https://doi.org/10.31513/linguistica.2020.v16nEsp.a38165
Edição Especial Comemorativa

REPEATED-NAME PENALTY: A MULTIFACTORIAL EFFECT

PENALIDADE DO NOME REPETIDO: UM EFEITO MULTIFATORIAL

Eva Vilma Aires Cabral Gondim[1]
Márcio Martins Leitão[2]
Matheus de Almeida Barbosa[3]

RESUMO

Este artigo investigou o processamento anafórico de nomes repetidos e pronomes plenos em Português Brasileiro (PB), focando aspectos teóricos e metodológicos relacionados ao efeito da Penalidade do Nome Repetido (PNR), que consiste em um aumento de custo de processamento de nomes repetidos quando comparados com pronomes no estabelecimento da correferência. Neste estudo, apresentaremos dois experimentos que conduzimos com a técnica de leitura automonitorada manipulando os fatores: tipo de retomada (nome repetido, pronome pleno); quantidade de antecedentes humanos (um ou dois antecedentes), e controlando a forma de segmentação/aferição do tempo de leitura (aferição do tempo de leitura apenas da retomada, aferição do tempo de leitura da sentença inteira que continha a anáfora). Os resultados encontrados apontaram para uma possível interação entre esses fatores influenciando a ocorrência ou não da PNR, fortalecendo a nossa hipótese geral de que essa penalidade é um efeito multifatorial, o que pode explicar os resultados divergentes encontrados na literatura sobre a PNR em PB.

PALAVRAS-CHAVE: Psicolinguística Experimental; Processamento Anafórico; Penalidade do Nome Repetido.

ABSTRACT

This article investigated the anaphoric processing of repeated names and overt pronouns in Brazilian Portuguese (BP), focusing on theoretical and methodological aspects related to the repeated-name penalty effect (RNP), which consists of an increase in the cost of processing repeated names when compared with pronouns in coreference resolution. In this study, we present two experiments that we conducted using the self-paced reading technique manipulating the factors: anaphor type (repeated name, overt pronoun); number of human antecedents (one or two antecedents) and controlling the segmentation type/presentation of critical segment (presenting the whole sentence at once, presenting the anaphor alone). The results showed a possible interaction between these factors influencing the occurrence or not of the RNP, strengthening our general hypothesis that this penalty has a multifactorial effect, which could explain the divergent results found in the RNP literature in BP.

KEYWORDS: Experimental Psycholinguistics; Anaphoric Processing; Repeated-Name Penalty.

[1] Mestre em Linguística pelo Programa de Pós-graduação em Linguística (PROLING) da Universidade Federal da Paraíba (UFPB). Contato: evagondim@hotmail.com.
[2] Doutor em Linguística pela Universidade Federal do Rio de Janeiro (UFRJ), professor da Universidade Federal da Paraíba (UFPB) e pesquisador 2 do CNPq. Contato: leitaomm@pesquisador.cnpq.br.
[3] Doutorando pelo Programa de Pós-graduação em Linguística (PROLING) da Universidade Federal da Paraíba (UFPB). Contato: matheusba@gmail.com.

Introduction

In discourse, relations between linguistic elements that provide textual cohesion are established. One of these cohesive mechanisms is anaphoric coreference that corresponds to the process in which a linguistic expression refers to another previously mentioned entity in the discourse. These referential expressions may be, depending on the language, a repeated name, an overt pronoun, an empty category/null pronoun, or others. Many scholars in the field of Psycholinguistics have sought to understand which factors influence the processing of certain anaphoric forms.

Many studies in several languages have shown that distinct anaphoric forms are processed differently, for example, evidencing that repeated-name anaphors are processed more slowly than pronouns when there is discourse salience in the antecedent-anaphor relation. This processing drawback of the repeated name was called the repeated-name penalty (RNP) by Gordon, Grosz and Gilliom (1993).

In Brazilian Portuguese (BP), there is divergence among the results of studies, which we will address in section 1; some of these studies find the RNP while others do not. However, if we closely analyze these studies, we realize that there are differences both in the linguistic stimuli used and in the methods. Thus, the main objective of the present study was to investigate how pronouns and repeated-name anaphors are processed in Brazilian Portuguese (BP), focusing on the repeated-name penalty effect. The general hypothesis is that this penalty has a multifactorial effect, that is, it results from the combination of several factors, both linguistic and methodological, and the variation of these factors is what would be behind the divergent results found in BP.

In section 1, we briefly address psycholinguistic studies on the RNP, focusing on BP studies. In section 2, we report two experiments using the self-paced reading paradigm to analyze coreferential processing of pronouns and repeated-name anaphors, focusing on the RNP in BP and the differences found in other studies. In section 3, we conduct a general discussion, and in section 4, we make our final remarks.

1. Psycholinguistic research on the repeated-name penalty

Gordon, Grosz, and Gilliom (1993) conducted research in the English language comparing anaphoric processing of repeated names and overt pronouns using the self-paced reading technique to test the prediction of the Centering Theory (GROSZ et al., 1995). This theory states that every sentence refers to an entity, called backward-looking center, thus allowing the connection with the previous clause. The pronoun is the preferred linguistic form to refer to the more salient antecedent because it promotes continuity and discursive coherence. Their results showed that the processing time of repeated-name anaphors, which corresponded to salient antecedents (those in subject position), was higher than the one for pronouns. The authors called this processing delay of repeated names for pronouns the repeated-name penalty (RNP).

Since the seminal study of Gordon et al. (1993), several psycholinguistic studies have focused on the RNP in different languages. The penalty between overt pronouns and repeated names has been confirmed in other studies conducted in English (GORDON & CHAN, 1995; CHAMBERS & SMYTH, 1998; KENISSON & GORDON, 1997) French (ERNST, 2007), Mandarin Chinese (YANG et al., 1999), Spanish (EGUSQUIZA et al., 2016), and Japanese (SHOJI et al., 2017). However, this effect was not found in studies conducted in Argentinean Spanish (GERLOMINI-LEZAMA, 2008, 2010) and in Brazilian Portuguese/BP (MAIA & CUNHA LIMA, 2011, 2012; MAIA, 2013; LIMA, 2015; ALMOR et al., 2017), which diverge from the research conducted by Leitão and colleagues (LEITÃO, 2005; QUEIROZ & LEITÃO, 2008; LEITÃO & SIMÕES, 2011) who found systematic RNP in BP. Thus, attention should be paid to the differences among the studies that have focused on the RNP in BP and have reported diverging results.

Therefore, to investigate the similarities and differences between the methods and linguistic stimuli used in the studies, we conducted a psycholinguistic study on the repeated-name penalty (RNP) in Brazilian Portuguese (BP).

1.1 Studies on the repeated-name penalty in Brazilian Portuguese

The pioneering study in Brazilian Portuguese (BP) of coreferential processing on the repeated-name penalty (RNP) was conducted by Leitão (2005) in his doctoral thesis. Based on the assumptions of the Centering Theory and the Information Load Hypothesis, the author investigated the processing of the anaphoric direct object in BP.

Leitão (2005) designed 5 online and offline experimental tasks to analyze the coreferential processing established by the anaphoric direct object in repeated names, lexical/overt pronouns, null pronouns, and NPs related to hypernymy and/or hyponymy and the antecedents. The author investigated anaphoric processing of repeated names versus overt pronouns focusing on the RNP, which will be described in Experiment 2.

Experiment 2 was performed using the online self-paced reading technique. The experimental items consisted of two coordinate sentences that contained an anaphor (repeated name or overt pronoun) in direct object position corresponding to an antecedent also in direct object position, therefore, in parallel structures. The sentences were divided into 10 segments and the reading time of referential expression was measured. The results of this research showed that the mean reading times of segments with referential expressions and overt pronouns was significantly lower than those with repeated names, showing that the RNP in BP occurs when referents are in the direct object position.

From this study, Leitão and colleagues conducted a series of experimental studies using the self-paced reading technique and the eye-tracking paradigm (LEITÃO, RIBEIRO and MAIA, 2012) to observe coreferential processing of overt pronouns and repeated names that confirmed the RNP in BP (QUEIROZ and LEITÃO, 2008; LEITÃO and SIMÕES, 2011; GONDIM and LEITÃO, 2012; LIMA, 2014; BARBOSA, GONDIM and LIMA, 2016) both for referents in subject position and referents in object position.

Queiroz and Leitão (2008) investigated the coreferential processing of overt pronouns versus repeated names and hypernyms versus hyponyms in subject position using the self-paced reading technique. Experiment 1 focused on the RNP. The linguistic stimuli used in this experiment consisted of two parallel coordinate structures, divided into nine segments, containing an anaphor (overt pronoun or repeated name) in subject position that referred to an antecedent in the same syntactic position to measure the reading time of the critical segment

corresponding to the anaphoric expression. According to the results, overt pronouns had a significantly lower processing time than repeated names, confirming the occurrence of the RNP.

Leitão, Ribeiro and Maia (2012) replicated the studies of Leitão (2005) and Queiroz and Leitão (2008) using the eye-tracking paradigm to test whether the RNP would occur using a different technique (not segmented and more accurate than the self-paced reading task). The results from the two experiments revealed that the duration of the first fixation in the region following the anaphor (spillover) was significantly higher in the conditions with repeated names than in the conditions with overt pronouns, both in subject and object position, corroborating the findings of Leitão (2005) and Queiroz and Leitão (2008) for the RNP.

Gondim and Leitão (2012) also replicated the first experiment conducted by Queiroz and Leitão (2008) using the self-paced reading technique to verify whether the method, particularly the presentation/segmentation of the linguistic items could influence the experimental results. The authors used the same experimental structures from the study but changed the segmentation of stimuli and what was measured in self-paced reading task, that is, instead of measuring the time to read only the anaphor, as described in Queiroz and Leitão (2008), they measured the entire critical sentence that contained the anaphoric element. The results showed that, despite using different methodological approaches, the RNP was observed.

To investigate if structural parallelism can influence the RNP, Barbosa, Gondim and Lima (2016) conducted an experiment analyzing the coreferential processing of overt pronouns and anaphoric repeated names using the self-paced reading technique. The experimental passages consisted of two juxtaposed sentences. The first sentence contained two antecedents, one in subject position and the other in object position, and the second sentence contained an anaphor that was either in subject position or in object position that referred to the antecedent in subject position. The stimuli were divided into 10 segments and the reading time of the critical segment that contained the anaphor (overt pronoun or repeated name) was measured. According to the analysis of the critical segment (containing the anaphor), RNP occurred both in sentences with structural parallelism and in sentences without structural parallelism. However, when analyzing the post-critical segment, the authors observed a possible spillover effect, as they found a RNP in sentences with structural parallelism, but they did not find it in sentences without structural parallelism.

However, recent studies conducted in BP focusing on the RNP (MAIA and CUNHA LIMA, 2011, 2012; MAIA, 2013; LIMA, 2015; ALMOR et al., 2017) have pointed out that this effect does not exist in our linguistic system, contradicting the results found in the above-mentioned studies. Based on the studies by Gerlomini-Lezama (2008) in Argentinian Spanish, Maia and Cunha Lima (2011) conducted, for the first time, a self-paced reading experiment to observe the anaphoric processing of repeated names, overt pronouns and null pronouns on the RNP in BP.

The experiment was based on the experimental stimuli used by de Yang et al. (1999). The experimental items consisted of three juxtaposed sentences: (1) the first sentence introduced two antecedents (proper names) of opposite genders, one in subject position and the other in object position; (2) the second sentence (critical one) contained two referents in subject (repeated name, overt pronoun, null pronoun) and object (repeated name, oblique pronoun) positions. Thus, according to the authors, "there was a relationship of syntactic parallelism between the first two sentences, that is, the syntactic functions of the entities in the initial sentence were maintained in the second sentence [...]" (MAIA and CUNHA LIMA, 2011, p. 7); (3) the third sentence did not mention the previously mentioned entities as the objective was not for the critical sentence to stand alone. The stimuli were presented as entire sentences and the reading time of the critical sentence containing the anaphors was measured.

According to Maia and Cunha Lima (2011), the results revealed that the processing time for sentences with anaphoric referents in subject position was higher for repeated names than for null pronouns. However, there were no significant differences in the mean reading times between repeated names and overt pronouns, neither for the anaphors in subject position nor for the ones in object position. The authors interpreted the results as indicative of the non-existence of the RNP in BP, considering only the comparison of repeated names and overt pronouns based on the characterization of this penalty proposed by Gordon et al. (1993)[4].

Maia and Cunha Lima (2012) decided to test the results of the research by Maia and Cunha (2011) using another experimental paradigm to expand the debate about the occurrence or not of the RNP in BP.

The authors conducted an experiment using the eye-tracking technique by using and adapting part of the experimental stimuli used by Maia and Cunha (2011). The experimental

[4] It is noteworthy that Yang et al (1999) and Gerlomini-Lezama (2008) presented a different proposal for the RNP than the one proposed by Gordon et al. (1993), admitting the occurrence of this effect for repeated names versus null pronouns for coreference resolution. Therefore, Maia and Cunha Lima (2011), even though they conducted an experiment along the lines of the study by Yang et al (1999), chose the parameter proposed by Gordon et al. (1993) for the RNP.

stimuli consisted of two juxtaposed sentences: the first sentence introduced two antecedents (a subject and an object) of different genders, and the second sentence (critical one) contained an anaphor[5] (repeated name or overt pronoun) that either referred to the subject or to object antecedent. The independent variables were as follows: antecedent salience (subject or object) and anaphor form (pronoun or repeated name). The dependent variables of the experiment were: duration of the first fixation and total fixation time. The regions analyzed were: (1) location of the anaphor; (2) location of the oblique pronoun; (3) verb position; (4) entire critical sentence.

According to Maia and Cunha Lima (2012), the results obtained in each region analyzed revealed that:

(1) anaphor region: (a) first fixation: there was no significant difference, neither in anaphor form nor in the antecedent salience or in the interaction among the variables. (b) total fixation times: same as the result of the first fixation. According to the authors, these results suggested that repeated names and overt pronouns are processed in a similar way.

(2) region of the oblique pronoun (immediately after the anaphor)[6]: (a) first fixation: there was a significant anaphoric effect, per participant, showing that the processing cost for overt pronouns was higher than for repeated names, but there was no significant difference in the antecedent salience nor in the interaction of the two factors; (b) total fixation times: a significant effect was confirmed in the anaphor form per participant and there was no significant effect on the antecedent salience, but there was an interaction among the variables within the experimental items. According to the authors' interpretation, these results “seem to point to the difficult nature of two pronouns in a sentence (a strong and weak one, in sequence) during the process of reference resolution” (MAIA and CUNHA LIMA, 2012, p. 121-122).

(3) region of the verb (after the oblique pronoun): (a) first fixation: there was a significant anaphoric effect, indicating that the processing cost for repeated names was higher than the one for overt pronouns, therefore, confirming the RNP; however, there was no significant effect on the antecedent salience and interaction among the variables. (b) total fixation times: there was no significant effect for any of the variables and there was no

[5] In the example of an experimental item in the article by Maia and Cunha Lima (2012), the critical sentence contains two possible referents: a subject anaphor (repeated name or overt pronoun) or an object anaphor (clitic pronoun). However, the authors only considered the referent in subject position.

[6] In the regions after the critical segment, the effect known as spillover can occur, that is, an effect that was expected for the critical region can be found.

interaction among them. Although the RNP was confirmed in the results of the first fixation, the authors argued as follows:

> This late effect, in which processing differences appear after the critical segment (in our case, the region of the main anaphoric term), is quite common in coreference processing studies (Camblin et al., 2007). Although it is natural, the effect does not confirm, however, the RNP as proposed by Gordon, Grosz and Gilliom (1993) and Kennison and Gordon (1997), since an essential element of the processing penalty is the syntactic salience of the antecedent; that is, the retrieval of the subject (higher syntactic position) using a repeated name would cause RNP, while the retrieval of an object (lower syntactic position) using the same repeated name would not cause the penalty. However, repeated names caused processing difficulties in the verbal region when retrieving both the salient (subject) and non-salient antecedents (object). (MAIA and CUNHA LIMA, 2012, p. 123)

(4) region of the entire critical sentence: the total reading times revealed that there was a significant effect on the antecedent salience per participant, favoring anaphors corresponding to salient antecedents; there was no significant effect in an anaphor form or in the interaction of factors.

Therefore, we understand that the RNP seems to have occurred in spillover regions, as found in English studies that observed the RNP in object position (GORDON and CHAN, 1995) and even with other different factors, such as Structural Parallelism (CHAMBERS and SMYTH, 1998), but Maia and Cunha Lima (2012) preferred to conclude that the results found in their study suggest the non-existence of this effect in BP.

Maia (2013) conducted six experiments for his dissertation focusing on RNP and OPP[7]. The first three experiments investigated RNP and OPP in BP using the self-paced reading technique. The fourth experiment was an offline grammaticality judgement task using stimuli from Experiment 3. The fifth experiment focused only on RNP in BP using the eye-tracking technique, and the sixth experiment analyzed the RNP effect in European Portuguese (EP). Among them, we will describe the experiments that focused on the RNP in BP using online techniques and we will only consider the results related to this effect.

Experiment 1, using the self-paced reading technique, was a replication of Gerlomini-Lezama's first experiment (2008) for Argentinian Spanish adapted for BP. The experimental items consisted of two juxtaposed sentences: the first sentence included two antecedents

[7] Gerlomini-Lezama (2008, 2010) investigated the RNP by analyzing the anaphoric correspondence of repeated names, overt pronouns, and null pronouns in Argentinian Spanish using the self-paced reading technique. The results revealed that overt pronouns were penalized (higher processing costs) compared to null pronouns, and the author called it the overt pronoun penalty (OPP). It should be noted that OPP is not the focus of this article.

(subject and object) of opposite genders; the second sentence (critical) contained two possible anaphors, one subject (repeated name, overt pronoun, null) and the other object (overt pronoun), but only one (subject) was considered by the author. The subject anaphor referred either to the subject antecedent or to the object antecedent. The independent variables were as follows: form of the referential expression (repeated name or overt pronoun) and antecedent salience[8] (subject, object). The dependent variable was the reading time of the entire critical sentence. The results showed a significant effect of antecedent salience, showing that in subject antecedent conditions were read faster than in object antecedent conditions, that is, a significant effect for the anaphor form was observed, suggesting that repeated names and overt pronouns are read similarly and the interaction among the variables suggest that one affected the level of the other. In this research, the author interpreted the results as indicative of the non-existence of RNP in BP.

Experiment 2 used the same experimental technique and the same stimuli used in the first experiment, changing the overt pronouns, in object position of the critical sentence, to oblique pronouns. The results were the same as those found in Experiment 1.

Experiment 3 also used the self-paced reading technique. The linguistic stimuli had the same syntactic configuration as those used in Experiment 2 but underwent a more rigorous control regarding the length of the critical sentences, semantic relationships between sentences, tense, and aspect of verbs. The results showed a significant effect of antecedent salience and interaction among variables, as in the previous experiments, but they did not reveal a significant effect for anaphor form, indicating that overt pronouns and repeated names are processed similarly. With these results, the author confirmed the non-occurrence of RNP in BP.

Experiment 5, using the eye-tracking technique, investigated the anaphoric processing of repeated nouns and overt pronouns for coreference resolution. The experimental passages had two juxtaposed sentences, two antecedents (subject and object) of different genders and an anaphor (subject) that alternately referred to the subject and object antecedents. The independent variables were anaphor form (repeated name or overt pronoun) and antecedent salience (subject, object). The dependent variables were duration of the first fixation and total fixation time. The following regions were analyzed: (1) anaphor; (2) verb (after the anaphor); (3) object/adjunct (after the verb); (4) entire critical sentence. The results did not show

[8] The variable that the author called the antecedent salience may be interpreted as structural parallelism, since there is parallelism when the referent is a subject, but there is no parallelism when it is an object.

significant differences between the conditions tested in any of the areas analyzed, unlike the previous experiments that revealed a significant effect for the salience variable of the antecedent. Once again, Maia (2013) confirmed the non-existence of RNP in BP.

Lima (2015) also investigated RNP in BP using two eye-tracking experiments to compare the anaphoric processing of repeated nouns, overt and null pronouns. As the results of the experiments did not reveal any significant differences between the reading times of repeated names and overt pronouns and considering the parameter proposed by Gordon et al. (1993) to interpret the RNP, the author concluded that this penalty was not confirmed under the conditions tested in his dissertation[9]. However, in this study, unlike all others, there was gender ambiguity between potential antecedents and the referents.

Almor et al. (2017) conducted a self-paced reading experiment to investigate RNP in BP. The authors observed coreferential processing of repeated names, overt pronouns, and anaphoric null pronouns. The results revealed that the processing cost for overt pronouns related to subject antecedents was significantly higher than for repeated names, but these, in turn, were more costly than null pronouns, "but the difference between null pronouns and anaphoric repeated names with a subject antecedent was not significant, suggesting that RNP in BP may be weaker than OPP" (ALMOR et al., 2017, p. 109). Based on these results, the authors confirmed the existence of RNP in BP, considering the penalty of repeated nouns compared to null pronouns, according to a parameter proposed by Gerlomini-Lezama (2008, 2010), although the difference among the reading times of these referential expressions, which referred to a subject antecedent, were not significant.

Given the above, there are differences in the experimental stimuli and several other factors, both linguistic and methodological, among the studies that focused on RNP in BP with diverging results. Thus, we conducted two experiments using the self-paced reading technique, which will be addressed below, to explore some of these factors: anaphor type, number of potential human antecedents, and segmentation type, and reading time with the purpose of clarifying the divergences among the previously mentioned studies.

[9] The results of the first experiment in the study by Lima (2015) showed that the processing of repeated names was more costly than null pronouns. However, the author did not interpret the RNP as proposed by Gerlomini-Lezama (2008, 2010), who understands this effect as the penalty of repeated names in relation to any other anaphoric forms.

2. Experiments

With the purpose of analyzing the anaphoric processing of pronouns and repeated names for coreference resolution, focusing on the repeated-name penalty effect in Brazilian Portuguese, we conducted two experiments using the self-paced reading technique and manipulated and isolated/controlled some of the factors that seemed to be relevant to this investigation.

2.1 Experiment 1

One of the differences among the studies that have focused on the RNP in Brazilian Portuguese (BP) is the number of human referents/antecedents introduced in the non-critical sentence. While the studies that reported the occurrence of this effect usually included only one antecedent in the initial sentences, the studies that argued for the non-existence of this penalty used initial sentences containing two antecedents. This difference was observed and proposed by Maia (2013), but it was not experimentally manipulated. Therefore, the purpose of the first experiment was to investigate the anaphoric processing of pronoun and repeated-name anaphors under experimental conditions containing one or two antecedents. The hypothesis is that we would find the RNP in conditions with one antecedent, but not in conditions with two antecedents. This prediction arises from the understanding that there is a relationship between anaphoric processing and working memory based on the Informational Load Hypothesis (ALMOR, 1999). Assuming that there would be an increase in focus on the antecedent in conditions when the initial sentences have a single antecedent in subject position, thus facilitating the occurrence of RNP, conditions with initial sentences with two antecedents, even though with different genders, could cause competition in anaphoric processing, making it difficult for this penalty to occur.

2.1.1 Method

a) Participants

Thirty-six undergraduate and graduate students (12 men) from the Federal University of Paraíba participated in this experiment. They were between 17 to 43 years of age (mean age = 24), and they were all native speakers of Brazilian Portuguese.

b) Material

The material consisted of 4 experimental sets (each set with 16 experimental sentences and 32 filler sentences). The experimental sentences were constructed by combining the independent variables: Anaphor form (overt pronoun/OP, repeated name/RN) and number of antecedents (one antecedent/1A, two antecedents/2A). All the linguistic stimuli were divided into 9 segments, followed by a probe word and a yes/no question.

The within-subjects, randomized 2x2 Latin square design allowed each participant to be exposed to all conditions and all types of structures, but without repetition of the experimental passages.

The experimental passages consisted of two juxtaposed sentences. The first sentence either introduced one antecedent in subject position or two antecedents - proper names with different genders - one in subject position and the other in object position. The second sentence began with an anaphoric element (repeated name or overt pronoun) in subject position, which referred to the antecedent in subject position.

Table 1 shows examples of the experimental sentences in each condition tested in this experiment.

Table 1 – Examples of experimental passages in Experiment 1.

Condition	**Experimental stimuli**								
	S1	S2	S3	S4	S5	S6	S7	S8	S9
2A + AP	**Ana$_i$**	substituiu	**Nei$_j$**	na gincana	da escola.	**Ela$_i$**	hoje	joga	na seleção.
	Ana$_i$	*replaced*	***Nei$_j$***	*in the school*	*games*.	**She$_i$**	*now*	*plays*	*for the national team.*
2A + RN	**Ana$_i$**	substituiu	**Nei$_j$**	na gincana	da escola.	**Ana$_i$**	hoje	joga	na seleção.
	Ana$_i$	*replaced*	***Nei$_j$***	*in the school*	*games*.	***Ana$_i$***	*now*	*plays*	*for the national team*
1A + AP	**Ana$_i$**	participou	bem	da gincana	da escola.	**Ela$_i$**	hoje	joga	na seleção.
	Ana$_i$	*participated*	*well*	*in the school*	*games*.	**She$_i$**	*Now*	*plays*	*for the national team*
1A + RN	**Ana$_i$**	participou	bem	da gincana	da escola.	**Ana$_i$**	hoje	joga	na seleção.
	Ana$_i$	*participated*	*well*	*in the school*	*games*.	***Ana$_i$***	*now*	*plays*	*for the national team*
	Probe word: ANA								

c) Procedure

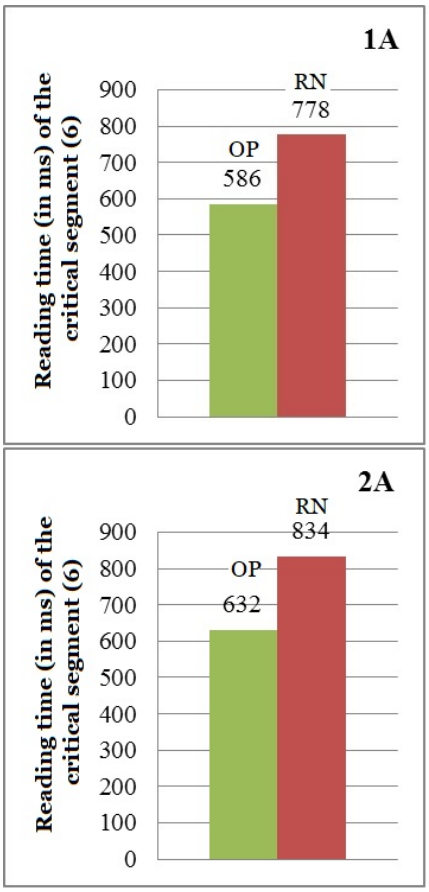
The self-paced reading experiment task was carried out in a sound-proof room (LAPROL - Language Processing Laboratory - Federal University of Paraíba) on an Apple MacBook computer using the PsyScope software (COHEN, MacWHINNEY, FLATT & PROVOST, 1993).

Each participant was first orally instructed by the researcher and the instructions were also displayed on the computer screen. Before the experimental task began, participants would take part in a training practice session to become familiar with the task. After understanding the instructions, the researcher left the room and the participant, alone in the room, started the task by pressing a key (l) and start reading the segments of the sentences on the computer screen in a non-cumulative way. Then, when a probe word was shown on the screen at the end of each sentence, the participant should answer whether it had appeared in the sentences by pressing one of the keys identified with the words “yes” and “no”. With this probe recognition task, we aimed to control the attention and understanding of the participants regarding the sentences they had read. The experimental task was conducted in a single session that lasted approximately 6 minutes, and the participants reported no difficulties performing it.

2.1.2 Results

First, we used the Boxplot tool to identify the outliers of the critical segment data (6). Overall, outliers (7.4%) were removed from all conditions. Graph 1 shows the mean reading times in milliseconds (ms) of the critical segment in the 4 conditions after removing the outliers.

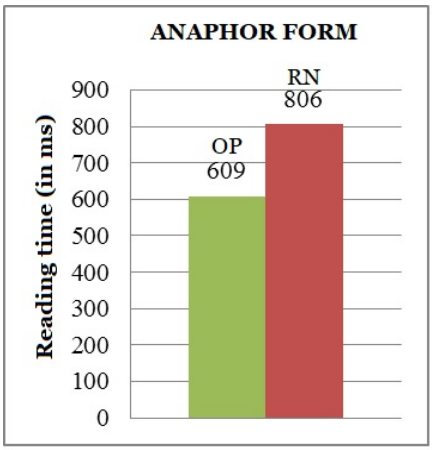
Graph 1 – Mean reading times of the critical segment for all conditions in Experiment 1.



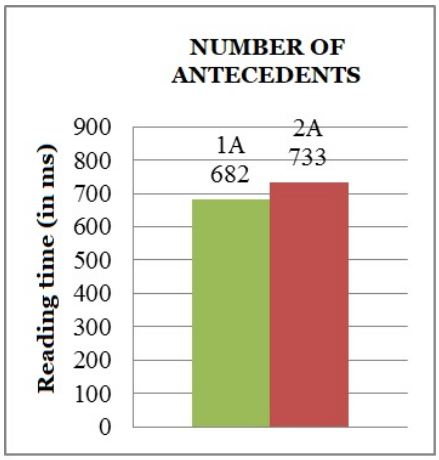
The test revealed a significant effect of the Anaphor Form ($F_1(1,31) = 55,35$; $p<0,05$; $F_2(1,14) = 35,17$; $p<0,05$), showing that the processing cost of repeated names was higher than overt pronouns (see Graph 2). There was also a significant effect for the Number of Antecedents per participants ($F_1(1,31) = 3,74$; $p<0,05$; $F_2(1,14) = 0,68$; $p = 0,40$), indicating that conditions with a single antecedent were easier to process than those with two antecedents (see Graph 3).

There was no significant effect of interaction between Anaphor Form and Number of Antecedents ($F_1(2,31) = 0,038$; $p<0,84$; $F_2(2,14) = 0,01$; $p = 0,90$), suggesting that one variable did not directly affect the other.

Graph 2 –Effect of the Anaphor Form in Experiment 1.



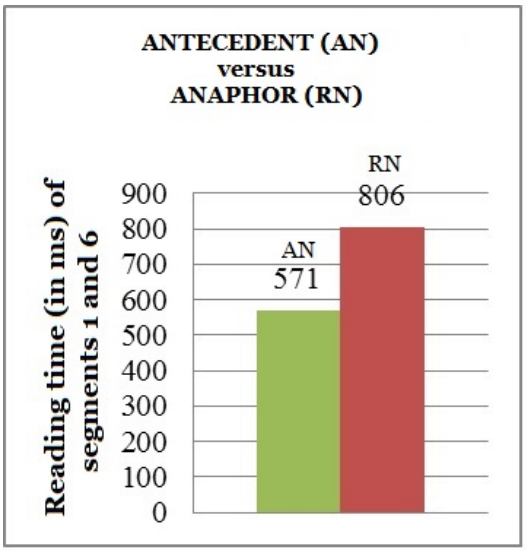
Graph 3 – Effect of the Number of Antecedents in Experiment 1.



Regarding the controversy related to the segmentation type/reading time of the critical segment corresponding only to the anaphor, MAIA (2013, p. 88) considers it "an effect of the lexical level at most". Thus, we chose to investigate and compare the reading times of the name antecedent (segment 1) with that of the repeated-name anaphor (segment 6) to confirm if anaphoric coreference was established. The *t*-test ($t(31) = -8.990$; $p<0.05$) revealed that the reading time of the referential expression was significantly lower than that of the antecedent (see Graph 4). Since the antecedent and anaphor corresponded to the same name, if only

considering lexical access, both elements should have the same processing cost. Thus, we predict that the higher processing cost of the anaphoric element in relation to the antecedent reflects anaphoric coreference resolution. This result reinforces the study by Vasconcelos (2012) who also found a lower reading times for the name antecedent when compared to the repeated-name anaphor.

Graph 4 –Antecedent versus Anaphor effect.



We then analyzed the data regarding the number of "yes" answers (correct answers) and "no" (incorrect answers) given to the probe word. The predominance of "yes" answers evidenced that the participants carefully read the linguistic stimuli, as shown in Table 2. The table shows the absolute and relative values of the responses to the probe word.

Table 1 – Answers to the probe task in Experiment 1.

Answers	**Conditions**			
	1A + OP	1A + RN	2A + OP	2A + RN
YES	139 (97%)	144 (100%)	133 (92%)	144 (100%)
NO	5 (3%)	0 (0%)	11 (8%)	0 (0%)
Total	144 (100%)	144 (100%)	144 (100%)	144 (100%)

2.1.3 Discussion

The results showed that the reading times in conditions with initial sentences with one antecedent (1A) were lower than those that introduced two antecedents (2A). In addition, the processing cost of the repeated-name anaphor (RN) was higher than overt pronouns (OP) in both conditions (1A and 2A), thus confirming the repeated-name penalty in the linguistic stimuli investigated in Experiment 1.

Thus, the results do not fully corroborate our initial hypothesis that predicted a RNP for the conditions with initial sentences containing a single antecedent. We believe that the emphasis given to an antecedent entity, without a competitor, based on the principles of the Information Load Hypothesis (ALMOR, 1999), could overload the working memory, contributing to the occurrence of this penalty, but it would not predict a RNP for conditions with two antecedent entities because, although with different genders, they could cause some type of competition, making it difficult for this effect to occur. However, the results showed that conditions containing two antecedents have a higher processing cost than those containing only one antecedent, confirming that the number of antecedents may be one of the factors playing a role in coreferential processing in the stimuli used in this study.

The results of this experiment also showed that the reading times of referential expressions with a repeated name were significantly higher than the reading times of the same name used as an antecedent. This higher cost in the reading time of the repeated name points to coreference resolution and, therefore, the controversy regarding the reading time of the anaphor alone could only be related to lexical access, corroborating a previous study that found results along the same lines (VASCONCELOS, 2012).

As the results revealed the occurrence of the RNP for both one and two antecedents, suggesting that the number of antecedents alone does not seem to be so relevant in relation to this effect, we observed other factors that, together with the factor number of antecedents, could have influenced these results, such as Segmentation Type/Reading Time. Based on this, we decided to conduct Experiment 1 again focusing (isolating/controlling) on this methodological factor, which is another distinction among the studies that have focused on the RNP in BP.

2.2 Experiment 2

Reflecting on the results of the first experiment led us to consider that the factor number of antecedents can influence the occurrence of RNP when combined with other factors, such as a Segmentation Type/Reading Time.

According to Nair and Almor (2006), the RNP is a phenomenon concerning discourse integrative processes that do not occur immediately upon encountering an anaphor. Thus, the most appropriate way of presenting linguistic stimuli would be sentence by sentence and not word by word. On the other hand, Ernst (1999) considers that the division and stimuli presentation in several segments or sentence segments in a self-paced reading task, by isolating and measuring the reading time of the anaphor alone, provides a better measure for anaphoric resolution because it allows a more online analysis. Furthermore, studies using the eye-tracking paradigm have evidenced the existence of the RNP in the first fixation duration in a spillover effect (LEITÃO, RIBEIRO & MAIA, 2012). In the region immediately after the anaphor, they found that the RNP can be observed in an anaphor alone or even during first reading, corroborating Ernst (2007), and in the initial processing of the discourse and not only in the late discursive integrative phase, as claimed by Nair and Almor (2006).

Aware of this issue, Gondim and Leitão (2012) replicated an experiment by Queiroz and Leitão (2008) modifying the segmentation type of the linguistic stimuli. Queiroz and Leitão (2008) divided the sentences into several segments (9 in total) and measured the reading time of the segment containing the anaphor (overt pronoun or repeated name), while Gondim and Leitão (2012) divided the linguistic stimuli into only two segments (two coordinate sentences) and measured the time of the whole critical sentence. The results found by Gondim and Leitão (2012) revealed that, despite using a different methodology, RNP was observed.

Although Gondim and Leitão (2012) investigated this methodological factor regarding the Segmentation Type/Reading Time, they used different experimental stimuli from the one we used in Experiment 1. The experimental items in the study by Gondim and Leitão (2012) consisted of coordinate sentences with one antecedent, while the stimuli in our first experiment consisted of juxtaposed sentences containing one or two antecedents. Based on this,

we aimed to investigate if the Segmentation Type/Reading Time could have influenced the RNP results, as investigated in the first experiment.

From a multifactorial perspective, we understand that several factors are integrated in discourse processing when measuring the reading time of a whole sentence and conditions with two antecedents can result in competition during anaphoric resolution, decreasing discourse salience of the antecedent and anaphor and making it difficult to observe the RNP. Therefore, our hypothesis for Experiment 2 was that the RNP would not be observed in conditions with 2 antecedents when measuring the entire critical sentence, but it would occur in conditions with 1 antecedent when measuring the entire critical sentence.

2.2.1 Method

a) Participants

The participants of this experiment were 32 undergraduate and graduate students (19 men) from the Federal University of Paraíba. They were between 18 to 32 years of age (mean age = 22.9), and they were all native speakers of Brazilian Portuguese.

b) Material

Experiment 2 used the same stimuli from Experiment 1 (4 sets containing 16 experimental passages and 32 fillers), slightly changing some of the linguistic stimuli to ensure greater control of the size of the critical sentences and segmentation type (the linguistic stimuli in this second experiment were divided only into 2 segments and those from the first experiment were divided into 9 segments).

Table 3 shows examples of experimental passages in each condition used in the second experiment.

Table 2 –Examples of experimental sentences in Experiment 2

Conditions	**Experimental passages**	
	S1 (Initial Sentence)	S2 (Critical Sentence)
2A + OP	**Bel$_i$** marcou **Ian$_j$** na foto do Facebook. ***Bel$_i$*** *tagged* ***Ian$_j$*** *in a photo on Facebook.*	**Ela$_i$** ainda colocou um comentário. ***She$_i$*** *even posted a comment.*
2A + RN	**Bel$_i$** marcou **Ian$_j$** na foto do Facebook. ***Bel$_i$*** *tagged* ***Ian$_j$*** *in a photo on Facebook.*	**Bel$_i$** ainda colocou um comentário. ***Bel$_i$*** *even posted a comment.*
1A + OP	**Bel$_i$** postou logo a foto no Facebook. ***Bel$_i$*** *soon posted a photo on Facebook.*	**Ela$_i$** ainda colocou um comentário. ***She$_i$*** *even posted a comment.*
1A + RN	**Bel$_i$** postou logo a foto no Facebook. ***Bel$_i$*** *soon posted a photo on Facebook.*	**Bel$_i$** ainda colocou um comentário. ***Bel$_i$*** *even posted a comment.*
	Probe word: BEL	

c) Procedure

We used the same experimental design from Experiment 1.

The self-paced reading technique was used, but with one difference from the first experiment: stimuli presentation was sentence by sentence to measure the reading time of the entire sentence.

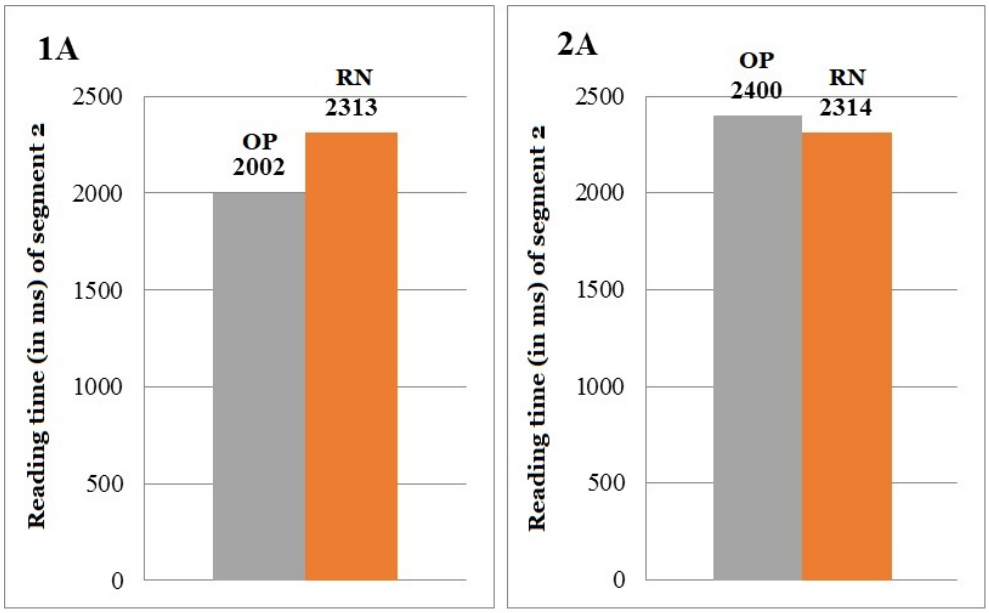
After instructions and a practice session, the participants started the experimental session by pressing the "L" key on the keyboard. Once the participants had pressed the key, they were presented the first sentence on the screen. Participants were instructed to press "L" key again to advance to the second sentence (the critical sentence). Soon after both experimental and filler passages, a probe word would appear. Participants were required to press keys labeled “yes” or “no” to answer if the probe word was in the sentences. The purpose of this task was to control the attention of the participants.

Each session lasted approximately of 6 minutes per participant. At the end, participants reported that did not have any difficulties performing it.

2.2.2 Results

As in the previous experiment, we used the Boxplot in the critical segment data (Segment 2) to identify the outliers. In total, 4.8% of outliers were removed. The mean reading times of the conditions resulting from this procedure are shown in Graph 5.

Graph 5 – Mean reading times of the critical segment for all conditions in Experiment 2.

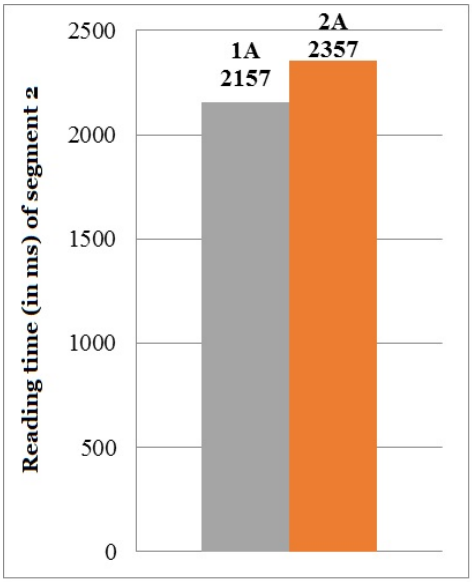


The mean reading times in the four conditions were analyzed using the statistical test ANOVA considering each participant and the items.

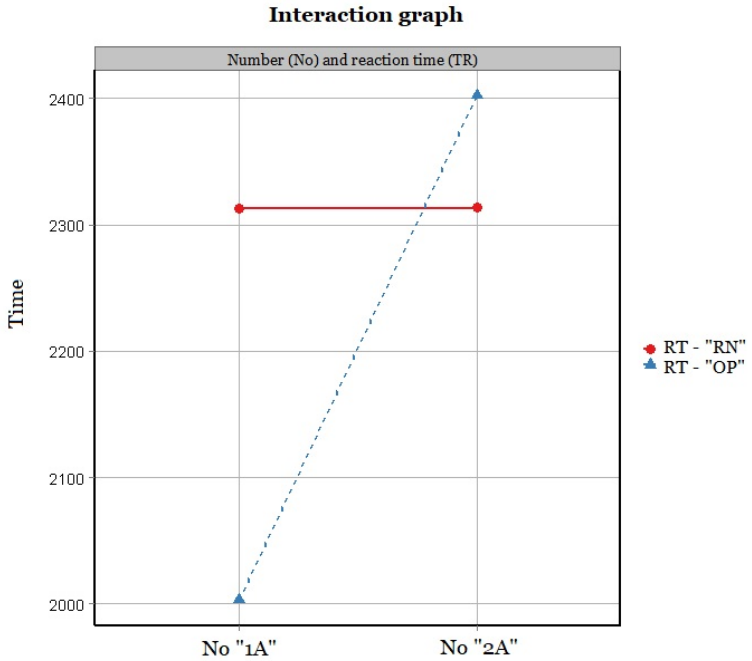
The ANOVA test revealed a significant effect for Number of Antecedents ($F_1(1,31) = 11,608$; $p<0,05$; $F_2(1,14) = 5,09$; $p<0,05$), showing that conditions with one antecedent were read faster than those with two antecedents (Graph 6), but there was no significant effect for the Anaphor Form ($F_1(1,31) = 0,818$ e $p= 0,36$; $F_2(1,14) = 1,56$; $p= 0,21$), indicating that there was no difference between the processing cost of overt pronouns and repeated names. There was also a significant interaction effect between Type of Sentence and Anaphor Form ($F_1(2,31) = 4, 970$; $p<0,05$; $F_2(2,14) = 5,05$; $p<0,05$), suggesting that in conditions with one antecedent, repeated-name anaphors had a higher processing cost than overt pronouns. In conditions with two antecedents the inverse occurred, as shown in Graph 7, which was also observed by the *t*-

test between conditions 1A OP (2002 ms) and 1A RN (2313 ms) (Graph 5), ($t(31) = 3.96$; $p<0.05$).

Graph 6 – Effect of the factor Number of Antecedents in the experiment.



Graph 7 - Interaction effect among variables in Experiment 2.



Regarding the accuracy of responses to the probe-word task at the end of each sentence, we found a predominance of "yes" answers (correct answers), indicating that

participants performed the experiment with attention. The absolute and percentage values of "yes" and "no" answers for each experimental condition are shown in Table 4.

Table 3 – Response values for the probe-word task in Experiment 2.

Answers	**Conditions**			
	1A + OP	1A + RN	2A + OP	2A + RN
YES	123 (96%)	126 (98%)	121 (95%)	127 (99%)
NO	5 (4%)	2 (2%)	7 (5%)	1 (1%)
Total	128 (100%)	128 (100%)	128 (100%)	128 (100%)

2.2.3 Discussion

The effect of the factor Number of Antecedents once again indicated that this factor seems to be relevant in coreferential processing. On the other hand, the factor Segmentation Type/Reading Time, which was controlled in this experiment, influenced the RNP, since this penalty only occurred in conditions with one antecedent, while in the first experiment it occurred in conditions with one and two antecedents. This finding reinforces the assumption of Ernst (2007) that the measurement of the anaphor alone provides a better measure of anaphoric coreference because it allows a more online analysis. These findings corroborate studies that have also found this effect when observing the initial processing time even without segmentation (LEITÃO, RIBEIRO and MAIA, 2012). In addition, this result shows that the divergences among the results found in BP may have occurred due to the difference in the number of antecedents as well as how the reading time was measured. This may have occurred because measuring the reading time in conditions with 2 potential antecedents, in which we found no effect, is similar to the experimental settings of studies that did not observe the RNP. At the same time, conditions with 1 antecedent are similar to the studies that observed the RNP, either by measuring the entire sentence (GODIM and LEITÃO, 2012) or by measuring the anaphor alone, as in Experiment 1.

3 General discussion

This paper aimed to investigate coreferential processing of overt pronoun and repeated names anaphors focusing on the RNP effect in Brazilian Portuguese (BP).

Previous studies conducted in BP that focused on the RNP based on the parameters proposed by Gordon et al. (1993) presented diverging results regarding the existence of the penalty in BP. While Leitão and colleagues have confirmed the existence of this effect in BP (LEITÃO, 2005; QUEIROZ & LEITÃO, 2008; LEITÃO & SIMÕES, 2011; GONDIM & LEITÃO, 2012; LEITÃO, RIBEIRO & MAIA, 2012; LIMA, 2014), recent studies have not found the occurrence of this penalty in Brazilian Portuguese (MAIA & CUNHA LIMA, 2011, 2012; MAIA, 2013; LIMA, 2015). Therefore, our purpose was to understand the origin of these divergences among the studies and investigate what could be the causing the controversial RNP results in BP. Thus, we conducted two self-paced reading tasks and manipulated, controlled, and related factors to observe the influence they could have on the RNP effect.

In Experiment 1, the anaphor form (repeated name or overt pronoun) and number of antecedents (initial sentences with one antecedent/1A or two antecedents/2A) were manipulated. The results showed a significant effect for the two manipulated variables. The reading times in conditions with 1A were lower than in conditions with 2A and the same occurred in conditions with overt pronouns when compared to repeated names in both conditions. Therefore, the RNP was found both in conditions with 1A as well as in conditions with 2A. Thus, our hypothesis was not confirmed, since the RNP was predicted only for conditions with 1A, considering that there is a relationship between coreferential processing and the working memory, based on the Informational Load Hypothesis (ALMOR, 1999), and assuming that the existence of only one antecedent (without competition) in discourse would facilitate the occurrence of the RNP. On the other hand, conditions with two antecedents could cause competition in coreference resolution, making it difficult for this penalty to occur. These results were interpreted as indicating that the number of antecedents seems to be relevant in coreferential processing, but it was not a determining factor regarding the RNP effect.

The results from the first experiment led us to reflect if the relationship between the factors Number of Antecedents and Segmentation Type/Reading Time may have influenced the results on the RNP. Based on this, we conducted a new round of Experiment 1 and isolated/controlled the factors Segmentation Type/Reading Time to investigate whether the Number of Antecedents together with other factors could interfere in the RNP results. In Experiment 2, instead of dividing the experimental passages into several segments and measuring the reading time of the anaphor alone, as we had done in the previous experiment, we divided the stimuli into two segments (two sentences) and measured the reading times of the entire critical sentence that contained the anaphor. The results of the second experiment elicited a significant effect for the Number of Antecedents, confirming the findings of the first experiment that conditions with 2A have higher processing costs than conditions with 1A. A significant interaction effect was also found, showing that the processing cost of overt pronouns was lower than repeated names in conditions with 1A (RNP), while the processing cost of repeated names was lower than overt pronouns in conditions with 2A. These results confirmed our hypothesis that the RNP would be observed in conditions with 1A and juxtaposed sentences when measuring the entire critical sentence because, although there were two factors (Sentence Type, Segmentation Type/Reading Time) that could hinder the occurrence of RNP, a single antecedent in the working memory would facilitate the occurrence of this effect. In view of these results, we understand that the effect of the factor Number of Antecedents once again indicated that this factor plays a role in coreference, and the interaction effect suggests that this factor seems to influence the results on the RNP. The factor Segmentation Type/Reading Time seemed to influence the RNP, since in Experiment 2 this penalty only occurred in conditions with 1A, while in the first experiment it occurred in conditions with both 1A and 2A, which is in agreement with studies that have also found this effect when observing the anaphor alone (ERNST, 1999; LEITÃO, RIBEIRO & MAIA, 2012).

The results of the two experiments seem to reinforce our general hypothesis that the RNP effect is multifactorial, that is, there are several factors that may be acting together and contributing or not to the occurrence of the RNP in coreference resolution in BP. Among the manipulated and controlled factors in the experiments, the factor Segmentation Type/Reading time influenced the RNP effect. In the experiments that measured the entire critical sentence, the RNP was only found conditions with one antecedent. In Experiment 1, though, which

measured the reading time of the anaphor alone, the penalty was found in both conditions with 1A and 2A.

However, there is controversy regarding the measurement of the anaphor alone. Maia (2013) considers that measuring the reading time of the anaphor alone only indicates lexical processing. In view of this, we compared the reading times of the antecedent with repeated-name anaphor in Experiments 1 to analyze if the anaphoric coreference was established. The results showed that the reading time of the anaphor was significantly higher than the reading time of the antecedent, thus reflecting the processing cost of anaphoric resolution. This may have occurred because since the antecedent and anaphor were established through the same name, they should have been processed in the same way if only lexical access was considered. These results corroborate the study by Vasconcelos (2012) that also found a higher reading time for the anaphoric element than for the antecedent.

These results indicate that two of the possibilities raised by Maia (2013) can help to explain the divergent results in BP. The first one is related to the division of experimental stimuli into several segments in self-paced reading tasks and measuring the reading time of the referential expression alone. This seems to provide a better measurement of anaphoric resolution, as argued by Ernst (2007), making it possible to observe RNP online or during initial processing, regardless of the number of antecedents. This penalty, as argued by Nair and Almor (2006), occurs only in a relatively late discourse integrative process and other studies have found RNP when observing the anaphors alone (ERNST, 2007; LEITÃO, RIBEIRO and MAIA, 2012). The second possibility is related to the existence of one or two potential human antecedents that seem to influence discourse salience and the occurrence or not of the RNP effect. When there is only one antecedent, discourse salience of the antecedent increases and this means that even when measuring the reading of the entire sentence, the RNP will be observed. In contrast, when there are two antecedents, in which RNP is not observed, there is interaction between anaphor form and number of potential antecedents.

In view of the above, we reviewed previous research conducted in BP, based on the assumptions of Gordon et al (1993) on RNP, and we sought to understand which factors could be causing the diverging results. We observed that the Segmentation Type/Reading Time is one of the differences among these studies and our findings suggest that this factor as well as the number of potential antecedents could have a strong influence on the RNP effect. Thus, we understand that these factors could explain the diverging results among the studies on RNP in BP. However, there are other differences and factors in the experimental stimuli used

in these studies that were not addressed in this paper, such as syntactic parallelism, presence or absence of sentence connectors, and null pronouns compared with overt pronouns and repeated names, which should be further investigated in other studies from a multifactorial perspective.

4 Final remarks

The main objective of this study was to analyze the coreferential processing of overt pronoun and repeated name anaphors in Brazilian Portuguese (BP), focusing on the effect known as the Repeated-Name Penalty (RNP), under the hypothesis that this penalty has a multifactorial effect that occurs due to several factors.

In the two experiments conducted in this research, several linguistic and methodological factors were manipulated and controlled to observe if they had any influence on the occurrence of the RNP. The results indicated that the relationship between the factors seems to interfere in the occurrence or not of the RNP, corroborating our general hypothesis that the RNP has a multifactorial effect.

Based on the results obtained in this study, we have summarized, as shown in Table 5, the main factors in the two experiments, in addition to the types of sentences that will be investigated in the future and listed the most relevant findings of studies on RNP in BP as well as the penalties observed in other languages. The table also contains information on the experimental technique and the methodology used and whether the RNP was observed.

Table 4 – Main factors investigated in this research related to other studies on coreference in the literature.

	STUDY	EXPERIMENTAL TECHNIQUE	FACTORS			RNP
			Reading time of referential expression	Number of antecedents	Sentence type	
EXPERIMENTS IN THIS RESEARCH	Exp. 1	self-paced reading	anaphor	1A or 2A	juxtaposed	yes (for 1A and 2A)
	Exp. 2	self-paced reading	whole sentence	1A or 2A	juxtaposed	yes (for 1A)
OTHER STUDIES IN BP	Gondim & Leitão (2012)	self-paced reading	whole sentence	1A	coordenate	yes
	Leitão, Ribeiro & Maia (2012)	reading during eye tracking	whole sentence	1A	coordenate	yes
	Maia & Cunha Lima (2011)	self-paced reading	whole sentence	2A	juxtaposed	no
	Maia & Cunha Lima (2012)	reading during eye tracking	whole sentence	2A	juxtaposed	no
	Maia (2013)	self-paced reading (experiments 1, 2 and 3) and reading during eye tracking (experiment 5)	whole sentence	2A	juxtaposed	no
STUDIES IN OTHER LANGUAGES	Gordon et al. (1993) (English)	self-paced reading	sentence split in 2 or 3 segments each with 2-6 words	1A	juxtaposed	yes
	Gordon & Chan (1995) (English)	self-paced reading	whole sentence	2A (Exp. 1 to 3) 1A (Exp. 4)	juxtaposed	yes
	Chambers & Smyth (1998) (English)	self-paced reading	whole sentence	2A	juxtaposed	yes
	Kennison e Gordon (1997) (English)	reading during eye tracking	whole sentence	2A	juxtaposed	yes
	Almor (1999) (English)	self-paced reading	anaphor	2A	cleft or pseudo-cleft	yes

	Yang et al (1999) (Mandarin Chinese)	self-paced reading	whole sentence	2A	juxtaposed	yes
	Ernst (2007) (French)	self-paced reading	anaphor	2A	juxtaposed	yes
	Egusquiza et al. (2016)	self-paced reading	anaphor	2A	juxtaposed	yes
	Shoji et al. (2017)	self-paced reading	whole sentence	2A	juxtaposed	yes

We included the factors Segmentation Type/Reading Time and Number of Antecedents because they were the most significant findings in this study and in other studies in the literature. As shown in Table 5, the RNP was found in all studies, both in BP and in other languages, when the anaphor was analyzed alone, irrespective of the number of antecedents and type of sentence. However, among the studies in BP that measured the reading time of the entire critical sentence, the RNP was found only in the following studies: (a) the second experiment of this research in conditions with juxtaposed sentences and only one antecedent; (b) in conditions with coordinate sentences and one antecedent (Gondim and Leitão, 2012); (c) in conditions similar to those used by Gondim and Leitão (2012) but using a different experimental technique (Leitão, Ribeiro and Maia, 2012). Therefore, it may be concluded that the Segmentation Type/Reading Time and number of potential human antecedents may be two of the possible answers for the differences among the studies conducted in BP. Anyhow, the RNP in BP between overt pronouns and repeated names, depending on the factors observed, can be confirmed.

It is worth mentioning that if we consider studies in Spanish, despite the differences, similar results were found. In the study by Gelormini-Lezama (2010) with Spanish speakers from Argentina, in which the entire sentence was measured, they did not find the RNP between overt pronouns and repeated names. However, in the study by Egusquiza et al. (2016) with European Spanish speakers, in which the reading time of referential expression was measured, the RNP was observed even with two potential antecedents, both in subject and object positions. Nevertheless, the RNP was found in studies conducted in English both when measuring the reading time of referential expression alone and when measuring the reading time of the entire critical sentence. In addition, the RNP has been observed in most studies only in subject position. These differences among the RNP results in different languages as well as the scarce reports on the factors addressed suggest that the RNP should

be investigated in more languages, from a multifactorial perspective, to obtain comparative data for better understanding the RNP phenomenon.

Thus, we hope this study will contribute to further the discussions on anaphoric processing, particularly those that concern the theoretical and methodological aspects regarding the RNP effect in BP.

REFERENCES

ALMOR, A. Noun-phrase anaphora and focus: the informational load hypothesis. *Psychological Review*. vol. 106, No 4, 748-765, 1999.

ALMOR, A.; MAIA, J. C.; CUNHA LIMA, M. L.; VERNICE. M.; GELORMINI-LEZAMA, C. Language processing, acceptability, and statistical distribution: a study of null and overt subjects in Brazilian Portuguese. *Journal of Memory and Language*, 92, 2017.

BARBOSA, M. A.; GONDIM, E. V. A. C.; LIMA, J. N. Influência do paralelismo estrutural na correferência de pronomes e nomes repetidos. In: *I Workshop Pronomes: Morfossintaxe, Semântica e Processamento*, ILUFBA - Instituto de Letras da UFBA, Salvador, 2016.

CHAMBERS, C.; SMYTH, R. (1998). Structural parallelism and discourse coherence: a test of centering theory. *Journal of Memory and Language*, 1998.

COHEN, J. D.; MacWHINNEY; B., FLATT, M., PROVOST, S. Psyscope: a new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments & Computers*, v. 25, n. 2, 1993.

ERNST, E. *Le traitement en temps réel de l'anaphore pronominale dans le langage écrit - Développement normal et dysfonctionnements. Apports de la théorie du Centrage*. Thèse de doctorat. Paris: Université Paris V, 2007.

EGUSQUIZA, N.; NAVARRETE, E.; ZAWISZEWSKI, A. Antecedent frequency effects on anaphoric pronoun resolution: Evidence from Spanish. *Journal of Psycholinguistic Research*, v. 45, 2016.

GELORMINI-LEZAMA, C. *Processing repeated names, overt pronouns and null reference in Spanish*. Dissertação de Mestrado. University of South Carolina, Columbia, 2008.

GELORMINI-LEZAMA, C. *The overt pronoun penalty:* a processing delay in Spanish anaphora comprehension. Tese de Doutorado. University of South Carolina, 2010.

GONDIM, E. V. A. C.; LEITÃO, M. M. Processamento correferencial e penalidade do nome repetido: investigando distinções metodológicas. In: II Workshop em Processamento Anafórico, Fortaleza. *Caderno de Resumos do II Workshop em Processamento Anafórico*, 2012.

GORDON, P. C; CHAN, D. Pronouns, passives and discourse coherence. *Journal of Memory and Language*, [S.I.], v. 34, n. 2, 1995.

GORDON, P. C; GROSZ, B. J.; GILLION, L. A. Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, [S.I.], v. 17, n. 3, 1993.

GROSZ, B. J., JOSHI, A. K.; WEINSTEIN, S. Centering: A Framework for Modeling the Local Coherence of Discourse. *Computational Linguistics*. v. 21, n. 2, 1995.

KENNISON, S. M.; GORDON, P. C. Comprehending referential expressions during reading: evidence from eye tracking. *Discourse Processes*, v. 24, n. 2-3, 1997.

LEITÃO, M.M. *O processamento do objeto direto anafórico no português brasileiro*. Rio de Janeiro: UFRJ/ Faculdade de Letras. Tese de Doutorado em Linguística, 2005.

LEITÃO, M.M; RIBEIRO, A. J. C.; MAIA, M. Penalidade do nome repetido e rastreamento ocular em português brasileiro. *Revista LinguíStica / Revista do Programa de Pós-Graduação em Linguística da Universidade Federal do Rio de Janeiro*, v. 8, n. 2, 2012.

LEITÃO, M.M; SIMÕES, A. G. A influência da distância no processamento correferencial de pronomes e nomes repetidos em português brasileiro. *Veredas On-line*, 2011.

LIMA, A. H. V. *A influência da retomada e da distância sintática no processamento de pronomes plenos e nulos em português brasileiro*. Programa de Pós-Graduação em Linguística da Universidade Federal do Ceará. Dissertação de Mestrado, 2015.

LIMA, J. C. *Paralelismo e foco estrutural no processamento da correferência de pronomes e nomes repetidos*. Dissertação de Mestrado. UFPB, 2014.

MAIA, J. C.; CUNHA LIMA, M. L. O processamento de expressões correferenciais em português brasileiro: nomes repetidos, pronomes plenos e pronomes nulos. *Revista do GELNE (UFC)*, v. 13, n. 1/2, 2011.

MAIA, J. C.; CUNHA LIMA, M. L. Processamento correferencial de nomes e pronomes plenos em PB: evidências de rastreamento ocular. *ReVEL*, edição especial n. 6, 2012.

MAIA, J. C. *O processamento de expressões correferenciais em português*. UFMG/Programa de Pós-Graduação em Estudos Linguísticos. Dissertação de mestrado em Linguística, 2013.

NAIR, V. A.; ALMOR, A. Measuring referential processes in sentences and discourse. In: Poster presented at the *19th Annual CUNY Conference on Human Sentence Processing*, New York, NY, March 23-25, 2006.

QUEIROZ, K.L.; LEITÃO, M.M. Processamento do sujeito anafórico em português brasileiro**.** *Veredas On-line*, 2008.

SHOJI, S.; DUBINSKY, S.; ALMOR, A. The Repeated Name Penalty, the Overt Pronoun Penalty, and Topic in Japanese. *Journal of Psycholinguistic Research*, v. 46, 2017.

VASCONCELOS, M.L. *Processamento da correferência em pacientes com afasia de expressão*. Programa de Pós-Graduação em Linguística da Universidade Federal da Paraíba. Dissertação de Mestrado, 2012.

YANG, C. L.; GORDON, P. C.; HENDRICK, R.; WU, J. T. Comprehension of referring expressions in Chinese. *Language and Cognitive Processes*, v. 14, n. 5/6, 1999.