Iconicity ratings for 10,995 Spanish words and their relationship with psycholinguistic variables

J. A. Hinojosa [1,2,3] · J. Haro [4] · S. Magallares [1] · J. A. Duñabeitia [3,5] · P. Ferré [4]

Abstract

The study of iconicity, or the resemblance between word forms and their meanings, has been the focus of increasing attention in recent years. Nevertheless, there is a lack of large-scale normative studies on the iconic properties of words, which could prove crucial to expanding our understanding of form–meaning associations. In this work, we report subjective iconicity ratings for 10,995 visually presented Spanish words from 1350 participants who were asked to repeat each of the words aloud before rating them. The response reliability and the consistency between the present and previous ratings were good. The relationships between iconicity and several psycholinguistic variables were examined through multiple regression analyses. We found that sensory experience ratings were the main predictor of iconicity, and that early-acquired and more abstract words received higher iconicity scores. We also found that onomatopoeias and interjections were the most iconic words, followed by adjectives. Finally, a follow-up study was conducted in which a subsample of 360 words with different levels of iconicity from the visual presentation study was auditorily presented to the participants. A high correlation was observed between the iconicity scores in the visual and auditory presentations. The normative data provided in this database might prove useful in expanding the body of knowledge on issues such as the processing of the iconic properties of words and the role of word-form associations in the acquisition of vocabularies. The database can be downloaded from https://osf.io/v5er3/.

Introduction

The debate surrounding the relationship between the sound and the meaning of words has a long historical tradition going back to early antiquity. In the work *On Interpretation*, Aristotle outlined his concept of a linguistic sign as an arbitrary convention between sounds and meanings. This idea became dogmatic when one of the founders of modern linguistics, Ferdinand de Saussure, established that a central property of natural language is the capacity of linguistic symbols to combine into limitless conventional forms of the sign. Thus, arbitrariness would allow unlimited possibilities for communication and explain form differences across languages to denote the same concepts (Lockwood & Dingemanse, 2015). However, this view has been challenged by the findings of Sapir (1929) and Köhler (1947) on mappings between vowel/consonant types and the shape or size of pictorial stimuli. A key observation was the *maluma/takete* effect, which refers to the association between nonce words and round and sharp shapes (Köhler, 1929; also known as the *Bouba/Kiki* effect, Ramachandran & Hubbard, 2001). Other studies have shown additional types of patterns in form–meaning associations, such as the use of individual phonemes in mapping motion, brightness, distance, or even emotion (Adelman, Estes, & Cossu, 2019; Cuskley, 2013; Sapir, 1929; Schmidtke & Conrad, 2018; Tanz, 1971; Thompson & Estes, 2011). Furthermore, the existence of words with vivid sensory links has been demonstrated in many languages (Dingemanse, Schuerman, Reinisch, Tufvesson, & Mitterer,

✉ J. A. Hinojosa
hinojosa@ucm.es

[1] Instituto Pluridisciplinar, Universidad Complutense de Madrid, Paseo Juan XXIII, 1, 28040 Madrid, Spain

[2] Dpto. Psicología Experimental, Procesos Cognitivos y Logopedia, Universidad Complutense de Madrid, Madrid, Spain

[3] Centro de Ciencia Cognitiva - C3, Universidad Nebrija, Madrid, Spain

[4] Universitat Rovira i Virgili, Department of Psychology and CRAMC, Tarragona, Spain

[5] Department of Languages and Culture, The Arctic University of Norway, Tromsø, Norway

2016; Vigliocco & Kita, 2006; Winter, Perlman, Perry, & Lupyan, 2017; see, Ahlner & Zlatev, 2010; Lockwood & Dingemanse, 2015; Sidhu & Pexman, 2018; Thompson and Do, 2019, for reviews).

In recent years the term iconicity has become the most fitting cover-all term to define the resemblance-based mapping between the form of a linguistic sign and the object or idea it represents (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015; Lockwood & Dingemanse, 2015; see also Dingemanse, 2018, Elsen, 2017 and Nielsen & Dingemanse, 2020, for theoretical considerations about other aspects of the relationship between word forms and meanings). The prototypical examples of iconic words or ideophones are onomatopoeias (words that phonetically resemble the sound that they describe, e.g., *plop*). It has been suggested that iconicity benefits language learning (Imai & Kita, 2014) and communication by making language more direct and vivid (Lockwood & Dingemanse, 2015). In fact, some theoretical views have emphasized that arbitrariness and iconicity are two co-existing aspects of language (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015; Dingemanse, Perlman, & Perniss, 2020; Lockwood & Dingemanse, 2015; Perniss & Vigliocco, 2014).

Early studies on iconicity (e.g., Davis, 1961; Miro, 1961; Taylor & Taylor, 1962) used nonwords as stimuli because they allow for the careful experimental control of linguistic variables. However, as noted by Lockwood and Dingemanse (2015), language properties that are found to be iconic based on evidence from these experiments might not resemble those that can be found in natural languages. Therefore, research needs to be conducted that uses existing words with different degrees of iconicity to investigate sound symbolism and how sensory properties modulate natural language processing. In recent years, a growing body of behavioral and neuroimaging research has revealed a variety of iconic effects in word processing. In this sense, iconicity in abstract words was found to elicit more "concrete" responses in abstract/concrete semantic decision tasks, which possibly reflects the activation of iconicity-related semantic activation that influenced participants to make incorrect "concrete" responses (Lupyan & Winter, 2018). Also, iconicity has been proven beneficial for the lexical processing of visually presented words in normal (Sidhu, Vigliocco, & Pexman, 2020) and aphasic (Meteyard, Stoppard, Snudden, Cappa & Visgglioco, 2015) individuals. Similarly, reduced N400 responses for words with iconic mapping between forms and meanings relative to arbitrary words were observed in an event-related potentials (ERPs) study, which suggests a processing advantage for highly iconic words (Peeters, 2016). Larger N400 effects have also been reported for onomatopoeias preceded by arbitrary words in a semantic relatedness task, suggesting enhanced semantic processing of onomatopoeias, which in turn allows for improved detection of the mismatch between primes and targets

(Vigliocco, Zhang, del Maschio, Todd, & Tuomainen, 2020). In another study, Lockwood and Tuomainen (2015) observed facilitated processing of iconic adverbs compared to arbitrary adverbs as indexed by enhanced P2 and late positive component effects. Finally, functional magnetic resonance imaging (fMRI) studies have shown that the processing of iconic words increases the activation of sensory brain regions relative to the processing of more arbitrary words (Hashimoto et al., 2006; Kanero, Imai, Okuda, & Matsuda, 2014), and that affective iconic words elicited enhanced amygdala activations, which were modulated by the activation of brain regions related to the processing of sound and meaning (Aryani, Hsu, & Jacobs, 2019). Overall, the literature summarized here suggests that these effects of iconicity on language processing are a promising avenue for future research.

Studies on language processing rely heavily on the availability of data sets from normative studies on a number of variables that have been found to impact language production and comprehension. To give just a few examples, large data sets are currently available in many languages for variables such as word frequency (e.g., Brysbaert & New, 2009; Duchon, Perea, Sebastián-Gallés, Martí, & Carreira, 2013), concreteness (e.g., Brysbaert, De Deyne, Voorspoels, & Storms, 2014; Coso, Guasch, Ferré, & Hinojosa, 2019), age of acquisition (e.g., Alonso, Fernández, & Díez, 2015; Kuperman, Stadthagen-González, & Brysbaert, 2012), familiarity (e.g., Guasch, Ferré, & Fraga, 2017; Stadthagen-González & Davis, 2006), valence (e.g., Monnier and Syssau, 2014; Warriner, Kuperman, & Brysbaert, 2013), imageability (e.g., Della Rosa, Catricalà, Vigliocco, & Cappa, 2010; Soares, Costa, Machado, & Comesaña, 2017), and sensory experience (e.g., Juhasz & Yap, 2013). The abundance of such data sets is at odds with the scarce number of normative studies on the iconic features of words. To date, iconicity ratings from only three studies are available to researchers. Perry, Perlman, and Lupyan (2015) collected iconicity ratings for 592 English and 638 Spanish words from the MacArthur-Bates Communicative Developmental Inventories and established a correlation between participants' scores for written and auditorily presented words. In a subsequent study, this data set was supplemented with additional scores for 2409 English words that were visually presented to participants (Winter et al., 2017). In these studies, the authors found a relationship between iconicity and sensory experience ratings (SERs) and age of acquisition (AoA): Highly iconic words were learned earlier and showed higher SERs than less iconic words. Also, onomatopoeias were found to be the most iconic words, followed by verbs and adjectives. Interestingly, whereas verbs were more iconic than nouns in English, these differences were not found in Spanish. Finally, using a different approach, Xiao and Treiman (2012) reported norms for 213 Chinese words. These authors presented trials with an English word or phrase together with two Chinese

characters to English participants who did not know Chinese. Participants were asked to guess which of the two Chinese characters corresponded to the English word or phrase. The proportion of correct responses for a given character was taken as a measure of its degree of iconicity.

From the literature reviewed above, it seems that normative studies on the iconic features of a large sample of words are still needed. By making these data sets available, researchers will be able to further investigate questions that might be relevant in psychological or educational contexts, such as the effects of iconicity in language processing or its role in the acquisition of new words by children or second language learners. Thus, in the present work, we conducted a normative study with the aim of collecting iconicity ratings for a large sample of Spanish words. Of note, the current study deals with the notion of subjective iconicity, which refers to the resemblance between word form and meaning as perceived by participants (Taylor & Taylor, 1965). When subjectively judging the resemblance between word forms and meanings, people may rely on heuristic processes that bias their judgments. In this vein, there is evidence demonstrating that words are perceived as fitting their referents based on heuristics that shape people's understanding of why objects have their names and are used to make sense of the world more generally (Cimpian & Salomon, 2014; Sutherland & Cimpian, 2015). Nonetheless, the question of what perceptuo-motor analogy motivates high subjective iconic relationships remains elusive. In contrast, objective iconicity is defined as the regularity in the distribution of sounds in a language (Taylor & Taylor, 1965). Objective measures of similarity between forms and their meanings investigate sound–meaning associations with a focus on identifying statistical regularities across languages (Blasi et al., 2016; Motamedi, Little, Nielsen & Sulik, 2019). However, the results of these studies demonstrate only that sound–meaning association distributions across languages are statistically reliable, and they do not provide evidence that these relationships are iconic in essence (Motamedi et al., 2019).

Words were visually presented and participants were asked to articulate each word before rating it. Based on prior observations (Lupyan & Winter, 2018; Perry et al., 2015; Winter et al., 2017), we also examined the relationship between iconicity and several psycholinguistic variables: AoA, SERs, concreteness, word length, and word frequency. Furthermore, we examined the relationship between iconicity and lexical class in light of prior findings pointing to differences in word ratings across word classes (Perry et al., 2015; Winter et al., 2017). Finally, we investigated whether the sensory modality of the stimulus presentation affected the participants' iconicity scores. To this end, we selected 360 words from the main study with different degrees of iconicity. These words were then presented auditorily to determine any possible differences that might arise between reading aloud and the auditory presentation of stimuli.

Materials and methods

Main study

Participants

The study involved 1350 native speakers of Spanish, all of whom were students at Universitat Rovira i Virgili (Tarragona, Spain), Universidad Complutense de Madrid (Madrid, Spain) or Universidad Nebrija (Madrid, Spain). Ninety-six participants were removed from the analyses because of atypical responses (the data cleaning procedure is described in the *Materials and procedure* section). The remaining 1254 participants had an average age of 25.25 ($SD = 7.97$, range = 17–56), 944 were women (75.28% of the sample), and 310 were men (24.72% of the sample). Participants received academic credits for their participation.

Materials and procedure

We selected a total of 10,995 words belonging to different grammatical categories from prior normative studies in Spanish (Ferré, Guasch, Martínez-García, Fraga, & Hinojosa, 2017; Stadthagen-González, Ferré, Pérez-Sánchez, Imbault, & Hinojosa, 2018): 6310 nouns, 2517 adjectives, 1625 verbs, 99 adverbs, 8 pronouns, 3 conjunctions, and 2 prepositions. There were also 214 words that belonged to two different grammatical categories (mainly words that could be both nouns and adjectives). Additionally, we included 87 onomatopoeias and 120 interjections, since these word categories are typically the most iconic (Winter et al., 2017). We selected words from a wide range of lexical frequencies and with different lengths, in order to achieve a word pool as representative as possible of the Spanish language. These words were visually presented and randomly distributed in 55 iconicity questionnaires, which were created and administered online using the TestMaker software (Haro, 2012). On average, each questionnaire included 200 words on 10 pages. Participants were asked to evaluate the iconicity of each printed word using a scale from 1 to 7, with 1 meaning very arbitrary (i.e. the sound of the word is not related at all to its meaning) and 7 meaning very iconic (i.e. the sound of the word is closely related to its meaning). Participants were asked to read each word aloud before assessing its iconicity. This was done to ensure that they took the phonology of the words into account in their judgments. The complete instructions, adapted from other studies on iconicity conducted in English (e.g., Perry et al., 2015; Winter et al., 2017), are provided in the appendix. Of note, prior studies (e.g., Perry et al., 2015) have used a scale ranging from −5 to 5, in which participants had to score −5 for "words that sound like the opposite of what they mean" and 0 for "words that do not sound like what they mean or the opposite". Since forms that

mapped onto the opposite of their meaning might be viewed as a special kind of non-arbitrary relationship (see Sidhu & Pexman, 2018, for a similar claim), and no prior theoretical proposals have addressed opposite relationships between word forms and meanings, we instead opted to focus on the arbitrary (score 1) and iconic (score 7) features of words. Nonetheless, we found a significant correlation between the ratings compiled in our study and those from prior studies (see the Results and Discussion section).

The participants' responses were examined to assess the reliability of the data. This process led us to exclude the responses of 96 participants. We removed the data from participants whose ratings showed a low correlation with the average ratings of all the participants who completed the same questionnaire (i.e., $r < .1$). Correlation values close to zero were interpreted as idiosyncratic response patterns, while negative values would indicate that the participant understood the iconicity scale in the reverse order. In addition, we removed the data from 32 participants who completed the same questionnaire twice, as well as the data from three participants who responded to fewer than 50% of the words on the questionnaire. After this process, 22.8 responses were obtained on average per questionnaire (range = 19–25, $SD = 1.89$). Each of the 10,995 words was rated by an average of 22.16 participants (range = 6–25, $SD = 2.52$). It should be noted, however, that participants had the option to indicate that they did not know the word or its meaning, which explains why some words had a low number of ratings. On average, there were 0.64 "don't know" responses for each word (range = 0–18, $SD = 1.73$). The Ns that appear in the descriptive statistics, and those included in the analyses, refer only to valid responses.

We also compiled word ratings for concreteness, AoA, and SERs from different databases for the purpose of exploring the relationship between iconicity and some relevant psycholinguistic variables. Concreteness indicates the degree of specificity of the meaning of the word (Paivio, Yullie, & Madigan, 1968), ranging from 1 (very abstract) to 7 (very concrete). AoA is an estimate of the age at which the speaker thinks that he/she learned the word (Carroll & White, 1973). It is rated on a scale ranging from 1 (before the age of 2) to 11 (at age 11 or later), including continuous values between 1 and 11 to reflect the exact age at which the word was learned (e.g., 4 indicates that the word was learned at the age of 4). SERs refers to the extent to which the word evokes a sensory or perceptual experience (Juhasz, Yap, Dicke, Taylor, & Gullick, 2011), on a scale ranging from 1 (low degree of sensory experience) to 7 (high degree of sensory experience). We obtained concreteness values for 3518 words from the databases of Duchon et al., (2013), Guasch et al., (2016), Haro, Ferré, Boada, and Demestre (2017), Hinojosa et al. (2016a), and Ferré, Guasch, Moldovan, and Sánchez-Casas (2012). AoA values for 2926 words were compiled from the databases of Alonso,

Fernandez, and Díez (2015), Haro et al. (2017), and Hinojosa et al. (2016b). SERs for 2481 words were obtained from the database of Díez-Álamo, Díez, Wojcik, Alonso, and Fernández (2019). We also compiled values of word frequency, number of letters, number of syllables, and grammatical category for 10,762 words from the Spanish lexical database EsPal (Duchon et al., 2013). The grammatical category classification was later manually reviewed by the authors to identify the words that could belong to more than one grammatical category, and to classify those words that were not found in EsPal (a total of 233 words did not appear in EsPal, mainly onomatopoeias, interjections, and compound words).

Auditory study

Participants

Forty-nine native Spanish-speaking students from Universitat Rovira i Virgili (Tarragona, Spain) participated in the auditory study. They had an average age of 21.56 years ($SD = 6.32$, range = 19–58), 33 were women (67.35% of the sample), and 16 were men (35.65% of the sample). No participants were excluded from the analyses. Participants received academic credits for their participation.

Materials and procedure

The auditory study was conducted after collecting the ratings by means of the visual presentation. We selected 360 words from the total set of 10,995 words. In this selection, we aimed to cover the entire range of iconicity values obtained in the visual study, from words which were considered not iconic at all to those considered highly iconic. The 360 words were distributed in two questionnaires of 180 words each. The instructions and scale were the same as those used in the visual mode, with the exception that participants were not asked to repeat the word aloud. The words were presented auditorily, one at a time. We used the Microsoft speech synthesis engine to convert the words to speech. We selected peninsular Spanish as the language for the speech synthesizer and an adult male voice type. Participants had to click on a button to hear each word and then rate its iconicity.

Results and discussion

Availability of the norms

The database can be downloaded as an Excel file from this link: https://osf.io/v5er3/. The file includes the following columns: *word* (Spanish word), *ico-m* (average iconicity of the word), *ico-sd* (standard deviation of the iconicity of the

word), *ico-n* (number of participants who rated the iconicity of the word), *ico-dn* (number of participants who indicated that they did not know the word or its meaning), *audio-m* (average iconicity of the word in the auditory modality), *audio-sd* (standard deviation of the iconicity of the word in the auditory modality), *audio-n* (number of participants who rated the iconicity of the word in the auditory modality), *audio-dn* (number of participants who indicated that they did not know the word or its meaning in the auditory modality), and *gcat* (grammatical category of the word).

Reliability, correlations with other psycholinguistic norms and predictive capacity of iconicity ratings

We calculated the intra-class correlation coefficient (ICC; Koo & Li, 2016) for each iconicity questionnaire to obtain the interrater reliability of the measure. To do this, we used the two-way random effects based on the absolute agreement of multiple raters (2,k). The ICCs were all statistically significant (all $ps < .001$), $M = .99$, $SD = .00$, range = .97–99, which strongly supports the reliability of the data.

Additionally, we compared our iconicity ratings with those of Perry et al. (2015). Although there were 238 Spanish words in common with that study, we selected only 197. We did this because some words in the Perry et al. study had a negative iconicity value because the authors used a scale ranging from −5 to 5. In that scale, negative values indicated that the sound of the word suggested the opposite of its meaning, 0 indicated that there was no relationship between the sound of a word and its meaning, and positive values indicated a congruent relationship between the sound of a word and its meaning. Hence, we excluded from the analyses the words that received a negative iconicity rating in Perry et al. (of note, a similar procedure was adopted by Sidhu & Pexman, 2018). The correlation between the ratings of the two databases was significant albeit low, $r = .29$, $p < .001$. It has been suggested that subjective iconicity arises from participants' own experience with the world and/or language (Occhino, Anible, Wilkinson, & Morfors, 2017). Individual susceptibility to the symbolic connotations of the words (Taylor & Taylor, 1965) and increased consistency of the mapping between word forms and meanings with age (Taylor & Taylor, 1962) also seem to play a role in how iconicity is subjectively perceived. Although the age of the participants was not reported in the Perry et al. study, age and/or individual differences might account for the low correlations between scores from their study and the current one. Methodological differences should be also considered, particularly regarding task instructions. In this sense, it is worth noting that Perry et al. asked participants to rate the stimuli on a scale varying from words that sound the opposite of what they mean to words that sound the same as what they mean. In contrast, in the current study we asked participants to score on a scale from a lack of resemblance to a close

relationship between words' sound and meaning. In sum, although significant, the low correlation between the two normative studies points to the need for additional research using similar methodological settings to test the contribution of individual differences in language-related factors to the perceived relationship between word forms and meanings.

Finally, we examined the predictive power of iconicity ratings in lexical decision response times (RTs). We obtained the lexical decision data from the Spanish megastudy of Aguasvivas et al. (2018) and computed the mean RT in response to each word. We selected only the responses of native Spanish speakers living in Spain and removed the experimental sessions with more than 15% of response errors, wrong responses, RTs below 200 ms and above 2000 ms, and RTs above 1.5 SD and below 1.5 SD from the mean of RTs of each experimental session. The RTs were introduced as a dependent variable in a stepwise multiple regression analysis, where we examined whether iconicity was able to predict RTs after controlling for the effect of different lexical variables. The predictor variables were, in addition to the iconicity ratings, word frequency, number of letters, bigram frequency, number of neighbors, and number of higher-frequency neighbors, all obtained from EsPal (Duchon et al., 2013). The resulting model included 9084 words and was able to significantly predict RTs, $F(4, 9079) = 822.27$, $p < .001$, $R^2 = .27$. The iconicity ratings showed a significant effect, facilitating the RTs, $\beta = -.03$, $p < .001$. This result suggests that iconicity has a facilitating effect on lexical decisions, after controlling for the effects of several other classic lexical variables.

Relationships between iconicity and lexico-semantic variables

The descriptive statistics and distribution of the variables included in this study are shown in Table 1 and Fig. 1, and the bivariate correlations between variables are presented in Table 2.

We conducted a multiple regression analysis to examine the relationship between iconicity and several lexico-semantic variables. Only those words with ratings available for all variables were included in the regression analysis ($n = 1088$). It should be noted that onomatopoeias and interjections were not included in this analysis, because there were no SERs, AoA, and concreteness values for many of them. Iconicity was the main dependent measure, and concreteness, AoA, SERs, word frequency, and number of letters and syllables were predictors. The variables were entered using the stepwise method. The variance inflation factor (VIF) and tolerance values showed that there were no multicollinearity problems (all VIF values were below 1.85, and the tolerance values were between .54 and .84). We also checked that model residuals were normally distributed. The resulting model was able to significantly predict the iconicity ratings, $F(5, 1082) =$

Table 1 Descriptive statistics for the variables examined in the study

	N	Range	Mean	SD	Skewness	Kurtosis
Iconicity (visual)	10,995	1.00–6.91	2.97	0.83	1.05	2.38
Iconicity (auditory)	360	1.44–6.91	2.78	0.98	1.82	4.01
Concreteness	3518	2.14–6.85	4.47	0.93	0.22	−0.66
AoA	2926	1.12–10.96	7.45	2.01	−0.39	−0.77
SERs	2481	1.47–6.31	3.42	0.79	0.49	−0.04
Word frequency	10,762	0.00–4.52	0.51	0.54	1.89	4.51
Number of letters	10,762	2.00–20	8.13	2.54	0.40	0.02
Number of syllables	10,762	1.00–8.00	3.37	1.07	0.37	0.01

15.56, $p < .001$, $R^2 = .07$. The model included SERs, concreteness, AoA, word frequency, and number of syllables (see Table 3 and Fig. 2). The variable number of letters did not reach statistical significance ($p = .67$), and thus was excluded from the model. SERs, which showed a positive relationship with iconicity, was the variable that explained the largest portion of variance in the model ($R^2 = .04$; i.e., 4%). The other variables—concreteness, age of acquisition, word frequency, and number of syllables—showed a negative relationship with iconicity.

The finding of SERs as the main predictor of iconicity suggests that the most iconic Spanish words are also those that contain richer sensory information. The contribution of SERs to iconicity has also been demonstrated for English words (Sidhu & Pexman, 2018; Winter et al., 2017) and highlights the contribution of information from multiple sensory modalities to iconicity. In contrast, we observed a negative relationship between iconicity and concreteness. Since SERs and concreteness are positively correlated but showed opposite effects as predictors of RTs, we performed an additional regression analysis with both variables as predictors and with iconicity as a criterion in order to rule out suppression effects in the main analysis. Specifically, we estimated whether the predictive capacity of SERs would increase when concreteness was introduced in the analysis. The

Table 2 Bivariate correlations between the variables examined in the study

	1	2	3	4	5	6	7
1. Iconicity (visual)	1	.02	−.12**	.24**	.03**	−.08**	−.09**
2. Concreteness		1	−.34**	.21**	.06**	−.27**	−.19**
3. AoA			1	−.34**	−.52**	.40**	.38**
4. SERs				1	.26**	−.08**	−.08**
5. Word frequency					1	−.33**	−.33**
6. Number of letters						1	.88**
7. Number of syllables							1

Note. Double asterisks (**) indicate that the correlation is significant at the .01 level ($p < .01$)

results showed that the beta coefficient of SERs increased from 0.207 to 0.215. Also, the zero-order correlation between concreteness and iconicity was .005 and the part correlation (after including SERs) was −.038. Therefore, since adding concreteness did not substantially increase the beta coefficient of SERs (a .008-point increment), and the difference between the part correlation and the zero-order correlation between concreteness and iconicity was minimal (both coefficients were very close to 0), the results of these analyses suggest that there were no suppression effects in the main regression analyses.

The finding of higher iconicity ratings for abstract words might explain the increased activation of iconicity-related information for abstract words during a semantic decisions task reported in prior studies (Lupyan & Winter, 2018). Also, although the contribution of concreteness to iconicity has not been examined before in normative studies, Winter et al. (2017) reported a negative relationship between iconicity and imageability. Considering that imageability is a variable that is highly correlated with concreteness, our results are in line with those of Winter et al. (2017), suggesting that words with fewer visual properties are more iconic. While this claim might be a priori at odds with the positive relationship between iconicity and SERs found here and in previous works, it should be kept in mind that iconicity effects are differently modulated by specific sensory modalities. For example, Winter et al. (2017) found that within the set of highly sensory words, those denoting visual meanings were the least iconic, while those denoting auditory and tactile meanings were the most iconic. Furthermore, some data suggest that sound-symbolic associations are grounded in auditory-visual feature integration (Kovic, Plunkett, & Wetermann, 2010).

In keeping with the suggestion that iconicity may facilitate word learning during childhood (e.g., Perniss & Vigliocco, 2014, or the "sound symbolism bootstrapping hypothesis," Imai & Kita, 2014; see Nielsen & Dingemanse, 2020, for a critical review of evidence for learning enhancement of iconic words), the results of our regression analyses also showed that AoA is linked to word iconicity. As in previous studies conducted in Spanish and English (Perry et al., 2015) as well as in
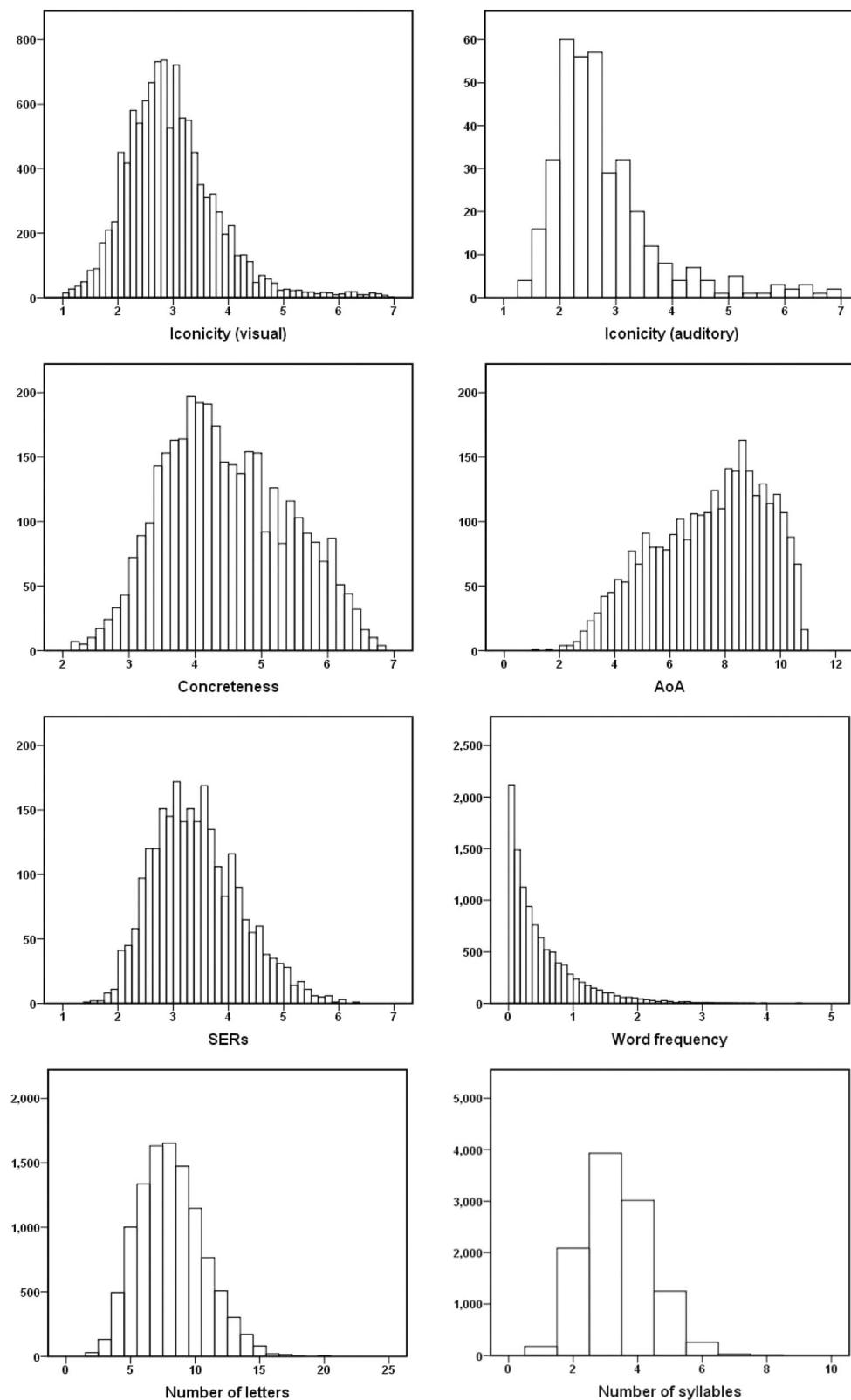
Fig. 1 Distribution of the variables examined in the study

British Sign Language (Thompson, Vinson, Woll, & Vigliocco, 2012), we found that words rated most iconic were learned first. Of note, the predictive capacity of AoA cannot be explained by the presence of onomatopoeias and interjections in the data set, because these word types were not included in the regression analysis.

Table 3   Coefficients of the multiple linear regression model

	b	95% CI	β	SE	t	p	Tolerance	VIF
Constant	3.17	(2.67; 3.65)	-	0.24	13.16	<.001	-	-
Sensory ratings	0.19	(0.14; 0.25)	0.21	0.03	6.53	<.001	0.84	1.19
Number of syllables	−0.10	(−0.15; −0.05)	−0.12	0.03	−3.62	<.001	0.82	1.22
Concreteness	−0.08	(−0.13; −0.03)	−0.10	0.02	−3.13	.002	0.80	1.25
Log frequency	−0.13	(−0.21; −0.04)	−0.10	0.04	−2.87	.004	0.65	1.53
AoA	−0.03	(−0.06; 0.00)	−0.09	0.01	−2.26	.024	0.54	1.84

Prior research has indicated that infants are sensitive to sound–meaning correspondences by four months of age (Ozturk, Krehm, & Vouloumanos, 2013). Infant vocabularies also tend to include a high proportion of onomatopoeias, which have inherent iconic properties (Laing, 2019). Additional evidence comes from the results of a recent study that examined the relationship between iconicity and child and adult word frequency measures (Perry, Perlman, Winter,
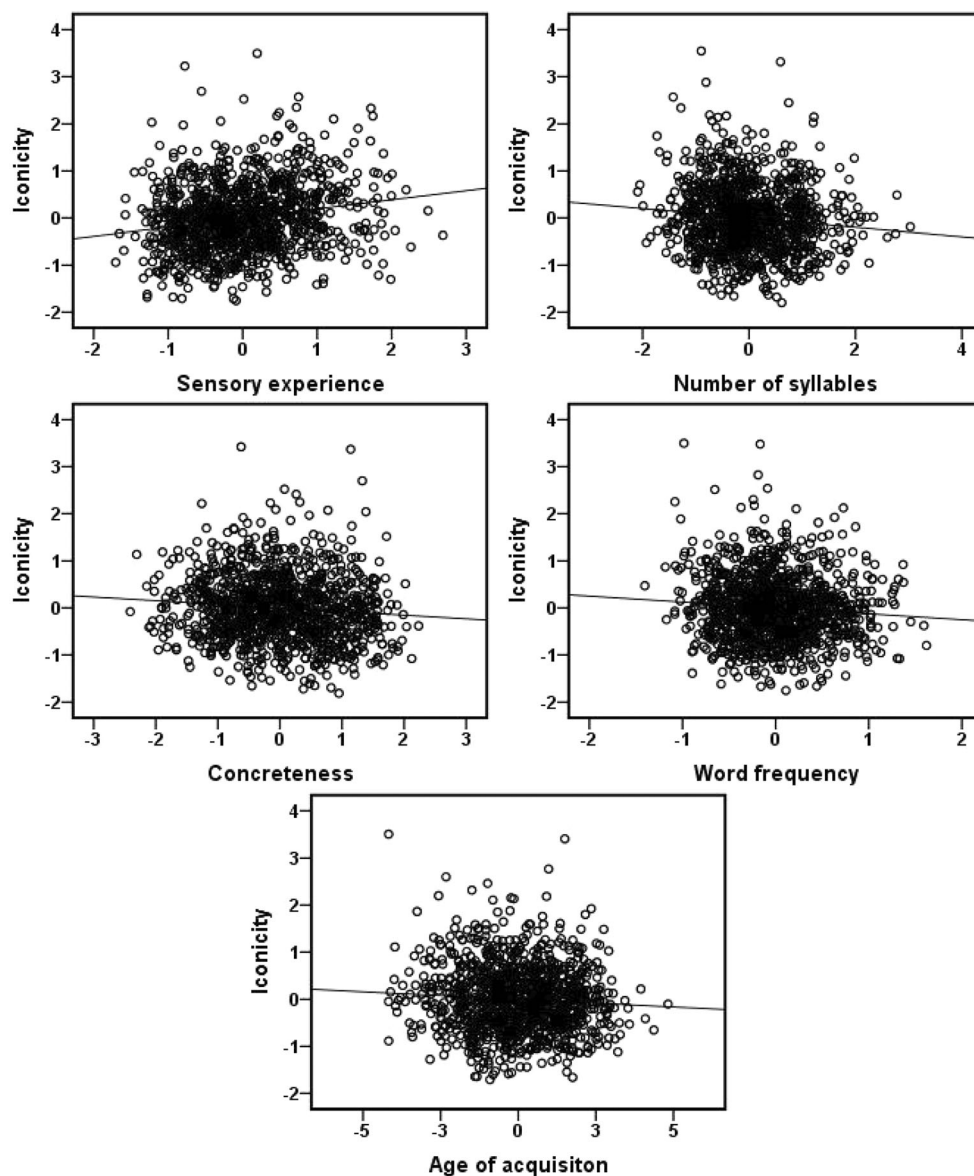

Fig. 2 Relationship between iconicity and each variable included in the multiple linear regression, when controlling for the other variables. Each dot represents a word, and the solid line shows the linear fit

Massaro, & Lupyan, 2018). The authors found a positive relationship between iconicity and frequency in children, suggesting that not only does iconicity help children learn new words, but also that, once learned, iconic words are more frequently used than non-iconic words. Importantly, an interaction with age emerged, such that the positive relationship between iconicity and frequency disappears as children get older, and the direction of this relationship is even reversed in older children and adults. This is in line with the finding of a negative relationship between frequency and iconicity in the current study, suggesting that adult speakers use low iconicity words more often than high iconicity words. Interestingly, Perry et al. (2018) found that this pattern was reversed in child-directed speech (i.e., when adults speak to young children, they use highly iconic words more frequently than words with a low degree of iconicity).

Iconicity and grammatical category

We conducted an analysis of variance (ANOVA) to examine whether iconicity ratings vary by grammatical category. Grammatical categories that were underrepresented in the data set were not included in the analysis, namely, prepositions ($N = 2$), conjunctions ($N = 3$), and pronouns ($N = 8$). Words that could belong to two different grammatical categories (e.g., noun and adjective) were also excluded from the analysis ($N = 214$). A total of 10,768 words were included in the ANOVA. Although there were high iconicity words and low iconicity words in all of the grammatical categories included in the study (see Table 4), the average iconicity ratings clearly differed between grammatical categories, $F(5, 10,767) = 463.19$, $p < .001$ (see Fig. 3). Onomatopoeias had the highest iconicity values ($N = 97$, $M = 5.6$, $SD = 1.36$, range = 2.05–6.91), followed by interjections ($N = 120$, $M = 5.17$, $SD = 1.07$, range = 1.60–6.82), adjectives ($N = 2517$, $M = 3.02$, $SD = 0.73$, range = 1.07–6.28), verbs ($N = 1625$, $M = 2.91$, $SD = 0.73$, range = 1.24–6.00), adverbs ($N = 99$, $M = 2.91$, $SD = 0.83$, range = 1.38–5.52), and nouns ($N = 6310$, $M = 2.88$, $SD = 0.75$, range = 1.00–6.61). Post-hoc comparisons revealed significant differences in iconicity between onomatopoeias and the other grammatical categories (all ps < .001), between interjections and the other

grammatical categories (all ps < .001), and between adjectives and nouns, verbs, and adverbs (all ps < .001). No differences in iconicity were observed between nouns, verbs, and adverbs (all ps > .05). These results reveal that onomatopoeias and interjections are perceived as the most iconic words in the lexicon. Considering that the number of onomatopoeias and interjections in the Spanish language is not very high (as compared to other types of words, such as nouns or verbs), it might be argued that iconicity is a marginal phenomenon in language. However, iconicity might have a broader influence, affecting word lexicalization, whereby words belonging to different grammatical categories are derived from onomatopoeias through morphology. Interestingly, those words are also considered as highly iconic, as in the case of the verb "*bufar*" (to snort, $M = 5.64$), which derives from "*buf*" ($M = 6.14$), or the noun "*gruñido*" (grunt, $M = 6$), which is related to the onomatopoeia "*grrr*" ($M = 6.61$).

The pattern of findings reported here is consistent with that observed in Perry et al. (2015, Exp. 4) for Spanish words. These authors found higher iconicity ratings for Spanish interjections and onomatopoeias in comparison with adjectives, verbs, nouns, and function words. They also found higher iconicity ratings for adjectives in comparison with verbs, nouns, and function words. While onomatopoeias and interjections are generally acknowledged as being the lexical categories that reflect direct sound-to-meaning mapping, the closer relationship between adjectives and iconicity relative to other word classes deserves further attention. One possibility is that adjectives often contain meanings for properties such as size, shape, repetition, intensity, and temporal unfolding, which have been closely related to iconicity (Dingemanse et al., 2015; Perlman, Little, Thompsom, & Thompsom, 2018). Our results also resemble those found in English, with the exception of verbs, which were found to be more iconic than nouns in that language (Perry et al., 2015; Winter et al., 2017). The lack of iconicity differences between verbs and nouns in Spanish has been attributed to the fact that verbs in Spanish are less expressive of manner of movement compared to those in English (Perry et al., 2015; Winter et al., 2017). In this regard, movement meanings show a close relationship with iconicity. For

Table 4 List of the most and least iconic words across grammatical categories (average iconicity value in parentheses)

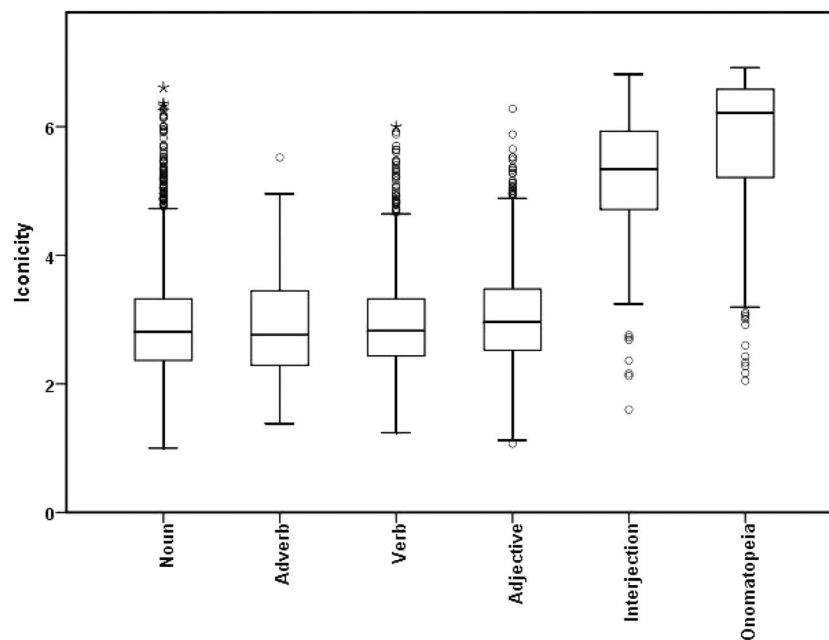|	Most iconic word	Least iconic word
Adjectives	*larguísima* ("very long") (5.88)	*disidente* ("dissident") (1.07)
Adverbs	*Más* ("more") (5.52)	*creces* ("by far") (1.38)
Interjections	*¡uf!* ("phew!") (6.82)	*¡ahó!* ("hey!") (1.60)
Nouns	*zigzag* ("zigzag") (6.61)	*gerencia* ("management") (1.00)
Onomatopoeias	*clic* ("click") (6.91)	*tachín* ("ta-da") (2.05)
Verbs	*aullar* ("to howl") (6.00)	*vender* ("to sell") (1.24)

Fig. 3 Distribution of iconicity values for each grammatical category

instance, according to the hierarchy proposed by Dingemanse (2012) to categorize how certain meanings are encoded in ideophone systems, meanings related to movements have the most common mapping onto sound after sound-to-sound mapping. To explore the contribution of movement meanings to iconicity in verbs, we used the normative data of San Miguel Abella and González-Nosti (2020). In that study, a large set of Spanish verbs were rated on a scale of 1 to 7 based on their motor content (i.e., the amount of mobility that the action described by the verb entails). There were 1430 words in common between the two databases. We split this word set between a high motor content subset and a low motor content subset, taking as a criterion the average point on the scale: verbs with motor ratings above 4 (e.g., *nadar*, to swim) were classified as high motor content verbs and those with motor ratings below 4 (e.g., *creer*, to believe) were classified as low motor content verbs. According to this classification, there were many more low motor content verbs ($n = 1250$) than high motor content verbs ($n = 180$). T tests revealed that high motor content verbs were more iconic than low motor content verbs. This was true both when the analysis included all the words in the subsets (average iconicity ratings: $M = 3.15$ and $M = 2.90$ for high motor content verbs and low motor content verbs, respectively), $t(1428) = 4.42$, $p < .001$, and when a random set of 180 low motor content words was selected (average iconicity rating: $M = 2.80$) for comparison with the 180 high motor content words, $t(358) = 4.26$, $p < .001$. The results of these analyses reveal that although Spanish verbs have low iconicity ratings overall, speakers perceive verbs whose meanings entail greater mobility as more iconic.

Comparison between iconicity ratings in the visual and auditory modalities

Finally, we examined whether the iconicity ratings from the visual and auditory modalities were correlated, focusing on the 360 words for which both types of ratings were available. The correlation between the two modalities was very high, $r = .69$, $p < .001$, and was very similar to that reported by Perry et al. (2015, $r = .61$), who also compared the written and auditory presentation for English words. This high correlation indicates that the iconicity scores are reliable, and suggests that participants relied on phonology to rate the words that were presented visually. It should be noted, however, that the auditory ratings were significantly lower ($M = 2.78$, $SD = 0.98$) than the visual ratings ($M = 3.65$, $SD = 0.83$), $t(359) = 22.84$, $p < .001$. Prior studies on iconicity have already shown that participants showed enhanced performance when judging the equivalence of word pairs from different languages (e.g., English, Chinese, Japanese, and Hebrew) if words were presented in the visual relative to the auditory modality (e.g., Brackbill and Little, 1957; Brown, Black, & Horowitz, 1955). Similarly, Oda (2000) found that English speakers were better at matching unfamiliar highly iconic Japanese words to English definitions when they read the words aloud themselves than when they were read out by a native speaker of Japanese (Oda, 2000). These authors speculated that articulating the words might increase their perceived iconicity. Also, it has been claimed that iconicity strongly relies on the expressive voice quality of speakers in speech (Ertel and Dorst, 1965). Therefore, synthesizing speech as in the current study might have homogenized these expressive cues, leading

to lower iconicity scores in the auditory relative to the visual presentation of words.

Conclusion

In this study, we report subjective iconicity ratings for a large set of Spanish words. Our results indicated that onomatopoeias and interjections were the lexical categories associated with higher iconicity ratings. Remarkably, high iconicity values were also found in other word categories such as verbs, nouns, adverbs and, particularly, adjectives. These findings argue against a language conception that is solely grounded in the arbitrariness between word forms and meanings (De Saussure, 2011). In agreement with this view, we also observed a close relation between iconicity and two lexico-semantic variables—concreteness and age of acquisition. Related to this, the results of our regression analyses showed that the words associated with higher sensory experiences are also more iconic. Interestingly, the negative relationship between concreteness and iconicity suggests that iconicity might play a role in the representation of abstract concepts. Finally, in agreement with prior results from experimental studies, we found that words acquired early in life are rated higher in iconicity. Overall, the data reported in this normative study are consistent with theoretical views assuming that both arbitrariness and iconicity cooperate in shaping language (Dingemanse et al., 2020; Lockwood & Dingemanse, 2015; Perniss & Vigliocco, 2014). The norms we provide here might be of use for researchers from different fields, particularly for those interested in psycholinguistics or language learning in educational contexts.

Appendix

Instructions for iconicity ratings (Spanish)

En este cuestionario usted deberá calificar distintas palabras en castellano en función de hasta qué punto el sonido de las palabras se parece a su significado.

Déjenos explicarle algo más sobre esto. Algunas palabras en castellano tienen un sonido que se parece a su significado. Por ejemplo, la palabra RONCAR suena como el ruido que alguien hace al respirar mientras duerme. La onomatopeya ¡PLAS! suena como el ruido que un objeto hace al caer al suelo. Otro ejemplo es la palabra CHICO, que suena como algo pequeño (en comparación con GIGANTE, que suena grande). Estas palabras son icónicas. Podría adivinar qué significan aunque no supiera hablar castellano. Otras palabras no son icónicas, es decir, son arbitrarias. Por ejemplo, no hay nada que suene canino o felino en las palabras PERRO y GATO. Si usted no supiera castellano, no podría adivinar su significado.

Para cada palabra que vea, califique en una escala del 1 al 7 cuán icónica cree que es, tomando 1 como muy arbitraria (es decir, el sonido de la palabra no tiene nada que ver con su significado) y 7 como muy icónica (es decir, el sonido de la palabra está muy relacionado con su significado). Antes de otorgarle una puntuación, repítase la palabra en voz alta, con el objetivo de prestar verdadera atención a cómo suena y compararlo con su significado. Siéntase libre de usar todo el rango de números, del 1 al 7; al mismo tiempo, no se preocupe por la frecuencia con la que usa un número en particular, siempre y cuando sea honesto en sus calificaciones. Si no conoce el significado de una palabra, elija la opción "No sé el significado de esta palabra". Trabaje con bastante rapidez, pero no sea descuidado en sus calificaciones, lo importante es que sea lo más preciso posible. Recuerde que no hay respuestas buenas o malas, ya que le estamos preguntando por su impresión acerca de las palabras.

Instructions for the iconicity questionnaires (English translation)

In this questionnaire you will be asked to rate the extent to which the sound of different Spanish words resembles their meaning.

Let us explain this a little further. Some Spanish words have a sound that resembles their meaning. For example, the word RONCAR ("SNORE") sounds like the noise someone makes when they breathe in their sleep. The onomatopoeia ¡PLAS! sounds like the noise an object makes when it falls to the ground. Another example is the word CHICO ("SMALL"), which sounds like something small (compared to GIGANTE ("GIANT"), which sounds big). These words are iconic. You could guess what they mean even if you didn't speak Spanish. Other words are not iconic, that is, they are arbitrary. For example, there is nothing that sounds canine or feline in the words PERRO ("DOG") or GATO ("CAT"). If you didn't speak Spanish, you wouldn't be able to guess their meaning.

Rate on a scale of 1 to 7 how iconic you think each word is, taking 1 as very arbitrary (i.e., the sound of the word is not related at all to its meaning) and 7 as very iconic (i.e., the sound of the word is very related to its meaning). Before giving a rating, please repeat the word aloud to pay close attention to how it sounds and to compare its sound to its meaning. Feel free to use the entire range of values, from 1

to 7. Don't worry about how often you use a particular value, as long as you are accurate in your ratings. If you don't know the meaning of a word, choose the "I don't know the meaning of this word" option. Try to answer quickly, but don't be careless with your ratings—it is important to be as accurate as possible. Remember that there are no right or wrong answers, because we are asking you about your impression of the words.

References

Adelman, J. S., Estes, Z., & Cossu, M. (2018). Emotional sound symbolism: Languages rapidly signal valence via phonemes. Cognition, 175, 122–130. https://doi.org/10.1016/j.cognition.2018.02.007

Aguasvivas, J. A., Carreiras, M., Brysbaert, M., Mandera, P., Keuleers, E., & Duñabeitia, J. A. (2018). SPALEX: A Spanish lexical decision database from a massive online data collection. *Frontiers in psychology*, 9, 2156. https://doi.org/10.3389/fpsyg.2018.02156

Ahlner, F., & Zlatev, J. (2010). Cross-modal iconicity: A cognitive semiotic approach to sound symbolism. Sign Systems Studies, 38(1/4), 298-348. https://doi.org/10.12697/SSS.2010.38.1-4.11

Alonso, M. A., Fernandez, A., & Díez, E. (2015). Subjective age-of-acquisition norms for 7,039 Spanish words. Behavior Research Methods, 47, 268–274. https://doi.org/10.3758/s13428-014-0454-2

Aryani, A., Hsu, C. T., & Jacobs, A. M. (2019). Affective iconic words benefit from additional sound–meaning integration in the left amygdala. Human Brain Mapping, 40(18), 5289-5300. https://doi.org/10.1002/hbm.24772

Blasi, D. E., Wichmann, S., Hammarström, H., Stadler, P. F. & Christiansen, M. H. (2016). Sound–meaning association biases evidenced across thousands of languages. Proceedings of the National Academy of Sciences 113(39), 10818–10823. https://doi.org/10.1073/pnas.1605782113

Brackbill, Y., & Little, K. B. (1957). Factors determining the guessing of meanings of foreign words. The Journal of Abnormal and Social Psychology, 54(3), 312-318. https://doi.org/10.1037/h0042411

Brown, R. W., Black, A. H., & Horowitz, A. E. (1955). Phonetic symbolism in natural languages. The Journal of Abnormal and Social Psychology, 50(3), 388-393. https://doi.org/10.1037/h0046820

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. Behavior Research Methods, 41(4), 977-990. https://doi.org/10.3758/BRM.41.4.977

Brysbaert, M., Stevens, M., De Deyne, S., Voorspoels, W., & Storms, G. (2014). Norms of age of acquisition and concreteness for 30,000 Dutch words. Acta Psychologica, 150, 80-84. https://doi.org/10.1016/j.actpsy.2014.04.010

Carroll, J. B., & White, M. N. (1973). Word frequency and age of acquisition as determiners of picture naming latency. Quarterly Journal of Experimental Psychology, 12(1), 85-95. https://doi.org/10.1080/14640747308400325

Cimpian, A., & Salomon, E. (2014). The inherence heuristic: An intuitive means of making sense of the world, and a potential precursor to psychological essentialism. Behavioral and Brain Sciences, 37(5), 461-527. https://doi.org/10.1017/S0140525X13002197

Ćoso, B., Guasch, M., Ferré, P., & Hinojosa, J. A. (2019). Affective and concreteness norms for 3,022 Croatian words. Quarterly Journal of Experimental Psychology, 72(9), 2302-2312. https://doi.org/10.1177/1747021819834226

Cuskley, C. (2013). Mappings between linguistic sound and motion. Public Journal of Semiotics, 5(1), 39-62. https://doi.org/10.37693/pjos.2013.5.9651

De Saussure, F. (2011). *Course in general linguistics*. Columbia University Press.

Davis, R. (1961). The fitness of names to drawings. A cross-cultural study in Tanganyika. British Journal of Psychology, 52(3), 259-268. https://doi.org/10.1111/j.2044-8295.1961.tb00788.x

Della Rosa, P. A., Catricalà, E., Vigliocco, G., & Cappa, S. F. (2010). Beyond the abstract—concrete dichotomy: mode of acquisition, concreteness, imageability, familiarity, age of acquisition, context availability, and abstractness norms for a set of 417 Italian words. Behavior Research Methods, 42(4), 1042-1048. https://doi.org/10.3758/BRM.42.4.1042

Díez-Álamo, A. M., Díez, E., Wojcik, D. Z., Alonso, M. A., & Fernandez, A. (2019). Sensory experience ratings for 5,500 Spanish words. Behavior research methods, 51(3), 1205-1215. https://doi.org/10.3758/s13428-018-1057-0

Dingemanse, M. (2012). Advances in the cross-linguistic study of ideophones. Language and Linguistics compass, 6(10), 654-672. https://doi.org/10.1002/lnc3.361

Dingemanse, M. (2018). Redrawing the margins of language: Lessons from research on ideophones. Glossa: A Journal of General Linguistics, 3(1), 4. https://doi.org/10.5334/gjgl.444

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. Trends in cognitive sciences, 19(10), 603-615. https://doi.org/10.1016/j.tics.2015.07.013

Dingemanse, M., Perlman, M., & Perniss, P. (2020). Construals of iconicity: experimental approaches to form–meaning resemblances in language. Language and Cognition, 12(1), 1-14. https://doi.org/10.1017/langcog.2019.48

Dingemanse, M., Schuerman, W., Reinisch, E., Tufvesson, S., & Mitterer, H. (2016). What sound symbolism can and cannot do: Testing the iconicity of ideophones from five languages. Language, 92(2), e117-e133. https://doi.org/10.1353/lan

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. Behavior research methods, 45(4), 1246-1258. https://doi.org/10.3758/s13428-013-0326-1

Elsen, H. (2017). The two meanings of sound symbolism. Open Linguistics, 3(1), 491-499. https://doi.org/10.1515/opli-2017-0024

Ertel, S., & Dorst, R. (1965). Expressive Lautsymbolik. Zeitschrift für experimentelle und angewandte Psychologie, 12(4), 557-569.

Ferré, P., Guasch, M., Martínez-García, N., Fraga, I., & Hinojosa, J. A. (2017). Moved by words: Affective ratings for a set of 2,266 Spanish words in five discrete emotion categories. Behavior Research Methods, 49, 1082-1094. https://doi.org/10.3758/s13428-016-0768-3

Ferré, P., Guasch, M., Moldovan, C., & Sánchez-Casas, R. (2012). Affective norms for 380 Spanish words belonging to three different semantic categories. Behavior Research Methods, 44(2), 395-403. https://doi.org/10.3758/s13428-011-0165-x

Guasch, M., Ferré, P., & Fraga, I. (2016). Spanish norms for affective and lexico-semantic variables for 1,400 words. Behavior Research Methods, 48, 1358–1369. https://doi.org/10.3758/s13428-015-0684-y

Haro, J. (2012). Testmaker: Aplicación para crear cuestionarios online [Computer software and manual]. Retrieved from http://jharo.net/dokuwiki/testmaker

Haro, J., Ferré, P., Boada, R., & Demestre, J. (2017). Semantic ambiguity norms for 530 Spanish words. Applied Psycholinguistics, 38(2), 457-475. https://doi.org/10.1017/S0142716416000266

Hashimoto, T., Usui, N., Taira, M., Nose, I., Haji, T., & Kojima, S. (2006). The neural mechanism associated with the processing of

onomatopoeic sounds. Neuroimage, 31(4), 1762-1770. https://doi.org/10.1016/j.neuroimage.2006.02.019

Hinojosa, J. A., Martínez-García, N., Villalba-García, C., Fernández-Folgueiras, U., Sánchez-Carmona, A., Pozo, M. A., & Montoro, P. R. (2016a). Affective norms of 875 Spanish words for five discrete emotional categories and two emotional dimensions. Behavior Research Methods, 48, 272–284. https://doi.org/10.3758/s13428-015-0572-5

Hinojosa, J. A., Rincón-Pérez, I., Romero-Ferreiro, M. V., Martínez-García, N., Villalba-García, C., Montoro, P. R., & Pozo, M. A. (2016b). The Madrid Affective Database for Spanish (MADS): Ratings of Dominance, Familiarity, Subjective Age of Acquisition and Sensory Experience. PLoS ONE, 11(5): e0155866. https://doi.org/10.1371/journal.pone.0155866

Imai M, Kita S. 2014 The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. Philosophical Transactions of the Royal Society B 369, 20130298. https://doi.org/10.1098/rstb.2013.0298

Juhasz, B. J., Yap, M. J., Dicke, J., Taylor, S. C., & Gullick, M. M. (2011). Tangible words are recognized faster: The grounding of meaning in sensory and perceptual systems. The Quarterly Journal of Experimental Psychology, 64(9), 1683–1691. https://doi.org/10.1080/17470218.2011.605150

Juhasz, B. J., & Yap, M. J. (2013). Sensory experience ratings for over 5, 000 mono-and disyllabic words. Behavior Research Methods, 45(1), 160-168. https://doi.org/10.3758/s13428-012-0242-9

Kanero, J., Imai, M., Okuda, J., Okada, H., & Matsuda, T. (2014). How sound symbolism is processed in the brain: a study on Japanese mimetic words. PLoS one, 9(5), e97905. https://doi.org/10.1371/journal.pone.0097905

Köhler, W. (1929). Gestalt psychology. New York, NY: Liveright.

Köhler, W. (1947). Gestalt psychology: An introduction to new concepts in modern psychology. New York, NY: Liveright.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine, 15(2), 155-163. https://doi.org/10.1016/j.jcm.2016.02.012

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. Behavior Research Methods, 44(4), 978-990.

Laing, C. (2019). A role for onomatopoeia in early language: evidence from phonological development. Language and Cognition, 11(2), 173-187. https://doi.org/10.1017/langcog.2018.23

Lockwood, G. & Dingemanse, M. (2015) Iconicity in the lab: a review of behavioral, developmental, and neuroimaging research into sound-symbolism. Frontiers in Psychology, 6:1246. https://doi.org/10.3389/fpsyg.2015.01246

Lockwood, G., & Tuomainen, J. (2015). Ideophones in Japanese modulate the P2 and late positive complex responses. Frontiers in Psychology, 6:933. https://doi.org/10.3389/fpsyg.2015.0093

Lupyan, G., & Winter, B. (2018). Language is more abstract than you think, or, why aren't languages more iconic?. Philosophical Transactions of the Royal Society B: Biological Sciences, 373(1752), 20170137. https://doi.org/10.1098/rstb.2017.0137

Meteyard, L., Stoppard, E., Snudden, D., Cappa, S. F., & Vigliocco, G. (2015). When semantics aids phonology: A processing advantage for iconic word forms in aphasia. Neuropsychologia, 76, 264-275. https://doi.org/10.1016/j.neuropsychologia.2015.01.042

Miron, M. S. (1961). A crosslinguistic investigation of phonetic symbolism. The Journal of Abnormal and Social Psychology, 62(3), 623-630. https://doi.org/10.1037/h0045212

Monnier, C., & Syssau, A. (2014). Affective norms for French words (FAN). Behavior Research Methods, 46(4), 1128-1137. https://doi.org/10.3758/s13428-013-0431-1

Motamedi, Y., Little, H., Nielsen, A., & Sulik, J. (2019). The iconicity toolbox: empirical approaches to measuring iconicity. Language and Cognition, 11(2), 188-207. https://doi.org/10.1017/langcog.2019.14

Nielsen, A. K., & Dingemanse, M. (2020). Iconicity in word learning and beyond: A critical review. Language and Speech, https://doi.org/10.1177/0023830920914339

Occhino, C., Anible, B., Wilkinson, E., & Morford, J. P. (2017). Iconicity is in the eye of the beholder: How language experience affects perceived iconicity. Gesture, 16(1), 100-126. https://doi.org/10.1075/gest.16.1.04occ

Oda, Hiromi. 2000. An Embodied Semantic Mechanism for Mimetic Words in Japanese. Ph.D dissertation. Bloomington: Indiana University.

Ozturk, O., Krehm, M., & Vouloumanos, A. (2013). Sound symbolism in infancy: Evidence for sound–shape cross-modal correspondences in 4-month-olds. Journal of Experimental Child Psychology, 114(2), 173-186. https://doi.org/10.1016/j.jecp.2012.05.004

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. Journal of Experimental Psychology, 76(1, Pt.2), 1–25. https://doi.org/10.1037/h0025327

Peeters, D. (2016). Processing consequences of onomatopoeic iconicity in spoken language comprehension. In 38th Annual Meeting of the Cognitive Science Society (CogSci 2016) (pp. 1632-1647). Cognitive Science Society

Perlman, M., Little, H., Thompson, B., & Thompson, R. L. (2018). Iconicity in Signed and Spoken Vocabulary: A Comparison Between American Sign Language, British Sign Language, English, and Spanish. Frontiers in psychology, 9, 1433. https://doi.org/10.3389/fpsyg.2018.01433

Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: from a world of experience to the experience of language. Philosophical Transactions of the Royal Society B: Biological Sciences, 369(1651), 20130300. https://doi.org/10.1098/rstb.2013.0300

Perry, L. K., Perlman, M., & Lupyan, G. (2015). Iconicity in English and Spanish and its relation to lexical category and age of acquisition. PloS one, 10(9); e0137147 https://doi.org/10.1371/journal.pone.0137147

Perry, L. K., Perlman, M., Winter, B., Massaro, D. W., & Lupyan, G. (2018). Iconicity in the speech of children and adults. Developmental Science, 21(3), e12572. https://doi.org/10.1111/desc.12572

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia – A window into perception, thought and language. Journal of Consciousness Studies, 8(12), 3–34.

San Miguel Abella, R.A., & González-Nosti, M. (2020). Motor content norms for 4,565 verbs in Spanish. Behavior Research Methods, 52, 447–454. https://doi.org/10.3758/s13428-019-01241-1

Sapir, E. (1929). A study in phonetic symbolism. Journal of Experimental Psychology, 12(3), 225–239. https://doi.org/10.1037/h0070931

Schmidtke, D., & Conrad, M. (2018). Effects of affective phonological iconicity in online language processing: Evidence from a letter search task. Journal of Experimental Psychology: General, 147(10), 1544–1552. https://doi.org/10.1037/xge0000499

Sidhu, D. M., & Pexman, P. M. (2018). Lonely sensational icons: semantic neighbourhood density, sensory experience and iconicity. Language, Cognition and Neuroscience, 33(1), 25-31. https://doi.org/10.1080/23273798.2017.1358379

Sidhu, D. M., Vigliocco, G., & Pexman, P. M. (2020). Effects of iconicity in lexical decision. Language and Cognition, 12(1), 164-181. https://doi.org/10.1017/langcog.2019.36

Soares, A. P., Costa, A. S., Machado, J., Comesaña, M., & Oliveira, H. M. (2017). The Minho Word Pool: Norms for imageability, concreteness, and subjective frequency for 3,800 Portuguese words. Behavior Research Methods, 49(3), 1065-1081. https://doi.org/10.3758/s13428-016-0767-4

Stadthagen-González, H., & Davis, C. J. (2006). The Bristol norms for age of acquisition, imageability, and familiarity. Behavior Research Methods, 38(4), 598-605. https://doi.org/10.3758/BF03193891

Stadthagen-González, H., Ferré, P., Pérez-Sánchez, M. A., Imbault, C., & Hinojosa, J. A. (2018). Norms for 10,491 Spanish words for five discrete emotions: Happiness, disgust, anger, fear, and sadness. Behavior research methods, 50(5), 1943–1952. https://doi.org/10.3758/s13428-017-0962-ydoi.org/10.3758/s13428-017-0962-y

Sutherland, S. L., & Cimpian, A. (2015). An explanatory heuristic gives rise to the belief that words are well suited for their referents. Cognition, 143, 228-240. https://doi.org/10.1016/j.cognition.2015.07.002

Tanz, C. (1971). Sound symbolism in words relating to proximity and distance. Language and Speech, 14(3), 266-276. https://doi.org/10.1177/002383097101400307

Thompson, A. L., & Do, Y. (2019). Defining iconicity: An articulation-based methodology for explaining the phonological structure of ideophones. Glossa: A Journal of General Linguistics, 4(1), 72. https://doi.org/10.5334/gjgl.872

Taylor, I. K., & Taylor, M. M. (1962). Phonetic symbolism in four unrelated languages. Canadian Journal of Psychology, 16(4), 344-356. https://doi.org/10.1037/h0083261

Taylor, I. K., & Taylor, M. M. (1965). Another look at phonetic symbolism. Psychological bulletin, 64(6), 413-427. https://doi.org/10.1037/h0022737

Thompson, P. D., & Estes, Z. (2011). Sound symbolic naming of novel objects is a graded function. Quarterly Journal of Experimental Psychology, 64(12), 2392-2404. https://doi.org/10.1080/17470218.2011.605898

Thompson, R. L., Vinson, D. P., Woll, B., & Vigliocco, G. (2012). The road to language learning is iconic: Evidence from British Sign Language. Psychological science, 23(12), 1443-1448. https://doi.org/10.1177/0956797612459763

Vigliocco, G., & Kita, S. (2006). Language-specific properties of the lexicon: Implications for learning and processing. Language and Cognitive Processes, 21(7-8), 790–816. https://doi.org/10.1080/016909600824070

Vigliocco, G., Zhang, Y., Del Maschio, N., Todd, R., & Tuomainen, J. (2020). Electrophysiological signatures of English onomatopoeia. Language and Cognition, 12(1), 15-35. https://doi.org/10.1017/langcog.2019.38

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. Behavior research methods, 45(4), 1191-1207. https://doi.org/10.3758/s13428-012-0314-x

Winter, B., Perlman, M., Perry, L. K., & Lupyan, G. (2017). Which words are most iconic? Iconicity in English sensory words. Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems, 18(3), 443–464. https://doi.org/10.1075/is.18.3.07win

Xiao, W., & Treiman, R. (2012). Iconicity of simple Chinese characters. Behavior Research Methods, 44(4), 954-960. https://doi.org/10.3758/s13428-012-0191-3

Open practice statement

The data and materials for all experiments are available in the following link: https://osf.io/v5er3/