DIFFERENCES BETWEEN GRAMMATICAL GENDER AND SEMANTIC GENDER IN PRONOMINAL ANTECEDENT RETRIEVAL IN BRAZILIAN PORTUGUESE

AS DIFERENÇAS ENTRE GÊNERO GRAMATICAL E GÊNERO SEMÂNTICO NA RECUPERAÇÃO DE ANTECEDENTES PRONOMINAIS EM PORTUGUÊS BRASILEIRO

Michele Calil dos Santos Alves[*]
michelecalil@gmail.com

Coreference is a syntactic dependency in which pronouns are bound to previous referents. The aim of this research is to provide more information on how pronominal antecedents are retrieved from memory, and, more precisely, to clarify the role of gender cues in pronominal antecedent retrieval in Brazilian Portuguese, granted that its speakers are used to rely on agreement cues during its processing once it has visible morphology. The results of two eye-tracking experiments conducted with native speakers of Brazilian Portuguese demonstrated that both binding structural constraints and gender morphological cues are equally important in antecedent retrieval in memory at early stages of coreference processing. This is evidence that binding structural constraints do not work as an initial filter, blocking the influence of structurally unacceptable antecedent candidates. In addition, the results indicated that semantic gender and masculine gender seemed to weigh more in memory than grammatical gender and feminine gender since structurally unacceptable candidates carrying the former types of gender caused more interference effects.

Keywords: Coreference processing. Gender cues. Memory retrieval. Brazilian Portuguese.

A correferência é uma dependência sintática em que pronomes são ligados a referentes mencionados previamente. O objetivo deste trabalho é fornecer mais informações sobre como os antecedentes pronominais são recuperados na memória, e, mais precisamente, esclarecer o papel das pistas de gênero na recuperação de antecedentes pronominais em português brasileiro, dado que seus falantes estão costumados a depender das pistas de concordância para processar a língua, uma vez que esta possui morfologia visível. Os resultados de dois experimentos de rastreamento ocular realizados com falantes nativos de português brasileiro demonstraram que tanto as restrições estruturais quanto as pistas morfológicas de gênero são igualmente importantes na recuperação dos antecedentes na memória nos estágios iniciais do processamento da correferência. Isto é evidência a favor de que as restrições estruturais não funcionam como um filtro inicial, bloqueando as influências de candidatos a antecedentes estruturalmente inaceitáveis. Além disso, os resultados indicaram que o gênero semântico e o gênero masculino parecem ter um peso maior na memória do que o gênero gramatical e o gênero feminino, já que os candidatos a antecedentes inaceitáveis estruturalmente que carregavam os primeiros tipos de gênero causaram mais efeitos de interferência.

---

[*] Universidade Federal do Acre, Brasil. ORCID: https://orcid.org/0000-0002-9514-6795

Palavras-chave: Processamento da correferência. Pistas de gênero. Recuperação na memória. Português brasileiro.

•

1. Introduction

In order to process language in real time, previously interpreted information must be kept at least momentarily in memory so that integration with novel upcoming material can take place rapidly (Lewis, Vasishth & van Dyke 2006). This way, memory can be considered one of the key factors in processing long distance dependencies such as coreference, in which pronouns are bound to antecedents that occupy linearly distant positions in the discourse.

Among other cues, coreference can be influenced by salience of the discourse entities involved in the context, agreement relations between antecedents and pronouns, and binding structural constraints of Principle B, which posits that the relation between antecedents and pronouns cannot be local (Chomsky 1981). Previous research that has investigated how those three factors play a role in binding processing is very contradictory. On the one hand, it has been claimed that structurally unacceptable candidates, that is, local candidates, cannot initially influence binding processing even in cases in which they are salient discourse entities and agree with the pronouns (Nicol & Swinney 1989; Sturt 2003; Leitão, Peixoto & Santos 2008; among others). On the other hand, other research has shown that structural constraints can be fallible as apparently structurally unacceptable candidates can be initially considered as potential antecedents if they are salient entities that feature-match the anaphoric expressions (Badecker & Straub 2002; Patil, Vasishth & Lewis 2016).

Languages with limited overt morphology like English might not be the most appropriate to study gender agreement. By comparing overt agreement marking in English and in Brazilian Portuguese, one notices that unlike the former, the latter has redundant gender agreement marking in most determiners, nouns, and adjectives, for example. In these terms, the present study tried to control for the different types of features that may exist under the category of gender in a rich visible morphology language such as Brazilian Portuguese. This way, this study aims at verifying whether different types of gender conveyed by pronominal antecedent candidates would influence the way they would be retrieved from memory. Agreement features may be more helpful in pronominal antecedent retrieval due to the looseness of its constraints. In other words, pronominal binding constraints (Principle B of Binding Theory) only posits antecedents must not be local, which is not a quite restrictive constraint. Thus, morphological cues could be very helpful in pronominal antecedent retrieval.

This way, the present research would fill a gap in the literature as it will provide not only one more piece of evidence to the puzzle of binding processing, which has shown very contradictory results, but it would also provide evidences of the role of gender cues in pronominal binding in a language with redundant visible gender morphology like

Brazilian Portuguese. It would also be examined whether speakers of Brazilian Portuguese tend to initially consider structurally unacceptable candidates as potential antecedents despite the fact they violate binding constraints.

(1)

O <u>arquiteto</u>	agradeceu	o <u>engenheiro</u>	que indicou **ele** justamente
The <u>architect</u>[masc]	thanked	the <u>engineer</u>[masc]	who fairly recommended **him**
para um dos	cargos mais cobiçados do país.		
for one of the	most desirable jobs in the country.		

For instance, in Example 1, according to the pronominal binding structural constraints, the pronoun *ele*, ('him'), refers to *arquiteto* ('architect[masc]'); however, there is another antecedent candidate in the sentence that also agrees in gender with the pronoun, *engenheiro* ('engineer[masc]'). The question is whether structurally unacceptable antecedents such as *engenheiro* ('engineer[masc]') would influence antecedent retrievals in memory. Moreover, another question is whether different genders (masculine or feminine), or even whether different types of gender (semantic gender or grammatical gender) conveyed by structurally unacceptable antecedent candidates would be responsible for any differences in the how coreference is processed.

In Brazilian Portuguese, most nouns with semantic gender vary in gender through compositional/derivational processes, for example, arquiteto, 'male architect', *versus arquiteta*, 'female architect', or europ**eu**, 'male European', *versus europeia*, 'female European'. However, there are other different types of gender variation. There are nouns whose gender is syntactically/grammatically determined such as the epicenes. For example, *vítima* ('victim[fem]') is grammatically feminine, but it can refer to either a male or a female referent. Moreover, there are nouns with lexically determined gender variation since they carry no morphological cues to indicate gender, as for example, *mulher*, 'woman', *versus homem*, 'man'. Finally, a third type of nouns is the bigenders, which are gender ambiguous and dependent on context, as for example, *turista*, 'male or female tourist', or *estudante*, 'male or female student'. Some bigender nouns are stereotyped biased, for example *recepcionista*, 'receptionist', is feminine-biased while *surfista*, 'surfist', is masculine-biased. This way, taken into account the richness of gender variation in Brazilian Portuguese, do different types of gender have different weights in memory, that is, different prominence levels in memory?

The aim of this research is to provide more information on how pronominal antecedents are retrieved from memory, and more precisely to clarify the role of gender cues in pronominal antecedent retrieval when gender morphology is overt. Since Brazilian Portuguese is a language with overt morphology, speakers of this language are used to rely on agreement cues to process language. Thus, the first hypothesis is gender morphological cues play a great role in pronominal antecedent retrieval in Brazilian Portuguese. This way, structurally unacceptable antecedent candidates that agree in gender with the pronouns would be considered as potential candidates, despite the fact they violate pronominal binding constraints. The second hypothesis is related to the fact different gender features would be encoded/retrieved in memory with different weights

(Dillon, Mishler, Sloggett & Pillips 2013; van Dyke & McElree 2011). Thus, memory can be affected by the prominence of gender features.

In order to test the hypotheses, two eye-tracking experiments were conducted with native speakers of Brazilian Portuguese. The eye-tracking technique is suitable for our purposes as it enables the researcher to examine the temporal course of language processing, including very early processing measures.

2. Content-addressable memory model

Lewis *et al.* (2006) questioned the processes in which working memory retrieves previous interpreted information and the constraints that may exist on those processes. They proposed a new model capable of explaining the content-addressable memory mechanism. According to this model, prior information that was previously interpreted is retrieved by a parallel search based on a set of grammatical cues generated by a target. This set of retrieval cues consist of several types, including structural, morphological, semantic, and discursive (among others).

According to Lewis and Vasishth (2005) and Lewis *et al.* (2006), the parallel search in memory can be affected by similarity-based interference. Similarity-based interference occurs when the overlap between the items in memory and the retrieval cues increase, reducing the strength of association between the cue and the target item, as a great number of items will be associated with the cue. Consequently, memory failure rates increase, and distractors, that is, candidates that partially-match the cues, can sometimes be misretrieved.

The content addressable memory model can also be used to explain how pronouns retrieve their antecedents in memory. For instance, in Example 1, by the time the pronoun *ele*, 'him', is encountered, a group of grammatical cues is generated in order to retrieve the antecedent. The antecedent must not be local[1], and it must be masculine and singular. After that, there is a parallel search in memory and two candidates that are similar to the cues generated by the target are found: *arquiteto* ('architect[masc]') and *engenheiro* ('engineer[masc]'). The former candidate is a perfect match; however, although the latter candidate is only a partial-match (it is masculine, but it is local), it can interfere with memory retrieval, the so-called similarity-based interference effect. Candidates like *engenheiro* ('engineer[masc]') are called distractors according to the content addressable model and, according to this model, distractors such as *engenheiro* ('engineer[masc]') can sometimes be erroneously retrieved as antecedents as a result of a failure caused by similarity-based interference effects.

---

[1] It is important to mention that the status of the [-local] feature can be questioned, as it seems awkward that languages would have this feature specified for each item. However, it is assumed that it is a relational feature that is only specified in binding dependencies.

2.1. Extended content-addressable memory model

Engelmann, Jäger and Vasishth (2015) realized the literature shows a great variability of results on content-addressable memory, and some of them are not even predicted by the model. Thus, the authors proposed an extension to the classic content addressable memory model in order to better explain already published results. They reviewed 69 experiments on reflexive-antecedent and subject-verb dependencies and presented the results of a computational model.

Engelmann *et al.* (2015) stated similarity-based interference can cause elevated reading times, which is called "inhibitory interference". Based on Lewis and Vasishth (2005), they explained the inhibitory effect is motivated by a competition between the target and the distractor. Since the amount of activation associated with a retrieval cue is shared between all matching items, the presence of competitors in memory will reduce item activation. Since retrieval speed is a function of item activation, reduced activation due to cue-matching distractor culminates in longer retrieval latency as compared to a condition without a cue-matching distractor.

Moreover, Engelmann *et al.* (2015) pointed out that, according to Lewis and Vasishth (2005), the similarity-based interference increases the probability of erroneously retrieving the partial-matching distractor. These occasional misretrievals are predicted to cause incorrectly formed dependencies, affecting comprehension in the respective trials. In special occasions, misretrievals of the distractor can lead to an observed speed-up in reading times means. This is called "intrusion".

However, Engelmann *et al.* (2015) claimed when target and distractor do not overlap in the manipulated feature in the distractor-match condition, no similarity-based interference is predicted. Nevertheless, because both target and distractor partially-match the retrieval cues, the probability of erroneously retrieving the distractor is predicted to increase. This causes shorter retrieval latencies in the distractor-match conditions. This speed-up effect is called "facilitatory interference".

3. The role of structural constraints in binding processing

3.1. Evidences of initial infallibility of structural constraints in binding processing

Nicol and Swinney (1989) examined the reactivation of anaphoric antecedents. They found out that immediately after the reflexive expressions only the structurally appropriate antecedent was reactivated, while the other referents were not significantly reactivated. The results for pronouns were similar to the results of reflexive expressions. Thus, the authors concluded that the reactivation of prior referents is restricted by grammatical constraints. Nicol and Swinney (1989) explained that only when binding constraints do not constrain the list of potential antecedents to a single one; pragmatic and other sentence or discourse processing procedures would come into play, but only at a later point in processing.

Sturt (2003) was concerned about two questions: i) to what extent sentence processing is affected by ungrammatical antecedents; ii) to what extent do binding principles act like a filter on the final interpretation of a sentence. He conducted an eye-tracking study to investigate the influence of inaccessible antecedents in reflexive binding when they are put strongly into discourse focus. Stereotypical subjects were used in order not to expose participants to ungrammatical sentences. His results showed that binding constraints were applied extremely early (at First Fixation and First Pass reading times). First Fixation and First Pass reading times were faster when the gender of the reflexive matched the stereotype of the accessible antecedent than when it did not, but they did not differ reliably as a function of whether the inaccessible antecedent matched the reflexive. However, reliable influences of the inaccessible antecedent at late measures were found (Second Pass in the second area after the reflexive). There were longer Second Pass times when the inaccessible antecedent mismatched the reflexive than when it did not. The author concluded that antecedents that were not initially considered by the binding principles could affect processing at a later stage.

Leitão *et al.* (2008) investigated the relationship between pronominal binding principles and phi-features (gender, number, and animacy) in coreference processing in Brazilian Portuguese. In the first experiment, there were structurally unacceptable antecedents in the sentences, and the results showed that the spillovers (regions after the pronouns) had longer reading times when structurally unacceptable antecedents in the sentence feature-matched the pronoun. However, in the second experiment, there was a structurally unacceptable candidate available in a preamble. Unlike the first experiment, the results of the second experiment did not show any differences among the conditions, although the reading times at the pronoun region were faster when compared to the first experiment. The authors suggested that when there are no structurally acceptable antecedent candidates available, as in the first experiment, candidates that feature-match the pronouns could be considered as potential antecedents even if they violate Principle B. However, when there is a structurally acceptable antecedent available, as in the second experiment, the search of an antecedent ends faster and the structurally unacceptable candidates are not taken into account.

3.2 Evidence of initial fallibility of the structural constraints in binding processing

Badecker and Straub (2002) investigated whether the content of structurally inaccessible NPs would influence pronoun processing.

(2) Sample of the materials in Badecker and Straub (2002):
    a) multiple match: *John thought that Bill owed him another chance to solve the problem.*
    b) accessible match: *John thought that Beth owed him another chance to solve the problem.*
    c) inaccessible match: *Jane thought that Bill owed him another chance to solve the problem.*
    d) no-match: *Jane thought that Beth owed him another chance to solve the problem.*

They observed longer reading times in the no-match condition than in the accessible match condition. The results also show faster reading times when there was a structurally

accessible antecedent than when there was an inaccessible antecedent. There was no difference between the multiple match and the accessible-match conditions. The authors concluded that gender was automatically used to identify the referent of a pronoun, and that the structurally accessible antecedents were also rapidly accessed. In contrast, inaccessible candidates were not blocked for an initial candidate set, as they influenced the evaluation process as soon as the pronoun was encountered.

Badecker and Straub (2002) concluded that binding-theory principles do not function as initial filters as reading times were longer when the grammatically inaccessible NPs agreed in gender (and number) with the pronoun or reflexive expressions. The authors supported the *interactive-parallel-constraint model*: the initial candidate set is composed of the focused discourse entities that are compatible with the lexical properties of the referentially dependent expression, while the grammatical constraints on interpretation operate quickly and effectively in the process of selecting from among these options.

Patil *et al.* (2016) argued that reflexive binding might be a very informative phenomenon to understand the role that grammatical and non-grammatical constraints play in memory. The structural constraints of reflexive binding are relatively clear, that is, the antecedents must be local (Chomsky 1981). They conducted an eye-tracking experiment increasing the strength of the inaccessible subject. They used an object pronoun within a relative clause where the inaccessible antecedents were the subjects of the clause. They found a significant main effect of interference in First Pass and in First Pass Regression Probability. The authors concluded that non-structural cues such as gender are crucial for antecedent retrieval so that gender agreement features must be included in the set of initial retrieval cues. Moreover, it seems their results are inconsistent with strict syntactic constraints on antecedent retrieval, and it seems reflexive binding is not infallible at initial processing stages, as the majority of previous research has shown.

4. Is masculine a default gender?

Corbett (1991) argued there may be constructions in which the target has to agree with a controller that is not specified for gender, as an infinitive clause, or when a choice of gender would force greater specificity than is possible or desirable for the speaker. For example, speakers may desire to refer to a child but be unable to select a gender agreement based on sex. Many languages solve this problem by using the regular gender form, which is often called *neutral agreement form* or *default agreement form*. However, neuter may not be the unmarked gender since almost all nouns denoting humans are masculine or feminine. Thus, the choice of neutral agreement may be understood as the selection of the gender that is most appropriate in semantic terms.

Corbett (1991) explained that when one analyzes the gender used to refer to pairs or larger groups, it is possible to discover interesting semantics of gender in a given language. For instance, in French, *les Américains*, 'the Americans', is used to denote males or both males and females. This is one semantic justification for the use of a particular gender resolution for conjuncts of different genders. Thus, masculine plural can

denote semantic gender neutralization in French. The same would happen in Brazilian Portuguese. For examples, the masculine form in the plural in *os americanos*, 'the Americans', is used to refer to males or both males and females; on the contrary, the feminine gender in the plural as in *as americanas* can only refer to females. In other words, since masculine gender (in the singular or in the plural) is unmarked in languages such as Spanish, French, and Brazilian Portuguese, it works as default gender, that is, it neutralizes semantic gender conflicts in some contexts of use.

Casado, Palma and Paolieri (2017) studied the influence of the sex of the participants in gender reference. They used three kinds of tasks in their study: word repetition, lexical decision, and gender decision. The authors only used transparent gender marked nouns in their experiments, that is, masculine nouns ending in -o and feminine nouns ending in -a. According to Casado *et al.* (2017), both female and male participants included female and male representations when hearing semantically gendered masculine nouns, which is evidence in favor of the fact that masculine works as the generic or default gender in Spanish. In other words, when a male speaker uses a masculine gendered word, both male and female listeners would think on either the male or the female referents, but the strength of activation would be drawn to the male referent. And that would be easier for a male listener than for a female listener. However, when a male speaker uses a feminine gendered word, it would be hard for listeners to think on either male or female referents, especially if the listener is male.

5. Comparing semantic and grammatical gender

Vigliocco and Franck (1999) was interested in determining whether Italian and French agreement production would be sensitive to the distinction between grammatical and semantic gender. The authors hypothesized that for semantic gender there would be a match between the syntactic and conceptual gender, while for grammatical gender, there would be gender specified by syntactic features only. Thus, they predicted more errors of agreement for nouns with grammatical gender than for nouns with semantic gender. In other words, semantic information would help correct agreement since it provides redundant compatible information.

Vigliocco and Franck (1999) conducted a series of four experiments eliciting gender agreement errors between subjects and predicative adjectives. The participants were presented to an adjective and then a sentential fragment. They were instructed to repeat the fragment and complete it with the adjective informed. All fragments were composed by subject head noun followed by an embedded modifier prepositional phrase (local noun). As expected, the authors reported more gender agreement errors between subjects and predicative adjectives for subject head nouns with grammatical gender than with semantic gender in Italian and French. Less errors for nouns with semantic gender reveals redundant information ensures accuracy and allows a more efficient encoding.

6. The present study

Experiments 1a and 1b aimed at investigating how antecedent retrievals work in Brazilian Portuguese during the initial stages of coreference processing in order to verify whether structural constraints of Principle B would initially block the influence of structurally unacceptable antecedent candidates in memory. In order words, the main purpose of this study is to check whether there is initial fallibity of structural binding constraints in Brazilian Portuguese, which is a language with rich gender variation. Therefore, different gender features would be compared: grammatical and semantic gender, as well as feminine and masculine gender. By comparing these gender features, it would be possible to verify whether semantic gender would weigh more in memory than grammatical gender, and whether masculine would weigh more in memory than feminine.

On one hand, according to Vigliocco and Franck (1999), semantic gender would be more easily processed than grammatical gender since the former is redundantly gender specified. In other words, they claimed semantic gender is gender-specified both syntactically and semantically, while grammatical gender is only gender-specified syntactically. And because semantic gender is redundant, it might be easier to be retrieved in memory, that is, it might be more prominent in memory than grammatical gender. On the other hand, according to Casado *et al.* (2017), semantic gendered masculine nouns include both female and male representations, which means masculine works as a default gender in Brazilian Portuguese. And because masculine carries both gender representations, it also might be easier to be retrieved in memory, that is, it might be more prominent in memory than feminine.

Since semantic gender and masculine gender were predicted to weigh more in memory than grammatical gender and feminine gender, it was expected that structurally unacceptable candidates carrying the former kinds of features would be responsible for slower coreference processing. A reason for that lies in the fact structurally unacceptable candidates with semantic gender or masculine gender would be more preferable candidates than structurally unacceptable candidates with grammatical gender or feminine gender. A greater preference would mean larger interference effects, that is, more competition with the structurally acceptable antecedents and, consequently, slower coreference processing, as predicted by the content-addressable memory model (Engelmann *et al.* 2015; Lewis & Vasishth 2005).

Both Experiments 1a and 1b tested for grammatical and semantic genders; however, Experiment 1a tested for those types of gender in the feminine and Experiments 1b in the masculine as illustrated in Table 1.

Table 1. Experiments 1a and 1b.

Eye-tracking experiment	Type of gender	Structurally unacceptable antecedent candidates examples
Experiment 1a	Feminine semantic gender	*psicóloga* ('female therapist')
Experiment 1b	Feminine grammatical gender	*pessoa* ('person')
	Masculine semantic gender	*engenheiro* ('male engineer')
	Masculine grammatical gender	*gênio* ('genius')

6.1. Participants

Thirty-two native speakers of Brazilian Portuguese (twenty-six female and ten male, age average of 22) participated in Experiment 1a; and thirty-six (twenty-two female and fourteen male, age average of 22) participated in Experiment 1b. All participants were randomly invited to participate in this experiment as volunteers. They were undergraduate students and had normal or corrected-to-normal vision. All participants were naive in relation to the object of study of the experiment and signed a consent form giving permissions to the experimenter to publish the results.

6.2. Materials and design

The independent variables of the experiment were: a) structurally acceptable antecedent matching the gender of the pronoun, which is a factor that is directly related to Principle B; and b) distractor matching the gender of the pronoun, which is factor that relies purely on morphological agreement cues. This way, the experimental design was 2x2, with four main conditions. We also controlled for the distractor type of gender; therefore, half of the experimental trials contained distractors with semantic gender and the other half contained distractors with grammatical gender. It should be noticed that all distractors of Experiment 1a were feminine and all distractors of Experiment 1b were masculine.[2]

The experimental trials were arranged into four lists using a Latin Square. Each list was pseudo-randomized and contained twelve experimental items and twenty-four fillers. Each and every trial was accompanied by a comprehension question. Like in Example 1, the experimental trials were composed by embedded third-person-singular pronouns *ele*, 'him'/*ela*, 'her', with pronominal antecedents (masculine/feminine common nouns)

---

[2] The experimental materials of each experiment consisted of 48 sentences distributed in 4 conditions: (i) structurally acceptable antecedent mismatching the pronoun_distractor matching the pronoun, (ii) structurally acceptable antecedent mismatching the pronoun_distractor mismatching the pronoun, (iii) structurally acceptable antecedent matching the pronoun_distractor matching the pronoun, (iv) structurally acceptable antecedent matching the pronoun_distractor mismatching the pronoun. It was not possible to consider *type of gender* as part of the conditions; otherwise the experiment would be too long, with 8 conditions. Besides that, it would not be possible to find enough distractors with [+animate] grammmatical gender, as they are not numerous in Brazilian Portuguese. And even if it would be possible to find enough nouns with grammatical gender, some of them would be very low frequent. Thus, it was decided to work with few, but frequent nouns with grammatical gender. This way, type of gender was not manipulated, but controlled. In other words, half of the distractors carried grammatical gender (6 experimental items) and the other half carried semantic gender (6 experimental items). Consequently, each participant would see the same condition with the same type of gender twice.

followed by distractors, which are local antecedent candidates that cannot be considered as structurally acceptable antecedents due to Principle B structural constraints. Thus, the structurally acceptable antecedents should be the preferable antecedent candidates.

The experimental trials contained two regions of interest. The critical region was the pronoun, which contained the pronouns *ele*, 'him', or *ela*, 'her', formed by 3 characters. Before the pronoun, there was a relative pronoun *que* (who), which introduces the relative clause, followed by a transitive verb (approximately 5–6 characters). After the pronoun, there was a spillover region that contained an adverb of manner (approximately 9–11 characters). Since the critical region is too small (only 3 characters), it was expected that any processing difficulties readers could have at the pronoun region would spread to the subsequent region (the spillover region).

A sample of the materials of Experiment 1a can be found in Table 2. Brackets delimit the regions of interest. One may check the Appendix for a complete list of the materials used in both experiments.

Table 2. Sample of the materials for distractors with feminine semantic gender by regions of interest in Experiment 1a

	Antecedent mismatch	Antecedent match
Distractor match	O <u>bailarino</u> admira a <u>psicóloga</u> que ajudou [**ela**] [gentilmente] depois de uma das fases mais difíceis na vida. *The <u>dancer$_{[masc]}$</u> admires the <u>therapist$_{[fem]}$</u> who gently helped **her** after one of the most difficult phases in life.*	A <u>bailarina</u> admira a <u>psicóloga</u> que ajudou [**ela**] [gentilmente] depois de uma das fases mais difíceis na vida. *The <u>dancer$_{[fem]}$</u> admires the <u>therapist$_{[fem]}$</u> who gently helped **her** after one of the most difficult phases in life.*
Distractor mismatch	A <u>bailarina</u> admira a <u>psicóloga</u> que ajudou [**ele**] [gentilmente] depois de uma das fases mais difíceis na vida. *The <u>dancer$_{[fem]}$</u> admires the <u>therapist$_{[fem]}$</u> who gently helped **him** after one of the most difficult phases in life.*	O <u>bailarino</u> admira a <u>psicóloga</u> que ajudou [**ele**] [gentilmente] depois de uma das fases mais difíceis na vida. *The <u>dancer$_{[masc]}$</u> admires the <u>therapist$_{[fem]}$</u> who gently helped **him** after one of the most difficult phases in life.*
Question	A psicóloga é cruel com seus pacientes? *Is the therapist$_{[fem]}$ cruel with her patients?*	

The on-line dependent variables for both Experiments 1a and 1b are the following reading measures at the pronoun and at the spillover region: (a) First Fixation, duration of the first fixation in a word or region of interest; and (b) First Pass, sum of the durations of all fixations on a word or region before leaving it to the right or to the left. These measures correspond to be the very beginning of language processing.

6.3 Procedure

The experiment was conducted at a psycholinguistics laboratory in Brazil. The eye-tracker used in this experiment was *Eye Link 1000* and the experiment was programmed

and conducted on *Eye Track 7.10m*[3] software. All trials were typed in font *Monaco* size 12. The participants were instructed to seat comfortable and were given written and oral task instructions. The instructions screen is illustrated in Figure 1.[4]

```
During this test, you'll silently read several sentences.
Each sentence will be followed by a comprehension question.
As soon as you finish reading each sentence, press the right button in the
joystick to go to the comprehension question for that sentence.
To answer the question, press the left button for YES and the left button
for NO.
Before each sentence, you'll have to fixate your eyes at a black square on
the left corner of your screen.
By doing this, the sentence will appear in the screen immediately.
Please, press the right button to start the Practice Session.
```

Figure 1. Instructions screen for Experiments 1a and 1b.

After receiving the instructions, the calibration process would start followed by a short practice with six filler sentences so that the experimenter would check whether the participants understood the task and were performing it at a natural speed. Each participant performed one of the four lists of the experiment, which were pseudo-randomized by *Eye Track* software. The experiment duration was of twenty minutes approximately.

6.4. Analysis

The eye-tracking data was analyzed using the following tools: *Visual EDF to ASC*, to convert the .EDF files that *Eye Link 1000* generates; *Robodoc.py*[5], to clean eye blinks and long saccades (longer than 80ms); *Question_acc.py*[6] to compute the comprehension questions accuracy; *EyeDry*[7] to compute the reading measures; and *R* for the data exploration and statistical analysis.

Some experimental trials had to be excluded due to eye blinks and long saccades at the regions of interest (15% in Experiment 1a, and 21% in Experiment 1b). Moreover, 2 participants were excluded from analysis in Experiment 1a due to very slow reading as they trespassed the time limit (4 seconds) in all trials, including fillers, for this reason, 6% of the experimental data also had to be excluded.

Experiment 1a tested for feminine gender and Experiment 1b tested for masculine gender. Running two separate experiments with different participants was a solution

---

[3] The primary developers of *Eye Track* were David Stracuzzi and Jeff Kinsey and it is conceptually based on software written by Saarbruken and provided to UMASS by Christoph Scheepers. *Eye Track* can be downloaded for free on https://blogs.umass.edu/eyelab/software/.

[4] The participants received the instructions in Brazilian Portuguese, but we translated them to English for the purpose of this paper.

[5] *Rododoc.py* is a python script created by Adrian Staub and Chuck Clifton, and the 2016 version was revised by Jesse Harris. It can also be downloaded on https://blogs.umass.edu/eyelab/software/.

[6] *Question_acc.py* is a python script that comes with Robodoc.py utils to check questions accuracy and their reaction times.

[7] *EyeDry* was created by Chuck Clifton and can be downloaded on https://blogs.umass.edu/eyelab/software/.

found for the issues that a large experiment might cause in data and in participants. However, Experiment 1a and 1b were analyzed together. And as Experiment 1a and 1b were run with different participants, a *between-subjects* analysis was chosen.

The data was normalized with log-transformations and the means were centered. After that, linear mixed effects models (lmes) were created with the help of *lmerTest*[8] package in order to analyze the role of each independent variable in the results. The fixed effects of the lmes were: a) antecedent matching the gender of the pronoun (match/mismatch); b) distractor matching the gender of the pronoun (match/mismatch); c) distractor type of gender (semantic/grammatical); and d) distractor gender (masculine/feminine). On the other hand, the random effects were the participants and the items.

Full random lme models (with random slopes for each of the random effects) were created to analyze the data. However, the models with random slopes did not converge for the spillover regions in the First Pass. As trimming the outliers was not an option in this specific study due to the amount of data already excluded before, the lme models with random intercepts were run especially for these measures instead. For space reasons, only the statistically significant results were reported in the text. One may check the Appendix for a complete list of the lme results.

After running the lme models, post-hoc tests were conducted in order to figure out which pairwise comparisons among the conditions of the experiments were statistically significant. Differences between the least squared means (lsmeans), which are predicted marginal means of lme models, were estimated with the help of *emmeans*[9] package.

Aiming to clarify the statistically significant results, histograms with error bars were created using *ggplot2*[10] package. In order to facilitate the comprehension of the graphs, all statistically significant pairwise comparisons were indicated in the figures with "*".

6.5. Results

Including experimental trials and fillers, it should be mentioned that the participants answered the comprehension questions with an average of accuracy of 88% in Experiment 1a and 93% in Experiment 1b, which means that the participants were paying attention to the task and reading the sentences properly.

Results for each of the eye-tracking measures will be reported according to the regions of interest investigated.

---

[8] Alexandra Kuznetsova, Per Bruun Brockhoff and Rune Haubo Bojesen Christensen (2015). lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-29 on https://CRAN.R-project.org/package=lmerTest.
[9] Russel Lenth (2019). emmeans: estimated marginal means, aka Least-Squares Means. R package version 1.4. https://CRAN.R-project.org/package=emmeans.
[10] H.Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

6.5.1. First fixation duration

The results of the lme model are illustrated in Table 3. It was not found any statistically significant results for the spillover region; this is the reason why only the results at the pronoun were reported in the text.

Table 3. Lme results for First Fixation Duration at the pronoun.

Parameter	Pronoun			
	β_0	SE	t-value	p-value
(Intercept)	-0.64	0.31	-0.20	0.84
Distractor x Type of gender x Gender	-1.65	0.49	-3.35	0.0008
Antecedent x Distractor x Type	-0.97	0.49	-1.95	0.051
Distractor x Type of gender[a]	1.42	0.67	2.12	0.03

Note. [a]Post-hoc tests failed in predicting lsmeans for this interaction.

In post-hoc tests, the interaction Distractor x Type of gender x Gender was found statistically significant in one of the pairwise comparisons. The differences between the lsmean*s* ($\beta_0 = -0.57$, SE = 0.193, t-value = -2.94, $p = 0.004$) indicated that in the sentences in which the distractors matched the pronouns, reading times at the pronoun were longer for masculine semantic gender (lsmean = 5.53, SE = 0.06) than for feminine semantic gender (lsmean = 5.31, SE = 0.07).

The differences between the lsmeans pointed out that the interaction Antecedent x Distractor x Type of Gender was statistically significant in a couple of pairwise comparisons, although the results of the lme showed a trend towards significance as shown in Table 3.

Figure 2 illustrates this interaction Antecedent x Distractor x Type of Gender x Gender.
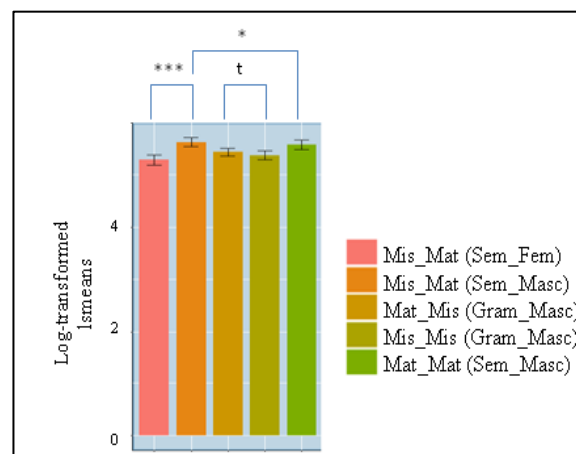

Figure 2. First Fixation Duration at the pronoun.

As can be seen in Figure 2, the differences between the lsmeans ($\beta_0 = -0.34$, SE = 0.09, t-value = -3.47, $p = 0.0006$) indicated that reading times in sentences in which antecedents and pronouns mismatched, followed by matching distractors, were longer when the distractors carried masculine semantic gender (lsmean = 5.63, SE = 0.08) than when they

carried feminine semantic gender (lsmean = 5.29, SE = 0.09). The same interaction also showed reading times at the pronoun were faster when antecedents matched the pronouns (lsmean = 5.43, SE = 0.07; lsmean = 5.37, SE = 0.08) than when they mismatched (lsmean = 5.63, SE = 0.08; lsmean = 5.58, SE = 0.09). This occurred in sentences with matching distractors with masculine semantic gender (β_0 = 0.19, SE = 0.87, t-value = -2.19, p = 0.03) and in sentences with mismatching distractors with masculine grammatical gender respectively, in a trend towards significance in the latter case (β_0 = -0.20, SE = 0.10, t-value = -1.95, p = 0.052).

6.5.2. First pass duration

6.5.2.1. Pronoun

The results of the lme model for First pass duration at the pronoun are illustrated in Table 4.

Table 4. Lme results for First Pass Duration at the pronoun.

Parameter	Pronoun			
	β_0	SE	t-value	p-value
(Intercept)	0.27	0.22	1.24	0.21
Distractor	-0.56	0.29	-1.9	0.059
Antecedent x Distractor	1.07	0.39	2.67	0.007
Antecedent x Distractor x Type	0.68	0.39	1.73	0.08

On one hand, the post-hoc tests did not show a statistically significant difference in pairwise comparisons regarding the variable Distractor at the pronoun. On the other hand, they showed statistically significant differences involving the interactions Antecedent x Distractor and Antecedent x Distractor x Type. Moreover, they indicated statistically significant pairwise comparisons between the lsmeans that were not found statistically significant, as the interaction Antecedent x Distractor x Gender.

Statistically significant differences between the lsmeans concerning the interaction Antecedent x Distractor at the pronoun is illustrated in Figure 3.
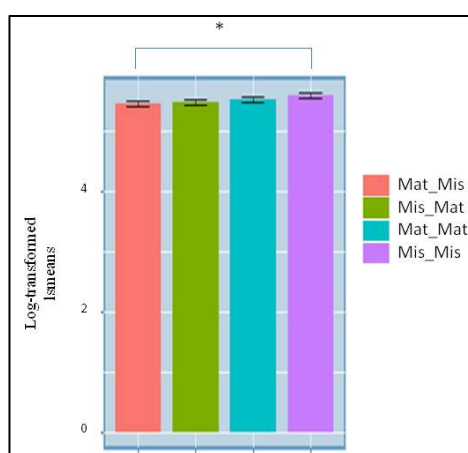

Figure 3. First Pass Duration at the pronoun by antecedents and distractors.

According to Figure 3, sentences in which pronouns and distractors mismatched, reading times were longer when antecedents also mismatched the pronouns (lsmean = 5.59, SE = 0.04) than when antecedents matched the pronouns (lsmean = 5.46, SE = 0.04), (β_0 = -0.57, SE = 0.193, t-value = -2.94, p = 0.004).

The interaction Antecedent x Distractor x Type is portrayed in Figure 4.
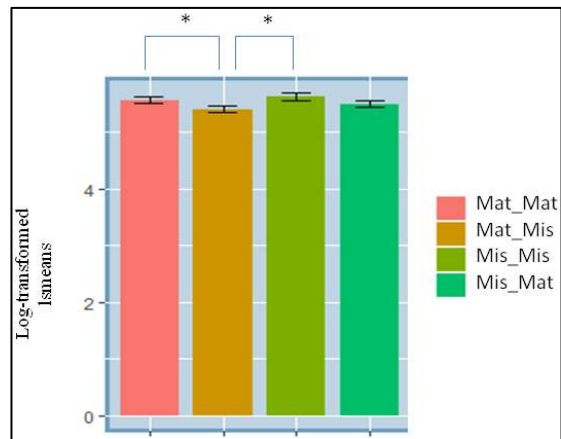

Figure 4. First Pass Duration at the pronoun by antecedents and distractors with grammatical gender.

As can be seen in Figure 4, when antecedents matched the pronouns, reading times at the pronoun were longer when distractors with grammatical gender also matched the pronouns (lsmean = 5.57, SE = 0.06) than when they mismatched them (lsmean = 5.40, SE = 0.06), (β_0 = 1.63e-01, SE = 8.27e-02, t-value = 1.97, p = 0.04). Moreover, reading times at the pronoun in sentences in which distractors with grammatical gender mismatched the pronouns were longer following antecedents that also mismatched the pronouns (lsmean = 5.63, SE = 0,06) than following antecedents that matched the pronouns (lsmean = 4.40, SE = 0.06), (β_0 = -2.20e-01, SE = 8.51e-02, t-value = -2.59, p = 0.01).

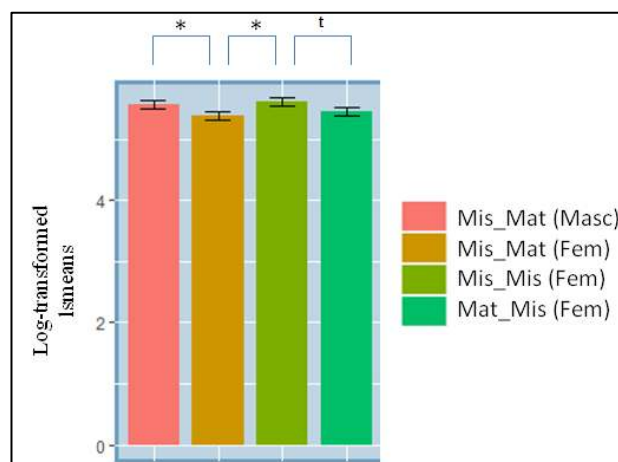The interaction Antecedent x Distractor x Gen is portrayed in Figure 5.

Figure 5. First Pass Duration at the pronoun by antecedents, distractors and gender.

As Figure 5 illustrates, when antecedents mismatched the pronouns, reading times at the pronoun were longer for masculine distractors that matched the pronouns (lsmean = 5.56, SE = 0.07) than for feminine distractors that matched the pronouns (lsmean = 5.38, SE = 0.07), (β_0 = -1.97e-01, SE = 9.84e-02, t-value = -2.00, p = 0.04). Furthermore, when antecedents mismatched the pronouns, reading times at the pronoun were longer with feminine distractors that also mismatched the pronouns (lsmean = 5.61, SE = 0.06) than when they matched the pronouns (lsmean = 5.38, SE = 0.07), (β_0 = -2.33e-01, SE = 8.60e-02, t-value = -2.71, p = 0.007). Finally, reading times at the pronoun were longer in sentences in which feminine distractors mismatched the pronouns following antecedents that also mismatched the pronouns (lsmean = 5.61, SE = 0.06) than following antecedents that matched the pronouns (lsmean = 5.45, SE = 0.06) in a trend towards significance (β_0 = -1.63e-01, SE = 8.51e-02, t-value = -1.91, p = 0.056).

6.5.2.1. Spillover

The results of the lme model for First Pass Duration at the pronoun +1 region are illustrated in Table 5.

Table 5. Lme results for First Pass Duration at the spillover.

Parameter	Spillover			
	β_0	SE	t-value	p-value
(Intercept)	0.31	0.19	1.63	0.10
Type of Gender	-0.58	0.24	-2.40	0.01*

Post-hoc tests did not show a statistically significant difference between the lsmeans regarding the variable Type of Gender at the spillover. However, this variable appears statistically significant in a couple of pairwise comparisons in post-hoc tests.

The difference of the lsmeans shows the interaction Antecedent x Type of Gender is statistically significant as can be seen in Figure 6.
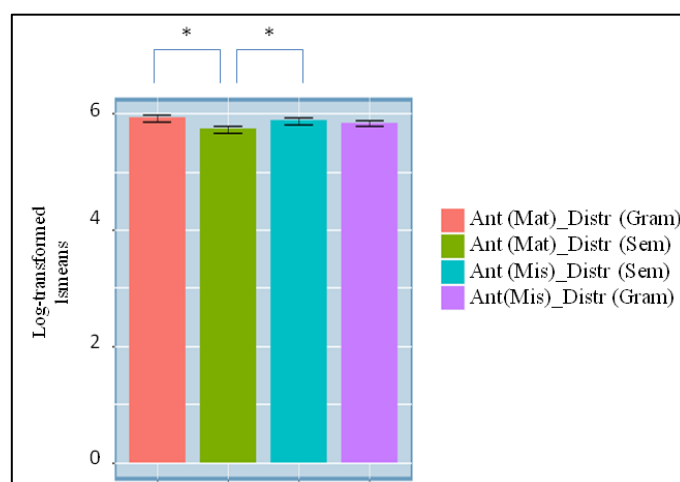

Figure 6. First Pass Duration at the spillover by antecedents and type of gender of the distractors.

When antecedents matched the pronouns, reading times at the spillover were faster for semantic distractors (lsmean = 5.73, SE = 0.05) than for grammatical distractors (lsmean = 5.92, SE = 0.05), (β_0 = 0.18, SE = 0.07, t-value = 2.66, *p* = 0.01). Moreover, reading times at the spillover were longer in sentences with semantic distractors following antecedents that mismatched the pronouns (lsmean = 5.87, SE = 6.02) than in sentences in which the antecedents matched them (lsmean = 5.73, SE = 0.05), (β_0 = -0.14, SE = 0.05, t-value = -2.50, *p* = 0.01).

Post-hoc tests also revealed the Distractor x Type of gender interaction was statistically significant as it can be noticed in Figure 7.
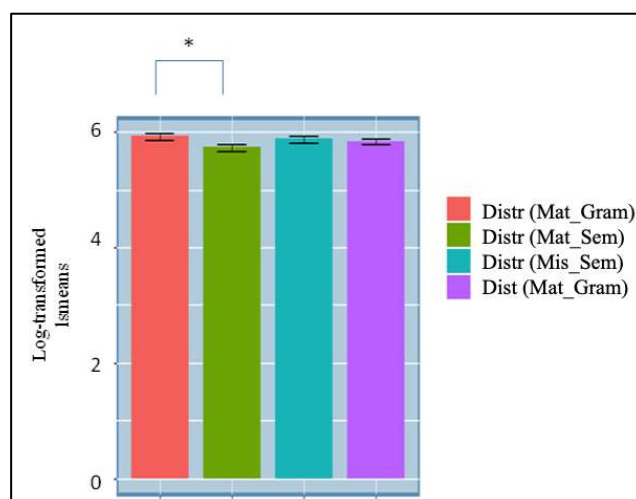

Figure 7. First Pass Duration at the spillover by distractors and type of gender.

Differences in the lsmeans showed reading times at the spillover were faster when distractors with semantic gender matched the pronouns (lsmean = 5.77, SE = 0.05) than when distractors with grammatical gender did (lsmean = 5.92, SE = 0.05), (β_0 = 0.14, SE = 0.06, t-value = 2.14, *p* = 0.04).

The interaction Antecedent x Distractor x Type of gender was also found statistically significant as shown in Figure 8.
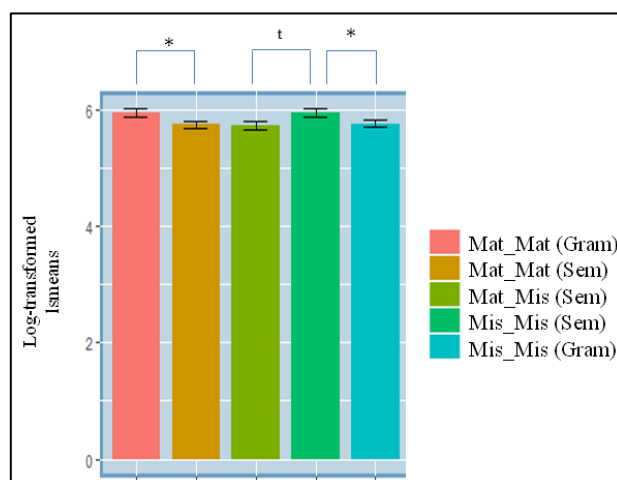

Figure 8. First Pass Duration at the spillover by distractors and type of gender.

On one hand, the differences between the lsmeans revealed when both antecedents and distractors matched the pronouns, distractors with semantic gender were responsible for faster reading times at the spillover (lsmean = 5.74, SE = 0.06) than distractors with grammatical gender (lsmean = 5.93, SE = 0.07), (β_0 = 0.19, SE = 0.09, t-value = 0.01 p = 0.03). On the other hand, when both antecedents and distractors mismatched the pronouns, distractors with semantic gender were responsible for longer reading times at the spillover (lsmean = 5.94, SE = 0.06) than distractors with grammatical gender (lsmean = 5.76, SE = 0.06).

Moreover, reading times at the spillover were faster when distractors mismatched the pronouns following matching antecedents (lsmean = 5.73, SE = 0.06) than following mismatching antecedents (lsmean = 5.94, SE = 0.06) in a trend towards significance (β_0 = 0.17, SE = 0.09, t-value = 1.96, p = 0.053).

7. Discussion

First Fixation Duration is an eye measure that corresponds to the very beginning of language processing. The results of this study indicated that at the very beginning of coreference processing, agreement between pronouns and antecedents is crucial for antecedent retrieval. In sentences in which the distractors mismatched the pronouns, reading times were faster when antecedents matched the pronouns than when antecedents mismatched them. This is evidence in favor of the importance of agreement in antecedent retrievals (*cf.* Figure 2).

The results of First Fixation Duration at the pronoun indicated that, at the very beginning of coreference processing, antecedent retrieval is also influenced by gender. When antecedents mismatched the pronouns, reading times in sentences in which masculine distractors that matched the pronouns were longer in comparison to feminine distractors that matched the pronouns (*cf.* Figure 2). According to Engelmann *et al.* (2015), what happened is called "inhibitory interference", that is, since the antecedents mismatched the pronouns, both antecedents and distractors were partial-matches, which increased the competition between the antecedents and the distractors. This explains the longer reading times in this case. Masculine caused more inhibitory interference than feminine, because masculine is the default gender. In other words, masculine carries the conceptual interpretation of both male and female referents (Casado *et al.* 2017), which makes masculine distractors preferred in comparison to feminine distractors. This is evidence that masculine weighs more in memory (it is more prominent) than feminine.

Finally, First Fixation results at the pronoun also showed that when distractors [with masculine semantic gender] matched the pronouns, reading times were faster when antecedents matched the pronouns than when they mismatched (*cf.* Figure 2). This result reflects the effects of binding structural constraints in antecedent retrieval. In order words, in order to be retrieved as antecedents, it not enough to only match the pronouns, as the distractors in this case did, but it is crucial to respect the binding structural constraints. Therefore, antecedents that matched the pronouns and respected binding structural constraints were quickly retrieved in memory. This result was also found by most of

studies in the field such as Nicol and Swinney (1989), Sturt (2003), Leitão *et al.* (2008), among others.

First Pass Duration is also considered to be an early eye tracking measure as it occurs before reanalysis processes. Thus, it describes early processing stages as well as First Fixation Duration. It was noticed that sentences in which antecedents matched the pronouns followed by distractors that also matched the pronouns, reading times were slowed down in comparison to sentences in which only antecedents matched the pronouns (*cf.* Figure 4). According to the content-addressable memory model (Lewis & Vasishth 2005), the overlap of cues between the antecedents and the distractors might have increased the competition among them, which explains the longer reading times. Engelmann *et al.* (2015) called this an "inhibitory interference". It is relevant to mention that this result contradicts studies such as Nicol and Swinney (1989), Sturt (2003), and Leitão *et al.* (2008), which argued in favor of the infallibility of binding constraints at initial processing stages. Therefore, according to the results found in this research, structural constraints do not block the influence of structurally unacceptable candidates that match the pronouns at early stages of coreference processing. It should be noticed that this also contradicts Badecker and Straub (2002) since they did not find any interference from structurally unacceptable candidates when there is a matching antecedent in the sentence. On the contrary, this piece of evidence is in line with the results found by Patil *et al.* (2016), which reported effects of structurally unacceptable candidates in First Pass and in First Pass Regression Probability, which are early processing measures as well.

Another evidence of the importance of gender agreement was that in sentences without matching distractors, the presence of a matching antecedent would facilitate antecedent retrieval in comparison to sentences in which no matching candidate is available (*cf.* Figure 3 and 4), which replicates the result found in First Fixation Duration. Thus, it means gender agreement is necessary in antecedent retrievals since the beginning of coreference processing.

Moreover, the results in First Pass at the pronoun indicated that when antecedents mismatched the antecedents in gender, but there were [feminine] matching distractors in the sentence, reading times were faster than if there were [feminine] mismatching distractors (*cf.* Figure 5). Similarly, Badecker and Straub (2002) reported this as an "accessible match condition", which is faster than a "no-match condition". This piece of result is in agreement with most studies in the field such as Nicol and Swinney (1989), Leitão *et al.* (2008), and Sturt (2003). In other words, when there is not a good match between antecedent and pronouns, structurally unacceptable candidates would be taken into account. According to Engelmann *et al.* (2015), this is a "facilitatory interference". In this case, misretrievals rates increases due to the fact that both antecedents and distractors are partial-matches in relation to the pronouns. Thus, distractors could be misretrieved by memory. However the question is that this was expected to happen at late stages of processing, as a last resource of finding an antecedent in the sentence. However, differently from the studies mentioned above, this phenomenon occurred here at the beginning of processing, which means that binding structurally constraints involving

antecedents and pronouns do ever work as a filter, not even in the beginning of antecedent retrieval.

Finally, in First Pass at the pronoun, when there were not any matching antecedents in the sentences, masculine distractors that matched the pronouns in gender would be responsible for longer reading times than feminine distractors (*cf.* Figure 5). For Engelmann *et al.* (2015), this is called "inhibitory interference", that is, since the antecedent mismatched the pronouns, both antecedents and distractors were partial-matches. Consequently, they both compete for retrieval, which caused the elevated reading times in this case. This is also evidence in favor of the idea that masculine gender is the default gender, that is, its conceptual representation includes male and female referents (Casado *et al.* 2017), which might be the reason why masculine is preferable over feminine. In other words, masculine is a stronger competitor than feminine, and that might justify longer reading times for masculine distractors than for feminine distractors.

The results in First Pass at the spillover showed both gender agreement and binding structural constraints are of paramount importance in order to retrieve antecedents in memory, especially when there is not any other candidate that matches the pronouns. It was found that when distractors mismatched the pronouns, reading times were faster when antecedents matched the pronouns than when they mismatched them (*cf.* Figure 8). This replicated the results found in First Fixation and First Pass Durations.

The First Pass results at the spillover also showed that when antecedents matched the pronouns, reading times for distractors with semantic gender were faster in comparison to grammatical gender (*cf.* Figure 6). Similarly, reading times were faster when distractors with semantic gender matched the pronouns in comparison to grammatical gender (*cf.* Figure 7). Likewise, when both antecedents and distractors matched the pronouns, distractors with semantic gender were responsible for speeding up the reading times in comparison to distractors with grammatical gender (*cf.* Figure 8). This last piece of result suggested that distractors have been taken into account by memory even though there were suitable antecedents in the sentences. This contradicts the studies of Nicol and Swinney (1989), Sturt (2003), and Leitão *et al.* (2008), which argued in favor of the infallibility of structural constraints at initial stages of processing. However, this result dialogues with Patil *et al.* (2016). Thus, structural constraints do not initially block the influence of structurally unacceptable candidates when they match the pronouns, even when there is already a suitable antecedent in the sentences. According to Engelmann *et al.* (2015), this is an "intrusion effect", that is, the similarity-based interference between the antecedent and the distractors, increase the chances of distractors misretrievals, which explains the faster reading times in this case. Moreover, it was easier for distractors with semantic gender to be retrieved rather than the distractors with grammatical gender because semantic gender has both semantic and syntactic representations. According to Viggliocco and Franck (1999), the redundancy of gender information makes semantic gender to be more easily retrieved in memory than grammatical gender.

However, First Pass at the spillover showed that when both antecedents and distractors mismatched the pronouns, reading times in sentences with distractors with semantic gender were longer in comparison to grammatical gender (*cf.* Figure 8). This

indicates that despite mismatching the pronouns, the distractors have been taken into account by memory. This apparently challenges the content-addressable memory model, which argues that only cue-matching candidates would be considered as potential candidates in memory. What happens in this specific case is that there was no candidate that matched the pronouns; therefore, memory seemed to seriously evaluate all the possibilities, including the ones that mismatched the searched cues. This is a novel piece of evidence in the literature.

The results found in Experiments 1a and 1b confirmed the hypotheses of this work. First, it seems gender cues play an important role in antecedent retrievals in memory. Gender agreement not only allows candidates that respect binding structural constraints to be retrieved as antecedents, but also candidates that violate these constraints. Therefore, it seems gender agreement is as important for coreference processing as binding principles are. Second, the results indicated that memory can be influenced by language in terms of being able to perceive differences that exist between morphological features. In this study, it was shown that semantic gender and masculine are more prominent in memory, that is, they can be more easily retrieved in memory since they are preferred over grammatical gender and feminine. Semantic gender is more easily retrieved in memory because it carries both semantic and syntactic gender information (Viggliocco & Franck 1999), whereas grammatical gender only carries syntactic gender information. Similarly, masculine is more easily retrieved in memory because it is the default gender, which means that it carries both male and female representations, working as a "neuter" gender, whereas feminine is a marked gender feature, referring only to female referents (Casado *et al.* 2017; Corbett 1991). This way, semantic gender and masculine gender, as expected, are more prominent, that is, weigh more in memory than grammatical gender and feminine gender.

This work brought novel contributions to understand the process of antecedent retrievals in memory, as it showed evidences of a rich gender variation language such as Brazilian Portuguese. The results reported here challenges most of the studies in the field, showing that binding structural constraints are not as important as we thought. Binding structural constraints do not work as a filter, initially blocking the influence of structurally unacceptable antecedents, but it operates together with [gender] morphological cues. Both structural constraints and gender cues are equally important in coreference processing so that structurally unaccepted candidates can be misretrieved from memory when they match the morphological cues generated by the pronoun. Furthermore, this paper showed memory perceives semantic and grammatical gender, as well as masculine and feminine gender as distinct features as language normally does. This is a novel contribution to the comprehension of how memory operates and works.

Moreover, it seems that the content-addressable memory model (Lewis & Vasishth 2005) as well as the extended content-addressable memory model (Engelmann *et al.* 2015) were found adequate to explain the results found here, except for the last piece of finding. The model still needs to explain how memory operates when no candidates match the cues generated during the encoding phase. It seems that in this case, even the mismatching cues might be taken into account by memory as a last resource strategy. Still, more studies are needed in order to better comprehend this.

References

Badecker,W. & Straub, K. (2002). The processing role of structural constraints on the interpretation of pronouns and anaphora. *Journal of Experimental Psychology Learning Memory and Cognition*, *28*, 748–769. https://doi.org/10.1037/0278-7393.28.4.748

Casado, A., Palma, A. & Paolieri, D. (2017). The influence of sex information on gender word processing. *Journal of Psycholinguistic Research*, *47* (3), 1–27. https://doi.org/10.1007/s10936-017-9546-3

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Corbett, G. (1991). *Gender*. UK: Cambridge University Press.

Dillon, B., Mishler, A., Sloggett, S. & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, *69* (2), 85–103. https://doi.org/10.1016/j.jml.2013.04.003

Engelmann, F., Jäger, L. & Vasishth, S. (2015). *The determinants of retrieval interference in dependency resolution: Review and computational modeling*. [Manuscript submitted for publication]. Retrieved from https://pdfs.semanticscholar.org/73bc/a9b53475d14b9f0a0fbeeee9469f25003ccb.pdf

Leitão, M., Peixoto, P. & Santos, S. (2008). *Processamento da co-referência intra-sentencial em português brasileiro*. Veredas, *12* (2), 50–61.

Lewis, R. L. & Vasishth, S. (2005). An Activation-Based Model of Sentence Processing as Skilled Memory Retrieval. *Cognitive Science*, *29* (3), 375–419. https://doi.org/10.1207/s15516709cog0000_25

Lewis, R.L., Vasishth, S. & van Dyke, J.A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Science*, *10*, 447–454. https://doi.org/10.1016/j.tics.2006.08.007

Nicol, J. & Swinney, D.A. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguists Research*, *18* (1), 5–19. https://doi.org/10.1007/BF01069043

Patil, U., Vasishth, S. & Lewis, R. (2016). Retrieval interference in syntactic processing: The case of reflexive binding in English. *Frontiers of Psychology, 7*, 329. https://doi.org/10.3389/fpsyg.2016.00329

Sturt, P. (2003). The time-course of the application of binding constraints in reference. *Journal of Memory and Language*, *48* (3), 542–562. https://doi.org/10.1016/S0749-596X(02)00536-3

van Dyke, J. & McElree, B. (2011). Cue-dependent interference in comprehension. *Journal of Memory and Language*, *65* (3), 247–263. https://doi.org/10.1016/j.jml.2011.05.002

Vigliocco, G. & Franck, J. (1999). When sex and syntax go hand in hand: Gender agreement in language production. *Journal of Memory and Language*, *40* (4), 455–478. https://doi.org/10.1006/jmla.1998.2624

APPENDIX

List of the experimental sentences of the Experiment 1a

Distractors with semantic gender

1) O perito/A perita acatou a delegada que chamou ela/ele novamente por volta do meio dia no escritório.
    A delegada chamou alguém por volta do meio dia no escritório?
2) O cenógrafo/A cenógrafa despediu a fotógrafa que ofendeu ela/ele puramente por causa de fofocas nos bastidores.
    Houve fofocas nos bastidores?
3) O bibliotecário/A bibliotecária seguiu a secretária que guiou ela/ele brevemente através do corredor do grande gabinete real.
    Há um corredor grande no gabinete real?
4) O cozinheiro/A cozinheira agradeceu a faxineira que acudiu ela/ele rapidamente no meio da confusão daquele restaurante.
    A confusão foi na padaria?
5) O bailarino/A bailarina admira a psicóloga que ajudou ela/ele gentilmente depois de uma das fases mais difíceis de sua vida.
    A psicóloga é cruel com seus pacientes?
6) O cabeleireiro/A cabeleireira detestou a coordenadora de RH que olhou ela/ele lentamente dos pés a cabeça durante a entrevista.
    O presidente da empresa foi quem conduziu a entrevista?

Distractors with grammatical gender

1) O veterinário/A veterinária reconheceu a pessoa que feriu ela/ele fortemente por trás da cabeça momento do assalto.
    O ferimento foi por trás da cabeça no momento do assalto?
2) O comissário de bordo/A comissária de bordo processou a estrela de cinema que xingou ela/ele ferozmente devido ao um mal entendido.
    A estrela da cinema xingou devido a um mal entendido?
3) O pedagogo/A pedagoga elogiou a criança que tocou ela/ele totalmente com uma das lições mais belas de coragem.
    A criança deu uma bela lição de coragem?
4) O advogado/A advogada hostilizou a visita que tratou ela/ele rudemente na frente de alguns convidados na festa.
    A visita foi rude no café da manhã?
5) O criminoso/A criminosa ameaçou a vítima que notou ela/ele de repente próximo ao prédio alguns dias depois do crime.
    A vítima foi assassinada?
6) O enfermeiro/A enfermeira conhecia a criatura que matou ela brutalmente na frente da casa da família.
    O crime aconteceu na frente do shopping?

List of the experimental sentences of the Experiment 1b

Distractors with semantic gender

1) A empregada/O empregado elogiou o proprietário que pagou ele/ela legalmente dentro da lei e das exigências do sindicato.

      O proprietário pagou dentro da lei e das exigências do mercado?

2) A bióloga/O biólogo ligou para o geógrafo que buscou ele/ela felizmente antes da enchente hoje à tarde.

      Houve uma enchente hoje?

3) A aluna/O aluno questionou o coreógrafo que julgou ele/ela duramente apenas pelo físico e não pelo talento.

      O julgamento foi duro?

4) A estagiária/O estagiário criticou o técnico que treinou ele/ela claramente aquém dos critérios estabelecidos pela empresa.

      O treinamento foi de acordo com os critérios estabelecido pela empresa?

5) A prefeita/O prefeito procurou o deputado que avisou ele/ela sabiamente acerca do futuro incerto do país.

      É certo o futuro do país?

6) A arquiteta/O arquiteto agradeceu o engenheiro que indicou ele justamente para um dos cargos mais cobiçados no mercado.

      O engenheiro foi indicado para um cargo ruim?

Distractors with grammatical gender

1)A arqueóloga/O arqueólogo recebeu o ente querido que ouviu ele/ela alegremente ao longo de todo a conversa.

      Houve uma conversa em família?

2) A fazendeira/O fazendeiro denunciou o ser que roubou ele cruelmente acima de qualquer suspeita por longos anos.

      O roubo aconteceu por muito anos?

3) A milionária/O milionário recompensou o gênio que alertou ele/ela severamente a respeito dos efeitos da crise na economia.

      O alerta era a respeito de uma crise na economia?

4) A médica/O médico reconheceu o indivíduo que golpeou ele/ela seriamente sem ter ao menos chance de defesa.

      O indivíduo de defendeu?

5) A menina/O menino abraçou o anjo que salvou ele/ela bravamente apesar de todos os perigos.

      A criança morreu?

6) A fazendeira/O fazendeiro sacrificou o bicho que mordeu ele/ela vorazmente perto do celeiro ontem pela manhã.

      O animal foi poupado?

Results of the lme models

Table 1. Lme results for First Fixation Duration.

Fixed effects	Estimate	SE	t-value	p-value
Pronoun				
(Intercept)	-0.64	0.31	-0.20	0.84
Antecedent	0.02	0.30	0.09	0.92
Distractor	-0.37	0.50	-0.75	0.45
Type of gender	-0.20	0.35	-0.58	0.56
Gender	-0.24	0.32	-0.76	0.44
Antecedent x Distractor	0.49	0.37	1.31	0.18
Antecedent x Type of gender	-0.14	0.39	-0.36	0.71
Distractor x Type of gender	1.42	0.67	2.12	0.037
Antecedent x Gender	0.005	0.40	0.01	0.98
Distractor x Gender	0.52	0.38	1.37	0.16
Type of gender x Gender	0.51	0.37	1.37	0.17
Antecedent x Distractor x Type of gender	-0.97	0.49	-1.95	0.051
Antecedent x Type of gender x Gender	0.60	0.51	1.16	0.24
Distractor x Type of gender x Gender	-1.65	0.49	-3.35	0.0008
Spillover				
(Intercept)	0.06	0.20	0.32	0.74
Antecedent	0.10	0.25	0.39	0.74
Distractor	-0.22	0.35	-0.61	0.53
Type of gender	-0.36	0.24	-1.50	0.13
Gender	-0.04	0.26	-0.16	0.86
Antecedent x Distractor	-0.05	0.29	-0.19	0.84
Antecedent x Type of gender	0.13	0.33	0.39	0.69
Distractor x Type of gender	0.59	0.50	1.18	0.23
Antecedent x Gender	-0.02	0.33	-0.07	0.93
Distractor x Gender	0.23	0.22	0.81	0.41
Type of gender x Gender	0.28	0.33	0.89	0.37
Antecedent x Distractor x Type of gender	-0.19	0.41	-0.47	0.63
Antecedent x Type x Gender	-0.09	0.44	-0.20	0.83
Distractor x Type of gender x Gender	-0.49	0.40	-1.22	0.22

Table 2. Lme results for First Pass Duration.

Fixed effects	Estimate	SE	t-value	p-value
Pronoun				
(Intercept)	0.27	0.22	1.24	0.21
Antecedent	-0.49	0.30	-1.63	0.10
Distractor	-0.56	0.29	-1.90	0.059
Type of gender	-0.37	0.28	-1.31	0.19
Gender	-0.29	0.28	-1.03	0.30
Antecedent x Distractor	1.07	0.39	2.67	0.07
Antecedent x Type of gender	0.22	0.39	0.56	0.57
Distractor x Type of gender	0.68	0.39	1.73	0.08
Antecedent x Gender	0.44	0.41	1.09	0.27
Distractor x Gender	0.39	0.38	1.04	0.29
Type of gender x Gender	0.34	0.37	0.91	0.35
Antecedent x Distractor x Type of gender	-0.59	0.54	-1.08	0.28
Antecedent x Distractor x Gender	-0.58	0.55	-1.04	0.29

Antecedent x Type of gender x Gender	0.20	0.53	0.39	0.69
Distractor x Type of gender x Gender	-0.40	0.51	-0.87	0.38
Antecedent x Distractor x Type of gender x Gender	-0.31	0.75	-0.41	0.67
Spillover				
(Intercept)	0.31	0.19	1.62	0.10
Antecedent	-0.03	0.23	-0.14	0.88
Distractor	-0.24	0.23	-1.04	0.29
Type of gender	-0.58	0.24	-2.40	0.01
Gender	-0.27	0.24	-1.12	0.26
Antecedent x Distractor	-0.25	0.33	-0.77	0.44
Antecedent x Type of gender	0.18	0.32	0.56	0.56
Distractor x Type of gender	0.07	0.32	0.24	0.81
Antecedent x Gender	-0.03	0.31	-0.11	0.91
Distractor x Gender	0.38	0.31	1.22	0.22
Type of gender x Gender	0.42	0.30	1.38	0.16
Antecedent x Distractor x Type of gender	0.81	0.46	1.74	0.08
Antecedent x Distractor x Gender	0.05	0.43	0.12	0.90
Distractor x Type of gender	-0.002	0.43	-0.006	0.99
Antecedent x Type of gender x Gender	-0.08	0.43	-0.19	0.84
Distractor x Type of gender x Gender	-0.62	0.61	-1.01	0.31