The relevance of receptive vocabulary in reading comprehension

ABSTRACT

Purpose: To characterize the performance of students from the 5th year of primary school, with and without indicative of reading and writing disorders, in receptive vocabulary and reading comprehension of sentences and texts, and to verify possible correlations between both. Methods: This study was approved by the Research Ethics Committee of the institution (no. 098/13). Fifty-two students in the 5th year from primary school, with and without indicative of reading and writing disorders, and from two public schools participated in this study. After signing the informed consent and having a speech therapy assessment for the application of inclusion criteria, the students were submitted to a specific test for standardized evaluation of receptive vocabulary and reading comprehension. The data were studied using statistical analysis through the Kruskal–Wallis test, analysis of variance techniques, and Spearman's rank correlation coefficient with level of significance to be 0.05. A receiver operating characteristic (ROC) curve (was constructed in which reading comprehension was considered as gold standard. Results: The students without indicative of reading and writing disorders presented a better performance in all tests. No significant correlation was found between the tests that evaluated reading comprehension in either group. A correlation was found between reading comprehension of texts and receptive vocabulary in the group without indicative. Conclusion: In the absence of indicate of reading and writing disorders, the presence of a good range of vocabulary highly contributes to a proficient reading comprehension of texts.

RESUMO

Objetivo: Caracterizar o desempenho de escolares de 5º ano do Ensino Fundamental, com e sem indicativos de alterações de leitura e escrita, em vocabulário receptivo e compreensão leitora de orações e textos, e verificar possíveis correlações entre ambas. Métodos: Pesquisa aprovada pela Comissão de Ética em Pesquisa da Instituição (nº 098/13). Participaram deste estudo 52 crianças do 5º ano do Ensino Fundamental de 2 escolas públicas, com e sem indicativos de alterações de leitura e escrita. Após assinatura do Termo de Consentimento Livre e Esclarecido e avaliação fonoaudiológica para aplicação dos critérios de inclusão, os sujeitos realizaram uma prova específica para avaliação padronizada do vocabulário receptivo e compreensão leitora. Foi realizada análise estatística dos dados, utilizando o teste de Kruskal–Wallis, técnica de análise de variância e coeficiente de correlação de Spearman com nível de significância de 0.05. Foi construída uma curva ROC, na qual a compreensão de leitura foi considerada padrão-ouro. Resultados: Os escolares sem indicativos de alterações de leitura e escrita apresentaram melhor desempenho em todas as provas. Não foi encontrada correlação significativa entre as provas que avaliaram a compreensão leitora em ambos os grupos. Foi encontrada associação entre a compreensão leitora de textos e o vocabulário receptivo no grupo sem indicativos. Conclusão: Na ausência de indicativos de alterações de leitura e escrita, a presença de um bom vocabulário contribui significativamente para uma compreensão de leitura de textos proficiente.
INTRODUCTION

Reading comprehension (RC) refers to the capacity of understanding content expressed in reading material and putting it together with information from the reader’s world knowledge. It is considered a complex multidimensional process that involves the integral view of the graphic material, structural interpretation of the content, the integration of such information in the memory, and inference\(^{(1,2)}\). There is, then, the union of two factors so that the adequate comprehension of the material read can occur: the existence of good receptive vocabulary and the presence of the abilities of automated decoding\(^{(3)}\).

The learning process of reading is a process initiated by decoding, development of fluency, passing through the comprehension of words and sentences, and reaching the level of textual comprehension\(^{(4)}\). RC of texts has a unique importance in the process of acquisition, sharing, and construction of knowledge\(^{(5)}\).

When evaluating decoding and RC, the level of education should be taken into consideration, as should factors involved in these competencies, such as the ability to recognize printed word and make the graphophonemic association; activate representative structures attributing meaning to what is written; and understand the message, recognizing the grammatical structures, considering the word order and the use of punctuation marks\(^{(6,7)}\).

The receptive vocabulary is related to the perception and the processing of information and it corresponds to the words whose concept and meaning were developed\(^{(1,2,7,8)}\). Studies\(^{(2,3)}\) have shown that vocabulary can be a relevant variable for RC because it influences the processes of decoding, speed, and accuracy of reading.

The National Curriculum Parameters (NCPs) are quality references elaborated by the Federal Government to guide school staff (primary and secondary schools) in the execution of their tasks. On the basis of what is presented in the NCPs, it is possible to know which abilities the students should have according to their level of education. Regarding RC in the fifth year of Primary School (PS), students are expected to be able to autonomously read texts from different genres and of different lengths, choosing strategies to comprehend the explicit message, as well as the implicit content of the text\(^{(9)}\).

In Brazil, studies conducted to understand the possible relationships between the learning of reading and writing abilities and the metalinguistic capacities are common\(^{(10,11)}\). However, only developed metalinguistic abilities and good reading fluency do not guarantee the success in comprehension of the material read. Other linguistic abilities such as the efficiency of word recognition and vocabulary knowledge, the understanding of grammatical structures, and inferences are relevant abilities, which are developed and reach different competency levels with the advancement of age, education level, practice, and reading stimulation. This way, they should be investigated during the development process of RC\(^{(12)}\).

While reviewing the literature for this study, no Brazilian studies were found that analyzed and correlated the receptive vocabulary and RC of sentences and texts\(^{(10)}\). The investigation of this possible relation in subjects with and without indicatives of reading and writing disorders is relevant because, based on the studies obtained, a pattern of normality can be defined, making an articulation between the clinical and educational practices possible.

This way, this study aimed to characterize the performance of students from the fifth year of PS, with and without indicatives of reading and writing disorders, in receptive vocabulary and RC of sentences and texts, and investigate the possible correlation between both.

METHODS

This study was initiated after the approval from the Research Ethics Committee of the institution, under no. 098/13. This study was conducted in two public schools of regular education in the west side of São Paulo, with similar Basic Education Development Indexes (IDEB) (4.4 and 4.7). Children from two schools were recruited, being that in the first school contacted, it was not possible to complete the corpus for the study.

Initially, 84 children from the fifth year of PS from school A were asked to participate in the study (average age = 10 years and 9 months). From these, 35 students had the informed consent signed by their legal guardians and 32 fulfilled the inclusion criteria for this study.

With the intention of amplifying the corpus of the study, school B was contacted and 38 children were invited to participate in the study. From these, 20 fulfilled the inclusion criteria and totaled in 52 subjects that comprised the research corpus.

The inclusion criteria for this study were as follows: consent of the legal guardians by signing the TCLE; signing the consent; be enrolled in and attending the fifth year of PS at the schools that were researched; and absence of complaints or indicatives of visual/auditory disorders, neurological, behavioral, and/or cognitive disorders (investigated through surveys filled out by teachers and the subjects’ legal guardians).

Among the 52 subjects, 28 (13 females and 15 males) comprised the group without indicatives of reading and writing disorders (GOI) and 24 comprised the group with indicatives of reading and writing disorders (GWI) (9 males and 15 females).

All the tests were proctored individually, on different days. Although the student’s educational curricular activities were taken into consideration, they remained at school only during one period and, when asked to come before or after school, they did not. Therefore, students were pulled out of class, one by one, preferably when they had finished the activity given by their teachers. Loss of curricular content by the subjects was not observed, seeing that they only remained with the researchers for a short time, making it possible to keep up with the class activities.

The School Performance Test (SPT) was used as an instrument to separate subjects into both groups, using the standard score for fifth year/former fourth grade of PS. The GOI had the children whose performance was ≥112 points whereas in the GWI were those whose performance was <112 points and an alphabetic-stage writing level.
**Experimental Evidence**

Subjects were evaluated, individually, in terms of receptive vocabulary, sentence reading comprehension (SRC) and text reading comprehension (TRC) following the proctoring and grading instructions described on the tests.

For the evaluation of receptive vocabulary, the Vocabulary Test with Images USP (TVFUSP) was used. It contains 139 items and the child must hear a word and select the image that corresponds to it among four options. The passing grade established by the authors for children from the fifth year of PS (of low socioeconomic level) was 120.73 points.

On another day, subjects took the SRC and the TRC tests, numbers 9 and 10 of item IV of the PROLEC Test. The PROLEC Sentences (PROLEC-O) is made up of 12 sentences that express simple orders that should be executed. The passing grade established for the fourth grade (fifth year) students of PS was ≥11.81 points. The PROLEC Texts (PROLEC-T) is made up of four short texts (two narrative and two expository texts) whose meaning should be extracted to answer the four oral questions for each text. The passing grade established for the fifth year students of PS is ≥14.01 points.

**Statistical Analysis**

The distribution of the PROLEC-T grade among the groups with and without indicatives in both schools was compared through the Kruskal-Wallis test. To locate the differences indicated in the Kruskal-Wallis test, the procedure proposed by Dunn was applied, and the numbers for p were corrected by the inequality of Bonferroni. The same procedures were adopted to analyze the PROLEC-O.

To measure the correlation between the PROLEC-T and the PROLEC-O, Spearman’s correlation coefficient was calculated.

The technique for variance analysis was applied with the intention of comparing the grades of the TVFUSP in both groups and both schools. The suppositions necessary to apply the technique were verified through the analysis of the waste.

Spearman’s correlation coefficient was adopted as a correlation measurement between the PROLEC-O and the PROLEC-T and the TVFUSP. The association between the classifications of both tests was verified through goodness-of-fit tests.

Considering the PROLEC-T as gold standard, a receiver operating characteristic (ROC) curve was constructed for TVFUSP.

In the hypothesis tests, a significance level of 0.05 was set. Because the statistical analysis did not indicate significant differences among the performances obtained from the subjects of the two schools, the researchers opted for joining the samples in only two groups (GOI and GWI), disregarding the school of origin.

**RESULTS**

The medians observed in the TVFUSP and the GOI were higher than in the GWI (Table 1). The data were confirmed through the variance analysis technique, which indicated a significant difference between the average grades in the TVFUSP in both groups (p<0.001).

In the PROLEC-O analysis, the averages and medians of the GOI were observed to be higher than those of the GWI (Table 1), a fact verified through the Kruskal–Wallis test, which indicated a difference between the grades of both groups in the PROLEC-O (p=0.003). Although the majority of children from both groups were indicated as “average” or “above average,” the percentage of children above average in the GOI (78.6%) was found to be higher than those in the GWI (37.5%).

<table>
<thead>
<tr>
<th>Group without indicatives (n=28)</th>
<th>Mean (SD)</th>
<th>Median</th>
<th>Min-Max</th>
<th>Expected (5th year of PS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TVFUSP*</td>
<td>121.0 (21.6)</td>
<td>125</td>
<td>15–136</td>
<td>≥120</td>
</tr>
<tr>
<td>PROLEC sentences**</td>
<td>11.6 (0.9)</td>
<td>12</td>
<td>8–12</td>
<td>≥11</td>
</tr>
<tr>
<td>PROLEC texts**</td>
<td>13.3 (2.1)</td>
<td>14</td>
<td>7–16</td>
<td>≥14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group with indicatives (n=24)</th>
<th>Mean (SD)</th>
<th>Median</th>
<th>Min-Max</th>
<th>Expected (5th year of PS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TVFUSP*</td>
<td>117.4 (5.8)</td>
<td>118.5</td>
<td>104–127</td>
<td>&lt;120</td>
</tr>
<tr>
<td>PROLEC sentences**</td>
<td>10.1 (3.1)</td>
<td>11</td>
<td>0–14</td>
<td>&lt;11</td>
</tr>
<tr>
<td>PROLEC texts**</td>
<td>9.3 (3.7)</td>
<td>9.5</td>
<td>0–14</td>
<td>&lt;14</td>
</tr>
</tbody>
</table>

*Variance analysis; **Kruskal-Wallis test

Caption: TVFUSP = receptive vocabulary test through images from USP; PROLEC = Reading Comprehension test for sentences and texts; SD = standard deviation; Min = minimum; Max = maximum; PS = Primary School.

Per example of what occurred with the PROLEC-O, the results obtained regarding the PROLEC-T show averages and medians higher in the GOI (Table 1), with significant differences among the grades distributed in the PROLEC-T in both groups (p<0.001), which is confirmed by the Kruskal-Wallis test. It was noted that 91.7% children in the GWI were classified as “below average,” whereas the majority of those in the GOI were classified as “average” or “above average.”

The coefficient value of Spearman’s correlation of the grades in the PROLEC-T and PROLEC-O in the GWI was r=0.39 (p=0.052) and that in the GOI was r=0.29 (p=0.139). Therefore, no significant correlation was observed between the grades in the two tests in both groups, yet the p-value obtained in the GWI was close to the significant value (0.05).

In Figures 1 and 2, diagrams of dispersion between the PROLEC-O and the TVFUSP and between the PROLEC-T and the TVFUSP are presented, respectively. In these diagrams, joint behavior of the PROLEC and the TVFUSP can be viewed, for children with and without indicatives of disorders. Below each figure, the values observed for Spearman’s correlation coefficient between the PROLEC and the TVFUSP are presented.
In Figure 1 and through Spearman’s correlation coefficient between the PROLEC-O and the TVFUSP, no significant correlation was observed between the two groups, yet the p-value obtained in the GWI (p=0.073) was close to 0.05.

However, in Figure 2 and through Spearman’s correlation coefficient, a significant positive correlation was observed between the PROLEC-T and the TVFUSP only in the GOI.

With the intention of getting a cutoff point for the TVFUSP, such as the result from a child could point to the presence or lack of indicatives of reading and writing disorders, an ROC curve was constructed (Figure 3).

The PROLEC-T was considered gold standard and its results were classified in this study under two categories: adequate (“average” or “above average”) and inadequate (“below average”).

It was observed that all the children classified as “adequate” in the PROLEC-T were classified as “adequate” in the TVFUSP; from the children classified as “inadequate” in the PROLEC-T, 60.6% were also classified in this category in the TVFUSP.

The point closest to the upper left corner of the ROC curve corresponds to the maximum value for sensitivity (91%) and specificity (89%) reaching the cutoff point of the TVFUSP (125.15). Consequently, the children with a score of 126 or higher in the TVFUSP were classified as having an adequate performance in the TRC and those with 125 points or less as having an inadequate performance.

DISCUSSION

RC is a very complex process that involves various other related processes and abilities, once it is not necessary merely to decode but to attribute a meaning to the material read. Such an interpretation depends on innumerable variables, such as metalinguistic and cognitive abilities, and short-term verbal and vocabulary memory. However, the role of each of these components in the RC process is still not very clear. Therefore, the intention of this study was to characterize the performance of fifth year students from PS, with and without indicatives of reading and writing disorders, and to investigate possible correlations between both.

In terms of receptive vocabulary (TVFUSP), higher averages were observed for the GOI in relation to the GWI. Such findings corroborate a study\(^{(18)}\) in which it was verified that a good performance in perception and oral language production is an indicator of intelligence and development in formal academic abilities.

Another study\(^{(19)}\) also found close results, when investigating the relationship between reading, working memory, intelligence quotient, and vocabulary in children from the second to fifth grade of public school with an underprivileged psychosocial sample. The results showed that there is a narrow relationship between working memory, vocabulary, and other cognitive processes in the reading performance of these students. In addition, when the children with intellectual deficits were excluded from the sample, the vocabulary and the visual-spatial working
memory greatly contributed to the RC performance of texts (retelling and questions) of these children.

Stimulating world knowledge in these children so that they may comprehend information based on the context in which it is presented (inference), through oral or written language, is one of the main forms of acquiring vocabulary. Learning a new word implies in learning its form, meaning, and use\(^{20}\). This way, the findings of this study, in addition to corroborating the mentioned studies, call the attention of the educators to the necessity of working on lexical amplification, especially with the children who present a more hindered linguistic performance because students in the GOI had a better performance than those in the GWI.

In terms of SRC (PROLEC-O), a significant difference was observed in the distribution of the scores between both groups, with superior results for the GOI. Similar results were found by researchers\(^{21,22}\) when analyzing the reading process of students with learning disorders when compared with children with adequate school performance. In these studies, children with a better performance obtained good results in RC. This way, the importance of adequate development of the aspects that influence RC (world knowledge, vocabulary, inference processes, syntactic and semantic integration within the sentences, visual recognition of the words, short-term memory, monitoring, and metacognition) can be pointed out so that the information being transmitted may make sense and the reader can be capable of recognizing the role of each of the words and the way they are organized in a clause\(^{23,24}\).

As for the results from TRC (PROLEC-T), it can be affirmed that this test was a good indicator of students with and without indicatives of reading and writing disorders, confirming the division of subjects in the groups GOI and GWI through the SPT, because the difference between the two groups was bigger than that in the PROLEC-O. Therefore, these results showed that the SRC test was not an instrument for distinguishing the performance of fifth year students, contrary to what was found in another study\(^{25}\), conducted with students in the initial phase of literacy.

In RC, in addition to recognizing a word, extracting its meaning, and establishing its role within the clause, the reader must be able to integrate into their memory the various information in the clauses, put together their world knowledge and make inferences to appropriate the content read\(^{1,4}\). Seeing that the students with indicatives present a deficiency in orthographic representations, use the phonological route more frequently during reading, and take longer for the decoding of the material read to occur, they can be hindered in the use of the memory and integration of information, which justifies their difficulty in understanding texts\(^{22,26}\).

No positive correlation was observed between the PROLEC-O and the PROLEC-T for the GWI (p=0.052) as well as for the GOI (p=0.139); however, the children from the GWI had a worse score in both tests. This suggests that the children with indicatives of disorder tend to have a below average performance in the reading of sentences and, consequently, texts. This fact was already expected, especially in TRC, because it demands cognitive processes and more complex world knowledge.

In Figure 1, the absence of significant correlation between the PROLEC-O and the TVFUSP, in both groups, suggests that the vocabulary does not seem to significantly interfere in SRC, especially the ones with a simple order, regardless of the group the child belongs to. However, the children from the GWI presented vocabulary and SRC, which were inferior to the GOI.

The data in Figure 2 and Spearman’s correlation coefficient indicate a significant correlation between the PROLEC-T and the TVFUSP only in the GOI. These data corroborate recent studies\(^{22,27}\), which describe that the knowledge of vocabulary is one of the best indicators of reading mastery because it relates to the ability of obtaining new details from reading texts. In addition, the quantity of familiar vocabulary words is one of the most significant elements to discern the level of complexity of a text because the more vocabulary a child knows on a graphic material, the easier it will be to interpret the message on it.

The findings of this study bring an original contribution by showing that, in addition to the metalinguistic abilities and the mastery of the rules for converting graphemes/phonemes and spelling rules, vocabulary is also a determining and predicting variable for the development and consolidation of TRC for students in the fifth year of PS.

Another innovative result from this study was the analysis of the ROC curve, which showed high sensitivity and specificity for the cutoff point, indicating that from 125.12 points there would be a higher probability of a child having his/her performance considered adequate in the PROLEC-T, confirming that the receptive vocabulary is a predictor of performance in children from the fifth year in PS in TRC. Therefore, a parameter is established for vocabulary performance, which is important for professionals who work with children at the end of PS.

The data obtained through this study were delivered (in the format of individual and collective reports) and discussed with the teachers and administrative staff from the two schools where the data were collected. It is relevant to state the enormous receptiveness from the members of the staff from both schools, who showed an interest in the study and in suggestions of applicability for the data in the context of education. Once again, the benefits of speech therapists and teachers working together are evidenced. In addition, all students with indicatives for reading and writing disorders were sent to have a speech therapist evaluation with the aim of textual interpretation, with the intention of clarifying the presence or lack of presence of this learning problem in reading and writing.

It is important to point out the implications of the findings of this study, not only for the clinical speech therapy work but also for the educational context. This way, by using strategies of lexical amplification, the teachers will be providing students with RC enhancement, without being glued to the main objective of the pedagogical activities. However, having activities of textual interpretation, with the intention of lexical amplification, can enhance oral language and RC of school texts, having in mind that the students in the GOI obtained better performances in vocabulary and TRC.
Thus, this study substantially contributed to the scholars of the reading and writing area by showing important evidence of vocabulary for TRC. This way, it can be inferred that the vocabulary should be highlighted in the class planning (with typical children who have indicative of reading and writing disorders), as well as in the speech therapy clinic for the children with a learning issue.

CONCLUSION

In this study, students in the GOI present a better performance in receptive vocabulary and in SRC and TRC.

The positive correlation between the receptive vocabulary and TRC only in students belonging to the GOI indicates the relevance of knowing the words and their meanings for the consolidation of a proficient TRC.

The data in this study allowed the researchers to speculate that from the moment a child is capable of understanding written language, a relationship of reciprocity is consolidated between reading and vocabulary.

*AFON was responsible for the collection, tabulation, and analysis of the data and for drafting of the manuscript; AJCS contributed to the analysis and the interpretation of the data and with the drafting and editing of the manuscript; MSC directed the research design, contributed to the analysis and interpretation of the data, and to the drafting, supervision, and editing of the manuscript.

REFERENCES