Volume 27, n.1 (2023) e40735

Veredas
Revista de Estudos Linguísticos
https://periodicos.ufjf.br/index.php/veredas/index

The influences of prosodic phrasing and constituent length on garden-path sentences in Brazilian Portuguese

As influências do fraseamento prosódico e comprimento de constituintes sobre frases *garden-path* no Português Brasileiro

Vitor Gabriel Caldas[1]

Universidade Federal do Rio de Janeiro, Brasil

RESUMO

Este artigo explora o papel da prosódia no processamento de frases *garden-path* no português brasileiro. A estrutura investigada apresenta um verbo *dicendi* com duas orações iniciadas por "que" ("A aluna$_i$ disse à professora$_j$ que $_{-j}$ estava atrasada que $_{-i}$ precisaria sair da sala"). A primeira oração com "que" é temporariamente ambígua, apresentando análise de oração completiva (OC) ou oração relativa (OR). Assumimos que o fraseamento prosódico atua na resolução da ambiguidade, corroborando modelos de processamento em que a prosódia opera antes da sintaxe. Uma tarefa do tipo *completion* foi conduzida e 40 participantes ouviram a porção inicial de sentenças e escolheram uma resposta entre duas opções para completar o fragmento ouvido. Os resultados mostraram que a prosódia influenciou a análise dos participantes sobre os estímulos auditivos e eles escolheram a opção compatível com o fraseamento prosódico ao qual eles tinham sido expostos.

PALAVRAS-CHAVE:

Fraseamento prosódico. Processamento de frases. *Garden-path*. Tarefa de *completion*.

ABSTRACT

This paper explores the role of prosody in the processing of garden-path sentences in Brazilian Portuguese. The structure investigated presents a *dicendi* ('saying') matrix verb with two that-clauses (A aluna$_i$ disse à professora$_j$ que $_{-j}$ estava atrasada que $_{-i}$ precisaria sair da sala - 'The student told the teacher that was late that she would need to leave the room'). The first that-clause is temporarily ambiguous between a Complement Phrase (CP) and a Relative Clause (RC) analysis. We assume that prosodic phrasing influences the resolution of the ambiguity, supporting prosody-first accounts of processing. A Cross-modal Completion Task was carried out and 40 participants listened to the initial portion of sentences and chose an answer between two options to complete the fragment heard. The results showed that prosody influenced subjects' parsing of the auditory stimuli and they chose the option compatible with the prosodic phrasing they had been exposed to.

Recebido em: 03/04/2023

Aceito em: 25/07/2023

[1] E-mail: vitorgabrielish@gmail.com |ORCID: https://orcid.org/0000-0002-1934-9072.

CC BY SA

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

KEYWORDS:
Prosodic phrasing. Sentence processing. Garden-path. Completion task.

1. Introduction

This study aims at investigating the role of prosody in sentence processing testing garden-path sentences in Brazilian Portuguese (BP)[2]. Garden-path sentences present a temporary ambiguity which leads individuals to a misinterpretation. At some point in the sentence, individuals encounter some syntactic information that is inconsistent with the previous linguistic representation they had built of the sentence. When this happens, we say that a garden-path effect occurred. After this garden-path, individuals are expected to engage in reanalysis and fix the problem.

The syntactic structure investigated here was studied by Silva (2018) and inspired our research. The author carried out a reading experiment using an eye-tracker equipment to explore the syntactic processing and final interpretation of sentences by middle school and undergraduate Brazilian students. The investigated sentences present a *dicendi* ('saying') matrix verb with two that-clauses. The first clause is temporarily ambiguous between a Complement Phrase (CP) and a Relative Clause (RC) analysis. See the following example.

(1) A aluna$_i$ disse à professora$_j$ que –$_j$ estava atrasada que –$_i$ precisaria sair da sala
'The student told the teacher that was late that she would need to leave the room'

In the above sentence, the RC "que estava atrasada" ('that was late') is temporarily ambiguous and it is firstly misparsed as a CP of the matrix verb "disse" ('told'). This ambiguity has, at least, two reasons: (1) the word "que" ('that') can be either the complementizer of a CP or the relative pronoun that introduces an RC, and (2) since Portuguese is a language that allows null subjects, CPs can have the word "que" ('that') followed by an empty category, presenting, therefore, the same string of an RC. In comparison to a language that does not allow null subjects, such as English, "that was late" can only be analyzed as an RC.

[2] The research findings reported in this paper come from an experiment carried out in the author's doctoral dissertation (Caldas, 2022), which was supervised by Professors Carolina Serra (UFRJ) and Marcus Maia (UFRJ), and defended in the Graduate Program in *Letras Vernáculas* (PPGLEV-UFRJ). This study was also supervised by Professor Sun-Ah Jun (UCLA), during the author's doctoral internship (CAPES-PrInt Grant, Process n° 88887.508099/2020-00) at the University of California, Los Angeles (UCLA).

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

But why do we analyze the ambiguous clause as a CP at first? There might be a few reasons. The Garden-Path Model (Frazier, 1979) would explain this by using the Minimal Attachment (MA) principle, which states that the parser should not postulate any potentially unnecessary syntactic nodes. RCs are optional and, therefore, cannot be predicted by the reader/listener. In this way, the parser prefers to analyze the ambiguous clause as a CP. We could also explain this preference with an argument-first strategy. Since saying verbs like "dizer" ('say'), "contar" ('tell'), "falar" ('speak') select a CP, whenever we encounter a that-clause after these verbs, we firstly parse it as a CP in order to satisfy the argument structure of the verb.

In the experiment reported here, we wanted to explore the influence of prosody on listeners' initial syntactic structuring of the sentence. The assignment of syntactic structure given by prosody has been shown by previous works on early/late closure ambiguities (Kjelgaard & Speer, 1999; Fonseca, 2012). The novelty here is that we are studying a Minimal Attachment (MA) construction.

We believe that, depending on the initial prosodic phrasing of the sentences, the parser will either analyze the ambiguous clause as a CP or as an RC. In broader terms, we assume that prosody is prior to syntax during the course of processing and that the prosodic representation determines the syntactic structure of sentences. In order to test these hypotheses, we carried out a Cross-modal Completion Task[3], in which participants listened to the initial portion of sentences and were presented with two alternative answers to complete them on a computer screen. Considering that our task does not provide real-time online measures, we do not strictly adopt a specific model of processing. However, we support prosody-first accounts of processing such as Schafer's (1997) and Blodgett's (2004). The details of our experiment are presented in the third section of this paper; before that, we briefly summarize some studies that somehow inspired our research.

2. Previous studies

Hirose (2003) investigated a Japanese structure that is similar to the one studied here and carried out two experiments which inspired our experimental design. The sentences presented an ambiguous analysis between a matrix clause and a relative clause. See the example below.

[3] I am very grateful to Professor Jesse Harris (UCLA), who gave me the idea for this experiment, during my doctoral internship at the University of California, Los Angeles (UCLA).

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

(2) Mori'sita-ga si'nyaku-o kokoro'kara sinyoosita yuuji'ntati-ni . . .

'Morisita$_{\text{-Nom}}$ new medicine$_{\text{-Acc}}$ truly trusted friends$_{\text{-Dat}}$'

(Hirose, 2003, p. 168)

Considering that the processor is incremental and follows the Minimal Attachment principle, during the first parsing, the parser would initially analyze the word string up to the first verb (shinyooshita - 'trusted') as a single clause (associated with the meaning "Morishita truly trusted the new medicine"). When the parser encounters the head noun (yuuji'ntati - 'friends'), the analysis of the clause as a matrix clause is not supported and reanalysis towards a relative clause should occur. The opening of the relative clause may happen either before or after the accusative NP si'nyaku-o ('new medicine'). If the matrix verb is monotransitive like a'tta ('met'), the relative clause opens before the accusative NP (Early Opening/EO). If the matrix verb is ditransitive like mi'seta ('showed'), the relative clause begins after the accusative NP (Late Opening/LO).

Researchers have long agreed that the LO structure is more costly to process than its EO counterpart. However, this view contradicts native-speaker intuitions. Hirose (2003) proposes that this disagreement may exist due to length/prosodic effects on ambiguity resolution. In order to account for this, the linguist carried out two offline sentence completion questionnaires to test length effects of sentence-initial subject phrase on subjects' preference in clause boundary ambiguity.

Experiment 1 compared 12 pairs of sentence fragments containing single proper name subjects to conjoined proper name subjects and experiment 2 compared other 12 pairs of sentence fragments with either a single name subject or a full name subject. The tasks consisted of paper-and-pencil sentence-completion questionnaires, in which subjects had to complete each fragment by writing in words, so as to create a grammatical sentence. Subjects' responses were coded either as EO or LO, based on the content implied and the verb used. Responses with monotransitive verbs were coded as EO and ditransitive verbs were coded as LO.

In both experiments, the results showed a preference for EO independently of the conditions. However, in the long subjects (conjoined and full names), the preference for EO was even greater and the statistical analysis showed that the likelihood of an EO interpretation significantly exceeded 50% only for the long subjects. This result supported the hypothesis that length has an effect on the choice between EO and LO structures.

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

The author's hypothesis that long subjects affect syntactic parsing decisions is based on the prosodic organization of Japanese. In Japanese, two intonationally defined prosodic constituents are established: minor phrase and major phrase. Minor phrases are grouped into major phrases. The author proposes that differences in major phrasing may affect the resolution of EO/LO clause boundary ambiguity. In the items used in the experiments, both full name and conjoined name subjects are theoretically composed by two minor phrases, which would, therefore, form one major phrase, whereas single name subjects contained only one minor phrase.

Clifton, Carlson & Frazier (2006) presented three experiments that tested their Rational Speaker Hypothesis (RSH), which claims that (1) speakers are self-consistent, employing intonation in a manner consistent with their intended message and (2) listeners interpret intonation by assuming that speakers do not make prosodic choices without some reason (and are, therefore, rational). Considering this, the authors aimed to investigate whether the presence of an Intonational Phrase (IP) boundary becomes less informative to the listener when there is more than one reason for its existence.

Prosodic boundaries can be used in speech either to signal the syntactic structure of the sentence or as a matter of phonological weight. Speakers tend to place prosodic boundaries before and after long constituents. In this case, the function of a prosodic boundary that flanks a long constituent would be ambiguous and, therefore, less informative to the listener when compared to a boundary that flanks a short constituent whose existence could only signal syntax. In order to test this hypothesis, the authors carried out three auditory questionnaires. In the first two tasks, subjects listened to sentences containing short or long constituents marked by early or late boundaries. See the following examples.

(3) **Short constituent, early boundary:** $[\text{Pat}]_{IP}$ or $[\text{Jay and Lee}]_{IP}$ convinced the bank president to extend the mortgage.
(4) **Short constituent, late boundary:** $[\text{Pat or Jay}]_{IP}$ and $[\text{Lee}]_{IP}$ convinced the bank president to extend the mortgage.
(5) **Long constituent, early boundary:** $[\text{Patricia Jones}]_{IP}$ or $[\text{Jacqueline Frazier and Letitia Connolly}]_{IP}$ convinced the bank president to extend the mortgage.
(6) **Long constituent, late boundary:** $[\text{Patricia Jones or Jacqueline Frazier}]_{IP}$ and $[\text{Letitia Connolly}]_{IP}$ convinced the bank president to extend the mortgage.

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

(Clifton, Carlson & Frazier, 2006, p. 855)

In these sentences, the prosodic phrasing employed in the sentence reflects its syntactic structure. The early boundary conditions are associated with the "(X) or (Y and Z)" interpretation and the late boundary conditions with the "(X or Y) and (Z)" interpretation.

The mean percentage of early boundary interpretations ("(X) or (Y and Z) analysis") was the highest (81,7%) for the condition with a short constituent and an early boundary. The mean RT for the same condition was also the fastest (3,490ms) among all conditions. The differences between early boundary choices for short and long conditions were significant, showing that prosody had a greater effect on the short constituents. These results support the authors' hypothesis on the informativeness of the prosodic boundaries. When the boundary flanks long constituents, it is less informative about syntax.

In the third task, the authors explored the presence versus absence of an IP boundary before short and long constituents. See the material set below.

(7) **Short adverb phrase, no boundary:** Susie learned that Bill telephoned last night.
(8) **Short adverb phrase + boundary:** Susie learned that Bill telephoned$_{IP}$ last night.
(9) **Long adverb phrase, no boundary:** Susie learned that Bill telephoned late last night after the general meeting.
(10) **Long adverb phrase + boundary:** Susie learned that Bill telephoned$_{IP}$ late last night after the general meeting.

(Clifton, Carlson & Frazier, 2006, p. 857)

The conditions contain either a short or a long adverb phrase that can modify the matrix verb "learned" or the complement verb "telephoned" and the presence or absence of an IP boundary before the adverb phrase. According to the RSH, the condition with the short adverb phrase and the boundary should present the highest occurrence of high attachments, since the purpose of the boundary in this case is purely syntactic.

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

A main effect for the interaction between the presence versus absence of the IP boundary before the adverb phrase and the length of the constituent was found. The differences between the number of high attachment choices showed a main effect only for the short conditions, in which the presence of the IP boundary significantly increased the percentage of matrix modification interpretations (24,4% vs. 39,5%). These results support the RSH, showing that the flanking of short constituents is more informative about syntax and, therefore, affects sentence comprehension.

Fonseca (2012) carried out a self-paced listening experiment testing early/late closure ambiguities in subordination and NP coordination structures, in both Brazilian and European Portuguese, as the examples below illustrate.

(11) **Subordination**

Enquanto Maria costurava as camisas ouviu a propaganda na TV.

'While Mary mended the shirts (she) heard the propaganda on TV.'

(12) **NP Coordination**

A mãe castigou o Paulo e o Bruno trancando os meninos no quarto.

'The mom grounded Paul and Bruno locking up the boys in the bedroom.'

(Fonseca, 2012, p. 200)

One of the objectives of this experiment was to test whether prosody was able not only to provoke the garden-path effect by misleading the parser to a misinterpretation analysis, but also to preclude processing difficulty by guiding the parser to the correct analysis. All the experimental sentences were recorded in two prosodic conditions: one in which the prosodic structure is incongruent with the syntactic structure (GP prosody) and one in which prosody coincides with syntax (Reanalysis prosody).

A main effect of prosody was found in the Early Closure (EC) and Late Closure (LC) versions of the sentences, both in Brazilian and in European Portuguese. The GP prosody conditions presented longer Reaction Times (RTs) compared to the Reanalysis condition both in the EC and LC sentences.

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

These results are aligned with the ones found by Kjelgaard & Speer (1999) testing similar sentences in English and support the hypothesis about the influence of prosody on syntactic parsing. The results found for the sentences in the LC condition show that GP prosody guides the parser to place a syntactic boundary earlier in the sentence, contradicting a preferred structure and, therefore, the LC principle. The author provides reliable evidence that prosody is accessed early during the course of processing by listeners and guides the parser's syntactic structuring by either deceiving the parser to construct a misleading syntactic analysis (generating the garden-path effect) or conducting the parser to assign the correct syntactic structure.

Silva (2018) studied the same type of garden-path sentences investigated in this paper. The author carried out a reading experiment using an eye-tracker equipment to explore the syntactic processing and final comprehension of both garden-path (GP) and unambiguous (NGP) sentences by middle school and undergraduate students in Brazil. The garden-path sentences presented an ambiguous analysis between a complement clause and a relative clause. The disambiguated sentences were used as a control for comparison with the ambiguous sentences. See the examples below.

(13) **GP:** "A aluna$_i$ disse à professora$_j$ que $_{-j}$ chegou tarde que $_{-i}$ precisava sair mais cedo."
'The student told the teacher that was late that she needed to leave earlier.'

(14) **NGP:** "A aluna disse à professora que chegou tarde porque perdeu o primeiro ônibus."
'The student told the teacher that was late because she missed the first bus.'

(Silva, 2018, p. 72)

For the analysis of syntactic processing, the author measured the total fixation duration (TFD) and the fixation count (FC) in three regions of interest. The results showed a main effect of the factor Type of sentence (GP vs NGP) for both the TFD and the FC measures. The TFD means were significantly longer during the reading of both the ambiguous clause and the second *that*-clause of the GP sentences for both the middle school and undergraduate students. The FC means were also significantly higher in the CP of the GP sentences for the two groups of subjects. This data provides nice evidence that participants fell into a garden-path during the reading of the CP of the GP sentences.

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

In order to assess the subject's final interpretation of the sentences, a comprehension question with two alternative answers was presented after the reading of each sentence. The rate of error responses was significantly higher in the GP sentences for both groups of subjects.

Silva (2018) showed that the reading of the investigated structure leads subjects to a garden-path effect. But what if they listened to these sentences pronounced with different patterns of prosodic phrasing? Would prosody be able to guide listeners to different syntactic parsings? To answer these questions, we carried out the experiment reported in the next section.

Almeida *et al*. (2021) performed an offline auditory questionnaire to analyze the role of prosody in the resolution of the global ambiguity in sentences of the type NP1 - V - NP2 - Adverb of place - Adverb of intensity - Attribute. See the example below.

(15) O guitarrista recebeu o baterista no quarto bastante drogado.
'The guitarist received the drummer in the room very drugged.'

(Almeida *et al*., 2021, p. 7)

In the example, the adjective "drogado" ('drugged') can either be attached locally, modifying the low NP "baterista" ('drummer'), or non-locally, modifying the high NP "guitarrista" ('guitarist'). In order to test the influence of prosodic cues, both in isolation and in conjunction, on disambiguation, four conditions were created: (i) neutral (N), no prosodic cues; (ii) focus (F), a /H+L*/ pitch accent on the subject and on "bastante" ('very'); (iii) pause (P), a 200-millisecond pause before "bastante" ('very'); (iv) focus and pause (FP), both prosodic cues present.

Participants' task was to listen to the auditory stimuli as many times as they wanted and select, on a five-point Likert scale, the option that best matched their judgment regarding who the adjective referred to (NP1 or NP2). The adoption of a Likert scale is a powerful paradigm to measure participants' confidence in their responses depending on the prosodic cues they were exposed to.

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

The results showed a preference for non-local attachment judgments in all analyzed conditions, supporting the first-mention factor (Cuetos e Mitchell, 1988). However, the choice of assigning the adjective to the NP1 increased, in a statistically significant way, in the prosodically manipulated conditions. The judgment mean for the NP1 in the Neutral condition was significantly different compared to the means of the three other conditions (F, P, and FP), but the differences among the means of the conditions containing the prosodic cues were not significant. This means that, for the NP1, no prosodic cue was more important than the other nor their conjunction.

The data regarding the NP2 was very similar to the one found for the NP1. The judgment mean in the Neutral condition was significantly different compared to the means of the three other conditions (F, P, and FP). However, a significant difference was found between the F and P conditions.

3. Experiment: Cross-modal Completion Task

As mentioned in the Introduction, in this study, participants listened to the initial portion of sentences and were presented with two alternative answers to complete them on a computer screen. They were instructed to choose the best option to complete each fragment, so as to create a grammatical sentence. This paradigm provides measures of processing on subjects' initial parsing and serves as an indication of whether prosody influenced participants' comprehension of sentences.

The objectives of this experiment were to assess (1) if the prosodic phrasing of the initial portion of the garden-path sentence (the word string up to the second that-clause) is able to guide listeners' initial parsing of the sentence, favoring one syntactic analysis over the other, and (2) if, when constituent length is increased and the function of a prosodic boundary becomes ambiguous (Clifton *et al*., 2006), listeners are sensitive to this and, therefore, accept other constructions for the sentences.

The main hypothesis of this experiment is that the placement of prosodic boundaries is accessed by listeners, determining the initial parsing they build of the sentence. To better explain this hypothesis, we take the sentence fragment below.

(16) [A aluna]$_{\text{subj-DP}}$ [disse]$_{\text{verb}}$ [à professora]$_{\text{dat-PP}}$ [que estava atrasada]$_{\text{that-clause}}$
'The student told the teacher that was late'

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

If the dative prepositional phrase (PP) “à professora” (‘the teacher’) is phrased together with the that-clause “que estava atrasada” (‘that was late’), the parser will analyze it as a restrictive RC. The parser follows this analysis because, since both the PP and the that-clause are grouped in the same IP, prosody tells it that the that-clause should be attached to the antecedent noun “professora” (‘teacher’).

On the other hand, if an IP boundary occurs before the that-clause, the parser will analyze it as a CP. The prosodic break between the PP and the that-clause tells the parser that the upcoming clause should not be attached to the previous noun and, therefore, non-local attachment should happen.

In this experiment, we also wanted to test length effects on the initial parsing. It is well established by the Rational Speaker Hypothesis (Clifton *et al*., 2006) that, when constituent length is increased, the informativeness of a prosodic boundary becomes ambiguous. After long constituents, prosodic boundaries may either signal syntactic structure or occur as a matter of phonological weight. Considering that, observe the sentence fragment below.

(17) [A aluna]$_{\text{subj-DP}}$ [disse]$_{\text{verb}}$ [à professora de português da escola]$_{\text{dat-PP}}$ [que estava atrasada]$_{\text{that-clause}}$

‘The student told the Portuguese teacher of the school that was late’

As we can see, the initial portion of the sentence (up to the that-clause) is long. The function of an IP boundary before the that-clause would, therefore, be ambiguous to the listener. In this case, we hypothesize that maybe the effect of the prosodic break on favoring the parsing of the upcoming clause as a CP would be weaker compared to when the portion before the that-clause is short. On the other hand, when prosody favors an RC parsing of the ambiguous clause, the IP boundaries are placed in the same positions both in the short and long conditions. However, we have to consider that the IP that maps the RC will be shorter or longer and this may also account for a weaker effect of prosody when it flanks a longer constituent.

3.1 Experimental design

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

The experiment presented a 2x2 within-subjects factorial design. The independent variables of this experiment were (1) the prosodic phrasing of the initial portion of the garden-path sentences and (2) the length of the constituents before the that-clause. Variable (1) was divided into two levels: (i) the prosodic phrasing favoring the parsing of the ambiguous clause as an RC and (ii) the prosodic phrasing favoring the parsing of the ambiguous clause as a CP. Variable (2) was divided into two other levels: (i) short and (ii) long. The short condition is composed by a single noun subject + a matrix *dicendi* verb + a single dative noun + the ambiguous that-clause (always formed with the verbal form "estava" - 'was'). The long condition was modified so that the dative noun always has two prosodic words attached to it, which can be either adjectives or prepositional phrases. The four conditions are illustrated in Table 1.

Table 1: Experimental conditions

Conditions	Examples
RC-Short (RCS)	**A aluna disse%[4] à professora que estava atrasada%** 'The student told the teacher that was late'
CP-Short (CPS)	**A aluna disse à professora% que estava atrasada%** 'The student told the teacher that was late'
RC-Long (RCL)	**A aluna disse% à professora de português da escola que estava atrasada%** 'The student told the Portuguese teacher of the school that was late'
CP-Long (CPL)	**A aluna disse à professora de português da escola% que estava atrasada%** 'The student told the Portuguese teacher of the school that was late'

[4] The "%" symbol indicates the placement of an IP boundary.

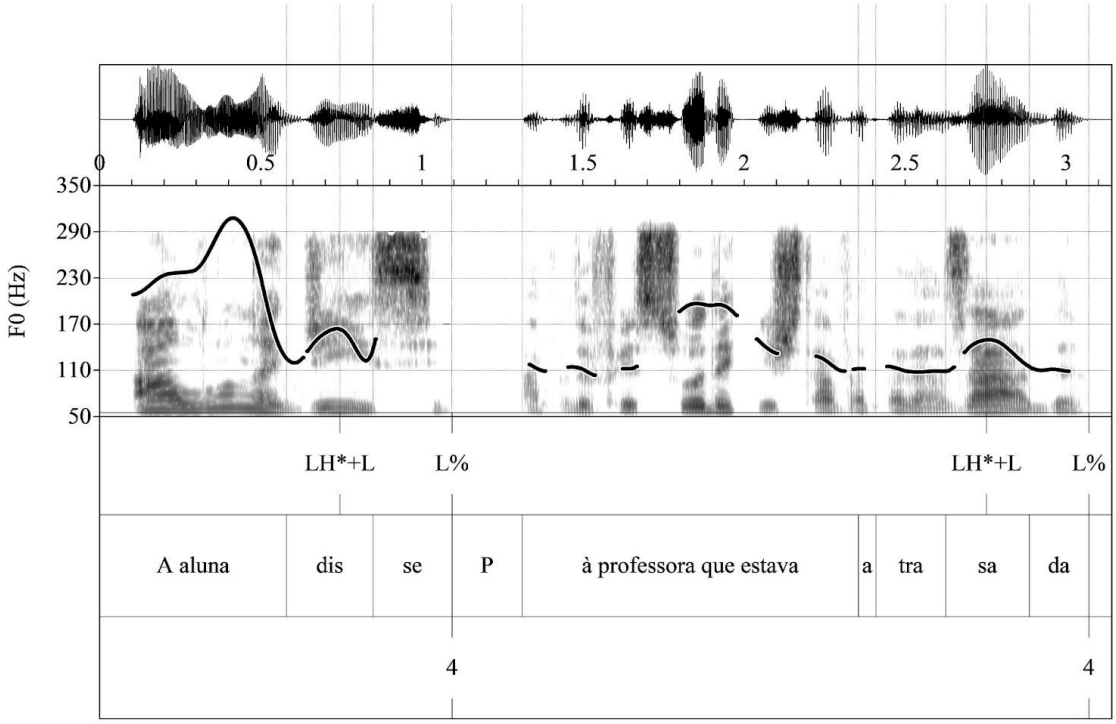
Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

The sentences were recorded by the author of this study, who is a native speaker of BP from Rio de Janeiro, with training in prosody. They were first recorded as complete sentences, i.e. containing the second that-clause that disambiguates them, for example "A aluna disse à professora que estava atrasada que precisaria sair da sala" ('The student told the teacher that was late that she would need to leave the room'). Then, the recordings were edited on the software of acoustic analysis PRAAT Version 6.2.14 (Boersma e Weenink, 2022) so that it only contained the word string that goes from the beginning of the sentence until the end of the ambiguous clause. The intonational transcription followed the tonal inventory proposed by Intonational Phonology (Ladd, 1996), within the Autosegmental-Metrical framework, and the P-ToBI system (Frota *et al*., 2015). All the recordings are publicly available on the link: https://github.com/vitorgabrielish/completionrecordings.git

The conditions favoring the RC analysis presented an IP boundary after the verb of the matrix clause and another one at the end of the that-clause. This prosodic phrasing maps the RC and its antecedent in a single IP and is consistent with the analysis of the ambiguous clause as a restrictive RC. We consider this prosodic phrasing as the optimal grouping to indicate the restrictive RC analysis and the syntactic structure of the sentence as a whole, since it prosodically marks a constituent that undergoes syntactic movement (Frota e Vigário, 2001). The IPs were marked by a /LH*+L L%/ nuclear contour and a 200ms manipulated pause after the first IP. Figures 1 and 2 illustrate the prosodic phrasing and intonational contours of two experimental sentences in the RCS and RCL conditions.

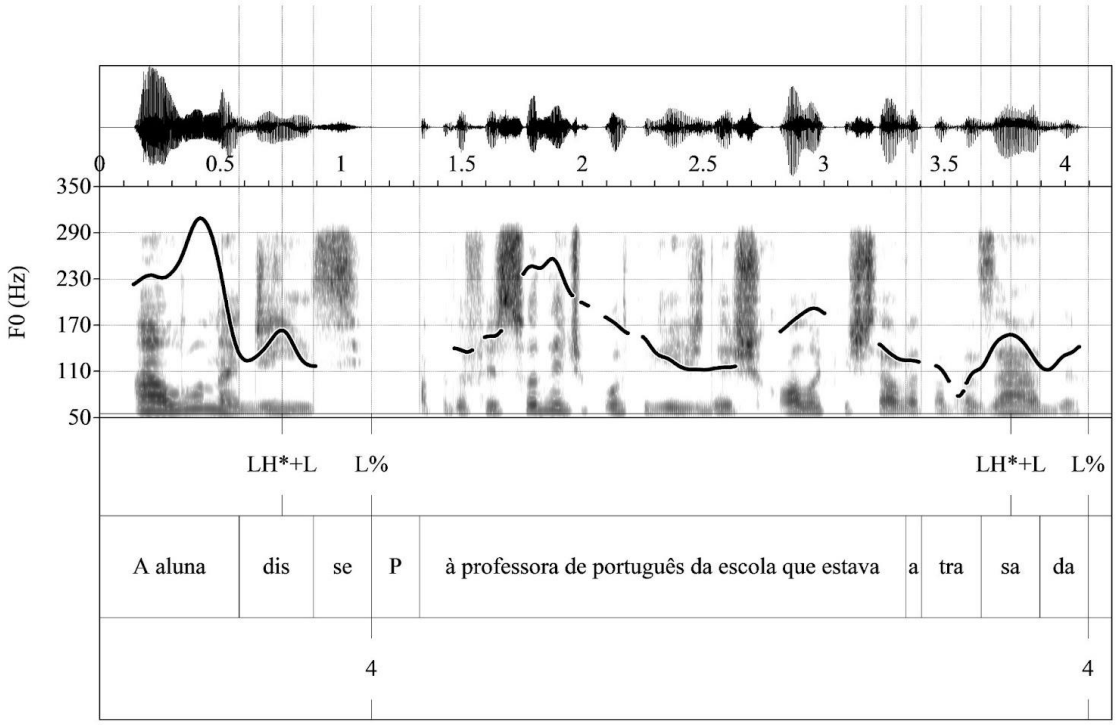
Figure 1 – Example of the prosodic phrasing of the RCS condition

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735



Source: created by the author

Figure 2 – Example of the prosodic phrasing of the RCL condition

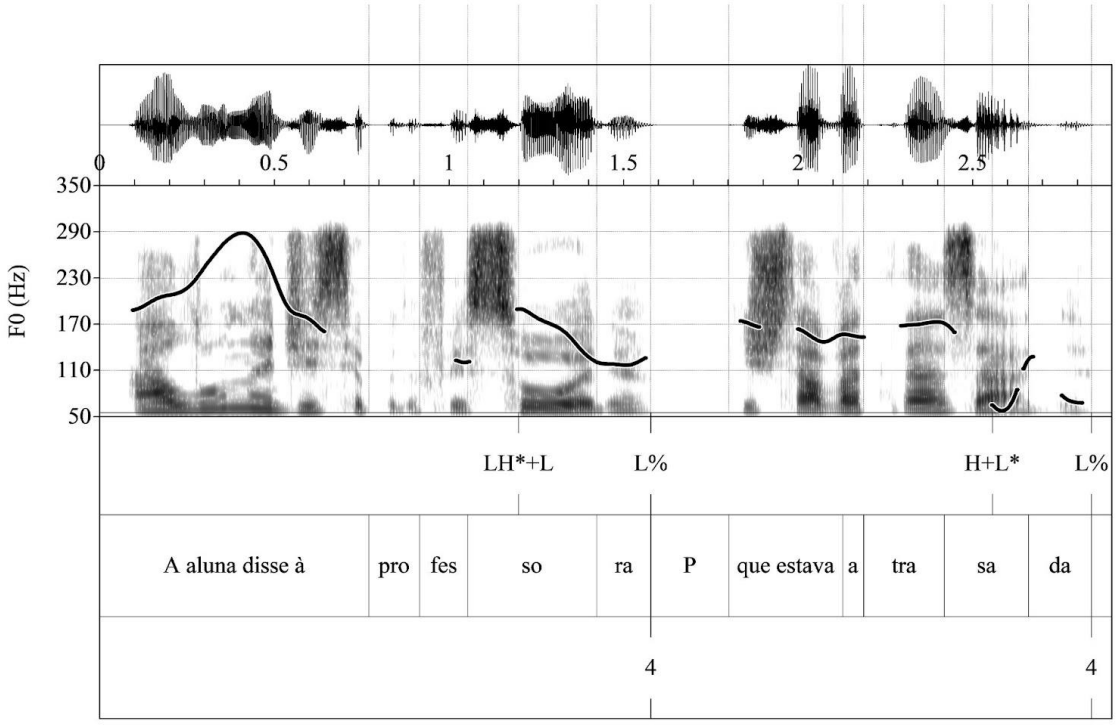


Source: created by the author

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

The conditions favoring the CP analysis presented an IP boundary before the ambiguous clause and another one at the end of it. This prosodic phrasing maps the potential RC and its antecedent into different IPs and, along with the intonational contours marking them, it is consistent with a CP parsing of the ambiguous clause. The first IP was marked by a /LH*+L L%/ nuclear contour and the second IP was marked by a /H+L* L%/ nuclear contour, which is the typical contour associated with the end of an IP in a neutral declarative sentence in BP (Cunha, 2000; Moraes, 2008; Serra, 2009; Silvestre, 2012). We also added a 200ms manipulated pause after the first IP. Figures 3 and 4 illustrate the prosodic phrasing and intonational contours of two experimental sentences in the CPS and CPL conditions.

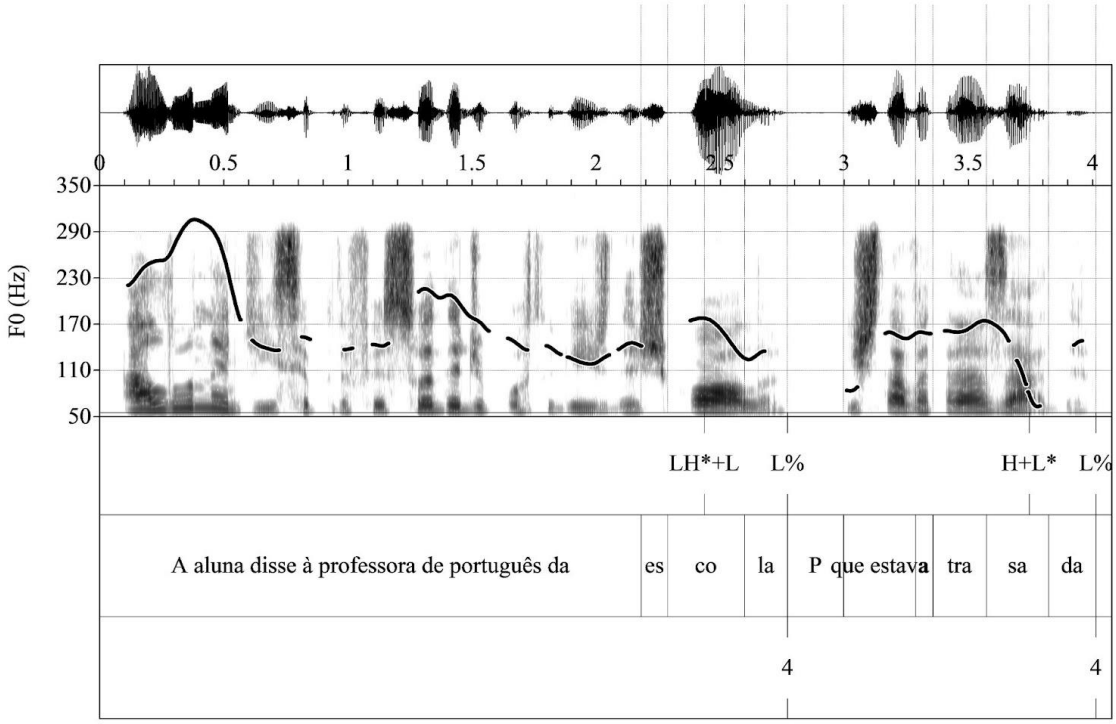
Figure 3 – Example of the prosodic phrasing of the CPS condition



Source: created by the author

Figure 4 – Example of the prosodic phrasing of the CPL condition

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

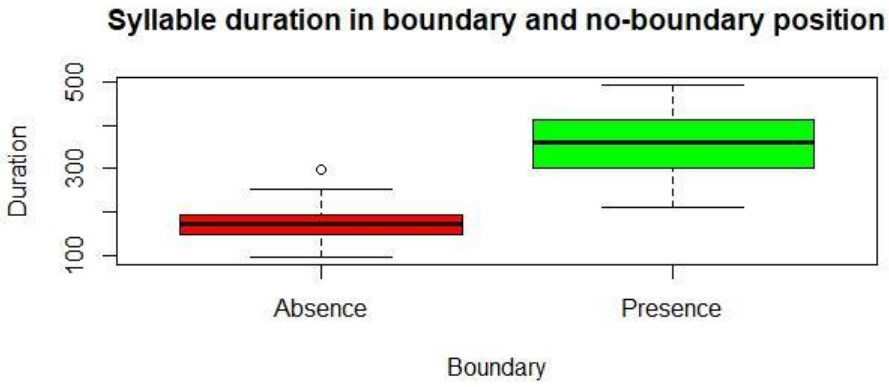


Source: created by the author

In order to verify syllable lengthening, we measured the durations of the stressed syllables of the words in boundary positions and compared it with the durations of the stressed syllables of the same words in no-boundary positions. The mean duration of the syllables in boundary position was 358ms, whereas the mean duration of the syllables in no-boundary position was 173ms. We performed a paired t-test in the software RStudio Version 4.1.1 (R Core Team, 2021) to compare these means and a highly significant difference was found (t = -21.605, df = 79, p-value < .001). Figure 5 shows the durations in the presence *vs.* absence of a boundary.

Figure 5 – Boxplot of the duration of the stressed syllables in no-boundary and boundary positions

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735



Source: created by the author

The boundary tones, the 200-millisecond manipulated pause and syllable lengthening were, therefore, the prosodic cues that marked the IP boundaries. These cues were identified and marked in the recordings by the author of this study and then verified by his doctoral supervisor.

3.2 Predictions

The dependent variables of the experiment were the rates of responses given by participants in each one of the four conditions and the reaction times (RTs) to select an answer.

We expect that, in the RC conditions, participants will choose the alternative that provides them a complement clause to complete the sentence. If subjects are guided by the prosodic phrasing of the fragment they listen to and build an initial RC parsing of the that-clause, they will be expecting an option that allows them to create a grammatical sentence, satisfying the argument structure of the main verb. We also predict that the effect of prosody will have a different weight depending on constituent length (short or long). Although the IP boundaries are placed in the same positions, both in the short and long conditions, the IP mapping the RC will be shorter or longer. This weaker effect could be observed in higher error rates to choose an option.

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

We also predict different results in the CP conditions depending on length. As explained in the hypotheses, we hypothesized that maybe the effect of the prosodic boundary on favoring the parsing of the upcoming clause as a CP would be weaker in the long conditions. In the short conditions, we expect that participants will choose the alternative that does not present a complement clause to complete the sentence. In the long conditions, we also expect subjects to choose options that do not contain a complement clause to complete the sentence. However, considering the weaker effect of the prosodic boundary on signaling syntactic structure in these conditions, we expect a weaker effect of prosody in the long conditions when compared to the short ones. Again, this weaker effect could be observed in higher error rates to choose an option.

3.3 Method

Subjects

40 native speakers of BP (10 in each list of the experiment), being 30 of them women, from 18 to 30 years old (24 on average), participated in the experiment as volunteers. All the subjects were undergraduate students of Letras (Languages and Literature) courses at the Federal University of Rio de Janeiro and were naive about the purposes of the experiment.

Materials

We created 20 sets of experimental sentences inspired by the stimuli used by Silva (2018) (check the Appendix for the full list of experimental sentences). For this reason, the data was not analyzed for acceptability, since it had already been used in a previous study that produced reliable results. The materials were equally distributed in a Latin Square design across four lists of the experiment. 40 filler sentences were interspersed among the experimental items in the four lists.

Procedure

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

The experiment was carried out on the online platform *PennController for Internet Based Experiments* – PCIbex (Zehr e Schwarz, 2018). Participants were first presented to the instructions of the experiment and then two test trials to practice and understand how the experiment worked. In the task, they listened to the auditory stimuli while an image of a speakerphone was kept on their computer screen. Immediately after the audio finished, the image of the speakerphone disappeared and two alternative answers marked with "A)" and "B)" appeared on the screen. They had to read the options and choose the best one to complete each fragment by pressing either the key "A" or "B" on their keyboard. One of the alternatives always presented a CP and the other always had an adverbial phrase initiated by "porque" ('because') or "para" ('to/for'). The other words of the answers were either the same or adjusted for verb tenses. We decided to do this in order to avoid that participants' choices differed because of semantic aspects of the options. The position of the alternatives was counterbalanced so that each one of them appeared half of the time on the left (option A) and the other half on the right (option B). Figure 6 shows the screen with the two alternative answers to complete the experimental sentence "A aluna disse à professora que estava atrasada" ('The student told the teacher that was late').

Figure 6 – Screen with the two alternative answers

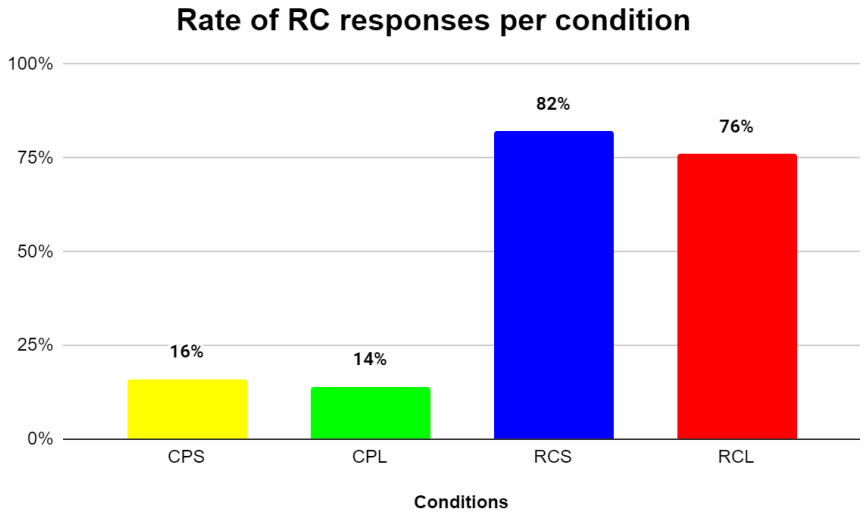
Source: created by the author

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

3.4 Results

The PCIbex platform registered participants' choices to complete the initial fragment heard and the RTs to choose the option. Figure 7 shows the rate of RC responses depending on each one of the four conditions.

Figure 7 – Rate of RC responses per condition



Source: created by the author

The graph clearly shows that the rate of RC responses increased in the RCS and RCL conditions, suggesting that the RC prosody favored the RC parsing of the ambiguous clause, independently of length.

In order to analyze this data, we performed a generalized linear mixed-effects regression model for binomial data (Baayen *et al*., 2008) in RStudio (R Core Team, 2021)[5]. This model allows us to see the correlation effect between the dependent variable (in this case, the rate of responses) and the independent variables (prosody and length). Therefore, the model included participants' answers as a function of the factors Prosody and Length as fixed effects, and participants and items as random effects (Formula: Analysis ~ Prosody * Length + (1|Participant) + (1|Item)). The CP Prosody and Short Length were selected as the reference values (Intercept).

[5] We used the "glmer" function from the "lme4" package of RStudio to run the model.

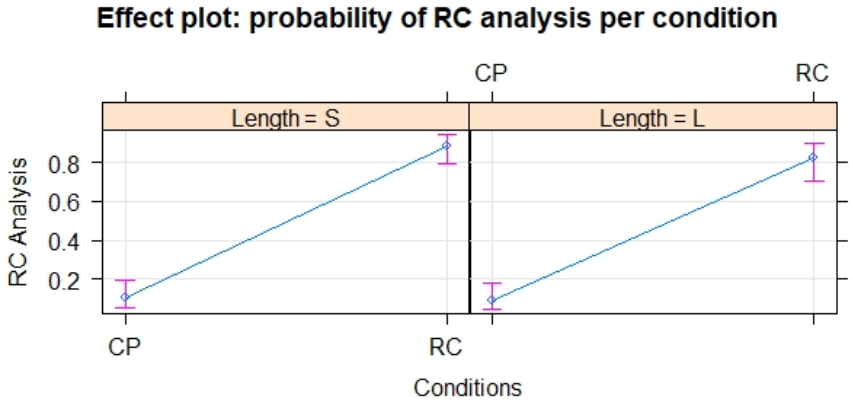
Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

The statistical analysis revealed a highly significant difference between the CP prosody and RC prosody on the rate of responses. The CP prosody favored the CP analysis and the RC prosody favored the RC analysis. No significant difference was found between the Short and Long conditions. No significant interaction was found between the factors Prosody and Length. The summary of the statistical test can be seen in Table 2 and Figure 8 shows the effect plot of the probability of an RC analysis per condition is also shown.

Table 2: Statistical analysis of participants' analysis in the four conditions

Predictors	Odds Ratios	CI	P
Intercept	0.12	0.06 - 0.25	**<0.001**
Prosody [RC]	63.39	22.93 - 175.22	**<0.001**
Length [Long]	0.85	0.33 - 2.18	0.741
Prosody [RC] * Length [Long]	0.69	0.19 - 2.55	0.583

Figure 8 – Effect plot of the probability of RC analysis per condition

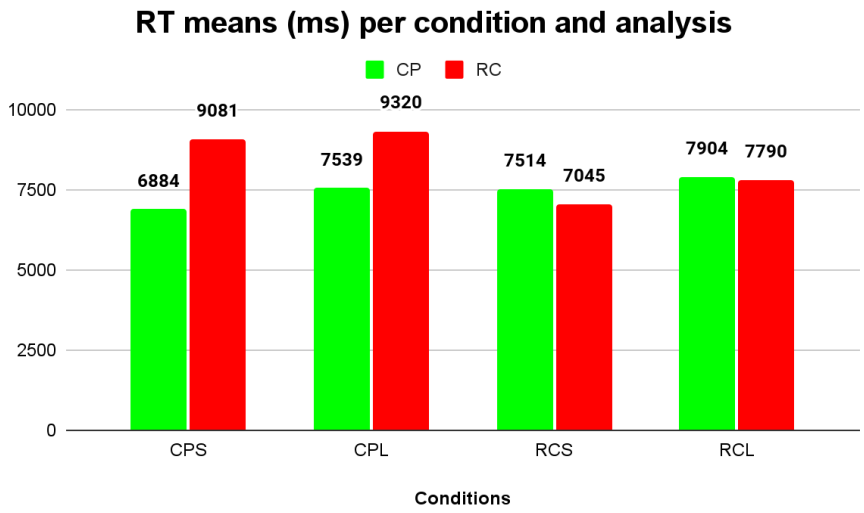


Source: created by the author

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

We analyzed participants' RTs considering participants' analysis (CP or RC) of the ambiguous clause in each one of the four conditions. Our hypothesis was that RTs would be lower when the answer given was consistent with the prosody of the auditory stimulus, whereas the RTs would be higher when the answer given did not match the prosody of the auditory stimulus. In other words, RTs would be lower when choosing a right answer and higher when choosing a wrong answer. These RTs comprise the time they spent to read the options and to choose one of them. The RTs above 13636ms were considered outliers and, therefore, dropped from the analysis[6]. Figure 9 shows this data.

Figure 9 – RT means (ms) per condition and analysis



Source: created by the author

[6] In order to identify the outliers, we used the quartiles method (Formula: Q3 + (Q3 - Q1)*1.5).

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

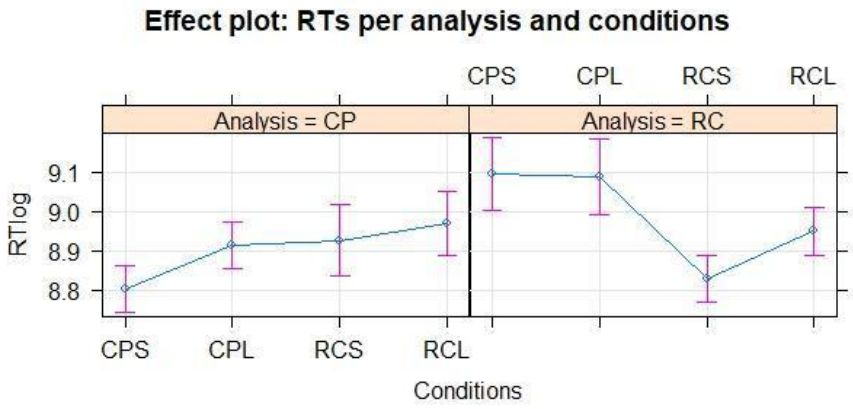
The RT means show that participants spent less time to choose an answer that was consistent with the prosody they were exposed to and took a longer time to choose an answer that contradicted the prosody they were exposed to, independently of length. A linear mixed-effects regression model in RStudio (R Core Team, 2021) was performed to verify these differences. The RTs were log-transformed for the purpose of normality. The model included the log-transformed RTs as a function of participants' analysis and conditions as fixed effects, and participants and items as random effects (Formula: RT_Log ~ Analysis * Conditions + (1|Participant) + (1|Item)). The CP Analysis and the CPS condition were selected as the reference values (Intercept). The statistical analysis revealed that the RTs for the CP analysis were significantly lower than the ones found for the RC analysis, independently of the conditions. A highly significant interaction was also found between the RC analysis and the RCS and RCL conditions, indicating that the RTs of the RC analysis were significantly lower when participants were exposed to an RC prosody, independently of the length. The summary of the statistical test is reproduced in Table 3 and Figure 10 shows the effect plot of the RTs per analysis and conditions.

Table 3 – Statistical analysis of the log-transformed RTs per condition and analysis

Predictors	Estimates	CI	p
Intercept	8.80	8.74 - 8.86	**<0.001**
Analysis [RC]	0.29	0.21 - 0.38	**<0.001**
Analysis [RC] * Condition [RCS]	-0.39	-0.51 - -0.27	**<0.001**
Analysis [RC] * Condition [RCL]	-0.31	-0.42 - -0.20	**<0.001**

Figure 10 – Effect plot of the log-transformed RTs per condition and analysis

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735



Source: created by the author

4. Conclusions

The results provide nice evidence that prosody was able to guide listeners' parsing of the ambiguous clause, independently of length. When prosody favored a CP analysis of the clause, participants mostly chose adverbial phrases to complete the initial fragment they previously listened to. When prosody favored an RC analysis of the clause, participants mostly chose complement phrases to complete the initial fragment they heard.

In the CP prosody, one could say that the CP analysis is consistent with the MA principle (Frazier, 1979) and with an argument-first strategy to parse the ambiguous clause as a CP in order to satisfy the argument structure of the *dicendi* matrix verb. However, the RC prosody clearly favored an RC parsing of the ambiguous clause, contradicting both MA and an argument-first strategy.

Following the Rational Speaker Hypothesis (Clifton *et al*., 2006), we wanted to explore whether long conditions would have a weaker effect on participants' answers and/or RTs in comparison to the short conditions. We did not find a strong effect of length neither on the participants' analysis of the ambiguous clause nor on the RTs. It seems like prosody fulfilled its mission to guide the parser both in the short and long conditions.

By adopting a simple auditory questionnaire, we showed that prosody determined the syntactic structure of the sentence. Also, our data seems to support the view that prosody operates before syntax during the course of processing, supporting Schafer's (1997) and Blodgett's (2004) prosody-first accounts of processing.

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

One limitation of our study is that, unlike Almeida et al. (2021), the prosodic cues were only seen in conjunction. Therefore, we were not able to determine the influence of each cue on the resolution of the ambiguity, which is one of the steps of our future research.

References

ALMEIDA, R.; OLIVEIRA, M.; COZIJN, R. A influência da prosódia da fala na resolução de ambiguidade sintática: um estudo de processamento de sentença. Cad. Est. Ling., Campinas, v.63, p. 1-23, 2021.

BAAYEN, R. H.; DAVIDSON, D. J.; BATES, D. M. Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language, v.59, p. 390-412, 2008.

BLODGETT, A. The interaction of prosodic phrasing, verb bias, and plausibility during spoken sentence comprehension. PhD dissertation, The Ohio State University, Columbus, 2004.

BOERSMA, P.; WEENINK, D. Praat: doing phonetics by computer [Computer program]. Version 6.2.14, 2022. <http://www.praat.org/>

CALDAS, V. G. The role of prosody in the processing of garden-path sentences: online and offline evidence from production and comprehension. Tese (Doutorado em Língua Portuguesa) - Faculdade de Letras, Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2022.

CLIFTON, C.; CARLSON, K.; FRAZIER, L. Tracking the what and why of speakers' choices: Prosodic boundaries and the length of constituents. Psychonomic Bulletin & Review, v.13, n.5, p. 854-861, 2006.

CUETOS, F.; MITCHELL, D. Cross-linguistic differences in parsing: restrictions on the use of the late closure strategy in Spanish. Cognition, v. 30, p. 73-105, 1988.

CUNHA, C. S. Entoação regional no português do Brasil. Tese (Doutorado em Língua Portuguesa) - Faculdade de Letras, Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2000.

FONSECA, A. A prosódia no parsing: evidências experimentais do acesso à informação prosódica no input linguístico. Tese (Doutorado em Linguística) - Faculdade de Letras, Universidade Federal de Minas Gerais. Belo Horizonte, 2012.

FRAZIER, L. On comprehending sentences: Syntactic parsing strategies. PhD dissertation, University of Connecticut. 1979.

FROTA, S.; VIGÁRIO, M. Efeitos de Peso no Português Europeu. In: *Saberes no Tempo. Homenagem a Maria Henriqueta Costa Campos*. Lisboa: Colibri, p. 315-333, 2001.

FROTA, S.; OLIVEIRA, P.; CRUZ, M.; VIGÁRIO, M. P-ToBI: Tools for the transcription of Portuguese prosody. Lisboa: Laboratório de Fonética, CLUL/FLUL, 2015.

HIROSE, Y. Recycling Prosodic Boundaries. Journal of Psycholinguistic Research, v.32, n.2, p.167-195, 2003.

KJELGAARD, M.; SPEER, S. Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. Journal of Memory and Language, n.40, p. 153–194, 1999.

LADD, D. R. Intonational phonology. Cambridge: CUP, 1996.

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

MORAES, J. A. The pitch accents in Brazilian Portuguese: analysis by synthesis. Proceedings of Speech Prosody. Campinas: Brasil, p. 389-398, 2008.

R CORE TEAM. R: The R Foundation for Statistical Computing Platform, 2021. <http://www.r-project.org/index.html>.

SCHAFER, A. Prosodic Parsing: The role of prosody in sentence comprehension. PhD dissertation, University of Massachusetts, Amherst, MA: GLSA. 1997.

SERRA, C. R. Realização e percepção de fronteiras prosódicas no Português do Brasil: Fala espontânea e Leitura. Tese (Doutorado em Língua Portuguesa) - Faculdade de Letras, Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2009.

SILVA, E. Rastreamento ocular de orações temporariamente ambíguas: reflexões sobre os efeitos garden-path e good-enough na leitura de alunos de curso fundamental e superior. Dissertação (Mestrado em Linguística) - Faculdade de Letras, Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2018.

SILVESTRE, A. P. S. A Entoação Regional dos Enunciados Assertivos nos Falares das Capitais Brasileiras. Dissertação (Mestrado em Língua Portuguesa) - Faculdade de Letras, Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2012.

ZEHR, J.; SCHWARZ, F. PennController for Internet Based Experiments (IBEX), 2018. <https://doi.org/10.17605/OSF.IO/MD832>.

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

Appendix:

Short 1: A aluna disse à professora que estava atrasada.
Long 1: A aluna disse à professora de português da escola que estava atrasada.
Options: A) que perdeu o ônibus. B) porque perdeu o ônibus.

Short 2: A enfermeira assegurou ao médico que estava no quarto.
Long 2: A enfermeira assegurou ao médico do hospital público que estava no quarto.
Options: A) que cuidava do doente. B) para cuidar do doente.

Short 3: A médica confirmou à paciente que estava chegando.
Long 3: A médica confirmou à paciente da clínica de estética que estava chegando.
Options: A) que cancelou a consulta. B) para fazer a consulta.

Short 4: A secretária contou ao faxineiro que estava doente.
Long 4: A secretária contou ao faxineiro da companhia de limpeza que estava doente.
Options: A) que não iria trabalhar. B) para ir ao trabalho.

Short 5: A vendedora falou para a cliente que estava chorando.
Long 5: A vendedora falou para a cliente da loja de roupas que estava chorando.
Options: A) que tinha se machucado. B) porque tinha se machucado.

Short 6: O advogado alertou ao juiz que estava confuso.
Long 6: O advogado alertou ao juiz do tribunal federal que estava confuso.
Options: A) que o fórum está cheio. B) porque o fórum está cheio.

Short 7: O bombeiro respondeu ao morador que estava no prédio.
Long 7: O bombeiro respondeu ao morador do primeiro andar que estava no prédio.
Options: A) que apagaria o incêndio. B) para apagar o incêndio.

Short 8: O chefe garantiu ao empregado que estava na sala.
Long 8: O chefe garantiu ao empregado da empresa de tecnologia que estava na sala.
Options: A) que poderiam conversar. B) para poderem conversar.

Short 9: O crítico alegou ao cineasta que estava revoltado.
Long 9: O crítico alegou ao cineasta arrogante de bigode que estava revoltado.
Options: A) que seus atores são ruins. B) porque seus atores são ruins.

Short 10: O deputado comunicou ao prefeito que estava no congresso.
Long 10: O deputado comunicou ao prefeito da cidade do Rio que estava no congresso.
Options: A) que começaria a votação. B) para começar a votação.

Short 11: O devoto declarou ao padre que estava na missa.
Long 11: O devoto declarou ao padre da igreja da cidade que estava na missa.
Options: A) para buscar o perdão. B) que buscaria o perdão.

Veredas – Revista de Estudos Linguísticos | E-ISSN: 1982-2243 | v.27, n.1, 2023 e40735

Short 12: O diretor avisou ao ator que estava fotografando.
Long 12: O diretor avisou ao ator da peça de teatro que estava fotografando.
Options: A) porque a cena está ótima. B) que a cena está ótima.

Short 13: O gerente explicou ao cliente que estava preocupado.
Long 13: O gerente explicou ao cliente da suíte de luxo que estava preocupado.
Options: A) porque a água acabou. B) que a água acabou.

Short 14: O leitor relatou para o repórter que estava descansando.
Long 14: O leitor relatou para o repórter do jornal da cidade que estava descansando.
Options: A) para voltar ao trabalho. B) que voltou ao trabalho.

Short 15: O juiz negou ao deputado que estava com raiva.
Long 15: O juiz negou ao deputado do estado do Paraná que estava com raiva.
Options: A) porque o réu gritou. B) que o réu gritou.

Short 16: O músico afirmou ao pintor que estava estudando.
Long 16: O músico afirmou ao pintor da galeria de arte que estava estudando.
Options: A) para reger a orquestra. B) que regeria a orquestra.

Short 17: O passageiro advertiu ao motorista que estava com pressa.
Long 17: O passageiro advertiu ao motorista do carro de luxo que estava com pressa.
Options: A) porque queria chegar ao trabalho. B) que queria chegar ao trabalho.

Short 18: O patrão reclamou com o pedreiro que estava cansado.
Long 18: O patrão reclamou com o pedreiro de chapéu de palha que estava cansado.
Options: A) porque a obra não terminava. B) que a obra não terminava.

Short 19: O repórter informou ao morador que estava perdido.
Long 19: O repórter informou ao morador do bairro do centro que estava perdido.
Options: A) porque não conhecia o bairro. B) que não conhecia o bairro.

Short 20: O vizinho comentou com o porteiro que estava nervoso.
Long 20: O vizinho comentou com o porteiro do turno da noite que estava nervoso.
Options: A) porque o ladrão fugiu. B) que o ladrão fugiu.