

Spanish norms for affective and lexico-semantic variables for 1,400 words

Marc Guasch¹ · Pilar Ferré¹ · Isabel Fraga²

© Psychonomic Society, Inc. 2015

Abstract Studies of semantic variables (e.g., concreteness) and affective variables (i.e., valence and arousal) have traditionally tended to run in different directions. However, in recent years there has been growing interest in studying the relationship, as well as the potential overlaps, between the two. This article describes a database that provides subjective ratings for 1,400 Spanish words for valence, arousal, concreteness, imageability, context availability, and familiarity. Data were collected online through a process involving 826 university students. The results showed a high interrater reliability for all of the variables examined, as well as high correlations between our affective and semantic values and norms currently available in other Spanish databases. Regarding the affective variables, the typical quadratic correlation between valence and arousal ratings was obtained. Likewise, significant correlations were found between the lexico-semantic variables. Importantly, we obtained moderate negative correlations between emotionality and both concreteness and imageability. This is in line with the claim that abstract words have more affective associations than concrete ones (Kousta,

Vigliocco, Vinson, Andrews, & Del Campo, 2011). The present Spanish database is suitable for experimental research into the effects of both affective properties and lexico-semantic variables on word processing and memory.

Keywords Normative data · Valence · Arousal · Concreteness · Imageability · Context availability · Familiarity

There is a long history in cognitive psychology of work on the processing differences between concrete and abstract words. A concreteness effect (i.e., an advantage for concrete words) has commonly been reported, using a variety of experimental tasks and paradigms, such as word naming (De Groot, 1989), lexical decision (Binder, Westbury, McKiernan, Possing, & Medler, 2005), acquisition of new vocabulary (De Groot & Keijzer, 2000), or free recall (Fliessbach, Weis, Klaver, Elger, & Weber, 2006; Romani, McAlpine, & Martin, 2008). These differences in processing have not only been observed in behavioral data, but also with electrophysiological measures (i.e., event-related potentials, ERPs; see Barber, Otten, Kousta, & Vigliocco, 2013, for an overview). It is worth mentioning, however, that a reversed concreteness effect has also been reported in neuropsychological patients (e.g., Yi, Moore, & Grossman, 2007), as well as in healthy people (e.g., Barber et al., 2013; Kousta, Vigliocco, Vinson, Andrews, & Del Campo, 2011).

Two main theories have been developed in recent decades to account for these processing differences: the dual-coding theory (Paivio, 1986) and the context availability theory (Schwanenflugel & Shoben, 1983). According to the dual-coding theory, concrete words are represented in two different systems: a verbal system, and a nonverbal system based on sensory representations. Conversely, abstract words, because of their lack of sensory referents, would only be represented in

Electronic supplementary material The online version of this article (doi:10.3758/s13428-015-0684-y) contains supplementary material, which is available to authorized users.

✉ Marc Guasch
marc.guasch@urv.cat

¹ Research Center for Behavior Assessment and Department of Psychology, Universitat Rovira i Virgili, Tarragona, Spain

² Cognitive Processes & Behavior Research Group, Department of Social Psychology, Basic Psychology, and Methodology, Universidade de Santiago de Compostela, Santiago de Compostela, Spain

the verbal system. As a consequence, the semantic representations would be richer for concrete words than for abstract words. In contrast, Schwanenflugel and Shoben (1983) proposed that there is a single coding system for both types of words and that the difference between their representations is quantitative, because concrete words would have more and stronger associations to contextual knowledge than abstract words. That is, contextual information would be more easily retrieved for concrete than for abstract words.

In relation to the above proposals, it has been demonstrated that concrete words are assigned higher ratings of imageability and context availability than abstract words (Altarriba, Bauer, & Benvenuto, 1999). Furthermore, there are significant correlations between imageability and concreteness (Paivio, 1971) and between context availability and concreteness (Schwanenflugel, Harnishfeger, & Stowe, 1988). Therefore, the reported experimental concreteness effects might be explained by differences in either imageability or context availability, in that the two variables are often confounded. In order to test the claims of Paivio (1986) and Schwanenflugel and Shoben (1983), it is necessary to tease apart the contributions of imageability and context availability to concreteness effects. To that end, a manipulation of one of these two variables (e.g., imageability) should be conducted in experiments in which the other variable (e.g., context availability) is controlled. Several studies have used a similar approach. For instance, both Schwanenflugel et al. (1988) and Levy-Drori and Henik (2006) demonstrated that the typical concreteness advantage in lexical decision times was present when concrete and abstract words differed in context availability, but that the effect disappeared when concrete and abstract words were matched in this variable, supporting the proposal of Schwanenflugel and Shoben. However, further analyses of Levy-Drori and Henik's data suggested that familiarity might have contributed to their results. In particular, these authors found that familiarity predicted lexical decision times significantly. Importantly, they also found that, among the set of words that differed in context availability, there was a positive correlation between concreteness and familiarity (i.e., concrete words were more familiar than abstract words), whereas among the set of words matched in terms of context availability, that correlation was negative (i.e., concrete words were less familiar than abstract words). Therefore, we cannot discount that familiarity might have been behind the disappearance of the concreteness effect. Likewise, familiarity might have acted as a confounding factor in other studies in the field (e.g., Binder et al., 2005) in which concrete and abstract words have been matched in objective frequency. However, objective frequency and subjective frequency (i.e., familiarity) may refer, at least in part, to different characteristics of the words (Barca, Burani, & Arduino, 2002). The importance of controlling for familiarity in further research on the effects of concreteness, then, is clear.

Another relevant study here is that of Kousta et al. (2011), who found that abstract words were processed faster than concrete words in a lexical decision task (i.e., a reversed concreteness effect) when the words were matched in both context availability and imageability. These findings support neither the dual-coding theory nor the context availability theory. According to Kousta et al., the reversed concreteness effect could be a consequence of abstract words having more affective associations than concrete words. This conclusion was supported by two additional experiments in which the authors demonstrated that this advantage for abstract words disappeared when all of the words in a lexical decision task were neutral in valence. Considering the effects of familiarity and emotionality reported in the studies above, as well as the other variables that can affect word processing and memory (e.g., length, lexical frequency, and number of orthographic/phonological neighbors), it is clear that rigorous control of the experimental materials is necessary before firm conclusions can be reached on the differences between concrete and abstract words. Such control is only possible when large sets of words are available with known values for all of these variables. The aim of the present work, then, has been to provide subjective ratings for a large set of Spanish words for variables related to concreteness (i.e., concreteness, imageability, and context availability), as well as for familiarity and affective variables.

Closely related to the discussion above, there is an extensive field of research on emotional word processing in which concreteness has not typically been taken into account. Indeed, over the last decade many studies have been conducted in which the affective properties of words have been manipulated. The two most frequently examined variables have been valence and arousal, which are considered the two dimensions that define the structure of emotion from a dimensional perspective (e.g., Lang, 1995; Russell, 2003). *Valence* describes the extent to which an emotion is pleasant or unpleasant, whereas *arousal* refers to its degree of activation. It has been repeatedly demonstrated that emotional content affects word processing in a variety of experimental tasks and paradigms, such as lexical decision (e.g., Vinson, Ponari, & Vigliocco, 2014), naming (e.g., Kuperman, Estes, Brysbaert, & Warriner, 2014), emotional Stroop tasks (e.g., Eilola, Havelka, & Sharma, 2007), valence judgments (e.g., Estes & Verges, 2008), short-term memory tasks (e.g., Majerus & D'Argembeau, 2011), and long-term memory tasks (e.g., Ferré, Sánchez-Casas, & Fraga, 2013). Emotionality also has a neural signature, as revealed by ERP data (see Citron, 2012, for an overview). However, despite widespread evidence suggesting an advantage in processing and memory for emotional words, discrepant findings also exist, revealing that other variables might modulate the effects of emotionality. Some of these variables have been identified, such as valence (i.e., whether words are positive or negative; Unkelbach, Fiedler,

Bayer, Stegmüller, & Danner, 2008), the degree of semantic relatedness between words (e.g., Talmi & Moscovitch, 2004), and the type of encoding during memory tasks (e.g., Ferré, Fraga, Comesaña, & Sánchez-Casas, 2015). With regard to concreteness, only recently have studies emerged that sought to test Kousta et al.'s (2011) claim that affective information is more relevant in the representation and processing of abstract than of concrete words. The common approach of these studies has been to orthogonally manipulate concreteness and emotional content, with discrepant findings: Whereas some authors have found an interaction between concreteness and emotional content, revealing a stronger emotionality effect for abstract than for concrete words (Ferré, Ventura, Comesaña, & Fraga, 2015; Kaltwasser, Ries, Sommer, Knight, & Willems, 2013; Yao & Wang, 2013, 2014), others have not (Kanske & Kotz, 2007).

To make the aforementioned theories testable, it is necessary to have large sets of stimuli with values for emotional as well as semantic variables, such as those related to concreteness. Information about these variables can easily be found in English, a language for which large databases currently exist. For instance, the database of Brysbaert, Warriner, and Kuperman (2014) provides concreteness ratings for 40,000 English word lemmas, whereas that of Warriner, Kuperman, and Brysbaert (2013) includes affective ratings for 13,915 English lemmas. In contrast, no such large corpus exists for Spanish. Concerning emotionality, to our knowledge there are only four affective databases from which researchers interested in emotional word processing can select their stimuli: the Spanish adaptation of ANEW (Redondo, Fraga, Padrón, & Comesaña, 2007); the affective norms of Ferré, Guasch, Moldovan, and Sánchez-Casas (2012); the affective norms of Redondo, Fraga, Comesaña, and Perea (2005); and the affective norms of Hinojosa et al. (2015). The four databases together contain affective ratings for 2,525 different words. Regarding concreteness and its associated variables, both Ferré et al. (2012) and Hinojosa et al. (2015) collected concreteness ratings for all of their stimuli. In addition, the database of Ferré et al. (2012) contains values for familiarity. However, the Spanish adaptation of ANEW (Redondo et al., 2007) provides concreteness, imageability, and familiarity ratings for only 612 of its 1,034 words. There are other sources of concreteness values for Spanish words, such as the studies of Vega and Fernández (2011), with 730 words, and EsPal (Duchon, Perea, Sebastián-Gallés, Martí, & Carreiras, 2013), with ratings for concreteness, imageability, and familiarity for 6,500 words. Familiarity was also collected for 820 Spanish words in another normative study (Moreno-Martínez, Montoro, & Rodríguez-Rojo, 2014). To the best of our knowledge, no database in Spanish reports ratings for context availability. Thus, a limitation of the existing databases is that those focused on lexico-semantic variables either do not provide affective ratings for their stimuli (Moreno-Martínez et al.,

2014) or directly incorporate those contained in ANEW (EsPal; Duchon et al., 2013). Conversely, most affective databases (e.g., ANEW; Redondo et al., 2007) do not include lexico-semantic data for all words. One consequence of this gap is that researchers interested in the study of both types of variables and their possible interactions do not easily find affective and lexico-semantic ratings for the same words across databases.

The aim of the present work was to obtain a large corpus of Spanish words suitable for experiments investigating the effects of affective as well as semantic properties on word processing and memory. We provide subjective ratings for 1,400 words for valence, arousal, concreteness, imageability, context availability, and familiarity. The main value of such a database is that it will enable the manipulation and/or control of relevant variables, and thus help shed further light on processing of the semantic and affective properties of words. Additionally, it will make possible the study of the relationship between lexico-semantic features and affective variables, as some authors have recently explored in other languages, such as Citron, Weekes, and Ferstl (2014) for French.

Method

Participants

A total of 826 participants (678 women, 148 men) with a mean age of 21.54 years ($SD = 4.56$) took part in the study. They were undergraduate students at the Universitat Rovira i Virgili ($N = 567$) in Tarragona (northeastern Spain) and the Universidade de Santiago de Compostela ($N = 259$; northwestern Spain). All were highly fluent speakers of Spanish and had normal or corrected-to-normal vision. They participated in the study either voluntarily or in exchange for extra course credits. Participants were recruited mainly from the degree programs in psychology, but several other degrees were also involved (nursing, education, audiovisual communication, pedagogy, public relations, and journalism). Word ratings were collected in three different time windows, pertaining to different academic semesters: October 2012–December 2012, November 2013–January 2014, and April 2014.

Materials

The criterion to select the sample of 1,400 words used in this study was that they ranged across the concreteness dimension according to the intuition of the authors. Additionally, in order to optimize the use of this database for different types of psycholinguistic studies, we did not discard words either because of their lexical frequency (i.e., we did not exclude low-frequency words) or because they belonged to a particular part

of speech (i.e., we did not restrict our stimuli to nouns). According to the NIM search engine (Guasch, Boada, Ferré, & Sánchez-Casas, 2013), the selected words have frequencies per million ranging from 0.18 to 414.44 ($M = 24.57$, $SD = 38.41$) and a length range of 3–13 letters ($M = 7.41$, $SD = 2.06$). Concerning parts of speech, our set contains 94.29 % nouns, 24.43 % verbs (including both infinitive and conjugated forms), 17.00 % adjectives, and 0.36 % of other categories (i.e., adverbs and interjections). Note that the total sum exceeds 100 % because 32.79 % of the words belong simultaneously to two or more parts of speech.

Procedure

Each of the 1,400 words was rated according to six variables (i.e., valence, arousal, concreteness, imageability, context availability, and familiarity). We constructed 16 questionnaires for each variable (a total of 96), with a set of words assigned randomly to each one. Each version of the questionnaire contained between 75 and 100 words and was completed by at least 20 participants (range: 20–28, $M = 21.77$, $SD = 1.32$). Each participant, depending on the course credit involved in their case, completed between one and four questionnaires assigned to him or her randomly, with no participant completing the same questionnaire more than once. The questionnaire took about 10 min, and participants were instructed to complete the different versions assigned to them in separate sessions.

Words were presented and data were collected online using TestMaker (Haro, 2012), a free PHP-based application to generate questionnaires for online use. At the beginning of each questionnaire, participants were presented with a short explanation page, followed by a section in which personal information about age, sex, and course was collected. Then they received detailed instructions about the particular variable that they were going to rate subjectively. The words on each questionnaire were divided into five pages, with a maximum of 20 words per page. Words were presented in a column at the left side of the screen, with a Likert scale at the right. Labels on the top of the scale were displayed on each page as a reminder of the meanings of the values.

The Likert scales used to assess concreteness, imageability, context availability, and familiarity included seven points, as with the scales that had been used to develop EsPal (Duchon et al., 2013) and other Spanish databases (Hinojosa et al., 2015; Moreno-Martínez, Montoro, & Rodríguez-Rojo, 2014). In this way, our values can be directly compared to those reported in these databases. This strategy has the advantage of allowing researchers to select experimental stimuli from different sources, in the knowledge that a particular value (e.g., a concreteness value) can be interpreted in the same way regardless of the database from which it is taken. This was also the reason for using a 9-point scale in the assessment of the

affective variables (i.e., valence and arousal). In the normative studies investigating the affective properties of words, the most-used scale has been the Self-Assessment Manikin (SAM; Bradley & Lang, 1994), which is composed of 9 points accompanied by characters depicting the different anchor points. For the sake of comparability, and following the common procedure in the field (e.g., Ferré et al., 2012; Montefinese, Ambrosini, Fairfield, & Mammarella, 2014; Redondo et al., 2007; Soares, Comesaña, Pinheiro, Simões, & Frade, 2012; Söderholm, Häyry, Laine, & Karrasch, 2013), we adopted SAM as the rating scale, too. Participants rated the valences of words on a 9-point Likert scale ranging from *completely sad*(1) to *completely happy*(9). Arousal was rated on a scale that ranged from *completely calm*(1) to *completely energized*(9).

Concreteness was defined as the degree of specificity of a word's content, following the procedure used by Duchon et al. (2013). Participants were asked to rate the words on a 7-point Likert scale ranging from *very abstract*(1) to *very concrete*(7) words. We provided examples for the two anchor points, where *objeto* ("object") served as a sample of the abstract words, and *percha* ("hanger") as a sample of the concrete words.

Regarding imageability and context availability, we used instructions similar to those employed by Altarriba et al. (1999). Thus, imageability was defined as the ease or difficulty of deriving a mental image from the content of the word. For instance, whereas it is easy to form an image of *bandera* ("flag") in our minds, it is more difficult to form an image from the word *caridad* ("charity"). These two examples were provided to the participants, who were asked to rate the words on a 7-point Likert scale ranging from 1 (a minimum level of imageability) to 7 (a maximum level of imageability). With respect to context availability, this was defined as the ease or difficulty in associating each word with a context in which the word might appear. As examples, we provided the word *llorar* ("cry") for high context availability, assuming that the notion of a baby crying in the crib would come immediately to our minds. Conversely, the word *herencia* ("inheritance") was provided as a sample of low context availability. Again, we used a 7-point Likert scale, with 1 being the low-availability anchor and 7 the high-availability one.

Finally, the instructions to rate familiarity were also adapted from other studies (Stadthagen-Gonzalez & Davis, 2006). Participants rated words on a 7-point Likert scale, where 1 meant the minimum level and 7 the highest level of familiarity. To perform their ratings, participants were instructed to take into account the extent to which they knew the meaning of a word, as well as the frequency with which they used it. The word *mano* ("hand") was used as an example of a very familiar word, whereas *quark* ("quark") was the example of a very unfamiliar word.

Supplementary material

The resulting database is available as [supplementary material](#) to this article. It is provided in an Excel spreadsheet with the 1,400 words sorted by alphabetical order in Spanish, along with their English translations. Each word is accompanied by its part of speech according to the Spanish-language dictionary of the Real Academia Española (Spanish Royal Academy) from 2001 (22nd edition): Nouns are coded as “N,” adjectives as “A,” verbs as “V,” interjections as “I,” and adverbs as “B.” Regarding the normative data, the six variables are presented in the following order: valence (VAL), arousal (ARO), concreteness (CON), imageability (IMA), context availability (AVA), and familiarity (FAM). Three columns are provided for each variable: the mean value for the word (M), the standard deviation (SD), and the sample size (N).

Results and discussion

Data trimming

Before extracting the final ratings, the data from all of the filled-out questionnaires were examined for aberrant or random response patterns. We visually inspected a scatterplot of each participant's response against the average response value of all participants, excluding those responses with a pattern with almost no variation (e.g., cases in which a participant assigned the same value to almost all words). We also computed a personal correlation coefficient between each participant's data and the mean. Questionnaires with values near zero (i.e., suggesting random answers) or below zero (i.e., suggesting that the participants had understood the scale in the opposite direction from the one intended) were discarded. This led to the exclusion of 46 questionnaires (2.12 % of the total); these questionnaires were replaced when their sample size did not reach the minimum of 20 respondents.

Accuracy, reliability, and validity of the measures

We computed the mean of the standard errors of each word for each of the six variables, in order to establish the accuracy of the measures in relation to their respective sample sizes. At a confidence level of 90 %, the error margin for each variable was ± 0.46 for valence, ± 0.73 for arousal, ± 0.56 for concreteness, ± 0.54 for imageability, ± 0.67 for context availability, and ± 0.54 for familiarity. Thus, these values can be taken into consideration when deciding the cut points for creating experimental subgroups of words (e.g., in the selection of subgroups of negative, neutral, and positive words based on valence ratings).

Furthermore, to assess the interrater reliability of the measures, we calculated the intraclass correlation coefficients

(ICCs) for all of the questionnaires. Since there were 16 different questionnaires for each variable, we present here the data from their mean ICCs by variables (see Table 1).

Overall, the interrater reliabilities were high for all of the variables examined. Focusing on particular variables, valence and imageability were the most consistently rated, showing both the highest ICCs (.97 and .95, respectively) and the lowest percentages of variation among questionnaires (0.83 % and 2.06 %, respectively). In contrast, context availability and familiarity were the variables with least agreement between the raters (ICCs of .85 for both variables) and across questionnaires (6.50 % and 5.98 % variation, respectively). Regarding the two affective variables, valence had a higher interrater reliability than arousal, which is a common pattern in affective databases: There is greater consensus in valence than in arousal ratings (e.g., Eilola & Havelka, 2010; Monnier & Syssau, 2014; Redondo et al., 2007; Soares et al., 2012).

Apart from accuracy and reliability, it was appropriate to assess the validity of the measures. A common approach here is to compare these values, when possible, with those obtained from other sources. It should be noted that there are no normative data for context availability in Spanish. Thus, all ratings of context availability in the present database are novel, and we were not able to compare our ratings with other sources. In contrast, many words in our database had already been rated in previous studies on one or more of the remaining five variables. The numbers of words not previously rated in Spanish were 781 for valence, 781 for arousal, 265 for concreteness, 313 for imageability, and 256 for familiarity. We note that although ratings were already available for a large number of words in some variables, an advantage of the present work is that it provides values obtained in a homogeneous way for the six variables, and in the same database. As we observed above, this redundancy allowed us to assess the validity of our ratings by comparing them with those of the normative studies with overlapping words. For the affective ratings (i.e., valence and arousal), there was a high correlation between our values and those of the Spanish adaptation of ANEW (Redondo et al., 2007), both for valence, $r(324) = .97$, $p < .001$, and arousal, $r(324) = .84$, $p < .001$. Similarly,

Table 1 Mean (M), standard deviation (SD), coefficient of variation (CV), and range for the intraclass correlation values of the questionnaires for each variable

	M	SD	CV	Range
Valence	.97	.01	0.83 %	.96–.99
Arousal	.88	.04	4.65 %	.78–.93
Concreteness	.93	.04	4.59 %	.82–.97
Imageability	.95	.02	2.06 %	.92–.98
Context availability	.85	.06	6.50 %	.72–.95
Familiarity	.85	.05	5.98 %	.72–.93

correlations were high with the ratings of Redondo et al. (2005) for both valence, $r(183) = .95, p < .001$, and arousal, $r(183) = .88, p < .001$, as well as with the data of Ferré et al. (2012) [valence, $r(66) = .90, p < .001$; arousal, $r(66) = .83, p < .001$] and Hinojosa et al. (2015) [valence, $r(132) = .97, p < .001$; arousal, $r(132) = .78, p < .001$]. Concerning concreteness, imageability, and familiarity, we used EsPal (Duchon et al., 2013) as the main comparison. Words in common correlated $r(1101) = .88, p < .001$, on concreteness; $r(1069) = .87, p < .001$, on imageability; and $r(1102) = .69, p < .001$, on familiarity. We also compared our concreteness ratings with those of Hinojosa et al. (2015), obtaining a correlation of $r(132) = .82, p < .001$. Concerning familiarity, we also considered the study by Moreno-Martínez, Montoro, and Rodríguez-Rojo (2014), obtaining a correlation of $r(132) = .76, p < .001$, for the 134 words that the two databases have in common. As can be seen from the findings above, the results of the correlational analyses indicate a high consistency in the ratings of the variables considered in the present work, despite differences among the studies in terms of procedures and sample sizes.

Evaluation of the normed variables with lexical decision times

To assess the capacity of our normed variables to predict word recognition performance, we computed a linear regression analysis considering the reaction time (RT) values of the 580 overlapping words between our database and that of González-Nosti, Barbón, Rodríguez-Ferreiro, and Cuetos (2014). The dependent variable was the RT to a lexical decision task, whereas the predictors were log frequency, word length (in number of letters), and the six subjective variables rated in the present database.

The R^2 of the model was .54, $F(8, 579) = 85.32, p < .001$. The results showed that the highest standardized regression coefficient (beta) was for log frequency ($\beta = -.31, t = -8.78, p < .001$), with a facilitatory effect over RTs, followed by word length ($\beta = .30, t = 9.69, p < .001$) as an inhibitory effect. From our rated variables, the highest beta value corresponded to subjective familiarity ($\beta = -.27, t = -6.27, p < .001$), with a facilitatory effect over RTs. Concreteness ($\beta = .18, t = 3.46, p < .005$) and context availability ($\beta = -.13, t = -2.63, p < .01$) also showed significant beta values: the first as an inhibitory variable, and the second as a facilitative one. The remaining three variables (i.e., valence, arousal, and imageability) were not significant predictors of RTs (all $ps > .05$). Table 2 shows the correlations between the variables entered into the regression model and RTs, length, and log frequency.

Taken together, these results indicate that the best predictors of RTs are the objective and subjective frequency-related variables (i.e., log frequency and subjective familiarity) and word length, plus modest contributions from concreteness and

Table 2 Pearson correlation coefficients between response latencies (RTs), word length, and log frequency and the six assessed variables

	RT	Length	Log Frequency
RT	–		
Length	.44***	–	
Log frequency	–.59***	–.23***	–
Concreteness	–.04	–.21***	–.08*
Imageability	–.21***	–.32***	.06
Context availability	–.43***	–.18***	.36***
Familiarity	–.58***	–.16***	.55***
Valence	–.19***	–.06	.16***
Arousal	.10**	.19***	.08*

* $p < .05$, ** $p < .01$, *** $p < .001$

context availability. The affective variables and imageability are not predictive at all. However, the results of these analyses should be considered with caution. As can be seen in Table 2 and Table 5 (below), there is a considerable amount of multicollinearity among the variables. For instance, concreteness and imageability are highly correlated, as are log frequency and familiarity. This fact might hinder the interpretation of the roles of the individual predictors in the model, but it would not undermine the validity of our norms because, as we have already seen, there is substantial consistency between our normative data and those from other, similar databases.

Descriptive statistics of the results

Descriptive statistics for the six variables included in the database are presented in Table 3, including data for two relevant indices in psycholinguistics: word length and word frequency per million (available in NIM; Guasch et al., 2013).

Exploration of the relationships between variables

Relationship between affective variables: Valence and arousal

First of all, we explored the relationship between valence and arousal, since previous studies developing affective databases in different languages have commonly reported that these two variables are related (e.g., Bradley & Lang, 1999; Eilola & Havelka, 2010; Ferré et al., 2012; Kanske & Kotz, 2010; Redondo et al., 2005; Redondo et al., 2007; Soares et al., 2012; Võ et al., 2009). To this end, we carried out a regression analysis with valence as the independent measure and arousal as the dependent one. The relation between the two variables was clearly quadratic, $R = .64, F(2, 1397) = 473.47, p < .001$, since this trend accounted for 40.40 % of the variance, whereas a linear relation accounted for only 19.30 % of the variance. Figure 1 shows this typical U-shaped relation between the mean valence and arousal ratings in a two-dimensional affective space. It is notable that a relationship

Table 3 Descriptive statistics of the ratings for valence, arousal, concreteness, imageability, context availability, and familiarity, and of the psycholinguistic indices of length and word frequency per million

	Mean	SD	Min	Max	1st Quartile	Median	3rd Quartile
Valence	4.88	1.70	1.00	8.91	3.78	5.05	5.91
Arousal	5.08	1.34	1.50	8.40	4.00	4.85	6.15
Concreteness	4.56	1.36	1.55	7.00	3.38	4.60	5.71
Imageability	4.61	1.69	1.20	7.00	3.05	4.64	6.29
Context availability	4.84	1.10	1.14	6.86	4.10	4.95	5.71
Familiarity	5.50	0.91	1.65	7.00	4.95	5.64	6.18
Length	7.41	2.06	3.00	13.00	6.00	7.00	9.00
Frequency per million	24.57	38.41	0.18	414.44	4.62	11.19	27.00

of this kind has been the common pattern in the previously mentioned studies (Bradley & Lang, 1999; Eilola & Havelka, 2010; Ferré et al., 2012; Kanske & Kotz, 2010; Redondo et al., 2005; Redondo et al., 2007; Soares et al., 2012; Vö et al., 2009).

As can be seen, there is a clear tendency for arousal to increase along with emotional content, in both the positive and negative domains. Congruently, words that can be considered as neutral in valence tend to also have the lowest arousal values.

To explore this relationship further, we conducted a two-step cluster analysis with Schwarz's (1978) Bayesian information criterion as the clustering criterion, to determine the optimal organization of the items into groups whose members are more similar to each other than to those of the other groups. For valence ratings, the cluster analysis determined that the

best division of the 1,400 words was into three groups, with a silhouette coefficient of 0.7 (Rousseeuw, 1987), which indicates a good tightening of the grouping in three levels. Negative words ranged from 1 to 3.71 ($M = 2.48$, $SD = 0.68$), neutral words from 3.73 to 5.82 ($M = 4.97$, $SD = 0.50$), and positive words from 5.83 to 8.91 ($M = 6.91$, $SD = 0.76$; see Table 4 for the proportions of words falling within each category). As can be seen, the cut points between categories are near to the intuitively used values resulting from splitting an n -point Likert scale into three portions (e.g., 1–3.66 for negative, 3.67–6.33 for neutral, and 6.34–9 for positive on a 9-point Likert scale), but in some ways the use of a cluster analysis sharpens this division. After separating the words into the three groups as mentioned, we computed the pairwise correlation between valence and arousal within each group (see Fig. 1). In the negative range, the correlation was

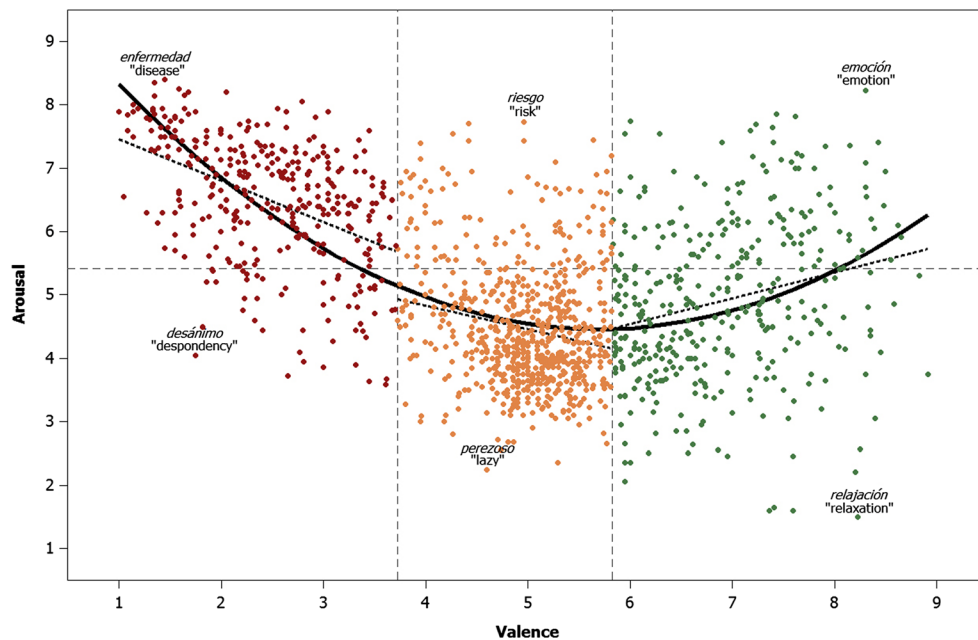


Fig. 1 Valence ratings plotted against arousal ratings (i.e., means for each word across participants), categorized into negative, neutral, and positive words (vertical lines), and into low- and high-arousing words (horizontal

line). Examples of extreme words and their English translations are included

Table 4 Numbers and proportions of words into each combination of valence groups (negative, neutral, and positive) and arousal groups (low and high), according to the division obtained from the cluster analysis

	Arousal					
Valence	Low		High		Total	
Negative	51	3.64 %	291	20.79 %	342	24.43 %
Neutral	578	41.29 %	108	7.71 %	686	49.00 %
Positive	245	17.50 %	127	9.07 %	372	26.57 %
Total	874	62.43 %	526	37.57 %	1,400	100 %

moderate and negative, $r(340) = -.45, p < .001$, whereas in the positive domain it was positive but lower in magnitude than that for the negative words, $r(370) = .24, p < .001$. This result reflects the wider variation in arousal present in the high-valence domain. For neutral words, the correlation was negative and smaller than those for the other two types of words, although it was still significant, $r(684) = -.19, p < .001$. This result indicates that even in the neutral zone the more negative words tend to be the more arousing ones.

The same type of analysis was carried out for the arousal ratings, leading to interesting findings. In contrast to the valence data, the cluster analysis showed that the optimal clustering of arousal ratings was into only two conglomerates (again the silhouette coefficient was of 0.7, but for a solution with just two levels). The range was 1.50–5.41 ($M = 4.19, SD = 0.02$) for low-arousal words, and 5.43–8.40 ($M = 6.57, SD = 0.03$) for the high-arousal ones. This division reflects the fact that whereas it makes sense to split words into three valence groups, with arousal a three-portion division adds nothing in comparison to a simple division into two sets. This finding is in line with the proposal of Vö et al. (2009), who consider that whereas valence is a dimension that includes positive, negative, and neutral words, arousal is better represented as a unipolar dimension ranging from low- to high-arousing words.

Finally, we divided our sample into the three valence groups across the two groups of arousal, and computed the proportion of words falling in each cell (see the reference lines on Fig. 1, and Table 4).

The proportion of words pertaining to the combination of each level of valence with each level of arousal is coherent with the quadratic relation observed between the two variables (i.e., the U-shaped regression line). The most populated region is that corresponding to neutral words with low arousal (e.g., *perezoso*, “lazy”), followed by high-arousal negative words (e.g., *enfermedad*, “disease”). In the positive domain of valence, words are more evenly distributed across the two levels of arousal. The lowest percentage corresponds to low-arousing negative words (e.g., *desánimo*, “despondency”). A similar distribution has been reported by Söderholm et al. (2013) with a sample of 420 Finnish nouns.

Relationship between lexico-semantic variables: Concreteness, imageability, context availability, and familiarity As we stated in the introduction, concreteness, imageability, and context availability are three deeply related variables from a theoretical point of view, and they are often confounded. In order to explore whether the usual pattern of relations was also observed in the present database, we computed the Pearson correlations among these variables (see Table 5 and Fig. 2).

As can be seen, the most strongly correlated variables were concreteness and imageability, indicating that as concreteness increases, it also raises the ease of forming a mental image depicting the meaning of the word. In fact, the scatterplots of concreteness against imageability (see Fig. 2) show that the most populated areas are the extreme ones—that is, concrete words that can be easily imagined (e.g., *mechero*, “lighter”) and abstract words that are difficult to capture with a mental picture (e.g., *asunto*, “issue”). However, we would note that our database also includes words with a low level of concreteness that can be imagined with ease (e.g., *gente*, “people”) and words rated as concrete but difficult to depict (e.g., *tuberculosis*, “tuberculosis”). Thus, despite the high correlation between concreteness and imageability, there are enough words in the database to allow an orthogonal manipulation of the two variables. Such a manipulation is crucial as a means of testing one of the most pervasive theories to account for the concreteness effect: the dual-coding theory (Paivio, 1986). According to this theory, imageability is the source of the concreteness effect: Concrete, but not abstract, words benefit from visual coding because only the referents of the former can be imagined. Logical predictions are that the concreteness advantage will decrease or disappear if concrete words are not easy to imagine, but that abstract words might show an advantage when they are easy to imagine. The existence of concrete words with low imageability values and of highly imageable abstract words in the present database allows for research into this issue.

Next, imageability and context availability also showed a high and positive correlation, suggesting that the easier it is to imagine the content of a word, the faster is access to an appropriate context for its use. Finally, concreteness and context availability showed a moderate but highly significant positive correlation, indicating that concrete words are more easily associated with a context than are abstract words. These results are in agreement with those observed in previous studies (i.e., Altarriba et al., 1999; Paivio, 1971; Schwanenflugel et al., 1988), and therefore confirm the adequacy of our database for the study of processing differences between concrete and abstract words in Spanish.

Finally, we explored the relationship between word familiarity and the concreteness, imageability, and context availability measures. As we stated in the introduction, the relationship between familiarity and concreteness has perhaps led to some confusion in the literature regarding the concreteness

Table 5 Pearson correlation coefficients between the affective and lexico-semantic variables

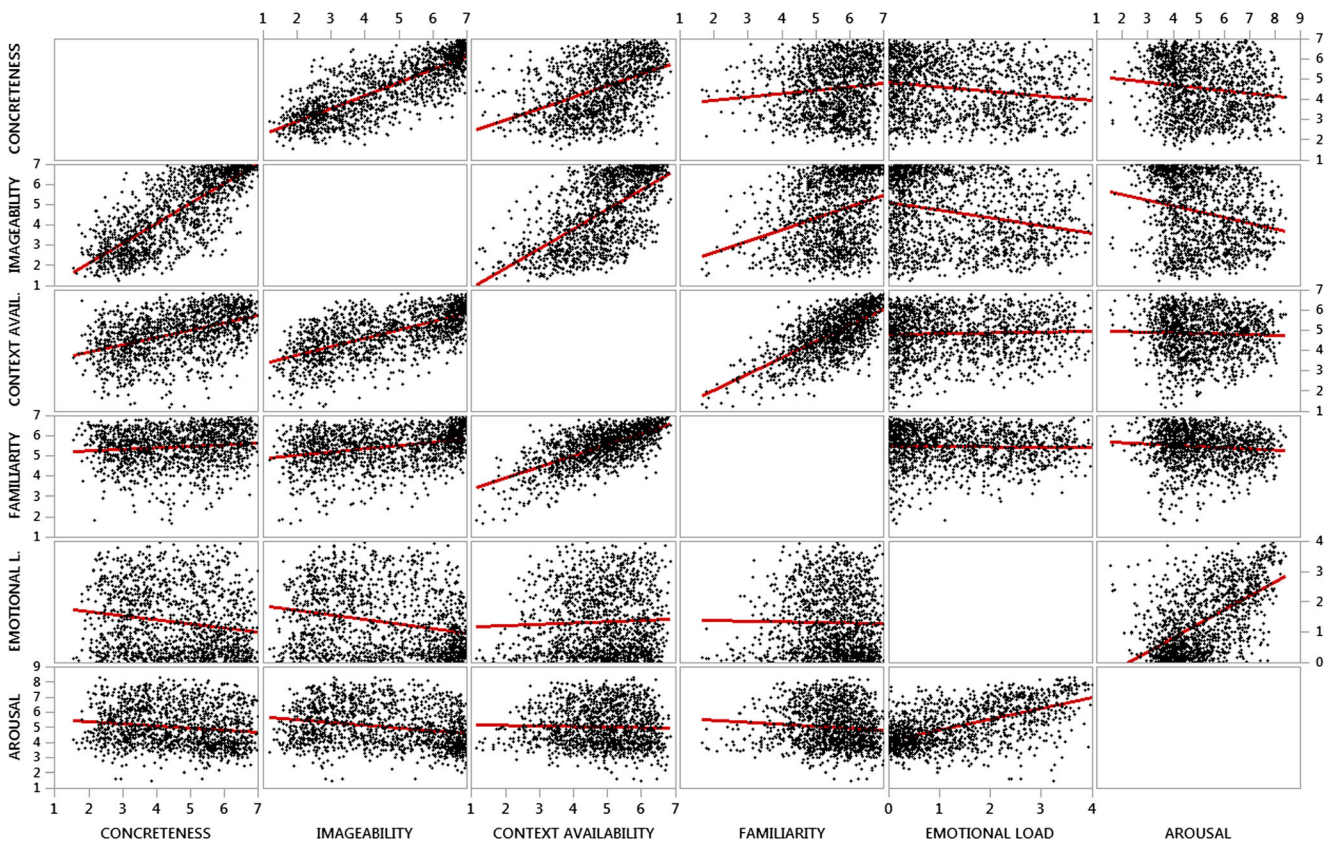
	Concreteness	Imageability	Context Availability	Familiarity	Emotional Load	Arousal
Concreteness	–	.80*	.46*	.11*	–.18*	–.14*
Imageability		–	.63*	.31*	–.24*	–.22*
Context availability			–	.67*	.05	–.03
Familiarity				–	–.02	–.09*
Emotional load					–	.58*
Arousal						–

* $p < .001$

effect (Levy-Drori & Henik, 2006). Taking the above into consideration, it was important to know whether these variables are correlated in the present database. The correlation between familiarity and context availability proved to be the highest, suggesting that highly familiar words are more easily associated with a context than unfamiliar words. In addition, familiarity and concreteness correlated low but significantly. The correlation between familiarity and imageability was somewhat higher and similar to the value of .40 obtained by Citron et al. (2014). These results suggest that concrete and highly imageable words tend to be more familiar than abstract and low imageable ones. Thus, as several authors have pointed out (e.g., Levy-Drori & Henik, 2006), it is appropriate to

control for familiarity in studies on the processing of concrete and abstract words.

Relationship between affective and lexico-semantic variables Because one of the aims of the present work was to provide researchers with affective and semantic (nonaffective) variables that could be manipulated or controlled in experiments, we also examined how these two types of measures were related. The knowledge of the pattern of relationships between these two types of variables is also relevant for our understanding of how the affective and lexico-semantic properties of words are represented in our semantic memory. Thus, we computed Pearson correlations between

**Fig. 2** Scatterplot matrix with linear regression lines of the affective and lexico-semantic variables

these variables (see Table 5 and Fig. 2). We note that emotional valence is distributed across a continuum ranging from negative, to neutral, to positive words. Words at both extremes of the continuum (negative and positive) can be considered as having a high emotional load, whereas words in the middle range would be neutral (without emotional load). Thus, a direct correlation between valence ratings and the other, semantic variables would not be appropriate to examine if emotional words (either positive or negative) are more concrete or imageable than neutral ones. To overcome this problem, we computed a new variable (i.e., *emotional load*) by subtracting 5 (i.e., the middle point of the 9-point Likert scale) from each valence rating and taking the absolute value of the result. After this conversion, the emotional load ratings had a mean value of 1.32 ($SD = 1.08$) and ranged from 0 to 4, where the lowest value corresponded to a completely neutral word, and a value of 4 to a very emotionally loaded one (regardless of whether it belonged to the positive or the negative domain). The Pearson correlation between emotional load and the other variables was then computed. Emotional load showed a negative correlation with concreteness, as well as with imageability. However, the correlation with context availability was not significant. Concerning arousal, we found negative correlations with both concreteness and imageability, but not with context availability. These results suggest that the more concrete and imageable a word is, the less emotionally loaded it is. Thus, although the above correlations are moderate, they would be in line with the ideas of Kousta et al. (2011), who suggested that abstract words have more affective associations than concrete words do.

Conclusions

In this article, a database that provides subjective ratings for both lexico-semantic (concreteness, imageability, context availability, and familiarity) and affective (valence and arousal) properties of 1,400 Spanish words is presented. Although research aimed at studying the effects of lexico-semantic variables and affective dimensions on word processing and recall has traditionally run in different directions, in recent years interest has been growing in studying the relationship between the two. For instance, the potential interaction between concreteness and emotionality has attracted the attention of researchers (e.g., Kousta et al., 2011). Our main aim here was to provide an instrument that would facilitate experimental research into the effects on word processing and memory of both types of variables simultaneously, as well as their potential overlap.

After collecting subjective ratings through current standard procedures, the relationships between (a) affective variables (i.e., valence and arousal), (b) lexico-semantic variables, and (c) affective and lexico-semantic variables were analyzed.

Concerning affective variables, in recent years a considerable amount of research has investigated the issue of how affective properties influence word recognition. Briefly, a remarkable number of studies have shown advantages in processing and memory for emotional words with respect to neutral words (e.g., Ferré et al., 2013; Kuperman et al., 2014; Vinson et al., 2014). The two variables *par excellence* in this domain, valence (pleasantness) and arousal (intensity), were examined in the present study. On the one hand, our results show the typical U-shaped relation between valence and arousal that has been found in previous studies (e.g., Bradley & Lang, 1999; Eilola & Havelka, 2010; Ferré et al., 2012; Kanske & Kotz, 2010), with a quadratic relation that confirms the tendency for arousal to increase as emotional content does (see also Redondo et al., 2005; Redondo et al., 2007; Soares et al., 2012; Vö et al., 2009). On the other hand, the present results also point to the fact that, whereas valence is a dimension that includes positive, negative, and neutral words, arousal might better be conceptualized as a unipolar dimension that ranges from low- to high-arousing words (Vö et al., 2009).

Focusing on lexico-semantic factors, concreteness and imageability were the most strongly correlated variables in the present database. Likewise, imageability and context availability showed a high and positive correlation. Finally, concreteness and context availability showed a moderate but significant positive correlation. In short, these results confirm previous findings in English (Altarriba et al., 1999; Paivio, 1971, 1986; Schwanenflugel et al., 1988; Schwanenflugel & Shoben, 1983) supporting the view that the greater the concreteness, the greater the imageability and ease of accessing an appropriate context for the word, and also the view that the greater the ease of imagining a word, the greater the ease of accessing an appropriate context. As regards familiarity, the analyses conducted showed a high correlation between this variable and context availability, a moderate correlation between familiarity and imageability, and a low but significant correlation between familiarity and concreteness. Indeed, these correlations confirm that concrete and high-imageable words are usually more familiar than abstract and low-imageable ones. These results show the importance of taking into account familiarity in research on the effects of concreteness, since it might have acted as a confounding factor in previous studies (e.g., Levy-Drori & Henik, 2006). Overall, the findings on lexico-semantic variables show the adequacy of the present database for studying differences in processing and recall between concrete and abstract words in Spanish, since it provides researchers with a large set of words that enable the manipulation and control of relevant variables for particular studies.

One variable that has shown a notable contribution to some of the effects related to concreteness is the affective content of words. In particular, Kousta et al. (2011) considered that the “reversed concreteness effect” (i.e., abstract words being

faster and more accurately processed than concrete words) reported recently in several studies (Barber et al., 2013; Kousta et al., 2011) could be explained by the fact that abstract words are more emotionally loaded than concrete ones. In the present study, we examined the relationship between affective and lexico-semantic variables. Both emotional load (i.e., positive and negative words) and arousal showed moderate negative correlations with concreteness and imageability, whereas they did not correlate significantly with context availability. On the one hand, these correlations, although modest, support Kousta et al.'s claim that the more concrete and imageable a word is, the less emotionally loaded and arousing it is—that is, that abstract words tend to be more emotionally loaded. On the other hand, these findings also show the adequacy of the present database for studying the interaction of affective and lexico-semantic variables, in that it allows them to be manipulated orthogonally, an approach that has been adopted in recent research (e.g., Ferré, Ventura, et al., 2015b; Yao & Wang, 2013, 2014).

In summary, the present database provides subjective ratings for 1,400 Spanish words for both affective properties and lexico-semantic variables. Descriptive statistics for valence, arousal, concreteness, imageability, context availability, and familiarity are supplied in an Excel file as [supplementary material](#) to this article. The analyses carried out confirm the reliability and consistency of the present data. Moreover, results regarding both the lexico-semantic and affective variables seem to confirm the patterns of relationships that are commonly found in each specific domain. Finally, our results provide some support for the idea that abstract words might be more emotionally loaded (Kousta et al., 2011) than concrete ones. With this database, future research in Spanish will be able to begin bridging the gap between those studies that have examined how either the lexico-semantic variables or the emotional properties of words affect word processing and memory.

Author note This research was funded by the Spanish Ministry of Economy and Competitiveness (Grant Nos. PSI2012-37623, PSI2012-32834) and by the Autonomous Government of Catalonia (Grant No. 2014SGR-1444).

References

- Altarriba, J., Bauer, L. M., & Benvenuto, C. (1999). Concreteness, context availability, and imageability ratings and Word associations for abstract, concrete, and emotion words. *Behavior Research Methods, Instruments, & Computers*, 31, 578–602. doi:10.3758/BF03200738
- Barber, H. A., Otten, L. J., Kousta, S. T., & Vigliocco, G. (2013). Concreteness in word processing: ERP and behavioral effects in a lexical decision task. *Brain and Language*, 125, 47–53. doi:10.1016/j.bandl.2013.01.005
- Barca, L., Burani, C., & Arduino, L. S. (2002). Word naming times and psycholinguistic norms for Italian nouns. *Behavior Research Methods, Instruments, & Computers*, 34, 424–434. doi:10.3758/BF03195471
- Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T., & Medler, D. A. (2005). Distinct brain systems for processing concrete and abstract concepts. *Journal of Cognitive Neuroscience*, 17, 905–917. doi:10.1162/0898929054021102
- Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The Self-Assessment Manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25, 49–59. doi:10.1016/0005-7916(94)90063-9
- Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings* (Technical Report C-1, pp. 1–45). Center for Research in Psychophysiology, University of Florida.
- Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, 46, 904–911. doi:10.3758/s13428-013-0403-5
- Citron, F. M. M. (2012). Neural correlates of written emotion word processing: A review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and Language*, 122, 211–226. doi:10.1016/j.bandl.2011.12.007
- Citron, F. M. M., Weekes, B. S., & Ferstl, E. C. (2014). How are affective word ratings related to lexicosemantic properties? Evidence from the Sussex Affective Word List. *Applied Psycholinguistics*, 35, 313–331. doi:10.1017/S0142716412000409
- De Groot, A. M. B. (1989). Representational aspects of word imageability and word frequency as assessed through word association. *Journal of Experimental Psychology: Learning Memory and Cognition*, 15, 824–845. doi:10.1037/0278-7393.15.5.824
- De Groot, A. M. B., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning*, 50, 1–56. doi:10.1111/0023-8333.00110
- Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior Research Methods*, 45, 1246–1258. doi:10.3758/s13428-013-0326-1
- Eilola, T. M., & Havelka, J. (2010). Affective norms for 210 British English and Finnish nouns. *Behavior Research Methods*, 42, 134–140. doi:10.3758/BRM.42.1.134
- Eilola, T. M., Havelka, J., & Sharma, D. (2007). Emotional activation in the first and second language. *Cognition and Emotion*, 21, 1064–1076. doi:10.1080/02699930601054109
- Estes, Z., & Verges, M. (2008). Freeze or flee? Negative stimuli elicit selective responding. *Cognition*, 108, 557–565. doi:10.1016/j.cognition.2008.03.003
- Ferré, P., Fraga, I., Comesaña, M., & Sánchez-Casas, R. (2015a). Memory for emotional words: The role of semantic relatedness, encoding task and affective valence. *Cognition and Emotion*, 29, 1401–1410. doi:10.1080/02699931.2014.982515
- 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, 395–403. doi:10.3758/s13428-011-0165-x
- Ferré, P., Sánchez-Casas, R., & Fraga, I. (2013). Memory for emotional words in the first and the second language: Effects of the encoding task. *Bilingualism: Language and Cognition*, 16, 495–507. doi:10.1017/S1366728912000314
- Ferré, P., Ventura, D., Comesaña, M., & Fraga, I. (2015b). The role of emotionality in the acquisition of new concrete and abstract words. *Frontiers in Psychology*, 6, 976. doi:10.3389/fpsyg.2015.00976
- Fliessbach, K., Weis, S., Klaver, P., Elger, C. E., & Weber, B. (2006). The effect of word concreteness on recognition memory. *NeuroImage*, 32, 1413–1421. doi:10.1016/j.neuroimage.2006.06.007

- González-Nosti, M., Barbón, A., Rodríguez-Ferreiro, J., & Cuetos, F. (2014). Effects of the psycholinguistic variables on the lexical decision task in Spanish: A study with 2,765 words. *Behavior Research Methods*, 46, 517–525. doi:10.3758/s13428-013-0383-5
- Guasch, M., Boada, R., Ferré, P., & Sánchez-Casas, R. (2013). NIM: A Web-based Swiss army knife to select stimuli for psycholinguistic studies. *Behavior Research Methods*, 45, 765–771. doi:10.3758/s13428-012-0296-8
- Haro, J. (2012). Testmaker: Aplicación para crear cuestionarios online [Computer software and manual]. Retrieved May 15, 2013, from <http://jharo.net/dokuwiki/testmaker>
- Hinojosa, J. A., Martínez-García, N., Villalba-García, C., Fernández-Folgueiras, U., Sánchez-Carmona, A., Pozo, M. A., & Montoro, P. R. (2015). Affective norms of 875 Spanish words for five discrete emotional categories and two emotional dimensions. *Behavior Research Methods*. doi:10.3758/s13428-015-0572-5
- Kaltwasser, L., Ries, S., Sommer, W., Knight, R. T., & Willems, R. M. (2013). Independence of valence and reward in emotional word processing: Electrophysiological evidence. *Frontiers in Psychology*, 4, 168. doi:10.3389/fpsyg.2013.00168
- Kanske, P., & Kotz, S. A. (2007). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain Research*, 1148, 138–148. doi:10.1016/j.brainres.2007.02.044
- Kanske, P., & Kotz, S. A. (2010). Leipzig affective norms for German: A reliability study. *Behavior Research Methods*, 42, 987–991. doi:10.3758/BRM.42.4.987
- Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General*, 140, 14–34. doi:10.1037/a0021446
- Kuperman, V., Estes, Z., Brysbaert, M., & Warriner, A. B. (2014). Emotion and language: Valence and arousal affect word recognition. *Journal of Experimental Psychology: General*, 143, 1065–1081. doi:10.1037/a0035669
- Lang, P. J. (1995). The emotion probe: Studies of motivation and attention. *American Psychologist*, 50, 372–385. doi:10.1037/0003-066X.50.5.372
- Levy-Drori, S., & Henik, A. (2006). Concreteness and context availability in lexical decision tasks. *American Journal of Psychology*, 119, 45–65. doi:10.2307/20445318
- Majerus, S., & D'Argembeau, A. (2011). Verbal short-term memory reflects the organization of long-term memory: Further evidence from short-term memory for emotional words. *Journal of Memory and Language*, 64, 181–197. doi:10.1016/j.jml.2010.10.003
- Monnier, C., & Syssau, A. (2014). Affective norms for French words (FAN). *Behavior Research Methods*, 46, 1128–1137. doi:10.3758/s13428-013-0431-1
- Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2014). The adaptation of the Affective Norms for English Words (ANEW) for Italian. *Behavior Research Methods*, 46, 887–903. doi:10.3758/s13428-013-0405-3
- Moreno-Martínez, F. J., Montoro, P. R., & Rodríguez-Rojo, I. C. (2014). Spanish norms for age of acquisition, concept familiarity, lexical frequency, manipulability, typicality, and other variables for 820 words from 14 living/nonliving concepts. *Behavior Research Methods*, 46, 1088–1097. doi:10.3758/s13428-013-0435-x
- Paivio, A. (1971). *Imagery and verbal processes*. New York, NY: Holt, Rinehart & Winston.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford, UK: Oxford University Press.
- Real Academia Española. (2001). *Diccionario de la lengua española* (22nd ed.). Madrid, Spain: Espasa Calpe.
- Redondo, J., Fraga, I., Comesaña, M., & Perea, M. (2005). Estudio normativo del valor afectivo de 478 palabras españolas. *Psicológica*, 26, 317–326.
- Redondo, J., Fraga, I., Padrón, I., & Comesaña, M. (2007). The Spanish adaptation of ANEW (Affective Norms for English Words). *Behavior Research Methods*, 39, 600–605. doi:10.3758/BF03193031
- Romani, C., McAlpine, S., & Martin, R. C. (2008). Concreteness effects in different tasks: Implications for models of short-term memory. *Quarterly Journal of Experimental Psychology*, 61, 292–323. doi:10.1080/17470210601147747
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53–65. doi:10.1016/0377-0427(87)90125-7
- Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, 110, 145–172. doi:10.1037/0033-295X.110.1.145
- Schwanenflugel, P. J., & Shoben, E. J. (1983). Differential context effects in the comprehension of abstract and concrete verbal materials. *Journal of Experimental Psychology Learning Memory and Cognition*, 9, 82–102. doi:10.1037/0278-7393.9.1.82
- Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language*, 27, 499–520. doi:10.1016/0749-596X(88)90022-8
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461–464. doi:10.1214/aos/1176344136
- Soares, A. P., Comesaña, M., Pinheiro, A. P., Simões, A., & Frade, C. S. (2012). The adaptation of the Affective Norms for English words (ANEW) for European Portuguese. *Behavior Research Methods*, 44, 256–269. doi:10.3758/s13428-011-0131-7
- Söderholm, C., Häyry, E., Laine, M., & Karrasch, M. (2013). Valence and arousal ratings for 420 Finnish nouns by age and gender. *PloS one*, 8. doi:10.1371/journal.pone.0072859
- Stadthagen-Gonzalez, H., & Davis, C. J. (2006). The Bristol norms for age of acquisition, imageability, and familiarity. *Behavior Research Methods*, 38, 598–605. doi:10.3758/BF03193891
- Talmi, D., & Moscovitch, M. (2004). Can semantic relatedness explain the enhanced memory for emotional words? *Memory & Cognition*, 32, 742–751. doi:10.3758/BF03195864
- Unkelbach, C., Fiedler, K., Bayer, M., Steg Müller, M., & Danner, D. (2008). Why positive information is processed faster: The density hypothesis. *Journal of Personality and Social Psychology*, 95, 36–49. doi:10.1037/0022-3514.95.1.36
- Vega, M., & Fernández, A. (2011). Datos normativos de concreción de 730 palabras utilizadas por sujetos de habla castellana. *Psicológica*, 32, 171–206.
- Vinson, D., Ponari, M., & Vigliocco, G. (2014). How does emotional content affect lexical processing? *Cognition and Emotion*, 28, 737–746. doi:10.1080/02699931.2013.851068
- Vö, M. L. H., Conrad, M., Kuchinke, L., Urton, K., Hofmann, M. J., & Jacobs, A. M. (2009). The Berlin Affective Word List Reloaded (BAWL-R). *Behavior Research Methods*, 41, 534–538. doi:10.3758/BRM.41.2.534
- Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, 45, 1191–1207. doi:10.3758/s13428-012-0314-x
- Yao, Z., & Wang, Z. (2013). The effects of the concreteness of differently valenced words on affective priming. *Acta Psychologica*, 143, 269–276. doi:10.1016/j.actpsy.2013.04.008
- Yao, Z., & Wang, Z. (2014). Concreteness of positive word contributions to affective priming: An ERP study. *International Journal of Psychophysiology*, 93, 275–282. doi:10.1016/j.ijpsycho.2014.06.005
- Yi, H. A., Moore, P., & Grossman, M. (2007). Reversal of the concreteness effect for verbs in patients with semantic dementia. *Neuropsychology*, 21, 9–19. doi:10.1037/0894-4105.21.1.9