Anthea Santos, Beatriz Bermúdez-Margaretto, Carlos J. Alvarez and Alberto Dominguez*

The frequency of word gender as a variable for lexical access in Spanish

https://doi.org/10.1515/ling-2020-0054
Received March 23, 2020; accepted June 25, 2021; published online April 7, 2022

Abstract: The endings of Spanish nouns reflect gender with varying degrees of frequency and regularity. The most common and regular endings are *-o* for masculine nouns and *-a* for feminine nouns, *-o* being more frequent and less closely associated with a specific gender category (masculine) than *-a*. Pairs of words occurring with both gender categories differ in the frequencies with which they are used as masculine or feminine forms: *médic-o/médic-a* 'doctor' (m.)/ 'doctor' (f.) is a clear example of a masculine-dominant pair, whereas *enfermer-o/ enfermer-a* 'nurse' (m.)/ 'nurse' (f.) is a feminine-dominant pair. Adult readers of Spanish are faster in recognizing feminine forms of feminine-dominant pairs, and masculine forms of masculine-dominant pairs (Dominguez, Alberto, Fernando Cuetos & Juan Segui. 1999. The processing of grammatical gender and number in Spanish. *Journal of Psycholinguistic Research* 28(5). 485–498). This study aims to test the dominant frequency effect in third and sixth grade children, as well as in adults. Children were faster in recognizing masculine forms in masculine-dominant pairs, but not feminine forms in feminine-dominant pairs. Adults, by contrast, tended to respond faster to higher frequency words, irrespective of gender, indicating that they have independent representations for both genders. The dominance of masculine forms in children could be a consequence of the statistical distribution of gender dominance and regularity in Spanish. The

***Corresponding author: Alberto Dominguez**, Facultad de Psicología, Universidad de La Laguna, Campus de Guajara, 38205 San Cristobal de La Laguna, Tenerife, Canary Islands, Spain, E-mail: adomin@ull.es. https://orcid.org/0000-0002-4958-0179
Anthea Santos and Carlos J. Alvarez, Departamento de Psicología Cognitiva, Facultad de Psicología, Universidad de La Laguna, Tenerife, Canary Islands, Spain; and Instituto Universitario de Neurociencia (IUNE), Universidad de La Laguna, Tenerife, Canary Islands, Spain, E-mail: asantosa@ull.edu.es (A. Santos), calvarez@ull.es (C.J. Alvarez)
Beatriz Bermúdez-Margaretto, Departamento de Psicología Básica, Psicobiología y Metodología de las ciencias del comportamiento, Facultad de Psicología, Universidad de Salamanca, Salamanca, Spain; and Instituto Universitario de Neurociencia (IUNE), Universidad de La Laguna, Tenerife, Canary Islands, Spain, E-mail: bermudezmargaretto@usal.es

experience of skilled adult readers seems to make them less dependent on this statistical pattern.

Keywords: development of morphology; dominant gender; dual route model; frequency of the stem; morphological gender

1 Introduction

The dominant frequency manipulation paradigm has been used for more than 20 years to study the question of whether whole words or morphologically decomposed forms are stored in memory for grammatical categories such as gender and number (Baayen et al. 1997b). It has been traditionally related to the major models of morphological processing. The *Full Parsing* model (Taft 2004; Taft and Forster 1975) assumes that word stems are stored in speakers' and readers' lexical memory after separation and identification of gender or number markers. The model predicts that reaction times will be influenced by the cumulative frequency of both masculine and feminine forms, that is, by the frequency of the stem and not by the frequency of each masculine or feminine form. The *Full Listing* model, on the contrary, expects a direct influence of individual frequencies of each form, because whole word forms, feminine as well as masculine, will be stored in memory (Butterworth 1983; Mannelis and Tharp 1977; Rueckl et al. 1997). Finally, the *Dual Route* model (Baayen et al. 1997b; Schreuder and Baayen 1995) predicts the application of both procedures, direct and indirect, depending on several factors, such as the regularity of inflected forms and frequency of the stem in relation to the frequency of the whole word.

Nowadays, however, the perspective taken on morphological word processing is changing. Rather than investigating the nature of representations in memory and segmentation processes, recent studies focus on how the morphological relationships between words belonging to the same morphological paradigm (i.e., the same family) determine recognition times, which is indeed a more pragmatic goal (Marelli et al. 2020). From a linguistic point of view, gender is only functional for syntactic agreement and, in many languages, is a non-canonical entity with a morphological realization only with a probabilistic assignment (Corbett and Fedden 2018). It is important to take this fact into account to understand what we consider as gender in this study since we focus on a lexical-morphological but not the syntactic gender perspective. The divergence between form classes and gender categories of morphologically related words has been called "paradigm entropy" (Baayen et al. 2007), and the assumption is that the higher the entropy, the longer the reaction times (Marelli et al. 2020). Therefore,

more important than the debate concerning the application of morphosyntactic rules is the study of the frequency of a certain root or suffix, the size of a morphological paradigm, its inflectional consistency across an entire morphological family, and consistency in the relationship between endings and gender categories in a set of morphological relatives (Milin et al. 2009). Importantly, manipulating dominant frequency to study word processing fits well within this current perspective, since it considers not only the frequency of a morphological form but also that of the contrasting form, understanding gender as a system of forces in which masculine and feminine genders participate for the recognition of each particular form.

According to the dominant frequency technique, participants are presented with pairs of masculine and feminine forms of nouns and adjectives (each one presented to different participants, to avoid repetition effects), such as *médico/médica* 'doctor' (m./f.), in which the masculine form is more frequent than the feminine one, and *enfermero/enfermera* 'nurse' (m./f.), in which the feminine form is more frequent than the masculine. Participants complete a lexical decision task with these words. The rationale is that if the whole word frequency determines reaction times, the masculine form of masculine-dominant pairs will produce shorter lexical decision times than the feminine form of the same pairs, and, inversely, the feminine form of feminine-dominant pairs will produce shorter response times than the masculine form. By contrast, no difference between masculine and feminine forms is expected if the determinant frequency is not the individual frequency of each form, masculine or feminine, but the frequency of the stem (summed frequency of masculine and feminine forms), which would be regarded as an index of a shared representation.

Dominguez et al. (1999) tested the morphological models' predictions on adult Spanish readers using the dominant frequency paradigm with a lexical decision task. They found that the masculine form was recognized faster in masculine-dominant word pairs, whereas in feminine-dominant pairs, the feminine form was recognized faster. Spanish readers, they concluded, store independent representations for feminine and masculine forms, as supported by the *Full Listing model*. In other words, the correlation of forces in terms of gender relationships does not seem to play an important role.

These results could be determined by the amount of entropy of Spanish word endings reflecting the gender of the words, as the *double route model* suggests (Baayen et al. 1997a, 1997b; Schreuder and Baayen 1995). In this sense, the more frequent endings of Spanish nouns are *-a* for feminine nouns and *-o* for masculine nouns; however, there are exceptions to this tendency: there are both words ending in *-a* with masculine gender (e.g., *tema* 'topic'), and words ending in *-o* with feminine gender (e.g., *mano* 'hand'; see also Teschner and Russel 1984).

Furthermore, the *a*/*o*-endings are not exclusive to nouns and adjectives but appear on verbs as well, for which the gender information is irrelevant. Also, many other nouns use endings such as *-on*, *-iz*, *-ad*, *-e*, etc., determining gender with different degrees of regularity.

The determination of the category of number, on the contrary, is more reliable, since the consonantal ending *-s*, and the effects of singular/plural dominant manipulation, are different from those observed for gender. The plural forms of plural dominant pairs do not exhibit faster reaction times than singular ones (Dominguez et al. 1999), whereas the singular forms of singular-dominant pairs are recognized faster than plural forms. Similar results have been obtained for German (Baayen et al. 1997b), English (Sereno and Jongman 1997), French (New et al. 2004), Italian (Baayen et al. 1997a), and Dutch (Reifegerste et al. 2016), even in production tasks (Beyersmann et al. 2015).

These differences between gender and number, possibly due to differences in the consistency of the suffixes, indicate the relevance of distributional characteristics of languages, determining one-word processing strategy or another. If the differences in regularity between gender and number determine different processing strategies in adults' word recognition, children may be even more sensitive to the observed statistical distributions than adults, as has been shown for the production of irregular past tense forms. First, the child memorizes a handful of highly frequent irregular forms and produces them correctly; then, the child is massively exposed to regular forms and begins to over-apply the rule, doing overregulation rather than regularization (Marcus et al. 1992; Marcus et al. 1995).

The case of gender seems to be different. Even though the phonological cues of word endings are not reliable to identify noun gender, Spanish children acquire gender around the age of three (Mariscal 2008; Pérez-Pereira 1991), both for regular nouns with the typical endings *-a*/*-o* and for those with atypical endings, such as *-e*, *-ión*, *-ón*, *-ú*, etc., and even for irregular words like *mapa* 'map' (m.) and *mano* 'hand' (f.). This early acquisition of gender contrasts with the acquisition of the past tense and seems to be determined by the agreement between determiners and nouns, as Mariscal (2008) points out. Evidence for this is provided by the fact that the acquisition of determiners precedes the acquisition of gender. Even so, the study reported here avoids gender agreement aspects of development to study the development of isolated orthographic/phonological representations of words in memory using a lexical decision task at three different ages.

Another factor that might affect the creation of lexical representations for regular and irregular gender forms is the distribution of nouns and adjectives

ending in *-a*/*-o*. Masculine-dominant pairs such as *médico/médica* 'doctor' (m./f.) are much more frequent than feminine-dominant pairs such as *enfermero/enfermera* 'nurse' (m./f.). Table 1 shows the number of masculine- and feminine-dominant pairs of nouns, adjectives, and participles occurring with both *a*-and *o*-endings. Type and token counts show a much larger number of masculine-dominant pairs than feminine-dominant pairs in all lexical categories represented in NIM, a large database for Spanish (Guasch et al. 2013). These results agree with the observations made by linguists such as Harris (1985, 1991) and Fuchs et al. (2015), pointing out that masculine nouns generally end in *-o*, conforming to the unmarked gender, whereas the feminine is the marked gender. The gender of nouns ending in *-a* is consistently feminine; however, a noun ending in *-o* can refer to either a masculine or a feminine word, or both if the noun is plural. Thus, the unbalanced distribution of gender frequency dominance could determine word recognition during the early stages of language and reading acquisition.

The main goal of the present study is to analyze the influence of the frequency of each form (feminine/masculine) of a stem on recognition time using a lexical decision task. We aim to investigate this question from a developmental point of view, comparing the performance of adults and children of different ages. Many studies have explored how adults represent the morphological structure of words, but less numerous are those studying the interaction with developmental factors (Amenta and Crepaldi 2012; Marelli et al. 2020 for a revision). The present study

Table 1: Type count for masculine-dominant and feminine-dominant pairs of words and token counts for masculine (*-o*) and feminine (*-a*) words.

		Masculine-dominant		Feminine-dominant	
		Masculine	**Feminine**	**Masculine**	**Feminine**
Nouns	Type		145		41
	Token	7,119.50	2,030	378	1,358.82
Adjectives	Type		882		314
	Token	11,201.40	7,056	2,543.4	5,181
Participles	Type		194		44
	Token	7,876.40	3,045.80	1,104.4	3,722.40
Adj/nouns	Type		485		132
	Token	31,088.50	3,404.21	6,454.80	13,860

All words have masculine and feminine forms. Nouns (*gat-o/a* 'cat' [m./f.]) and adjectives (*rar-o/a* 'rare' [m./f.]) were only considered if they were grammatically unambiguous. Participles are verb forms with the function of adjectives (*clavad-o/a* 'fixed' [m./f.]). Adj/Nouns belong to both classes of words (*loc-o/a* 'madman', 'madwoman'). The number of masculine-dominant forms was greater than the number of feminine-dominant forms in all categories.

was carried out with children in the third grade (aged 8–9) and sixth grade (aged 11–12) of a Spanish primary school, two age groups that are at the initial stages of reading but already able to complete a lexical decision task. A third age group included adults over 19 years old. In the experimental task, masculine and feminine nouns and adjectives of the same stem were presented for recognition (each one to different groups of participants, thus avoiding repetition priming). One subset of words was masculine-dominant (e.g., *médico/médica* 'doctor' (m./f.)) and the other half was feminine-dominant (e.g., *enfermero/enfermera* 'nurse' (m./f.)). Participants were required to perform a lexical decision based on the stimuli. This task is probably the most preferred technique to determine lexical access in both adults (Paap et al. 1987) and children (Ehri and Wilce 1983; Perfetti and Hogaboam 1975) and has been demonstrated, together with naming tasks, to measure the influence of the morphological structure of pseudo-words and study the mode of representation during single word presentation (Burani et al. 2002). Moreover, the majority of experimental studies with this paradigm have been conducted using visual word recognition, which offers a more controlled experimental situation than production or auditory comprehension. For these reasons, we also chose a visual lexical decision task, to ensure comparability with adult data and generalize our results.

Three different hypotheses can be posited for the results of this manipulation: The first is that gender irregularity will determine lexical decision times, and hence responses will be faster for the feminine if it is the dominant form (*enfermera* 'nurse' [f.]) and for the masculine, if it is the dominant one (*medico* 'doctor' [m.]). This result, previously reported for adults (Domínguez et al. 1999) would be considered evidence for separate representations of masculine and feminine forms in children according to the *Full listing model*. The second hypothesis states that children will be more sensitive to the correlation between endings and gender, and the exposure to exceptions (irregularities) would not significantly weaken the massive correlation between *a*/*o*-endings and gender. This scenario would lead to similar reaction times for masculine and feminine forms of the same stem. The third hypothesis is that the representation of gender in children will be strongly conditioned by the distribution of feminine and masculine words in the language, that is, the imbalance between a large number of masculine-dominant pairs and a smaller number of feminine-dominant pairs. In this case, masculine should be the most accessible gender in the masculine-dominant-masculine pairs, while in feminine-dominant pairs, the feminine would produce reaction times similar to those of masculine forms.

2 Method

2.1 Stimuli and design

Forty masculine-feminine pairs of words (nouns and adjectives) were selected for the study. All nouns and adjectives were taken from the *Diccionario de Frecuencias del Castellano Escrito en niños de 6 a 12 años* (Dictionary of Frequency of Spanish Written Words in Children Aged 6–12, Martínez Martín and Garcia Pérez 2004). Twenty pairs were feminine-dominant, with the feminine form having a higher frequency than the masculine form, and twenty pairs were masculine-dominant, with masculine form being more frequent than feminine form. Participants were also presented with forty pseudo-words, half of them ending in *-a* and the other half ending in *-o*. Pseudo-words were created by changing a letter of an existing word. The number of substitutions in the initial, middle, and final part of words was balanced to avoid response biases. The last letter was never changed, so if the word ended in -a or -o, the created pseudo-word would keep it as the final letter.

All stimuli categories were matched in length. The difference in length between masculine-dominant and feminine-dominant items was not significant (t (19) = −0.483, p = 0.635). *T*-tests showed no differences in frequency between masculine-dominant and feminine-dominant items (t (19) = −0.22, p = 0.83), or between masculine-non-dominant and feminine-non-dominant items (t (19) = −0.484, p = 0.634). However, significant differences were obtained, as expected, between masculine-dominant and feminine-non-dominant items (t (19) = 5.03, p < 0.000) as well as between feminine-dominant and masculine-non-dominant items (t (19) = 4,00 p = 0.001). The mean values of all controlled and manipulated variables are shown in Table 2.[1]

Even though word frequencies were calculated using a children's dictionary, the same stimuli were used for adults in a third experiment, given that we intended to test our hypothesis with the same materials across ages. The lexical frequencies of the words were also extracted from a corpus of adults, LEXESP (Sebastián-Gallés et al. 2000) in order to observe whether they differed substantially from the frequencies in the children's corpus. The correlation was sufficient (r = 0.59, p < 0.001) to support the similarity between word frequency in both corpora. As for the frequency values obtained in children, the frequency values obtained in the adult population showed no differences between

1 The experimental materials and the reaction times are available in a repository of the Universidad de La Laguna RIULL, https://riull.ull.es/xmlui/handle/915/22240 (Santos et al. 2020).

Table 2: Mean, standard deviation and quartile values of lexical frequency and length in each stimuli category.

		Masculine		Feminine		Length
		Frequency				
		Adults	Children	Adults	Children	
Masculine-dominant	Mean	147.65	113.98	63.25	41.21	6.45
	S.D.	127.98	73.84	59.67	41.39	1.27
	1st quartile	58.75	58.29	24.25	6.07	5.25
	2nd quartile	101.00	89.57	35.50	25.33	6.00
	3rd quartile	213.5	163.89	96.75	67.21	7.00
Feminine-dominant	Mean	71.60	45.99	137.45	115.58	6.70
	S.D.	45.88	36.89	135.81	96.87	1.56
	1st quartile	28.25	12.04	53.00	50.40	5.00
	2nd quartile	87.50	43.81	78.50	91.37	6.00
	3rd quartile	102.00	64.80	175.50	123.54	8.00

The categories were initially matched for frequency and length in the children's database. The adult indexes showed no important differences from those of children.

masculine-dominant and feminine-dominant items (t (19) = −0.405, p = 0.69), or between masculine-non-dominant and feminine-non-dominant items (t (19) = −0.997, p = 0.331). However, significant differences were obtained, as expected, between masculine-dominant and feminine-non-dominant items (t (19) = 3.74, p < 0.001), as well as between feminine-dominant and masculine-non-dominant items (t (19) = 4.00, p < 0.001).

The factors introduced in the experiment allow us to compare words of different gender categories with the same stem. This is a necessary control of stimuli, given that differences in reaction times can be directly attributed to word gender, while stems remain invariable, and therefore, the absence of differences will be only due to sharing the same stem.

2.2 Participants

Fifty students in the third grade of primary school (29 boys and 21 girls) aged between 8 and 9 (mean 8;8, range 8;4–9; 4) and 52 students in the sixth grade of primary school (23 boys and 29 girls) aged between 11 and 12 years old (mean 11;7, range 11;4–12;4) were selected. None of the participants was enrolled in special education or reading support programs. Their academic marks were around average for their years. All were native Spanish speakers with normal or corrected to normal vision. They took part in the experiment voluntarily after permission from their

parents was received. Students came from two different schools in Santa Cruz de Tenerife, Canary Islands, Spain: the CEIP Tomé Cano and the CEIP Chapatal.

Seventy adult undergraduate students (62 females) from the University of La Laguna (Spain) participated in this study for course credits. All had normal or corrected to normal vision. Their ages ranged from 18 to 27 years old with a mean of 19.97.

2.3 Procedure

Two lists of 10 stimuli per category were given to different participants to avoid repetition of the same stem. Thus, if a participant received the masculine form of a word (e.g., *enfermero*), other participants received the feminine form (e.g., *enfermera*). Each list was matched as far as possible in lexical frequency and length to avoid spurious effects due to the distribution of stimuli.

Participants carried out the task individually in a quiet room of their school. A lexical decision task was presented. Participants were instructed to respond as quickly and accurately as possible, answering whether the stimulus was a word or a pseudo-word. They had to press the L key on the keyboard (labeled as "yes") with the index or middle finger of their right hand if a word appeared on the screen, and to press the S key (labeled as "no") with the index or middle finger of their left hand if a pseudo-word was displayed. Stimuli were presented using E-prime 2.0 software (Schneider et al. 2002). Both latency and accuracy indices were recorded.

The experiment started after participants completed 10 training items, and once the researcher ensured the correct understanding of the instructions. The sequence of stimuli presentation was as follows. First, an asterisk was presented as a fixation point for 1,000 ms; then, the target word was presented and remained on the screen until the participant's response. Stimuli were randomized over participants and presented in white uppercase Courier font on a dark background in the center of the screen, with a 70 Hz refresh rate. Each character covered approximately 0.38° of visual angle from a distance of 60 cm. The same sequence and display settings were used in both experiments.

3 Results

The experiment included two within-participant variables: Gender (masculine and feminine) and dominance (dominant and not dominant) and one between-participant factor, group of age (third grade 8-year-old participants, sixth grade 11-year-old participants and undergraduate students older than 18).

Analyses of RT were carried out only for correct responses. Responses exceeding 2.5 standard deviations above or below the mean for each participant were replaced by that cutoff value (2.8% of responses, including the errors). Table 3 shows mean reaction times as well as error rates (in parentheses) for each experimental condition.

All the analyses for both reaction times (RTs) and error rates were analyzed using linear mixed-effect modeling (Baayen 2008; Baayen et al. 2008), as implemented in the lme4 package (Bates and Maechler 2009) in the statistical software R 3.2. In particular, we used mixed-design analyses of variances (ANOVAs) with Satterthwaite approximation for degrees of freedom. The model was estimated following Barr et al. (2013) with the two repeated measure factors as fixed and random slopes by participant. We selected the specific model following the Barr et al. (2013) approach, in which the authors argue that linear mixed-effects models generalize best when a maximal random effects structure is included, and it is justified by the design. For this reason, no other models were compared. In addition, a random intercept was included in the tested model.

Contrast coding was used to estimate fixed effects of the model. Once estimates and significant effects were obtained, post-hoc comparisons were carried out by estimating the marginal means (least-squares means) of the model, as well as their standard errors. Relevant post-hoc comparisons of simple effects were conducted on the means and standard errors, with the emmeans R package (Lenth et al. 2020). This approach makes it is unnecessary to interpret the meaning of the intercept value in order to reconstruct the means from the parameters of the estimated model. All comparisons were corrected applying Hochberg's Type I error correction (Benjamini and Hochberg 1995; Hochberg 1988).

Results for RTs showed that the factor group was significant (F (2, 172) = 102.9, $p < 0.001$), indicating that third grade participants were slower than sixth grade

Table 3: Mean reaction times (and mean % of errors) obtained in Experiment 1.

	Masculine-dominant		**Feminine-dominant**	
	Masculine *médico* (doctor m.)	**Feminine *médica* (doctor f.)**	**Masculine *enfermero* (nurse m.)**	**Feminine *enfermera* (nurse f.)**
3rd grade	1,389 (2.6)	1,492 (5.4)	1,395 (4.4)	1,393 (4.8)
6th grade	1,148 (1)	1,217 (4.9)	1,193 (1)	1,167 (1.4)
Adults	615 (2.2)	639 (3.2)	624 (2.2)	615 (1.6)

participants (post-hoc analysis, t (171.9) = 3.7, p < 0.001) and adults (t (171.9) = 13.6); sixth graders were also slower than adults (t (171.8) = 9.7, p < 0.001). As far as gender is concerned, words of the -o-class were recognized faster than words of the -a-class (F (1, 177) = 9.47, p < 0.005). Dominance also reached significance, with dominant words producing faster RTs than non-dominant (F (1, 6,191) = 26.6, p < 0.001). Two interactions also yielded significance, gender × dominance (F (1, 6,266) = 26.63, p < 0.001) and group × dominance (F (2, 1,397) = 3.04, p < 0.05). Gender × dominance showed that in the masculine-dominant words, RTs were shorter for masculine words than for feminine words (t (634.8) = 5.7, p < 0.001), whereas the difference in gender was not significant for feminine-dominant words, t (617.42) = 1.33, p = 0.18. Group × dominance indicated that masculine-dominant pairs were slower than feminine-dominant pairs only in the third graders (t (164.4) = 2.3, p < 0.05), whereas no differences appeared between masculine-dominant and feminine-dominant items in sixth grade (t (157.15) = 0.442, p = 0.65 or in adults, t (127.30) = 0.64, p = 0.52) (see Figure 1).

To examine our results more carefully, we decided to analyze each group of participants separately, with the same two within-participants factors. In the case of third graders, there was a significant effect of gender (F (1, 50) = 5.06, p < 0.05), with shorter RTs for masculine words, and gender by dominance interaction (F (1, 1,760) = 8.42, p < 0.005), showing that the masculine of the masculine-dominant words produced shorter RTs (1,389 ms) than the feminine form (1,492 ms, t (156) = 3.61, p < 0.001, difference of 103 ms), whereas the gender difference (2 ms) was not significant for feminine-dominant words (1,393 and 1,395 ms, respectively (t < 1)).

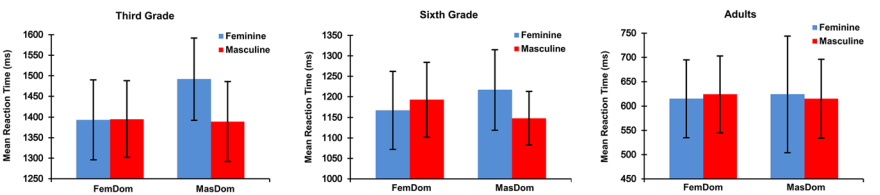

Figure 1: Evolution of the lexical decision on the words with dominant feminine gender and dominant masculine gender from third grade (left) and sixth grade (center) to the adult group (right). Note that the difference between masculine and feminine in the feminine-dominant is inexistent in the youngest children (left) but increases with age. On the contrary, the difference between masculine and feminine in masculine-dominant pairs seems to be maintained across the age groups.

For sixth graders, the same gender by dominance interaction (and no main effect) reached significance (F (1, 1,836) = 7.12, $p < 0.01$), showing the same pattern as in third graders. The difference (69 ms) between masculine (1,148) and feminine (1,217) in masculine-dominant pairs was significant (t (483) = 2.69, $p < 0.001$) whereas the difference in gender (1,193 ms for masculine and 1,168 ms for feminine words, difference of 25 ms) for feminine-dominant words did not reach statistical significance, t (462) < 1, although it was higher than in the third grade group.

In adults, only the interaction gender × dominance turned out to be significant (F (1, 2,591) = 12.26, $p < 0.001$). However, post-hoc analyzes did not show differences between genders in each dominance level.

The same statistical procedure and model were applied to analyze error rates but using the mixed model with logit family link function for binomial data. Only the factor group was significant ($\chi^2(2) = 14.85$, $p < 0.001$). In particular, third graders made more errors than sixth graders (post-hoc, $z = 3.6$, $p < 0.001$) and adults ($z = 3.3$, $p < 0.005$).

4 Discussion

The present study aims to investigate the influence of the frequency of each form (feminine/masculine) of a stem word. This question was addressed within a developmental perspective, comparing the performance of adults and children of different ages using the dominant frequency paradigm.

In general, our results show that masculine gender words were recognized faster than feminine words, indicating that mental representation of masculine forms allows faster recognition. Words with the dominant gender were also recognized faster, implying that the whole word frequency plays an important role in lexical access. The interaction between gender and dominance supports a different role of masculine and feminine gender. Masculine and feminine versions of the same words differ in the masculine-dominant pairs but not in the feminine-dominant pairs. The higher frequency of the feminine word *enfermera* 'nurse' (f.) over the masculine form, *enfermero* 'nurse' (m.) does not involve an advantage reflected in faster recognition. By contrast, when the higher frequency form corresponds to the masculine form, as in *médico* 'doctor' (m.), its recognition is faster than that of the feminine form *médica* 'doctor' (f.).

Importantly, the analyses carried out for each age group show differences between the two groups of children in comparison to the adult group. These results could support differences between the representation of masculine and feminine forms. Masculine forms are recognized quickly regardless of their lexical frequency; however, feminine forms only reach the latency responses of masculine

forms when they are frequent whereas their recognition is delayed when their frequency is non-dominant.

RTs obtained for sixth grade children, similar to the youngest participants, however, deserve some discussion. Although non-significant, a tendency toward faster responses for feminine forms of feminine-dominant items than for masculine forms of feminine-dominant items (25 ms) is observed, which would indicate a strengthening of feminine word representations that is later confirmed in adults in the cross-interaction. Obviously, sixth grade children were overall faster and more accurate in their responses than third grade children, which implies more efficient and better-developed general processes of reading.

A different picture emerges from the analyses of adults' RTs (see Figure 1). The cross-interaction is in the same direction as that previously obtained for adults by Dominguez et al. (1999). In the masculine-dominant pairs, masculine words produced shorter RTs (616 ms) than feminine ones (640 ms), whereas the opposite pattern was observed for feminine-dominant pairs: 624 ms for masculine words and 615 ms for feminine words. This result should be considered with caution since post-hoc analyses did not show differences between masculine and feminine, either in the masculine-dominant pairs or in the feminine-dominant pairs.

Overall, these results suggest that children show a bias toward the representation of masculine forms over feminine forms, which determines the lexical access for masculine and feminine forms of words. Adults, however, tend to balance the importance of feminine and masculine forms and do not assign a different status to one gender over the other, which is consistent with previous research (Dominguez et al. 1999). The developmental trend found in this study is in line with previous observations about regularity and participants' age during visual recognition of singular and plural words. For instance, Reifegerste et al. (2016) showed a different pattern of results when singular and plural dominant words were evaluated in younger and older readers of Dutch and German.

The interpretation of the present results in terms of classical models of morphological processing supports assumptions made by the *Dual Route model* regarding gender access and representation (Baayen et al. 1997b; Schreuder and Baayen 1995). According to this view, masculine words are accessed directly and represented as whole forms, whereas lexical representations of feminine words are accessed through their masculine representations. Although this study only intends to explore the formal representation of words in memory based on their gender, interactions between such formal representations and other more semantic aspects of gender must be also considered. In this sense, a plausible hypothesis for the dominant role of masculine forms in gender representation is that children could be more sensitive to gender as a semantic feature, as pointed

out by Harris (1985, 1991) and Fuchs et al. (2015); masculine nouns generally end in *-o*, the unmarked gender, whereas the feminine is the marked gender. Although with some exceptions, the gender of a noun ending in *-a* is consistently feminine; however, a noun ending in *-o*, can refer to either a masculine or a feminine word, or both if the noun is plural. For example, sentence (1) shows how a professional role such as *medicos* 'doctors', which includes a final masculine mark *-o* and a plural suffix *-s*, may refer equally to men and women. In a similar way, sentences with two subjects as in (2) may be followed by the anaphoric pronoun *ell-o-s* 'they', with a masculine ending *-o*, reflecting referents of both genders, even if one of the subjects is a woman.

(1) *L**o**s médic**o**s salvan vidas*
'Doctors save lives'
(2) *Juan y María fueron al cine; ell**o**s compraron palomitas*
'John and Mary went to the cinema; they (m) bought popcorn'

This indetermination of the final *-o* of words with a masculine referent, in the case of semantic biological gender, could lead children to generalize and use masculine forms as the basic lexical entry in the lexicon, representing masculine as a more general entity than feminine. This could explain why a feminine-dominant word does not benefit from a higher frequency when it is being processed.

Studies into children's gender acquisition have shown outcomes both in favor and against the unmarked gender hypothesis. Differences between the marked and unmarked gender were supported by Pérez-Pereira's (1991) data on Spanish word production in children. This author found a significant tendency of children aged 4–11 years old to use more masculine than feminine adjectives in agreement with non-words. Against this argument, the data by Hernández Pina (1984) found errors in spontaneous speech such as *la mot**a** rot**a*** instead of the correct exception *la mot**o** rota* 'the broken motorcycle': An irregular feminine word such as *moto* 'motorcycle' (f.) was regularized so that it agrees with the feminine determiner *la* in children between 1;09 and 2;01. Also, Mariscal (2008) found errors such as *nene mala* 'bad boy' in which the masculine noun *nene* 'boy' exhibits an agreement mismatch with the feminine adjective *mala* 'bad' (f.). In this case, the author states that children tend to learn their first adjectives in packages with specific nouns; for example, '*mala*' is an adjective typically matched with the noun *bruja* 'witch' (f.) in fairy tales. Therefore, '*mala*' will be used more frequently than *malo* 'bad' (m.), irrespective of the noun it combines with.

The field of visual word recognition also reports controversial results for the unmarked gender hypothesis. No differences between marked (feminine) and

unmarked (masculine) gender were found in adults (Dominguez et al. 1999); however, Alemán-Bañón and Rothman (2016) found differences between feminine and masculine in gender agreement, a morphological process depending more on syntactic comprehension. In an ERP study also conducted with adult participants, these authors found that the Left Anterior Negativity between 250 and 450 ms emerged only for feminine agreement violations of an adjective, e.g., in *Laura lavó un uniforme que parecía sucia* 'Laura washed a uniform (m.) that appeared dirty (f.)', but not for the unmarked masculine. However, these results should be treated with caution. Indeed, the authors acknowledge this effect as an atypical LAN, since it was a sustained overlap with the late negativity in the 500–1,000 ms windows. In the same vein, P600, a later component, emerged earlier for gender violations in feminine adjectives than in masculine adjectives. However, Barber and Carreiras (2005), using a similar syntactic paradigm, did not find a differential influence over these ERP components.

Although studies of the unmarked gender hypothesis have been inconclusive, Mariscal's (2008) work supports the importance of children's sensitivity to the characteristics of those words to which they are most exposed. The finding that visual access during visual word recognition is faster for masculine than for feminine words in young readers could be related to the distributional properties of the two genders in the language, as shown in Table 1. The statistical distribution of gender dominance suggests that children are more exposed to masculine-dominant pairs, and this predominant exposure to masculine forms may influence them in recognizing feminine forms through the corresponding masculine. Children are statistical collectors of regularities in their first years of experience with language. Thus, independent representations for feminine and masculine, marked and unmarked forms, would develop only after years of reading experience, as well as after repeated exposure to many different endings (*-a*, *-o*, *-z*, *-ón*, *-ú*, *-e*, *-ión*, *-ad*, etc.) and to the irregularities of words with only one gender, inconsistent endings (*mano* 'hand' [f.]), or ambiguous endings (*andando* 'walking'). This repeated exposure could increase the accessibility of irregular words, creating independent representations for feminine and masculine forms. The tendency of older children to respond faster to feminine-dominant words than to masculine-non-dominant words suggests an increase of the feminine-independent representations that are found later in adults.

About semantic gender, it must be noted that most of the stimuli used in our experiments define semantic gender, and more specifically, stereotypical gender. It has been found that stereotypical gender information is incorporated into the mental representation of the word assuming that, for example, *doctors* are represented as men and *nurses* are represented as women. This stereotypical

trait produces a direct effect on response latencies and amplitude of electrophysiological components, as evidenced by Cacciari and Padovani (2007) or Siyanova-Chanturia et al. (2012). The results obtained in these studies suggest that there is an asymmetry in morphological processing since male and female stereotypes affect processing differently: it is easier to pair a male stereotype with a female pronoun (e.g., *doctor – she*) than vice versa (e.g. *nurse – he*). The effects of stereotype incongruity of gender also occur in children, although the interaction effect has the opposite direction of that obtained with adult participants (Siyanova-Chanturia et al. 2015). Markedness and stereotypical gender seem to be plausible semantic aspects that can contribute to the access and representation of the gender paradigm. Therefore, the interaction between both variables should be addressed in future studies, disentangling the relative contribution of each factor to word recognition.

In summary, the present study suggests that word recognition is determined by the characteristics of the morphological paradigm to which the word belongs (Marelli et al. 2020). In this case, the gender paradigm, which is characterized by two decisive aspects: the first one, a certain degree of inconsistency in the relationship between the orthographic endings of words and their genders (Baayen et al. 2007; Marelli et al. 2015), and second, an imbalance in the relationship between masculine frequency, whose dominance is greater, and feminine gender, which is less frequently dominant. These factors, taken from a developmental perspective, seem to be more influential at the youngest ages.

Although tentative, the conclusions reached in this study are valuable for future research. This study tests access to independent representations of words with different gender categories. However, our results could be very different if gender access was measured in the presence of a determiner, a question that deserves further examination in future studies. Furthermore, it would be worthwhile to know more about the effects of frequency dominance in spoken language, which could be studied at even earlier ages. That paradigm could help to clarify the presence of U-shaped development effects in gender processing, as previously reported for the past tense studies.

Research funding: This research was supported by Ministerio de Ciencia e Innovación, Spain project PID2020-114246GB-I00 “Orthographic learning in a second language. Cross-linguistic and sensory-motor factors” and it was also carried out within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE University) in 2021. The first author was supported by an Ministerio de Universidades, Spain FPU grant FPU16/02523 from the Spanish Ministry of Education.

References

Alemán-Bañón, José & Jason Rothman. 2016. The role of morphological markedness in the processing of number and gender agreement in Spanish: An event-related potential investigation. *Language, Cognition and Neuroscience* 31(10). 1273–1298.

Amenta, Simona & Davide Crepaldi. 2012. Morphological processing as we know it: An analytical review of morphological effects in visual word identification. *Frontiers in Psychology* 12(3). 232.

Baayen, R. Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Baayen, R. Harald, Cristina Burani & Robert Schreder. 1997a. Effects of semantic markedness in the processing of regular nominal singulars and plurals in Italian. In Geert E. Booij & Jaap van Marle (eds.), *Yearbook of morphology 1996*, 13–34. Dordrecht: Kluwer.

Baayen, R. Harald, J. Davidson Doug & M. Bates Douglas. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59. 390–412.

Baayen, R. Harald, Ton Dijkstra & Robert Schreuder. 1997b. Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language* 37. 94–117.

Baayen, R. Harald, H. Wurm Lee & Joanna Aycock. 2007. Lexical dynamics for low-frequency complex words: A regression study across tasks and modalities. *The Mental Lexicon* 2(3). 419–463.

Barber, Horacio & Manuel Carreiras. 2005. Grammatical gender and number agreement in Spanish: An ERP Comparison. *Journal of Cognitive Neuroscience* 17. 137–153.

Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. TilySimilarly. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68. 255–278.

Bates, Douglas & Martin Maechler. 2009. *Lme 4: Linear mixed-effects models using S4 classes*. R package version 0.999375-27.

Benjamini, Yoav & Hochberg Yosef. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 57(1). 289–300.

Beyersmann, Elisabeth, Eleanor M. Dutton, Sohaila Amer, Niels O. Schiller & Britta Biedermann. 2015. The production of singular- and plural-dominant nouns in Dutch. *Language, Cognition and Neuroscience* 30(7). 867–876.

Burani, Cristina, Stefania Marcolini & Giacomo Stella. 2002. How early does morphological reading develop in readers of a shallow orthography? *Brain and Language* 81. 568–586.

Butterworth, Brian. 1983. Lexical representation. In Brian Butterworth (ed.), *Language production: Development, writing and other language processes*, vol. 2, 257–294. London: Academic Press.

Cacciari, Cristina & Roberto Padovani. 2007. Further evidence of gender stereotype priming in language: Semantic facilitation and inhibition in Italian role nouns. *Applied Psycholinguistics* 28. 277–293.

Corbett, Greville, G. & Sebastian Fedden. 2018. New approaches to the typology of gender. In Sebastian Fedden, Jenny Audring & Greville Corbet, G. (eds.), *Non-canonical gender systems*, 9–35. Oxford: Oxford University Press.

Dominguez, Alberto, Fernando Cuetos & Juan Segui. 1999. The processing of grammatical gender and number in Spanish. *Journal of Psycholinguistic Research* 28(5). 485–498.

Ehri, Linnea Carlson & Lee S. Wilce. 1983. Development of word identification speed in skilled and less-skilled beginning readers. *Journal of Educational Psychology* 75. 3–18.

Fuchs, Zuzanna, Gregory Scontras & Maria Polinsky. 2015. The differential representation of number and gender in Spanish. *The Linguistic Review* 32. 703–737.

Guasch, Marc, Boada Roger, Pilar Ferré & Rosa Sánchez-Casas. 2013. NIM: A Web-based Swiss Army knife to select stimuli for psycholinguistic studies. *Behavior Research Methods* 45. 765–771.

Harris, James Wesley. 1985. Spanish word markers. In Frank H. Nuessel (ed.), *Current issues in Spanish phonology and morphology*. Bloomington, IN: Indiana University Linguistics Club.

Harris, James Wesley. 1991. The exponence of gender in Spanish. *Linguistic Inquiry* 22(1). 27–62.

Hernández Pina, Fuensanta. 1984. *Teorías psicosociolingüísticas y su aplicación a la adquisición del español como lengua materna*. [Psychosociolinguistic theories and their application to the acquisition of Spanish as a mother tongue]. Madrid: Siglo XXI.

Hochberg, Yosef. 1988. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* 75. 800–802.

Lenth, Russell, Henrik Singmann, Jonathon Love, Paul Buerkner & Maxime Herve. 2020. Emmeans: Estimated marginal means. R package version 1.4. 4.

Mannelis, León & David Tharp. 1977. The processing of affixed words. *Memory & Cognition* 5. 690–695.

Marcus, Gary, Úrsula Brinkmann, Harald Clahsen, Richard Wiese & Steven Pinker. 1995. German inflection: The exception that proves the rule. *Cognitive Psychology* 29(3). 189–256.

Marcus, Gary F., Steven Pinker, Michael Ullman, Michelle Hollander, John Rosen, Fei Xu & Harald Clahsen. 1992. Overregularization in language acquisition. *Monographs of the Society for Research in Child Development* 57(4). 1–178.

Marelli, Marco, Simona Amenta & Davide Crepaldi. 2015. Semantic transparency in free stems: The effect of orthography-semantics consistency on word recognition. *Quarterly Journal of Experimental Psychology* 68(8). 1571–1583.

Marelli, Marco, Daniela Traficante & Cristina Burani. 2020. Reading morphologically complex words: Experimental evidence and learning models. In Vito Pirrelli, Ingo Plag & Wolfgang Dressler (eds.), *Word knowledge and word usage. A cross-disciplinary guide to mental lexicon*, 553–592. Berlin & Boston: De Gruyter Mouton.

Mariscal, Sonia. 2008. Early acquisition of gender agreement in the Spanish noun phrase: Starting small. *Journal of Child Language* 35. 1–29.

Martínez Martín, Jesús Antonio & Emma García Pérez. 2004. *Diccionario de frecuencias del castellano escrito en niños de 6 a 12 años*. [Dictionary of frequencies of Spanish written by children from 6 to 12 years old]. Salamanca: Servicio de Publicaciones de la Universidad Pontificia de Salamanca.

Milin, Petar, Durdević Dusica Filipović & Martín Fermín Moscoso del Prado. 2009. The simultaneous effects of inflectional paradigms and classes on lexical recognition: Evidence from Serbian. *Journal of Memory and Language* 60. 50–64.

New, Boris, Marc Brysbaert, Juan Seguí, Ludovic Ferrand & Kathleen Rastle. 2004. The processing of singular and plural nouns in French and English. *Journal of Memory and Language* 51. 568–585.

Paap, Kenneth R., James E. McDonald, Roger W. Schvaneveldt & Ronald W. Noel. 1987. Frequency and pronounceability in visually presented naming and lexical decision tasks. In

Max Coltheart (ed.), *Attention and performance XII: The psychology of reading*. Mahwah, NJ: Lawrence Erlbaum.

Perfetti, Charles & Thomas Hogaboam. 1975. Relationship between single word decoding and reading comprehension skill. *Journal of Educational Psychology* 4. 461–489.

Pérez-Pereira, Miguel. 1991. The acquisition of gender: What Spanish children tell us. *Journal of Child Language* 18(3). 571–590.

Reifegerste, Jana, Antje S. Meyer & Pienie Zwitserlood. 2016. Inflectional complexity and experience affect plural processing in younger and older readers of Dutch and German. *Language, Cognition and Neuroscience* 32(4). 471–487.

Rueckl, Jay G., Mikolinski Michelle, Michal Raveh, Caroline S. Miner & Frank Mars. 1997. Morphological priming, fragment completion, and connectionist networks. *Journal of Memory and Language* 36. 382–405.

Santos, Anthea, Beatriz Bermúdez-Margaretto, Carlos Javier Álvarez & Alberto Domínguez. 2020. Hoja de cálculo anexa al artículo "The frequency of word gender as a variable for lexical access in Spanish". Universidad de La Laguna. Available at: http://riull.ull.es/xmlui/handle/915/22240.

Sebastián-Gallés, Nuria, Fernando Cuetos, Manuel Carreiras & Antonia Marti. 2000. LEXESP. *Léxico informatizado del español*. [Spanish computerized lexicon]. Barcelona: Universitat de Barcelona Edicions.

Schneider, Walter, Amy Eschman & Anthony Zuccolotto. 2002. *E-prime user's guide*. Pittsburgh: Psychology Software Tools Inc.

Schreuder, Robert & R. Harald Baayen. 1995. Modeling morphological processing. In Lauri B. Feldman (ed.), *Morphological aspects of language processing*, 131–154. Mahwah, NJ: Lawrence Erlbaum.

Sereno, Joan & Jongman Allard. 1997. Processing of English inflectional morphology. *Memory & Cognition* 25(4). 425–437.

Siyanova-Chanturia, Anna, Francesca Pesciarelli & Cristina Cacciari. 2012. The electrophysiological underpinnings of processing gender stereotypes in language. *PLoS One* 7(12). e48712.

Siyanova-Chanturia, Anna, Paul Warren, Francesca Pesciarelli & Cristina Cacciari. 2015. Gender stereotypes across the ages: On-line processing in school-age children, young and older adults. *Frontiers in Psychology* 6. Art no. 1388.

Taft, Marcus. 2004. Morphological decomposition and the reverse base frequency effect. *The Quarterly Journal of Experimental Psychology* 57(4). 745–765.

Taft, Marcus & Kenneth Forster. 1975. Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior* 14. 638–647.

Teschner, Richard V. & William M. Russell. 1984. The gender patterns of Spanish nouns: An inverse dictionary-based analysis. *Hispanic Linguistics* 1. 115–132.