RESEARCH REPORT

Prosody and sentence processing in Brazilian Portuguese: a visual world paradigm study

Vitor Gabriel CALDAS [iD]
Universidade Federal do Rio de Janeiro (UFRJ)

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 alunai disse à professoraj 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 different prosodic phrasing patterns will be aligned with different syntactic analyses, guiding listeners to parse the ambiguous clause either as a CP or as an RC. A Visual World Paradigm study, using an eye-tracking equipment, was carried out to assess participants' processing and comprehension of the ambiguity. Thirty-two (32) undergraduate students listened to sentences, with either a Garden-Path (GP) prosody or a No Garden-Path (NGP) prosody, and visualized two pictures representing the CP and the RC analyses of the ambiguous clause on a computer screen while an eye-tracker monitored their eye movements. The results showed that, when participants listened to a GP prosody, there were more fixations at the CP pictures. When they listened to an NGP prosody, there were more fixations at the RC pictures. Also, there was a higher percentage of errors to a comprehension question when participants were exposed to the GP prosody conditions, indicating a higher level of difficulty.

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 alunai disse à professoraj 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). Postulamos a hipótese de que diferentes padrões de fraseamento prosódico estarão alinhados a diferentes estruturas sintáticas, guiando os ouvintes a analisarem a oração ambígua como OC ou OR. Um estudo de Paradigma do Mundo Visual, utilizando um equipamento de rastreamento ocular, foi realizado para avaliar o processamento e a compreensão da ambiguidade pelos participantes. Trinta e dois (32) estudantes de graduação ouviram sentenças, com prosódia Garden-Path (GP) ou Sem Garden-Path (SGP), e visualizaram duas imagens representando as análises OC e OR da oração ambígua em uma tela de computador enquanto um rastreador ocular monitorava os seus movimentos oculares. Os resultados mostraram que, quando os participantes ouviam uma prosódia de GP havia mais fixações nas imagens de OC. Quando eles ouviam uma prosódia SGP, havia mais fixações nas imagens OR. Além disso, houve maior percentual de erros nas perguntas de compreensão quando os participantes haviam sido expostos às condições de prosódia GP, indicando maior nível de dificuldade.

RESUMO PARA NÃO ESPECIALISTAS
Este artigo estuda como a entoação pode influenciar a interpretação de frases ambíguas no português brasileiro. A estrutura investigada apresenta duas orações iniciadas por "que" ("A aluna disse à professora que estava atrasada que precisaria sair da sala"). Ao lermos uma frase como essa, primeiramente, interpretamos que (1) "A aluna estava atrasada". No entanto, ao chegarmos à segunda oração com "que", entendemos que, na verdade, (2) "A professora estava atrasada". Postulamos a hipótese de que, a depender da entoação, somos levados a uma interpretação ou a outra. Desenvolvemos um experimento em que 32 estudantes de graduação ouviram

sentenças, com a entoação da interpretação (1) ou da interpretação (2), e, simultaneamente, visualizaram duas imagens representando as duas interpretações da oração ambígua em uma tela de computador enquanto um rastreador ocular monitorava os seus movimentos oculares. Os resultados mostraram que, quando os participantes ouviam a entoação (1), havia mais fixações nas imagens da interpretação (1). Quando eles ouviam a entoação (2), havia mais fixações nas imagens da interpretação (2). Além disso, houve maior percentual de erros nas perguntas de compreensão quando os participantes haviam sido expostos às condições da entoação (1), indicando maior nível de dificuldade.

Introduction

This paper presents a Visual World Paradigm study, using an eye-tracking equipment, that aimed at investigating the role of prosody in sentence processing testing garden-path sentences in Brazilian Portuguese (BP)[1]. The 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 example (1) below.

(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.

Although these syntactic properties explain the ambiguity at issue, we believe that prosody can have an influence on the parsing of the ambiguous clause. Our main hypotheses are: (i) the assignment of a prosodic phrasing consistent with the syntactic structure blocks the garden-path effect and resolves the ambiguity of the sentence and (ii) prosody can also provoke the garden-path effect, guiding

_____

the parser to analyze the ambiguous clause as a CP. See the prosodic phrasings below for hypotheses (i) and (ii), in which the % symbols represent intonational phrase (henceforth IP) boundaries.

(i)   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'

(ii)   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'

In (i) the noun "professora" ('the teacher') is phrased together with the RC "que estava atrasada" ('that was late'), therefore, the language processing system should not attach this that-clause to the previous matrix verb "disse" ('told') and parse it as a CP because they are in different IPs. Whereas in (ii) the opposite should happen, since the ambiguous clause is phrased in a single IP, the CP attachment is expected.

In order to investigate these hypotheses, we carried out a Visual World Paradigm study, using an eye-tracking equipment, in which participants listened to sentences with either a Garden-Path prosody (GP) or a No Garden-Path prosody (NGP) while visualizing two pictures representing the two analyses (CP *vs.* RC) of the ambiguous clause on a computer screen.

This paper is structured as it follows: (1) first, we briefly summarize some studies on the prosody-syntax interface in sentence processing; (2) then, we present the details of the experiment, reporting its experimental design, predictions, method, and results of both the online measure and the offline measures; and (3) finally, we draw some conclusions and reflect upon future research.

1. The prosody-syntax interface in sentence processing

The interest in the role of prosody in sentence processing is not restricted to psycholinguists. It is also important for phonologists who want to better understand the constitution of prosodic domains in the prosody-syntax interface. It is no wonder that one of the first studies to deal with the influence of prosody in ambiguous sentences was done by Nespor & Vogel (1986), phonologists who proposed the theory of Prosodic Hierarchy.

Nespor & Vogel (1986) propose that the Prosodic Hierarchy is composed by seven domains, from the lowest to the highest: syllable (σ), foot (Σ), phonological word (ω), clitic group (C), phonological phrase (φ), intonational phrase (IP) and phonological utterance (U). Each one of these constituents is defined based not only on phonological information but also morphological, syntactic and semantic. For example, the IP is mapped through the prosody-syntax interface, being influenced by phonological factors (e.g., rhythm and phrase length) and also by the syntactic structure of sentences.

Considering that the formation of IPs is dependent on syntax, Nespor & Vogel (1986) investigated how syntactically ambiguous structures are aligned with different prosodic structures. For example,

the Italian sentence "Marco ha guardato la ragazza col canocchiale" ('Marco looked at the girl with the binoculars' - NESPOR & VOGEL, 1986, p. 261) can mean (i) "The girl is holding the binoculars" or (ii) "Marco is holding the binoculars". If the intended meaning is (i), the prosodic phrasing should be ("Marco% ha guardato la ragazza col canocchiale"); if the intended meaning is (ii), an IP boundary should be added after "ragazza" ("Marco% ha guardato la ragazza% col canocchiale"). It is worth noting that, for some hearers, this sentence may remain ambiguous. What these two prosodic structures do is to help bias the possible interpretations.

In the late 1990s', Speer et al. (1996) and Kjelgaard & Speer (1999) studied early/late closure ambiguities and showed that when prosodic and syntactic structures matched, sentence processing was facilitated. See example (2) below.

(2) When Madonna sings the song is a hit.

To correctly parse this sentence, an IP boundary should be placed after the verb "sings". In this case, the prosodic structure is aligned with the syntactic structure. However, if we follow the late closure principle, parsing the DP "the song" as an object of the verb "sings", an IP boundary should be placed after this DP. This prosodic phrasing creates a mismatch between prosody and syntax, guiding listeners to a misleading syntactic analysis and showing an interference effect of prosody on sentence processing.

Clifton et al. (2006) proposed the Rational Speaker Hypothesis (RSH). According to the RSH, (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 want to investigate whether the presence of an IP boundary becomes less informative to the listener when there is more than one reason for its existence.

One of the structures studied by the authors was previously investigated by Carlson et al. (2001) and Clifton et al. (2002). The sentences contained either a short or a long adverb phrase that could modify the matrix verb (high attachment) or the complement verb (low attachment) of the clauses. See examples (3) and (4) below.

(3) Susie learned that Bill telephoned last night.
(4) Susie learned that Bill telephoned late last night after the general meeting.

The authors showed that an IP boundary before the short adverb phrase in (3) resulted in the highest occurrence of high attachments, since the purpose of the boundary in this case is purely syntactic. This result is explained by the RSH, since an IP boundary before the long adverb phrase in (4) can either signal the syntactic structure of the sentence or occur as a matter of phonological weight.

More recently, Fonseca & Silva (2022) carried out a Visual World Paradigm study, using an eye-tracking equipment, to investigate the influence of pitch accents and IP boundaries in the resolution of ambiguous adverb phrases. The examples below show a material set of the conditions used in the experiment.

(5) **Pitch accent on the first verb**

O tio do Marcos ESCUTOU que a Priscila cantou no banheiro da suíte.

'Marcos's uncle HEARD that Priscila sang in the suite bathroom.'

(6) **Pitch accent on the second verb**

O tio do Marcos escutou que a Priscila CANTOU no banheiro da suíte.

'Marcos's uncle heard that Priscila SANG in the suite bathroom.'

(7) **Pitch accent on the first verb + IP boundary**

O tio do Marcos ESCUTOU que a Priscila cantou% no banheiro da suíte.

'Marcos's uncle HEARD that Priscila sang% in the suite bathroom.'

(8) **Pitch accent on the second verb + IP boundary**

O tio do Marcos escutou que a Priscila CANTOU% no banheiro da suíte.

'Marcos's uncle heard that Priscila SANG% in the suite bathroom.'

In the task, participants listened to one of the experimental conditions while visualizing two pictures representing the two possible interpretations of the adverb phrase on a computer screen. After listening to the stimuli, subjects answered a comprehension question to verify their final processing of the sentences.

The results showed that when participants listened to the pitch accent on the first verb (conditions 5 and 7), there were more eye fixations to the picture that represented the attachment of the adverb phrase to the first verb. Whereas, when participants listened to the pitch accent on the second verb (conditions 6 and 8), there were more eye fixations to the picture that represented the attachment of the adverb phrase to the second verb. The answers to the comprehension questions also confirmed the results of the eye-tracking measures, showing the influence of prosody on sentence processing.

2. Experiment: Visual World Paradigm

The experiment reported in this paper consists of an eye-tracking study using the Visual World Paradigm (VWP). Eye-tracking experiments provide data in high temporal resolution on online language processing while participants perform relatively natural tasks (BERENDS ET AL., 2015). By analyzing people's eye movements, psycholinguists are able to better understand the cognitive processes going on inside people's minds moment-by-moment when they are either reading or listening to a sentence.

The main hypothesis of this experiment is that prosody will guide listeners' parsing of the sentence, favoring one syntactic analysis over the other. When prosody favors an RC parsing of the

ambiguous clause, no processing difficulty should happen. However, when prosody favors a CP parsing of the first that-clause, a garden-path effect should occur when the parser encounters the second that-clause. We take example (1) again to make it clearer.

(9)　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'

If the prosodic phrasing of this sentence is consistent with an RC parsing of the first that-clause, it should preclude processing difficulty, since the prosodic structure is aligned with the syntactic structure. If the prosodic phrasing of this sentence guides listeners to parse the ambiguous clause as a CP, it provokes a garden-path effect when the parser encounters the second that-clause, since this upcoming syntactic structure does not fit in the representation built by the parser. In other words, the parser had already satisfied the argument structure of the matrix verb, but the first that-clause was misparsed as a CP and now the parser needs to reanalyze it as an RC.

We also hypothesize that length may have an effect both when prosody facilitates processing (RC parsing) and when it provokes a garden-path and reanalysis should occur (CP parsing). See example (10) below.

(10)　A aluna$_i$ disse à professora$_j$ de português da escola que $_{–j}$ estava atrasada

que $_{–i}$ precisaria sair da sala

'The student told the Portuguese teacher of the school that was late

that she would need to leave the room'

The prosodic phrasing favoring the RC parsing will map a long constituent [à professora de português da escola que estava atrasada]IP ('the Portuguese teacher of the school that was late') into a single IP. Comparing to an IP that maps a short constituent, the effect of prosody on signaling syntactic structure may be weaker, since listeners' prosodic competence knows that an IP boundary should be placed before the that-clause and this may provoke processing difficulty and affect the correct parsing of the clause as an RC.

On the other hand, we hypothesize that the prosodic phrasing favoring the CP parsing in a long constituent (an IP boundary occurring before the first that-clause "... escola % que estava atrasada" - '... school % that was late') could facilitate the reanalysis of the ambiguous clause as an RC. We justify this idea based on the Rational Speaker Hypothesis (CLIFTON ET AL., 2006). As explained in the previous section, the RSH states 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, a listener who is well aware of the reasons for the existence of IP boundaries could reinterpret the placement of the IP boundary before the clause as a consequence of length.

2. 1 Experimental design

The experiment presented a 2x2 within-subjects factorial design. The independent variables of this experiment were (1) the prosodic phrasing of the garden-path sentences and (2) the length of the constituents before the ambiguous clause. Variable (1) was divided into two levels: (i) the prosodic phrasing that favors a CP parsing of the ambiguous clause (garden-path prosody - GP) and (ii) the prosodic phrasing that favors an RC parsing of the ambiguous clause (no garden-path prosody - NGP). Variable (2) was divided into two other levels: (i) Short (S) and (ii) Long (L). The short condition is composed by a single noun subject + a matrix *dicendi* verb + a single noun dative + 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. An experimental set with the conditions is illustrated below.

(9)   **GP-Short (GPS)**

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'

(10)   **No GP-Short (NGPS)**

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'

(11)   **GP-Long (GPL)**

A aluna disse à professora de português da escola% **que estava atrasada**% que precisaria sair da sala

'The student told the Portuguese teacher of the school% **that was late**% that she would need to leave the room'

(12)   **No GP-Long (NGPL)**

A aluna disse% **à professora de português da escola que estava atrasada**% que precisaria sair da sala

'The student told% **the Portuguese teacher of the school that was late**% that she would need to leave the room'

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. 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/VWPrecordings.git

The GP conditions 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). The final IP was marked by another /H+L* L%/ nuclear contour. We also added 200-millisecond manipulated pauses after the first and second IPs. Figures 1 and 2 illustrate the prosodic phrasing and intonational contours of two experimental sentences in the GPS and GPL conditions.
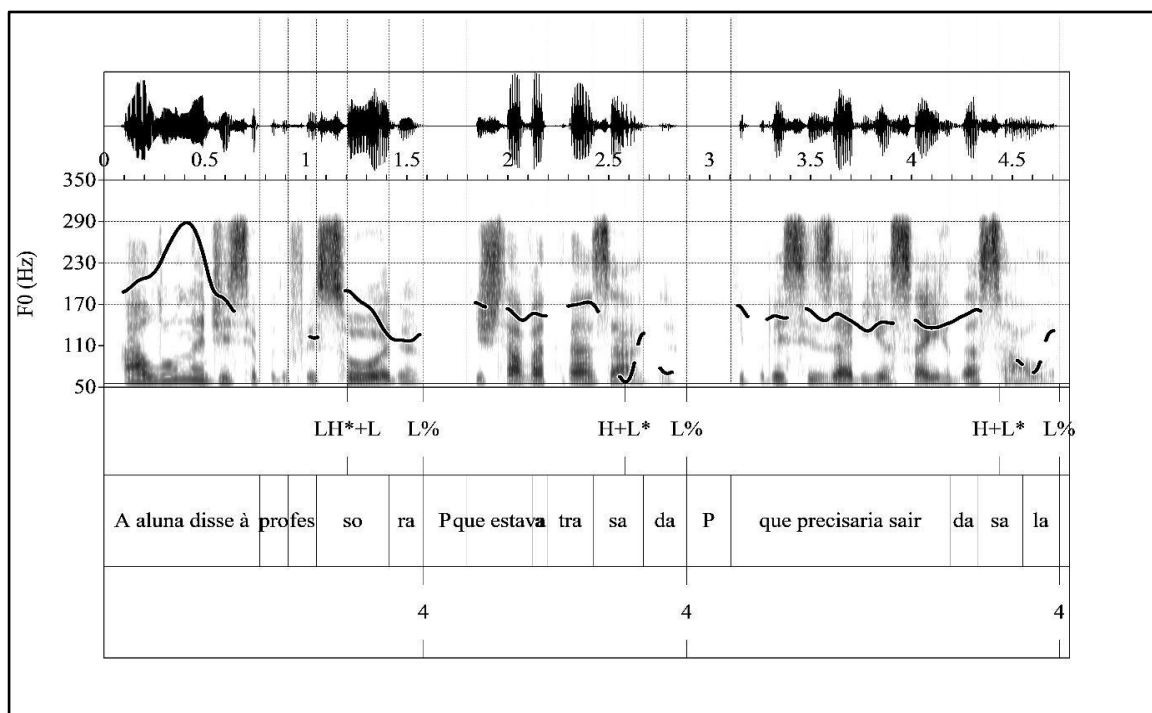

FIGURE 1 - Example of the prosodic phrasing of the GPS condition.
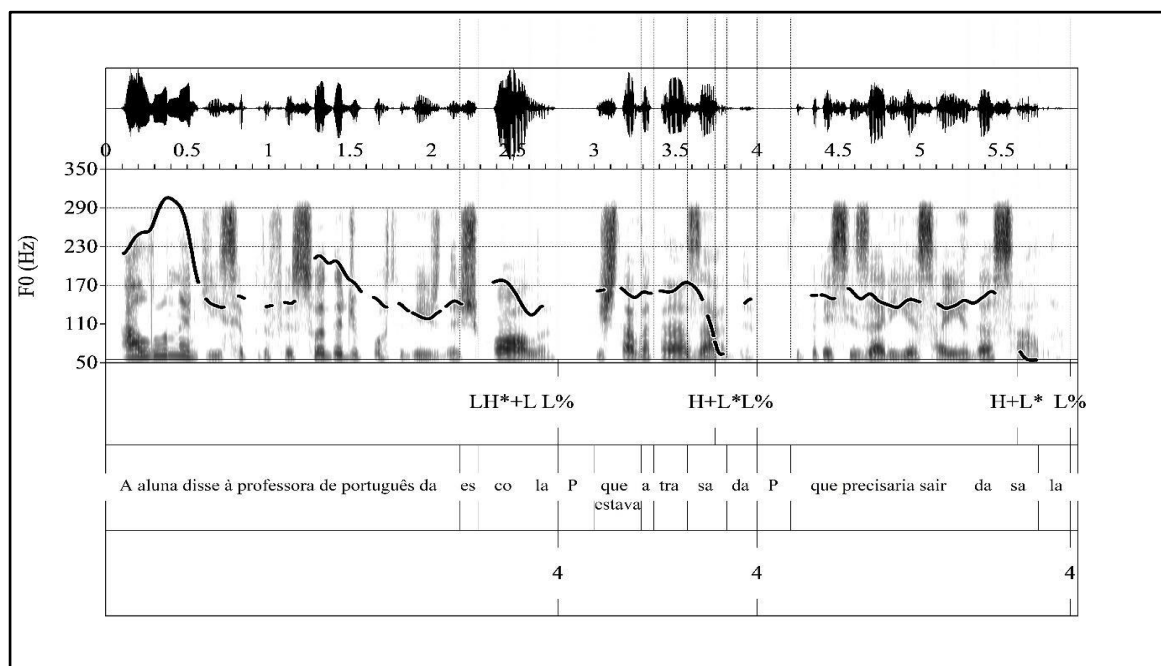Source: elaborated by the author

FIGURE 2 - Example of the prosodic phrasing of the GPL condition.
Source: elaborated by the author

The NGP conditions presented an IP boundary after the verb of the matrix clause and another one at the end of the first 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. The non-final IPs were marked by a /LH*+L L%/ nuclear contour and we also added 200-millisecond manipulated pauses after the first and second IPs. Figures 3 and 4 illustrate the prosodic phrasing and intonational contours of two experimental sentences in the NGPS and NGPL conditions.
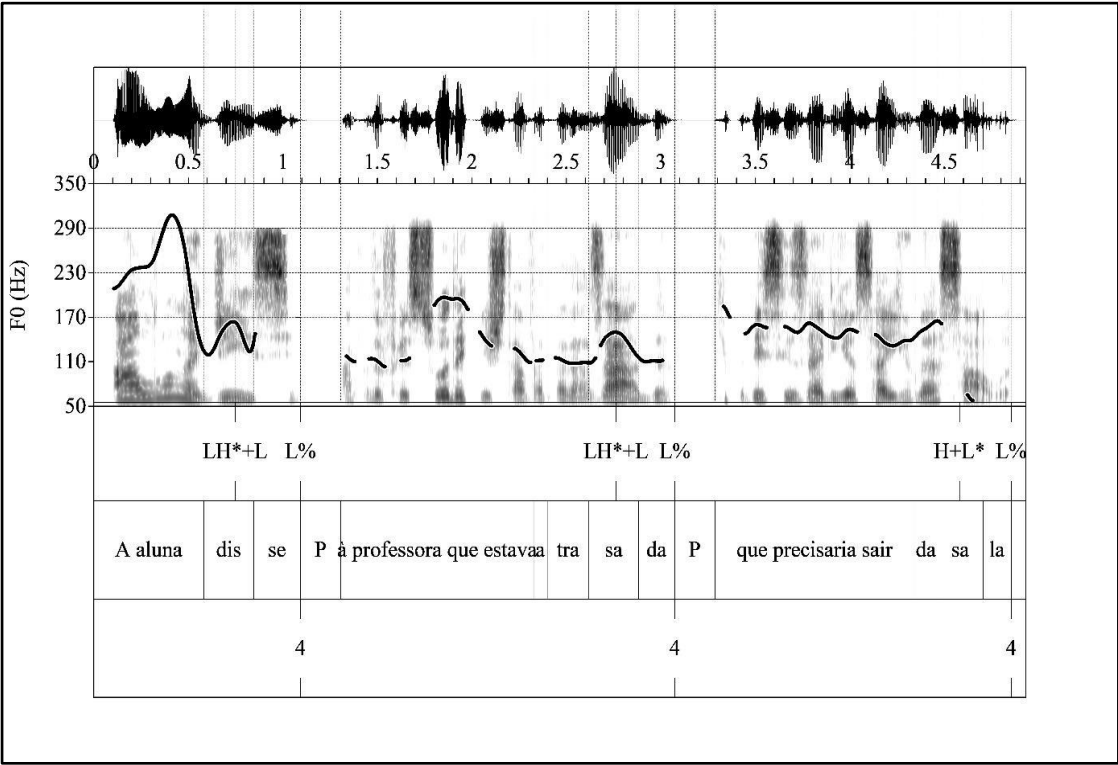
FIGURE 3 - Example of the prosodic phrasing of the NGPS condition.
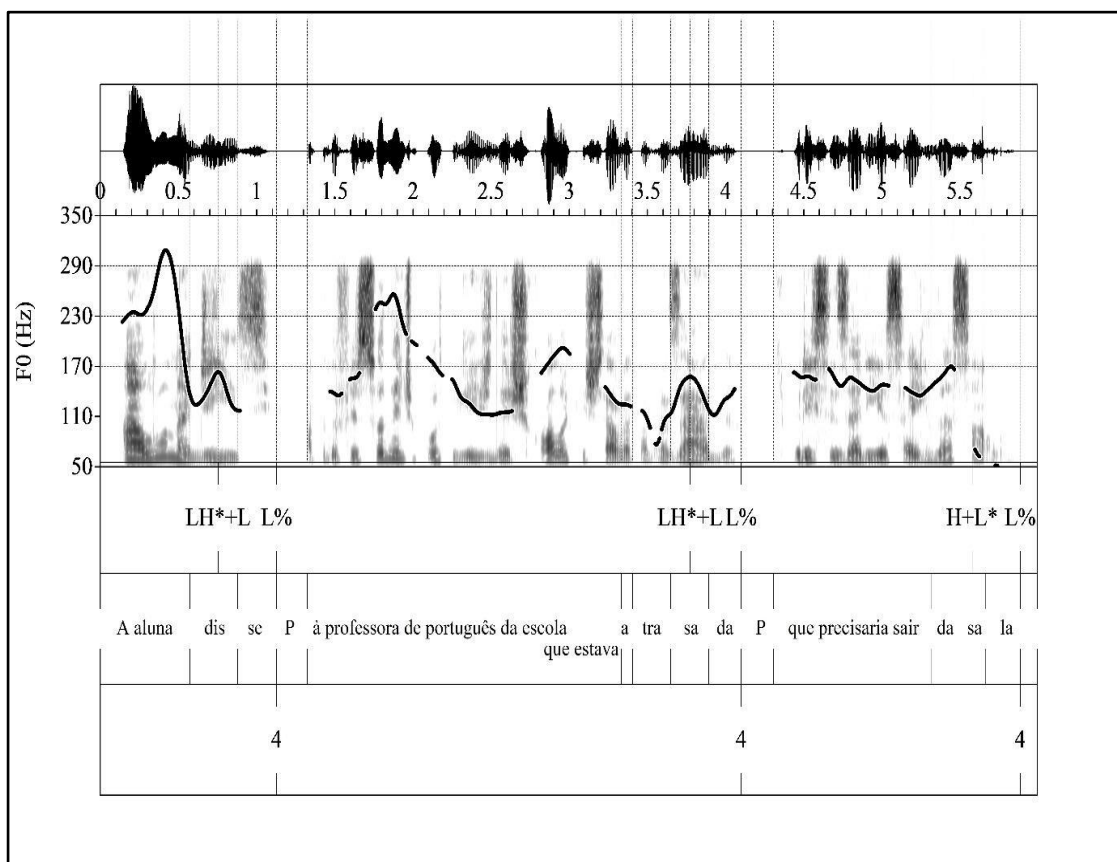Source: elaborated by the author

FIGURE 4 - Example of the prosodic phrasing of the NGPL condition.
Source: elaborated by the author

2.2 Predictions

We have different predictions for the online and offline measures. For the online measures, we expect that in the GP conditions there will be more fixations to the picture that represents a CP parsing of the first that-clause when participants are listening to this portion of the sentence. We also expect an oscillation pattern between the pictures when participants start listening to the second that-clause, which would indicate the garden-path effect. In the NGP conditions, we predict more fixations to the picture that represents an RC parsing of the first that-clause when participants start listening to this portion of the sentence and fix on this picture until the end of the sentence.

Our predictions based on length are related to the offline measures. For the GPL conditions, we expect that the reanalysis of the ambiguous clause as an RC may be simpler compared to the GPS conditions since the reinterpretation of the IP boundary as reflecting length is easier when the IP boundary occurs after a long constituent. This effect of length on reanalysis could be observed in lower error rates in the GPL conditions compared to the GPS conditions. For the NGPL conditions we have an opposite prediction. We expect that this condition will be more difficult not only because it contains more material to be processed, but also because the RC is mapped into a long IP and listeners' prosodic

competence knows that an IP boundary should be placed before the that-clause. This effect of length could be observed in lower error rates in the NGPS conditions compared to the NGPL conditions.

2.3 Method

<u>Subjects</u>

Thirty-two (32) native speakers of BP (8 in each version of the experiment), 24 years old on average, participated in the experiment as volunteers[2]. 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. They all had normal or corrected-to-normal vision.

<u>Materials</u>

We used twenty (20) quartets of experimental sentences like the ones presented above. The materials were equally distributed in a Latin Square design across four lists of the experiment. Forty (40) filler sentences were interspersed among the experimental items in the four lists.

As for the pictures used, each one of them represented a CP or an RC analysis of the ambiguous clause. Figure 5 depicts the CP parsing of the ambiguous clause in the sentence "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 earlier') and Figure 6 depicts the RC parsing for the same sentence.
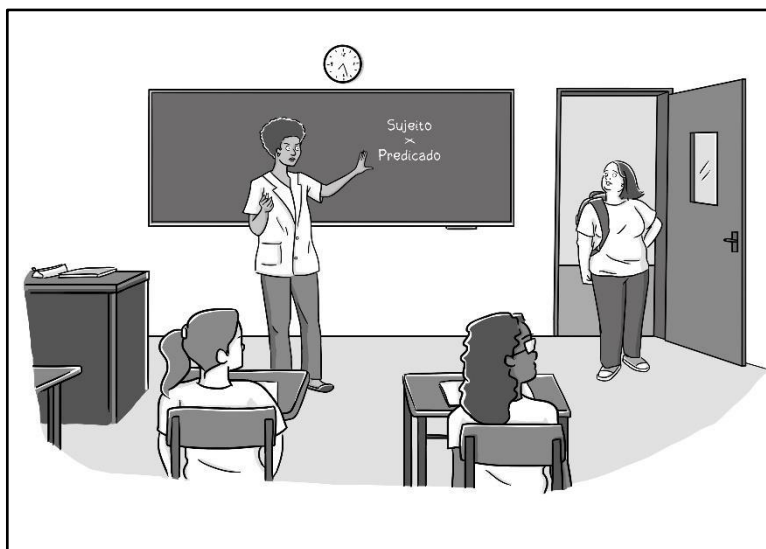

FIGURE 5 - Example of picture representing the CP analysis of the ambiguous clause.
Source: elaborated by the author

_____

[2] All the participants signed a declaration of informed consent before engaging in any research activities.
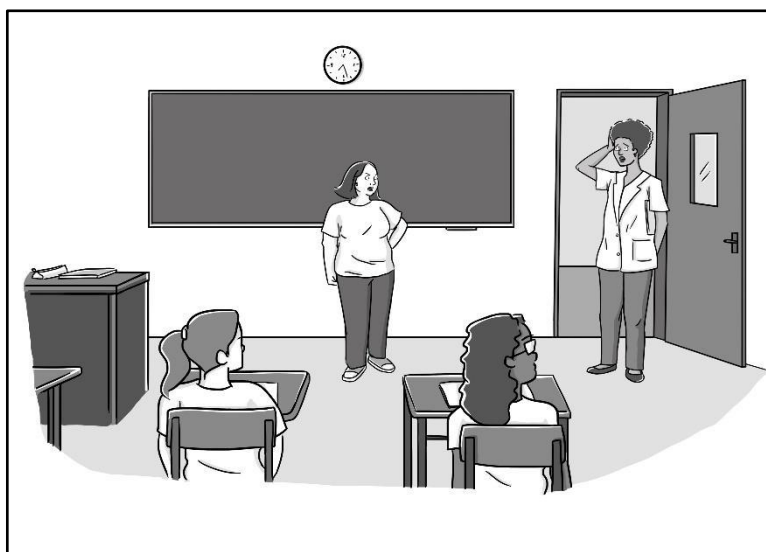
FIGURE 6 - Example of picture representing the RC analysis of the ambiguous clause.
Source: elaborated by the author

Procedures

The experiment was carried out at the *Laboratório de Psicolinguística Experimental da Universidade Federal do Rio de Janeiro* (LAPEX-UFRJ - 'Laboratory of Experimental Psycholinguistics of the Federal University of Rio de Janeiro'), using an SR Research Eyelink 1000 eye-tracker, which has a high precision camera with an accuracy of 1000 hertz. The eye-tracker was set up for a monocular recording, integrated to a 32" inches screen, with a resolution of 1920x1080 pixels. Subjects' head was stabilized using the tower mount of the eye-tracker and their eyes were calibrated before beginning the experiment with a five-point calibration system.

Participants were first presented with four test trials to practice and understand how the experiment worked. The first screen of the trials displayed the two pictures representing the two analyses of the ambiguous clause on the computer screen. Subjects were instructed to visualize these images only in order to familiarize themselves with them. After four seconds, the pictures disappeared, a fixation cross appeared at the center of the screen and participants had to stare at it. After one second, the cross disappeared and the same pictures presented before came back to the screen. At the same time the images appeared on the screen, participants listened to a sentence on headphones. They were instructed to look at the picture that best represented what they were listening to. When the audio finished, they had to choose the picture that best represented the sentence heard by pressing a blue or green key assigned on the computer keyboard. After choosing the image, a comprehension question with two alternative answers marked with "A)" and "B)" appeared on the screen, such as the example (15) shows.

(11)   Quem estava atrasada?   'Who was late?'

a)   A aluna 'The student'   b) A professora   'The teacher'

Again, they had to choose the option that best represented the sentence heard by pressing the blue or green key assigned on the computer keyboard.

Statistical Analysis

The statistical analysis of the results used linear mixed-effects regression model (BAAYEN ET AL., 2008) in the software RStudio (R CORE TEAM, 2021).

2.4 Results
2.4.1 Online Measure

The online measure collected for analysis was the Fixation Count (FC), which corresponds to the number of participants' eye fixations to each picture during the listening of each experimental sentence.

For the analysis of processing moment-by-moment, we defined four Regions of Interest (ROIs) based on time windows in the audio files of each sentence. These ROIs are exemplified below.

(9)   **ROI 1:** onset of the sentence to the offset of the matrix verb

"A aluna disse" ('The student told')

(10)   **ROI 2:** onset of the dative noun to the offset of the ambiguous clause

Short -> "à professora que estava atrasada" ('the teacher that was late')

Long -> "à professora de português da escola que estava atrasada" ('the Portuguese teacher of the school that was late')

(11)   **ROI 3:** the ambiguous clause

"que estava atrasada" ('that was late')

(12)   **ROI 4:** onset of the second that-clause to the end of the sentence

"que precisaria sair da sala" ('that she would need to leave the room')

The first ROI is important for us to see which picture subjects start looking at and if there is a shift when they start listening to the second ROI. The ROI 2 is important to see if participants are guided by prosody and analyze the ambiguous clause based on the prosodic phrasing they are listening to. The third ROI is the most important of all. This ROI contains only the ambiguous clause and it is where we can be sure of subjects' parsing of it. ROI 4 is important because it is where participants will build their final comprehension of the sentence and where we expect to see the garden-path effect in the GP conditions.

We will start reporting the results of the FC means at the CP and RC images per conditions in the first ROI. For the FCs, values equal to 0 and above 9.8 were dropped from the analysis. Graph 1 shows the FC means at each picture per condition in the first ROI.

FC means at CP and RC images in ROI 1 per condition
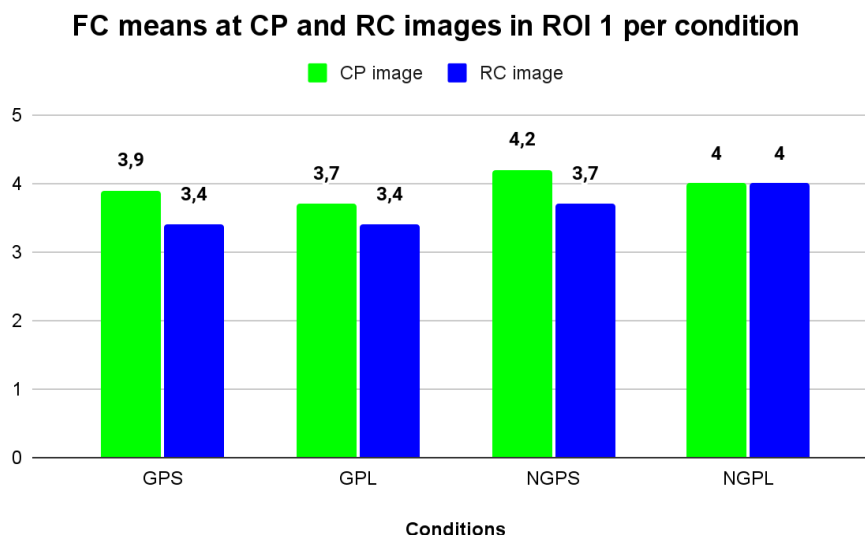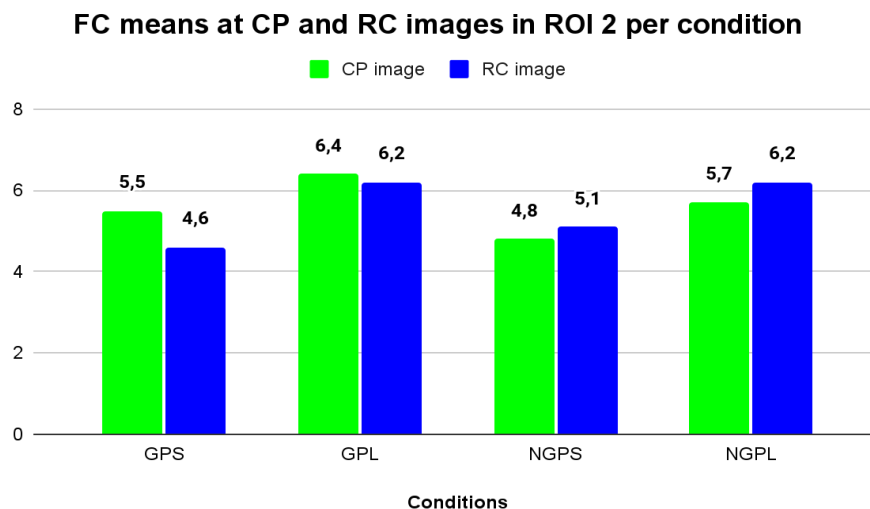

GRAPH 1 - FC means at the CP and RC images in ROI 1 per condition.
Source: elaborated by the author

The results of the FC means show that participants looked more at the CP image in all the conditions in the first ROI. In order to see if these differences were significant, we a performed linear mixed-effects regression model (BAAYEN ET AL., 2008) in RStudio (R CORE TEAM, 2021). The model included the FCs as a function of the conditions and pictures as fixed effects, and participants and items as random effects.

The statistical analysis of the FCs revealed only a significant effect of image type, indicating that the RC pictures received fewer fixations than the CP pictures (Estimates: -0.50, CI: [-0.96 ~ -0.04], p = 0.032).

In the second ROI, FCs equal to 0 and above 11.5 were dropped from the analysis. Graph 2 shows the FC means at each picture per condition in the second ROI.

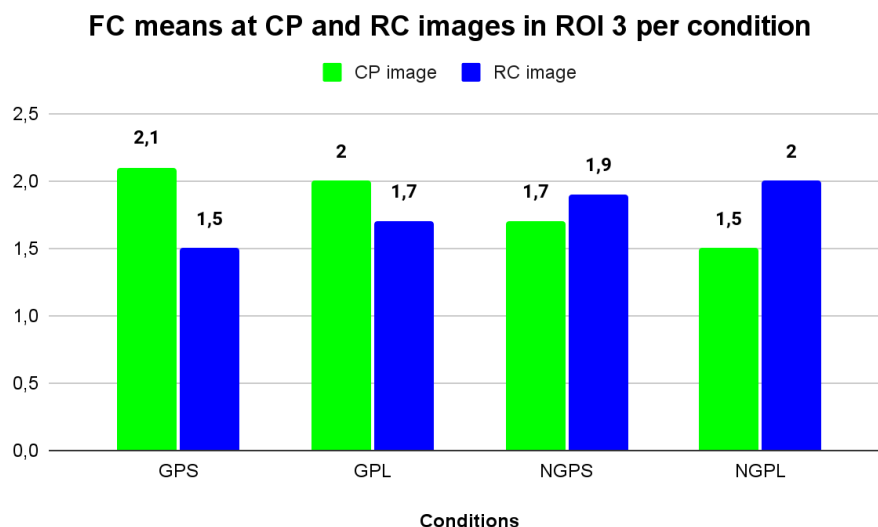## FC means at CP and RC images in ROI 2 per condition

GRAPH 2 - FC means at the CP and RC images in ROI 2 per condition.
Source: elaborated by the author

In this region, the FC means show that participants looked more at the picture that was consistent with the prosodic condition they were listening to. In the GP conditions, participants looked more to the CP picture, whereas, in the NGP conditions, participants looked more to the RC picture. Another linear mixed-effects regression model (BAAYEN ET AL., 2008) was performed in RStudio (R CORE TEAM, 2021). The model included the FCs as a function of the conditions and pictures as fixed effects, and participants and items as random effects.

The analysis revealed a significant difference between the GPS and the NGPS conditions, indicating fewer FCs in the NGPS conditions (Estimates: -0.73, CI: [-1.23 ~ -0.23], p = 0.004). A significant effect of image type was also found, indicating that the RC pictures received fewer fixations than the CP pictures (Estimates: -0.90, CI: [-1.40 ~ -0.40], p < 0.001). Again, there was a significant interaction between the NGPS condition and the RC picture (Estimates: 1.25, CI: [0.55 ~ 1.96], p = 0.001) and another between the NGPL condition and the RC picture (Estimates: 1.51, CI: [0.80 ~ 2.22], p < 0.001), indicating that the RC pictures received more fixations in the NGPS and NGPL conditions.

In the third ROI, FCs above 6 were dropped from the analysis. Graph 3 shows the FC means at each picture per condition in ROI 3.

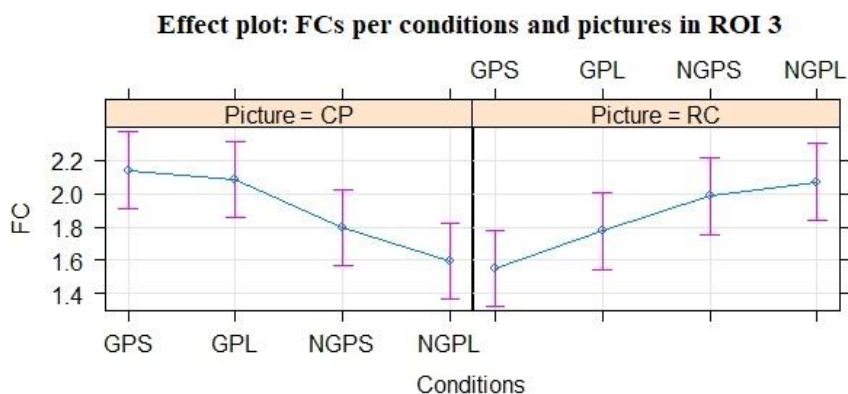## FC means at CP and RC images in ROI 3 per condition

GRAPH 3 - FC means at the CP and RC images in ROI 3 per condition.
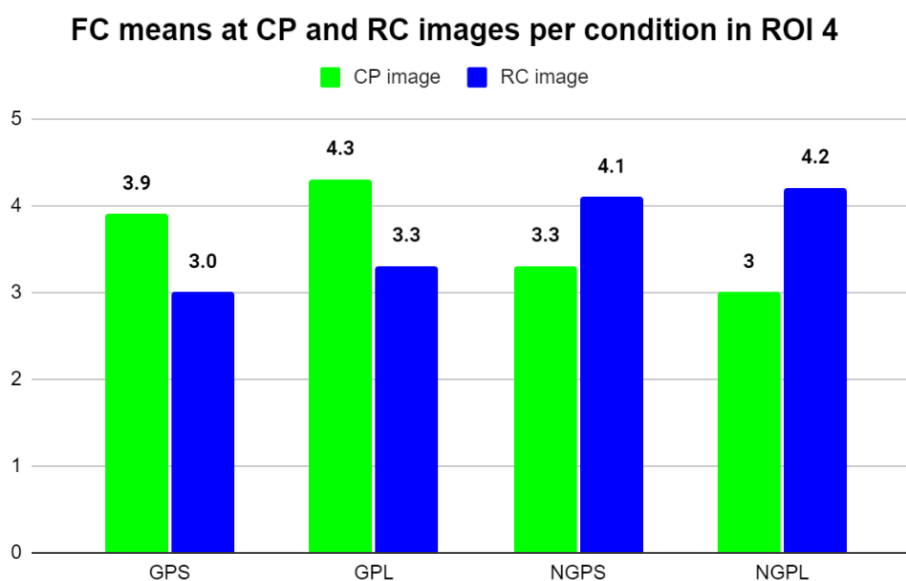Source: elaborated by the author

The FCs replicated the results found in the second ROI. Participants looked more to the picture that was consistent with the prosodic condition they were listening to. In the GP conditions, participants looked more to the CP picture. In the NGP conditions, participants looked more to the RC picture. Again, we analyzed this data using a linear mixed-effects regression model (BAAYEN ET AL., 2008) in RStudio (R CORE TEAM, 2021). The model included the FCs as a function of the conditions and pictures as fixed effects, and participants and items as random effects.

The analysis showed a significant effect of image type, indicating that the RC pictures received fewer fixations than the CP pictures (Estimates: -0.59, CI: [-0.91 ~ -0.27], p < 0.001). As well as in ROI 2, a significant difference between the GPS and the NGPS conditions was found, indicating fewer FCs in the NGPS conditions (Estimates: -0.35, CI: [-0.67 ~ -0.03], p = 0.034). Significant interactions were found between the NGPS condition and the RC picture (Estimates: 0.78, CI: [0.33 ~ 1.24], p = 0.001) and between the NGPL condition and the RC picture (Estimates: 1.06, CI: [0.61 ~ 1.52], p < 0.001), indicating that RC pictures received more fixations in these conditions. Graph 4 shows the effect plot of the FCs per conditions and pictures.

Effect plot: FCs per conditions and pictures in ROI 3

GRAPH 4 - Effect plot of the FCs per conditions and pictures in ROI 3.
Source: elaborated by the author

In the fourth ROI, FCs above 9.5 were dropped from the analysis. Graph 5 shows the FC means at each picture per condition in ROI 4.

**FC means at CP and RC images per condition in ROI 4**

GRAPH 5 - FC means at the CP and RC images in ROI 4 per condition.
Source: elaborated by the author

The FCs of the last region of the sentence confirm that participants were guided by prosody since the second ROI and looked more to the picture that matched the prosodic conditions to which they were exposed. A linear mixed-effects regression model (BAAYEN ET AL., 2008) was performed again in RStudio (R CORE TEAM, 2021). The model included the FCs as a function of the conditions and pictures as fixed effects, and participants and items as random effects.

The analysis replicated the results found in ROI 3. A significant effect of image type was found, indicating that the RC pictures received fewer fixations than the CP pictures (Estimates: -0.94, CI: [-

1.42 ~ -0.46], p < 0.001). Significant interactions were found between the NGPS condition and the RC picture (Estimates: 1.79, CI: [1.11 ~ 2.47], p < 0.001) and between the NGPL condition and the RC picture (Estimates: 2.14, CI: [1.46 ~ 2.82], p < 0.001), indicating that the RC pictures received more fixations in these conditions. Graph 6 shows the effect plot of the FCs per conditions and pictures.
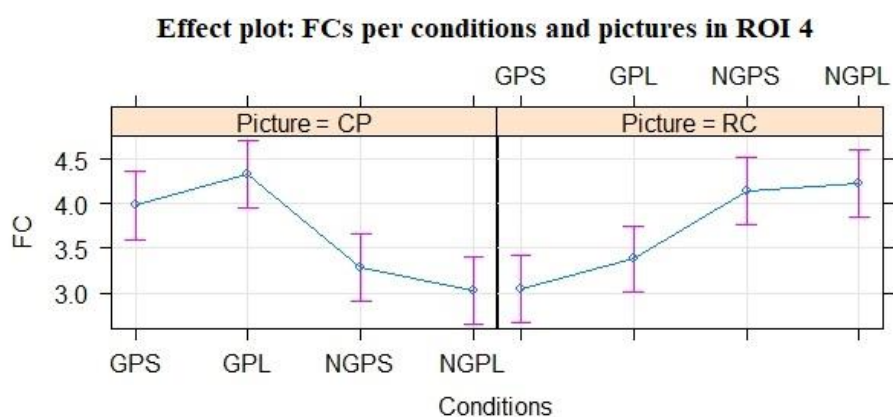

GRAPH 6 - Effect plot of the FCs per conditions and pictures in ROI 4.
Source: elaborated by the author

The results of the FCs confirm our hypotheses on the assignment of syntactic structure depending on prosodic phrasing presented in the Introduction. When participants listened to a GP prosody, there were more fixations at the CP pictures. When they listened to an NGP prosody, there were more fixations at the RC pictures.

2.4.2 Offline Measures

The offline measures collected for analysis were (1) the RTs to answer the comprehension questions and (2) the answers to the comprehension questions. Graph 7 shows the RT means per conditions and answers (right *vs.* wrong).

RT means (ms) per condition and answers

■ Right ■ Wrong
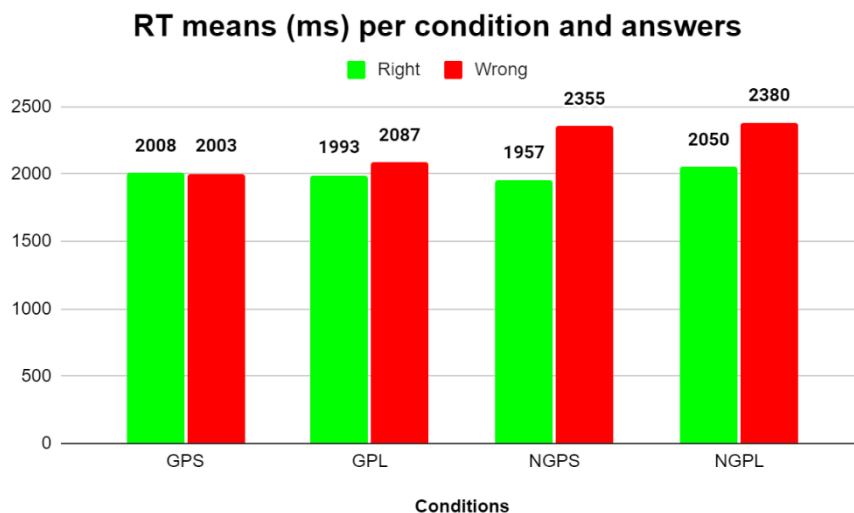
GRAPH 7 - RT means (ms) per condition and answers.
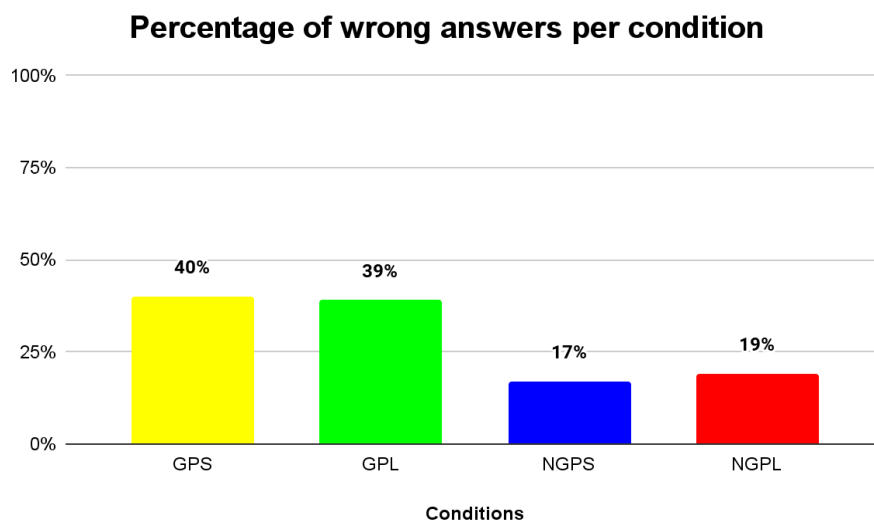Source: elaborated by the author

The graph shows that the RT means are higher when participants chose a wrong answer in the GPL and, especially, in the NGPS and NGPL conditions. In order to verify if these differences were significant, we performed a linear mixed-effects regression model (BAAYEN ET AL., 2008) in RStudio (R CORE TEAM, 2021). The model included log-transformed RTs as a function of participants' answers and conditions as fixed effects, and participants and items as random effects. No significant difference nor an interaction between the conditions and the answers were found. We also conducted a post-hoc Tukey HSD test, but no significant differences were found. However, Graph 8 shows that there was a tendency of RTs to be longer when participants chose a wrong answer in the NGPS and NGPL conditions.

Effect plot: RTs per answers and conditions
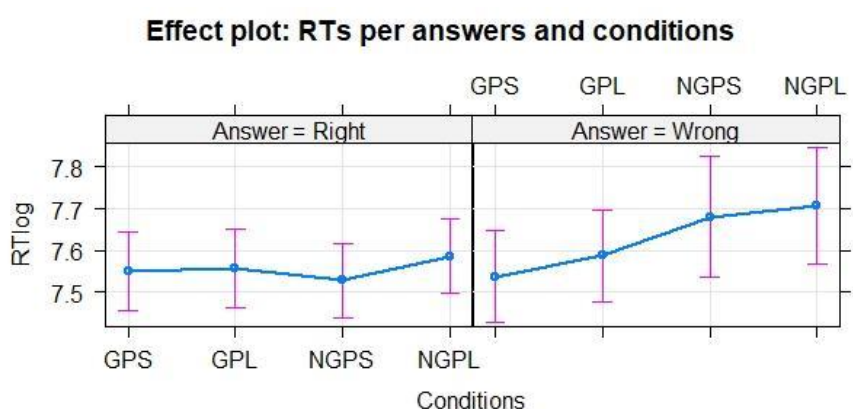
GRAPH 8 - Effect plot of RTs per answers and condition.
Source: elaborated by the author

We also analyzed the rates of wrong answers per condition. The data clearly shows a difference between the GP and the NGP conditions. The GP conditions presented a higher percentage of wrong answers compared to the NGP conditions. This result indicates that prosody guided participants' parsing of the ambiguous clause and the GP conditions were more difficult for them. Graph 9 shows the percentage of wrong answers per condition.

Percentage of wrong answers per condition

GRAPH 9 - Percentage of wrong answers per condition.
Source: elaborated by the author

This data was analyzed in a generalized linear mixed-effects regression model for binomial data (BAAYEN ET AL., 2008) in RStudio (R CORE TEAM, 2021). The model included participants' answers as a function of the factors Prosody and Length as fixed effects, and participants and items as random effects. The statistical analysis revealed a significant difference between the GP and NGP prosody (OR: 3.91, CI: [1.68 ~ 9.12], p = 0,002), indicating that the chances of a right answer are significantly higher in the NGP conditions. No significant interaction between Prosody and Length was found. Graph 10 shows the effect plot of the probability of choosing a right answer per condition.

Effect plot: Probability of Right Answers per condition
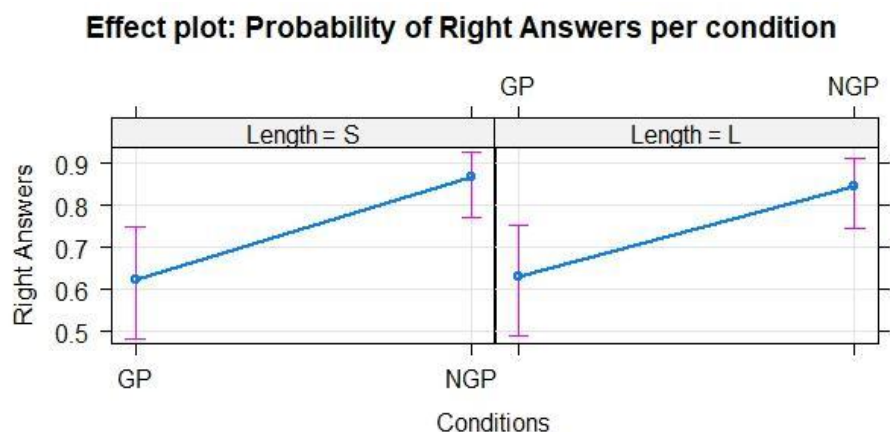
GRAPH 10 - Effect plot of the probability of right answers per condition.
Source: elaborated by the author

3. Conclusions

The results of both the online and offline measures clearly reveal the influence of prosody on the participants' processing of the sentences. Prosody was able to guide the parsing of the sentences at a very early stage of processing and determine the syntactic structure that subjects build of the sentences. These results are aligned with previous works on the prosody-syntax interface in sentence processing, such as the ones briefly reviewed in section 1, and confirm our hypotheses on the assignment of syntactic structure depending on prosodic phrasing presented in the Introduction.

In the second and third ROIs, the FCs to the pictures provided robust online evidence of the influence of prosody on subjects' syntactic parsing. When participants were exposed to the GP prosody, they looked more at the CP pictures. When they were exposed to the NGP prosody, they looked more at the RC images. This data shows the role of prosody at a very early stage of processing. In the case of the NGP prosody, it contradicts the Minimal Attachment principle (FRAZIER, 1979) and the preference for arguments over adjuncts. These principles did not operate in the presence of prosodic cues that pointed to a different syntactic structure from the preferred one predicted by these strategies.

The last ROI was important to check participants' final comprehension of the sentence and it was where we expected to see the garden-path effect in the GP conditions. The FCs replicated the results found in ROI 3. Participants kept looking at the same pictures they were looking at when they listened to the ambiguous clause. In the GP conditions, we expected to find an oscillation pattern between the two pictures, indicating that participants fell into a garden-path. However, it seems like participants did not try to look at the other picture in order to solve the ambiguity. Perhaps, subjects preferred to analyze the two that-clauses as two CPs in coordination. This analysis could even be more consistent with the prosodic phrasing they have been exposed to.

The results of the offline measures also showed the impact of prosody on participants' final comprehension of the sentences. Although there was no significant difference among the RTs comparing

the conditions and the answers given, the NGP conditions presented slower RTs when participants chose the wrong answer. This could indicate some hesitation when subjects chose the wrong answer in these conditions, whereas, when they chose a wrong answer in the GP conditions, the RTs were very similar, showing that they were not necessarily sure of the answer.

Finally, the analysis of the answers chosen reveals the influence of prosody on participants' final comprehension of the sentences. The error rate was significantly higher in the GP conditions (GPS - 40% and GPL - 39%) compared to the error rate in the NGP ones (NGPS - 17% and NGPL - 19%). This result is similar to the one found by Silva (2018). The author found an error rate of 36.7% in comprehension questions testing the same structure in an eye-tracking reading experiment for university-level students. It was in the analysis of the answers chosen that we expected to see an effect of length. However, no effect of length was found.

Furthermore, we believe that the percentage of wrong answers in the GP conditions may reflect a good-enough effect (FERREIRA ET AL., 2002), in which the partial comprehension of the sentence occurs due to the initial erroneous parsing that persists in the linguistic representation. This is a speculation that could not be fully investigated in this study, but constitutes an interesting topic for future research.

Additional information

Editors

Luma Miranda
Afiliation: Eötvös Loránd University
ORCID: https://orcid.org/0000-0002-5529-0338

Manuella Carnaval
Afiliation: Federal University of Rio de Janeiro
ORCID: https://orcid.org/0000-0002-4321-5859

Carolina Gomes da Silva
Afiliation: Federal University of Paraíba
ORCID: https://orcid.org/0000-0002-1490-0814

EVALUATION ROUNDS

Evaluator 1: Andressa Christine Oliveira da Silva
Affiliation: Federal University of Juiz de Fora
ORCID: https://orcid.org/0000-0002-3011-7723

Evaluator 2: Denise Cristina Kluge
Affiliation: Federal University of Rio de Janeiro
ORCID: https://orcid.org/0000-0003-4656-7902

EVALUATOR 1

O artigo "Prosody and Sentence Processing in Brazilian Portuguese: a Visual World Paradigm Study" busca investigar o papel da prosódia no processamento de frases ambíguas, do tipo garden-path, no português brasileiro. O estudo utiliza a metodologia experimental em Psicolinguística, com a técnica Paradigma do Mundo Visual.

Com relação à parte estrutural do trabalho:

1) o resumo é claro, apresentando o objetivo, a relevância do tema, a metodologia e os resultados finais;

2) a introdução combina com o objetivo que é apresentado no resumo;

3) a metodologia está clara e explica em detalhes como o experimento foi construído e aplicado;

4) os resultados e a discussão do estudo estão explicados de forma clara; fiz apenas uma sugestão que considero importante para os gráficos 4 e 6.

5) as conclusões estão feitas de forma sucinta e retomam o que foi discutido ao longo do artigo.

Fiz algumas sugestões de escrita e algumas recomendações para que o artigo fique o mais claro possível para os leitores, tanto leigos quantos os da área. O arquivo está em anexo.

No geral, o trabalho traz resultados interessantes e importantes para o estudo da interface sin-taxe-prosódia, bem como contribui para ampliar os estudos com o uso do rastreamento ocular no português brasileiro, especialmente com o uso da metodologia de Paradigma do Mundo Visual. A área de Psicolinguística no Brasil ainda carece de estudos que utilizem esta técnica no Brasil. Portanto, re-comendo fortemente a publicação do artigo.

EVALUATOR 2:

De forma geral o artigo está bem escrito e organizado e traz um título adequado ao trabalho e ao tema proposto. O resumo descreve de forma clara e compreensível o estudo realizado incluindo objetivos, metodologia e resultados.

A introdução, de forma geral, apresenta, de forma clara, os objetivos, o contexto e os principais conceitos abordados ao longo da pesquisa. No entanto, sugiro guiar um pouco mais o leitor indicando em quais seções determinados conceitos serão apresentados/discutidos posteriormente no texto.

Quanto à seção de metodologia, o texto apresenta de forma clara o contexto da pesquisa, os participantes, os testes desenvolvidos e o procedimento de coleta. Sugiro indicar, nessa seção, quais testes e programa foram utilizados para a análise estatística.

Os resultados são apresentados de forma clara e os gráficos e figuras contribuem para a apresentação e discussão dos dados. Ressalto que as referências usadas para discutir os dados são apropriadas e relevantes.

Sugiro retomar as duas hipóteses da tua pesquisa na seção de análise dos resultados, bem como nas considerações finais. Além disso, sugiro incluir as limitações do trabalho e sugestões para pesquisas futuras a partir da pesquisa realizada.

AUTHOR'S ANSWER

O trabalho revisado atende aos pedidos de correção e/ou recomendações dos pareceristas tanto em relação a questões de forma quanto de conteúdo do manuscrito.

Em relação às sugestões da parecerista A, foram realizadas todas as modificações de texto recomendadas. Quanto à sugestão que consta na página 22, o autor decidiu remover as referências às quais faz menção por julgá-las fora do escopo da pesquisa relatada.

Em relação às sugestões da parecerista B, foram realizadas as seguintes modificações: (1) inclusão de um parágrafo detalhando a estrutura do artigo na Introdução; (2) acréscimo de um período sobre os testes e programa foram utilizados para a análise estatística na seção de Metodologia; (3) retomada das hipóteses apresentadas na Introdução nas seções de Análise de Resultados e de Conclusões; e (4) inclusão de um parágrafo sobre as limitações do trabalho e sugestões para pesquisas futuras a partir da pesquisa realizada nas Conclusões.

Interest conflict

O autor não tem conflitos de interesse a declarar.

Research Protocol and Pre-Registration

Data Availability Statement

Os dados, códigos e materiais que suportam os resultados deste estudo estão disponíveis para consulta sob demanda em drive institucional de responsabilidade do autor.

Research with Human Beings

Financing Source

REFERENCES

BAAYEN, Rolf; DAVIDSON, Douglas; BATES, Douglas. Mixed-effects modeling with crossed random effects for subjects and items. **Journal of Memory and Language**, v.59, p. 390–412, 2008.

BERENDS, Sanne; BROUWER, Susanne; SPRENGER, Simone. Eye-tracking and the visual world paradigm. *In*: SCHMID, M. S. *et al.* (Authors). **Designing Research on Bilingual Development**: Behavioral and Neurolinguistic Experiments. Switzerland: Springer International Publishing, p. 55-78, 2015.

CALDAS, Vitor. The role of prosody in the processing of garden-path sentences: online and offline evidence from production and comprehension. 166 f. Tese (Doutorado em Língua Portuguesa) - Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2022.

CARLSON, Katy; CLIFTON, Charles Jr.; FRAZIER, Lyn. Prosodic boundaries in adjunct attachment. **Journal of Memory and Language**, 45, p. 58-81, 2001.

CLIFTON, Charles Jr.; CARLSON, Katy; FRAZIER, Lyn. Informative Prosodic Boundaries. **Journal Language and Speech**, 45, p. 87-114, 2002.

CLIFTON, Charles Jr.; CARLSON, Katy; FRAZIER, Lyn. Tracking the what and why of speakers' choices: Prosodic boundaries and the length of constituents. **Psychonomic Bulletin & Review**. v.13 (5), p. 854-861, 2006.

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

FERREIRA, Fernanda; BAILEY, Karl; FERRARO, Vittoria. Good-enough representations in language comprehension. **Current Directions in Psychological Science**, 11(1), p. 11-15, 2002.

FONSECA, Aline; SILVA, Andressa. Os estudos da interface sintaxe-prosódia na psicolinguística. *In*: MAIA, Marcus. **Psicolinguística**: diversidades, interfaces e aplicações. São Paulo: Contexto, p. 256-271, 2022.

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

FROTA, Sónia; OLIVEIRA, Pedro; CRUZ, Marisa; VIGÁRIO, Marina. P-ToBI: tools for the transcription of Portuguese prosody. Lisboa: Laboratório de Fonética, CLUL/FLUL, 2015.

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

LADD, Dwight Robert. **Intonational phonology**. Cambridge: CUP, 1996.

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

NESPOR, Marina; VOGEL, Irene. **Prosodic Phonology**. Foris: Dordrecht. 1986.

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

SERRA, Carolina. 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, Emily. 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, Aline. 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.

SPEER, Shari; KJELGAARD, Margaret; DOBROTH, Kathryn. The Influence of Prosodic Structure on the Resolution of Temporary Syntactic Closure Ambiguities. **Journal of Psycholinguistic Research**, Vol. 25, No. 2, p. 249-271, 1996.