

Cloze test as an instrument to measure reading proficiency: linguistic and non-linguistic factors

O teste de Cloze como instrumento de medida da proficiência em leitura: fatores linguísticos e não linguísticos

Kátia Nazareth Moura de Abreu

Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
kabreu00@gmail.com

Daniela Cid de Garcia

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
cid.daniela@gmail.com

Katharine de Freitas P. N. A. da Hora

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
freitas.katharine@gmail.com

Cristiane Ramos de Souza

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
tianeramoss@gmail.com

Abstract: This article aims to present a review of the literature concerning the use of the Cloze test as a tool to measure reading performance, discussing its limits and possibilities; to access the sensibility of the test for linguistic and non-linguistic variables; and to present a research agenda that considers the use of original qualitative designs to access different aspects of reading. To verify the sensibility of the test, we manipulated the variables of word type (function words and content words), text length (short and long), and reading experience (measured

in terms of the level of schooling). We observed the performance of Elementary-level and High School students. The results show that Elementary-level students had more difficulty with function words and with long texts, which indicates that the test is sensible to linguistic and non-linguistic factors. From the results and analyses, we draw a discussion of how the Cloze test can be useful to evaluate different cognitive processes involved in reading.

Keywords: reading comprehension; Cloze test; education; content words; function words.

Resumo: Este artigo tem como objetivos apresentar uma revisão da literatura do uso do teste de Cloze como ferramenta para medir o desempenho em leitura, discutindo seus limites e possibilidades, verificar a sensibilidade do teste para variáveis de natureza linguística e não linguística, e apresentar uma agenda de pesquisa que considere o uso de desenhos qualitativos originais para acessar diferentes aspectos da leitura. Para verificar a sensibilidade do teste, manipularam-se as variáveis tipo de palavra (lexical e funcional), tamanho de texto (curto e longo) e experiência leitora (medida em termos de nível de escolaridade). Observou-se o desempenho de alunos do Ensino Fundamental e do Ensino Médio. Os resultados demonstraram que os alunos do Ensino Fundamental apresentaram mais dificuldade com palavras funcionais e com os textos longos, indicando que o teste é sensível a fatores linguísticos e não linguísticos. A partir dos resultados e das análises, discute-se como esse teste pode ser útil para avaliar diferentes processos cognitivos envolvidos na leitura.

Palavras-chave: compreensão leitora; teste de Cloze; ensino; palavras lexicais; palavras funcionais.

Received on January 5, 2017

Accepted on April 10, 2017

1 Introduction

The understanding of a text may vary according to the reading circumstances and may depend on complex factors present on the text surface, as well as in the activation by the reader, of a set of knowledge implicit in the reading act. Reading requires the use of perception and cognition processes; knowledge of the language, grammar, and topic; as well as the knowledge of how to read. Furthermore, reading involves a combination of visual and non-visual information. It is an interaction between the reader and the text (SMITH, 2003).

This article aims to provide a literature review on the use of the Cloze test in order to discuss its reach as a reading comprehension measurement tool, to determine its sensitivity to linguistic and non-linguistic factors – given that reading is an amalgamate of factors of different natures – and to present an agenda on how this type of tests may be qualitatively designed so as to access different abilities involved in the reading act.

The Cloze test is a technique consisting of a text from which every fifth word is missing, and has been replaced with a line of similar size. This technique has been used in Brazil for the past forty years as a reading comprehension diagnostic instrument – which assesses reading comprehension, interprets the results, and classifies the participants as to performance levels – and as an intervention instrument – which develops reading comprehension ability by allowing participants the opportunity to predict the missing words based on their previous knowledge and mastering of linguistic structures.

To assess the impact of factors such as word type, text length, and reader's level of education, we developed a Cloze test for sixth graders and tenth graders. The variables were selected so as to include linguistic knowledge elements (the difference between function and lexical words), as well as elements related to more general cognitive processing, such as attention, memory, and motivation. In addition, groups of different education levels were included to determine if their reading experience would have any impact on performance.

Considering the linguistic factor of word type, it is known that, from the information processing standpoint, there is a value differentiation between lexical and function words, regarding their contribution to recovering the input meaning (VAN PATTEN, 1996).

Lexical words present greater semantic content, while function words carry, above all, grammatical information; therefore, psycholinguistic studies tend to establish a difference in representation and processing between these two categories. Lexical words also tend to be longer and represent an open inventory, constituting a larger set than that of function words. Therefore, the competition among the possible lexical words in a given position within a sentence tends to be considerably greater in most cases. From the lexical access standpoint, studies show that processing is easier for function words. This difference is attributed to certain factors, such as the word predictability and familiarity, and may be affected by frequency (SEGALOWITZ; LANE, 2000) or by the word being presented separately or in a meaningful context (TAFT, 1990). Acquisition studies reveal that children start the production process by using lexical words, which is the most prominent type of word. Nevertheless, where perception is concerned, studies reveal that children are sensitive to the presence of function words and that they have expectations regarding these words' distribution within the sentence (VAN HEUGTEN; CHRISTOPHE, 2015). As children acquire function words, they define distribution patterns for co-occurrence with the lexical word category. The acquisition of function words, therefore, plays a relevant role in language acquisition, contributing to the recognition process of the grammatical categories in the language (BRAINE, 1987; BRUSINI *et al.*, 2017, 2016). The literature discusses whether this contribution is the result of a linear distribution analysis or the result of a hierarchic structural relation among the phrase elements. Jointly, these studies attest to the different natures of function and lexical words, which have an impact on language acquisition and in processing by adults.

According to sociolinguistic studies, linguistic and non-linguistic variables act in an integrated manner. Such studies indicate that non-linguistic variables may favor or inhibit the use of a given linguistic form. The variable level of education, one of the non-linguistic variables addressed by this study, has been widely tested and studied in order to observe if such a level has any influence on speakers regarding the use of the standard variety of Brazilian Portuguese. Gonzales *et al.* (2007) stated that the level of education variable influences the use of different linguistic phenomena and that such an influence may be related to the social role played by the school.

It is known that

school leads to changes in the speech and writing of people who attend it and of discourse communities. It is determined, however, that it acts as the preserver of prestigious forms, in face of ongoing changing trends in these communities. (VOTRE, 2012, p. 51).

Pereira and Araújo (2016) investigated the influence of the education level variable on verb and subject agreement variation for the third person plural in Brazilian Portuguese. The authors noticed that most studies show that the prestigious variety is more frequently used in records of individuals with a higher education level than in those of individuals with a low education level. The authors state, in agreement with Cyranka and Pernambuco (2008), that being deprived of formal education prevents the individual from using certain prestigious linguistic forms, and, consequently, from circulating in several communicative situations in which these forms are used.

Parente *et al.* (2009), in a study about the evidence of the role played by education level in brain organization, have observed that there are behavioral changes, performance changes in neuro-psychological evaluations, and structural-functional changes detected in neuro-imaging exams of the human brain structure and functionality, according to the education level of the individuals tested. For the authors, education level is presented as a multidimensional variable, since factors such as reading and writing habits developed by individuals, due to personal interests, lifestyles, or professional needs, are also part of such a variable.

This study also manipulated another non-linguistic factor: text length. According to Kato ([1986], 2001), extracting meaning from a statement depends on the elements present in an individual's memory and on how memory operates. In reading a text, an individual activates different types of memory – short, medium, and long term. The relation between the type of memory activated and the text content determines if the reading strategy will be predominantly top-down – when the reader starts from previous knowledge to build the text meaning – or bottom up – in which the reader recognizes word by word, combining them to build larger units and generating meaning from the combination of these units.

Researchers state that short texts tend to favor a bottom-up reading, since a minimal memory requirement would allow for greater attention to details and would encourage word-by-word processing strategies (KINTSCH; VAN DIJK, 1978; SWAFFAR; ARENS; BYRNES, 1991). In addition, short texts are also seen as potentially more difficult, due to the reduced amount of information, while long texts would be more easily understood, given that they provide more content by means of which the meaning may be understood. However, the attention factor must be taken into account when considering the role played by text length in reading comprehension and in students' performances in reading texts. The experience of teachers in the classroom indicates a greater difficulty in concentration and a compromised performance in reading long texts, when compared to short texts. This finding is legitimated by the preponderance of relatively short texts in textbooks.

The next section presents a literature review of the Cloze test as a measurement tool in reading comprehension. The subsequent section presents an experiment that quantitatively manipulates the variables of linguistic and non-linguistic nature listed above, determining the sensitivity of the test to these variables. Finally, a discussion is presented on how the linguistic and non-linguistic variables may interact with reading proficiency, proposing a research agenda for future works, which may benefit from a qualitative approach to extracting items in Cloze tests.

2 The Cloze test

The term Cloze is derived from the English word closure (COLLINS, 2009), which, in turn, is the name given to a principle in Gestalt psychology. This term was used by Wilson L. Taylor in 1953 to name a procedure designed to measure the text's readability, which is based on Information Theory, Gestalt psychology, and the statistical notion of a random sample. In a seminal article, Taylor (1953, p. 416 *apud* ADELBERG; RAZEK, 1984, p. 111) defined Cloze as

a method of intercepting a message from a transmitter (writer or speaker), mutilating its language patterns by deleting parts, and so administering it to receivers (readers or listeners) that their attempts to make patterns whole again potentially yield a considerable number of cloze units.

Taylor (1953, p. 416 *apud* ADELBERG; RAZEK, 1984, p. 111) goes on to explain that a Cloze unit is

any single occurrence of a successful attempt to reproduce accurately a part deleted from a 'message' ... by deciding, from the context that remains, what the missing part should be.

Therefore, in an experiment using the Cloze technique, Information Theory principles may be detected in the idea that the sender and the receiver must share the same set of possible messages, Gestalt psychology principles may be detected in the idea of having the reader reconstruct a message and reach its understanding, as well as by filling in the blanks with previously suppressed words. Thus, the statistical notion of a random sample may be detected in the idea that different classes of words gain proportional representativeness in the gaps opened in the text. That is, the probability of choosing lexical words, for instance, would be proportional to the choice of function words. On this aspect, Adelberg and Razek (1984, p. 112) comment that the Cloze test appears to enable the measurement of both a lexical component and a functional component, depending on the type of word selected for suppression. Condemarin and Milicic (1988), however, argue that, although Taylor had based the Cloze technique on Gestalt, it is also considered to be a cognitive instrument by cognitivist scholars, since they assume that such a technique deals with cognitive processes, such as perception and memory, and see in Cloze exercises an interactive process between the reader and the text.

In its original format, the Cloze test provides for the selection of a text with approximately 250 words. After every five words, the fifth one is suppressed systematically. These words are replaced with blanks whose sizes are proportional to the words' sizes. Several studies, however, have used varied methodological forms to build Cloze tests. One of the proposals for changing the original version, which has already been able to bring together a significant number of researchers favorable to it, is based on the assumption that, to test specific aspects of textual comprehension, precise elements should be suppressed. According to Sadeghi (2014, p. 1), this proposal has already promoted the Cloze variety entitled "rational Cloze", as opposed to the "random Cloze". In

the rational Cloze test, omissions are based on determined criteria (non-mechanical), while in the random Cloze test, they are done randomly (mechanical).

Scholars who support the rational Cloze (ABRAHAM; CHAPELLE, 1992; SADEGHI, 2014) argued that, if the test purpose is to measure text comprehension, randomized suppression must be discarded in favor of rational suppression. It is possible to observe, therefore, that, depending on the application purpose of the Cloze test, different types of rational Cloze may be constructed, such as the lexical Cloze – in which lexical items of nouns, adjectives, verbs, and adverbs are omitted – and the grammatical Cloze – in which function items of prepositions, articles, and conjunctions are omitted (Santos *et al.*, 2009).

Another Cloze change proposal is related to the type of answer. In the classical version, answers are open-ended, that is, the participant may choose the word to fill in the blank correctly. In modified versions, answers are close-ended, that is, the participant will have to choose, among the offered alternatives, the one to fill in the blank correctly. In this case, we have the multiple choice Cloze and its variations, such as the maze Cloze, the paired Cloze, and the restricted Cloze (OLIVEIRA *et al.*, 2009). Supporters of this proposal argue that the use of close-ended questions makes the test more practical in terms of scoring.

Finally, this study observes the existing variation in the application model, which may be a collective application – as is the use of the Cloze test as a pedagogic tool by teachers in one or more classes – or an individual application – as is the use of the Cloze test as a diagnostic tool by psychologists in a patient's assessment. Oliveira *et al.* (2009) show us that the Cloze post-oral reading and the interactive Cloze are examples of the individual application model. The choice of an application model will always depend on the study's origin and on the professional who will use it.

According to Santos *et al.* (2007), there are two correction forms for the Cloze procedure, literal correction and synonymic correction.

The literal correction considers the right answer to be the correct completion of the exact word that was omitted, respecting the correct spelling and accent marks. The synonymic correction considers the correct answer to be the completion of the blank not only with the excluded

word, but also with a synonym of the omitted word. (SANTOS *et al.*, 2007, p. 42).

These same authors, however, in a 2009 article, present a third type of correction, the weighted correction. In this form, the evaluator weighs the word written by the participant, checking whether the word used is close to that expected or not, but the evaluator is not concerned with spelling (OLIVEIRA *et al.*, 2009).

Regardless of the type of correction used, the commonly used scoring system assigns one point per right answer. The participant's final score is obtained by the sum of points, which may vary between 0 and 25 points or between 0 and 50 points, or even between 0 and "n" points, depending on the existing number of blanks in the text. However, Santos *et al.* (2002, p. 550) warn that

[t]his interdependence between the scores and the test difficulty level is one of the limitations of classical psychometrics, in which scores are calculated by assigning the same weight to heterogeneous items in terms of complexity.

Therefore, a different scoring system measures the item's difficulty level. In this case, different values are attributed to each item, according to its degree of complexity. These values are defined by Portuguese language specialists, named "judges", who, after examining the suppressed items, indicate the difficulty levels, namely, easy, medium, moderate, and difficult. Thus, a variability is witnessed in the sum of the points, since items classified as easy are given fewer points, unlike items classified as difficult, which are given more points.

For the score reached by the participant, statistical analyses may be adopted to verify the influence of variables and the significance or insignificance of the statistical results. Another option is adopting the classification parameter adopted by Bormuth (1968) which establishes two reading levels: the instructional level and the independent level. The instructional level corresponding to a right percentage within the range between 45% and 57% of the text indicates that the reader has achieved sufficient comprehension, but additional external assistance (from the teacher or the textbook, for instance) is still required. The independent level corresponds to the right percentage above 57% of the text, and

indicates that the reader has reading comprehension autonomy. Santos *et al* (2002) have added yet another level, named the frustration level, which corresponds to the right percentage up to 44% of the whole text and indicates that the reader extracted little information from the reading, that is, was scarcely successful in understanding it. Santos *et al.* (2002, p. 550) also state, about the classification proposed by Bormuth (1968), that this “application to new texts may only be done if these texts have a difficulty level similar to that present in the texts used by Bormuth in his studies”.

Therefore, based on the Cloze description provided here, it may be concluded that this is a type of experiment whose preparation is economical and whose application is simple (pencil and paper); however, this is an experiment that requires careful selection of texts in defining the blanks and careful correction/scoring.

Regarding this topic, Suehiro (2013, p. 224) warns that:

Regarding text structuring, in addition to word omission parameters, the question of difficulty and of the items or words omitted must also be taken into account. Regarding difficulty, the topic addressed by the text is also an aspect to be considered, given that performance differences have been identified in association with this variable; therefore, under this perspective, the familiarity with the topic increases the probability of right answers provided, which introduces an intervening variable in the comparison among the performances achieved by individuals when the technique is applied (COHEN, 1975; OLIVEIRA; BORUCHOVITCH; SANTOS, 2009; PAGE, 1975).

For this study, a Cloze test was performed, whose purpose was to investigate if the processing of a text during reading may be modulated by a linguistic factor, such as the difference between lexical and function words, and by non-linguistic factors, such as text length, which may have a subsequent impact on processing factors, such as attention and working memory. The purpose of the word type variable is to determine the role played by linguistic factors that may be intrinsic to the item or related to its distribution regarding other items in constructing a proposition's

meaning. In addition, this study also verifies how these factors may be affected by education level. For such a purpose, the test was conducted with two different groups: 6th grade students and 10th grade students, since studies indicate that the age/education level is a factor to be taken into account, regarding linguistic knowledge.

3 The use of the Cloze test

Widely used in the educational sphere, the Cloze test is a reading comprehension diagnostic tool. This tool has been used by different researchers, both in basic approaches to the cognitive processing of sentences, and in more applied dimension studies, with practical implications on language teaching (FEDERMEIER *et al.*, 2007; SANTOS, 2004; BALIZA; SILVA, 2015). At first, one might think that this tool is limited to vocabulary checking, but the test may also be used in loose sentences, in which the verification is local, or in texts, in which the verification is macrostructural.

Since it was created, the Cloze technique has been considered a reliable resource to evaluate written text comprehension and has been used in a wide variety of research on readability studies, as well as on studies that seek to evaluate the influence of the syntactic position of words within the sentence. In addition to these applications, Cloze has also been used in foreign language proficiency exams and entrance exams at several universities (CUNHA, 2009). Nowadays, this test is being used by reading and language researchers, from a multidisciplinary standpoint, which includes areas such as Psychology, Communication, Speech Therapy, Linguistics, and Education, containing two main branches: one branch focusing on the evaluation of psycholinguistic mechanisms that affect the written language comprehension and another branch focusing on the establishment of a diagnosis of difficulties associated with written language comprehension (SANTOS *et al.*, 2009).

A study developed by Baliza and Silva (2015), using the Cloze technique, observed the influence of the education level on reading comprehension, among other factors. This test was applied to 7th and 8th grade students. The authors observed that the students who had a higher education level understand what they read better, when compared to other studies.

Other studies combine the Cloze technique and academic performance perception. Oliveira, Boruchovitch & Santos (2009) investigate the relation among reading comprehension, actual academic performance, and self-perceived academic performance in 5th grade students and verified basic level reading comprehension skills. They demonstrate that, although students at this education level achieve the minimum learning ability, they do not reach the reading proficiency level. Upon observing the relation between academic performance in the Portuguese Language discipline and reading comprehension, the authors verify that students with better performance in the discipline also exhibit better reading comprehension skills. The study also reveals that students who have a negative self-perception of their performance achieved a lower score in the Cloze test, and that the same fact occurred inversely, which may suggest that students are aware of their own academic performance. Finally, they point out the need for intervention proposals to be adopted to prepare the diagnosis and the remedy of specific difficulties in reading comprehension.

Similarly, focusing on reading comprehension in elementary school, Joly and Piovezan (2012) engaged in a study designed to evaluate the *Programa Informatizado de Leitura Estratégica* (PILE - Strategic Reading Computer Based Program) developed by Joly (2008), a program designed to promote reading comprehension in students from the 5th to the 9th grade. The program content consists of a series of story sections in which the Cloze Oriented System (JOLY, 2006, 2009) was applied and in which the difficulties associated with the number of words that appear between the gaps and the additional clues, such as blank space size and word bank, are shown in a gradual sequence, using the Cloze Oriented System (COS) as a text organization criterion based on difficulty the level. The results obtained not only prove the Program's effectiveness, but also point towards the importance of the intervening proposal in order to improve reading comprehension performances. Furthermore, they corroborate the effectiveness of COS in intervention programs (JOLY, 2007, 2009), as well as the Cloze procedure as a diagnosis and remedy technique, as previously demonstrated by Santos *et al.* (2002), Santos (2004), and Cunha (2009).

Joly *et al.* (2014), in continuing studies on reading comprehension, prepared a reading comprehension assessment study for students in the 6th to the 9th grades, using the COS. In this study, variables such as gender,

age, and education level were taken into account to verify if there were differences in reading comprehension. The result analysis indicated a significant difference between genders – since girls performed better – and between education levels – given that students in higher grades exhibited a better performance. Results did not prove to be significant regarding age.

Suehiro and Boruchovitch (2016) also conducted a study directed toward elementary school students, specifically with students from the 6th to the 9th grade, and used gender, education level, and age as variables. After the analysis, the results revealed that the participants' reading comprehension would be rated as basic, which was not compatible with the education level. Similar to the data presented by Joly *et al.* (2014), there was a significant difference between the genders, since girls performed better, as well as among the education levels, given that students at higher grades exhibited better performances. The discrepancy between the studies was shown in the age variable, since, in Joly *et al.* (2014), this aspect did not prove to be significant, unlike the description provided by Suehiro and Boruchovitch (2016), in which older students demonstrated a better performance.

One study, conducted by Oliveira, Cantalice & Freitas' (2009) involving high school students, exhibited similar purposes to those previously presented, despite having the specific goal to explore right answers for the items. The issue of items is particularly important in the Cloze test, considering that the accuracy in completing the omitted item possibly does not only depend on the participants' reading ability, but also on the difficulty of the item to be completed.

Therefore, in this study, the items consisted of eight word classes (pronoun, adjective, adverb, article, conjunction, preposition, noun, and verb) and one sub-class (adjective phrase), and they were classified as easy, of medium difficulty, and difficult. The results indicated that students possibly demonstrated difficulties in reading and understanding what they were reading, and that this reading comprehension is below the basic level, which was an unexpected result for this education level. The authors emphasized that these results are exploratory and pointed out that, regarding the item difficulty, the results showed that the categories with higher right answer rates were prepositions and nouns, while the one with a lower right answer rate was the adjective phrase. The

authors emphasize that these data are exploratory and point towards the convenience of performing a more accurate analysis.

In consonance with what this study intends to present, the studies mentioned above took linguistic and non-linguistic variables into account. It is important to point out, however, that, if, on the one hand, the set of works reported is far from consisting of an exhaustive list of studies with Cloze exercises in elementary and high school, on the other hand, they illustrate the acceptability and the credibility achieved by this psycholinguistic procedure in the academic sphere.

The technical relevance of Cloze resides in its contribution to investigating reading comprehension and performance problems at all education levels. Its usefulness may be observed in the number of publications dealing with its use, whether working as reading comprehension diagnostic element or assisting in the development of intervention and remedy programs, as proven by Suehiro's (2013) data in describing the scientific production about the Cloze test in psychology journals between 2002 and 2012.

4 Accessing the Cloze test sensitivity to linguistic and non-linguistic factors

This experiment was designed to determine the sensitivity of the Cloze test to linguistic and non-linguistic factors. For such purpose, the variables of word type (function and lexical), text length, and participant education level were manipulated. We adopted the rational procedure (ABRAHAM; CHAPELLE, 1992; SADEGHI, 2014) to create the blanks, given that the type of removed word (function or lexical) was manipulated. Therefore, in a given text, there will be a higher number of blanks for lexical words (focus) and a lower number of blanks for function words. Another given text will present the opposite, a higher number of blanks for function words (focus) and a lower number of blanks for lexical words.

In short, the independent variables tested were: (i) word type, which was divided into function and lexical words, and (ii) text length, which could be short or long. Only one dependent variable was used in this test: the Cloze test error rates were measured. **The independent variables constitute a 2x2 design, generating the conditions of Long**

Text with Function word (LTF), Long Text with Lexical word (LTL), Short Text with Function word (STF), and Short Text with Lexical word (STL) (Table 1).

TABLE 1 – Experiment design

6 th Grade			10 th Grade		
	Function	Lexical		Function	Lexical
Long Text	LTF	LTL	Long Text	LTF	LTL
Short Text	STF	STL	Short Text	STF	STL

Based on the data obtained by this experiment, we expected a significant performance difference among the groups regarding text length and word type. In relation to text length, it is expected that more experienced readers would be less affected by the attention difficulty caused by long texts. Regarding word type, Cain and Nash (2011) indicate that function word reading skills are modulated by the readers' age.

For the word type separately, we expected greater difficulty in completing blanks with function words, given that, considering factors such as word distribution and co-occurrence, it is probably harder for participants to access a function word based on a lexical word. This results from the possibility of function word options occurring close to a lexical word being more restricted than the lexical word options occurring close to a function word. Therefore, even if the predictability of a function word is greater in a given blank, if compared to a lexical word, due to the paradigm limited repertoire, the difficulty can be explained exactly by this limited number of options. Since the protocol used in the experiment accepts synonymous answers or answers that are coherent with the text meaning, a student is more likely to recognize a possibility within a larger repertoire than within a smaller one.

In relation to the long text variable, when considered separately, the literature suggests that the reading comprehension performance may be facilitated by text length, since longer texts would provide more information and clues for interpretation. However, it is believed that the

role played by factors such as attention and motivation may significantly affect the reader's performance, which would lead to the prediction of greater difficulty in understanding longer texts, especially in the group with a lower education level.

Participants

This study included 88 students from a public school in the state of Rio de Janeiro, located in the city of São Gonçalo, who had regular attendance, 44 of which were 6th graders, while the other 44 students were 10th graders. The average age of 6th-graders was 12 years of age, considering the standard deviation of 1.38, while for 10th graders the average age was 17 years of age, considering the standard deviation of 0.74.

Material

The test material consisted of texts extracted from textbooks corresponding to the grades attended by the research participants. Fictional texts, of the chronicle genre, were selected, given that this is one of the first textual genres with which students have contact in elementary school, and considering Joly's (2014, p. 239) recommendation, stating that:

The numerous Cloze technique variation possibilities favor the use of this technique with different content texts, whether specific and associated with school disciplines, or of literary nature, such as those used in the present study.

Furthermore, it is a more everyday and confessional genre, which deals with routine topics. With a linguistic structure close to that of the narrative and with simple language, this genre is viewed as more accessible not only to all types of readers, but also to all age brackets in general. The texts selected to make up the versions of the Cloze test were three chronicles by Carlos Drummond de Andrade, entitled "*O assalto*" (The robbery), "*No aeroporto*" (At the airport), and "*Recalcitrante*" (Recalcitrant), together with a travel log entitled "*O segredo de Mioko*" (Mioko's secret) by Marco Clanflone. The criteria used for selecting these texts also took into account the observations made by Coelho (2000) on

text types and text genres that are appropriate to students, according to their age group and grade.

The text “*O assalto*” (The robbery) had 218 words and 25 blanks; the text “*No aeroporto*” (At the airport) had 515 words and 50 blanks; the text “*Recalcitrante*” (Recalcitrant) had 629 words and 50 blanks; and the text “*O segredo de Mioko*” (Mioko’s secret) had 394 words and 30 blanks. The blanks were intentionally opened and the marked space was the same for all blanks.

Procedures

After the initial contact with the school administration, when the formal request to conduct the research was made, a cooperation agreement was signed. The school administration set the days and times for the classes that would be the most appropriate for the experiment application. The application was, therefore, conducted collectively, in the classroom, in compliance with the schedule provided by the school.

Prior to the start of the test, the researchers explained the entire procedure clearly and answered any possible questions about the task to be performed. The texts were handed out to the students, in a classroom situation, by the researchers. The tests were applied at different times during a regular class day, among other activities included in the teachers’ lesson plans for the different groups. The participants were handed a text containing one of the four conditions tested (LTL, LTF, STL, STF) and were expected to fill in the blanks with a word that ensured text coherence. In addition, the participants answered identification questions, providing the following information on the Cloze test sheet: name, age, and class number.

The test was conducted in a quiet, well-ventilated and well-lit environment. The activity’s average duration was 50 minutes per classroom, although a time limit for students to complete the task had not been set.

5 Analysis

Correction criteria

Two criteria can be adopted traditionally for the correction of items in the Cloze test: the literal criterion (or objective) and the synonymous criterion (or subjective). According to the first criterion, the items that correspond to the words used in the original text are considered to be correct. According to the second criterion, nevertheless, the items that ensure content coherence are considered to be correct. In addition to these criteria, the weighted criterion (OLIVEIRA *et al.*, 2009) is also used, in which the items that approach the completion ideal are considered to be correct, regardless of spelling or grammatical accuracy.

In this study, performance assessment was conducted based on the weighted criterion and consisted of scoring one point for each correct answer. Words that completed the sentences coherently were considered to be correct. To avoid excessive subjectivity in the correction, the answers were assessed by the researchers and submitted to the evaluation of judges (FONSECA *et al.*, 2011), that is, Portuguese language specialists who examined and defined the difficulty level of the items (OLIVEIRA *et al.*, 2009). The maximum number of correct answers was 50 (long text) or 25/30 (short text), corresponding to the number of omissions in each text version.

Data analyses

At first, each tested individual's correct and incorrect answers were counted. After data for each participant had been verified, the data was then counterbalanced in such a way that all conditions were quantitatively equivalent for statistical analysis. All data were equaled to 50 in order to match the data in the long text. Therefore, for a total of 50 blanks, the number of incorrect answers provided by student 1, for instance, under the STF condition was 24, which corresponded to 12 errors out of 25 answers. Such a procedure was repeated for the data from all students, in both elementary school and high school, to allow for a more homogeneous and reliable analysis.

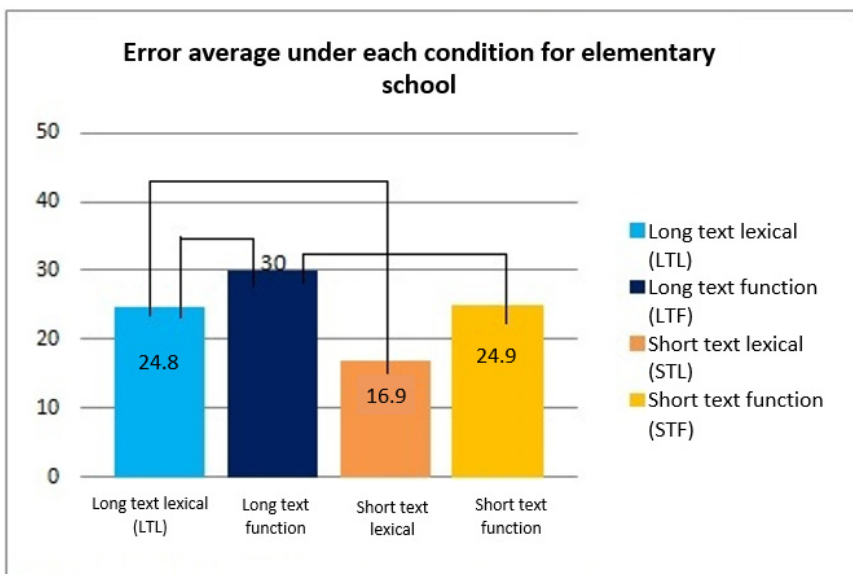
After this step, the statistical analyses were performed using the R computer program. Two tests were conducted: T test and ANOVA. Two types of analyses were conducted: the Within Subject analysis, in which data were compared per group, and the Between Subject analysis, in which the data were compared between the tested groups.

6 Results

As mentioned above, this work was designed to verify the sensitivity of the Cloze test to the linguistic factors of text type (long or short) and word type (lexical or function), and to non-linguistic factors (education level). The test was conducted with 6th graders and 10th graders.

Chart 1, which shows the error average in filling in the blanks for elementary school students, reveals that students had quite high error averages under all conditions. In a comparison among the conditions tested, it is possible to observe, visually, that students encountered more difficulties in filling in the blanks than expected when they required a function word and when the texts were long. Regarding the word type (lexical or function), students presented a higher number of errors when function words were involved, regardless of the text length.

CHART 1 – Filling in the blank error averages



A statistical analysis (T-test), using error averages for each condition, revealed that the comparison between the LTL and STL conditions is highly significant, reaching $t(10)=9.58$, $p<0.0001^{***}$. It also revealed that, the STL and STF conditions were significant, reaching $t(10)=4.81$, $p<0.0007$. The comparison between the LTL and LTF conditions also revealed a statistical significance, reaching $t(10)=2.65$, $p<0.02$. The significance in the comparison between LTF and STF proved to be minimal, reaching $t(10)=3.03$, $p<0.01$.

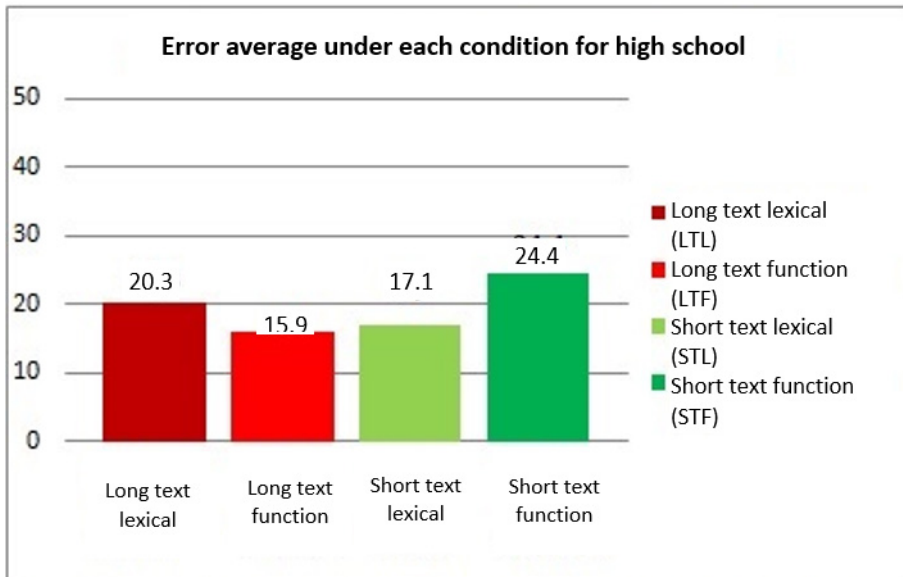
The ANOVA significance test resulted in highly significant values for the Text Type variable $F(1.10) = 73.2$ $p<0.000007$ as well as for the Word Type variable $F(1.10)=20.5$ $p<0.001$. Therefore, the Within Subject analysis conducted with elementary school students led to the conclusion that the Word Type and Text Length variables, put into question in the present article, have great influence on the students' performance.

The same test was conducted with 10th graders, and the results obtained were different from those found for the group of 6th graders. The STF condition showed the highest error rates (Chart 2).

Different from that which occurred with elementary school students, error averages under each condition for high school students were very similar.

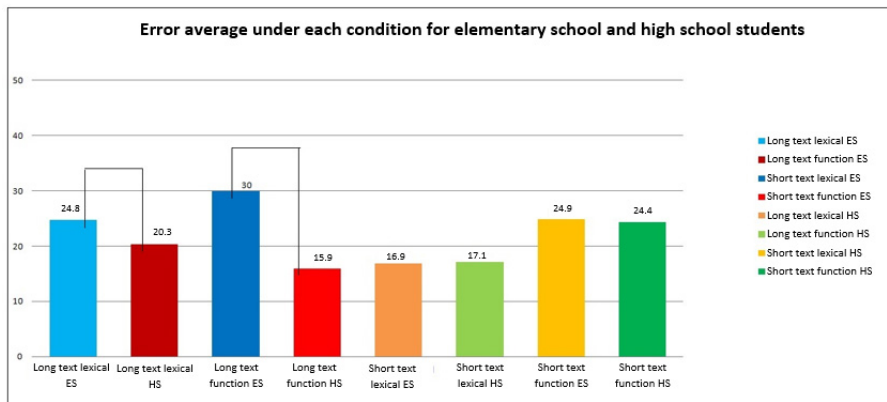
A statistical analysis revealed that the comparison among the conditions did not present significance, since the error averages were very similar under all conditions.

CHART 2 – Filling in the blank error averages for high school



In the Between Subject statistical analysis (the comparison between the two tested groups – elementary school and high school students), the statistical analysis (T-test) was conducted first. The T-test revealed that the comparison between the long text conditions, under which the students were to fill in the blanks with function words, was significant ($t(20)=4.43$, $p < 0.0003$), and the elementary school students presented the highest error averages. The result was not significant in long texts in which students were to fill in the blanks with lexical words, reaching $t(20)=1.93$, $p < 0.06$. However, there is strong indication of a trend, showing that the application of the test with a higher number of students is relevant.

CHART 3 – Filling in the blank error averages for both groups



The ANOVA significance test (Chart 3) presented significant results for the variable word type $F(1.80)=6.31, p<0.01$, function words presenting the highest error rates, and for the variable education level $F(1.80) = 8.65, p<0.004$, elementary schools presenting the highest error rates.

7 Discussion and research agenda

According to Kato ([1986] 2001), written language comprehension involves an array of knowledge types. In a first analysis, written language may be understood as the transcription of a language. In this sense, after the code is understood, reading proficiency should be as effective as language comprehension. However, written language comprises rules and specificities that transcend what would be a simple language transcription and requires cognitive factors, such as access to short and long-term memories and the ability to recognize and establish cohesive relations among different significant units. These units tend to vary in complexity according to the textual genre and demand a higher level of formal effort as these genres distance themselves from the reader's vernacular language.

Therefore, the measurement of reading proficiency requires the ability to measure these different aspects in a controlled manner. The experiment described in the previous section showed that the Cloze test might be the way to manipulate different aspects related to reading comprehension. This study detected differences between linguistic nature aspects (difference between function and lexical words) and between non-linguistic nature aspects (text length and education level).

The experiment was designed so as to be as faithful to the Cloze test creation patterns as possible, using the rational procedure to simply extract function and lexical words, without a qualitative regard on which function items and which lexical items would be extracted. The results revealed a difference regarding the variable of word type, which has the merit of showing that the Cloze test may be sensitive to this linguistic nature variable. However, from the reading processing standpoint, this sensitivity becomes more relevant if it is possible to evaluate what this difference can tell about reading comprehension, such as, for instance, determining the type of cognition involved in reading that was accessed to fill in the blanks.

Due to the multifaceted nature of the processes involved in reading, measuring the relation between comprehension tests and the cognitive processes evaluated by them is not a simple task. Specifically regarding the cognitive processes involved in the resolution of the Cloze test, there is a debate regarding the test primarily accessing a contextual global knowledge or an immediate local processing – which would activate the previous knowledge – in preparing the answers (SANTOS *et al.*, 2002). As Sousa and Hubner (2015, p. 37) point out, “although [the Cloze test] exhibits an extensive correlation with reading comprehension tests, there are questions about what it measures exactly”.

A brief qualitative analysis of some items in the texts used in the reported experiments is provided below, to identify the nature of the participants’ difficulty. The goal is to show that there are specificities in each item that may access different aspects of the reading processing. Each error that occurs both in function and in lexical words was analyzed qualitatively, item by item. This type of analysis reveals the specificities in each item, showing that the skills required to fill in a blank may have different natures, even if they are two items of the same kind (lexical or function). This type of analysis is useful to inform future researchers, so that original experimental designs may be created in which it is possible to

specifically manipulate the type of cognition involved, accounting for the possible intervening variables that are present in a category distinction, such as function and lexical. The following examples show some aspects that may be explored based on these analyses and can be manipulated in a controlled manner in future tests:

- 1) The blank in “*sacar do _____*” (draw the _____) was filled in by participants with the lexical items *cão, carro, bolso, cofre, homem, balcão, banco e caixa* (dog, car, pocket, safe, man, counter, bank, and box).

In example 1, the original text item is *revólver* (gun), a lexical word. As observed, the student has difficulty filling in the blank when only the phrasal level is considered, and is successful in recognizing the syntactic and semantic restrictions of the structure, which would license an interpretation “remove from x”. However, in the idiom “*sacar do revólver*” (draw the gun), the structure is different, interpreting “take x”, in which the word *gun* is part of the object of the verb *draw* and plays a different thematic role from any option used by the students participating in the test. This item, therefore, evaluates an erudition not expected from the common speaker and does not exactly reveal a problem in linguistic processing, since participants were successful in filling in the phrase with possible items. The use of these possible items reveals, however, a difficulty by the participant in terms of reading proficiency. That is, even if they are all possible items to complete the phrase, and even if they make sense within the semantic field of the text, some of these lexical items would not be acceptable, taking the construction of coherence into account in a more macrostructural analysis.

Therefore, the participants do not seem to have problems in using their linguistic knowledge to fill in the text blanks, which reveals that the difficulty is related to factors extrinsic to the language, possibly associated with the textual structure or the ignorance of certain linguistic structures foreign to their language. A qualitative analysis thus enables one to separate the type of error made by the student so that specific pedagogical actions can be performed.

A preliminary analysis of frequent errors indicates that, in general, inexperienced readers do not demonstrate problems in completing simple phrases. Furthermore, they do not demonstrate great difficulties

in identifying the predominant text topic, a skill that only depends on the activation of a semantic field by means of a set of lexical items, without necessarily having a cohesive structure among them. However, inexperienced readers face the challenge of identifying larger significant units, such as clauses or complex phrases – which leads to difficulty in generally understanding the text as a cohesive unit.

It is also noteworthy, in the analysis of errors made in filling in the blanks with function words, that the occurrence of errors was greater in structures with greater structural complexity or which depend on a larger vocabulary (examples 2 and 3 below), demanding a higher literacy level – an observation that is legitimized by the difference encountered between the groups with different education levels.

- 2) Structures with right answers: “um homem”; “o jardim”; “da casa” (a man; the garden; of the house).
- 3) Structures with wrong answers: “ele astúcia” (he slyness) (rather than “de astúcia” (of slyness)); “ele súplica” (he plea) (rather than “a súplica” (the plea)); “disse que era da saúde pública ele indagou” (said that he was with the public health agency he asked) (rather than “disse que era da saúde pública e indagou” (said that he was with the public health agency and asked)); “prometeu a nunca mais tentaria assaltar” (promised to never again would try to rob) (rather than “prometeu que nunca mais tentaria assaltar” (promised that he would never try to rob again)).

To fill in the blank in “ astúcia” (slyness), the participant must know the lexical item “astúcia” (slyness), with low frequency in this age group. Even less frequent is the possibility of combination with the preposition “de” (of) to form the expression “de astúcia” (of slyness). In addition, it is possible that, by not recognizing the lexical item stem, the participant may have recognized the final sequence of letters in the word as verb inflection, which explains why one might fill in the blank with the item “ele” (he), which would function as the subject of the alleged verb. In the case of “súplica” (plea), the participant may have failed to recognize the difference between the verb form (not accentuated) and the noun form (accentuated) of this word. Once again, if these are the cases, the error in filling in the blanks results from not recognizing a lexical

item, rather than from a failure in linguistic processing. In “*prometeu a nunca mais tentaria assaltar*” (promised to never again would try to rob) the participant failed to recognize the subordinate clause being used as internal argument for the verb “*prometeu*” (promised), and used a preposition “*a*” (to) rather than the conjunction “*que*” (that). The structures present in example 2, however, although dealing with function words, are simpler, and the content words are not sufficiently known by elementary school students.

The superior performance by high school students may be due to the longer time of exposure to more complex structures, encouraged in the schooling process, as well as by the increased vocabulary, which may be a factor related to both age and educational level.

We emphasize the relevance of associating the Cloze tests with chronometric measures, which enable accessing the item difficulty level in a more controlled manner. For instance, it is assumed that experienced readers would also possibly find it difficult to complete the blank presented in example 1 with the lexical item “*revólver*” (gun), even if, given time to think, they would end up answering correctly.

In view of this brief analysis of specific items, it is possible to determine that the Cloze test may be a useful tool to evaluate different aspects of reading proficiency. This brief analysis indicates that the test, if developed consciously and using the rational process with open-ended answers, can be useful in revealing different cognitive processes, for instance, if the participant would be accessing contextual global knowledge or previous knowledge (more local and immediate). This is only possible, however, if the test is developed in an informed and qualitative manner, controlling the definition of the blanks more to access the knowledge being evaluated, but less in terms of binary categories, whether linguistic or not. Therefore, it is possible to more accurately identify if the error occurs due to problems with specific linguistic knowledge, due to reading proficiency problems, or even by a blank in the linguistic literacy process. The error of each nature requires different pedagogical actions.

8 Conclusion

This study is designed to discuss the possibilities of using the Cloze test as a tool to measure aspects related to reading proficiency, determining its sensitivity to linguistic nature and non-linguistic nature aspects, and entertaining the possibility of creating original and qualitative designs, so as to identify the aspects evaluated by each omitted item in a more granular manner.

A literature review was presented, in which the study sought to describe the Cloze technique itself, from its traditional version to versions with procedural variations. The Cloze test was also presented in national productions designed for elementary and high school.

The test sensitivity was measured by means of a study that manipulated the variables of word type, text length, and education level. The experiment involved two groups of participants, one of them consisting of 6th graders, and the other of 10th graders. The results obtained for the 6th-grader group confirmed the hypotheses that participants would encounter more difficulty to understand the long text, since they would have to concentrate for a longer period of time, and would encounter more difficulty to access function words, which may easily be explained by the students' age group and the time of exposure to these items.

For the participants in 10th grade, the initial hypothesis was that participants would find the short text easier. However, the Cloze test result error average was higher in the Short Text (24.4) than it was in the Long Text (20.3). Even if it is not a significant difference, the result was unexpected. A possible explanation is that the long text has provided more information to the participant, who had more background to access and fill in the test blanks more accurately.

In the between group comparison, the best result was found in the 10th-grader group. This fact was already expected, given not only the time period that separates both groups, but also the number of school years, considering that the 10th graders have spent three more years in contact and interaction with the academic environment.

Finally, the present study highlighted the importance of developing Cloze tests that enable one to access specific aspects related to reading proficiency, so as to identify what is being evaluated by the test. Therefore, it is possible to identify and work on the reading problem nature exhibited by the participant. In this light, the Cloze test may be

consolidated as an important and versatile tool, capable of accessing reading proficiency at more local levels, which may reveal linguistic processing problems or a mere unfamiliarity of lexical items, or that may reveal literacy gaps in reading at macrostructural levels.

Acknowledgments

The authors wish to thank the CIEP 246 team and Teacher Adalgisa Cabral de Faria for authorizing the development of the research. The authors also wish to thank their colleague, Professor Marília Costa (UFRJ) who dedicated time to read the previous version of this study, as well as the anonymous reviewers for their varied suggestions.

References

- ABRAHAM, R. G.; CHAPELLE, C. A. The meaning of Cloze test scores: An item difficulty perspective. *The Modern Language Journal*, Wiley Online Library, v. 76, n. 4, p. 468-479, 1992.
- ADELBERG, A. H.; RAZEK, J. R. The Cloze Procedure: A Methodology for Determining the Understandability of Accounting Textbooks. *The Accounting Review*, American Accounting Association, v. 59, n. 1, p. 109-122, Jan. 1984.
- BALIZA, A. A.; SILVA, D. V. Avaliação da Compreensão em Leitura em Estudantes do Ensino Fundamental. *Educare, Revista Científica de Educação*, UNASP, v.1, n.1, p. 93-114, 2015. ISSN 2447-5432.
- BORMUTH, R. J. Cloze test readability: criterion reference scores. *Journal of Educational Measurement*, Wiley Online Library, v. 5, n. 3, p. 189-196, 1968.
- BRAINE, M.D.S. What is learned in acquiring word classes: A step toward an acquisition theory. In: MacWHINNEY, B. (Ed.). *Mechanisms of language acquisition*. Hillsdale, NJ: Erlbaum, 1987. p.65-87.
- BRUSINI, P.; DEHAENE-LAMBERTZ, G.; VAN HEUGTEN, M.; DE CARVALHO, A.; GOFFINET, F.; FIVET, A. C.; CHRISTOPHE, A. Ambiguous function words do not prevent 18-month-olds from building accurate syntactic category expectations: an ERP study. *Neuropsychologia*, Elsevier, v. 98, p. 4-12, April 2017. Doi: <http://dx.doi.org/10.1016/j.neuropsychologia.2016.08.015>.

BRUSINI, P.; DEHAENE-LAMBERTZ, G.; DUTAT, M.; GOFFINET, F.; CHRISTOPHE, A. ERP evidence for on-line syntactic computations in 2-year-olds. *Developmental Cognitive Neuroscience*, Elsevier, v. 19, p. 164-173, 2016. Doi: <https://doi.org/10.1016/j.dcn.2016.02.009>.

CAIN, K.; NASH, H. The influence of connectives on young readers' processing and comprehension of text. *Journal of Educational Psychology*, American Psychological Association, v. 103, p. 429-441, 2011. Doi: <https://doi.org/10.1037/a0022824>.

CEREJA, W. R.; MAGALHÃES, T. C. *Português: linguagens*. 6º ano. Ed. reform. São Paulo: Atual, 2009.

COELHO, N. N. *Literatura Infantil: teoria, análise, didática*. São Paulo: Moderna, 2000.

COHEN, J. H. The effect of content area material on Cloze test performance. *Journal of Reading*, International Literacy Association, v. 19, n. 3, p. 247-250, 1975.

COLLINS. *Escolar Plus Dictionary: English/Portuguese*. [s.l.]: Heinle Cengage Learning, 2009.

CONDEMARÍN, M.; MILICIC, N. *Teste de Cloze: aplicações psicopedagógicas*. Santiago de Chile: Editorial Andrés Bello, 1988.

CUNHA, N. de B. Diagnóstico de compreensão de leitura por meio do teste de Cloze. In: CONGRESSO DE LEITURA DO BRASIL, 17., 2009, Campinas. Anais... Campinas: Unicamp, 2009. v.1, p. 238-247.

CYRANKA, L. F.; PERNAMBUCO, D. L. C. A língua culta na escola: uma interpretação sociolinguística. *Instrumento: Revista de Estudo e Pesquisa em Educação*, Juiz de Fora, v. 10, p. 17-28, Jan./Dec. 2008.

DIAS, N. M.; LEÓN, C. B. R.; PAZETO, T. de C. B.; MARTINS, G. L. L.; PEREIRA, A. P. P.; SEABRA, A. Avaliação da leitura no Brasil: revisão da literatura no recorte 2009-2013. *Revista Psicologia: Teoria e Prática*, São Paulo, v. 18, n. 1, p. 113-128, Jan./Apr. 2016.

FEDERMEIER, K. D.; WLOTKO, E. W.; OCHOA-DEWALD, E.D.; KUTAS, M. Multiple effects of sentential constraint on Word processing. *Brain Research*, Elsevier, v. 1146, p. 75-84, 2007. Doi: <https://doi.org/10.1016/j.brainres.2006.06.101>.

FONSECA, R. P.; CASARIN, F. S.; OLIVEIRA, C. R.; GINDRI, G.; SOARES, E. C. S.; ORTIZ, K. Z.; PARENTE, M. A. M. P.; SCHERER, L. C. Adaptação de instrumentos neuropsicológicos verbais: um fluxograma de procedimentos para além da tradução. *Interação em Psicologia*, Universidade Federal do Paraná, v. 15, n. especial, p. 59-69, 2011. Doi: <https://doi.org/10.5380/psi.v15i0.25374>.

GONZALEZ, C. A.; QUADROS, E. S.; SCHWINDT, L. C. S.; TOLEDO, E. E. A influência da variável escolaridade em fenômenos fonológicos variáveis: efeitos retroalimentadores da escrita. *Revista Virtual de Estudos da Linguagem – Revel*, [s.l.], v. 5, n. 9, Aug. 2007.

INFANTE, U. *Textos: leituras e escritas: literatura, língua e redação*. São Paulo: Scipione, 2000. v. 1.

JOLY, M. C. R. A. *Sistema Orientado do Cloze*. Relatório de pesquisa apresentado à Universidade São Francisco, Itatiba, São Paulo, 2006.

JOLY, M. C. R. A. The validity of Cloze Oriented System (COS): a correlation study with an electronic comprehension test and a reading attitude survey. *Psicologia Escolar e Educacional*, Campinas, v. 11, n. esp., p. 49-58, 2007.

JOLY, M. C. R. A. *Programa Informatizado de Leitura Estratégica (PILE) – Pesquisa em desenvolvimento*. Itatiba, SP: Universidade São Francisco, 2008.

JOLY, M. C. R. A. Estudos com o Sistema Orientado de Cloze para o ensino Fundamental. In: SANTOS, A. A. A.; BUCHOVITCH, E.; OLIVEIRA, K. L. (Ed.). *Cloze: um instrumento de diagnóstico e intervenção*. São Paulo: Casa do Psicólogo, 2009. p. 119-144.

JOLY, M. C. R. A.; PIOVEZAN, N. M. Avaliação do Programa Informatizado de Leitura Estratégica para estudantes do ensino fundamental. *Paidéia*, Ribeirão Preto, v. 22, n. 51, p. 83-90, Apr. 2012. Doi: <https://doi.org/10.1590/S0103-863X2012000100010>.

JOLY, M. *et al.* Avaliação da Compreensão de Leitura pelo Sistema Orientado de Cloze (SOC). *Fractal: Revista de Psicologia*, Universidade Federal Fluminense, v. 26, n. 1, p. 223-242, Jan./Apr. 2014.

KATO, M. A. *No mundo da escrita: uma perspectiva psicolinguística*. 7. ed., 3ª impressão. São Paulo: Ática, [1986], 2001.

KINTSCH, W.; VAN DIJK, F. A. Toward a model of text comprehension and production. *Psychological Review*, American Psychological Association, v. 85, p. 363-339, 1978.

KLEIMAN, A. *Texto e leitor: aspectos cognitivos da leitura*. Campinas: Pontes, 1989.

OLIVEIRA, K. L.; BORUCHOVITCH, E.; SANTOS, A. A. A. A técnica de Cloze na avaliação da compreensão em leitura. In: SANTOS, A. A. A.; BORUCHOVITCH, E.; OLIVEIRA, K. L. *Cloze: um instrumento de diagnóstico e intervenção*. São Paulo: Casa do Psicólogo, 2009. p. 47-77.

OLIVEIRA, K. L.; CANTALICE, L. M.; FREITAS, F.A. Compreensão em leitura no ensino médio: análise de acertos por item. In: SANTOS, A. A. A.; BORUCHOVITCH, E.; OLIVEIRA, K. L. *Cloze: um instrumento de diagnóstico e intervenção*. São Paulo: Casa do Psicólogo, 2009. p. 165-185.

PAGE, W. D. The post-oral reading Cloze test: New link between oral reading and comprehension. *Journal of Reading Behavior*, Chicago, v. 7, n. 4, p. 383-389, 1975. Doi: <https://doi.org/10.1080/10862967509547155>.

PARENTE, M. A. M. P.; SCHERER, L. C.; ZIMMERMANN, N.; FONSECA, R. P. Evidências do papel da escolaridade na organização cerebral. *Revista Neuropsicologia Latinoamericana* [online], Sociedad Latinoamericana de Neuropsicologia, v. 1, n. 1, p. 72-80, 2009. Available at: http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S2075-94792009000100009&lng=pt&nrm=isso. Accessed on: 3 Jan. 2017. ISSN 2075-9479.

PEREIRA, M.; ARAÚJO, A. Considerações acerca da variável escolaridade e sua influência sobre a variação entre verbo-sujeito na 3ª pessoa do plural no português brasileiro. *PERcursos Linguísticos*, Vitória - ES, v. 6, n. 12, 2016. ISSN: 2236-2592.

SADEGHI, K. Phrase Cloze: a better measure of reading? *The Reading Matrix*, Reading Matrix Inc., v. 14, n. 1, p. 76-94, Apr. 2014.

SANTOS, A. A. A. *Diagnóstico de compreensão de leitura por meio do teste de Cloze*. Programa de Pós Graduação Stricto Sensu em Psicologia. Itatiba: Universidade São Francisco, 2002.

SANTOS, A. A. A. O Cloze como técnica de diagnóstico e remediação da compreensão em leitura. *Interação em Psicologia*, UFPR, v. 8, n. 2, p. 217-226, jul./dez. 2004. Doi: <https://doi.org/10.5380/psi.v8i2.3257>.

SANTOS, A. A. A.; BORUCHOVITCH, E.; OLIVEIRA, K. L. Análise da fidedignidade entre dois tipos de pontuação do teste de Cloze. *Psicologia em Pesquisa*, Universidade Federal de Juiz de Fora, v. 1, n. 1, p. 41-51, Jan./Jun. 2007.

SANTOS, A. A. A.; BORUCHOVITCH, E.; OLIVEIRA, K. L. *Cloze: um instrumento de diagnóstico e intervenção*. São Paulo: Casa do Psicólogo, 2009. 380 p.

SANTOS, A. A. A.; PRIMI, R.; TAXA, F. O. S.; VENDRAMINI, C. M. M. O teste de Cloze na avaliação da compreensão em leitura. *Psicologia: Reflexão e Crítica*, Universidade Federal do Rio Grande do Sul, v. 15, n. 3, p. 549-560, 2002. Doi: <https://doi.org/10.1590/S0102-79722002000300009>.

SEGALOWITZ; S. J.; LANE, K. Lexical access of function versus content words. *Brain and Language*, Elsevier, v. 75, n. 3, p. 376-389, Dec. 2000.

SMITH, F. *Compreendendo a leitura: uma análise psicolinguística da leitura e do aprender a ler*. Tradução de Daise Batista. 4. ed. Porto Alegre: Artmed, 2003.

SOUSA, L. B. de; HUBNER, L. C. Desafios na avaliação da compreensão leitora: demanda cognitiva e leiturabilidade textual. *Revista Neuropsicologia Latinoamericana*, Sociedad Latinoamericana de Neuropsicologia, v. 7, n. 1, p. 34-46, 2015.

SUEHIRO, A. C. B. Produção Científica sobre o Teste de Cloze. *Psicologia Escolar Educacional*, Maringá, v. 17, n. 2, p. 223-232, Dec. 2013. Doi: <https://doi.org/10.1590/S1413-85572013000200004>.

SUEHIRO, A. C. B.; BORUCHOVITCH, E. Compreensão em Leitura em Estudantes do Terceiro e Quarto Ciclos do Ensino Fundamental. *Psico-USF*, Itatiba, v. 21, n. 3, p. 561-572, Dec. 2016.

SWAFFAR, J. K.; ARENS, K. M.; BYRNES, H. *Reading for meaning: an integrated approach to language learning*. Englewood Cliffs, N.J.: Prentice Hall, 1991.

TAFT, M. Lexical processing of functionally constrained words. *Journal of Memory and Language*, Elsevier, v. 29, p. 245-257, 1990. Doi: [https://doi.org/10.1016/0749-596X\(90\)90075-B](https://doi.org/10.1016/0749-596X(90)90075-B).

TAYLOR, W. L. Cloze procedure: a new tool for measuring readability. *Journalism Quarterly*, Questia Trusted Online Research, v. 30, p. 415-433, 1953.

VAN HEUGTEN, M.; CHRISTOPHE, A. Infants' acquisition of grammatical gender co-occurrences. *Infancy*, Wiley Online Library, v. 20, p. 675-683, 2015.

VAN PATTEN, B. *Input processing and grammar instruction: theory and research*. Norwood, NJ: Ablex, 1996.

VOTRE, S. J. Relevância da variável escolaridade. In: MOLLICA, M. C.; BRAGA, M. L. (Org.). *Introdução à Sociolinguística: o tratamento da variação*. São Paulo-SP. Editora Contexto, 2012. p. 51-57.