

RESEARCH ARTICLE

Sentence contexts and cloze probabilities for Brazilian Portuguese children and adolescents

Natalia Freitas Rossi^{1†*}, Catarina Fernandes^{2‡}, Célia Sofia Moreira³, Célia Maria Giacheti¹, Bianca Bortolai Sichieri¹, Ana Patrícia Pinheiro^{2,4}, Adriana Sampaio²

1 Faculty of Philosophy and Sciences, São Paulo State University (UNESP), Speech, Language and Hearing Science Department, Campus of Marília, São Paulo, Brazil, **2** Psychological Neuroscience Lab, CIPsi, School of Psychology, University of Minho, Braga, Portugal, **3** Faculty of Sciences and Center of Mathematics (FCUP & CMUP), University of Porto, Porto, Portugal, **4** Faculty of Psychology, University of Lisbon, Lisbon, Portugal

‡ These authors share first authorship on this work.

* nataliafreitasrossi@yahoo.com.br



OPEN ACCESS

Citation: Rossi NF, Fernandes C, Moreira CS, Giacheti CM, Sichieri BB, Pinheiro AP, et al. (2020) Sentence contexts and cloze probabilities for Brazilian Portuguese children and adolescents. *PLoS ONE* 15(7): e0236388. <https://doi.org/10.1371/journal.pone.0236388>

Editor: Koji Miwa, Nagoya University, JAPAN

Received: February 21, 2019

Accepted: July 6, 2020

Published: July 30, 2020

Copyright: © 2020 Rossi et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: This research was supported by Grants from Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior (CAPES), and by a Post-Doctoral Fellowship at the University of Minho (Process BEX 9573/11-9). This study was partially conducted at the Psychology Research Centre (PSI/01662), University of Minho, and supported by the Instituto Nacional de Ciência e Tecnologia sobre Comportamento, Cognição e Ensino, with

Abstract

Purpose

In this study we investigated a set of 100 sentence contexts and their cloze probabilities to develop a database of linguistic stimuli for Brazilian Portuguese children and adolescents. The study also examined age-related changes on cloze probabilities, and specified the predictor effects of age and cloze probabilities on idiosyncratic responses and errors (semantic, syntactic, and other errors). Finally, the study also aimed to shed light on cultural effects on word generation by comparing Brazilian and Portuguese sentence databases.

Method

361 typically developing monolingual Brazilian speakers, with ages ranging from 7 to 18 years, participated in the study. The cloze task was composed by 100 sentence contexts, grounded on the European Portuguese database. Responses were classified as valid (correct) or invalid (semantic, syntactic, and other-type errors). Statistical analyses were based on mixed-effects logistic models.

Results

Sixty-three sentences met criteria for high cloze probabilities, 30 for medium cloze, and 7 for low cloze. Age was a significant predictor of idiosyncratic responses, semantic and syntactic errors: older participants were less likely to produce idiosyncratic responses, as well as semantic and syntactic errors. Cloze probability values were concordant in the Brazilian and Portuguese databases for 31 out of 49 (83.7%) high-cloze sentences and for 7 low-cloze sentences.

Conclusion

In this study we have provided a database with cloze probability values for a set of 100 sentence-final word contexts for Brazilian Portuguese children and adolescents. Results showed that both age and sentence contextual level predicted sentence final word

funding of the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, process no 465686/2014-1), Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, process no 2014/50909-8), Portuguese Foundation for Science and Technology and the Portuguese Ministry of Science, Technology and Higher Education (UID/PSI/01662/2019), through the national funds (PIDDAC) and by national funds from the Portuguese Foundation for Science and Technology (Grant PTDC/PSI-PCL/115316/2009). CSM was partially supported by CMUP (UID/MAT/00144/2013), which is financed by national funds through FCT (Portugal) under the project with reference UIDB/00144/2020.

Competing interests: the authors have declared that no competing interests exist.

completion. Older participants were more likely to choose more consistently the same final word, with the contextual level of a given sentence also contributing to the final word selection. Age should be controlled for in future studies probing semantic processing with this set of sentences.

Introduction

The language comprehension system changes across development and depends on a complex interaction between neurobiological and sociocultural factors [1, 2]. It is not surprising that sentence comprehension is one of the most sophisticated cognitive abilities of the human brain [3], relying on a well-orchestrated interplay of several cortical brain regions [4]. Specifically, the brain must recognize a word and its lexical form (spoken or written), assign meaning to that word (semantic processing), retrieve specific grammatical rules (syntactic processing), and relate the meaning to its broader context (pragmatic processing). Language comprehension is, therefore, one of the major challenges that children must master to become proficient language users, remaining a fundamental cognitive ability throughout life [3].

In particular, semantic processing refers to the ability to access acquired knowledge and to make sense out of otherwise arbitrary linguistic symbols [5], which relies on critical interactions with the semantic memory system. Semantic memory represents the internally organized knowledge about word meanings and their relationships [6, 7], as well as the conceptual knowledge of the world [6]. Therefore, it is essential for language comprehension [8], and for the functional use of language [9].

Sentence comprehension recruits a distributed network of brain regions, including the left inferior frontal gyrus, the right superior temporal gyrus, the left middle temporal gyrus, and the left posterior temporal region [10]. The selection and retrieval of lexical representations is supported by the inferior frontal cortex, whereas the anterior temporal cortex and angular gyrus support the integration of lexical input into the larger units of sentences or discourse [11]. Although this network is observed at 7 years, substantial neurodevelopmental changes occur underpinning semantic processing until adulthood [12].

Classical models of semantic processing [13–15] highlight the importance of semantic context in sentence comprehension, as the meaning of a sentence is more than the sum of the meanings of its individual words [16]. Contextual effects on language comprehension are observed very early in development [17], since context is required to solve ambiguities in sentence structure [2]. Sentence completion tasks have been used to probe oral and written language comprehension [18–20] in typically developing individuals and clinical populations [21–23]. Specifically, in the cloze task, participants are asked to complete a sentence with the first word that comes to their mind [24]. This task allows computing a cloze probability value for the final word that completes a given sentence context, which represents the proportion of individuals who provide the same final word to the sentence given the preceding words [18, 20, 25, 26].

In this context, Bloom and Fischler's (1980) [27] sentence completion norms have been widely used in different languages, such as British English [28], French [29], Latino-American Spanish [30] and European Portuguese [31], and with different age groups [30, 31]. Additionally, cloze probabilities have supported the selection of stimuli in controlled experimental tasks with event-related potentials [18]. Despite the widespread utility of the Bloom and Fischler's (1980) norms [27], some studies evince differences in cloze probabilities among different languages [28] and age groups [32, 33], which highlight the need to address cross-cultural

comparisons and different age intervals. Accordingly, cultural and linguistic specificities such as level of education or vocabulary and grammatical rules are among the most cited limitations [27]. A robust body of evidence demonstrated that cultural contexts are critical factors that modulate language production and comprehension [34] and influence which semantic information is learned, used [35] and organized [36]. Neurophysiological studies also supports an interaction between cultural identity and semantic processing during sentence comprehension [3].

Of note, most sentence completion norms are restricted to adults [13, 18, 28, 29, 32, 33]. To the best of our knowledge, only two studies relying on cloze tasks have tested children and adolescents, with Mexican Spanish (ages of 9 and 12) [30], and European Portuguese (ages of 6 and 11) samples [31]. Even though Rodríguez-Camacho et al. (2011) [30] did not explore age effects, the authors highlighted the need to consider developmental processes in word recognition processes and language comprehension as databases relying on norms for younger and older adults are inappropriate to assess school-age children, who use less complex grammatical structures.

In the current study, sentence contexts were grounded on the European Portuguese database [31] for two main reasons: first, Portuguese is the official language in both countries (Brazil and Portugal), ensuring lexical and grammatical similarities but with different cultural contexts; second, the European database was developed for the target population of children and adolescents. Notwithstanding, even though Brazilian Portuguese shares a large part of the European Portuguese language heritage, some cultural specificities determine lexical differences between European and Brazilian Portuguese [37], including phonetic-phonological, morphological and syntactical levels [38]. Specifically, the pronominal system (“*você*” vs. “*tu*”), use of gerund vs. infinitive verb (“*está chovendo*” vs. “*está a chover*”), and morphological derivation (“*desenhista*” vs “*desenhador*”) are some of the main differences between Brazilian and European Portuguese [37]. Lexical differences include the use of different words for the same referent (e.g., “*autocarro*” vs. “*ônibus*” [bus]), the use of words with different meanings (e.g., “*cueca*” [underpants] vs. “*calcinha*” [panty]) and the presence of words derived from indigenous and African languages in Brazil [39]. Social and cultural factors may, therefore, influence the frequency of word use. For example, fish (“*peixe*”) and soup (“*sopa*”) are everyday food items used in Portugal whereas bean (“*feijão*”) and rice (“*arroz*”) are more commonly used in Brazil [39]. These social and cultural effects are observed also when comparing the psychometric properties of vocabulary assessment tools in Brazil and Portugal, with differences regarding normative data being observed for Brazilian Portuguese and European Portuguese [40]. These examples highlight the need to adapt stimulus material to each cultural context [40, 41] and are consistent with previous studies documenting modulatory effects of cultural and linguistic differences on cloze probability values [27, 28]. Although the main purpose of this study was not to provide a robust analysis of cultural-specific patterns (Brazil and Portugal) on cloze probabilities values, we have investigated the concordance between the Brazilian and the European Portuguese database in terms of cloze probabilities value.

Regarding developmental changes, Pinheiro et al. (2010) [31] described an age-effect for the cloze probability values, syntactic and semantic responses errors and idiosyncratic responses. There was a reduction in the numbers of idiosyncratic responses and semantic errors; and, nonetheless, and an increased consistency in the type of appropriate final words chosen to complete a specific sentence context. Authors argued that these findings support the idea that a very early mechanism is learned for extracting regularities in the environment and constructing common representations about things, indicating that consistency in word selection tends to increase during development.

In addition to cultural specificities of the cloze probabilities values, age was also considered an important variable to account for the current study for several reasons: 1) there is evidence that semantic networks mature during childhood and adolescence, which is reflected in age-related differences in language performance [42]; 2) previous behavioral studies have documented more consistent patterns of sentence context completion as a function of increased age [31–33], with a decreased number of idiosyncratic (i.e., words chosen by only one subject) and invalid responses (i.e., words that do not properly match their sentence contexts) [31]. Our prediction was that older participants would be less likely to generate idiosyncratic responses and errors (syntactic and/or semantic or other errors) due to advantages related to brain maturation, literacy, education and metalinguistic competences. In addition, we expected to identify more idiosyncratic responses and errors in younger participants for low-cloze probability sentences. This prediction is consistent with developmental changes in cognitive control (i.e., executive attention and inhibitory mechanisms) and sentence-level semantic processing that operate together with sentence constraint in determining word choice. This hypothesis implies that higher cloze probabilities would be associated with high-constraint sentences and enhanced executive control (leading to inhibition of other lexical possibilities) in older participants. The consolidation of lexical-syntactic and semantic processes in older individuals is expected to lead to advantages in the selection of a more accurate final word for a given sentence [31–33], reflected in an increased consistency in word selection [31], as well as in a decreased number of idiosyncratic (i.e., words chosen by only one subject) and invalid responses (i.e., words that do not properly match their sentence contexts) [31].

In sum, the primary goal of this study was to investigate a set of 100 sentence contexts and their cloze probabilities to develop a database of linguistic stimuli for Brazilian Portuguese children and adolescents, which was grounded on a European Portuguese database [31]. The current study also examined age-related changes on the most frequently selected sentence ending, as well as the predictor effects of age and cloze probabilities on the generation of idiosyncratic responses and errors (semantic, syntactic, and other errors). Finally, the study also aimed to provide a comparison of sentence constraint levels in the Brazilian and European Portuguese databases.

Materials and methods

Participants

Participants were 361 monolingual Brazilian Portuguese students with ages ranging from 7 to 18 years ($M = 9.94$ $SD = 2.33$; 171 females; 190 males): 290 were children aged from 7 to 11 years ($M = 9.01$, $SD = 1.09$) and 71 were adolescents aged between 12 and 18 years ($M = 13.73$, $SD = 2.19$). The following inclusion criteria were used: (a) age ranging from 6 to 18 years old; (b) parent authorization with signed consent form; (c) no previous experience of grade repetition; (d) no current history of poor performance in reading, writing, and mathematics; (e) no reported behaviour problems or spoken language difficulties by the teacher and confirmed by a speech-language pathologist; (f) ability to complete the task in the written modality; and (g) correct answers in at least three of five training blocks. The study was approved by Ethical Committee of Faculty of Philosophy and Sciences, São Paulo State University (UNESP), Campus of Marília, São Paulo, Brazil (process number 0526/2012). All guardians signed written consent forms.

First, all teachers received information about the main goals of the current study. Second, the informed consent was sent to 490 parents of students from primary to secondary school levels. Parents of 431 students signed the informed consent. For those, the teacher was requested to fill out a questionnaire regarding their students' language (oral and writing)

abilities, academic performance, behaviour and communication problems. From 431 consented students, 410 fit the inclusion criteria. Four speech-language pathologists individually screened these students in a 15-minutes session that aimed to exclude major speech and language problems. From these 410 students, 5 (6 years old) were excluded and referred for a more specific evaluation due to speech sounds distortions (3 of 5) and speech fluency problems (2 of 5). Finally, the cloze task was administered to 405 students: 23 did not complete the task, 14 failed in the training block (<three corrected answers), and 7 were not able to complete the task in the written modality (all students with 6 years old). The final sample was composed of 361 participants. Participants' socioeconomic level ranged from B1 to D, according to the Brazilian Association of Research Companies [43].

Procedure

Data were collected in a quiet room at the school where participants were recruited from March to June 2014. The cloze task was composed by 100 sentence contexts: 73 were selected from the Pinheiro et al. (2010) database with European Portuguese participants [31], and the remaining 27 sentences were developed following the procedure described in this prior study. All sentences had the same sentence length (four words per sentence) and syntactic structure (subject and direct verb in the present tense: SVO). When a target word (subject or verb) was not available or it was unusual in the local cultural vocabulary, it was replaced by another word with semantic equivalence. Some adaptations were necessary for 19 of the 73 sentence contexts from Pinheiro et al. (2010) database [31]. The 100 sentence contexts are described in Appendix Table 1 in [S1 Appendix](#). The 100 sentences were presented in written form (recording sheet), with text printed in black color on white paper (Times New Roman font, size 12). The recording sheet was composed of two columns (the stimulus column and column response) with 100 lines (one line for each sentence).

Participants were told that this was a task in which they would read sentences that were incomplete; they were instructed to complete them with the first word that came to their mind and that provided an appropriate ending for that sentence context. Specifically, the following instruction was provided: "*You will read sentences that are incomplete, that is they are missing a word at the end of the sentence. You will have to read carefully each one and complete the sentence with the first word that comes to your mind. This word must fit the rest of the sentence, that is, you will try to form a sentence that makes sense. Do you understand? Got it? Before we begin, a short practice session will take place*". The training block included five sentences that were not part of the experimental task, ensuring that participants were familiarized with the instructions and with the task. During training, interventions by the experimenter were permitted to allow the participants to fully understand the nature of the task. These interventions included alerting to the semantic (i.e., a final word choice leading to a semantic violation as in "The seamstress uses the *whistle*") and syntactic nature of potential errors (i.e., gender or number agreement with the object) or even to distraction, as well as providing positive feedback for appropriate words and negative feedback for inappropriate word selections. Participants were instructed to respond carefully since they were not allowed to use rubber.

Response coding

Responses were coded as valid or invalid, following Pinheiro et al. (2010) [31]. A valid final word was computed if it completed the sentence context appropriately for both levels, semantic and grammatical. Additionally, for each valid response we computed (1) the frequency of the valid final word for a given sentence context to determine the proportion of participants who generated the same final-word sentence (cloze probability); (2) idiosyncratic responses

(valid words generated by a single participant); and (3) invalid responses. The invalid responses were further categorized as: semantic error; syntactic error; other errors. Semantic errors represented final-word selections that did not fit the previous context in a semantically valid way, for example, if the final word had no semantic relationship with the subject or with the verb of the sentence (e.g., “A galinha põe milho”; “The chicken lays corn”). A syntactic error occurred when the final-word was syntactically invalid and violated grammatical rules, such as grammatical gender (male and female) and number (singular and plural) agreements (e.g., “A cabeleireira corta o ^{masculine article} franja ^{feminine noun}”; “The hairdresser cuts the bangs”). Other-type errors included: (1) final-word with both semantic and syntactic violations (e.g., “O padre reza a ^{feminine article} trabalho ^{masculine noun and incongruous}” – “The priest prays the work”); (2) using more than one final-word response (e.g., “O professor ensina o aluno a ler”; “The teacher teaches the student to read”) and (3) suffix error creating a word that does not exist (e.g., “O carpinteiro faz uma carpintação”; “The carpenter makes a carpentertion”).

Two judges independently coded the responses. When there was a disagreement among them, they first tried to reach a consensus; when not possible, a third judge was involved and provided the final decision. All judges were experts in language assessment and were external to the research team. Inter-rater agreement was higher than 90%. Cloze probability level was established according to the criterion proposed by Block and Baldwin (2010) [18]: low cloze was defined as 0%–33% (.0–.33), medium cloze as 34%–66% (.34–.66), and high cloze as 67%–100% (.67–1).

Data analyses

All statistical analyses were performed using the R statistical environment (RStudio, version 3.6.1, R Development Core Team, 2019) using the packages “glmmTMB” [44] and “effects” [45]. The statistical significance level was set at the conventional $\alpha = .05$. The R script is available in Supporting Information S3 Rossi_R.R file.

The original dataset contains all participants’ responses to 100 context sentences. More specifically, each column corresponds to a different sentence and each row corresponds to a different participant ($N = 361$ rows). After identifying the most frequent word for each sentence, the following (100) dichotomous variables were added to the original dataset: occurrence (1 = “yes”, 0 = “no”) of the expected response in each cloze sentence; occurrence of idiosyncratic responses; and occurrence of errors, one for each error type (semantic, syntactic, and other errors).

All columns encoding information about the occurrence of expected words were used to construct a new simpler dataset (“S2 Dataset Rossi_sentences.csv”), which was used to assess the predictor effect of age on the generation of the most frequent word (S1 Table in [S1 Appendix](#)), for each sentence separately. The original dataset was also used to construct a coded long format dataset (“S2 Dataset Rossi.csv”), with $N \times 100$ rows. Moreover, in order to establish a comparison between error types, we constructed another dataset having $N \times 100 \times 3$ rows and a single “Error” variable encoding information about the occurrence of invalid responses, whose error type was identified in another variable “Type” (semantic, syntactic, or other type).

To assess the predictor effect of age and of cloze probabilities on the generation of idiosyncratic and invalid responses, mathematical modeling tools were applied to each outcome, independently, using the “glmmTMB” R package [34]. Since all outcomes are dichotomous, mixed-effects logistic models with both participants and sentences random effects were performed. Indeed, logistic modeling is adequate to model a dichotomous outcome variable Y , i.e., a variable Y that only assumes two values: 0 or 1. Usually, the condition $Y = 1$ indicates an individual property in a population. In a logistic model, the odds of having $Y = 1$ are defined

as the ratio $[P(Y = 1)]/[P(Y = 0)]$ and reflect the likelihood for the condition $Y = 1$ (P means the probability; $P(Y = 0) = 1 - P(Y = 1)$). Odds can be defined as the expected number of events divided by the expected number of non-events [46]. For example, consider a simple logistic model with a unique continuous predictor X , and assume that b_0 and b_1 are the model coefficients, b_1 being associated with X . Then,

$$\log([P(Y = 1)]/[P(Y = 0)]) = b_0 + b_1 * X$$

This model allows to write the probability $P(Y = 1)$ as:

$$\exp(b_0 + b_1 * X) / [1 + \exp(b_0 + b_1 * X)]$$

In terms of interpretation, if $b_1 > 0$ ($b_1 < 0$), the coefficients of the model are interpreted as follows: one unit increase in X will increase (decrease) the odds of $Y = 1$ by $\exp(b_1)$ times.

The two main datasets and the R script used to perform all analyses are available in the Supplementary Material ([S1](#), [S2](#) Data sets and [S1 File](#)).

Results

1) Sentences analysis

As observed, 63 out of 100 sentence contexts met criteria for high cloze sentences, 30 were medium cloze probability, and 7 low cloze probabilities. Among the 63 high cloze sentences, 49 (77.8%) were from Pinheiro et al (2010) [31]. The cloze level was concordant in Brazil and Portugal databases for 31 out of 49 (83.7%) high-cloze sentences and for 7 low-cloze sentences. Among the remaining 26 medium cloze sentences there was no concordance between sentences from Brazil and Portugal (see Appendix–Table 1 in [S1 Appendix](#)).

For each sentence-context, we analyzed whether participants' age influenced the generation of the most frequent word. To this purpose, logistic models were conducted, considering the dichotomous variable encoding information about the use of the most frequent word (1 = "yes", 0 = "no") as outcome. The results of these (one hundred) models are summarized in the Appendix (Table 1 in [S1 Appendix](#)), from high to low cloze probability. When inspecting age-related differences in cloze probability for each sentence, 63% of the sentences showed no age effects. Among these, 47 (74.6%) met criteria for high cloze probability, 10 (15.9%) had medium cloze probability, and 6 (9.5%) obtained low cloze probabilities.

We also examined whether cloze probabilities predicted the occurrence of idiosyncratic and invalid responses. Mixed-effects logistic models with both participants and sentences random effects were performed. At a first sight, this analysis might seem trivial since, for higher cloze probabilities, there should necessarily be less idiosyncratic and invalid responses. However, it is important to point out that a similar conclusion does not hold for lower cloze probabilities: when participants do not generate the most frequent word, they do not necessarily provide an idiosyncratic or invalid response, i.e. it may happen that another valid response has been given. Moreover, this analysis is also useful to compare differences between the effects of cloze probabilities on the three error types.

The results showed that higher cloze probabilities were associated with less idiosyncratic responses (estimate = -2.67, $p < .001$). In the case of invalid responses, and in order to compare the three error types, the interaction between cloze probabilities and the error type was selected as fixed effect. The results showed that higher cloze probabilities were associated with less semantic errors (estimate = -3.35), syntactic errors (estimate = -3.74), and other errors (estimate = -6.25), all cases with $p < .001$. In general, when compared to semantic errors, participants generated more syntactic errors (estimate = 0.80, $p < .001$) and other errors (estimate = 0.94, $p < .001$). However, as cloze probabilities increase, the decreased tendency of

other error types is expected to be stronger (estimate = -2.93 , $p < .001$). Appendix Table 2 in [S1 Appendix](#) summarizes the results of these two models and [Fig 1](#) outlines these effects.

2) Participants' performance analysis

After analyzing the cloze sentence features, we investigated the predictor effect of age on participants' performance. More specifically, we examined participants' ability to generate expected words, idiosyncratic responses, and errors (semantic, syntactic, and other type).

Mixed-effects logistic models with both participants and sentences random effects were performed. For the most frequent word, the result showed a significant effect of age, with one year difference being associated with a 7% ($\exp(0.07) = 1.07$, $p < .001$) increase in the odds of using the most frequent word. [Fig 2](#) outlines this predictor effect and Appendix Table 3 in [S1 Appendix](#) summarizes the model.

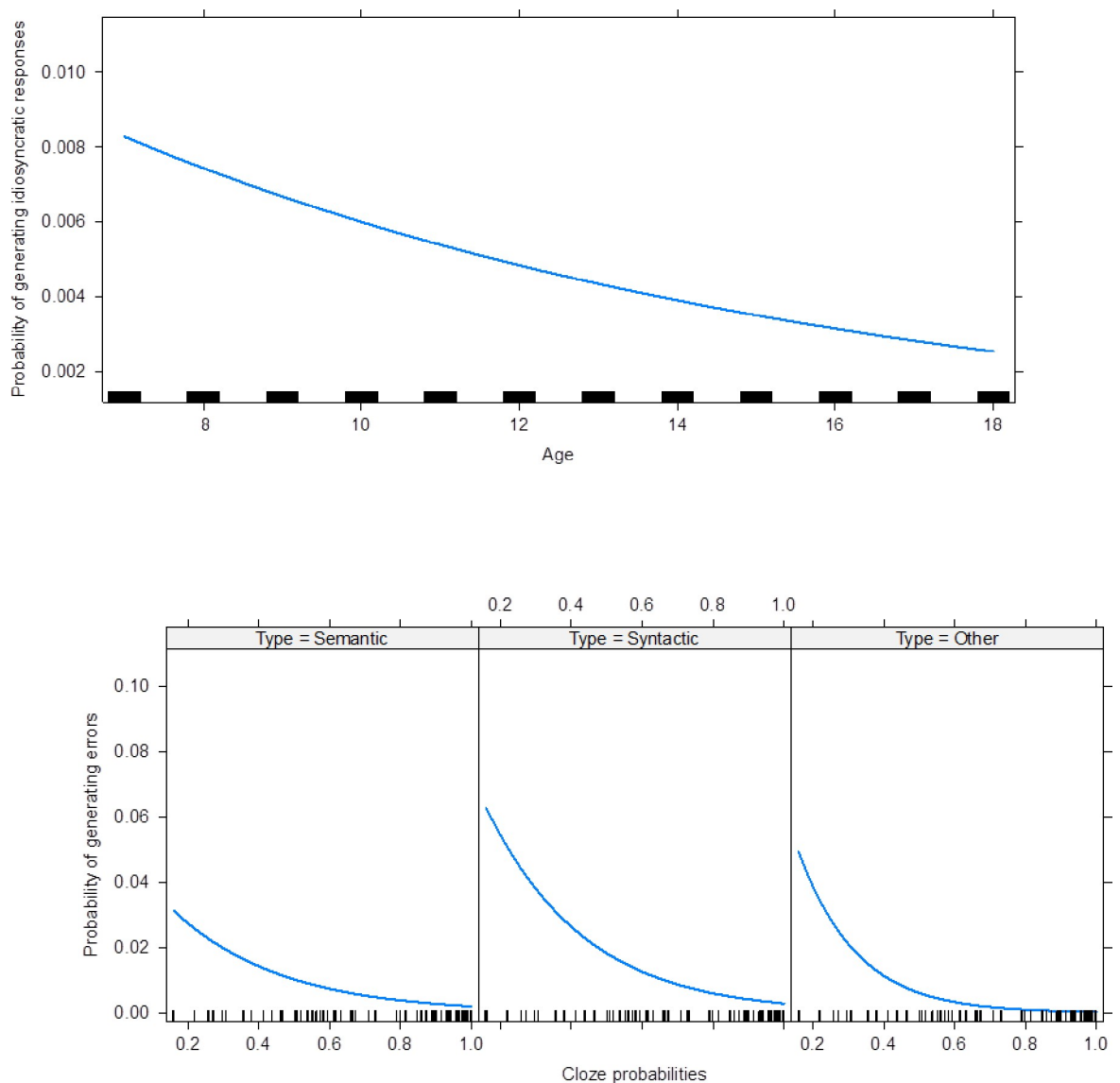


Fig 1. The predictor effect of cloze probabilities on the generation of idiosyncratic and invalid responses.

<https://doi.org/10.1371/journal.pone.0236388.g001>

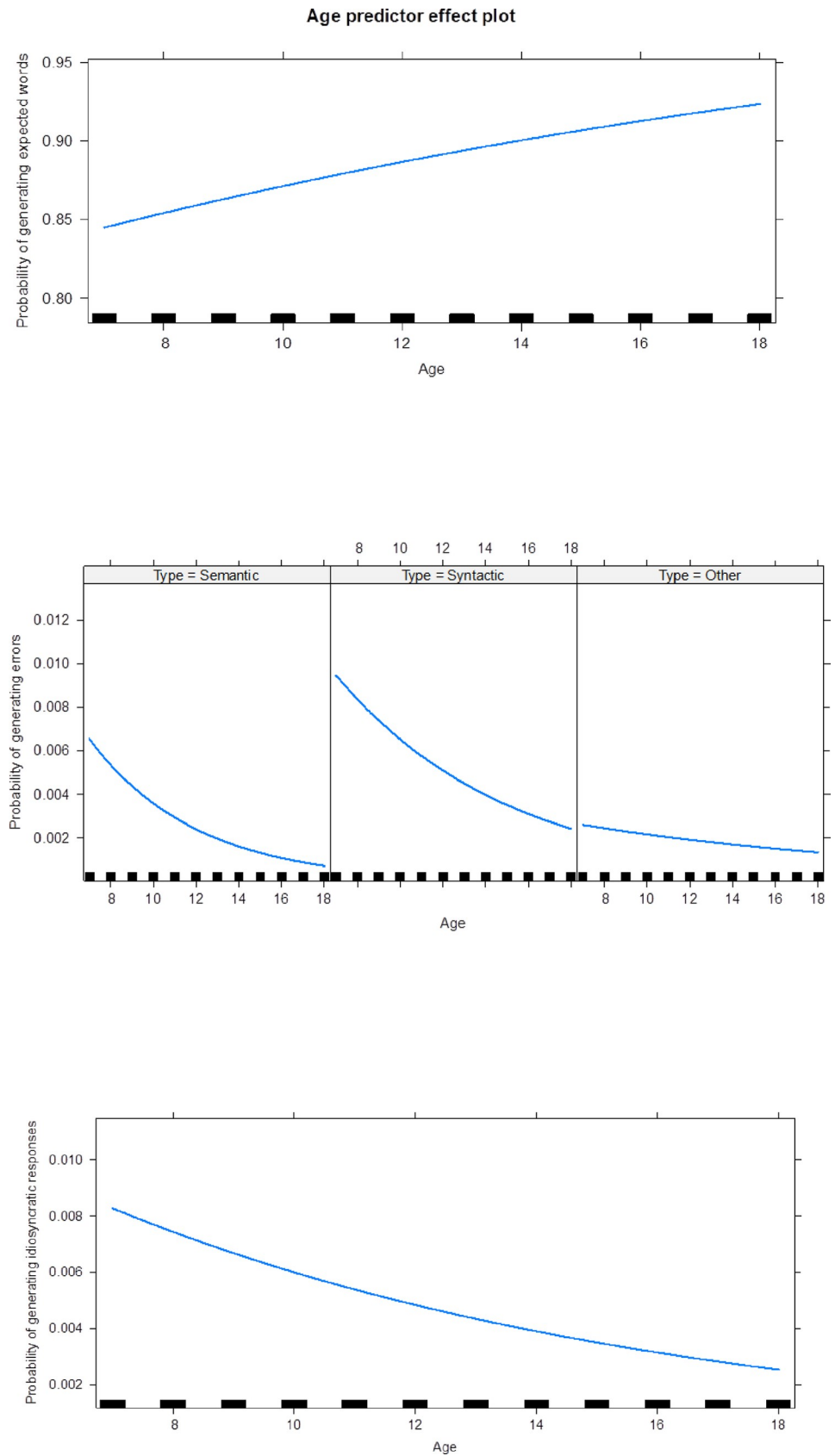


Fig 2. The predictor effect of age on the generation of expected answers, idiosyncratic responses, and errors.

<https://doi.org/10.1371/journal.pone.0236388.g002>

We also examined whether age had a predictor effect on the generation of idiosyncratic and invalid responses. We observed that older participants are less likely to generate idiosyncratic responses (estimate = -0.11, $p < .001$); more precisely, if we compare two participants with one year difference, the odds of generating idiosyncratic responses is expected to be 10% smaller for the older [$\exp(-0.11) = 0.90 = 1 - 0.10$, $p = .005$].

In the case of invalid responses, there were no significant differences between semantic and syntactic errors for the younger participants. The results showed that older participants are less likely to make semantic (estimate = -0.20) and syntactic errors (estimate = -0.12, $p < .001$), both with significance $p < .001$. For example, given two participants with a one year age difference, the odds of making semantic errors is expected to be 18% smaller for the older [$\exp(-0.20) = 0.82 = 1 - 0.18$, $p < .001$]. However, the variation of these two types of errors over time differs, with older participants showing a stronger decrease of semantic errors (estimate = 0.08, $p = 0.038$). Table 2 in [S1 Appendix](#) summarizes the results of the three models and [Fig 1](#) outlines these effects.

Participants made significantly less “other errors”, compared to semantic errors; for example, at 7 years old, the odds of making other error types is expected to be 82% smaller [$\exp(-1.74) = 0.18 = 1 - 0.82$, $p = .003$], compared to the occurrence of semantic errors. The decreasing tendency of making errors over age was not significant for “other errors”, i.e., one additional year of age is expected to reduce the generation of “other errors” in a non-significant amount (estimate = -0.06, $p = .098$). However, this non-significant difference becomes significant when dichotomizing the age scale by considering two groups—children and adolescents. A significant difference on “other errors” was observed between the two groups; more specifically, the odds of making other errors is expected to decrease 37% for an adolescents [$\exp(-0.47) = 0.63 = 1 - 0.37$, $p = .028$]. Notice that in this case we are not analyzing the effect of a one year old (continuous variable), but we are focusing on a group mean difference (categorical variable), which is a weaker constraint. The model describing these between group differences are shown in Appendix Table 4 in [S1 Appendix](#) (in order to better identify this difference in the results of the model, we used the reference level “other errors” of the predictor, so that the results are displayed using this error type as reference).

Discussion

In the current study we have analysed a set of 100 sentence contexts and their cloze probabilities in order to develop a database for Brazilian Portuguese children and adolescents. Additionally, we have specified age-related effects and the predictor effects of age and cloze probabilities on the generation of idiosyncratic responses and errors (semantic, syntactic, and other errors). We observed that a high proportion of sentences (63%) met high-cloze probability criteria, in accordance with Pinheiro et al. (2010) [31]. As most sentence contexts were high-constraint contexts, the number of final word possibilities was limited, thereby eliciting the same most frequent final word [16, 17, 28, 47, 48]. Age was found to affect the probability to generate the most frequent word in 63% of the sentence contexts, whereas no age effects were observed for 47 (76%) contexts that met criteria for high cloze probability. These results are line with previous studies documenting the effects of developmental variables in a cloze task [1, 17, 32].

Of note, the observed age-related changes on cloze probability were not explained by the proportion of semantic or syntactic errors that were mainly observed in younger participants. For example, in sentence number 8 (“The fireman extinguishes the. . .”), adolescents often chose the word “incêndio” (*fire*) instead of “fogo” (*fire*). In English language, these words have the same orthographic representation (“fire”), hence context is necessary to define whether

there is a fire situation outside control (represented by sentence number 8, for which the word “incêndio” would be more appropriate) or whether there is a minor event involving fire (for which the fireman do not need to be called, and the word “fogo” would be a more appropriate selection). In both cases, “fogo” instead of “incêndio” (fire in English) represent evidence of more general representations that were fitted within a specific context. These results are likely to represent developmental effects that reflect the refinement of the semantic system over age as a function of maturational processes and sociocultural factors (e.g., quality of language interactions, years of education) [31]. Accordingly, older participants are more likely to present maturational and sociocultural advantages over younger participants due to a longer and more diversified experience with the language system [31]. In fact, it is well established that an increase in the number of lexical items in semantic memory leads to richer semantic representations and vocabulary knowledge [49]. In the same line, vocabulary knowledge is important for the comprehension of single words, sentences, and narratives. Vocabulary size and vocabulary complexity are significantly increased with age [50–53]. Thus, the older participants are expected to use more refined words to fit the semantic context of a sentence [34, 54, 55], whereas children are expected to choose words reflecting more general representations, as suggested by our findings.

The current results are also consistent with those reported for European Portuguese regarding the number of idiosyncratic responses, semantic and syntactic errors [31]. In accordance, we observed that older participants were less likely to generate idiosyncratic responses and cloze errors, with no difference between semantic and syntactic errors. These findings are consistent with previous studies [31], where an increase in the commonality of responses with age and with years of education was documented [56]. Specifically, an increase of 5% per year of age was observed in the use of the most frequent final word for a given sentence context, and for each year a reduction between 10% and 16% is expected in the number of semantic errors, syntactic errors, and idiosyncratic responses.

Syntactic and semantic errors have been associated with limited semantic knowledge, with lexical gaps or weaker word representations [57]. We observed that most errors committed by younger participants were due to syntactic violations, caused by a violation of grammatical gender (e.g., in sentence number 92: “As pessoas comem no . . .”, “People eat on the . . .”). In Portuguese grammar, the article before the noun will determine both gender and number; i.e., for a given sentence, a masculine and singular noun is required for the final word in the example provided (e.g., “prato”; “plate”). However, younger participants have selected the word “table” (“People eat at the table”), which represents a syntactic violation. In the case of English grammar, both final word “plate” (prato) and “table” (mesa) are possible. In some cases, younger participants committed a double semantic and syntactic violation. For example, in sentence number 16 (“O jardineiro rega o . . .”, “The gardener waters the . . .”), younger participants sometimes completed the sentence with the word “flores” (flowers), which is a female and plural noun, even though the sentence requires a masculine and singular noun as final word.

Sociolinguistic studies conducted in Brazil have documented that plural matching is highly variable depending on the country [58]. The regular pattern of the verbal and nominal agreement is acquired over the years and is used when the individual spends more time in a formal educational context [58, 59]. Further, an electrophysiology study conducted in the Spanish language—which also has gender agreement between article and noun—showed an interaction between semantic congruity and gender agreement in the N400 component. A larger and more frontal N400 was observed for double violations (semantic and syntactic) compared to semantic violations alone. According to the authors, readers are able to anticipate and attend to the gender of both articles and nouns. Gender agreement and semantic congruity interact early in word processing to influence semantic integration of the noun into its sentence

context [60]. Moreover, the grammatical gender agreement rule represents one of the most difficult language aspects to acquire during language development [61], which may account for the current pattern of findings.

In addition to semantic and syntactic errors, children showed two additional types of errors: (1) completion responses that exceeded the use of a single word; and (2) changes in sentence structure, such as replacing the grammatical gender of the article preceding the object plus gender agreement change. These types of errors suggest that younger children have the conceptual representation of the sentence, but fail to choose a specific word related with it. This finding may be related to weaker inhibitory mechanisms operating together with the constraint level of a given sentence [62]. Low-constraint sentences are associated with a larger number of valid final word choices as more features are activated intra-lexically in semantic memory, which brings additional demands to the inhibitory system [62]. Inhibitory control is necessary to suppress the activation of automatic responses that are irrelevant for the task [63]. It is well established that inhibitory skills increase with age [64, 65] and that children show a less effective use of inhibitory mechanisms compared to adolescents, which develop rapidly during early school years [66, 67]. As such, children are expected to present more difficulties in inhibiting lexical items that are semantically related to the sentence-context in their semantic memory, mainly for low-constraint sentences. In accordance, when we included age and cloze probability as predictors in the error analysis, we identified an increased number of errors for younger participants in low cloze sentence contexts. However, the interaction between age and a sentence's constraint level appears to operate differently in adults [33] since age was not found to predict word selection in both high or low constraining sentence contexts.

Of note, the current study may also provide some insights into cross-cultural effects, namely by allowing comparisons of cloze probabilities between Brazilian and European Portuguese [31] children and adolescents. Specifically, most of our high cloze sentence contexts (31 of 49) displayed identical cloze probabilities to in Pinheiro et al. (2010) [31]. The remaining high cloze sentences from our study were classified as low cloze in Pinheiro et al. (2010) [31]. It is worth noting that the relationship between language and culture is complex. It is well known that the behaviors, beliefs, and customs of a given culture guide social interactions and rules that affect the communicative style and the development of language skills [68]. The current study did not control important cultural (educational level and socioeconomical status) and linguistic variables (lexical, syntactic, and morphological) to allow a more in-depth analysis of this relationship. The differences observed may be related to local, regional or more general country differences, impacting upon syntactic organization of sentences [69], as well as semantic lexicon development, and should be more systematically explored in future studies. Future studies should consider not only the cultural and linguistic differences between the two cultural contexts, but also the phonetic and orthographic differences between European vs. Brazilian Portuguese [38].

Finally, the results of the current study have important implications for experimental and clinical settings. The pool of sentence contexts described in our database and their cloze probabilities could be useful for future experimental studies probing language processing. In this realm, the selection of carefully controlled language stimuli often represents substantial challenge to researchers, relying on the use of norms for a given culture and age [70]. This database can be further used in electrophysiological and neuroimaging studies, in both typical and atypical developing populations in different developmental stages. In particular, investigating low and high cloze probability sentences with neuroimaging methods may shed light on the neural correlates of lexical and semantic processing [71]. Previous studies have documented a cloze

probability effect both on the N400 event-related potential [72] and on the BOLD signal, specifically in the superior frontal gyrus and the inferior frontal gyrus [73].

A main limitation of the current study is the unbalanced age sample, characterized by a higher number of adolescents compared to children. Future studies should include bigger samples composed by different age intervals.

Conclusions

The current study provides the first sentence database with cloze probabilities for Brazilian Portuguese. Both age and contextual constraint were found to predict the sentence completion response: final word selection was more consistent in older participants and for high-constraint sentence contexts. Even though similarities were observed for most sentence contexts in Brazilian and European Portuguese [31], some differences were also noted. These findings reinforce the importance of culture-specific norms. These norms are expected to stimulate experimental research on language processing at the sentence level.

Supporting information

S1 Dataset. Rossi_sentences.

(CSV)

S2 Dataset. Rossi.

(CSV)

S1 File. Rossi_R.

(R)

S1 Appendix.

(DOCX)

Author Contributions

Conceptualization: Natalia Freitas Rossi, Célia Maria Giacheti, Ana Patrícia Pinheiro, Adriana Sampaio.

Data curation: Natalia Freitas Rossi, Célia Maria Giacheti, Ana Patrícia Pinheiro, Adriana Sampaio.

Formal analysis: Natalia Freitas Rossi, Catarina Fernandes, Célia Sofia Moreira, Célia Maria Giacheti, Ana Patrícia Pinheiro, Adriana Sampaio.

Funding acquisition: Natalia Freitas Rossi, Adriana Sampaio.

Investigation: Natalia Freitas Rossi, Adriana Sampaio.

Methodology: Natalia Freitas Rossi, Catarina Fernandes, Célia Sofia Moreira, Célia Maria Giacheti, Ana Patrícia Pinheiro, Adriana Sampaio.

Project administration: Natalia Freitas Rossi.

Resources: Natalia Freitas Rossi.

Supervision: Natalia Freitas Rossi, Ana Patrícia Pinheiro, Adriana Sampaio.

Validation: Natalia Freitas Rossi, Catarina Fernandes.

Visualization: Natalia Freitas Rossi, Catarina Fernandes.

Writing – original draft: Natalia Freitas Rossi, Catarina Fernandes, Célia Maria Giacheti, Ana Patrícia Pinheiro, Adriana Sampaio.

Writing – review & editing: Natalia Freitas Rossi, Catarina Fernandes, Célia Sofia Moreira, Célia Maria Giacheti, Bianca Bortolai Sichieri, Ana Patrícia Pinheiro, Adriana Sampaio.

References

1. Johnson MH. Functional brain development in humans. *Nature Reviews Neuroscience*. 2001; 2(7):475–83. <https://doi.org/10.1038/35081509> PMID: 11433372
2. Bishop DV. *Uncommon Understanding (Classic Edition): Development and disorders of language comprehension in children*: Psychology Press; 2014.
3. Ellis C, Kuipers JR, Thierry G, Lovett V, Turnbull O, Jones MW. Language and culture modulate online semantic processing. *Social cognitive and affective neuroscience*. 2015; 10(10):1392–6. <https://doi.org/10.1093/scan/nsv028> PMID: 25767190
4. Friederici AD. The cortical language circuit: from auditory perception to sentence comprehension. *Trends in cognitive sciences*. 2012; 16(5):262–8. <https://doi.org/10.1016/j.tics.2012.04.001> PMID: 22516238
5. Jobard G, Crivello F, Tzourio-Mazoyer N. Evaluation of the dual route theory of reading: a meta-analysis of 35 neuroimaging studies. *Neuroimage*. 2003; 20(2):693–712. [https://doi.org/10.1016/S1053-8119\(03\)00343-4](https://doi.org/10.1016/S1053-8119(03)00343-4) PMID: 14568445
6. Kidd E. The role of verbal working memory in children’s sentence comprehension: A critical review. *Topics in Language Disorders*. 2013; 33(3):208–23.
7. Tulving E. Episodic and semantic memory. *Organization of memory*. 1972; 1:381–403.
8. Kutas M, Federmeier KD. Electrophysiology reveals semantic memory use in language comprehension. *Trends in cognitive sciences*. 2000; 4(12):463–70. [https://doi.org/10.1016/s1364-6613\(00\)01560-6](https://doi.org/10.1016/s1364-6613(00)01560-6) PMID: 11115760
9. Démonet J-F, Thierry G, Cardebat D. Renewal of the neurophysiology of language: functional neuroimaging. *Physiological reviews*. 2005; 85(1):49–95. <https://doi.org/10.1152/physrev.00049.2003> PMID: 15618478
10. Hahne A, Eckstein K, Friederici AD. Brain signatures of syntactic and semantic processes during children’s language development. *Journal of Cognitive Neuroscience*. 2004; 16(7):1302–18. <https://doi.org/10.1162/0898929041920504> PMID: 15453981
11. Lau EF, Phillips C, Poeppel D. A cortical network for semantics:(de) constructing the N400. *Nature Reviews Neuroscience*. 2008; 9(12):920–33. <https://doi.org/10.1038/nrn2532> PMID: 19020511
12. Cummings A, Čeponienė R, Dick F, Saygin AP, Townsend J. A developmental ERP study of verbal and non-verbal semantic processing. *Brain Research*. 2008; 1208 137–49. <https://doi.org/10.1016/j.brainres.2008.02.015> PMID: 18387601
13. Collins AM, Loftus EF. A spreading-activation theory of semantic processing. *Psychological review*. 1975; 82(6):407.
14. Collins AM, Quillian MR. Retrieval time from semantic memory. 1969.
15. Loftus EF. Activation of semantic memory. *The American Journal of Psychology*. 1973:331–7.
16. Kellas G, Paul ST, Martin M, Simpson GB. Contextual feature activation and meaning access. *Understanding word and sentence*. 1991:47–71.
17. Holcomb PJ, Coffey SA, Neville HJ. Visual and auditory sentence processing: A developmental analysis using event-related brain potentials. *Developmental Neuropsychology*. 1992; 8(2–3):203–41.
18. Block CK, Baldwin CL. Cloze probability and completion norms for 498 sentences: Behavioral and neural validation using event-related potentials. *Behavior Research Methods*. 2010; 42(3):665–70. <https://doi.org/10.3758/BRM.42.3.665> PMID: 20805588
19. Kutas M, Hillyard SA. Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*. 1980; 207(4427):203–5. <https://doi.org/10.1126/science.7350657> PMID: 7350657
20. Kutas M, Hillyard SA. Brain potentials during reading reflect word expectancy and semantic association. *Nature*. 1984; 307(5947):161–3. <https://doi.org/10.1038/307161a0> PMID: 6690995
21. Kimble MO, Kaufman ML, Leonard LL, Nestor PG, Riggs DS, Kaloupek DG, et al. Sentence completion test in combat veterans with and without PTSD: Preliminary findings. *Psychiatry Research*. 2002; 113(3):303–7. [https://doi.org/10.1016/s0165-1781\(02\)00229-9](https://doi.org/10.1016/s0165-1781(02)00229-9) PMID: 12559486

22. Kircher TT, Bulimore ET, Brammer MJ, Williams SC, Broome MR, Murray RM, et al. Differential activation of temporal cortex during sentence completion in schizophrenic patients with and without formal thought disorder. *Schizophrenia research*. 2001; 50(1–2):27–40. [https://doi.org/10.1016/s0920-9964\(00\)00042-6](https://doi.org/10.1016/s0920-9964(00)00042-6) PMID: 11378312
23. Nebes RD, Brady CB. The effect of contextual constraint on semantic judgments by Alzheimer patients. *Cortex*. 1991; 27(2):237–46. [https://doi.org/10.1016/s0010-9452\(13\)80128-1](https://doi.org/10.1016/s0010-9452(13)80128-1) PMID: 1879152
24. Staub A, Grant M, Astheimer L, Cohen A. The influence of cloze probability and item constraint on cloze task response time. *Journal of Memory and Language*. 2015; 82:1–17.
25. Taylor WL. “Cloze procedure”: A new tool for measuring readability. *Journalism quarterly*. 1953; 30(4):415–33.
26. Taylor WL. Recent developments in the use of “cloze procedure”. *Journalism quarterly*. 1956; 33(1):42–99.
27. Bloom PA, Fischler I. Completion norms for 329 sentence contexts. *Memory & cognition*. 1980; 8(6):631–42.
28. Arcuri SM, Rabe-Hesketh S, Morris RG, McGuire PK. Regional variation of cloze probabilities for sentence contexts. *Behavior Research Methods, Instruments, & Computers*. 2001; 33(1):80–90.
29. Robichon F, Besson M, Fata F. Completion norms for 744 French linguistic contexts of differing formats. 1996.
30. Rodríguez-Camacho MA, Prieto-Corona B, Bravo M, Marosi-Holzberger E, Bernal-Hernández J, Yáñez-Télez G. Normas de terminación para la palabra final de oraciones en español para niños mexicanos. *Avances en psicología latinoamericana*. 2011; 29(2):258–75.
31. Pinheiro A, Soares A, Comesaña M, Niznikiewicz M, Gonçalves Ó. Sentence-final word completion norms for European Portuguese children and adolescents. *Behavior Research Methods*. 2010; 42(4):1022–9. <https://doi.org/10.3758/BRM.42.4.1022> PMID: 21139169
32. Lahar CJ, Tun PA, Wingfield A. Sentence-final word completion norms for young, middle-aged, and older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*. 2004; 59(1):P7–P10. <https://doi.org/10.1093/geronb/59.1.p7> PMID: 14722333
33. Hamberger MJ, Friedman D, Rosen J. Completion norms collected from younger and older adults for 198 sentence contexts. *Behavior Research Methods, Instruments, & Computers*. 1996; 28(1):102–8.
34. Geeraerts D, Grondelaers S, Bakema P. The structure of lexical variation: Meaning, naming, and context: Walter de Gruyter; 2012.
35. Yoon C, Feinberg F, Hu P, Gutchess A, Hedden T, Chen HM, et al. Category norms as a function of culture and age: comparisons of item responses to 105 categories by American and Chinese adults. *Psychological Aging*. 2004; 19:379–93.
36. Ji L-J, Zhang Z, Nisbett RE. Is it culture or is it language? Examination of language effects in cross-cultural research on categorization. *Journal of Personality and Social Psychology*. 2004; 87(1):57. <https://doi.org/10.1037/0022-3514.87.1.57> PMID: 15250792
37. Barreiro A, Wittmann LH, Jesus-Pereira M. Lexical differences between European and Brazilian Portuguese. *INESC Journal of Research and Development*. 1996; 5(2):75–101.
38. Frota S, Cruz M, Fernandes-Svartman F, Collischonn G, Fonseca A, Serra C, et al. Intonational variation in Portuguese: European and Brazilian varieties. In: Prieto F, editor. *Intonation Romance*. Oxford: University Press 2015. p. 235–83.
39. Biderman MTC. O Português Brasileiro e o Português Europeu: Identidade e contrastes. *Revue belge de philologie et d'histoire*. 2001; 79(3):963–75.
40. Cáceres-Assenço A, Ferreira S, Santos A, Befi-Lopes D. Aplicação de uma prova brasileira de vocabulário expressivo em crianças falantes do Português Europeu. *CoDAS* 2018; 30(2).
41. Ferreira SCA, Cruz-Santos A. Processos de designação e substituição semântica usados por crianças falantes de Português Europeu numa prova de vocabulário. *Audiology-Communication Research*. 2018; 23.
42. Brauer J, Friederici AD. Functional neural networks of semantic and syntactic processes in the developing brain. *Journal of Cognitive Neuroscience*. 2007; 19(10):1609–23. <https://doi.org/10.1162/jocn.2007.19.10.1609> PMID: 18271736
43. Brasil C. Critério de classificação econômica Brasil. Associação Brasileira de Empresas de Pesquisa (ABEP). 2013.
44. Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, et al. glmmTMB: balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*. 2017; 9(2):378–400.

45. Fox J, Weisberg S. Visualizing fit and lack of fit in complex regression models with predictor effect plots and partial residuals. *Journal of Statistical Software*. 2018; 87(9):1–27.
46. Allison PD. *Logistic regression using SAS: Theory and application*. second ed. Cary, NC, USA: SAS Institute Inc; 2012.
47. Schneider W, Pressley M. *Memory development between two and twenty*: Psychology Press; 2013.
48. Schwanenflugel PJ, Shoben EJ. The influence of sentence constraint on the scope of facilitation for upcoming words. *Journal of Memory and Language*. 1985; 24(2):232–52.
49. Gathercole SE, Willis CS, Emslie H, Baddeley AD. Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental psychology*. 1992; 28(5):887.
50. Carneiro MP, Albuquerque P, Fernandez A, Esteves F. Normas de associação livre de 16 palavras portuguesas para crianças de diferentes faixas etárias. *Laboratório de Psicologia*. 2004; 2(1):49–76.
51. Munson B, Swenson CL, Manthei SC. Lexical and phonological organization in children: Evidence from repetition tasks. *Journal of Speech, Language, and Hearing Research*. 2005; 48:109–24.
52. Storkel HL. Restructuring of similarity neighbourhoods in the developing mental lexicon. *Journal of Child Language*. 2002; 29(2):251–74. <https://doi.org/10.1017/s0305000902005032> PMID: 12109371
53. Storkel HL. Developmental differences in the effects of phonological, lexical and semantic variables on word learning by infants. *Journal of Child Language*. 2009; 36(2):291–321. <https://doi.org/10.1017/S030500090800891X> PMID: 18761757
54. Marsh EJ, Tversky B. Spinning the stories of our lives. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*. 2004; 18(5):491–503.
55. Principe GF, Smith E. The tooth, the whole tooth and nothing but the tooth: How belief in the tooth fairy can engender false memories. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*. 2008; 22(5):625–42.
56. Rosenzweig MR. Experience, memory, and the brain. *American Psychologist*. 1984; 39(4):365. <https://doi.org/10.1037//0003-066x.39.4.365> PMID: 6372557
57. McGreor K, Friedman R, Reilly R, Newman R. Semantic representation and naming in young children. *Journal of Speech, Language, & Hearing Research*. 2002; 332–346:332–46.
58. Lima WF. A variação da concordância verbal em textos escolares escritos. *ALFA: Revista de Linguística*. 2001; 45.
59. Capellari ET, Zilles AMS. A marcação de plural na linguagem infantil: estudo longitudinal. *Revista da ABRALIN Rio de Janeiro Vol 1, n 1 (jul 2002)*, p 185–218. 2002.
60. Wicha NY, Moreno EM, Kutas M. Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*. 2004; 16(7):1272–88. <https://doi.org/10.1162/0898929041920487> PMID: 15453979
61. Loerts H, Wieling M, Schmid MS. Neuter is not common in Dutch: Eye movements reveal asymmetrical gender processing. *Journal of psycholinguistic research*. 2013; 42(6):551–70. <https://doi.org/10.1007/s10936-012-9234-2> PMID: 23229089
62. Simpson GB, Peterson RR, Casteel MA, Burgess C. Lexical and sentence context effects in word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1989; 15(1):88.
63. Canet-Juric L, Urquijo S, Richards MM, Burin DI. Predictores cognitivos de niveles de comprensión lectora mediante análisis discriminante. *International Journal of Psychological Research*. 2009; 2(2):99–111.
64. Best JR, Miller PH, Naglieri JA. Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and individual differences*. 2011; 21(4):327–36. <https://doi.org/10.1016/j.lindif.2011.01.007> PMID: 21845021
65. Xu F, Han Y, Sabbagh MA, Wang T, Ren X, Li C. Developmental differences in the structure of executive function in middle childhood and adolescence. *PLoS One*. 2013; 8(10).
66. Davidson MC, Amso D, Anderson LC, Diamond A. Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*. 2006; 44(11):2037–78. <https://doi.org/10.1016/j.neuropsychologia.2006.02.006> PMID: 16580701
67. Macdonald JA, Beauchamp MH, Crigan JA, Anderson PJ. Age-related differences in inhibitory control in the early school years. *Child Neuropsychology*. 2014; 5:509–26.
68. Minami M. *Culture-specific language styles: The development of oral narrative and literacy* Cromwell Press; 2002.
69. Larina T. Culture-specific communicative styles as a framework for interpreting linguistic and cultural idiosyncrasies. *International Review of Pragmatics*. 2015; 7(2):195–215.

70. Rossi NF. Uso das técnicas de ressonância magnética funcional e eletroencefalografia nos estudos sobre o desenvolvimento da linguagem In: Giacheti CM, editor. Avaliação da fala e da linguagem Mrília, São Paulo: Oficina Universitária; 2016. p. 109–32.
71. Federmeier KD, Wlotko EW, De Ochoa-Dewald E, Kutas M. Multiple effects of sentential constraint on word processing. *Brain research*. 2007; 1146:75–84. <https://doi.org/10.1016/j.brainres.2006.06.101> PMID: [16901469](https://pubmed.ncbi.nlm.nih.gov/16901469/)
72. DeLong KA, Urbach TP, Groppe DM, Kutas M. Overlapping dual ERP responses to low cloze probability sentence continuations. *Psychophysiology*. 2011; 48(9):1203–7. <https://doi.org/10.1111/j.1469-8986.2011.01199.x> PMID: [21457275](https://pubmed.ncbi.nlm.nih.gov/21457275/)
73. Zhu Z, Bastiaansen M, Hakun JG, Petersson KM, Wang S, Hagoort P. Semantic unification modulates N400 and BOLD signal change in the brain: A simultaneous EEG-fMRI study. *Journal of Neurolinguistics*. 2019; 52:100855.