Lexical and semantic age-of-acquisition effects on word naming in Spanish

Robert Davies · Analía Barbón · Fernando Cuetos

Published online: 21 November 2012
© Psychonomic Society, Inc. 2012

Abstract We report a study of the factors that affect reading in Spanish, a language with a transparent orthography. Our focus was on the influence of lexical semantic knowledge in phonological coding. This effect would be predicted to be minimal in Spanish, according to some accounts of semantic effects in reading. We asked 25 healthy adults to name 2,764 mono- and multisyllabic words. As is typical for psycholinguistics, variables capturing critical word attributes were highly intercorrelated. Therefore, we used principal components analysis (PCA) to derive orthogonalized predictors from raw variables. The PCA distinguished components relating to (1) word frequency, age of acquisition (AoA), and familiarity; (2) word AoA, imageability, and familiarity; (3) word length and orthographic neighborhood size; and (4) bigram type and token frequency. Linear mixed-effects analyses indicated significant effects on reading due to each PCA component. Our observations confirm that oral reading in Spanish proceeds through spelling–sound mappings involving lexical and sublexical units. Importantly, our observations distinguish between the effect of lexical frequency (the impact of the component relating to frequency, AoA, and familiarity) and the effect of semantic knowledge (the

impact of the component relating to AoA, imageability, and familiarity). Semantic knowledge influences word naming even when all the words being read have regular spelling–sound mappings.

Keywords Spanish · Reading · Naming · Age of acquisition · Orthographic transparency · Curvilinear

A general account of the cognitive reading system must describe and explain the effects of word attributes on reading in different orthographies. Orthographies may differ in a range of ways, but there has been a long-standing interest in how reading varies in relation to differences in orthographic transparency (e.g., Frost, Katz, & Bentin, 1987). The Spanish orthography is transparent because orthography–phonology mappings are completely rule governed across the language (Cuetos & Barbón, 2006). In comparison, the English orthography is quasi-regular, having both rule-governed and exceptional orthography–phonology mappings (Plaut, McClelland, Seidenberg, & Patterson, 1996). Current evidence suggests that reading is affected by a mix of factors that appears to be similar in kind, whether more opaque or more transparent orthographies are examined. Our aim was to bring evidence to bear on whether reading effects would be similar, also, in the specific components. We were especially concerned with the impact on reading of the key lexical factors—frequency, age of acquisition (AoA), and imageability (seen to affect reading in English; e.g., Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Cortese & Khanna, 2007)—and with the question of whether previous observations of the AoA effect on reading in Spanish (Cuetos & Barbón, 2006) can be taken to reveal the influence of semantics on word naming in a transparent orthography.

Electronic supplementary material The online version of this article (doi:10.3758/s13421-012-0263-8) contains supplementary material, which is available to authorized users.

R. Davies (✉)
Department of Psychology, Faculty of Health and Life Sciences,
Oxford Brookes University,
Harington Campus, Gipsy Lane,
Oxford OX3 0BP, UK
e-mail: r.davies@brookes.ac.uk

A. Barbón · F. Cuetos
University of Oviedo,
Asturias, Spain

Frequency and AoA effects in reading in Spanish and other languages

Researchers have observed that reading performance is affected by knowledge about words—in addition to influences due to knowledge about word constituents—across a range of variation in orthographic transparency. There have been numerous reports of the effects of word frequency and AoA on reading in transparent orthographies (frequency in Dutch, Brysbaert, Lange, & Van Wijnendaele, 2000; Brysbaert, Van Wijnendaele, & de Deyne, 2000; Ghyselinck, Lewis, & Brysbaert, 2004; frequency in Italian, Barca, Burani, & Arduino, 2002; Bates, Burani, D’Amico, & Barca, 2001; Burani, Arduino, & Barca, 2007; Paizi, Burani, & Zoccolotti, 2010; frequency in Persian, Baluch & Besner, 1991; AoA in Spanish, Cuetos & Barbón, 2006; frequency in Serbo-Croat, Frost et al., 1987; AoA in Turkish, Raman, 2006; and frequency in Turkish, Raman, Baluch, & Besner, 2004) in line with demonstrations seen in an opaque orthography such as English (Balota et al., 2004; Cortese & Khanna, 2007). However, there has been a vigorous debate over the interpretation of the AoA effect, especially over its relation to frequency and imageability effects, and over the locus (joint or separate) of the three effects (e.g., Brysbaert & Ghyselinck, 2006; J. Monaghan & Ellis, 2002; Zevin & Seidenberg, 2002). The outcome of these debates has implications for our assumptions about the architecture of the reading system.

One source of the debate surrounding the lexical influences on reading arises from the correlation between the lexical frequency and AoA variables. Words learned early in life will also tend to be experienced often. This correlation between the variables makes it difficult to distinguish the unique contribution of each to variance in a reading outcome measure, a situation termed *multicollinearity* (Cohen, Cohen, West, & Aiken, 2003). Some observations of an AoA effect may thus be due to a confound between the AoA and frequency of stimuli. And, for some previous reports on reading in English, a reanalysis of results suggests that this has been the case (Zevin & Seidenberg, 2002). It is possible that the AoA effect previously reported in Spanish, by Cuetos & Barbón (2006), was actually a frequency effect. Consistent with that possibility, in research in Italian, an orthography similar to Spanish, all observations on healthy adult reading indicate an effect of frequency, but not of AoA (Barca et al., 2002; Bates et al., 2001; Burani et al., 2007).

One approach to identifying the factors that affect reading, when they are correlated, is to attack the multicollinearity directly by orthogonalizing variables through principal components analysis (PCA; see Cohen et al., 2003). Burani and colleagues (Barca et al., 2002; Bates et al., 2001; Burani et al., 2007) followed this approach, reporting PCAs that showed that high proportions of variance in word attribute

variables were related to a set of easily interpretable underlying components. Frequency measures and, to a lesser extent, AoA and familiarity (where available; in Barca et al., 2002; Bates et al., 2001) loaded heavily on a “frequency” component. AoA also loaded, together with imageability and familiarity (where available), on a “semantic” component. In each study, multiple regression analyses using the orthogonal components as predictors showed that variance in Italian word-naming latencies was explained by the effect of the “frequency,” but not the “semantic,” component. Thus, it may be that the AoA effect reported for Spanish arises from the variance shared by AoA and frequency. This possibility is consistent with the idea (Brysbaert & Ghyselinck, 2006) that variance in AoA—and therefore, the AoA effect—stems from a combination of variance due to distinct underlying factors that encompass (1) frequency-related variance and (2) variance relating to word meaning and AoA.

One aim of the present study was to test the possibility that the Spanish AoA effect, of the sort reported by Cuetos and Barbón (2006), derives from frequency-related variance. We examined this possibility by conducting a PCA of item attribute variables and using orthogonalized factors to estimate the influences affecting Spanish word naming. We hypothesized that if the AoA effect observed by Cuetos and Barbón were actually a frequency effect—the AoA and frequency variables used in that study were “raw” variables—then we should find (1) that a PCA analysis of item attributes yields a “frequency” factor with loadings on AoA and frequency measures and a “semantic” factor with loadings on AoA and imageability and (2) that there are significant effects of the “frequency,” but not the “semantic,” factors on Spanish word-naming latencies.

However, it seemed to us possible, alternatively, that the AoA effect previously reported by some of us reflected the unique contribution of both the frequency-dependent and the semantic AoA components. We should explain why a semantic AoA effect in Spanish reading could be expected to be unlikely before presenting our argument for why it can be predicted, given past research. Two lines of research predict that AoA effects should be reduced or null in reading when it involves consistent mappings, as it does for Spanish, a language with a regular orthography. The evidence comes from computational modeling and from experimental comparisons of AoA effects in reading and other tasks.

Age-dependent learning has been simulated using connectionist networks, but it has been shown over a series of studies that the effect of AoA is modulated by the predictability of input and output patterns (Ellis & Lambon Ralph, 2000; Lambon Ralph & Ehsan, 2006; J. Monaghan & Ellis, 2002; Zevin & Seidenberg, 2002). The general observation is that AoA effects are weak or null where mappings are predictable, where they are systematic or consistent as in orthography-to-phonology (OP) mappings in reading

(Lambon Ralph & Ehsan, 2006; J. Monaghan & Ellis, 2002; Zevin & Seidenberg, 2002), but are substantial where mappings are less predictable, where they are arbitrary, as in semantics-to-phonology (SP) mappings in picture naming (Ellis & Lambon Ralph, 2000; Lambon Ralph & Ehsan, 2006; Zevin & Seidenberg, 2002), or inconsistent or irregular, as in OP mappings in exception word reading in English (J. Monaghan & Ellis, 2002).

The most recent, and extensive, observations are reported by P. Monaghan and Ellis (2010). In the critical simulation, their developmental model, the training regime furnished to the model learner was carefully tailored to resemble children's reading experiences. The developmental model simulations showed that the age at which words are learned—their point of entry in the training regime—had a substantial effect on reading performance, accounting for variance over and above variance due to other factors, including word length, neighborhood size, spelling–sound consistency, and, most important, cumulative frequency.

Crucially, the P. Monaghan and Ellis (2010) simulations replicated the interaction between the effects of AoA and spelling–sound consistency reported by J. Monaghan and Ellis (2002) for adult word naming. AoA effects were stronger for inconsistent exception words than for consistent words. The simulations showed that the point of entry to training (the AoA) of a word had an effect because the impact of a word's presentation on the network's connection weights decreased over time (consistent with results reported by Ellis & Lambon Ralph, 2000). Connectionist models learn as a result of the accumulated modification of connection weights in response to the presentation of training stimuli. In the P. Monaghan and Ellis simulations, there was a reduction in plasticity—the degree to which weights could be modified—as training progressed. Early-acquired words had greater opportunity to optimize connection weights, benefiting performance in response to them and thus creating an AoA effect. Learning to read inconsistent words involves overcoming competition from alternate pronunciations; therefore, it requires greater changes among connection weights. For early-acquired inconsistent words, there is space for experience to create such changes, but for late-acquired words, there is not. In contrast, learning to read consistent words can borrow from knowledge about words with similar OP mappings. As a result, late-acquired consistent words are not penalized for late entry.

The P. Monaghan and Ellis (2010) simulations are important because they demonstrate AoA effects in a network containing only orthographic and phonological representations. They are also important because they show that AoA effects are reduced for consistent, as compared with inconsistent, mappings (as are frequency effects). These results imply that the effect of AoA on word naming in Spanish should be small (certainly, as compared with those seen for exception word reading in English) to the extent that word

naming in Spanish is governed by the organizing properties of (just the) OP mappings, because such mappings are regular in Spanish.

Language tasks involving semantics have been found to show stronger AoA effects than have tasks that can be completed using just OP mappings (P. Monaghan & Ellis, 2010; Zevin & Seidenberg, 2002). Brysbaert and colleagues (in observations on Dutch; Brysbaert, Lange, & Van Wijnendaele, 2000; Brysbaert, Van Wijnendaele, & de Deyne, 2000; Ghyselinck et al., 2004) reported larger AoA effects in lexical decision, association generation, and semantic classification—tasks that depend upon semantic processes—as compared with word naming. Cortese and Khanna (2007) observed a larger AoA effect in lexical decision than in word naming for a large sample of words in English (drawn from the set used by Balota et al., 2004), although they noted that the AoA effect in word naming, while small, was significant. The AoA effect has been found to be larger in picture naming than in word naming, for the same words, in Italian (Bates et al., 2001), Turkish (Raman, 2011), and English (Lambon Ralph & Ehsan, 2006). And, indeed, Brysbaert and Ghyselinck (2006), drawing together evidence from a series of cross-task comparisons, found that the AoA effect can be distinguished from the frequency effect only where semantic processing—perhaps involving competition between concepts for output—is required, as in picture naming (see, also, Belke, Brysbaert, Meyer, & Ghyselinck, 2005).

If AoA effects are greater in tasks involving semantics, a substantial body of experimental research suggests that semantic involvement should not be expected for word naming in Spanish. Previous research on reading in English suggests that a semantic influence will be observed only where phonological coding is difficult, as it is for low-frequency irregular exception words (e.g., Strain, Patterson, & Seidenberg, 1995). Whether the semantic influence is better characterized as owing to imageability or to AoA has been disputed, with observations of an imageability effect in word naming (Balota et al., 2004; Strain et al., 1995) appearing not to survive entry of AoA as a covariate in latency analyses (Cortese & Khanna, 2007; J. Monaghan & Ellis, 2002). That is not surprising given the typically high correlation between AoA and imageability, as well as the PCA results reported by Burani and colleagues that link AoA and imageability to an underlying “semantic” factor (Barca et al., 2002; Bates et al., 2001; Burani et al., 2007). However, the important point is that data from English suggest that a semantic influence on reading appears only for words with atypical—that is, irregular exception or inconsistent—spelling–sound mappings (see also Woollams, 2005).

This generalization is limited, however, by the findings of recent research on reading, especially in transparent orthographies. First, Raman (2006) reported that an AoA effect on word naming for high-imageability items can be detected in

Turkish, a wholly transparent orthography, even after controlling for item frequency. That is, Raman (2006) reported a frequency-independent AoA effect that, as we have noted, previous research suggested would appear to have a locus in semantic processing. Indeed, Raman and Baluch (2001) reported that highly skilled readers of Turkish showed an effect of imageability on oral reading for low-frequency words. It should also be noted that Baayen, Feldman, and Schreuder (2006) argued that the effects of frequency and rated familiarity in reading may in part be due to the way in which these variables capture semantic familiarity. This claim is supported by the way in which frequency and familiarity measures cluster in lexical space with semantic variables like rated imageability, AoA, and measures of morphological connectivity. Semantic familiarity may affect performance, in this account, as an organizing principle in a highly interactive reading system (see Raman & Baluch, 2001, for a cognate view).

Some research therefore suggests a broadly based involvement for semantics on word naming. This is consistent with findings from neuropsychological studies concerning patients who produced semantic reading errors or who showed the effect of imageability on reading accuracy in transparent orthographies such as Spanish (Davies & Cuetos, 2005; Davies, Cuetos, & Rodríguez-Ferreiro, 2010), Turkish (Raman & Weekes, 2005), and Welsh (Beaton & Wyn Davies, 2007). These results suggest that a semantic route for phonological coding is available to readers of transparent orthographies. That possibility would be consistent with theoretical accounts for the development of reading fluency in transparent orthographies (e.g., Wimmer, 1993, 2006) in which fluency increases with the establishment of large-grain mappings from spelling to sound. Such mappings could be construed to involve semantics, if evidence for a frequency-independent AoA effect were found for reading in a transparent orthography. Indeed, Balota et al. (2004) proposed that their observation of semantic effects in word naming in English indicated activation cascading from the semantic level. Such cascading would support fluent reading.

The present study

A frequency-independent AoA effect in Spanish would clearly be surprising, although it would be consistent with some previous findings. We hypothesized that if the AoA effect on Spanish word-naming latencies were actually a frequency effect, then (1) a PCA analysis of item attributes would yield a “frequency” factor with loadings on AoA and frequency measures and a “semantic” factor with loadings on AoA and imageability, and (2) a mixed-effects analysis would indicate significant effects of the “frequency,” but not the “semantic,”

factor. Such a frequency–AoA effect would not require postulation of semantic activation in word naming (Burani et al., 2007), although it would be consistent with the influence of lexical knowledge on phonological coding. However, previous reports of frequency-independent AoA effects in transparent orthographies (e.g., Raman, 2006) suggested that we could find a frequency-independent AoA effect on Spanish reading. Such an effect would currently be interpreted in terms of semantic involvement in lexical processing (Brysbaert & Ghyselinck, 2006). And that semantic interpretation would be consistent with the possibility of semantic activation cascading in word naming (Balota et al., 2004).

We investigated the impact of a range of factors on reading for a sample of 2,764 mono- and multisyllabic words. Our study presents an advance on previous work by addressing the multicollinearity of AoA, frequency, and imageability. We analyzed item attribute variables using PCA orthogonalized factors. We then used these factors to estimate the effects of frequency-dependent and frequency-independent AoA effects on word naming.

Method

Subjects

We tested 25 subjects (average age = 21 years); all were healthy young adults, recruited from undergraduates studying psychology at the University of Oviedo. All subjects were monolingual speakers of Spanish. All had normal or corrected-to-normal vision, and none reported diagnoses of previous neurological illness.

Materials

We selected 2,764 words from the LEXESP database for Spanish (Sebastián, Martí, Carreiras, & Cuetos, 2000). Frequency estimates in that database, which we used in our analyses, were calculated from the frequency of appearance of words in a corpus of about five million words of printed text. We selected those nouns, verbs, and adjectives in the LEXESP corpus vocabulary that were 3–10 letters in length, excluding all compounds. We transformed the LEXESP frequency values to $\log_{10}(\text{frequency per million} + 1)$ to ameliorate skew in the distribution of the variable. We also took values for subjective ratings of word imageability from the LEXESP database (ratings were made on a 7-point scale). We collected ratings of AoA for all words (on a 7-point scale). The sample of 2,764 words was split into five subsets of about 600. Each subset of words was administered to a group of 25 volunteers, with each subset of words rated by a different group of

volunteers. In total, AoA ratings were collected from 125 volunteers; all were undergraduates attending the University of Oviedo.

While Cuetos and Barbón (2006) did not find a significant effect of familiarity on Spanish word-naming latencies, for a relatively small word sample, previous large-scale analyses of reading have shown that an important influence on reading performance is captured by subjective ratings of familiarity (Baayen et al., 2006; Balota et al., 2004; Balota, Pilotti, & Cortese, 2001; Barca et al., 2002; Bates et al., 2001; Cortese & Khanna, 2007), warranting its inclusion in our analyses. Values for rated familiarity (on a 7-point scale) were also drawn from the LEXESP database.

Studies have consistently found that words with larger orthographic neighborhoods are read more quickly (e.g., Andrews, 1989, 1992; Balota et al., 2004; Peereman & Content, 1995; see the review by Andrews, 1997), including in Spanish (Cuetos & Barbón, 2006). Estimates of orthographic neighborhood size (N) were taken from Pérez, Alameda, and Cuetos (2003). The Pérez et al. estimates of N counted as neighbors only those words with a frequency of 1+ occurrences per million in the Alameda and Cuetos (1995) Spanish print corpus of two million words. We transformed the N values to $\log_{10}(N + 1)$ to ameliorate skew.

The effect on oral reading latencies of word length measured in letters has been routinely observed in word naming in transparent orthography languages: in German (Ziegler, Perry, Jacobs, & Braun, 2001), in Italian (Barca et al., 2002; Burani et al., 2007; De Luca, Barca, Burani, & Zoccolotti, 2008; Zoccolotti et al., 2005), and in Spanish (Cuetos & Barbón, 2006). Recent studies have indicated that syllable length also affects performance in visual word recognition (New, Ferrand, Pallier, & Brysbaert, 2006; see also Butler & Hains, 1979) and in word naming (Jared & Seidenberg, 1990; Yap & Balota, 2009; see also Ferrand, 2000). Values for word length in letters, phonemes, and syllables were taken from the LEXESP database.

Evaluations of the influence on reading of the formal orthographic characteristics of words must take care to distinguish the effects of word length or neighborhood size from the potential confounding effect of the frequency of sublexical units (Andrews, 1997; Weekes, 1997). We drew estimates of the average bigram type and token frequency of words from the BuscaPalabras database (Davis & Perea, 2005), where bigram frequency is calculated with respect to the LEXESP corpus. The bigram frequency measures yielded by BuscaPalabras are length and position sensitive. The type frequency for a bigram is the number of different words that include that bigram. This is calculated with respect to words of the same length, with bigrams in the same position. The token frequency for a bigram is the sum of the frequencies of each of the words (of the same length) that include that bigram (in the same position). We drew

summary bigram frequency measures for each word. The average bigram token frequency for a word is the mean given by the sum of the token frequencies of the bigrams constituting that word, divided by the number of bigrams in the word. The average bigram type frequency for a target word is the mean given by the sum of the bigram type frequencies for the bigrams in that word, divided by the number of bigrams in the word. We transformed the bigram token frequency values to the $\log_{10}(\text{bigram token frequency} + 1)$ to ameliorate skew.

Item characteristics are summarized in Table 1, in which we report the mean, SD , minimum, and maximum values for each variable.

Subjective ratings of AoA may be argued to reflect multiple influences on introspective processes, just as much as they can be taken to estimate the ages at which words were learned. Therefore, we checked whether the subjective rated AoA estimates that we used were related to the objective AoA estimates obtained and reported by Álvarez and Cuetos (2007). Álvarez and Cuetos asked children of varying ages to produce word responses to pictured object stimuli (following Morrison, Chappell, & Ellis, 1997). If the children could name the picture, they were assumed to have acquired the word. Objective and subjective AoA estimates were available for only 180 items, insufficient for the former to be used in the analysis of reading latencies. However, we found that the correlation between objective and subjective AoA values was .47 ($p < .001$). This shows that the rated AoA values are a good reflection of the age at which words were learned.

We note that our analyses examined the possibility that the effects of frequency and length were curvilinear. Baayen et al. (2006) found a significant curvilinear effect of frequency, where the frequency effect diminishes for more frequent items, for English lexical decision and oral reading

Table 1 Summary of item characteristics

Variable	Mean	SD	Min.	Max.	N
Logfrequency	1.23	0.49	0.08	2.98	2,764
AoA	4.07	1.26	1.16	6.94	2,764
Imageability	4.65	1.21	1.64	7.00	2,764
Familiarity	5.25	1.07	1.12	7.00	2,764
Letters	6.55	1.81	3.00	10.00	2,764
Phonemes	6.45	1.85	2.00	11.00	2,764
Syllables	2.76	0.83	1.00	5.00	2,764
Log N	0.46	0.40	0.00	1.54	2,764
BF_TP	59.49	37.66	1.00	202.50	2,764
LogBFTK	2.71	0.31	0.66	3.94	2,764

Logfrequency = $\log_{10}(\text{LEXESP frequency per million} + 1)$; $\log N$ = $\log_{10}(N + 1)$, where N = Pérez, Alameda, and Cuetos (2003) orthographic neighborhood size; BF_TP = bigram type frequency; $\log BFTK$ = $\log_{10}(\text{BFTK} + 1)$, where BF_TK = bigram token frequency

latencies (see also Balota et al., 2004). In addition, there is some evidence that the length effect accelerates for longer words (New et al., 2006; Yap & Balota, 2009; but see Baayen et al., 2006). Thus, we examined both linear and nonlinear effects of frequency and length.

Data reduction

The intercorrelation of predictors is an important concern in the analysis of psycholinguistic data. We present a summary in Table 2 of the intercorrelation of variables. Where groups of predictor variables are, as here, highly correlated—where factors are collinear—it is difficult to distinguish in a regression analysis the unique contribution of each variable in accounting for observed variance. This difficulty is associated with problems concerning the interpretation of effects and the lack of stability between samples in their relative importance (Baayen, 2008). Our aim was to identify the important influences on reading from a range of potential factors. We followed Baayen et al. (2006; also, Barca et al., 2002; Bates et al., 2001; Burani et al., 2007) by orthogonalizing predictors using PCA and then performing regression analyses using these extracted PC scores.

A diagnostic measure of collinearity, the condition number (Belsley, Kuh, & Welch, 1980) for these key psycholinguistic variables factors equals 83 if we include just logfrequency, AoA, imageability, familiarity, logN, letters, syllables, phonemes, BFTP, and logBFTK in our assessment. This indicates a dangerous level of multicollinearity (Baayen,

2008). We conducted a PCA over these variables to derive orthogonalized predictors (see Table 3 for a summary of the factor loadings). The PCA resulted in four components that each accounted for more than 5 % and together accounted for 84 % of predictor variance. We label these factors (1) orthographic.form, (2) frequency, (3) semantic, and (4) bigram.frequency, reflecting their relationship to the raw predictor variables, as shown in the factor loadings summary.

In addition, our regression analyses included predictors representing the phonetic characteristics of word initials, as well as the stress patterns of words. This was done to capture variance associated with voice key biases and with stress patterns (Kessler, Mullenix, & Treiman, 2002; Spieler & Balota, 1997), although such coding may also capture variation in the ease of implementation of different phonological codes (Balota et al., 2004). We coded the phonetic characteristics of word initials using a commonly employed scheme where dichotomous variables represented the presence or absence of 13 phonetic features: vowel, alveolar, bilabial, dental, fricative, glottal, labiodental, liquid, nasal, palatal, stop, velar, and voiced. In addition, we coded the accent/stress class of words. Most words in the language (and in our sample) are stressed on the penultimate syllable (the *llana* or paroxytone class); the remaining words are stressed on the ultimate syllable (the *aguda* or oxytone class) or on the antepenultimate syllable (the *esdrújula* or proparoxytone class).

Since the variables coding for word-initial phonetic characteristics and accent/stress class were found to capture

Table 2 Summary of Pearson's *r* bivariate correlations between key psycholinguistic variables (*N* = 2,764 in all comparisons)

		AoA	Imageability	Familiarity	Letters	Phonemes	Syllables	LogN	LogBFTK	BF_TP
Logfrequency	<i>r</i>	-.40	.02	.52	-.16	-.14	-.15	.11	.23	-.13
	<i>p</i> (two-tailed)	<.001	.335	<.001	<.001	<.001	<.001	<.001	<.001	<.001
AoA	<i>r</i>		-.57	-.66	.35	.36	.36	-.33	-.16	.15
	<i>p</i> (two-tailed)		<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
Imageability	<i>r</i>			.33	-.28	-.30	-.25	.27	.01	-.21
	<i>p</i> (two-tailed)			<.001	<.001	<.001	<.001	<.001	.489	<.001
Familiarity	<i>r</i>				-.13	-.13	-.14	.15	.16	.00
	<i>p</i> (two-tailed)				<.001	<.001	<.001	<.001	<.001	.939
Letters	<i>r</i>					.98	.84	-.70	-.13	.64
	<i>p</i> (two-tailed)					<.001	<.001	<.001	<.001	<.001
Phonemes	<i>r</i>						.83	-.69	-.12	.63
	<i>p</i> (two-tailed)						<.001	<.001	<.001	<.001
Syllables	<i>r</i>							-.58	-.19	.49
	<i>p</i> (two-tailed)							<.001	<.001	<.001
LogN	<i>r</i>								.28	-.31
	<i>p</i> (two-tailed)								<.001	<.001
LogBFTK	<i>r</i>									.30
	<i>p</i> (two-tailed)									<.001

Table 3 Summary of principal components analysis (varimax rotation; factors extracted account for more than 5 % of predictor variance) of key psycholinguistic variables, showing loadings of raw variables on components, rotated component matrix

	Component			
	1	2	3	4
Factor name	Orthographic .form	Frequency	Semantic	Bigram. frequency
Logfrequency	-.11	.92	-.07	.06
AoA	.24	-.49	-.72	-.07
Imageability	-.19	-.05	.90	-.07
Familiarity	.00	.73	.50	.09
Letters	.96	-.08	-.12	.04
Phonemes	.96	-.07	-.14	.04
Syllables	.87	-.10	-.11	-.05
Log <i>N</i>	-.76	-.01	.21	.29
LogBFTK	-.17	.16	-.02	.90
BF_TP	.66	-.11	.00	.61

Extraction method, principal component analysis; rotation method, varimax with Kaiser normalization

overlapping portions of item variance in preliminary analyses (condition number = 7×10^{12}), we again extracted orthogonalized component scores in a PCA on phonetic coding and stress variables. The factor loadings for this PCA are reported in Table 4. We extracted seven components, each accounting for more than 5 % of predictor variance, together accounting for about 85 % of total variance among word-initial and accent/stress predictors. We used these PCA-extracted components

(labeled initialstress_x, where $x = 1-7$) as predictors in our regression analyses. The loadings of the extracted components on the word-initial stress coding variables are given in Table 4.

It can be seen that the factors relate to the following: for initialstress_1, +vowel, -velar, +voiced, -occlusive features; for initialstress_2, +alveolar, -occlusive, +liquid; for initialstress_3, +oxytone, -paraoxytone; for initialstress_4, -velar, +bilabial, +nasal, -vowel; for initialstress_5, +velar, -dental, -fricative, +voiced; for initialstress_6, +palatal, -fricative; and for initialstress_7, +proparoxytone. We will not present an account of the effects of initial stress components, because they were entered as control variables.

In sum, we conducted two different PCAs to derive orthogonalized predictors whose use allowed us to substantially exclude multicollinearity for our predictor set. The condition number for the initial stress factors plus the key psycholinguistic components equalled 1.4, indicating no multicollinearity among the predictors.

Apparatus and procedure

The 2,764 words were divided into six experimental lists. Words in each list were further split into six blocks: four blocks of 461 items and two blocks of 460 items. The order of presentation of stimulus blocks, as well as the order of stimuli within blocks, was randomized for each subject.

Stimuli were presented and responses recorded using DMDX (Forster & Forster, 2003) on a Windows XP desktop computer. Subjects were seated 50 cm from the display screen. Words were presented in Arial 10-point type and

Table 4 Summary of principal components analysis (varimax rotation; factors extracted if eigenvalues > 1) of all variables coding word-initial phonetic characteristics and word stress/accent type, showing loadings of raw variables on components, rotated component matrix

Initialstress_x	Component						
	1	2	3	4	5	6	7
Vowel	.75	-.33	.05	-.41	.26	-.13	.01
Alveolar	.05	.93	.00	-.02	-.07	-.11	-.01
Palatal	-.01	.08	-.04	-.09	-.03	.67	-.03
Velar	-.68	-.11	-.01	-.38	.39	-.29	-.05
Bilabial	-.16	-.22	-.02	.85	.15	.08	.04
Dental	-.06	-.23	.01	-.10	-.79	.23	-.01
Voiced	.75	.08	.03	.17	.43	.15	.01
Occlusive	-.75	-.43	.01	.08	.13	.34	.01
Nasal	.20	.08	-.04	.75	.09	-.26	-.04
Fricative	-.09	.27	-.04	-.17	-.74	-.43	-.01
Liquid	.09	.81	.02	-.09	.20	.37	.01
Oxytone	.01	.01	.97	-.03	.01	-.02	-.22
Paroxytone	-.02	-.01	-.96	.03	-.01	.04	-.28
Proparoxytone	.02	.00	.03	.00	.01	-.04	1.00

Extraction method, principal component analysis; rotation method, varimax with Kaiser normalization

subtended between 2.29 ° and 6.43 ° of visual angle (2–5.5 cm). Testing was conducted in a sound-attenuated, dimly lit room. Response latencies were registered by the DMDX software voice key, and one of us sat with subjects during testing to record errors.

An experimental trial had the following sequence of events: (1) A blank gray screen was presented for 512 ms; (2) a black asterisk was presented at the center of the screen for 512 ms; and (3) the target word replaced the asterisk and was presented for 1,536 ms. Each test session lasted 30 min. Each subject completed one list per day. All lists were completed in 8 weeks.

Results

We recorded a total of 69,100 responses. We analyzed only response times (RTs) to correct responses, excluding 3,527 data-points: latencies pertaining to incorrect responses (432 observations), latencies that fell outside the limits set by < 200 ms or > 1,500 ms (2,114 observations), and latencies (981 observations) that fell outside the limits set by, for each subject, the mean $\pm$ 3 SDs of their latencies. The exclusions equalled 5.1 % of the total number of observations, leaving 65,573 observations for analysis.

Data analysis strategy

We transformed latencies to $\log_{10}(\text{RT})$ to ameliorate skew in the RT distribution (Baayen et al., 2006). We analyzed the logRTs using linear mixed-effects modeling. In every model, all predictors were entered simultaneously. As was discussed by Baayen, Davidson, and Bates (2008), mixed-effects modeling incorporates the estimation of replicable “fixed” effects on performance due to key psycholinguistic variables like frequency, while taking into account “random” effects due to unexplained variation between subjects or between items. In the present study, models were fit using the lme4 package (Bates, 2005; version 0.999375-42) in R (R Development Core Team, 2012). We will report estimated coefficients and standard errors of effects of explanatory psycholinguistic variables. Following Baayen (2008; Baayen et al., 2008), we report Markov chain Monte Carlo (MCMC) derived *p*-values. The *p*-values were derived from 10,000 MCMC samples. We note that MCMC- and *t*-derived *p*-values largely coincide for samples as large as our data set (Baayen et al., 2008).

A question arises due to the fact that (1) we presented a large number of stimuli, and (2) in our use of mixed-effects modeling, we analyzed a large number of raw RT observations, rather than a smaller number of averaged by-items latencies. The question is whether this approach might be more prone to indicate (unwarranted) significant effects—that is, whether the increased size of the data set could result in an

increased type I error rate. We think the rationale for collecting responses to a large sample of words has been persuasively argued by Balota et al. (2004). In examining reading performance for a sample of 2,764 words, we avoided the restrictions—on the range of variables, especially—imposed by seeking to select items varying according to a factorial design. Our analyses therefore revealed a more comprehensive picture of reading in Spanish. Perhaps more to the point, we note that Baayen et al. (2008) reported analyses of the numbers of times different analytic methods detected an effect in simulated data runs. Mixed-effects analyses with random effects of items and subjects reported as significantly different (from zero) effects that actually had coefficients of zero about as often as by-items regressions, with both analysis methods shown to have a type I error rate close to the nominal type I error rates of $p = .05$ or $.01$. Critically, however, the mixed-effects analyses were found to detect nonzero effects consistently more often than by-items mean regressions.

Mixed-effects modeling results

We stepped through a series of models. First, assuming the same random effects of subjects and items on intercepts, we compared models differing in fixed effects: a model (model 1) with just initialstress factors; a model (model 2) with initialstress factors plus linear effects due to the orthographic.form, frequency, semantic, and bigram.frequency factors; and lastly, a model (model 3) with the same factors as model 2 but adding restricted cubic splines for the frequency and orthographic.form factors to examine the evidence for the presence of curvilinear effects of frequency and length (the orthographic.form factor loads heavily on length).

We contrasted the utility of the more complex models by comparing them with the simpler (nested) models, using the likelihood ratio test (LRT; Pinheiro & Bates, 2000). Note that, following the discussion in Pinheiro and Bates, the models were fitted using the REML=FALSE setting in lmer. Comparing models 1 and 2, models with initialstress factors but differing in whether they did or did not include key psycholinguistic factors like orthographic.form, the LRT statistic was significant, $\chi^2(4) = 1,007$, $p = 2 \times 10^{-16}$. Comparing models 2 and 3—that is, models with initialstress and key psycholinguistic components but differing in whether they did or did not use restricted cubic splines to fit the orthographic.form and frequency effects—the LRT statistic was significant, $\chi^2(2) = 23$, $p = 1 \times 10^{-5}$.

We evaluated whether the inclusion of random effects was necessary in the final model (model 3), using LRT comparisons between models with the same fixed effects structure but differing random effects. Here, following Pinheiro and Bates (2000; see also Baayen, 2008), models were fitted using the REML=TRUE setting in lmer. We compared models that included (1) both random effects of subjects and items, as

specified for model 3; (2) just the random effect of subjects; and (3) just the random effect of items. The difference between models (1) and (2) was significant, $\chi^2(1) = 185, p = 2 \times 10^{-16}$, indicating that inclusion of an item effect was justified. The difference between models (1) and (3) was significant, $\chi^2(1) = 17,388, p = 2 \times 10^{-16}$, indicating that inclusion of a subject effect was justified.

We then evaluated whether it would be justified to include the random effects of subjects on the slopes of the effects of the psycholinguistic factors. That is, we examined whether there was significant variation between subjects in the shape of the effects. We tested this proposition by fitting a series of models to the data set, each differing from model 3 only in the addition of a random effect of subjects on one psycholinguistic component. In a series of pairwise LRT comparisons, we found that the addition of random slopes terms to model 3 improved fit only for the orthographic.form, $\chi^2(3) = 286, p < 2 \times 10^{-16}$, and frequency, $\chi^2(3) = 38, p = 3 \times 10^{-8}$ effects, not for the semantic and bigram.frequency effects ($p > .05$ in both LRT comparisons). Overall, a model with the same fixed effects as model 3, with random effects of subjects and items on intercepts and random slopes for orthographic.form and frequency, fit the data better than did a model with the same fixed effects and just random effects on intercepts, $\chi^2(6) = 342, p < 2 \times 10^{-16}$. Critically, in this model, the effects of the orthographic.form, frequency, semantic, and bigram.frequency factors remained significant and in the same direction as for model 3.

It is not regarded as optimal to include nonsignificant effects in the final model. Whether or not the model included random intercepts and slopes, the fixed effect of the initialstress_4 component was not significant. Therefore, we refitted the final model without it. We did this twice because the lmer function does not calculate MCMC-derived p -values for models including random slopes. Thus, in the present article, we computed and report estimated coefficients and p -values for fixed effects in the final model with just significant fixed effects (cf. model 3) and just random effects of subjects and items on intercepts. A summary of this model is reported in Table 5. However, we also fitted a model with just significant fixed effects (cf. model 3) and random intercepts, as well as random slopes, and give a summary of this model in the Supplementary Materials. In neither model did the removal of the nonsignificant effect change the significance or direction of the other effects.

The effects in the final model are clearly shown by partial effects plots (see Fig. 1). We note that the plots are all scaled over the same range of y values so that the partial effects plots indicate relative effect size. It can be seen that the largest effects are due to the initialstress_6, frequency, and orthographic.form factors, although the semantic effect is also quite large. The curvilinear effects of frequency and orthographic.form are clearly evident. The frequency effect decreases in

magnitude for more frequent, more familiar, earlier-acquired words. The orthographic.form effect increases for longer words with fewer neighbors.

The analyses we have discussed to this point have not directly included information about the grammatical class or the morphological type of the stimuli. Previous research has demonstrated that reading performance is influenced by morphological factors. For example, Baayen et al. (2006) showed that English word-naming latencies are significantly affected by a word's inflectional entropy (Moscoso del Prado Martin, Kostic, & Baayen, 2004), a measure of morphological family size.

It is important that we establish whether the effects we have found persist when we control for variation in word class and morphological complexity. At present, information-theoretic measures of a word's morphological characteristics, measures like inflectional entropy, are not available for Spanish. We therefore added two simple measures to our data set to capture critical aspects of item variation: a variable coding for grammatical class, and a variable coding for whether an item was morphologically simple, a derivation, or an inflection.¹ As before, we compared REML=FALSE mixed-effects models varying in fixed effects. We did this twice, once for the final model with random intercepts, once for the final model with both random intercepts and random slopes. In both cases, we found that the addition of the coding variables—the word class and the morphological type factors—was justified by improved capacity of the model to fit the data [for both comparisons, $\chi^2(8) = 26, p = .001$]. We report a summary of the random intercepts model including the word class and morphological coding variables in the Supplementary Materials. However, we note that the effect of grammatical class stems from the fact that nouns and verbs are read significantly more slowly than adjectives ($p < .05$). We also note that morphological derivations are read significantly more slowly than morphologically simple words or inflected words ($p < .05$). Critically, addition of these factors did not change the direction or significance of the other effects.

Discussion

We report a study examining the influence of AoA on oral reading in Spanish. Consistent with findings in a number of languages, we observed significant effects relating word-naming latencies to (1) the phonetic characteristics of word-initial phonemes and word stress pattern (the initialstress PCA factors), (2) the length of items and their orthographic similarity to other words (the orthographic.form and bigram.frequency factors), and (3) the familiarity of words (the frequency factor). What is newsworthy about our results is

¹ We are grateful to Dr. Veronica Moreno for help in adding these data.

Table 5 Summary of linear mixed effects model of log (RT) including all initial-stress principal component analysis (PCA) factors and all psycholinguistic PCA component predictors: Final model

Fixed effects:				
	Estimate	SE	<i>t</i>	<i>p</i> MCMC
(Intercept)	2.6934	0.0083	322.6	.0001
Initialstress_1	−0.0016	0.0004	−4.6	.0001
Initialstress_2	0.0011	0.0003	3.3	.0010
Initialstress_3	−0.0019	0.0004	−5.4	.0001
Initialstress_5	0.0008	0.0003	2.3	.0176
Initialstress_6	0.0036	0.0003	10.4	.0001
Initialstress_7	0.0007	0.0003	2.1	.0340
rcs(orthographic.form ,3)orthographic.form	0.0071	0.0009	7.6	.0001
rcs(orthographic.form, 3)orthographic.form'	0.0032	0.0012	2.7	.0060
rcs(frequency, 3)frequency	−0.0102	0.0009	−11.5	.0001
rcs(frequency, 3)frequency'	0.0043	0.0011	3.9	.0001
Semantic	−0.0034	0.0004	−9.6	.0001
Bigram.frequency	−0.0015	0.0004	−4.2	.0001
Random effects:				
Groups			Variance	SD
Item (intercept)			0.0001	0.0099
Subject (intercept)			0.0017	0.0414
Residual			0.0053	0.0730

Number of observations: 65,573; groups = 2,764 items and 25 subjects

the finding, in addition, of the effect of (4) a factor (the semantic factor) related to variables, imageability, familiarity, and AoA, whose effects previous research has linked to lexical semantic processing. We will discuss our account of this semantic effect below but first will address the other findings, because they set the context for our explanation.

We found a curvilinear effect of length, captured through the use of restricted cubic splines to detect both linear and nonlinear effects of orthographic.form characteristics. Words with high values in the PCA-derived form factor are longer and have fewer neighbors. The effect on oral reading latencies of word length measured in letters (Balota et al., 2004; Weekes, 1997; Ziegler et al., 2001; Zoccolotti et al., 2005), in which longer words elicit longer naming latencies, is conventionally argued to reflect the serial phonological encoding of graphemes (Weekes, 1997). Our observation of a curvilinear length effect replicates, to some extent, previous findings in other languages (Ferrand et al., 2010; New et al., 2006). In such studies, latencies decrease as short words get longer (two to four letters), do not change for middle lengths, and increase as longer words get longer (nine letters or more). We found simply that the length effect was greater for longer words.

Potential explanations for a greater length effect for longer words have been proposed by New et al. (2006) to include (1) the decrease in visual acuity with increased distance from fixation in the letter string and (2) the possibility that longer

words will require refixation. A potential explanation for the difference in the shape of the curve, comparing our report and previous reports, may derive from the difference between the tasks used in our study and those used in previous studies. While we analyzed the impact of length (orthographic.form) on word naming, Ferrand et al. (2010) and New et al. (2006) examined the effect of length on visual word recognition (lexical decision). In particular, it seems likely that the centrality of phonological coding to word naming may have ensured the appearance of an accelerating length effect in our data that contrasts with the U-shaped length effect reported previously (Ferrand et al., 2010; New et al., 2006). The shape of the curvilinear length effect in word naming can be hypothesized to reflect the combined impact of the factors identified by New et al. and the time cost of serial phonological encoding of graphemes (Weekes, 1997).

Studies have consistently found that words with larger orthographic neighborhoods are read more quickly (Andrews, 1997). A connectionist account of reading would explain the facilitatory impact of neighborhood size in English by assuming that training for words with similar patterns will strengthen shared connections and that this advantage will be greater for members of larger neighborhoods (Seidenberg & McClelland, 1989). Our study contributes further evidence that orthographic neighborhood size affects reading in transparent orthographies (cf. Barca et al., 2002; Bates et al., 2001; Burani et al., 2007).

Fig. 1 Plots showing the partial effects of predictors in the final model. Predictors are orthogonalized factors derived through principal components analyses. Note that we back-transformed the $\log_{10}(\text{RT})$ values to raw RTs so estimated effects are more easily interpretable

The orthographic.form effect was observed together with the bigram.frequency effect, which captured the influence on naming performance of the type and token frequency of the constituent bigrams of words. Andrews (1997) considered the possibility that the effect of neighborhood size—typically, facilitatory in word naming—may be due to the high frequency of spelling–sound correspondences in words with many neighbors. The observation of both form and bigram.frequency effects reveals distinct phonological coding processes, operating over both lexical and sublexical (at minimum, bigram and grapheme) units. These processes are evidently facilitated by similarity between the target item and other words with which it may share letters, bigrams, or more complex features.

The effect of word frequency is commonly taken to indicate the influence of lexical knowledge on reading (within the framework of, for example, the dual-route cascaded model;

Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). In connectionist models of reading, no lexical representations are postulated, but the frequency effect is explained by the impact of word-specific experience on the weights on connections between representations in a distributed network (Plaut et al., 1996; Seidenberg & McClelland, 1989; Zevin & Seidenberg, 2002). In connectionist simulations, AoA and frequency effects arise naturally where networks of distributed representations learn through a gradual process of connection weight change in reaction to the kind of interleaved, cumulative experience seen in real-life reading development (P. Monaghan & Ellis, 2010). Frequent exposure to a pattern strengthens representations that are required to respond to it, benefitting performance.

A novel contribution of our study is to extend to Spanish the observation of nonlinear frequency effects previously reported for English (Baayen et al., 2006) and French (Ferrand et al., 2010). We found that the frequency effect

is smaller for more frequent words. A curvilinear frequency effect can be ascribed to the decreasing gains associated with practice in connectionist accounts of reading. As the adaptation of connections by experience drives network efficiency toward asymptote, the further effect of practice will have less impact (Zevin & Seidenberg, 2002).

It is noteworthy that a curvilinear frequency effect of similar form is apparent in both word naming (in our study and in Baayen et al., 2006) and visual word recognition (Baayen et al., 2006; Ferrand et al., 2010). This contrasts with the difference in the form of the curvilinear length effect, comparing our word-naming results with previous lexical decision results, discussed in the foregoing. The contrast between the task invariance (to some extent) of the shape of the curvilinear frequency effect and the apparent task dependence of the shape of the curvilinear length effect may be taken to highlight a distinction between the loci of the effects in the reading system. At minimum, while the length effect may reflect multiple factors concerning peripheral processes (impact of decreasing acuity on visual decoding or number of graphemes on phonological encoding), the frequency effect arguably reflects more central processes involving lexical representations in a localist system (Coltheart et al., 2001) or subsisting in the experience-dependent connection weights of a distributed system (e.g., Zevin & Seidenberg, 2002).

A part of the effect of the frequency factor was due to the contribution of word familiarity, consistent with previous research linking the two variables (Baayen et al., 2006). The impact of familiarity may owe something to its reflection of semantic attributes, whether meaningfulness (Balota et al., 2001) or semantic familiarity (Baayen et al., 2006). This possibility is demonstrated by the observation, in the present and in previous studies (Barca et al., 2002; Bates et al., 2001), that both rated familiarity and rated AoA load on both the frequency and the semantic components of PCA analyses of predictor variance. Most interesting, we found that a semantic factor—capturing the impact of AoA, imageability, and familiarity—affected word-naming latencies over and above both word and bigram frequency effects. The observation of a frequency-independent AoA effect is inconsistent with the hypothesis that previous reports of an AoA effect in Spanish (Cuetos & Barbón, 2006) were due to a confound between frequency and AoA—that is, that the Spanish AoA effect was actually a frequency effect. Our results now show that a frequency effect can be identified in both Spanish and Italian. The addition of the frequency-independent AoA effect serves to constrain the contrast between Spanish and Italian results precisely to the observation of an AoA effect in Spanish but not Italian.

There are a range of possible methodological reasons for the contrast, but they are difficult to evaluate without an experimental investigation in which item and subject characteristics are matched. One issue is highlighted by the

results of the analysis to check the impact of grammatical and morphological coding variables. In the Italian studies cited (Barca et al., 2002; Bates et al., 2001; Burani et al., 2007), all stimuli were morphologically simple nouns. The language sample for the present study included nouns, verbs, and adjectives, some of which were morphological derivations or inflections. The results of the analysis including variables coding for this variation demonstrated that both grammatical class and relative morphological complexity influenced word-naming latencies. However, the analyses also showed that the size and direction of the frequency and semantic effects were unaffected by the addition of these variables. Future research would elucidate cross-linguistic influences on reading by exploring the impact of word class and variation in morphological connectivity (Baayen et al., 2006).

An alternative account for the cross-linguistic contrast is suggested by the observation that, while the Spanish (Cuetos & Barbón, 2006) items were all high-imageability object names, the Italian items included abstract words. It is possible that the difference in item imageability was critical because reading high-imageability items was associated with a greater reliance on semantics for phonological coding and that it was that greater reliance on semantics that resulted in a larger AoA effect in the Spanish than in the Italian data. Consistent with this view, Raman's (2006) report of an AoA effect on word naming in Turkish concerns performance in response to high-imageability words varying in AoA but matched on frequency. In Cuetos and Barbón (2006) and Raman (2006), we have, then, observations of an AoA effect on word naming in two different transparent orthographies in two studies with the common factor that stimuli consisted of high-imageability words.

It is commonly assumed that imageable words are processed more semantically because they have richer semantic representations and connections and activate semantic features to a larger extent than do lower imageability items (Jones, 1985; Pexman, Hargreaves, Siakaluk, Bodner, & Pope, 2008). Thus, high-imageability words may induce access to semantics and activate semantic mediation in reading more than do low-imageability words (Buchanan, Westbury, & Burgess, 2001; Pexman, Lupker, & Hino, 2002). This would allow the AoA effect to appear in reading, to the extent that it is linked to semantic involvement in phonological coding.

The finding that the AoA effect is stronger for words that are more likely to be read semantically—high-imageability items—is consistent with the observations of several researchers, as noted above. In our results, the semantic effect is smaller than the frequency effect. Therefore, it may be that where the language sample is relatively small, the effect of raw frequency may be found to dominate. It may be that a sample of the order of 626 words (as in Barca

et al., 2002) is not sufficient to detect the influence of a relatively small AoA effect (or of a semantic effect linked to AoA). It may be that previous language samples in Italian have, by chance, sampled more words (e.g., abstract items) that are unlikely to induce semantic reading. These conjectures warrant further research; however, the key point is that the influence of semantics on word naming is more general than has previously been surmised. This is because the present study shows the influence of semantics on Spanish word naming for a larger and (on average) less imageable set of items than does that employed by Cuetos and Barbón (2006).

The effect of the semantic factor influenced Spanish reading latencies independently of the effect of the frequency. It has been argued, notably by Brysbaert and Ghyselinck (2006) and by Bates et al. (2001), that the effect of AoA—that part of it that can be distinguished from the effect of frequency—may reflect the involvement of semantics. This view is consistent with the results of our PCA and with the results of similar analyses reported by Burani and colleagues (Barca et al., 2002; Bates et al., 2001; Burani et al., 2007). In those analyses, raw frequency, familiarity, and AoA variables load heavily on a “frequency” factor, while AoA, imageability, and familiarity load heavily on a “semantic” factor. The PCA results are congruent with Brysbaert and Ghyselinck’s proposal that the AoA effect is “partly frequency-related, partly frequency-independent.” These authors argued that the frequency-independent AoA effect might have its locus either at the connections to and from semantics (the arbitrary mappings hypothesis; Zevin & Seidenberg, 2002) or at the semantic level (due to the importance of AoA to semantic knowledge development; Steyvers & Tenenbaum, 2005). We would argue that previous research leads us to the view that a frequency-independent effect of AoA indicates that semantics is involved in oral reading in Spanish. Given the requirements of word naming, we propose that this influence is likely to be located at the mapping between semantics and phonology.

We suggest that our observation of a semantic effect on word naming in Spanish not only tells us something about reading in transparent orthographies, but also has interesting implications for our understanding of the influence of semantics in reading in general. These implications stem from the contrast between our findings and the assumptions of existing accounts of semantic effects in reading in English.

Semantic involvement in word naming in English has been observed in most studies only with semantic information operationalized as imageability and with the effect of imageability observed only for low-frequency words with atypical (irregular or inconsistent exception) OP mappings (Strain et al., 1995; Woollams, 2005). The interaction between imageability and OP typicality has

been explained by supposing (Strain et al., 1995) (1) that words with atypical OP mappings were phonologically coded in a slow and noisy fashion, (2) that connectionist models of reading aloud afforded the possibility that semantic activation could cascade to affect phonological activation, (3) that high-imageability words have richer and more readily computable meanings, and thus (4) that low-frequency exception words were coded sufficiently slowly for semantics to influence phonology, so that, if those words were high in imageability, a semantic influence would be felt.

All words might activate semantics from orthographic processing, but the account just outlined predicts a more prominent semantic influence in word naming for low-frequency words with atypical OP mappings. This expectation is supported by computational simulations. In the connectionist framework, a network can learn to read all English words accurately in the absence of a semantic contribution (Plaut et al., 1996; Seidenberg & McClelland, 1989; see also P. Monaghan & Ellis, 2010). However, simulations reported by Plaut et al. (1996) show that if a semantic influence on phonology is permitted, the network will develop a division of labor in which the phonology of low-frequency OP atypical words will be jointly activated by OP and orthography-to-semantics-to-phonology mappings. That strong prediction follows from the idea that a network will learn to use all available information, learning from experience to read as efficiently as possible.

Nevertheless, Balota et al. (2004) reported that word-naming latencies for a large sample of words were affected by imageability even after controlling for effects of frequency and spelling–sound consistency. These authors suggested that the effect of imageability reflects the cascaded activation of meaning, early in lexical processing. This activation of semantics contributes to the activation of (phonological) output representations in word naming, apparently irrespective of OP typicality. As we have discussed, previous research by Raman and Baluch (2001) has also shown that imageability affects word-naming latencies (for highly skilled readers) in Turkish, a language with entirely regular OP mappings. Overall, our findings appear to be consistent with a line of research in which the influence of semantics on reading may be detected quite broadly across variation in OP mapping predictability. Such an influence can be envisaged to occur as the consequence of activation cascading forward from semantics to phonology in a highly interactive reading system (Balota et al., 2004; Raman & Baluch, 2001).

Acknowledgments This investigation was funded by Grant MCI-PSI2009-09299 from the Spanish government.

References

- Alameda, J. R., & Cueto, F. (1995). *Diccionario de frecuencias de las unidades lingüísticas del Castellano* [Frequency dictionary of the linguistic units of Castellano]. Oviedo: Servicio de Publicaciones de la Universidad de Oviedo.
- Álvarez, B., & Cueto, F. (2007). Objective age of acquisition norms for a set of 328 words in Spanish. *Behavior Research Methods*, 39, 377–383.
- Andrews, S. (1989). Frequency and neighbourhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15, 802–814.
- Andrews, S. (1992). Frequency and neighbourhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 234–254.
- Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighbourhood conflicts. *Psychonomic Bulletin and Review*, 4, 439–461.
- Baayen, R. H. (2008). *Analyzing linguistic data*. Cambridge University Press.
- 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.
- Baayen, R. H., Feldman, L. B., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language*, 55, 290–313.
- Balota, D., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, 133, 283–316.
- Balota, D. A., Pilotti, M., & Cortese, M. J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory & Cognition*, 29, 639–647.
- Baluch, B., & Besner, D. (1991). Visual word recognition: Evidence for strategic control of lexical and nonlexical routines in oral reading. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 17, 644–652.
- Barca, L., Burani, C., & Arduino, L. S. (2002). Word naming times and psycholinguistic norms for Italian nouns. *Behaviour Research Methods, Instruments and Computers*, 34, 424–434.
- Bates, D. M. (2005). Fitting linear mixed models in R. *R News*, 5, 27–30.
- Bates, E., Burani, C., D'Amico, S., & Barca, L. (2001). Word reading and picture naming in Italian. *Memory and Cognition*, 29, 986–999.
- Beaton, A. A., & Wyn Davies, N. (2007). Semantic errors in deep dyslexia: Does orthographic depth matter? *Cognitive Neuropsychology*, 24, 313–323.
- Belke, E., Brysbaert, M., Meyer, A. S., & Ghyselinck, M. (2005). Age of acquisition effects in picture naming: Evidence for a lexical-semantic competition hypothesis. *Cognition*, 96, B45–B54.
- Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity*. New York: Wiley.
- Brysbaert, M., & Ghyselinck, M. (2006). The effect of age-of-acquisition: Partly frequency-related; partly frequency-independent. *Visual Cognition*, 13, 992–1011.
- Brysbaert, M., Lange, M., & Van Wijnendaele, I. (2000). The effects of age-of-acquisition and frequency-of-occurrence in visual word recognition: Further evidence from the Dutch language. *European Journal of Cognitive Psychology*, 12, 65–85.
- Brysbaert, M., Van Wijnendaele, I., & de Deyne, S. (2000). Age-of-acquisition effects in semantic processing tasks. *Acta Psychologica*, 104, 215–226.
- Buchanan, L., Westbury, C., & Burgess, C. (2001). Characterising semantic space: Neighborhood effects in word recognition. *Psychonomic Bulletin and Review*, 8, 531–544.
- 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.
- Butler, B., & Hains, S. (1979). Individual differences in word recognition latency. *Memory and Cognition*, 7, 68–76.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioural sciences* (3rd ed.). Mahwah: Lawrence Erlbaum Associates.
- 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., & Khanna, M. M. (2007). Age of acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables: An analysis of 2,342 words. *The Quarterly Journal of Experimental Psychology*, 60, 1072–1082.
- Cueto, F., & Barbón, A. (2006). Word naming in Spanish. *European Journal of Cognitive Psychology*, 18, 415–436.
- Davies, R., & Cueto, F. (2005). Acquired dyslexia in Spanish: A review and some observations on a new case of deep dyslexia. *Behavioural Neurology*, 16, 85–101.
- Davies, R. A. I., Cueto, F., & Rodríguez-Ferreiro, J. (2010). Recovery in reading: A treatment study of acquired dyslexia in Spanish. *Aphasiology*, 24, 1115–1131.
- Davis, C. J., & Perea, M. (2005). BuscaPalabras: A program for deriving orthographic and phonological neighbourhood statistics and other psycholinguistic indices in Spanish. *Behavior Research Methods*, 37, 665–671.
- De Luca, M., Barca, L., Burani, C., & Zoccolotti, P. (2008). The effect of word length and other sublexical, lexical, and semantic variables on developmental reading deficits. *Cognitive and Behavioral Neurology*, 21, 227–235.
- del Prado, M., Martin, F., Kostic, A., & Baayen, R. H. (2004). Putting the bits together: An information theoretical perspective on morphological processing. *Cognition*, 94, 1–18.
- Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26, 1103–1123.
- Ferrand, L. (2000). Reading aloud polysyllabic words and nonwords: The syllabic length effect reexamined. *Psychonomic Bulletin and Review*, 7, 142–148.
- Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., et al. (2010). The French lexicon project: Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods*, 42, 488–496.
- Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, 35, 116–124.
- 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 multi-task investigation. *Acta Psychologica*, 115, 43–67.
- Jared, D., & Seidenberg, M. S. (1990). Naming multisyllabic words. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 92–105.
- Jones, G. V. (1985). Deep dyslexia, imageability, and ease of predication. *Brain and Language*, 24, 1–19.
- Kessler, B., Treiman, R., & Mullenix, J. (2002). Phonetic biases in voice key response time measurements. *Journal of Memory and Language*, 47, 145–171.
- Lambon Ralph, M. A., & Ehsan, S. (2006). Age of acquisition effects depend on the mapping between representations and the

- frequency of occurrence: Empirical and computational evidence. *Visual Cognition*, 13, 928–948.
- Monaghan, J., & Ellis, A. W. (2002). What exactly interacts with spelling-sound consistency in word naming? *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28, 183–206.
- Monaghan, P., & Ellis, A. W. (2010). Modeling reading development: Cumulative, incremental learning in a computational model of word naming. *Journal of Memory and Language*, 63, 506–525.
- Morrison, C. M., Chappell, T. D., & Ellis, A. W. (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*, 50, 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 and Review*, 13, 45–52.
- Paizi, D., Burani, C., & Zoccolotti, P. (2010). List context effects in reading Italian words and nonwords: Can the word frequency effect be eliminated? *European Journal of Cognitive Psychology*, 22, 1039–1065.
- Peereman, R., & Content, A. (1995). Neighborhood size effect in naming: lexical activation or sublexical correspondences? *Journal of Experimental Psychology: Learning, Memory and Cognition*, 21, 409–421.
- 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.
- Pexman, P. M., Hargreaves, I. S., Siakaluk, P. D., Bodner, G. E., & Pope, J. (2008). There are many ways to be rich: Effects of three measures of semantic richness on visual word recognition. *Psychonomic Bulletin and Review*, 15, 161–167.
- Pexman, P. M., Lupker, S. J., & Hino, Y. (2002). The impact of feedback semantics in visual word recognition: Number of features effects in lexical decision and naming tasks. *Psychonomic Bulletin and Review*, 9, 542–549.
- Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-plus (Statistics and Computing)*. 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 (2012). R: A language and environment for statistical computing (Version 2.15.1) [Computer software]. Vienna: R foundation for statistical computing.
- Raman, I. (2006). On the age-of-acquisition effects in word naming and orthographic transparency: Mapping specific or universal? *Visual Cognition*, 13, 1044–1053.
- Raman, I. (2011). The role of age of acquisition in picture and word naming in dyslexic adults. *British Journal of Psychology*, 102, 328–339.
- Raman, I., & Baluch, B. (2001). Semantic effects as a function of reading skill in word naming of a transparent orthography. *Reading & Writing. An Interdisciplinary Journal*, 14, 599–614.
- Raman, I., Baluch, B., & Besner, D. (2004). On the control of visual word recognition: Changing routes versus changing deadlines. *Memory and Cognition*, 32, 489–500.
- Raman, I., & Weekes, B. (2005). Acquired dyslexia in a Turkish-English speaker. *Annals of Dyslexia*, 55, 79–104.
- Sebastián, N., Martí, M. A., Carreiras, M. F., & Cuetos, F. (2000). *LEXESP, Léxico informatizado del español [Computerized lexicon of Spanish]*. Barcelona: Ediciones de la Universitat de Barcelona.
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96, 523–568.
- Spieler, D. H., & Balota, D. A. (1997). Bringing computational models of word naming down to the item level. *Psychological Science*, 8, 411–416.
- Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, 29, 41–78.
- Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 21, 1140–1154.
- Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *The Quarterly Journal of Experimental Psychology*, 510A, 439–456.
- Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular reading system. *Applied Psycholinguistics*, 14, 1–33.
- Wimmer, H. (2006). Don't neglect reading fluency! *Developmental Science*, 9, 447–448.
- Woollams, A. (2005). Imageability and ambiguity effects in speeded naming: Convergence and divergence. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 31, 878–890.
- Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language*, 60, 502–529.
- Zevin, J. D., & Seidenberg, M. S. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory and Language*, 47, 1–29.
- Ziegler, J. C., Perry, C., Jacobs, A. M., & Braun, M. (2001). Identical words are read differently in different languages. *Psychological Science*, 12, 379–384.
- Zoccolotti, P., De Luca, M., Di Pace, E., Gasperini, F., Judica, A., & Spinelli, D. (2005). Word length effect in early reading and in developmental dyslexia. *Brain and Language*, 93, 369–373.