Effects of the psycholinguistic variables on the lexical decision task in Spanish: A study with 2,765 words

**María González-Nosti · Analía Barbón ·
Javier Rodríguez-Ferreiro · Fernando Cuetos**

Abstract In order to explore the role of the main psycholinguistic variables on visual word recognition, several mega-studies have been conducted in English in recent years. Nevertheless, because the effects of these variables depend on the regularity of the orthographic system, studies must also be done in other languages with different characteristics. The goal of this work was to conduct a lexical decision study in Spanish, a language with a shallow orthography and a high number of words. The influence of psycholinguistic variables on latencies corresponding to 2,765 words was assessed by means of linear mixed-effects modeling. The results show that some variables, such as frequency or age of acquisition, have significant effects on reaction times regardless of the type of words used. Other variables, such as orthographic neighborhood or imageability, were significant only in specific groups of words. Our results highlight the importance of taking into account the peculiarities of each spelling system in the development of reading models.

Keywords Spanish · Lexical decision · Age of acquisition · Word frequency · Length

M. González-Nosti · F. Cuetos
Department of Psychology, University of Oviedo, Oviedo, Spain

A. Barbón
Department of Experimental Psychology, University of Granada, Granada, Spain

J. Rodríguez-Ferreiro
Department of Basic Psychology, University of Barcelona, Barcelona, Spain

M. González-Nosti (✉)
Facultad de Psicología, Pza. Feijoo,
s/n, 33003 Oviedo, Spain
e-mail: gonzaleznmaria@uniovi.es

Background

Visual word recognition is undoubtedly a basic component of reading and one of the most investigated processes in psycholinguistics. As a result of the interest in this issue, several computational models of reading have been developed (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Plaut, McClelland, Seidenberg, & Patterson, 1996). These models have been useful in understanding reading acquisition and dyslexic disorders caused by brain damage (that is, acquired dyslexia).

Until recently, most studies of visual word recognition, and the models constructed from those studies, were conducted in English. In the last few years, there has been an increase in the number of investigations conducted in other languages, due to the finding that the properties of writing systems influence visual word recognition (Frost, Katz, & Bentin, 1987). In languages with deep orthographies, such as English, wherein there are irregular words that do not conform to grapheme–phoneme rules, readers may use processes different from those used by readers of transparent languages, in which all the words conform to the rules. The most popular model of reading, the dual-route model developed by Coltheart (1981), proposed two ways of reading words: the lexical route, which allows reading irregular familiar words as they are stored in a lexicon or mental dictionary, and the sublexical route, which involves translating the graphemes into phonemes, which is useful for reading unknown words. In transparent languages, the sublexical route would be sufficient to read any word (Ziegler & Goswami, 2005). However, numerous experiments conducted in languages with transparent orthographic systems suggest the existence of a lexical reading in these languages as well, although perhaps less used than in languages with deep orthographies. Frost et al. (1987) found that lexical frequency affected the lexical decision task in Serbo-Croatian, a language with a shallow orthography.

However, the word-naming task was hardly affected by this lexical variable. Similarly, Burani, Arduino, and Barca (2007) found that in Italian, the word naming task was affected by lexical variables, while the lexical decision was affected by both lexical and semantic variables. In Spanish, Cuetos and Barbón (2006) made a regression analysis and found that the age of acquisition (AoA) was one of the best predictors of reading times.

Many experiments have also been carried out in Spanish with the lexical decision and reading-aloud methodologies. As is normally seen in English, these experiments have revealed effects of the main lexical and sublexical variables, like lexical frequency, AoA, imageability (Alija & Cuetos, 2006), and orthographic neighborhood (Carreiras, Perea, & Grainger, 1997). Effects of sublexical variables, such as number of letters and syllables (Acha & Perea, 2008) or syllable frequency, have also been found (Carreiras, Álvarez, & de Vega, 1993).

All these experiments have been carried out using factorial designs, in which certain variables are manipulated and the rest remain controlled. Nevertheless, some problems with factorial designs have been detected in recent years. Balota, Cortese, Sergent-Marshall, Spieler, and Yap (2004) identified five problems: (1) It is difficult to select words that vary in only one dimension, since most of the variables are highly correlated with each other (the most frequent words tend to be short, more concrete, and acquired at an early age), so it is not easy to get large samples of stimuli in each cell; (2) expectations of the experimenters can influence the selection of items, since they may have implicit knowledge about the most influential variables that can lead them to select the words that best fit their hypotheses; (3) the experimenters tend to choose items with extreme values in the variables under study, and this experimental manipulation can be perceived by the participants, thus inducing specific strategies (possibly the same word is read differently if it appears in a list of words of high frequency or low frequency); (4) in factorial designs, continuous variables are used as dichotomous variables, reducing the validity of statistical analysis (the words are not high or low frequency, or short or long, but are distributed along a continuum of frequency or length); and (5) it is possible that the conclusions drawn from factorial designs, in which there are few stimuli in each cell, are limited to that small number of stimuli and are not transferable to all the stimuli in general.

Due to these criticisms, an increasing number of studies are being carried out using regression analysis on data gathered with wide ranges of stimuli, in order to test the actual weight of each variable in the participant's performance. In recent years, several mega-studies in English have been published applying this procedure (Balota et al., 2007; Spieler & Balota, 1997; Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995).

In Spanish, there is one study of this type with 2,764 words (Davies, Barbón, & Cuetos, 2013), which uses the methodology of word naming. However, although it seems important to compare results obtained from different methodologies, to our knowledge, studies based on the lexical decision task have not yet been conducted. The goal of this study was to investigate the role of the main lexical and sublexical variables in lexical selection in Spanish. One of these variables is lexical frequency, whose influence on reading has been widely tested both in languages with a transparent orthography system and in those with an opaque one (for a review, see Ghyselinck, Lewis, & Brysbaert, 2004). The most common way to measure the frequency of words is to count the number of occurrences in written text samples. However, technological advances, and especially the Internet, have enabled the development of corpora of oral frequencies obtained from movie subtitles, allowing researchers to select the most appropriate type of lexical frequency to investigate oral or written language.

Another important lexical variable is AoA. There is increasing behavioral evidence that shows that it is a variable with as much significance as lexical frequency (for a review, see Juhasz, 2005). There are two methods to obtain the AoA of a word. The most common is to ask a group of adult participants to estimate the approximate age at which they learned the said word. The measures obtained by this method are called *subjective* AoA. The objective AoA is obtained by checking the average age at which children could use a word correctly to name an object (Morrison & Ellis, 2000). Nevertheless, there is some controversy regarding this methodology, since some authors consider that it involves using a performance measure from children (the average age of children who can correctly name a picture) to predict a performance measure in adults (response latency and accuracy in word naming). As such, a number of studies have shown that subjective and objective AoA are very strongly related (Álvarez & Cuetos, 2007; Morrison & Ellis, 2000; Morrison, Ellis, & Chappell, 1997); thus, most authors use subjective measures, which are easier to obtain.

Imageability is a variable closely related to semantics, since it is generally agreed that the meanings of highly imageable words contain some sort of additional semantic information. The imageability of a word influences its recognition, since words with high imageability ratings elicit shorter lexical decision times (Schwanenflugel, Harnishfeger, & Stowe, 1988). Measures of imageability are obtained by estimations of subjects on a scale of 1 to 7, where the lowest value corresponds to *a concept very difficult to imagine* and the highest to *easily imaginable words*.

Another factor affecting the recognition of words is the capacity of a word to activate other, similar terms, which is related to the variable orthographic neighborhood. Two words are neighbors to each other when they differ in only a grapheme, preserving both the length and the order of the remaining graphemes. There have been two different manipulations on orthographic neighborhood: one being the density, which is the number of neighbors a word has, and the other the frequency of the neighbors. The latter raises some controversy because of its relationship with lexical

frequency, since high-frequency words tend to have lower frequency neighbors, while low-frequency words tend to have higher frequency neighbors (Frauenfelder, Baayen, Hellwing, & Schreuder, 1993).

Finally, length is considered a sublexical variable that reflects the number of grapheme–phoneme conversions to be made during reading. Which is to say, the longer a word is, the longer it takes to recognize it, and the greater the probability of making a mistake in the process. The effect of word length measured in letters has been observed both in languages with transparent and in those with opaque writing systems (Acha & Perea, 2008; Juphard, Carbonnel, & Valdois, 2004; Martens & de Jong, 2006). Particular attention has been paid to these five variables—frequency, length, imageability, AoA, and orthographic neighborhood—whose effects in Spanish could be quite different from those found in opaque writing systems. Differences in the roles of these variables would imply the need for an adaptation of computer models of reading to be applied to transparent orthographic systems.

The data will be analyzed by means of linear mixed-effects modeling. Such models are increasingly used because they allow for the generalization of the results beyond the sample, due to the inclusion of random effects in the analysis.

In addition, we aimed to provide a large sample of words with their reaction times (RTs) and their values in the main psycholinguistic variables to be used by other researchers.

Method

Participants

The sample size for this research was decided taking into account those used in other studies with similar characteristics. Thirty-six first-year psychology students of the University of Oviedo participated in the experiment in exchange for course credits. Their ages ranged from 17 to 23 years, the mean being 18.6 years old. All were native Spanish speakers, and their vision was normal or corrected-to-normal.

Stimulus

A total of 5,530 stimuli, 3 to 10 letters long, were presented to the participants. Half of them were the 2,765 Spanish words taken from the study by Davies et al. (2013). The other half were legal pseudowords, formed by changing one letter of other words matched in frequency with those used in the study. In addition to letter length, we took into account the values of the stimuli for subjective AoA, written lexical frequency, imageability, and number of orthographic neighbors. The frequency values, number of orthographic neighbors, and length were obtained from the database of Pérez, Alameda, and Cuetos (2003). Imageability was gathered from LEXESP

(Sebastián, Martí, Carreiras, & Cuetos, 2000). AoA data were obtained from subjective questionnaires answered by a group of 25 psychology students who did not participate in the experimental task. These questionnaires consisted of a 7-point Likert scale in which 1 corresponded to *ages between 0 and 2 years old*, 2 to *ages between 2 and 4*, and so on up to 7, which corresponded to *ages over 12 years old*.

The stimuli were randomly divided into six different groups of about 922 words each. Every group had an equal number of words and pseudowords, and stimuli with different length values were evenly distributed across the six groups.

Procedure

The experiment was run in sound-attenuated booths in the basic psychology laboratory of the University of Oviedo. Participants were tested individually and went through the six groups of stimuli in six different sessions on different days. Six practice items were presented at the beginning of each session, followed by 922 experimental items divided into four blocks of 230–231 stimuli each. Rest periods were introduced between blocks. The order of word presentation in each block and of the blocks in each session was randomized for each participant.

DMDX software was used to run the experiment. The stimuli were presented in Arial 12-point font, in white lower-case letters in the center of a black screen of a PC computer. Each experimental item was preceded by an asterisk that appeared for 500 ms in the center of the screen as a fixation point. Then the experimental stimulus was presented for up to 2 s. The participant was instructed to press one of two keys: Z if the stimulus was a pseudoword or M if it was a word. If the participant did not respond during that time, the next trial started. The program recorded both RTs and errors. In the instructions, the participants were asked to perform the task as quickly and accurately as possible.

Results

Of the total 199,080 registered responses, 26,620 errors produced by the participants were removed from the analysis (14,754 in words and 11,866 in pseudowords), which represented 13.37 % of the total ($SD = 4.91$ %). Extreme latencies, defined as responses that were 3 standard deviations above or below the participant's average, were also eliminated; they represented a total of 1,130 responses (0.57 %). The average RT of the participants was 566 ms, with a standard deviation of 44 and a range of 451–751 ms. For words, the average RT was 548 ms, the standard deviation was 44, and the range was between 451 and 751 ms. The average RT for pseudowords was 584 ms, the standard deviation was 37, and the range was 481–721 ms. Figure 1 shows average RTs for both words and pseudowords of
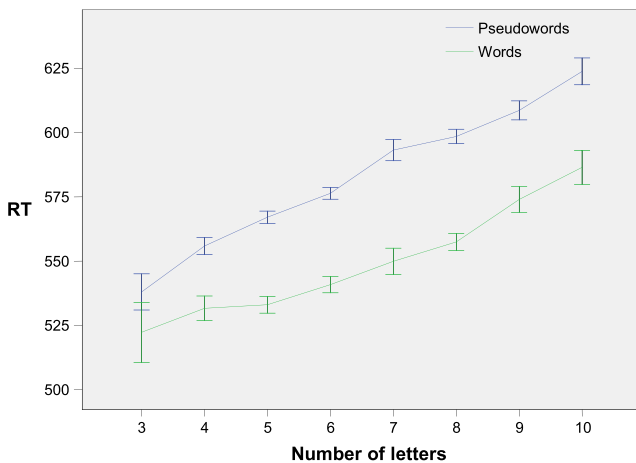
Fig. 1 Means and confidence intervals (95%) of the response latencies (RTs) according to the length of the stimuli

Table 1 Correlation between the response latencies (RTs) and the psycholinguistic variables

	Log Freq	Log AoA	Log Imag	Log Length	Ort Neigh	Log RT
Log Freq	1					
Log AoA	−.416**	1				
Log Imag	.041*	−.533**	1			
Log Length	−.213**	.362**	−.259**	1		
Ort Neigh	.124**	−.298**	.221**	−.644**	1	
Log RT	−.541**	.602**	−.291**	.360**	−.250**	1

Note. Log Freq = logarithm of the frequency; Log AoA = logarithm of the age of acquisition; Log Imag = logarithm of the imageability; Log Length = logarithm of the length; Ort Neigh = number of orthographic neighbors; Log RT = logarithm of the reaction times.

* Correlation is significant at the .05 level (bilateral).

** Correlation is significant at the .01 level (bilateral).

different lengths. Latency values for every word used in the study can be found in the Appendix.

In this study, we made use of mixed-effects multiple regression models with random intercepts for subjects and items and the predictors presented above as fixed effect factors or covariates. The construction of the mixed models follows the methodology used in the works by Baayen, Davidson, and Bates (2008), Kuperman, Schreuder, Bertram, and Baayen (2009) and Pinheiro and Bates (2000). The statistical analyses were performed using R statistical programming open code software (R Development Core Team, 2011, version 2.15).

The first step of the analysis was to test the normality of the variables through the Kolmogorov–Smirnov test. None of them fit the normal curve, although according to Zuur, Ieno, Walker, Saveliev, and Smith (2009), Sokal and Rohlf (1995), Zar (1999), and Fitzmaurice, Laird, and Ware (2004), given our sample size, the lack of normality is not a problem. Nevertheless, we proceeded to perform logarithmic transformations to reduce the influence of outliers on the RTs and the psycholinguistic variables, with the exception of the orthographic neighborhood, since it contained null values.

In the correlation analysis, shown in Table 1, none of the variables were highly correlated with the others.

A mixed-effects model was built considering the subject and the item (word) as random effects and length, AoA, orthographic neighborhood, frequency, and imageability as predictors (along with all first-order interactions). The transformation of the RT was deemed a response variable. After building the model, outliers (points farther than 2.5 standard deviations from the residual error) were identified and eliminated (the number of outliers was 1,797, 2.11 % of the total), and the model was readjusted, assuming a stepwise method for obtaining the final model (see Table 2).

The results showed that all variables, except length, influenced the RT, although significant interactions were

found between length and the AoA and orthographic neighborhood. The interaction found in the analysis between AoA and frequency shows that the AoA effect is more evident with low-frequency words. Something similar happens with imageability, since the effect of this variable is also greater with low-frequency words and words with late AoA. Regarding the orthographic neighborhood, this variable affects the RTs more for low-imageability words. Plots derived from the model can be observed in Fig. 2.

Analysis based on length

The length of a word is a central variable in visual word recognition, especially in transparent languages such as Spanish; however, in most of the factorial experiments conducted thus far, this variable has been controlled, rather than manipulated. Therefore, it seems important to perform analyses based on this variable. In order to determine whether word length affects the influence of the rest of the psycholinguistic variables in the RT, all the items were divided into two groups: short words, with stimuli of 3–6 letters, and long words, with stimuli of 7–10 letters. The short words group consisted of a total of 1,554 items. The average RT of participants in this group was 536 ms, with a standard deviation of 41 and a range of 451–657. The long words group consisted of 1,211 items. The mean latency in this group was 564 ms, with a standard deviation of 42 and a range of 472–751.

Using the same procedure as in the previous model, we obtained correlations between the transformed variables for both short- and long-word groups. Again, no high correlations were observed, so we proceeded to build the mixed-effects model, considering the main predictors and their first-order interactions as fixed effects and the subject and item as random effects (see Table 3). Outliers were removed (*n* = 1,026, 2.15 % for the short-

Table 2 Model for lexical decision reaction times (RTs)

Fixed Effects

Variable	Estimate	MCMCmean	HPD95lower	HPD95upper	pMCMC
Intercept	5.8162	5.8179	5.7466	5.8910	.001
Log Freq	0.1870	0.1855	0.1293	0.2342	.001
Log AoA	0.2252	0.2237	0.1760	0.2633	.001
Log Imag	0.1970	0.1962	0.1555	0.2358	.001
Ort Neigh	0.0167	0.0166	0.0123	0.0206	.001
Log Freq:Log AoA	−0.0986	−0.0978	−0.1142	−0.0783	.001
Log Freq:Log Imag	−0.0935	−0.0930	−0.1186	−0.0707	.001
Log AoA:Log Imag	−0.1166	−0.1161	−0.1368	−0.0884	.001
Log AoA:Log Length	0.0435	0.0437	0.0366	0.0495	.001
Log Imag:Ort Neigh	−0.0040	−0.0040	−0.0056	−0.0025	.001
Ort Neigh:Log Length	−0.0068	−0.0068	−0.0088	−0.0050	.001

Fixed Effects

Estimate	Name	SD	MCMCmedian	MCMCmean	HPD95lower	HPD95upper
Item	(Intercept)	0.0441	0.0405	0.0405	0.0390	0.0419
Subject	(Intercept)	0.0885	0.0792	0.0799	0.0735	0.0916
Residual		0.1521	0.1524	0.1524	0.1516	0.1531

Note. MCMC = Monte Carlo Markov chain; HPD95lower = lower boundary of the 95 % highest posterior density interval; HPD95upper = upper boundary of the 95 % highest posterior density interval; pMCMC = *p* values estimated by the MCMC method using 1,000 simulations; Log Freq = logarithm of the frequency; Log AoA = logarithm of the age of acquisition; Log Imag = logarithm of the imageability; Log Length = logarithm of the length; Ort Neigh = number of orthographic neighbors; Log RT = logarithm of the reaction times; SD = standard deviation.

word group and *n* = 769, 2.05 % for the long-word group), and the model was readjusted.

As in the general model, significant effects were found in the short-word group for all variables except length. The observed interactions were also consistent with those found in the previous analysis, although an interaction between imageability and orthographic neighborhood was found, which was not obtained in the general model. The data indicated that the RTs were affected more by the orthographic neighborhood when words had lower imageability values.

In the long-word group, on the other hand, the length was significant, but the number of orthographic neighbors was not. The interactions obtained showed that, as in the general model, the effect of AoA on RTs was more evident among low-frequency words, while the imageability effect was greater for low-frequency and early acquired words.

Discussion

This study, the first of its kind conducted in Spanish, with a sample of 5,530 words and pseudowords with lengths of 3–10 letters, was carried out to ascertain the influence of the main lexical and sublexical variables in a lexical decision task. The results show that all the variables included in the analysis, except length, influenced the RTs.

These data partially coincide with those obtained in previous studies carried out with different methodologies. Alija and Cuetos (2006) used a factorial design to evaluate variables influencing the lexical decision task and, as in the present study, found significant effects of frequency and AoA, but not of imageability. Cuetos, Barbón, Urrutia, and Domínguez (2009) used the evoked response potentials paradigm to investigate the effect of these three variables in a word-naming task. The results of their study were consistent with those obtained by Alija and Cuetos: Lexical frequency affected the recognition process in an early phase (between 175 and 360 ms), while an AoA effect was observed at a later stage (400–610 ms). Imageability did not produce any effect in any time window tested. These results are remarkable since the imageability variable has a semantic basis, and therefore, its effects tend to be stronger in the lexical decision task in which meaning is thought to be used to discriminate between words and nonwords (Chumbley & Balota, 1984).

By contrast, when reading aloud, one only needs to convert an orthographic representation into a phonological representation. The semantic information is not required for this, particularly in transparent languages in which the correspondence between grapheme and phoneme is complete. Imageability, therefore, should not be an important variable in the word-naming task. This is what Cuetos and Barbón (2006) found in a regression analysis study conducted in Spanish. The best predictors of RTs were AoA and, contrary to what was found in the present study, length. Neither frequency nor imageability was significant. Davies et al. (2013), however, in a similar study also using the word-naming task,
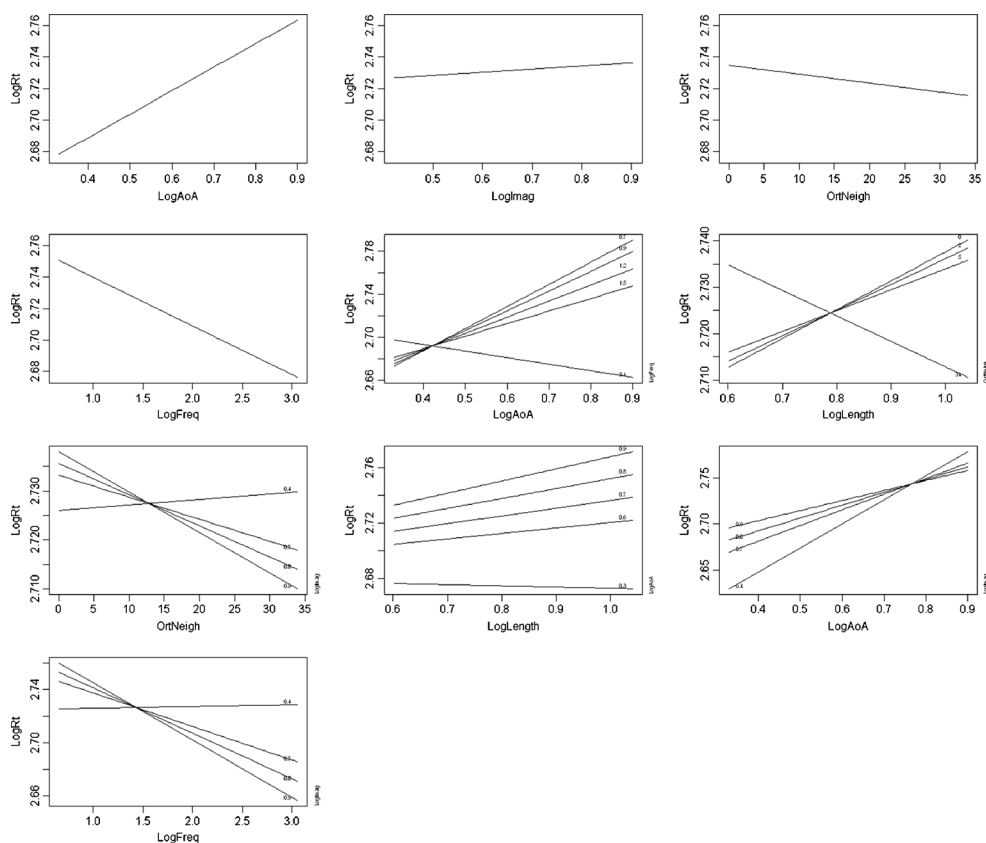
Fig. 2 Individual effects and interactions affecting lexical decision reaction times (RTs)

found that semantic knowledge (which grouped AoA, imageability, and familiarity) influenced the RTs despite the high consistency between orthography and phonology in Spanish. A similar study by Cortese and Schock (2012) in English explored the influence of imageability and AoA in lexical decision and word naming. The results showed that both variables were significant predictors of the two tasks, but the imageability effect did not depend on the consistency between orthography and phonology. The results obtained from this study and ours indicate that semantic activation may influence the generation of a phonological code and that semantics plays an important role in the recognition of words.

Some of the interactions obtained in the mixed-effects model are in the expected direction and are consistent with those obtained in similar studies conducted in Spanish (Cuetos & Barbón, 2006). AoA barely affects high-frequency words, since the RTs tend to be low regardless of the age at which the words have been acquired. By contrast, among the low-frequency words, the effects of AoA are much more evident, since the RTs for the early acquired words are lower than the RTs for the words acquired at a later age.

Something similar happens with the significant interactions between imageability and frequency and imageability and AoA. The data indicate that there is little variability in the high-frequency or early acquired words, since in these cases,

the RTs always tend to be short, regardless of the imageability. However, there is a greater influence of imageability in the low-frequency or late acquired items, with the RTs for the high-imageability words being lower than the RTs for the low-imageability words.

The interaction between imageability and orthographic neighborhood is similar, with the effect of the number of orthographic neighbors on RTs being greater for low-imageability words.

Orthographic neighborhood was significant only in the short-word group, which is not surprising, since the short words have many more orthographic neighbors than the long words.

The only predictor that was not significant in the analysis was length. However, the interactions between this variable and frequency and AoA indicate that there is an effect of length on RTs, although in very frequent or early acquired words, it is masked and cannot be observed. Length is considered a sublexical variable because it reflects the application of the grapheme–phoneme rules, so that the more graphemes a word has, the more time is needed for its recognition. In fact, in the mixed-effect regression analysis carried out after splitting the stimuli into short and long words, a significant length effect was found only in the long-word group, which reflects the fact that the short words can be processed at a glance

regardless of whether they have three or six letters, while the long words cannot. It is not surprising that a sublexical variable like this one has substantial effects in a transparent language like Spanish. However, this variable is ignored in many investigations. Therefore, most of the studies based on factorial experiments control the words' length, using short words of four or five letters, instead of manipulating the length in the experimental design. Perhaps this is appropriate for studies conducted in English, but we must not forget that the average length of Spanish words is around eight letters.

In this study the "U" effect of length, as described by New, Ferrand, Pallier, and Brysbaert (2006) for English, was not found. In New et al.'s study, the response latencies decreased with word length between three and five letters; in words of five to eight letters, the latencies did not change, and from eight letters on, the latencies increased in length. In our study, however, latencies increased linearly, with an average increase of 9 ms per letter, although it was not uniform, since the latency increase was mild between words with four to six letters and was much larger for words with seven or more

Table 3 Model for short- and long-word groups

Short Words

Fixed Effects

Variable	Estimate	MCMCmean	HPD95lower	HPD95upper	pMCMC
Intercept	5.9182	5.9173	5.8192	5.9975	.001
Log Freq	0.1407	0.1402	0.0812	0.2046	.001
Log AoA	0.2381	0.2380	0.1875	0.2872	.001
Log Imag	0.1267	0.1270	0.0792	0.1762	.001
Ort Neigh	0.0089	0.0089	0.0046	0.0132	.001
Log Freq:Log AoA	−0.0921	−0.0916	−0.1122	−0.0709	.001
Log Freq:Log Imag	−0.0707	−0.0704	−0.1028	−0.0428	.001
Log AoA:Log Imag	−0.0732	−0.0733	−0.1058	−0.0460	.001
Log AoA:Ort Neigh	−0.0024	−0.0024	−0.0036	−0.0010	.001
Log Imag:Ort Neigh	−0.0040	−0.0040	−0.0059	−0.0019	.001

Random Effects

Estimate	Name	SD	MCMCmedian	MCMCmean	HPD95lower	HPD95upper
Item	(Intercept)	0.0401	0.0371	0.0371	0.0351	0.0390
Subject	(Intercept)	0.0853	0.0779	0.0757	0.0586	0.0943
Residual		0.1531	0.1534	0.1534	0.1525	0.1543

Long Words

Fixed Effects

Variable	Estimate	MCMCmean	HPD95lower	HPD95upper	pMCMC
Intercept	5.3087	5.3093	5.1888	5.4294	.001
Log Freq	0.3008	0.2999	0.2095	0.3904	.001
Log AoA	0.4406	0.4415	0.3706	0.5085	.001
Log Imag	0.3229	0.3239	0.2600	0.3946	.001
Log Length	0.1436	0.1424	0.1143	0.1715	.001
Log Freq:Log AoA	−0.1435	−0.1425	−0.1799	−0.1020	.001
Log Freq:Log Imag	−0.1228	−0.1231	−0.1600	−0.0862	.001
Log AoA:Log Imag	−0.1946	−0.1954	−0.2352	−0.1536	.001

Random Effects

Estimate	Name	SD	MCMCmedian	MCMCmean	HPD95lower	HPD95upper
Item	(Intercept)	0.0475	0.0434	0.0434	0.0410	0.0459
Subject	(Intercept)	0.0937	0.0779	0.0808	0.0645	0.1030
Residual		0.1499	0.1503	0.1503	0.1491	0.1514

Note. MCMC = Monte Carlo Markov chain; HPD95lower = lower boundary of the 95 % highest posterior density interval; HPD95upper = upper boundary of the 95 % highest posterior density interval; pMCMC = *p* values estimated by the MCMC method using 1,000 simulations; Log Freq = logarithm of the frequency; Log AoA = logarithm of the age of acquisition; Log Imag = logarithm of the imageability; Log Length = logarithm of the length; Ort Neigh = number of orthographic neighbors; Log RT = logarithm of the reaction times; SD = standard deviation.

letters. Regarding pseudowords, the progression was 12 ms, and the increase was more uniform.

In sum, these results seem to suggest that reading in Spanish, as in deep orthography languages, depends on a mixture of lexical and sublexical strategies. Contrary to the hypotheses by Frost et al. (1987) that, in transparent orthographies, the lexical route is not necessary, because the sublexical route is fast and accurate enough, this study confirms the relevance of the lexical route in Spanish, as shown by the effects of AoA and lexical frequency. However, in light of our data, we cannot confirm whether the degree of involvement of the lexical route is similar to that in opaque languages like English, for which cross-linguistic studies would be necessary.

This study has shown that there are some peculiarities in the Spanish orthographic system, which entails different processing strategies. In particular, Spanish words are mostly polysyllabic, with an average length much larger than in English. Therefore, the effects of length may be stronger in our language, especially in long words. It has also been found that the effects of length in Spanish are linear, contrasting with the inverted U-shape effects obtained in English. Consequently, reading models must take into account the peculiarities of the different spelling systems if they want to be universal.

Perhaps the most important finding of this study is that the results obtained in factorial designs, which use a small number of words with very specific characteristics, cannot always be extrapolated to the rest of the words, since they could be empowering certain variables. Orthographic neighborhood, usually presented as an important variable in many studies, seems to play an important role only in short words. Imageability, on the other hand, which does not seem to be a relevant variable in transparent languages, appears to be relevant, especially when reading low-frequency or late acquired words. As such, possibly many of the results obtained with factorial designs should be reviewed.

References

Acha, J., & Perea, M. (2008). The effects of length and transposed-letter similarity in lexical decision: Evidence with beginning, intermediate, and adult readers. *British Journal of Psychology, 99,* 245–264.

Alija, M., & Cuetos, F. (2006). Efecto de las variables léxico-semánticas en el reconocimiento visual de palabras. *Psicothema, 18,* 485–491.

Álvarez, B., & Cuetos, F. (2007). Objective age of acquisition norms for a set of 328 words in Spanish. *Behavior Research Methods, 39,* 377–383.

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

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology. General, 133,* 283–316.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftus, B., ... Treiman, R. (2007). The English lexicon project: A user's guide. *Behavior Research Methods, 39,* 445–459.

Burani, C., Arduino, L. S., & Barca, L. (2007). Frequency, not age of acquisition, affects Italian word naming. *European Journal of Cognitive Psychology, 19,* 828–866.

Carreiras, M., Álvarez, C. J., & de Vega, M. (1993). Syllable frequency and visual word recognition in Spanish. *Journal of Memory and Language, 32,* 766–780.

Carreiras, M., Perea, M., & Grainger, J. (1997). Effecs of orthographic neighborhood in visual word recognition: cross-task comparisons. *Journal of Experimental Psychology: Learning, Memory and Cognition, 23,* 857–871.

Chumbley, J. I., & Balota, D. A. (1984). A word's meaning affects the decision in lexical decision. *Memory & Cognition, 12,* 590–606.

Coltheart, M. (1981). Disorders of reading and their implications for models of normal reading. *Visible Language, 15,* 245–286.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256.

Cortese, M. J. & Schock, J. (2012). Imageability and age of acquisition effects in disyllabic word recognition. *The Quarterly Journal of Experimental Psychology,* 1–27.

Cuetos, F., & Barbón, A. (2006). Word naming in Spanish. *European Journal of Cognitive Psychology, 18,* 415–436.

Cuetos, F., Barbón, A., Urrutia, M., & Domínguez, A. (2009). Determining the time course of lexical frequency and age of acquisition using ERP. *Clinical Neurophysiology, 120,* 285–294.

Davies, R., Barbón, A., & Cuetos, F. (2013). Lexical and semantic age-of-acquisition effects on word naming in Spanish. *Memory & Cognition, 41,* 297–311.

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2004). *Applied Longitudinal Analysis, Wiley Series in Probability and Statistics*. Hoboken: Wiley.

Frauenfelder, U. H., Baayen, R. H., Hellwing, F. M., & Schreuder, R. (1993). Neighborhood density and frequency across languages and modalities. *Journal of Memory and Language, 32,* 781–804.

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographic depth: A multilingual comparison. *Journal of Experimental Psychology. Human Perception and Performance, 13,* 104–115.

Ghyselinck, M., Lewis, M. B., & Brysbaert, M. (2004). Age of acquisition and the cumulative-frequency hypothesis: A review of the literature and a new multitask investigation. *Acta Psychologica, 115,* 43–67.

Juhasz, B. J. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin, 131,* 684–712.

Juphard, A., Carbonnel, S., & Valdois, S. (2004). Lenght effect in reading and lexical decision: Evidence from skilled readers and a developmental dyslexia participant. *Brain and Language, 55,* 332–340.

Kuperman, V., Schreuder, R., Bertram, R., & Baayen, R. H. (2009). Reading polimorphemic Dutch compounds: Toward a multiple route model of lexical processing. *Journal of Experimental Psychology. Human Perception and Performance, 35,* 876–895.

Martens, V. E., & de Jong, P. F. (2006). The effect of word length on lexical decision in dyslexic and normal reading children. *Brain and Language, 98,* 140–149.

Morrison, C. M., & Ellis, A. W. (2000). Real age of acquisition effects in word naming and lexical decision. *British Journal of Psychology, 91,* 167–180.

Morrison, C. M., Ellis, A. W., & Chappell, T. D. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *Quarterly Journal of Experimental Psychology, 50A,* 528–559.

New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006). Re-examining the word length effect in visual word recognition: New evidence from the English lexicon project. *Psychonomic Bulletin & Review, 13,* 45–52.

Pérez, M. A., Alameda, J. R., & Cuetos, F. (2003). Frecuencia, longitud y vecindad ortográfica de las palabras de 3 a 16 letras del Diccionario de la Lengua Española (RAE, 1992) [Frequency, length and orthographic neighbourhood of 3 to 16 letter words in the Dictionary of the Spanish Language (RAE, 1992)]. *Revista Electrónica de Metodología Aplicada, 8,* 1–10.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-Effects Models in S and S-Plus*. New York: Springer.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. E. (1996). Understanding normal and impaired reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56–115.

R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/. Accessed 3 Sept 2013.

Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language, 27,* 499–520.

Sebastián, N., Martí, M. A., Carreiras, M. F., & Cuetos, F. (2000). *LEXESP, Léxico informatizado del español [Computerized lexicon of Spanish]*. Barcelona: Edíciones de la Universitat de Barcelona.

Sokal, R. R., & Rohlf, F. J. (1995). *Biometry. The principles and practice of statistics in biological research*. New York: W. H. Freeman.

Spieler, D. H., & Balota, D. A. (1997). Bringing computational models of word naming down to the item level. *Psychological Science, 8,* 411–416.

Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology. General, 124,* 107–136.

Zar, J. H. (1999). *Biostatistical analysis*. Upper Saddle River: Prentice Hall.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131,* 3–29.

Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., & Smith, G. M. (2009). *Mixed Effects Models and Extensions in Ecology with R (Statistics for Biology and Health)*. New York: Springer.