correlation was found between TSWRF and AGE.

Table 2

Pearson correlation matrix calculated for the cognitive variables, age, reading and writing for the sample of third-year students

	AGE	TSWRF	VOC	$\sqrt{\text{RAVEN}}$	$\text{Inv}\sqrt{\text{PA}}$
AGE	1				
TSWRF	.17	1			
VOC	.38**	.41**	1		
$\sqrt{\text{RAVEN}}$	-.06	.45**	.49**	1	
$\text{inv}\sqrt{\text{PA}}$	-.09	-.59**	-.43**	-.46**	1

Note. ** 1% (significance level).

Note. Degree of association in absolute figures: between .1 and .3 = weak; between .3 and .5 = moderate; higher or equal to .5 = strong (Cohen, 1988).

What the fifth-year students are concerned, transformations were needed in the variables AGE, TSWRF, VOC, GR, E% and OR. Their new versions were In_AGE , TSWRF_cube , VOC_square , In_GR , $\sqrt{\text{E\%}}$ and In_OR , respectively. The descriptive statistics and the meaning of the changed variables were described in Table 1. The variable PA went through a change, as its influence on the reading ability had to be investigated using hierarchical regression and did not comply with the homoscedasticity requisite (constant variation in PA along the evolution of the TSWRF scale observed in a scatterplot between the two variables) required for this instrument to produce proper results. Therefore, PA was transformed into PA_sqrt_PA , complying with the orientations by Hair et al. (2010).

The fact that the D2 scale is divided in four components (GR, RL, E% and OR) motivated their synthesis into a single variable, in view of the strong correlations among their transformed coefficients (Table 3). The tool used for this purpose was the principal component analysis. As In_GR and $\sqrt{\text{E\%}}$ contain information about RL, the use of the latter was discarded in the reduction process, thus avoiding redundancy. The resulting variable was named D2_FACTOR .

Among the associations presented in Table 3, the variable TSWRF_cube shows a moderate correlation with VOC_square and a strong correlation with PA_sqrt_PA , both with a positive signal, evidencing that, in the sample, high

levels of reading expertise accompany high vocabulary and phonological awareness scores. The correlation between VOC_square and PA_sqrt_PA is positive and moderate. On the opposite, the variable In_AGE is weak or insignificantly correlated with the variables TSWRF_cube , VOC_square and PA_sqrt_PA ; in addition, the synthesis-variable D2_FACTOR shows similar correlation levels with In_AGE , TSWRF_cube and VOC_square . A moderate and negative correlation was also observed between D2_FACTOR and PA_sqrt_PA , which seems to indicate that higher levels of the first variable accompany lower levels of the second. Among the components of D2_FACTOR , the correlation found between PA_sqrt_PA and $\sqrt{\text{E\%}}$ appoint that higher rates of errors committed by the subjects who perform the D2 test are associated with lower phonological awareness scores. The remaining components are weak and negatively correlated with PA_sqrt_PA .

After studying the correlations among the variables, the objective was to verify the other variables’ power to explain the reading skills. The statistical procedure used for this purpose was hierarchical regression. The ability of one variable to explain the others is given by its contribution to the R^2 coefficient (ranges between 0 and 1), which measures how much of the variance in the dependent variable is explained by the joint dispersion of the independent variables, equal to 0 for absence of explanation and 1 for the situation in which all independent variables listed fully explain any deviation in the dependent variable.

Table 3

Pearson correlation matrix calculated for the cognitive variables, age, reading and writing for the sample of fifth-year students

	ln_AGE	TSWRF_cube	VOC_square	ln_GR	RL	sqrt_E%	ln_OR	PA_sqrt_PA	D2_FACTOR
ln_AGE	1								
TSWRF_cube	-0,05	1							
VOC_square	-.06	.38**	1						
ln_GR	.27	-.06	.01	1					
RL	.29*	.07	.21	.82**	1				
sqrt_E%	.12	-.13	-.31*	.58**	.04	1			
ln_OR	-.01	.03	-.12	.53**	.13	.78**	1		
PA_sqrt_PA	-.21	.58**	.41**	-.24	.00	-.38**	-.18	1	
D2_FACTOR	.14	-.06	-.17	.79**	.35*	.91**	.90**	-.31*	1.00

Note. **1% (significance level); * 5% (significance level).

Table 4

Hierarchical regression statistics for the group of third-year students using TSWRF as the dependent variable and AGE, sqrt_RAVEN, VOC and inv_sqrt_PA as independent variables

Model	Indep. Var.	Dependent Variable: TSWRF					
		Standard. coefficients	P-value <i>t</i> test	P-value <i>F</i> Test	<i>R</i> ²	<i>R</i> ²	P-value change
1	AGE	.17	.19	.19	.03	.03	.19
2	AGE	.19	.09	.00	.24	.21	.00
	sqrt_RAVEN	.46	.00				
3	AGE	.11	.38	.00	.26	.02	.18
	sqrt_RAVEN	.36	.01				
	VOC	.20	.18				
4	AGE	.11	.34	.00	.41	.15	.00
	sqrt_RAVEN	.22	.10				
	VOC	.08	.58				
	inv_sqrt_PA	-.45	.00				
5	AGE	.14	.17	.00	.41	.17	.00
	sqrt_RAVEN	.25	.03				
	inv_sqrt_PA	-.46	.00				
6	AGE	.11	.34	.00	.41	.00	.58
	sqrt_RAVEN	.22	.10				
	inv_sqrt_PA	-.45	.00				
	VOC	.08	.58				

In Table 4, the hierarchical regression analysis results are displayed for the third-year students. The dependent variable was TSWRF and the independent variables were AGE sqrt_RAVEN, VOC and inv_sqrt_PA, inserted sequentially.

Departing from the independent variable AGE, with a very small *R*² (.03), the variable sqrt_RAVEN was added, which provoked an increase by .21 in *R*². When inserted in the model, whether before or after inv_sqrt_PA, the variable VOC did not significantly contribute to the power

to explain TSWRF. The standardized VOC coefficients are not statistically significant either. Nevertheless, the variable inv_sqrt_PA contributes with *R*² coefficients between .15 and .17, depending on the order of the variable's insertion in the regression. Its standardized coefficients in all models in which it is present are statistically significant and negative, confirming the positive association between reading and phonological awareness, excluding the effects of the other variables involved.

For the fifth-year students, the hierarchical regression results were displayed in Table 5. The variables *ln_AGE* and *D2_FACTOR* jointly contribute very little to explain the variance in *TSWRF_cube* (independent variable), with R^2 equaling .02. When inserted before *PA_sqrt_PA*, the variable *VOC_square* contributes more to explain the independent variable ($R^2 = 0.19$) than when it is inserted

after *PA_sqrt_PA* ($R^2 = 0.05$). The coefficient of that variable was only statistically significant in the model when it did not appear together with *PA_sqrt_PA*. The contribution of that variable was greater ($R^2=0.37$) when inserted before *VOC_square* and smaller ($R^2 = 0.23$) when inserted afterwards. The coefficients of the variable were statistically significant and positive.

Table 5
Hierarchical regression statistics for the group of fourth-year students using *TSWRF_cube* as the dependent variable and *ln_AGE*, *D2_FACTOR*, *VOC_square* and *PA_sqrt_PA* as independent variables

Model	Indep. Var.	Dependent variable: <i>TSWRF_cube</i>					
		Standard. coefficients	P-value <i>t</i> test	P-value <i>F</i> Test	R^2	R^2	P-value change
1	<i>ln_AGE</i>	.04	.80	.80	.00	.00	.80
2	<i>ln_AGE</i>	.05	.72	.67	.02	.02	.39
	<i>D2_FACTOR</i>	-.13	.39				
3	<i>ln_AGE</i>	.07	.61	.02	.21	.19	.00
	<i>D2_FACTOR</i>	-.06	.69				
	<i>VOC_square</i>	.44	.00				
4	<i>ln_AGE</i>	.16	.18	.00	.44	.23	.00
	<i>D2_FACTOR</i>	.07	.56				
	<i>VOC_square</i>	.24	.07				
	<i>PA_sqrt_PA</i>	.56	.00				
5	<i>ln_AGE</i>	.17	.17	.00	.39	.37	.00
	<i>D2_FACTOR</i>	.06	.63				
	<i>PA_sqrt_PA</i>	.65	.00				
6	<i>ln_AGE</i>	.16	.18	.00	.44	.05	.07
	<i>D2_FACTOR</i>	.07	.56				
	<i>PA_sqrt_PA</i>	.56	.00				
	<i>VOC_square</i>	.24	.07				

It seems that, in both groups, phonological awareness contributed more than vocabulary to explain the reading. The models in which the explanatory variables were exclusively related to AGE and D2 (fifth year) and to AGE (third year) were not statistically significant.

Discussion

The results of the comparison between the initial (third year) and final years (fifth year) indicated that the mean phonological awareness and vocabulary scores were higher for the fifth than for the third year. Similar results were found in T.P. Alloway and R.G. Alloway (2010) and indicate that children's exposure to educational environments can be important to develop skills required to acquire reading and writing, as appointed in F.C. Capovilla (2011).

Nevertheless, the tasks developed in the school environment do not seem to have been sufficient to expand the participants' knowledge in reading skills. This conclusion

is due to the fact that, when comparing *TSWRF_b* between both groups, using quantile regression, the *YEAR* variable was not statistically significant. This finding indicates that, in the course of the education process, the participants did not evolve in the acquisition and development of reading. The standardized *TSWRF* scores for both years support this evidence as, in approximately 44% of the sample, the students present low reading levels according to the test. This means saying that the gross reading score increased from the third to the fifth year but that, in qualitative terms, the students can be considered bad readers in both years. In addition, as the gross reading scores evolved associated with higher phonological awareness and vocabulary scores, the effects of the year, excluding the correlations of phonological awareness and vocabulary, are not statistically significant. Therefore, it can be inferred that the evolution in the phonological awareness and vocabulary scores derived from other activities than the formal teaching of reading (Lundberg, Larsman, & Strid, 2012).

Two hypotheses can be raised to explain the lack of evolution in the participants' knowledge of reading skills. The first can be related to the participants' attendance of school activities (not assessed in this study). The main economic activity of the community the school is located in is the catching of a type of shellfish, whose preparation for sale is mainly executed by elderly people and children. While the younger children stay in school, the older ones do not regularly attend the school activities, provoking dropout and return rates that culminate in a year/age distortion and compromise these participants' learning process. School Census data for this school indicate 43.5% of age/year distortion for the fifth year (INEP, 2014).

Another hypothesis for this result, which can be associated with the first, is related to the actual teaching environment (methods employed and pedagogical policy for example), which may not be effective to teach reading. This hypothesis rests on the indicators of the school this sample was taken from. The score on the official Portuguese language test SAEB/Prova Brasil in 2011, for example, was inferior to 150 points. This score is equivalent to Level 1 on the SAEB Portuguese Language Proficiency Scale. At this level, the student merely locates explicit information in short texts and possesses minimal skills to recognize structural categories of narratives (INEP, 2014).

As regards the intra-series analyses, the results indicate a strong correlation between reading and phonological awareness, as described in the literature by Bandini et al. (2013) and Justi and Roazzi (2012) for the third as well as the fifth year. This correlation is positive, that is, the higher the reading scores, the greater the phonological awareness skills, supporting the hypothesis that phonological awareness and reading skills interact in a two-way sense, as proposed by Lundberg et al. (1988).

In addition, moderate correlations were observed between reading and vocabulary and between vocabulary and phonological awareness for both years. Similar results were described in Berendes et al. (2010), F.C. Capovilla (2011) and Justi and Roazzi (2012). These findings strengthen the theory about the mutual influence between the skills, as described by Bandini et al. (2013).

An interesting information observed in this study was that the variables phonological awareness and reading fluency were weak and non-significantly correlated with the AGE variable in both years, that is, AGE does not seem to affect reading fluency and phonological awareness. These findings suggest that training seems to be an important factor to improve these skills. According to Morais (1996), phonological awareness and reading skills cannot be learned only with age.

On the other hand, the vocabulary variable seems to vary in terms of the influence of the AGE variable. These findings verified a weak and non-significant correlation between age and vocabulary for the fifth year and a moderate correlation for the study participants in the third-year of elementary school. Similar findings on the evolution of the vocabulary as age advances can be observed in the studies by Santos and Befi-Lopes (2012). These findings suggest that the vocabulary evolves more

in the initial years of education.

The results appointed moderate correlations between phonological awareness and intelligence for the third year. These findings differ from the findings by Justi and Roazzi (2012), who reported that intelligence does not affect phonological awareness. Nevertheless, the use of different forms to assess the phonological awareness may have provoked this divergence. In the study by Justi and Roazzi (2012), only two phonological awareness measures were used (subtraction of phonemes and phonemic manipulation) while, in this study, the assessment of phonological awareness proposed by A.G.S. Capovilla and F.C. Capovilla (1998) was used, which includes ten phonological awareness skills, thus expanding the chances of assessing the relation between those skills and intelligence.

The intellectual performance measures for the third year also presented moderate correlation with reading and vocabulary skills. These findings correspond to the findings described in the studies by F.C. Capovilla (2011) and Gorman (2012), respectively, who highlight the role of intelligence as an important factor in vocabulary acquisition.

The correlations between the focused attention test and the phonological awareness skills were only assessed for the fifth year due to the application age of the test. The findings appointed moderate and negative correlations, that is, the higher the error rate on the D2 test, the lower the phonological awareness score. No studies were found that relate both skills, but this relation can be explained by the fact that the phonological awareness skills require not only good auditory perception performance to execute the task, but also attention to work with abstract concepts, such as phonemes. Therefore, the larger the number of errors on the D2 test, the lesser the participant's attention and the worse the phonological awareness score.

The hierarchical regressions conducted among the variables and presented in Tables 4 and 5 demonstrated that there are strong associations between phonological awareness and reading fluency and that the phonological awareness skills contribute more to the development of reading than the vocabulary skills, as presented by Bandini et al. (2013).

In view of the above, this study has demonstrated that the relations among skills that are considered prerequisites for reading acquisition are maintained, even when the children have a low socioeconomic status. Nevertheless, the participants' low reading performance in both years, despite the increase in the gross reading scores, phonological awareness and vocabulary, show that the non-effective teaching environment has raised limits for the expansion of the readers' learning capacity.

References

- Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology, 106*(1), 20-29. doi:10.1016/j.jecp.2009.11.003

- Angelini, A. L., Alves, I. C. B., Custódio, E. M., Duarte, W. F., & Duarte, J. L. M. (1999). *Matrizes progressivas coloridas de Raven: Escala especial. Manual* [Coloured progressive matrices of Raven: Special scale. Manual]. São Paulo, SP: CETEPP.
- Associação Brasileira de Empresas de Pesquisa. (2012). *Critério de classificação econômica: Brasil* [Criteria for economic classification: Brazil]. Retrieved from <http://www.abep.org/criterio-brasil>
- Bandini, H. H. M., Santos, F. H., & de Souza, D. G. (2013). Levels of phonological awareness, working memory, and lexical knowledge in elementary school children. *Paidéia (Ribeirão Preto)*, 23(56), 329-337. doi:10.1590/1982-43272356201307
- Berendes, K., Schnitzler, C. D., Willmes, K., & Huber, W. (2010). The importance of phonemic awareness and vocabulary for written language performance in elementary school. *Sprache, Stimme, Gehör*, 34(3), e48-e55. doi:10.1055/s-0029-1246203
- Brickenkamp, R. (2000). *Teste D2: Atenção concentrada: Manual, instruções, avaliação e interpretação* [Concentrated attention d2 Test: Manual, instruction, evaluation and interpretation]. São Paulo, SP: CETEPP.
- Cain, K., & Bignell, S. (2014). Reading and listening comprehension and their relation to inattention and hyperactivity. *British Journal of Educational Psychology*, 84(Pt 1), 108-124. doi:10.1111/bjep.12009
- Capovilla, A. G. S., & Capovilla, F. C. (1998). Prova de consciência fonológica: Desenvolvimento de dez habilidades da pré-escola à segunda série [Phonological awareness: Test development of ten preschool skills to the second series]. *Temas Sobre Desenvolvimento*, 7(37), 14-20.
- Capovilla, F. C. (2011). *Teste de vocabulário por figuras USP- TVfUSP: Normalizado para avaliar a compreensão auditiva de palavras dos 7 aos 10 anos* [Vocabulary test figures USP TVfUSP: Normalized to assess the listening of words from 7 to 10 years]. São Paulo, SP: Memnon.
- Capovilla, F. C., & Capovilla, A. G. S. (2010). *Teste de competência de leitura silenciosa de palavras e pseudopalavras* [Quiet reading proficiency testing words and pseudo words]. São Paulo, SP: Memnon.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. 2nd ed. Mahwah, NJ: Erlbaum.
- Engel, P. M. J., Santos, F. H., & Gathercole, S. E. (2008). Are working memory measures free of socio-economic influence? *Journal of Speech, Language, and Hearing Research*, 51(6), 1580-1587. doi:10.1044/1092-4388(2008/07-0210)
- Franceschini, S., Gori, S., Ruffino, M., Pedrolli, K., & Facchetti, A. (2013). A causal link between visual spatial attention and reading acquisition. *Current Biology*, 22(9), 814-819. doi:10.1016/j.cub.2012.03.013
- Gorman, B. K. (2012). Relationships between vocabulary size, working memory, and phonological awareness in Spanish-speaking English language learners. *American Journal of Speech-Language Pathology*, 21(2), 109-123. doi:10.1044/1058-0360(2011/10-0063)
- Hair, J. F., Jr., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Upper Saddle River, NJ: Prentice Hall.
- Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. (2014). *Indicadores educacionais*. Retrieved from <http://portal.inep.gov.br/indicadores-educacionais>
- Justi, C. N. G., & Roazzi, A. (2012). A contribuição de variáveis cognitivas para a leitura e a escrita no Português Brasileiro [The contribution of cognitive variables for reading and writing in Brazilian portuguese]. *Psicologia: Reflexão e Crítica*, 25(3), 605-614. doi:10.1590/S0102-79722012000300021
- Kendeou, P., van den Broek, P., Helder, A., & Karlsson, J. (2014). A cognitive view of reading comprehension: Implications for reading difficulties. *Learning Disabilities Research & Practice*, 29(1), 10-16. doi:10.1111/ldrp.12025
- Lundberg, I., Frost, J., & Petersen, O. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly*, 23(3), 263-285.
- Lundberg, I., Larsman, P., & Strid, A. (2012). Development of phonological awareness during the preschool year: The influence of gender and socio-economic status. *Reading and Writing*, 25(2), 305-320. doi:10.1007/s11145-010-9269-4
- Macedo, C. S., Andreucci, L. C., & Montelli, T. C. B. (2004). Alterações cognitivas em escolares de classe socio-econômica desfavorecida: Resultados de intervenção psicopedagógica [Cognitive function evaluation in school-age children from economically impoverished community: Results of enriched education program]. *Arquivos de Neuro-Psiquiatria*, 62(3b), 852-857. doi:10.1590/S0004-282X2004000500021
- Morais, J. (1996). *A arte de ler* [Art of reading]. São Paulo, SP: UNESP.
- Santos, M. T. M., & Befi-Lopes, D. M. (2012). Vocabulary, phonological awareness and rapid naming: Contributions for spelling and written production. *Jornal da Sociedade Brasileira de Fonoaudiologia*, 24(3), 269-275. doi:10.1590/S2179-64912012000300013

Heloisa Helena Motta Bandini is a Professor of the Universidade Estadual de Ciências da Saúde de Alagoas.

Carmen Silvia Motta Bandini is a Professor of the Universidade Estadual de Ciências da Saúde de Alagoas and Centro Universitário Cesmac.

Adhemar Ranciaro Neto is a Professor of the Universidade Federal de Alagoas.

Received: Sep. 26, 2015

1st Revision: May. 31, 2015

Approved: Sep. 26, 2016

How to cite this article:

Bandini, H. H. M., Bandini, C. S. M., & Ranciaro Neto, A. (2017). Relations Between Reading, Vocabulary and Phonological Awareness in low-Income Children. *Paidéia (Ribeirão Preto)*, 27(68), 314-323. doi: 10.1590/1982-43272768201709