On the representation of Portuguese gender-inflected words in the mental lexicon

Letícia M. Sicuro Corrêa,* Diogo A. de A. Almeida, and Renata Sobrino Porto

Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, RJ 22453-900, Brazil

Accepted 2 December 2003
Available online 29 January 2004

Abstract

This study aims at verifying whether Portuguese gender-inflected nouns and adjectives are represented as full forms as suggested by Spanish data (Dominguez, Cuetos, & Segui, 1999). A series of lexical decision experiments is reported. Grammatical gender, frequency dominance, and grammatical category are manipulated and cumulative frequency is controlled. The results do not provide support for a full form representation of gender-inflected words. They suggest that grammatical category, or the nature of the inflectional process involved (lexical or syntactic), affects the way words are represented and accessed. Shorter recognition latencies were obtained for nouns drawn from Feminine dominant gender-inflected pairs than from Masculine dominant pairs whereas a tendency in the opposite direction was observed in adjectives. The effect of frequency dominance appears, nevertheless, to be restricted to feminine nouns. The data are compatible with the view that masculine nouns and adjectives are represented as gender-unmarked forms. These results are discussed in relation to current dual-access models of word recognition and to the notion of "interpretability" of lexico-syntactic features, as put forward in the Minimalist Program of Generative Linguistics.

Keywords: Lexical representation; Lexical access; Gender; Dominance frequency; Interpretability; Portuguese

1. Introduction

This paper is concerned with the representation of gender-inflected words in the mental lexicon, and the recognition of Portuguese nouns and adjectives is investigated. Spanish data (Dominguez, Cuetos, & Segui, 1999) suggest that gender-inflected nouns and adjectives are represented as full forms, in contrast with number-inflected adjectives and nouns. This paper is intended to verify the extent to which this conclusion holds for Portuguese, as far as gender is concerned, since its gender system is highly similar to the Spanish one.

One of the major issues in the psycholinguistic literature concerns the access and the representation of inflected forms. Derived forms are likely to achieve a full form representation since the derivation process is not completely regular and derived words may acquire shades of meaning, that are unpredictable on the basis of the meanings of their constitutive parts. Highly frequent derived words are expected to be directly accessed and it is a consensus that irregular forms must be independently represented (Pinker, 1991). As for regular forms, a number of studies with different experimental paradigms have presented evidence for word decomposition—Sonnenstuhl, Eisenbeiss, and Clahsen (1999), on the basis of priming effects; Janssen and Penke (2002), on basis of the agreement errors of agrammatic aphasics; and Penke et al. (1997), with ERP data. Regularity, nonetheless, does not appear to be the only factor determining the way inflected forms are represented and accessed.

Baayen, Dijkstra, and Schreuder (1997) have argued that the specific properties of affixes and the word formation patterns in the language need to be taken into account in a word-identification model. They list a number of parameters that may affect morphological processing, such as productivity, semantic, and phonological transparency, frequency dominance, among others.
Frequency dominance refers to the relative frequency of the surface forms of an inflected pair. For instance, in a language with singular and plural forms, a pair of number-inflected words is singular-dominant if the singular form is more frequently used than the plural one, and it is plural-dominant, if it is the plural form that occurs more frequently. Frequency of use is a reliable predictor of the speed of the recognition of monomorphic words. For complex words, surface frequency effects can be taken as evidence for the recognition of complex words as full forms, as predicted by the Full Listing Model of word representation (Butterworth, 1983). The intricate pattern of results usually obtained in the identification of complex words makes it more likely, though, to be interpreted in the context of dual-access models (Caramazza, Laudanna, & Romani, 1988; Schreuder & Baayen, 1995; Taft, 1994).

In a study conducted in Spanish (Dominguez et al., 1999), the surface frequency of gender-inflected and number-inflected words was manipulated and the cumulative frequency of the stem (the sum of the surface frequency of each member of the inflected pair) was controlled in lexical decision experiments. The results of the experiment on gender present shorter latencies for the dominant forms irrespectively of gender. No effect of cumulative frequency was obtained in a follow-up experiment in which surface-frequency was manipulated. These results were interpreted as indicating that masculine and feminine words are independently stored in the mental lexicon. Further evidence for the independent representation of gender-inflected words has not yet been provided, though. Apart from follow-up studies by Dominguez, Cuetos, and Segui (2000), most of the research on gender-inflected words is concerned with the role of the orthographic pattern of masculine and feminine words in gender decision tasks (see Taft & Meunier, 1998, for a review).

As for number, the Spanish study obtained shorter latencies for the singular form in the singular-dominant condition though no dominant frequency effect in the plural-dominant pair was obtained. These results differ from those of Sereno and Jongman (1997), on the basis of English, in which a reverse effect of surface dominance was obtained. They also differ from Baayen et al.’s (1997) results, in which a surface frequency effect was obtained in the recognition of -en inflected plural Dutch nouns. These conflicting cross-language results can possibly be accounted for on the basis of the particular properties of the morphology of the language. The subcategorization ambiguity of the -en affix in Dutch (also a verbal affix) has, for instance, been presented as a factor that would make a parsing route particularly costly in the identification of these plural nouns (Baayen et al., 1997). If particular properties of the morphology of the language are a major predictor of dominance frequency effects, the predictions derived from the Spanish data can be tested on the basis of Portuguese data, especially when gender-inflected words are considered. The number system of Portuguese admits a number of dialectal variations and it is possible that, at least in Brazilian Portuguese, the morphological marking of number is crucially identified in the determiner. Comparing number inflected words may, therefore, add cross-language variable data to current findings before predictions can be made for cross-language differences in a principled way. The gender systems of Portuguese and Spanish are, on the contrary, highly similar, making it possible for predictions derived from results on Spanish to be directly tested in Portuguese.

The conclusions of the Spanish study raise a number of questions as far as gender is concerned. According to these data, masculine and feminine members of gender inflected pairs of nouns and adjectives have a full form representation. It is not, however, clear whether this conclusion is warranted. First of all, pairs of masculine and feminine animate and inanimate nouns were used in that study. However, gender inflection is a lexical process that appears to be restricted to animate nouns. Pairs of masculine and feminine inanimate nouns that apparently share a lexical base can hardly be taken as morphologically related in languages such as Spanish and Portuguese. Hence, it is not clear to what extent these pairs have contributed to the view that gender-inflected words are independently represented. Furthermore, the gender-inflected words in the Spanish study included both nouns and adjectives, and grammatical category was not manipulated. Gender inflection in nouns and adjectives are different morphological processes. The inflectional process that gives rise to gender inflected nouns is strictly lexical, that is, the feminine gender affix adds to animate nouns semantic information concerning a subset of the members of the class denoted by the unmarked masculine form (i.e. the subclass of female individuals). Gender inflection in adjectives, in contrast, is the morphological expression of the syntactic process of agreement. It is possible that this difference affects the way nouns and adjectives are represented and accessed. It is necessary therefore to constraint the concept of gender inflection to a regular morphological process and to manipulate grammatical category in order to verify the generality of the conclusions derived from Spanish.

In the next section, the gender morphology of Portuguese is presented in relation to Spanish. A series of lexical decision experiments is subsequently reported. Experiment 1 is similar to Dominguez et al.’s (1999) gender experiment in which Frequency dominance in a gender-inflected pair and Gender were manipulated. Rather than replicating that study, though, only [+animate] nouns were used and Grammatical category was introduced as an independent variable. The results of Experiment 1 introduce further questions, which three
follow-up studies aimed at clarifying. In particular, the following possibilities were investigated: (i) that feminine dominance in a gender-inflected pair affects the recognition of nouns regardless of gender; (ii) that the recognition of feminine adjectives is affected by the frequency of the masculine form. These results are discussed in relation to current dual-access models and in relation to a model of language, in which the interpretability of lexico-syntactic features plays a major role in the interaction between language and performance systems.

2. Gender inflection in Portuguese

Gender inflection in Portuguese, as in Spanish, can be characterized as a process whereby the gender affix –a is added to a masculine base, thereby causing alteration in meaning. When masculine nouns have an unstressed final vowel –o or –e, the vowel is suppressed as in gato/gats (cat), mestre/mestres (master/mistress). When the noun has a word final stressed vowel, -i or –u as in guril/guri (boy/girl), perul/ peru (turkey), or a consonant ending, as in professor/professora (teacher), the feminine gender morpheme is simply adjoined to the masculine base. Hence, from a morphological point of view, the masculine form can be characterized as an unmarked form whereas the feminine form is the marked one.

For historical reasons, there is a tendency for masculine nouns (animate or inanimate) to have an –o ending, and for feminine nouns (animate or inanimate) to have an –a ending, both in Portuguese and in Spanish. Gender variable adjectives, as well as participial forms, invariably have an unstressed –o ending in Portuguese and Spanish. There is, therefore, a quasi-regular phonological pattern that correlates –o ending and –a ending with masculine and feminine forms, respectively, regardless of the animacy feature of the noun. In Dominguez et al.’s study (1999), these endings were taken as masculine and feminine gender affixes. In this study, it is considered that there is a single gender affix, the feminine –a.

From a semantic point of view, the masculine noun, either singular or plural, can be underspecified for natural gender both in generic and non-specific reference. Whereas aos alunos (Portuguese) may refer to both male and female students, as alunas (fem) can only refer to female students. The addition of the gender feminine morpheme to a masculine noun restricts the class denoted by it to the subset of individuals of the feminine sex. A feminine noun may, nevertheless, incorporate additional semantic features. For instance, the Portuguese word empregada (employee, masc.) denotes the class of all possible employees, whereas the feminine form empregada-a may either refer to any female employee or to the particular class of domestic female employees. The same applies to Spanish. Hence, in both languages, gender inflection in nouns is similar to a derivational process. It is possible that the meaning alterations that result from gender inflection give rise to independent representations to feminine inflected nouns.

In both Portuguese and Spanish, there are pairs of inanimate nouns such as fruto (masc)/fruta (fem) (fruit), jarro (masc)/jarra (fem) (jug) which may be considered as gender-inflected pairs (Camara, 1970). There is, nevertheless, no evidence that the feminine form is morphologically related to the masculine one from an etymological point of view. Moreover, the lack of productivity of word formation by means of the addition of –a to inanimate nouns and the sort of semantic relationship that holds between the members of these pairs hardly justify their being considered to be gender-inflected pairs. For instance, fruta (an edible fruit with particular characteristics) is conceptually subordinate to fruto (part of the vegetable), which might be viewed as a natural gender distinction. It is difficult, nevertheless, to identity a similar conceptual relationship in pairs such canola (Portuguese), canola (Spanish) (pipe/sugar cane). In the Spanish study, these sort of inanimate masc/fem pairs constituted 30% of the feminine dominant words. It is not clear, therefore, whether the full form effect obtained can be ascribed to the presence of these pairs.

From a syntactic point of view, the gender of the noun controls agreement with syntactically related constituents. In both languages, the targets of agreement are determiners, adjectives, and participial forms. The inflectional process in adjectives is, therefore, essentially syntactic whereas it is strictly lexical in nouns. The feminine gender infix –a in nouns introduces semantic information that may affect the meaning of the whole word form. The gender infix –a in adjectives, rather than directly conveying semantic information, indicates that the word can be in grammatical agreement with a feminine noun. If, therefore, feminine nouns are independently represented due to the additional semantic features the inflected word may acquire, the same sort of independent representation would not be expected for feminine adjectives.

3. Experiment 1

This experiment was aimed at verifying whether frequency dominance determines the speed of the recognition of gender-inflected nouns and adjectives. An effect of frequency dominance independent of gender, with dominant forms being recognized more quickly, would suggest that it is the surface (inflected) form of the word rather than the morphemes that compose it that are immediately recognized. This effect would favor the hypothesis that inflected words are represented as full
forms, as suggested on the basis of the Spanish data (henceforth, *frequency dominance hypothesis*).

The same variables used in Dominguez et al. (1999), namely *Frequency Dominance in the gender-inflected pair* (Masculine Dominance (MD) and Feminine Dominance (FD)) and *Gender* (Masculine and Feminine), were manipulated here. Grammatical category was also manipulated. Half of the participants were presented to Nouns and the other half to Adjectives in a lexical decision task. The experimental conditions were the following:

- **MD Nouns/Adjectives**: Masculine Nouns/Adjectives from the MD pairs
- **Md Nouns/Adjectives**: Masculine Nouns/Adjectives from the FD pairs
- **Fd Nouns/Adjectives**: Feminine Nouns/Adjectives from the MD pairs
- **FD Nouns/Adjectives**: Feminine Nouns/Adjectives from the FD pairs

The following predictions can be made on the basis of the *frequency dominance hypothesis* and previous considerations:

If frequency dominance in a gender-inflected pair affects the recognition of nouns and adjectives regardless of gender, as suggested by the Spanish data, then a main effect of *Frequency Dominance* should be obtained, with Masculine words presenting the shortest latencies in the MD conditions and Feminine words presenting the shortest latencies in the FD conditions.

If Grammatical Category or the nature of inflectional process affects word recognition, then a *Frequency Dominance* effect is expected to be obtained in the Noun but not with Adjective Group. It is considered that Feminine nouns are more likely to present an effect of *Frequency Dominance* due to possible meaning alterations that the dominant forms may incorporate.

3.1. Method

3.1.1. Participants

Seventy-four native Brazilian Portuguese-speaker adults from PUC-Rio (Catholic University of Rio de Janeiro) community volunteered to participate as experimental subjects (48 females, 26 males). All of them had normal or corrected-to-normal sight. The participants were equally divided into two groups, according to the manipulation of Grammatical Category.

3.1.2. Stimuli

Fifty-eight pairs of gender-inflected words were used: Twenty-four nouns and twenty-four adjectives. Four experimental lists were created: Two lists of nouns and two of adjectives for the Noun and Adjective group, respectively. Each list contained twenty-four target words: twelve masculine words, six MD, and six Md, and twelve feminine words, six FD, and six Fd. A mixed design was used. Each participant of the Noun and of the Adjective group was exposed to only one member of each inflected pair. That is, each participant was exposed to each lexical base only once. The experimental lists of each group were randomly assigned to the participants by means of the computer program created for stimuli presentation. Masculine and feminine words were extracted from a corpus of written language obtained from four major newspapers for the period of a month, consisting of 1,129,365 word tokens (LAPAL database). This corpus was complemented by NURC database (Duarte, 1996), a corpus of the oral language as spoken by the educated community of Rio de Janeiro, consisting of 369,285 word tokens. The whole database consisted of about 1.5 million words. Gender-inflected words were initially classified as *noun* or *adjective* on the basis of their context of occurrence. They were then submitted to classification in isolation to one hundred educated people (undergraduate and graduated students or professionals). Only those words consistently classified as either preferentially noun or preferentially adjective by at least 90% of the judges were pre-selected to the experimental lists. 114 pairs of gender-inflected nouns and 96 pairs of gender-inflected adjectives, which had a surface frequency difference equal or higher than 3/1 million, were pre-selected. Surface frequency was defined as the mean frequency of the word in the databases equated to a 1 million-word reference base. These pairs were submitted to four groups of 75 native speakers to perform a judgment test on the relative frequency of the members of the masc–fem pairs. They had to mark, on a chart containing a list of gender-inflected nouns and adjectives, the member of the pair that they considered to be the most frequent in usual language use. A similar procedure proved reliable to access frequency on Dutch monomorphemic words, as reported by Schreuder and Baayen (1997). The results of the subjective judgment task were compared with the surface frequency of the members of the pairs. The dominance frequency judgment corresponded to actual frequency dominance for 83% of the pairs classified as nouns. As for adjectives, the subjective judgment corresponded to actual frequency dominance in only 5.21% of the pairs, giving rise to only five feminine dominant adjectives. For this reason, the experimental lists in the adjective condition included pairs whose dominance frequency was determined solely on the basis of the actual surface frequency. The mean cumulative frequency was 90.1 for nouns and 87.8 for adjectives. The mean surface frequencies were matched as much as possible: 76.7 for MD nouns, 66.7 for FD nouns, 23.4 for Md nouns, 13.5 for Fd nouns, 51.7 for MD adjectives, 49.1 for FD adjectives, 36.2 for Md adjectives, 38.6 for Fd adjectives. The mean difference of surface frequencies was 53.2 for nouns and 13.2 for adjectives (due to the overall smaller difference between the surface forms of adjectives). The
mean number of syllables in each condition was 3.1. Given that only 6 monomorphemic nouns satisfying the established criteria were identified, the morphological structure of the words (simple or complex base) was also counterbalanced. Fifty pseudo-words were included in each list. These were morphophonologically legal nouns and adjectives for the Noun and Adjective groups, respectively. The form of the pseudo-words was counterbalance as a function of their likelihood to be either masculine or feminine and the number of syllables was controlled as in the set of the target words. In addition, a list of fifteen practice items included seven words and eight pseudo-words (see Appendix A for the lists of the target words).

3.1.3. Procedure
The participants received standard lexical decision instructions and were individually tested in a noise-proof cabin. Each trial began with the presentation of a five-asterisk fixation points in the middle of the screen for 500 ms, followed immediately by the stimulus presentation centered at the same position. Stimuli remained on the screen until the subject responded to it or until 3000 ms after its onset. Following lexical decision (or time-out), a 100 ms blank was inserted between trials. Stimuli were presented on a SVGA monitor in white lowercase Arial font on a dark background and responses were collected on a button box connected to the parallel port of a PC AMD Athlon 900 MHz, located on the outside of the cabin. The experiment was generated by a computer program written in EXPE6 (Pallier, Dupoux, & Jeannin, 1997). Subjects were also instructed to respond according to their handedness. Response buttons were exchanged for half of the participants.

3.2. Results and discussion
The recognition time measures were submitted to by-subject and by-item Analyses of Variance with a 2 (gender) \( \times 2 \) (frequency dominance) \( \times 2 \) (grammatical category) factorial design in which the first two factors are repeated measures and the last one a grouping factor. A significant main effect of Grammatical Category was obtained in both analyses: \( F_1(2, 72) = 13.40 \ p < .001 \) (Means: 778.3 ms for Nouns and 655.16 ms for Adjectives), \( F_2(2, 10) = 25.01 \ p < .001 \), (Means 774.26 ms for Nouns and Adjectives 655.16). The effect of the interaction between Gender and Grammatical Category approached the significance level in the by-subject analysis \( F_1(2, 72) = 3.09 \ p < .08 \), and it was significant in the by-item analysis; \( F_2(2, 10) = 6.29 \ p = .03 \) (Means: 757.61 ms for Masc Nouns; 790.92 ms for Fem Nouns; 659.91 ms for Masc Adj; 651.41 ms for Fem Adj.). A significant effect of the interaction between Frequency Dominance and Grammatical Class was also obtained in both analyses: \( F_1(2, 72) = 10.07 \ p < .01 \) (Means: 799.62 ms for Nouns in the MD pair; 756.96 ms for Nouns in the FD pair; 646.6 ms for Adjectives in the MD pair and 663.6 ms for Adjectives in the FD pair); \( F_2(2, 10) = 7.96 \ p < .02 \). (Means: 794.08 ms for Nouns in the MD pair; 754.62 ms for Nouns in the FD pair; 646.55 ms for Adjectives in the MD pair; 663.76 ms for Adjectives in the FD pair). Table 1 presents the cell means of the by-subject analysis. The response errors were also submitted to a \( 2 \times 2 \times 2 \) ANOVA and no significant effects were obtained. The mean number of errors for actual words was 1.94 (max score = 24), showing that participants made accurate lexical decisions.

The present results do not support the frequency dominance hypothesis. The main effect of Grammatical Category and the interactions between Gender and Grammatical Category and between Frequency Dominance and Grammatical Category indicate that the grammatical category to which the word belongs (noun or adjective), or the sort of inflectional process involved (lexical or syntactic), affects the way gender-inflected words are represented and accessed.

The main effect of Grammatical Category shows shorter recognition times for adjectives than for nouns. It is possible that gender information is not computed in the recognition of adjectives since the information conveyed by the gender inflection is essentially syntactic and does not affect the meaning of the stem. The figures in Table 1 show that in the Noun Group, words drawn from masculine dominant (MD) pairs present longer latencies than words drawn from the feminine dominant (FD) pairs. As for Adjectives, a tendency in the opposite direction is observed. These differences are particularly manifested in Feminine words.

A planned comparison was carried out and the following results were obtained. The difference between FD and Fd nouns (Feminine nouns from FD pairs and from MD pairs, respectively) was significant in the direction predicted by the frequency dominance hypothesis (\( t(36df) = 2.54 \ p < .02 \) (one-tailed)); the difference between MD and Md nouns (Masculine nouns from MD and FD pairs, Table 1 Mean decision time (ms) for gender-inflected words as a function of grammatical category, frequency dominance and gender (\( n = 74 \))

<table>
<thead>
<tr>
<th>Grammatical category</th>
<th>Total mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Noun</td>
</tr>
<tr>
<td>Frequency dominance</td>
<td></td>
</tr>
<tr>
<td>in the gender-inflected pair</td>
<td>MD</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Masc.</td>
<td>777.75</td>
</tr>
<tr>
<td>Fem.</td>
<td>821.50</td>
</tr>
<tr>
<td>Total mean</td>
<td>799.63</td>
</tr>
</tbody>
</table>
respectively) nouns, in the opposite direction, approached the significance level ($t(36df) = 1.94, p = .06$ (one-tailed)). As for adjectives, the unpredicted difference between FD and Fd adjectives was significant ($t(36) = 2.87, p < .02$ (two-tailed)).

The present results are compatible with the view that it is feminine nouns that are likely to be sensitive to an effect of Frequency Dominance, possibly due to the fact that inflection is a lexical process that may give rise to alterations in meaning.

As for adjectives, the direction of the means may indicate that the unmarked masculine form guides lexical search. Notice, in this regard, that 94.8% of the subjective judgments on the dominant frequency of gender-inflected adjectives in the preparation of the stimuli favored the masculine unmarked form, regardless of the actual surface frequency of the members of the pair.

An effect of Feminine Dominance in the recognition of nouns regardless of Gender was unpredicted. Experiment 2 intended to assess the reliability of this effect.

4. Experiment 2

The effect of the interaction between Frequency Dominance and Grammatical Category in Experiment 1 raised the possibility of Feminine Dominance in a gender-inflected pair of nouns determining the speed of the recognition time regardless of gender. This experiment was intended to verify whether this effect is significant when Nouns are exclusively considered. Pairs of gender-inflected animate nouns were distinguished as a function of Frequency Dominance. Two experimental conditions were created: FD, feminine dominance, referring to gender-inflected pairs in which the feminine form is the dominant one; MD, masculine dominance, referring to gender-inflected pairs in which the masculine form is the dominant one. Each condition contained an equal number of feminine and masculine nouns. If Feminine Dominance determines recognition times, then shorter latencies should be obtained in the FD condition. A standard lexical decision task was used.

4.1. Method

4.1.1. Participants

Twenty-four native Brazilian Portuguese-speaker from undergraduate or graduate courses in PUC-Rio volunteered to participate as experimental subjects (17 females, 7 males). All of them had normal or corrected-to-normal sight.

4.1.2. Stimuli

Two lists containing twenty gender-inflected nouns from the database of written language were created. Each list included ten nouns from FD pairs, and ten nouns from MD pairs, being half masculine and half feminine nouns. The lists were created in a mixed design so that each participant was exposed to a single list and to only one of the members of the gender-inflected pairs—either the masculine or the feminine. The mean surface frequencies were 27.9 for feminine words and 7.0 for masculine words in the FD pairs, and 28.7 for masculine words and 7.3 for feminine words in the MD pairs. The word pairs were all matched for number of syllables (mean 3 in all conditions); and cumulative frequency (mean 34.8 in FD 1 and 36.0 in Fd). The resulting list was balanced for the morphological structure of the words (simple or complex). Sixty pseudo-words were also used. As in Experiment 1, the pseudo (possible masc/fem) words were morphophonologically legal (possible masculine and feminine) nouns with the number of syllable controlled and practice items included words and pseudo-words. (See Appendix B for the lists of the target words).

4.1.2. Procedure

The same as Experiment 1.

4.2. Results and discussion

The recognition time measures were submitted to by-subject and by-item one-way Analyses of Variance with repeated measures. A significant effect of Frequency Dominance was obtained: ($F_1(1,23) = 5.44, p < .03$ $F_2(1, 9) = 5.84 p < .04$). The response errors were also submitted to a one-way ANOVA, but no significant effect was obtained. The mean response error was 2.19 (max score = 24). Table 2 presents the mean recognition time in the by-subject analysis.

Words from the feminine dominant (FD) pairs had shorter recognition latencies than words from the MD pairs. Since these words included both masculine and feminine form, Feminine Dominance appears to be a factor affecting the access to nouns from gender-inflected pairs regardless of gender, as suggested by Experiment 1. FD nouns are few in the language (only 26 in the 1,500,000 database in contrast with 88 MD nouns). This unbalance may give rise to a high level of resting activation for the FD forms. It is not clear, however, the extent to which the access to Md nouns would be facilitated in relation to MD nouns. The follow-up Experiment 3 aims at clarifying this point.

<table>
<thead>
<tr>
<th>Frequency dominance in the gender-inflected pair</th>
<th>Total mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>MD</td>
<td>676.15</td>
</tr>
<tr>
<td>FD</td>
<td>643.19</td>
</tr>
<tr>
<td>659.67</td>
<td></td>
</tr>
</tbody>
</table>
5. Experiment 3

In this experiment, only masculine nouns from pairs of Masculine dominant (MD) and FD pairs were used in a lexical decision task. The Frequency dominance hypothesis would predict an effect of Frequency Dominance with MD nouns (nouns from MD pairs) presenting shorter latencies than Md nouns (nouns from the FD pairs), as suggested by the Spanish data; the Feminine dominance hypothesis would predict shorter latencies for Md nouns due to an effect of the level of activation of the feminine nouns of gender-inflected pairs.

5.1. Method

5.1.1. Participants

Thirty-five native Brazilian Portuguese-speaker adults from PUC-Rio community volunteered to participate as experimental subjects (25 females, 10 males). All of them had normal or corrected-to-normal sight.

5.1.2. Stimuli

A list of 20 nouns, including masculine and feminine words from the word pairs that satisfied the criteria of Experiment 1, was created. These nouns were taken from the written database (LAPAL database). Half of the masculine and half of the feminine nouns were distributed in two experimental conditions. Sixty pseudowords were also used. As in Experiment 1, the pseudo words were morphophonologically legal (masculine/feminine) nouns and practice items included words and pseudo-words. The mean surface frequencies were 14.3 for masculine words and 3.1 feminine words in the MD pairs, and 13.5 for feminine words and 3.6 for masculine words in the FD pairs. The word pairs were all matched for number of syllables (mean 3 in all conditions); and cumulative frequency (mean 17.4 in MD and 17.1 in FD). The resulting list was balanced for the morphological structure of the words (simple or complex) (see Appendix C for the list of the target words).

5.2. Results and discussion

The response latencies were submitted to a by-subject and a by-item one-way analysis of variance with repeated measures. No significant results were obtained $F_1(1, 34) = 1.21, p = .3; F_2(1, 9) = 0.74, p = .4$. Table 3 presents the means of the by-subject analysis.

The present results support neither the Frequency dominance nor the Feminine dominance hypotheses. They are compatible with the view that the unmarked masculine form is not affected by Frequency Dominance. This view fits with the claim that masculine forms are not gender inflected. Rather, they are likely to be represented as gender unmarked forms. The level of resting activation of FD forms does not appear to facilitate the recognition of their masculine counterparts in relation to the MD predominant forms in the language. It would at most neutralize an effect of frequency dominance in the direction predicted by the frequency dominance hypothesis.

6. Experiment 4

This experiment contrasts the recognition of gender inflected (feminine) nouns and adjectives as a function of Frequency Dominance in a within-subject design. It was intended to avoid that subjects expectations concerning the grammatical category of the words to be recognized affected the way the information conveyed by the feminine affix was processed. Grammatical Category (nouns and adjective) and Frequency Dominance (Masculine Dominance (MD) and Feminine Dominance (FD)) were then the independent variables. Four conditions were obtained.

<table>
<thead>
<tr>
<th>Frequency dominance in gender-inflected pair</th>
<th>Total mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>MD</td>
<td>772.84</td>
</tr>
<tr>
<td>FD</td>
<td>749.17</td>
</tr>
<tr>
<td>Total mean</td>
<td>761.01</td>
</tr>
</tbody>
</table>

FD Nouns: Feminine Nouns from FD pairs; Fd Nouns: Feminine Nouns from MD pairs; FD Adjectives: Feminine Adjectives from FD pairs; Fd Adjectives: Feminine Adjectives from MD pairs.

On the basis of the results of Experiment 1, a main effect of Grammatical Category is expected as well as an effect of the interaction between Grammatical Category and Frequency Dominance, with shorter latencies for FD nouns and Fd adjectives.

6.1. Method

6.1.1. Participants

70 native Brazilian Portuguese-speaker adults from PUC-Rio community volunteered to participate as experimental subjects (44 females, 26 males). All of them had normal or corrected-to-normal sight.

6.1.2. Stimuli

An experimental list containing twenty-four feminine words (twelve nouns and twelve adjectives) was created with words selected from the pre-selected pairs of the written database, following the same criteria as Experiment 1. The mean surface frequencies were 50.2 for FD nouns and 47.2 for FD adjectives, and 12.5 for Fd nouns and 18.5 for Fd adjectives. The word pairs were all matched for number of syllables (mean of 3.0), cumulative frequency (mean of 72.5 in FD nouns, 71.0 in Fd
nouns, 73.2 in FD adjectives, and 70.8 in Fd adjectives), and balanced for morphological structure. 72 morphophonologically valid pseudo-nouns and adjectives were also used. As in the other experiments, practice items included valid words and pseudo-words (see Appendix D for the list of the target words).

6.1.3. Procedure
The procedure was the same as in the previous experiments.

6.2. Results and discussion
The recognition time measures were submitted to by-subject and by-item Analyses of Variance in a 2 (grammatical class) × 2 (frequency dominance) factorial design in which the two factors are repeated measures. The main effect of Grammatical Category was not significant ($F_1(1, 69) = 2.06$, $p = .12$; $F_2(1, 5) = 1.74$, $p = .3$). A significant effect of the interaction between Grammatical Category and Frequency Dominance was obtained both in the by-subject and in the by-item analysis: ($F_1(1, 69) = 5.00$, $p < .04$); ($F_2(1, 5) = 8.02$, $p < .04$). A marginal effect of interaction was obtained between gender and dominance in the by-subject analysis: ($F_1(1, 34) = 4.18$, $p = .05$). A planned comparison was carried out and a significant difference FD nouns and Fd nouns was obtained ($t(69df) = 0.50$, $p = .01$). The response errors were also submitted to a 2 × 2 ANOVA with repeated measures and no significant effect was obtained. The error rate was 1.74 (maximum = 24). Table 4 presents the means of the by-subject analysis.

These results are compatible with the view that the nature of the inflectional process affects word recognition times. Frequency dominance affects the recognition of nouns and adjectives in the direction predicted by the results of Experiment 1 and this effect is particularly manifested in nouns. This result is compatible with the view that masculine forms are represented as unmarked forms and that gender information is not relevant for the identification of adjectives. In this experiment, the pattern of shorter latencies for adjectives than for nouns obtained in Experiment 1 was not obtained. Methodological factors may account for this difference. In the present experiment, subjects had to recognize both adjectives and nouns in the task whereas grammatical category was a grouping factor in Experiment 1. It is possible that the exclusive recognition of adjectives and pseudo-adjectives has contributed to the shorter latencies in Experiment 1. In the present experiment, word categorization may have contributed to the equalization of the means across Grammatical Class conditions.

7. General discussion and conclusion
The present results on the basis of Portuguese data do not support the conclusions derived from the study on the basis of Spanish data. Gender-inflected Portuguese nouns and adjectives are not represented as full independent forms. Only FD nouns appear to be accessed on the basis of a full form representation. Since the gender morphology of Portuguese and Spanish are highly similar, it is likely that methodological differences explain the different results. The present study, unlike the Spanish one, constrained the definition of gender inflection to nouns with an animate feature and to those elements that agree with nouns, such as adjectives. The testing material of the Spanish study included pairs of inanimate nouns whose members can hardly be morphologically related to each other. It is possible then that these pairs have contributed to the frequency dominance effect obtained. Moreover, the present results suggest that the recognition of adjectives is conducted by means of the masculine unmarked form. The inclusion of adjectives and nouns in the experimental lists of Spanish words may have, therefore, favored a tendency for shorter latencies to be obtained in MD words in the Spanish study (Dominguez et al., 1999, 2000).

The results of Experiment 1 and of its follow-ups strongly suggest that nouns and adjectives are represented and accessed in different ways. Information concerning grammatical category appears to be taken into account in the recognition of isolated words. Gender inflection in adjectives, rather than in nouns, does not affect the meaning of the stem. The strictly syntactic character of the gender inflection in adjectives is likely to account for the fact that gender does not affect the recognition of words from this grammatical category. As for nouns, the fact that FD words are few and may incorporate additional semantic features may favor their acquiring an independent representation. Experiment 1 suggested that there is a tendency for words drawn from FD pairs to be easier to recognize than words drawn from MD pairs, the reverse tendency being observed in adjectives. Experiment 2 replicated this finding for nouns though Experiment 3 showed that this effect does not imply that Md nouns are easier to recognize than MD nouns. Experiment 4 confirmed the pattern obtained in Experiment 1 though suggesting that mascu-

| Grammatical Class | Frequency dominance in der-inflected pair | Total mean |
|-------------------|----------------------------------------|
|                   | MD | FD                                    |
| Noun              | 765.26 | 724.57  | 744.92  |
| Adjective         | 756.63 | 766.38  | 765.51  |
| Mean:             | 760.95 | 745.48  |
line dominance does not make MD feminine adjectives easier to recognize than Md feminine adjectives. It appears therefore that frequency dominance does not affect the recognition of masculine nouns and adjectives. Both masculine nouns and masculine/feminine adjectives would be recognized on the basis of a gender unmarked form. How would current models of word recognition explain the present results?

The Augmented Addressed Morphology model (AAM) (Caramazza et al., 1988) presents an obligatory direct route, which is supplemented by parsing routes if it fails to identify the target representation. Different resting activation levels would account for surface frequency effects. FD nouns would have a higher resting activation level due to their additional semantic properties. Full parsing would be predicted for the feminine non-dominant nouns. This model makes the correct prediction for feminine nouns. A direct route to the unmarked form would account for the recognition of masculine nouns and adjectives, provided that the representation of adjectives contained information related to the fact that the morphological expression of gender is a mere indication of grammatical agreement and does not bear upon semantic interpretation.

The dual model discussed in Taft (1994) was intended to explain immediate prefix stripping in the recognition of written words. In this model, it is the parsing route that is mandatory, by means of which the stem is segmented. So, for words with the cumulative frequency controlled, surface effects would be accounted for in the central lexicon. It would be the combinations between stem and affixes that would give rise to surface frequency effects. If -a is considered to be the gender affix added to an unmarked base, the frequent combination between stem (unmarked base) and affix in FD words would give rise to the representation of a full form. The frequency of these combinations would additionally activate the stem (or the unmarked base), which could prime the recognition of the stem of masculine forms. This model would, therefore, account for a feminine dominance effect in nouns. This model would also predict a gender effect with longer latencies for feminine than for masculine non-dominant nouns. Assuming that the masculine form is gender unmarked, the model would not predict surface frequency effects for masculine nouns. The model would not predict a frequency effect for adjectives either, if the unmarked form could inform on the syntactic nature of the inflectional process.

According to the parallel dual route model by Schreuder and Baayen (1995), a direct route would map the access representation of the input (lexeme) with an integration node (lemma), linked to semantic and syntactic representations. This route and a parsing route would be immediately engaged in the recognition process, working in parallel. The latter would segment the input signal, license morphological combinations and compute the meaning combinations on the basis of the constitutive parts. The optimal route would provide the faster recognition response. The preferential route would be, in principle, the parsing route, which would avoid the cost of storing a form. The model is adaptive in the sense that if a parsing route proves to be costly, a full form representation is stored. Gradual adaptation would account for the immediate access to the dominant feminine forms since it would become an optimal route, once a full form representation was established. The model would also explain the overall tendency for words from FD to present faster latencies on the basis of an activation feedback system. As for the recognition of the unmarked forms, the direct route would provide faster recognition. This model would, therefore, account for the faster recognition of masculine nouns and for absence of gender effect in adjectives if, again, the identification of the unmarked form provides information bearing on the fact that gender inflection stems from a syntactic process.

In sum, current models of lexical access would account for the results obtained here provided masculine nouns are represented as unmarked forms and adjectives contain information concerning the strictly syntactic nature of the gender inflectional process. Taft’s and Schreuder & Baayen’s models would also account for the overall effect of feminine dominance in nouns, though this effect would not be enough for determining a faster recognition for masculine non-dominant forms.

The crucial differences between the Taft’s and Schreuder & Baayen’s models is that the latter takes the two routes in parallel whereas the former works serially. Taft’s accounts for surface effects in the central lexicon whereas in Schreuder & Baayen these would be by products of feedback and adaptive systems. It is not clear whether a central lexicon would be sensitive to frequency effects. These effects are apparently more compatible with low level form identification (Baayen et al., 1997). The fact that both routes work in parallel in Schreuder & Baayen’s model would explain the possibility of access to the less restrictive meaning of a FD form (i.e. the subclass of female individuals) when the more restrictive one (i.e. female individuals with a particular property) does not apply.

It appears, therefore, that current models of lexical access would account for the present results, provided that the identification of an unmarked form can inform on specific properties of grammatical gender in nouns and in adjectives. This sort of information is captured by the concept of interpretability, as a property of lexical features, which is crucial in recent generative linguistic theory.

According to the Minimalist Program (Chomsky, 1995), language is a system that includes a computational component and a lexicon, which specifies the semantic, the phonological and the formal or lexi-co-syntactic features of words. The output of a linguistic derivation (or
computation) is a linguistic expression. This expression consists of a pair (Phonetic Form and Logical Form), which constitute interfaces of the language system with the so-called performance systems. One of these levels presents instructions to be read by perceptual and articulatory systems (or equivalents), and the other, instructions to intensional/conceptual systems. These interface levels enable, therefore, a linguistic expression, computationally derived, to be perceived/comprehended in reception as well as conceptualized/articulated in production. The language computational system operates upon lexico-syntactic features by merging bundles of features (lexical items) into higher level constituents. Lexico-syntactic features can be either interpretable by performance systems at the interface or levels or uninterpretable. The latter are only useful for computation and are deleted in a linguistic derivation (or computation) before the interface levels are reached. Depending on parametric choices, the presence of some of these deleted features becomes transparent in the inflectional morphology of a particular language.

As far as the lexico-formal feature of gender is concerned, it would be interpretable in nouns but uninterpretable in the elements that agree with nouns. So, in the case of nouns, the gender affix provides interface information that can be read by conceptual systems on the basis of interface information, namely, information pertaining to natural gender. In the case of adjectives, the fact that gender is an uninterpretable feature would make gender uninterpretable at the interfaces. The information provided by the gender affix of adjectives would be a sort of residual trace of syntactic agreement, irrelevant for semantic interpretation.

If the recognition of an unmarked form can inform on the interpretability of the gender feature of the word, then the differences between the access routes taken in the recognition of feminine nouns and adjectives can be accounted for. It appears, therefore, that a model of lexical access needs to incorporate a model of language in which interpretability is a property of lexical features, in order to account for the representation of gender-inflected words.

Acknowledgments

The authors acknowledge the participation of Natália Paterson in the early stages of this study. They also thank an anonymous reviewer for the suggestion of follow-up Experiment 3, which helped to clarify the interpretation of the data.

Appendix A

Target words of Experiment 1 (the numbers denote the frequencies of the masculine or the feminine form):


Fd Nouns: cenógrafo (scenographer—0), funcionária (member of the staff—37), pomba (pigeon—5), prima (cousin—12), proprietária (owner—12), rata (mouse—0).

Fd Nouns: costureira (dressmaker—25), datilógrafa (typist—5), empregada (employee—127), esposa (wife—30), noiva (bride—76), viúva (widow—16).

**Experimental List 2:** Md Nouns: cenógrafo (scenographer—5), funcionário (member of the staff—193), pombinha (pigeon—19), primo (cousin—80), proprietário (owner—27), rato (mouse—28). Md Nouns: costureiro (dressmaker—16), datilógrafo (typist—5), empregado (employee—141), esposo (husband—0), noivo (groom—8), viúvo (widow—11).

Fd Nouns: aluna (student—74), curandeira (healer—0), fotógrafa (photographer—0), garotinha (boy/girl—96), pata (duck—14), urso (bear—6).

Fd Nouns: bruxa (witch—3), cozinheira (cooker—18), dona (owner—200), enfermeira (nurse—12), moça (young woman—217), sogra (mother-in-law—37).

**Experimental List 3:** Md Adjectives: bonito (handsome—38), esportivo (sportive—46), estranho (strange—44), famoso (famous—38), interno (internal—71), positivo (positive—56).

Md Adjectives: absoluto (absolute—22), amplo (broad—30), elétrico (electric—30), falso (false—21), profundo (deep—21), sagrado (holy—23).

Fd Adjectives: belo (beautiful—82), cheio (full—55), curto (short—66), diverso (diverse—70), moderno (modern—74), religioso (religious—60).

Fd Adjectives: aberto (open—91), antigo (ancient—98), básico (basic—57), científico (scientific—45), externo (external—48), frio (cold—35).

**Experimental List 4:** Md Adjectives: bela (beautiful—55), cheia (full—47), curta (short—44), diversa (diverse—66), moderna (modern—69), religiosa (religious—49).

Md Adjectives: aberta (open—113), antiga (ancient—106), básica (basic—63), científica (scientific—52), externa (external—64), fria (cold—76).

Fd Adjectives: bonita (beautiful—33), esportiva (sportive—30), estranha (strange—33), famosa (famous—33), interna (internal—30), positiva (positive—34).

Fd Adjectives: absoluta (absolute—30), ampla (broad—43), elétrica (electric—59), falsa (false—24), profunda (deep—36), sagrada (holy—34).

Appendix B

Target words of Experiment 2 (the numbers denote the frequencies of the masculine or the feminine form):

**Experimental List 1:** FD: aeromoça (flight attendant—2), bruxa (witch—5), cozinheira (cooker—2),
Appendix C

Target words of Experiment 3 (SM denotes the surface frequency of the masculine forms, and SD denotes the difference of surface frequency between the masculine and the feminine forms):

**Experimental List: MD:** bandido (bandit—SM = 12, FD = 11), biólogo (biologist = SM = 10, FD = 6), cenógrafo (scenographer—SM = 2, FD = 2), companheiro (fellow—SM = 22, FD = 18), escravo (slave—SM = 14, FD = 11), fotógrafo (photographer—SM = 19, DF = 19), herdeira (heirress—SM = 9, FD = 6), síndico (manager of a condominium—SM = 7, FD = 0), sobrinho (nephew—SM = 14, FD = 13), vizinho (neighbor—SM = 40, FD = 29). **FD:** aeromoço (flight attendant—SM = 0, FD = 5), bruxo (witch—FM = 1, FD = 4), cozinhheiro (cooker—SM = 2, FD = 7), dançarino (dancer—SM = 1, FD = 3), empregada (employee—SM = 15, FD = 6), enfermeiro (nurse—SM = 2, FD = 6), esposo (husband—SM = 0, FD = 24), moça (young woman—SM = 10, FD = 32), noiva (groom—SM = 15, FD = 5), viúvo (widow—SM = 3, FD = 16).

**Appendix D**

Target words used in Experiment 4 (SF denotes the surface frequency of the feminine forms, and SD denotes the difference of surface frequency between the masculine and the feminine forms):

**Experimental List: FD Nouns:** empregada (employee—SF = 21, FD = 6), esposa (wife—SF = 24, FD = 24), moça (young woman—SF = 42, FD = 32), psicóloga (psychologist—SF = 39, FD = 24), sainha (saint—SF = 156, FD = 65), viúva (widow—SF = 19, FD = 16). **Fd Nouns:** biólogo (biologist—SF = 4, FD = 6), comissária (flight attendant—SF = 0, FD = 7), cunhada (sister-in-law—SF = 4, FD = 1), escrava (slave—SF = 3, FD = 1), herdeira (heirress—SF = 3, FD = 6), ursa (bear—SF = 0, FD = 8). **FD Adjectives:** aberta (open—SF = 4, FD = 71), divina (divine—SF = 25, FD = 18), elétrica (electric—SF = 59, FD = 40), imensa (immense—SF = 39, FD = 24), monetária (monetary—SF = 44, FD = 26), pésima (terrible—SF = 22, FD = 8). **Fd Adjectives:** conclusiva (conclusive—SF = 4, FD = 4), corrupta (corrupt—SF = 4, FD = 1), decisiva (decisive—SF = 11, FD = 5), ingénua (naive—SF = 2, FD = 3), mista (mixed—SF = 8, FD = 2), nociva (noisome—SF = 2, FD = 2).

References


