

## **A behavioral study to investigate the processing routes of grammatical gender in Brazilian Portuguese**

### ***Um estudo comportamental para investigar as rotas do processamento do gênero gramatical do português brasileiro***

Natália C. A. de Resende

Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte, Minas  
Gerais / Brasil

natcarol2000@gmail.com

Mailce Borges Mota

Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina / Brasil

CNPq

mailce.mota@ufsc.br

**Abstract:** In this paper, we present a behavioral study aimed at investigating whether inanimate grammatical gender transparent nouns (feminine nouns ending in *-a* and masculine nouns ending in *-o*) and inanimate grammatical gender opaque nouns (other endings) in Brazilian Portuguese are processed by the same or by distinct cognitive mechanisms. To this purpose, 19 subjects participated in a grammatical gender agreement task in two conditions: between a determiner (definite article) and a noun (condition 1), and a noun and an adjective (condition 2). Factors such as frequency of the nouns and adjectives (high vs. low) as well as their phonological form (gender transparent vs. gender opaque) were manipulated. Results show that, in both conditions, frequency is a strong predictor of faster reaction times, suggesting that gender transparent and gender opaque forms can be stored in memory. This pattern holds for both conditions. We interpret this finding as evidence for the single-mechanism view of language processing.

**Keywords:** grammatical gender; Brazilian Portuguese.

**Resumo:** Este artigo apresenta um estudo comportamental que teve como objetivo investigar se substantivos inanimados do português do Brasil, transparentes quanto ao seu gênero gramatical (substantivos femininos terminados em *-a* e masculinos terminados em *-o*) e opacos (outras terminações) são ou não processados por mecanismos cognitivos distintos. Para tanto, uma tarefa de concordância de gênero foi executada por 19 sujeitos em duas condições: entre um artigo definido e um substantivo (condição 1) e entre um substantivo e um adjetivo (condição 2). Fatores como a frequência dos substantivos e adjetivos (alta *vs.* baixa) e forma fonológica (transparente *vs.* opaca) foram manipulados. Os resultados mostram que, em ambas as condições, a frequência é um forte preditor de respostas mais rápidas, sugerindo que tanto formas transparentes quanto formas opacas podem ser armazenadas na memória. Este padrão foi encontrado em ambas as condições. Estes resultados foram interpretados como evidência para a visão unitária de processamento da linguagem.

**Palavras-chave:** gênero gramatical; português do Brasil.

Recebido em: 10 de dezembro de 2016.

Aprovado em: 6 de abril de 2017.

## 1 Introduction

One of the most intriguing phenomena of the human mind is its ability to learn and process language. Although language processing seems quite simple given the speed and automaticity with which the human mind deals with it, how the brain stores and processes language regularities and irregularities has been largely debated within the Psycholinguistic research field (for example, PINKER, 1991, 1999; ULLMAN, 2001; BYBEE, 1985, 1995; MACWHINNEY, 2001; CHOMSKY, 1981; JACKENDOFF, 2002).

On one view, linguistic processing relies on a dual mechanism (PINKER, 1991, 1999; ULLMAN, 2001): a mental lexicon specialized in the storage of simplex/irregular lexical forms (PINKER, 1991, 1999; ULLMAN, 2001) and a different mechanism specialized in the processing of complex/decomposable lexical forms. On a second view (e. g. BYBEE, 1985, 1995; MCWHINNEY, 2001), simplex as well as complex forms are retrieved from and processed by a single associative memory mechanism.

Both the dual and the single mechanism views discuss the factors that influence the processing of complex and simplex forms. According to dual models of language processing, only simplex/idiosyncratic forms are sensitive to frequency effects. Regular/complex forms, however, usually are not subject to such effect since they are processed by a cognitive system specialized in the computation of rules. The single mechanism view, on the other hand, predicts that both categories of lexical items (simplex and complex forms) are subject to frequency effects as they are all processed and retrieved from an associative lexical memory.

A series of studies (FLEISCHHAUER; CLAHSSEN, 2012; VAN DER LELY; ULLMAN, 2001; PRADO; ULLMAN, 2009, PRASADA; PINKER, 1993; CLAHSSEN, ROTHWEILER; WOEST, 1992; MARCUS *et al.*, 1995) has been proposed to investigate the processing routes of regular vs. irregular inflectional morphology focusing, mainly, on the verbal morphology of languages. For the English language, for instance, many studies (NEWMAN; ULLMAN, 2007; ULLMAN, 2001, 2004, 2005; PRASADA; PINKER, 1993) have shown that the processing of high-frequency and low-frequency regular forms are not correlated with their frequencies, but that speakers of English process high-frequency irregular verbs faster than low-frequency irregular verbs. These findings suggest that speakers of English activate a computational mechanism to process regular forms (by concatenating a base to a suffix: *walk + ed* in language production or decomposing the items into their morphological constituents in language comprehension), but rely on memory to process irregular forms. Research on languages other than English has shown a similar pattern (BOWDEN, 2007; MORGAN-SHORT *et al.*, 2010; CLAHSSEN; EISENBEISS; SONNENSTUHL, 1997). Other findings, however, contradict such claims. Alegre and Gordon (1999), Baayens *et al.* (1997) and Baayen, Tweedie and Schreuder (2002), for instance, have shown that regular forms are also sensitive to surface frequency effects, suggesting that they can also be stored in memory as simplex forms.

Given the debate in the literature, the question that arises is: If regular forms can also be stored, what kinds of structures can be stored and what kinds of structures can be computed? An answer to this question is still far from being trivial. We believe that not only verbal morphology can provide good insights to this debate. In this paper, we test the hypotheses of the dual and single mechanism models by investigating another aspect of morphology: the processing of grammatical gender of

inanimate nouns. The language of our investigation is Portuguese, more specifically Brazilian Portuguese which, so far, has not received much attention on the part of researchers interested in investigating the routes of language processing. The research question we address in the present study is the following: Is grammatical gender agreement in Brazilian Portuguese governed by a dual or by a single cognitive mechanism?

Before presenting a detailed description of the method adopted in the present study, we think it is worth presenting a review of the literature focusing on the cognitive mechanisms involved in the processing of grammatical gender and a brief (but complete) outline of the grammatical gender of Portuguese with the purpose of clarifying how the grammatical gender system of Portuguese can provide subsidies for a study on the balance between storage and computation.

### 1.1. Literature Review

Grammatical gender, unlike semantic gender, does not refer to the biological sex of the referent, but represents, instead, a formal, intrinsic feature of inanimate nouns (CORBET, 1991). In Portuguese, as in other Romance languages, animate and inanimate nouns are distributed into gender categories: feminine and masculine. For instance, the word *cultura* (“culture”) does not have an animate referent in the world, but it is a feminine noun. Similarly, the word *pensamento* (“thought”) does not have an animate referent in the world, but it is a masculine noun.

In Portuguese, as a general rule, nouns ending in *-a* tend to be feminine, and nouns ending in *-o* tend to be masculine. However, many other nouns with other endings can be of either grammatical gender. Hence, a computational mechanism could be at work to process the grammatical gender rule (nouns ending in *-a* are feminine, nouns ending in *-o* are masculine) and a distinct mechanism, an associative memory, would be responsible for storing and processing the grammatical gender of opaque forms.

Analogous to the debate focusing on the routes of verbal morphology processing, researchers interested in the processing of grammatical gender have also examined whether the grammatical gender of inanimate nouns is represented and processed by a dual or a single cognitive mechanism. A number of studies (e.g. AFONSO *et al.*, 2014; GOLLAN; FROST, 2001; HOHFELD, 2006) in languages other than Portuguese suggest that there is a dual route to process

grammatical gender. A study in Hebrew has offered evidence in favor of such claim. Gollan and Frost (2001) have shown that response times to gender regular/transparent nouns were faster than the response times to gender opaque/irregulars forms. This effect was also observed in error rates. The authors observed that, in Hebrew, the phonological form (the ending of the nouns) of either regular or irregular nouns also influences the processing of grammatical gender decisions. Based on this finding, they argued that there are two routes for the processing and representation of grammatical gender in Hebrew: a lexically based route and a computational route. The lexically based route would be responsible for nouns whose grammatical gender is stored as information of the lemma of each noun. The computational route, on the other hand, would be responsible for the processing and representation of the nouns whose grammatical gender information can be recovered from their surface (phonological/orthographic) form.

Another very recent source of evidence in favor of a dual-mechanism for the processing of grammatical gender comes from Spanish. Afonso et al. (2013) employed a gender decision task and a masked priming task in order to assess two factors: first, whether words ending *-o*, typical of masculine nouns, and words ending in *-a*, typical of feminine nouns, are more relevant to gender assignment than other gender correlated endings such as *-ad* and *-ón*. For example, in Spanish, forms ending in *-ad* are associated with the feminine gender category (*ciudad* – “city”) and forms ending in *-ón* are associated with the masculine gender category (*embrión* – “embryo”, *bastión* – “stick”). Second, they investigated whether determiners are retrieved from the mental lexicon in conjunction with lexical items. In the gender assignment task, results clearly show that gender is easily assigned to masculine nouns ending in *-o*, and feminine nouns ending in *-a*, but participants were more prone to errors when the nouns were irregular, i.e., feminine nouns ending in *-o*, and masculine nouns ending in *-a*, and also when the words were opaque, that is, when the endings of the nouns were different from *-a* or *-o*. In addition, participants were faster to respond to regular forms than irregular or opaque forms. As regards the question of whether determiners facilitate gender assignment and whether they are retrieved together with the noun, the results of the masked priming task showed that participants were, again, faster in gender decisions for nouns ending in *-o* or *-a* than for nouns with other endings. Researchers observed that the definite article was relevant only when the nouns were opaque forms, suggesting

that the article is retrieved with the noun just for opaque forms that do not end in *-o* or *-a*. The authors interpreted these results as evidence for a dual-mechanism operating in gender assignment processing as well as in gender agreement processing.

Some recent neuroimaging studies have contributed to the dual route hypothesis as well. Hernandez *et al.* (2004, 2007) and Padovani *et al.* (2005) have found in a gender decision task that gender opaque forms show greater activation in Brodmann's area BA45 than gender transparent forms. Caffarra *et al.* (2015) have found different ERP components for gender transparent and gender opaque forms during sentence agreement processing.

The consistency of the results for different languages reveals that the processing of regular (transparent) forms seems to be quite different from the processing of opaque forms. These results seem to be analogous to results reported by studies on verbal morphology of English, which report differences between the processing of regular and irregular past tense formation. However, an issue that remains unresolved concerns the fact that while some studies report effects of frequency for gender transparent forms, other studies (TAFT; MEUNIER, 1998; DESROCHERS *et al.*, 1989) have shown that transparent/regular grammatical gender nouns can be stored.

Taft and Meunier (1998), for instance, indicated in a gender decision task that the phonological cues of French nouns play an important role in gender decisions. Participants were faster in regular nouns than in irregular ones. Nevertheless, effects of regularity and frequency also played a role in error rates and an interaction between regularity and frequency in reaction and error rates was not found. These results suggest that although phonological cues facilitate gender assignment, regular/transparent gender nouns can be stored in the lexical memory. These results, however, contradict those of a seminal study for French carried out by Tucker *et al.* (1968). French speakers were exposed to nonwords which had grammatical gender correlated endings. They found that the proportion of gender category (masculine or feminine) an item received was related to the proportion to which the ending was associated with the masculine or feminine gender category. Thus, their experiment showed that gender assignment to nonword was the product of phonological cues present in the end of the items and that gender assignment can be independent of lexical access.

In a gender assignment task, Bates *et al.* (1995) showed that participants were faster and more accurate in Italian nouns with highly predictable gender endings, thus providing evidence for effects of phonological cues as participants were faster in phonologically transparent nouns (masculine nouns ending in *-o*, and feminine nouns ending in *-a*) than in opaque nouns (other endings). Although the study controlled for frequency effects (high vs. low) and semantic features (abstractness vs. concreteness), results did not reveal any significance for these factors on the gender assignment task. Based on these findings, Bates *et al.* (1995) claim that there are two stages for gender assignment: First, the word is retrieved from the mental lexicon, and after that, a post-lexical check takes place in order to assign gender to the word. Faster reaction times to regular phonological cues are explained as convergence of information provided by the word ending and gender of the lexical item. Longer reaction times to irregular words are explained as a mismatch between the phonological cues and the grammatical gender.

In summary, the studies reviewed above show that phonological cues facilitate gender processing. Some researchers interpret this finding as the activation of a computational mechanism to process transparent grammatical gender lexical items. Nevertheless, other studies show that the frequency of items can influence gender processing, suggesting that memory plays a role independent of the form of the nouns. Thus, the debate remains open. Aiming to contribute to the discussion on the duality vs. singularity of grammatical gender processing, we carried out a behavioral study to investigate whether grammatical gender of transparent nouns and grammatical gender of opaque nouns in Brazilian Portuguese are processed by the same or by distinct cognitive mechanisms. In our study, we took into account the above mentioned frequency effects rationale, that is, if we find that frequency effects influence the processing of gender opaque nouns and do not influence the processing of transparent nouns, then our results suggest that regular/transparent gender nouns are computed while gender opaque forms are retrieved from memory. Based on the results encountered by Afonso *et al.* (2014) for Spanish, we hypothesize that the processing of grammatical gender in Portuguese is governed by a dual cognitive mechanism.

## 2 Grammatical gender in Portuguese

In Portuguese, every single noun has an intrinsic gender. Nouns in Portuguese can be either masculine or feminine. In most nouns in Portuguese, gender information can be retrieved from their phonological forms: nouns ending in *-a* are, generally, feminine, and nouns ending in *-o* are, generally, masculine. For instance, *menina* (“girl”) is a feminine noun and *menino* (“boy”) is a masculine noun. Similarly, the words *casa* (“house”) and *cama* (“bed”) are feminine nouns, whereas *carro* (“car”) and *corpo* (“body”) are masculine nouns, although they do not stand for an animate referent in the world. There are also other endings typical of feminine and masculine nouns. Examples of those endings are: nouns ending in *-agem* and *-ade* are, always, feminine nouns: *garagem* (“garage”), *imagem* (“image”), *idade* (“age”), *cidade* (“city”). Words ending in *-or* and *-ema* are, usually, masculine nouns: *amor* (“love”), *vapor* (“steam”), *problema* (“problem”), *cinema* (“cinema”), *sistema* (“system”).

It is important to emphasize that the general rule (feminine nouns ending in *-a*, and masculine nouns ending in *-o*) is the productive rule. Some studies for Brazilian Portuguese (e.g. NAME, 2002) have shown that children assign feminine gender to nonwords ending in *-a*, and masculine gender to nonwords ending in *-o*. This study also shows the same pattern for adults in gender assignment tasks. However, grammatical gender cannot always be predicted from a word’s endings. Actually, a huge amount of nouns are opaque as regards its grammatical gender. In other words, gender information cannot always be retrieved from a word’s phonological structure. In fact, in Portuguese there are nouns whose phonological form contradicts their gender, for instance, *dia* (“day”), *problema* (“problem”), *sofá* (“sofa”), *razão* (“sense”), *mão* (“hand”). As we can notice, although words such as *dia*, *problema*, and *sofá* end in *-a*, they are masculine, whereas words such as *mão* and *razão*, although ending in *-o* are feminine nouns. There are also other word endings which do not provide any cue for gender assignment such as *leite*<sub>masc</sub>, *abacaxi*<sub>masc</sub>, *noite*<sub>fem</sub>, among many others.

Given this puzzling scenario, the question that arises is: How do native speakers of Portuguese know the gender of opaque forms (such as words as *leite*, *abacaxi*, *noite*) if gender assignment cannot be based on phonological cues? There are two possible answers to this question: The first answer could be that there are two different strategies

for gender agreement or gender assignment. This hypothesis relies on the fact that word endings function as a cue to gender agreement so that native speakers apply a very straightforward rule: if the word ends in *-a*, then it is feminine; if the word ends in *-o*, then it is a masculine noun. When such rule is not applicable, speakers rely on memory of the word's grammatical gender. This first answer implies that there is a dual route for gender assignment and gender agreement. The second answer is that grammatical gender assignment and agreement are based on the memory of the gender of each noun. Thus, native speakers of Portuguese simply know the gender of each noun. Therefore, gender assignment and agreement depend on memory. This answer implies that there is only one strategy for gender assignment and gender agreement and that there is no rule-governed behavior, which is in line with the single-mechanism account of language processing.

In the next section, we present the method employed in the present study to examine whether grammatical gender in Brazilian Portuguese is governed by a dual or by a single cognitive mechanism.

### 3 Method

In order to investigate whether a single or a dual route governs the processing of grammatical gender in Brazilian Portuguese, we employed a grammatical gender agreement task. Although studies on grammatical gender have traditionally used the gender assignment paradigm (e.g. TAFT; MEUNIER, 1998; AFONSO *et al.*, 2014; BATES *et al.*, 1995), we adopted the grammatical gender agreement task for two main reasons: first, because instead of gender assignment, the grammatical gender phenomenon is manifested in real language by means of gender agreement, i.e., in natural language, speakers make use of grammatical gender information for agreement purposes, not for gender assignment purposes. Second, since we intended to analyze the differences of grammatical gender processing between two conditions – article + noun (Condition 1) and noun + adjective (Condition 2) – it would not be possible to attribute gender to two words of different types such as noun and adjective at the same time since adjectives in Portuguese are either neutral (e.g. *feliz* – “happy”) or inflected by gender (e.g. *bonit-o*<sub>masc</sub> or *bonit-a*<sub>fem</sub> – “beautiful”).

The research question we address by conducting a grammatical gender agreement task is: Is the processing of grammatical gender governed by a single or a dual mechanism in i) the article + noun and ii) in the noun + adjective condition? We hypothesize that grammatical gender agreement processing is governed by a dual route and that we will find similar patterns in the article + noun agreement and in the noun + adjective agreement conditions. In both conditions, we do not expect to find frequency effects for the agreement of transparent forms but only for the opaque forms. A lack of frequency effects in transparent forms processing suggests that they are decomposed in their form in order to have gender agreement assignment. Instead of recurring to memory of the item's gender, speakers rely on the phonological cues, that is, the word's endings, to perform gender agreement both with the determiners and with the adjectives. On the other hand, since no phonological cues are provided by the opaque forms and the phonological cues of the opaque forms contradict their gender, speakers rely on memory of the item's gender to perform gender agreement. In what follows, we describe our task design.

### 3.1 Task Design

**Materials:** The entire set of stimuli comprised 112 items in total. We used the *Corpus Brasileiro* available online<sup>1</sup> to select 32 nouns (16 transparent and 16 opaque nouns) for condition 1: agreement between the definite article and the noun. We also selected 32 nouns (16 transparent and 16 opaque) to construct the stimuli for condition 2: agreement between the noun and the adjective. In this paper, we will refer to the items selected for condition 2 as *bigrams*. We considered transparent forms feminine words ending in *-a*, and masculine words ending in *-o*; we considered opaque forms all other word endings, including feminine words ending in *-o* and masculine words ending in *-a*. Half of the transparent nouns (8 items) and half of the opaque nouns (8 items) selected for condition 1 and condition 2 were high-frequent nouns (310-4608 per million) and the other half were low-frequent nouns (1-17 per million). We also tried to approximate the number of feminine and masculine nouns selected as much as possible. To compose condition 1, we selected seven masculine and nine feminine opaque nouns as well as nine masculine and seven

---

<sup>1</sup> [www.linguateca.pt](http://www.linguateca.pt)

feminine transparent nouns. To compose condition 2, we selected eight masculine and eight feminine opaque nouns as well as eight masculine and eight feminine transparent nouns.

The experiment also included 48 fillers (nonwords). The nonwords were phonologically legal letter strings. They were constructed by exchanging a vowel phoneme or a consonant phoneme of existing gender transparent and gender opaque nouns, e.g. \**plefe* (correct form = *plebe*) or \**visma* (correct form = *cisma*).

To construct the bigrams of condition 2, we looked up in the corpus the adjective most likely to co-occur with the noun selected. Besides manipulation of the frequency of the noun of a bigram, we also manipulated the frequency of occurrence in the corpus of the adjectives following the nouns in a bigram. The adjectives of the bigrams selected were either high-frequent or low-frequent adjectives so that we obtained frequent and infrequent nouns co-occurring with frequent and infrequent adjectives. As it was not always possible to find adjectives co-occurring with infrequent nouns in the *Corpus Brasileiro*, we used Google search to find adjectives likely to co-occur with the infrequent nouns selected. The adjectives of the bigrams selected from Google were also controlled by frequency in our analysis. The control of the frequency of the two elements of a bigram was adopted to investigate which one matters for the gender agreement: the frequency of the head of the noun phrase or of the adjective itself.

### 3.2 Apparatus

The experiment was programmed on *E-prime* software for psychological experiments. Items were presented in the center of the screen in a white lowercase size 18 against a black background of a LG Computer Monitor (Flatron L1755S).

### 3.3 Participants

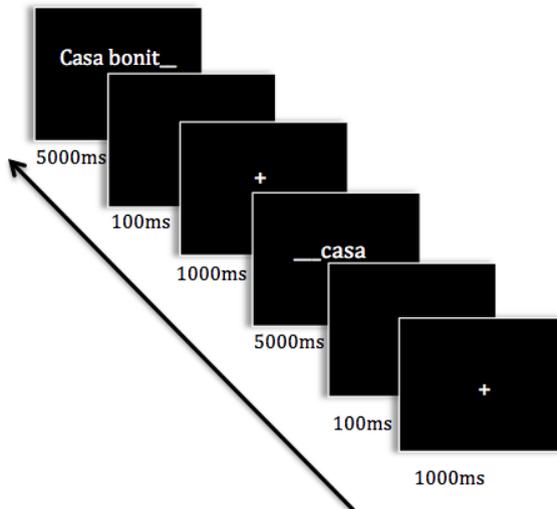
Participants were 19 right-handed native speakers of Brazilian Portuguese (9 female,  $M_{Age} = 31$ , range $_{Age} = 21-44$ ), recruited among graduate and undergraduate students at the Federal University of Santa Catarina (UFSC), Brazil. Subjects did not report previous register of neurological disorders. All participants reported having normal or corrected-to-normal vision. Subjects provided informed consent to

participate as volunteers in the present study. The study was approved by the research ethics committee of UFSC.

### 3.4 Procedure

The set of stimuli was presented visually and each subject performed the grammatical gender agreement task individually. The experimenter gave the task instructions to each participant while they appeared on the screen. The task consisted of a visual grammatical gender agreement in two conditions. Subjects were asked to push button 1 for letter *A* or button 5 for letter *O* which corresponded to the articles “A” and “O” preceding each noun or the ending of the adjectives to be inflected. Each trial was preceded by a fixation cross ‘+’ of 1000ms and the items remained on the screen up to 5000ms or until any button was pushed. We introduced 10 practice items before the beginning of each experimental session to make sure all participants were able to perform the main task properly. Figure 1 illustrates the presentation of the stimuli.

FIGURE 1 – Design of the experiment



NOTE: After the fixation cross, participants saw a blank screen followed by stimuli of condition 1 and condition 2. They then performed a gender agreement task. Stimuli were presented in different random order to each participant.

## 4 Results

### 4.1 Description of the data

Although we have coded participant's response for accuracy and reaction times, in this paper our analysis will be based on reaction times only. Inaccurate responses were removed from the dataset. We obtained 608 data points per condition, a total of 1216 data points. Table 1 displays description of reaction times collected for each category of condition 1 and condition 2.

TABLE 1 – Descriptive data per category

| Conditions  | Category                       | Mean   | SD    |
|-------------|--------------------------------|--------|-------|
| Condition 1 | Transparent frequent nouns     | 890ms  | 251.4 |
|             | Transparent infrequent nouns   | 1647ms | 688.8 |
|             | Opaque frequent nouns          | 1033ms | 304.8 |
|             | Opaque infrequent nouns        | 1184ms | 462.4 |
| Condition 2 | Transparent frequent bigrams   | 1563ms | 557.5 |
|             | Transparent infrequent bigrams | 1549ms | 582.8 |
|             | Opaque frequent bigrams        | 1413ms | 611.7 |
|             | Opaque infrequent bigrams      | 1777ms | 621.3 |

In condition 1, the results show that participants were faster to respond to transparent frequent nouns than to transparent infrequent nouns (890ms vs. 1647ms)  $t = -14.243$ ,  $df = 266.86$ ,  $p\text{-value} < 2.2e-16$ . Similarly, participants were faster to respond to opaque frequent nouns than to respond to opaque infrequent nouns (1033ms vs. 1184ms)  $t = -2.9157$ ,  $df = 283.25$ ,  $p\text{-value} = 0.003$ . These results show that frequency is a strong predictor for response speed regardless of the form of the nouns.

For condition 2, data description shows that participants were much faster to respond to opaque frequent bigrams than to respond to opaque infrequent bigrams (1413ms vs. 1777ms)  $t = -5.8663$ ,  $df = 295.75$ ,  $p\text{-value} = 1.19$ , suggesting that memory plays an important role in the processing of agreement between opaque forms and adjectives. On the other hand, reaction times for gender agreement between

frequent and infrequent transparent nouns and adjectives are very close (1549ms vs.1563ms)  $t = 0.4102$ ,  $df = 270.7$ ,  $p\text{-value} = 0.68$ , suggesting that a computational mechanism could be at work for agreement with transparent forms. Surprisingly, however, it is possible to notice that in condition 2 participants were slightly faster in agreement with transparent infrequent bigrams than in agreement with transparent frequent ones. Although the difference is statistically irrelevant ( $p > 0.6$ ), these results might show that participants were faster in agreement with transparent infrequent bigrams because this category of items imposed more processing costs than agreement with transparent frequent bigrams, since nouns were first accessed in memory and, after that, were decomposed to perform agreement, while infrequent bigrams were only decomposed to perform agreement.

Since our task design contains two different conditions of gender agreement and each condition presents different types of factors that might affect reaction times, we adopted the *Linear Mixed Effects Models* approach with crossed effects of subjects and items for the statistical analysis to obtain a better picture of the data. The next section describes how the data analysis on reaction times was carried out.

#### 4.1.1 Data analysis

The *Mixed Effects* models as well as data pre-processing were run in *Language R*. For the *Mixed Effects* models, we used the *lme4 package* (BATES, 2005; R Development Core Team, 2008). Prior to the data analysis, we screened the reaction times for extreme outliers ( $RT < 400ms > 3000ms$ ). The outliers were removed from the data set of each category from condition 1 and condition 2 separately using *Grubbs test* for one or two outliers (GRUBBS, 1950), package *Outliers*. We removed 12 data points from condition 1 (1,97%) and 15 data points from condition 2 (2,4%). After cleaning the dataset, histograms showed that the distribution of reaction times of condition 1 and condition 2 were left-skewed. In order to minimize the skewness of the distribution, we log-transformed reaction times.

As will be shown in the next section, we ran separate models for condition 1 and condition 2. This procedure was adopted in order to investigate more closely the effects of predictors for each condition. The fixed effects we included in the models of condition 1 were: lemma

frequency of nouns (continuous), number of graphemes of nouns (continuous), sex/gender of participants (categorical), grammatical gender of the nouns, that is, whether the noun was feminine or masculine (categorical), number of graphemes of nouns (continuous) and number of phonemes of nouns (continuous). For condition 2, we included the following fixed effects: lemma frequency of the nouns (continuous), lemma frequency of adjectives (continuous), sex/gender of participants (categorical), grammatical gender of the nouns, that is, whether the noun of the bigram was feminine or masculine (categorical), and number of graphemes of bigrams (continuous). The number of graphemes of nouns and bigrams were included in the models as control factors in order to maximize statistical power by reducing residuals.

In each model, we included subjects and items (nouns for condition 1 and bigrams for condition 2) as random effects. In order to determine which random slope should be included in the models, we adopted a forward procedure, starting with a minimum model. Then, we included all possible random slopes for subject and items (words and bigrams) one by one. Models converged easily after scaling the continuous predictors. Following Baayen et al. (2008), model comparisons were made using likelihood ratio test to verify if the random slopes and interactions added to the minimum model were justified. As a result, the random slopes included were removed from the models of condition 1 and condition 2. P-values reported here were obtained by means of ANOVA type 3 which tests for the presence of significant interactions.

Table 2 displays the results of the model that best explains reaction times of gender agreement between article and nouns (condition 1). Table 3 exhibits results of the model that best explains reaction times of gender agreement between nouns and adjectives (condition 2). Table 1 and Table 2 show only factors that appear to have justifiable inclusion<sup>2</sup> in the final model according to results obtained after models comparison by means of likelihood ratio test.

---

<sup>2</sup> Following a standard procedure (see BARR *et al.*, 2013), we constructed various models for comparison purposes, using the likelihood ratio test. This test indicates the factors and interactions that are more likely to influence the dependent variable. Thus, the models presented in this paper do not always include all factors and interactions tested since they are the models that best fit our dataset.

TABLE 2 – Model that best explains reaction times for condition 1

| <b>Fixed effects</b>        | <b>Estimate</b>  | <b>Std.Error</b> | <b>t</b>              | <b>Pr(&gt;F)</b> |
|-----------------------------|------------------|------------------|-----------------------|------------------|
| Intercept                   | 2.94986          | 0.06337          | 46.55                 | 2.2e-16 ***      |
| Lemma Frequency             | 0.00433          | 0.01456          | 0.30                  | 0.7685           |
| Category                    | 0.02056          | 0.02028          | 1.01                  | 0.3198           |
| Number of graphemes         | 0.01499          | 0.01247          | 1.20                  | 0.2402           |
| Participants' Gender (Male) | 0.02773          | 0.03472          | 0.80                  | 0.4355           |
| Nouns' Grammatical Gender   | -0.02450         | 0.02111          | -1.16                 | 0.2563           |
| Lemma Frequency:category    | -0.14243         | 0.02338          | -6.09                 | 1.898e-06 ***    |
| <b>Random Effects</b>       |                  |                  |                       |                  |
| <b>Groups</b>               | <b>name</b>      | <b>Variance</b>  | <b>Std. Deviation</b> | <b>Corr</b>      |
| <b>subjects</b>             | <b>intercept</b> | 0.005334         | 0.07303               |                  |
| <b>Word</b>                 | <b>Intercept</b> | 0.002329         | 0.04826               |                  |
| Residual                    |                  | 0.011867         | 0.10893               |                  |

TABLE 3 – Model that best explains reaction times for condition 2

| Fixed effects                     | Estimate         | Std.Error | t              | Pr(>F)      |
|-----------------------------------|------------------|-----------|----------------|-------------|
| Intercept                         | 3.253e+00        | 1.060e-01 | 30.695         | 0.002588 ** |
| Lemma Frequency (nouns)           | -6.614e-06       | 1.985e-06 | -3.332         | 0.002588 ** |
| Category                          | -4.639e-02       | 2.896e-02 | 0.61           | 0.121367    |
| Number of graphemes               | -5.688e-03       | 9.800e-03 | -0.58          | 0.568410    |
| Participants' Gender              | 2.077e-02        | 4.111e-02 | 0.50           | 0.619894    |
| Lemma Frequency (nouns): Category | 5.949e-06        | 2.303e-06 | 2.583          | 0.015618 *  |
| Grammatical gender                | 2.338e-02        | 2.187e-02 | 1.069          | 0.295044    |
| Random Effects                    |                  |           |                |             |
| Groups                            |                  | Variance  | Std. Deviation | corr        |
| <b>subjects</b>                   | <b>Intercept</b> | 0.007654  | 0.08749        |             |
| <b>Bigram</b>                     | <b>Intercept</b> | 0.003083  | 0.05553        |             |
| Residual                          |                  | 0.010900  | 0.10440        |             |

The model of condition 1 shows a significant interaction between category (transparent nouns vs. opaque nouns) with frequency (high vs. low). This interaction means that frequency affects categories of nouns differently. The interaction plot of Figure 2 illustrates the effect of frequency on reaction times of gender agreement between articles and nouns. Importantly, frequency of lemma itself and category itself did not reach significance. Factors such as number of graphemes and grammatical gender of the nouns (masculine vs. feminine) did not influence reaction times. Participants' gender (male vs. female) did not influence reaction times. We did not find a three-way interaction of gender, frequency, and category ( $\beta=0.010096$ ,  $t=0.53$ ), which suggests that reaction times of both men and women were equally influenced by interaction between frequency and category of nouns. We did not find a three-way interaction of grammatical gender, frequency, and category ( $\beta=4.528e-02$ ,  $t=0.55$ ) either, suggesting that reaction times for both masculine and feminine nouns were equally influenced by interaction between frequency and category of nouns.

As regards condition 2, results show a main effect of lemma frequency of nouns and a significant interaction between lemma frequency of nouns and category. As frequency of adjectives did not reach significance neither did a three-way interaction of lemma frequency of nouns, lemma frequency of adjectives, and category, our final model includes only the interaction between lemma frequency of nouns and category. As shown in Table 3, similar to condition 1, there is an interaction between lemma frequency of nouns and category, indicating that lemma frequency of nouns influenced reactions times of transparent and opaque bigrams differently. However, the effect of lemma frequency in condition 2 is more prominent than in condition 1, especially on opaque forms. The interaction plot of Figure 3 illustrates this effect. Concerning the other predictors included in our final model of condition 2 (participants' gender, grammatical gender of nouns, and number of graphemes of bigrams), as in condition 1, those predictors failed to reach significance as revealed by p-values presented in Table 3.

FIGURE 2 – Interaction Plot Condition 1 – Transparent forms x Opaque forms

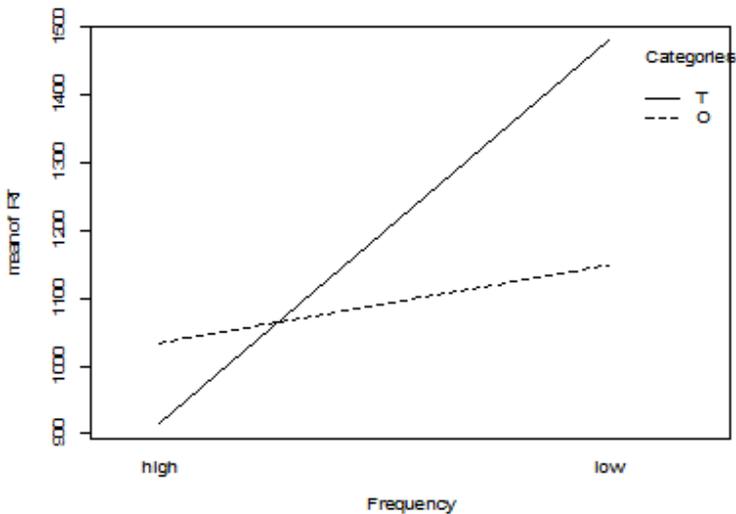
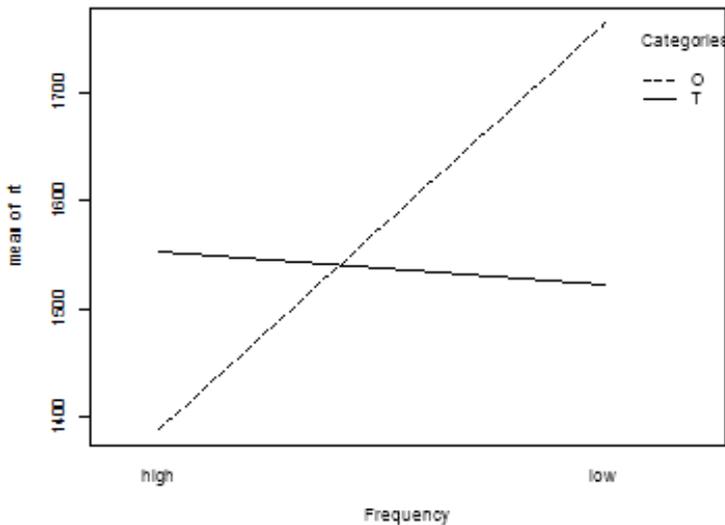


FIGURE 3 – Interaction Plot Condition 2 – Transparent forms x Opaque forms



## 5 Discussion

In the present study, we investigated the processing routes of grammatical gender agreement in Brazilian Portuguese in two conditions: between an article and a noun (condition 1) and between a noun and an adjective (condition 2). Results suggest that grammatical gender agreement processing between article and noun differs from that between noun and adjective. We can notice, by observing reaction times (see Table 1), that agreement between noun and adjectives elicits more processing costs, as reflected by longer reaction times, than agreement between article and nouns. It is our contention that, as subjects had more information to process in condition 2 than in condition 1 (for instance, a higher number of phonemes and graphemes, two items for retrieval, items decomposition, and concatenation), reaction times were longer as a result of the fact that more information demands more time to be processed.

In condition 1, while an interaction between frequency and form category plays a role in reaction times, frequency and form category alone fail to reach significance. This result suggests that form category itself is not a predictor of reaction times and neither is frequency itself. In fact,

there is a combination of both factors, but with differences concerning the influence of frequency on reaction times between form categories. This is very clear when we compare the reaction times of frequent transparent forms and frequent opaque forms. Participants were faster in gender agreement for frequent transparent forms compared to frequent opaque forms, infrequent transparent forms and infrequent opaque forms, replicating previous findings for other languages such as Italian (BATES et al., 1995), French (TAFT; MEUNIER, 1998; DESROCHERS et al., 1989; TUCKER et al., 1968), German (HOHFELD, 2006), Hebrew (GOLLAN; FROST, 2001) and Spanish (AFONSO et al., 2013). The faster reaction times to respond to frequent transparent forms than to respond to frequent opaque forms suggest that the phonological cues of the transparent forms yield a processing advantage as compared to a processing disadvantage for gender agreement with opaque forms, which, in turn, suggests that phonological cues facilitate gender processing.

Although in condition 1 our results indicate that high-frequency transparent forms can be stored, replicating results of Taft and Meunier (1998) and Desrochers et al. (1989), our findings suggest that low-frequency items might be decomposed. One of the clues to support this assumption is that, in condition 1, it is possible to observe that participants were slower to process infrequent transparent forms than infrequent opaque forms. A possible explanation for this phenomenon would be an increase in the processing costs while decomposing transparent infrequent nouns to retrieve gender information. This finding suggests that, on the one hand, high-frequency transparent forms are fully accessed in memory, but infrequent items might be decomposed. Thus, results of condition 1 suggest that computation would be applicable only when access to items in memory is difficult. Our assumption is based on the fact that other studies (e.g. Laine, Vainio and Hyönä, 1999) have shown that decomposition and concatenation processes are more demanding than lexical access, resulting in longer reaction times.

In condition 2, results also show that frequency of nouns plays a role in reaction times. As in condition 1, our final model shows a significant interaction between lemma frequency of the nouns and category, suggesting that frequency influences word categories differently. However, in condition 2, as illustrated by the interaction plot of Figure 3, it is possible to observe that the influence of lemma frequency of the nouns is stronger on opaque forms than on transparent forms, indicating

that opaque forms are stored and transparent forms might be computed. As previously mentioned, surprisingly, subjects were slightly slower to perform gender agreement with high-frequent transparent forms than with low-frequent transparent forms. Following Bates et al. (1995) rationale, we maintain that this phenomenon might be related to a three-stage processing strategy for high-frequent transparent forms – first lexical access, second item decomposition, third affix concatenation – as opposed to a two-stage processing for the low-frequency items (decomposition followed by affix concatenation). How to interpret these results in the light of the dual mechanism model and single-mechanisms model?

One may argue that since we found different patterns for both types of phonological forms, a dual mechanism would be at work in the processing of grammatical gender between an article and a noun and a noun and an adjective in Portuguese. However, we claim, on the basis of the frequency effects that were observed influencing the processing of both transparent and opaque forms, that there is a single mechanism governing grammatical gender processing in condition 1. As regards condition 2, in our view, there are two mechanisms governing gender agreement processing between nouns and adjectives. However, our results also indicate that, even if a computational processing strategy is operating, items are still being lexically accessed. This interpretation is based on the slightly longer reaction times in high-frequent transparent nouns in condition 2, suggesting an *a priori* lexical access and further decomposition and concatenation, leading to longer reaction times.

Our experiment shows that both transparent and opaque gender forms can be stored, but computational processing can be activated under some circumstances: when subjects do not readily extract the items from the mental lexicon and, as in condition 2, when there is more information to be processed. It seems that when the strings are longer as they are in condition 2, subjects prefer to decompose them to perform agreement. They also prefer to decompose items that are not easily accessed in memory as occurred with infrequent gender transparent nouns in condition 1. As observed by Rastle and Davis (2008), although more time-consuming, decomposition is an easier strategy when items present clues on their surface form.

Our results indicate that memory is involved regardless of the form of the noun when it is a frequent noun, that is, subjects recognize

the item and look up for it in the mental lexicon. Thus, we suggest that it is not the case that a computational route is activated every time subjects face a transparent form. The first strategy is to search for the item in the mental lexicon. If the item cannot be found, a second strategy takes place: decomposition. Based on these findings, we argue that our results do not support the dual mechanism view of grammatical gender processing and, therefore, do not support our initial hypothesis that the processing of grammatical gender in Portuguese is governed by a dual cognitive mechanism. On the contrary, our experiments show that regular/transparent forms can be stored as whole words, as predicted by single mechanism models.

It is noteworthy that the number of participants selected and number of items selected limited the experiment presented here. An experiment containing a larger sample of participants and items could provide different results. Future research with more subjects and more levels of items frequency is necessary to verify whether very infrequent transparent nouns are in fact subject to decomposition. In order to disentangle this issue, it would be of great aid to test if medium-frequency gender transparent nouns are also subject to decomposition or if they can also be accessed in memory just as high-frequency gender transparent nouns.

### **Acknowledgement**

This research was supported by CAPES/NUFFIC Grant 051/13 – 23038.007129/2013-40.

### **References**

AFONSO, O.; DOMÍNGUEZ, A.; ÁLVAREZ, C. J.; MORALES, D. Sublexical and lexico-syntactic factors in gender access in Spanish. *Journal of Psycholinguistic Research*, Springer Link, v. 43, Issue 1, p. 13-25, Feb. 2014.

ALEGRE, M.; GORDON, P. Frequency Effects and the representational status of regular inflections. *Journal of Memory and Language*, Elsevier, v. 40, p. 41-61, 1999. <https://doi.org/10.1006/jmla.1998.2607>.

BAAYENS, H.; DIJKSTRA, T.; SCHREUDER, R. Singulars and Plurals in Dutch: Evidence for a parallel dual route model. *Journal of Memory and Language*, Elsevier, v. 37, p. 94-117, 1997. <https://doi.org/10.1006/jmla.1997.2509>.

BAAYEN, R. H.; DAVIDSON, D. J.; BATES, D. M. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, Elsevier, v. 59, p. 390-412, 2008. <https://doi.org/10.1016/j.jml.2007.12.005>.

BAAYEN, H. R., TWEEDIE, F. J., SCHREUDER, R. The subjects as a simple random effect fallacy: Subject variability and morphological family effects in the mental lexicon, *Brain and Language*, v. 81, p. 55-65, 2002. <https://doi.org/10.1006/brln.2001.2506>.

BARR, D.J.; LEVY, R.; SCHEEPERS, C.; TILY, H. J. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, Elsevier, v. 68, p. 255-278, 2013. <https://doi.org/10.1016/j.jml.2012.11.001>.

BATES, E.; DEVESCOVI, A.; PIZZAMIGLIO, L.; D'AMICO, S.; HERNADEZ, A. Gender and lexical access in Italian. *Perception & Psychophysics*, ResearchGate, v. 57, p. 847-862, 1995.

BOWDEN, H. W. *Proficiency and second-language neurocognition: a study of Spanish as a first and second language*. 2007. 241f. Tese (Doutorado) – Georgetown University, Washington D.C., 2007.

BYBEE, J. L. *Morphology: a study of the relation between meaning and form*. Philadelphia: John Benjamins, 1985. <https://doi.org/10.1075/tsl.9>.

BYBEE, J. L. Regular Morphology and the Lexicon. *Language and Cognitive Processes*, Taylor & Francis Online, v. 10, p. 425-455, 1995.

CAFFARRA, S., SYIANOVA-CHANTURIA, A., PESCIARELLI, F., VESPIGNANI, F., CACCIARI, C. Is the noun ending a cue to grammatical gender processing? An ERP study on sentences in Italian. *Psychophysiology*, Wiley Online Library, v. 52, n. 8, p. 1019-1030, 2015.

CHOMSKY, Noam. *Lectures on government and binding*. Dordrecht: Foris, 1981.

CLAHSEN, H.; ROTHWEILER, M.; WOEST, A. Regular and irregular in the acquisition of German noun plurals. *Cognition*, Elsevier, v. 45, p. 225-255, 1992. [https://doi.org/10.1016/0010-0277\(92\)90018-D](https://doi.org/10.1016/0010-0277(92)90018-D).

CLAHSEN, H.; EISENBEISS, S.; SONNENSTUHL, I. Morphological structure and the processing of inflected words. *Theoretical Linguistics*, De Gruyter, v. 23, p. 201-249, 1997.

CORBETT, G. *Gender*. Nova York: Cambridge University Press, 1991.

DESROCHERS, A.; PAIVIO, A.; DESROCHERS, S. L'effet de la fréquence d'usage des noms inanimés et de la valeur prédictive de leur terminaison sur l'identification du genre grammatical. *Revue Canadienne de Psychologie*, Société Canadienne de Psychologie, v. 43, p. 62-73, 1989. <https://doi.org/10.1037/h0084253>.

FLEISCHHAUER, E.; CLAHSEN, H. Generating Inflected Word Forms in Real Time: Evaluating the Role of Age, Frequency, and Working Memory. In: ANNUAL BOSTON UNIVERSITY CONFERENCE ON LANGUAGE DEVELOPMENT, 36<sup>th</sup>., Somerville, 2012. *Proceedings...* Somerville: Cascadilla Press, 2012.

GOLLAN, T.; FROST, R. Two routes to grammatical gender: Evidence from Hebrew. *Journal of Psycholinguistic Research*, Springer Lind, v. 30, n. 6, p. 627-651, 2001.

GRUBBS, F. Procedures for Detecting Outlying Observations in Samples. *Technometrics*, Taylor & Francis Online, v. 11, n. 1, p. 1-21, 1969.

HERNANDEZ, A. E.; HOFMANN, J.; KOTZ, S. A. Age of acquisition modulates neural activity for both regular and irregular syntactic functions. *NeuroImage*, Elsevier, v. 36, p. 912-923, 2007. <https://doi.org/10.1016/j.neuroimage.2007.02.055>.

HERNANDEZ A. E.; KORTZ A. S.; HOFMANN J.; VALENTIN V. V., DAPRETTO M.; BOOKHEIMER S. The neural correlates of grammatical gender decisions in Spanish. *Neuroreport*, LWW Journals, v. 15, n. 5, p. 863-866, 2004.

HOHFELD, A. Accessing grammatical gender in German: the impact of gender-marking regularities. *Applied Psycholinguistics*, Cambridge University Press, v. 27, p. 127-142, 2006.

JACKENDOFF, R. What's in the lexicon?. In: NOOTEBOOM, S.; WEERMAN, F.; WIJNEN, F. (Org.). *Storage and Computation in the Language Faculty*. Dordrecht: Kluwer Academic Publishers, 2002. p. 61-92. [https://doi.org/10.1007/978-94-010-0355-1\\_2](https://doi.org/10.1007/978-94-010-0355-1_2).

LAINE, M.; VIANIO, S.; HYÖDNÄ, J. Lexical access routes to nouns in a morphologically rich language. *Journal of Memory & Language*, Elsevier, v.40, p.109-135, 1999. <https://doi.org/10.1006/jmla.1998.2615>.

MACWHINNEY, B. The competition model: the input, the context, and the brain. In: ROBINSON, P. (Org.). *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press, 2001. p. 69-90. <https://doi.org/10.1017/CBO9781139524780.005>.

MARCUS, G. F.; BRINKMANN, U.; CLAHSSEN, H.; WIESE, R.; PINKER, S. German inflection: The exception that proves the rule. *Cognitive Psychology*, Elsevier, n. 29, p. 189-256, 1995.

MORGAN-SHORT, K.; SANZ, C.; STEINHAEUER, K.; ULLMAN, M.T; second language acquisition of gender agreement in explicit and implicit training conditions: an event-related potential study. *Language Learning*, Wiley Online Library, v. 60, 1, 154-193, 2010. <https://doi.org/10.1111/j.1467-9922.2009.00554.x>.

NAME, M. C. L. *Habilidades perceptuais e linguísticas no processo de aquisição do sistema de gênero no português*. Ph.D. Dissertation. Rio de Janeiro: Pontifícia Universidade Católica do Rio de Janeiro, 2002.

NEWMAN, A. J.; ULLMAN, M. T.; PANCHEVA, R.; WALIGURA, D. L.; NEVILLE, H. J. An ERP study of regular and irregular English past tense inflection. *NeuroImage*, Elsevier, v. 34, p. 435-445, 2007. <https://doi.org/10.1016/j.neuroimage.2006.09.007>.

PADOVANNI, R.; CALANDRA-BUONAURO, G.; CACCIARI, C.; BENUZZI, F.; NICHELLI, P. Gramatical gender in the brain: Evidence from an fMRI study on Italian. *Brain Research Bulletin*, Elsevier, v. 65, p. 301-308, 2005. <https://doi.org/10.1016/j.brainresbull.2004.11.025>.

PINKER, S. Rules of Language. *Science*, American Association for the Advancement of Science, v. 253, p. 530-535, 1991. <https://doi.org/10.1126/science.1857983>.

PINKER, S. Words and Rules. *Lingua*, Elsevier, v. 106, p.219-242, 1998. [https://doi.org/10.1016/S0024-3841\(98\)00035-7](https://doi.org/10.1016/S0024-3841(98)00035-7).

PINKER, Steven. *Words and rules: the ingredients of Language*. Londres: Weidenfeld and Nicolson, 1999.

PRADO, E.; ULLMAN, M. T. Can imageability help us draw the line between storage and composition? *Journal of Experimental Psychology: Language, Memory and Cognition*, v. 35, n.4, p.849-866, 2009. <https://doi.org/10.1037/a0015286>.

PRASADA, S.; PINKER, S. Generalizations of regular and irregular morphology. *Language and Cognitive Processes*, Taylor & Francis Online, v. 8, n. 1, p. 1-56, 1993.

RASTLE, K.; DAVIS, M. Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes*, Taylor & Francis Online, v. 23, p. 942-971, 2008.

R DEVELOPMENT CORE TEAM. R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. ISBN 3-9000051-07-0, URL: <http://www-R-project.org>, 2008.

TAFT, M.; MEUNIER, F. Lexical representation of gender: A quasiregular domain. *Journal of Psycholinguistic Research*, Springer Link, v. 27, n. 1, p. 23-45, 1998.

TUCKER, G. R.; LAMBERT, W. E.; RIGAUULT, A. A.; SEGALOWITZ, N. A psychological investigation of French speaker's skill with grammatical gender. *Journal of Verbal Learning and Verbal Behavior*, ResearchGate, v. 7, p. 312-316, 1968.

ULLMAN, M.T. Acceptability ratings of regular and irregular past-tense forms: Evidence for a dual-system model of language from word frequency and phonological neighbourhood effects. *Language and Cognitive Processes*, Taylor & Francis Online, v. 1, n. 14, p. 47-67, 1999. <https://doi.org/10.1080/016909699386374>.

ULLMAN, M.T. The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, Springer Link, v. 30, n. 1, p. 37-69, 2001.

ULLMAN, M.T. Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *Cognition*, v. 92, n. 1-2, p. 231-270, 2004. <https://doi.org/10.1016/j.cognition.2003.10.008>.

ULLMAN, M.T. A cognitive neuroscience perspective on second language acquisition: the declarative/procedural model. In: SANZ, C. (Org). *Adult second language acquisition*. Washington, DC: Georgetown University Press, 2005. p. 141-178.

VAN DER LELY, H. K. J.; ULLMAN, M. T. Past tense morphology in specifically language impaired and normally developing children. *Language and Cognitive Processes*, Taylor & Francis Online, v. 16, n. 2/3, p. 177-217, 2001.

## Appendix

TABLE 1 – Stimuli list of condition 1

| Item      | Grammatical gender | Lemma Frequency<br>(noun) | Form Category |
|-----------|--------------------|---------------------------|---------------|
| __ dia    | masc               | 2119                      | Opaque        |
| __ saúde  | fem                | 320                       | Opaque        |
| __ clave  | fem                | 1                         | Opaque        |
| __ coldre | masc               | 1                         | Opaque        |
| __ covil  | masc               | 1                         | Opaque        |
| __ dogma  | masc               | 3                         | Opaque        |
| __ filme  | masc               | 597                       | Opaque        |
| __ gente  | fem                | 822                       | Opaque        |
| __ lei    | fem                | 572                       | Opaque        |
| __ limiar | masc               | 6                         | Opaque        |
| __ mão    | fem                | 905                       | Opaque        |
| __ noite  | fem                | 616                       | Opaque        |
| __ ode    | fem                | 1                         | Opaque        |
| __ país   | masc               | 1461                      | Opaque        |
| __ plebe  | fem                | 1                         | Opaque        |
| __ prole  | fem                | 3                         | Opaque        |
| __ prelo  | masc               | 2                         | Transparent   |
| __ ata    | fem                | 10                        | Transparent   |
| __ átrio  | masc               | 1                         | Transparent   |
| __ bojo   | masc               | 5                         | Transparent   |
| __ casa   | fem                | 1062                      | Transparent   |
| __ casta  | fem                | 7                         | Transparent   |
| __ fleuma | fem                | 1                         | Transparent   |
| __ hora   | masc               | 480                       | Transparent   |
| __ hulha  | fem                | 1                         | Transparent   |
| __ jogo   | masc               | 689                       | Transparent   |
| __ jugo   | masc               | 2                         | Transparent   |
| __ livro  | masc               | 648                       | Transparent   |
| __ pessoa | fem                | 385                       | Transparent   |
| __ preço  | masc               | 407                       | Transparent   |
| __ tempo  | masc               | 1515                      | Transparent   |
| __ vida   | fem                | 1059                      | Transparent   |

TABLE 2 – Stimuli list of condition 2

| Items           | Grammatical gender | Form Category | Lemma Frequency (noun) | Lemma Frequency (Adj) |
|-----------------|--------------------|---------------|------------------------|-----------------------|
| feixe rápid__   | Masc               | Opaque        | 7                      | 291                   |
| lider inat__    | Masc               | Opaque        | 440                    | 4                     |
| base sólíd__    | Fem                | Opaque        | 417                    | 46                    |
| time barat__    | Masc               | Opaque        | 652                    | 103                   |
| rede comprid__  | Fem                | Opaque        | 411                    | 28                    |
| prece calm__    | Fem                | Opaque        | 8                      | 27                    |
| fronte baix__   | Fem                | Opaque        | 15                     | 433                   |
| morte súbit__   | Fem                | Opaque        | 486                    | 34                    |
| motim louc__    | Masc               | Opaque        | 7                      | 68                    |
| série nov__     | Fem                | Opaque        | 310                    | 2370                  |
| nuance nov__    | Fem                | Opaque        | 4                      | 2370                  |
| haste comprid__ | Fem                | Opaque        | 5                      | 28                    |
| âmbar clar__    | Masc               | Opaque        | 1                      | 428                   |
| matiz escur__   | Masc               | Opaque        | 6                      | 69                    |
| papel gross__   | Masc               | Opaque        | 468                    | 52                    |
| total alt__     | Masc               | Opaque        | 410                    | 1014                  |
| vulto pret__    | Masc               | Transparent   | 17                     | 148                   |
| mundo inteir__  | Masc               | Transparent   | 1076                   | 185                   |
| Água parad__    | Fem                | Transparent   | 421                    | 43                    |
| rua escondid__  | Fem                | Transparent   | 612                    | 6                     |
| gueto escur__   | Masc               | Transparent   | 10                     | 69                    |
| fava branc__    | Fem                | Transparent   | 5                      | 308                   |
| beca antig__    | Fem                | Transparent   | 1                      | 318                   |
| banco privad__  | Masc               | Transparent   | 720                    | 299                   |
| crina long__    | Fem                | Transparent   | 1                      | 427                   |
| coisa cert__    | Fem                | Transparent   | 1222                   | 700                   |
| obra acabad__   | Fem                | Transparent   | 722                    | 5                     |
| ano letiv__     | Masc               | Transparent   | 4608                   | 6                     |
| gnomo velh__    | Masc               | Transparent   | 1                      | 414                   |
| prazo long__    | Masc               | Transparent   | 532                    | 427                   |
| pompa suntuos__ | Fem                | Transparent   | 5                      | 3                     |
| clero antig__   | Masc               | Transparent   | 4                      | 318                   |