

Norms of Valence and Arousal for 14,031 Spanish Words

Hans Stadthagen-Gonzalez^{1*}, Constance Imbault², Miguel A. Pérez Sánchez³, & Marc Brysbaert⁴

¹University of Southern Mississippi, USA

²McMaster University, Canada

³Universidad de Murcia, Spain

⁴Ghent University, Belgium

*Contact Author:

Hans Stadthagen-Gonzalez

Department of Psychology

University of Southern Mississippi

730 East Beach Blvd.,

Long Beach, MS 39560

H.Stadthagen@usm.edu

Abstract

Most current models of research on emotion recognize *valence* (how pleasant a stimulus is) and *arousal* (the level of activation or intensity a stimulus elicits) as important components in the classification of affective experiences (Feldman Barrett, 1998; Kuppens, Tuerlinckx, Russell, & Feldman Barrett, 2012). Here we present a set of norms for valence and arousal for a very large set of Spanish words, including items from a variety of frequencies, semantic categories, and parts of speech, including a sub-set of conjugated verbs. In this regard, we found that there were significant but very small differences between ratings for conjugations of the same verb, validating the practice of applying ratings for infinitives to all derived forms of the verb. Our norms show a high degree of reliability and are strongly correlated with Redondo, Fraga, Padrón, and Comesaña's (2007) Spanish version of the influential ANEW norms (Bradley & Lang, 1999), as well as those from Warriner, Kuperman, and Brysbaert (2013), the largest available set of emotional norms for English words. Additionally, we include measures of word prevalence, that is, the percentage of participants that know a particular word for each variable (Keuleers, Stevens, Mandera, & Brysbaert, 2015). Our large set of norms in Spanish will not only facilitate the creation of stimuli and analysis of texts in that language, but will also be useful for cross-language comparisons and research on emotional aspects of bilingualism.

Keywords: valence; arousal; emotion; conjugated verbs

Since very early on in the research of emotion *valence* and *arousal* have been central to the classification of affective experiences (Wundt, 1912/1924; Osgood, Suci, & Tannenbaum, 1957). *Valence* is a subjective assessment that describes how pleasant a stimulus is (ranging from pleasant to unpleasant), while *arousal* refers to the subjective level of activation or intensity a stimulus elicits (ranging from quiet to active). While there is no absolute consensus about the basic components of affect (e.g., Barrett & Russell, 1999; Fontaine, Scherer, Roesch, & Ellsworth, 2007; Lang, 1995; Larsen & Diener, 1992; Reisenzein, 1994; Russell, 1980; Thayer, 1989; Watson & Tellegen, 1985), most current models recognize valence and arousal as fundamental axes of that construct (Feldman Barrett, 1998; Kuppens, Tuerlinckx, Russell, & Feldman Barrett, 2012). The dimensional model of emotions originally proposed by Wundt includes, along with valence and arousal, a third variable: *dominance*; a subjective assessment of how “in control” a word makes you feel. However, this third variable has not been widely used in the literature (Montefinese, Ambrosini, Fairfield, & Mammarella, 2014), is not a strong predictor of variance of affective judgments (Bradley & Lang, 1994; Lang, Bradley, & Cuthbert, 2008), and is strongly correlated with valence (Warriner, Kuperman, & Brysbaert, 2013).

In addition to their importance in the study of emotion itself, affective variables have been shown to play an important role in multiple areas of cognition; for example, they influence lexical processing (e.g.: Kanske & Kotz, 2007; Kousta, Vinson, & Vigliocco, 2009; Kuchinke, Võ, Hofmann, & Jacobs, 2007; Kuperman, Estes, Brysbaert, & Warriner, 2014; Ortigue et al., 2004; Scott, O’Donnell, Leuthold, & Sereno, 2009; Syssau & Laxén, 2012), they affect recall in short term- (Majerus & D’Argembeau, 2011; Mammarella, Borella, Carretti, Leonardi, & Fairfield,

2013; Monnier & Syssau, 2008) and long-term memory (Dewhurst & Parry, 2000; Kensinger & Corkin, 2003; Talmi & Moscovitch, 2004), and they modulate attention effects (Mathewson, Arnell, & Mansfield, 2008; Stormark, Nordby, & Hugdahl, 1995). The interest in the effects of emotional properties on language processing goes beyond single-words, with several studies looking into their effect on the morpho-syntactic processing of sentences. For example, Hinojosa et al. (2014) proposed that words with negative content can facilitate the processing of gender-agreement violations (at least in the early stages of processing). Similarly, Martín-Loeches et al. (2012) found evidence that the valence of adjectives modulated how sentences are processed at both the syntactic and semantic levels, but Díaz-Lago, Fraga, & Acuña-Fariña (2015) did not find an interaction in the processing of grammatical and emotional properties of words embedded within sentences.

Valence and arousal are, by their very nature, subjective characteristics that reflect experiences and associations with particular objects, events, and words. Values for these variables are usually collected by subjective ratings: Participants are exposed to different stimuli (usually lists of words) and are asked to state how happy or sad, how active or quiet, each stimulus makes them feel. Probably the most influential set of affective norms is Bradley & Lang's (1999) Affective Norms for English Words (ANEW). In ANEW, participants provided their ratings based on the Self-Assessment Manikin (SAM), a non-verbal scale that depicts humanoid figures representing different values along each emotional dimension. Although this method has been widely used in several replications of ANEW, Warriner et al (2013) showed that equivalent normed data can be obtained by using verbal anchors on a numerical scale.

Norms of affective properties of words have been published in multiple languages such as European Portuguese (Soares, Comesaña, Pinheiro, Simões, & Frade, 2012), Brazilian Portuguese (Kristensen, de Azevedo Gomes, Justo, & Vieira, 2011), French (Monnier & Syssau, 2013; Bonin, et al., 2003; Gilet, Grün, Studer, and Labouvie-Vief, 2012, among others), German (Kanske & Kotz, 2010; Lahl, Göriz, Pietrowsky, & Rosenberg, 2009; Vö et al., 2009; Vö, Jacobs, & Conrad, 2006), Polish (Imbir, 2015), Finnish (Söderholm, Häyry, Laine, & Karrasch, 2013; Eilola & Havelka, 2010), Italian (Montefinese, Ambrosini, Fairfield, & Mammarella, 2014) and Dutch (Moors et al., 2013). In Spanish, the largest set of emotional norms so far was published in Redondo, Fraga, Padrón, & Comesaña (2007), which includes valence, arousal, and dominance norms for the Spanish translation equivalents of the original 1034 items from ANEW using a method very similar to that in Bradley & Lang (1999). Similarly, Ferré, Guasch, Moldovan, & Sánchez-Casas (2011) also used Bradley & Lang's (1999) Self-Assessment Manikin to obtain valence and arousal norms for 380 words belonging to three semantic categories, namely, animals, people, and objects. Hinojosa et al.'s (2015) "Madrid Affective Database for Spanish" (MADS) represents an interesting development in that it includes not only norms for arousal and valence for 875 words, but also ratings on five discrete emotional categories: happiness, anger, sadness, fear, and disgust, thus allowing a more fine-grained study of the emotional properties of words. Hinojosa et al. (2015) also includes a useful table listing details of most previous sets of norms of affective characteristics in various languages.

The availability of compatible norms in several languages not only opens the possibility of conducting relevant research in that particular language, but it also allows for cross-language comparison studies (Campos & Astorga, 1988; Harris, Aycicegi, & Gleason, 2003; Russell, 1991), and are useful in the study of how bilingualism affects processing, whether in terms of lexical variables or in terms of the expression of emotion (e.g.: Altarriba & Canary, 2004; Opitz & Degner, 2012). Research in these areas is very active and has yielded interesting results both for bilingualism and the study of emotions. For example, there is evidence that in bilinguals, the activation of emotion is modulated by their level of proficiency in each language (Degner, Doycheva, & Wentura, 2011; Robinson & Altarriba, 2015). Bilinguals often report that emotional words carry a higher level of intensity in their first (or dominant) language, particularly regarding taboo or swear words (e.g.: Dewaele, 2004; Pavlenko, 2012), a difference also observed with measures of autonomic response (e.g.: Colbeck & Bowers, 2012; Harris, Ayçiçeği, & Gleason, 2003). On the other hand, there is evidence that when a bilingual is highly proficient in both languages, an equivalent emotional activation is found in each language, even for late bilinguals (e.g.: Eilola & Havelka, 2007). There may also be some cross-linguistic differences regarding emotion words. For example, there is some recent work (e.g.: Kazanas & Altarriba, in press) on the distinction between emotion words (e.g. joy/anger) and emotion-laden words (e.g. puppy/coffin). The authors found that in English emotion-word processing occurs more quickly than emotion-laden word processing, especially for positive words. However, this effect seems to be less robust in Spanish. Research on the relationship between emotional words and bilingualism also encompasses other areas of cognition such as memory. For example, it has been shown that recall for emotional words is better than for neutral words

in monolingual speakers (e.g., Kensinger & Corkin, 2003; LaBar & Phelps, 1998; MacKay & Ahmetzanov, 2005), but there is a debate as to whether that advantage is also present for the second language of a bilingual. Anooshian and Hertel (1994) did not find such advantage, while Aycicegi and Harris (2004) found that, under certain circumstances, the effect was even larger in the second language. As can be seen, the literature on bilingual processing of emotional words encompasses many different topics and is far from settled, so much remains to be done in that regard. What all these studies have in common is the need for information on the emotional characteristics of large sets of words in multiple languages. Large sets of emotional norms are necessary in order to create well controlled item sets in experimental studies, as well as for input in regression and other types of analysis.

In most of the norming studies mentioned before, the number of items ranges from a few hundred to just over a thousand. Recent years have seen the publication of larger sets of norms of different lexical characteristics, usually including several thousand words (e.g.: Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012; Warriner et al., 2013; Brysbaert, Stevens, De Deyne, Voorspoels, & Storms, 2014). The availability of such large norm databases has greatly facilitated the creation of stimulus sets, as well as making it possible to include such variables in automated analysis of text samples (Leveau, Jhean-Larose, Denhière, & Nguyen, 2012).

Large datasets in addition make it possible to study the relationship between the affective properties of words and other lexical or semantic variables. Warriner et al. (2013), for instance, reported that positive words have a higher frequency than negative words. They related this finding to the observation that the English language has a bias towards positivity or pro-social

benevolent communication interactions (see also Warriner & Kuperman, 2015). The relationship between word frequency and arousal tended to be negative (high frequency words are rated on average as calmer), although this overall relationship hid the fact that some of the high frequency words were very arousing (such as love, god, life, money, hell, kill, fuck, happy). Children seem to be taught positive, low arousing words first, as there is a negative correlation between age-of-acquisition (AoA) and valence together with a positive correlation between AoA and arousal. Finally, Warriner et al. (2013) reported that positive words tended to refer to concepts with high imageability.

In these large norming studies, data tends to be collected in a modular fashion, dividing up the stimuli list into blocks that can be rated in a single session, and introducing inter-rater reliability measures to allow for the integration of the blocks into a single list. This approach allows for the collection of ratings for very large word lists. So far the biggest set of affective norms for English words is Warriner et al. (2013), which provides ratings for valence, arousal, and dominance for nearly 14 thousand words. Our purpose in the current paper was to generate a set of norms for a comparable number of Spanish words using very similar procedures to Warriner et al. Our large database will also allow us to explore how valence and arousal ratings relate to each other and to other lexical variables, namely frequency, age of acquisition, familiarity, imageability, and concreteness that are available in Spanish for large sets of words. Warriner et al. (2013) found weak correlations between emotional variables and most lexical variables included in their analyses, but found interesting relationships between their emotional ratings and the above mentioned lexical characteristics.

An additional objective is to evaluate whether affective properties such as valence and arousal are fully determined by the lemma of a word. A morphologically rich language such as Spanish allows us to evaluate whether, for example, verb conjugations modify the affective properties of the infinitive form. As noted by Rivera, Bates, Orozco-Figueroa, and Wicha (2010), there are very few lexical databases that include information for inflected verbs. Warriner et al.'s (2013) emotional norms included only infinitive verbs in the assumption that emotional values generalize to inflected forms, but so far this assumption has not been empirically tested. Moreover, while English verbs undergo little inflection, in languages with morphologically-rich inflections such as Spanish, infinitives can be quite different from their conjugated forms. We were interested in whether there are differences in emotional ratings with regards to person (first versus third) and tense (present, future, and conditional). With regards to person, our interest arose in view of the literature on self-referential processing of emotional words (e.g. Northoff et al., 2006; Fossati et al., 2003) that points to differences in the processing of emotional words when they are framed as personal attributes of the participant versus another person. With regards to verb tense, we were interested in whether there would be differences in ratings between present (ongoing, immediate actions), future (certain but unfulfilled actions), and conditional (potential actions). It is also possible that such differences, if they exist, only emerge in the presence of larger differences in word form, that is, for irregular verbs, so we looked into that contrast as well. Determining whether different verb conjugations share rating values is important for two reasons: 1) to determine to what extent emotional ratings are dependent on the word-form or are shared by variations of a base word (a "lemma"), and 2)

more pragmatically, to determine whether it is necessary to include conjugated forms in these type of ratings or whether infinitives are enough.

Method

Participants

A total of 512 native speakers of Spanish took part in our study. All respondents were taking undergraduate psychology courses at the University of Murcia, Spain, and received extra credit for their participation. Participants had the choice to rate up to 3 “blocks” of words per variable, and to provide ratings for one or both variables (ratings for Arousal were collected at least 11 weeks after those for Valence). A total of 233 participants rated words for both Valence and Arousal while the rest only did so for one or the other variable. Valence ratings were obtained from 350 raters (of which 294 or 80% were female); of those, 79% completed three blocks of ratings, 9% two blocks, and 12% one block. Raters for Valence had an average age of 22 years 3 months (range= 18 to 62 years; SD= 6 years 1 month). Arousal ratings were obtained from 395 raters (of which 327 or 82.7% were female); of those, 56% completed three blocks of ratings, 21% two blocks, and 23% one block. Raters for Arousal had an average age of 22 years 5 months (range= 18 to 62 years; SD= 5 years 8 months). The dominance of female participants is in line with current practice in psychological experiments and is acceptable given the high correlation of valence and arousal ratings provided by males and females (Montefinese et al., 2014; Moors et al., 2013). Still, it is fair to say that the present data may not be the most optimal to predict the responses of an all-male participant group¹.

Materials

The 14,037 words included in this set of norms were assembled from four sources (with some items being present in more than one source): 1) All 1,034 words from the adaptation of the ANEW norms to Spanish from Redondo, Fraga, Padron, and Comesaña (2007). These words were used as “controls” for validation purposes. 2) In order to include words from a variety of semantic categories, we included 3509 words from Marful, Diez, & Fernandez (2014), a Spanish adaptation of Battig & Montague (1969). In that study, participants were asked to write as many words as they could for 56 semantic categories. For our word list we included all words with availability of 0.7 or more and excluded proper nouns (male and female), geographical names (provinces, countries, and cities), and institutional names (colleges and brands). If both singular and plural were present in the list, only the singular form was included. If only the plural was present, the plural form was included. 3) A sample of 104 verbs to explore the extent to which conjugated forms share emotional ratings with their infinitive form. The list included regular verbs in all three infinitive terminations (-ar, -er, -ir), as well as irregular verbs. We took care not to include verbs with conjugations that could also be a different part of speech (e.g. “juego” is both the conjugated verb “I play” and the noun “game”). For each verb we included its infinitive form as well as the following conjugations: present indicative first person, present indicative third person (which is also the first person imperative for regular verbs), conditional (first and third person share the same word form), and future first person for a total of 520 words. 4) The rest of the words (8974) were taken from EsPal (Duchon, Perea, Sebastián-Gallés, Martí, & Carreiras, 2013), a Spanish database that provides information about many lexical properties based on a very large corpus. In order to include content words with a variety of

grammatical functions, we included words with at least one count per million in EsPal's individual listings of verbs, nouns, adjectives, and adverbs (excluding adverbs ending in “-mente”, the equivalent of “ly” in English). Many of the words included in this list can function as more than one part of speech (e.g.: the same word form can be both a verb and a noun), but taking into account just the most frequent POS category for each word according to EsPal, our list had the following composition: Nouns = 61.8%, Adjectives = 21.3%, Verbs = 15.8%, Adverbs = 0.7%, Determiners = 0.2%, and Pronouns = 0.1%. An effort was made to eliminate foreign words, proper nouns, acronyms, and multi-word utterances from the list. In terms of frequency, the average Zipf value (van Heuven, Mandera, Keuleers, & Brysbaert, 2014) of the set according to EsPal was 3.65 (SD = 0.73, range = 0.51 to 7.46, median = 3.58). This means that our stimuli are nicely centered on the frequency range (Zipf values of 0-3 indicate low-frequency words, values of 4-7 high-frequency words).

Procedure

The stimuli were distributed across 43 blocks with between 351 and 354 words each. Each block consisted of 9 calibrator words, 41 control words from Redondo et al.'s (2007) Spanish adaptation of ANEW, and a random selection of the other words in the list. The calibrators consisted of 3 each of verbs, adjectives, and nouns spanning the full range of the scale in the Redondo et al. norms. A different set of calibrators was chosen for Valence and Arousal and presented at the beginning of each block in order to give participants a sense of the entire range of the stimuli that they would encounter. The calibrators for valence were (in ascending order of valence): *sangriento* (bloody), *funeral* (funeral), *contaminar* (to pollute), *mendigo*

(beggar), *afectar* (to affect), *martillo* (hammer), *liso* (smooth), *mejorar* (to improve), and *alegre* (happy). The calibrators for arousal were (in ascending order of valence): *dormir* (to sleep), *silla* (chair), *nublado* (cloudy), *árbol* (tree), *ahorrar* (to save money), *insecto* (insect), *capaz* (able), *guapo* (handsome), and *ganar* (to win). Control words were randomly interspersed with the other words on the list and used to assess the reliability of our ratings with regards to the Redondo et al. norms. The sets of control words could appear in more than one block and they were similar but not identical for Valence and Arousal blocks. All words in a block (except the calibrators) were randomized once and presented in the same order to each participant rating that block.

Each block was rated by 20 participants. Blocks with ratings that correlated with the mean ratings per items at less than .10 were removed and replaced with ratings from a different participant (2.7% of the collected blocks; 26 blocks Valence; 20 blocks Arousal). These changes are already reflected in the participant information provided above.

Participants accessed the blocks online through Qualtrics. They first completed a brief demographic questionnaire and were then given written instructions for the relevant variable. The instructions given were similar to those of the ANEW norms, the exact wording in Spanish as well as an English translation are provided in the appendix. Most participants rated each block in 30 minutes or less. Participants were asked to rate each word on a 9-point scale as follows: Valence from 1 = “infeliz” (unhappy) to 9 = “feliz” (happy), and Arousal 1= “tranquilo(a)” (quiet) to 9 = “exitado(a)” (excited). There was a further option to indicate that they did not know the word (“No conozco la palabra”). Following Redondo et al., the direction

of the scales is reversed from those used in ANEW, a change intended to make the rating process more intuitive to participants. Each screen displayed approximately 24 words, and the scale (1 to 9) and anchors (*infeliz – feliz; tranquilo – excitado*) were displayed at the top of each group of six words. Following Warriner et al. (2013), we decided not to include the Self-Assessment Manikins (SAM) that were used in the ANEW study and other similar norms; in the Results section, we show that our numerical ratings correlated highly with the SAM ratings from ANEW.

Results and Discussion

For Valence, only 222 words were rated less than 10 times and 3.2% of responses were removed due to participants not recognizing the word. For Arousal, only 183 words were rated less than 10 times and 3.5% of responses were removed due to participants not recognizing the word.

Both valence and arousal were positively skewed, with respective average ratings of 5.2 (SD = 1.27) and 5.3 (SD = 1.50) exceeding the mid-point of the measurement scale (refer to table 1 for valence and arousal extremes). This finding is consistent with other rating studies (Warriner et al. , 2013) and, for valence, supports the bias towards positive word types well established in other languages (cf. Warriner & Kuperman, 2015). Histograms for the distribution of ratings for valence and arousal are shown in Figure 1. Valence ratings had a larger range than arousal; average valence ratings range from 1.1 to 8.8, whereas average arousal ratings range from 1.4 to 8.4. The ratings for both valence and arousal are relatively consistent across participants

(SD=1.27 and 1.50, respectively) when compared to other norming studies of valence and arousal (SD=1.68 and 2.30, respectively) (Warriner et al., 2013).

[Figure 1 about here]

Valence ratings were relatively stable across the valence spectrum (Figure 2, left panel), whereas variance in the arousal ratings is the highest in the lowest-arousal words and decreases linearly as arousal increases (Figure 2, right panel).

[Table 1 about here]

[Figure 2 about here]

Figure 3 shows the relationship between valence and arousal. As seen in previous studies (e.g.: Warriner et al., 2013), the two variables have a U-shaped relationship; very positive and very negative words are highly arousing, whereas neutral words are less arousing ($r^2=0.314$, $p<0.001$).

[Figure 3 about here]

To test whether the correlations with frequency, AoA, and imageability reported by Warriner et al (2013) for the English language also hold for Spanish, we made use of the word frequency data as well as the subjective ratings for Familiarity, Imageability, and Concreteness reported by Duchon et al. (2013) and the AoA ratings reported by Alonso, Fernandez, and Diez (2015). In general, our data show the same patterns as those reported by Warriner et al. (2013; see also Montefinese et al., 2014). As frequency, familiarity, imageability or concreteness of a word increase, the valence rating of that word increases, but the arousal rating decreases. One could

speculate that this is due to the “familiarity” or “mere exposure” effect, which has been shown in multiple domains: frequent exposure to a particular stimulus leads to more positive evaluations of it (Bornstein & Craver-Lemley, 2004; Murphy & Zajonc, 1993; Zajonc, 2001; Zajonc, 1968 among many others, also see Bornstein, 1989 for a meta-analysis). Along the same lines, frequent items could be taken as being more commonplace and therefore less exciting or arousing. This trend is reversed when looking at age of acquisition, words that are learned earlier in life result in higher valence ratings and lower arousal ratings. Words learned later in life are more arousing and more negative (see figure 4 and table 2 for correlations). An explanation could be that children are somehow shielded from “unpleasant” words, such as taboo words and insults, which are only acquired later in life, or at least that people norming words for age-of-acquisition tend to believe so when they provide their ratings.

[Figure 4 about here]

[Table 2 about here]

Reliability of our norms

In order to determine the reliability of our methods, we split the participants randomly into two groups and calculated their mean rating for each word. After recording the correlation between these two groups of participants, we performed this task 99 times to get a set of 99 correlations. The mean correlation coefficient provided us with the measure of split-half reliability, which for valence amounted to 0.88 (0.87 and 0.89, 95% confidence interval) and for arousal amounted to 0.75 (0.73 and 0.76, 95% confidence interval), suggesting that our methods are highly reliable.

Comparison with other norms

Our norms are highly correlated to other Spanish word ratings, including Hinojosa et al. (2015) (for 636 words in common: $r=0.97$, $p<0.001$ and $r=0.71$, $p<0.001$ for valence and arousal, respectively) and Redondo et al. (2007) (for 1031 words in common: $r=0.98$, $p<0.001$ and $r=0.75$, $p<0.001$ for valence and arousal, respectively). We translated the words of our list into English² and found 9403 one-word translation equivalents in common with Warriner et al.'s (2013) list. Previous studies (e.g.: Eilola & Havelka, 2010; Montefinese et al., 2014; Redondo et al., 2007; Soares et al., 2012; Warriner et al., 2013), have shown that ratings for both valence and arousal are relatively consistent across languages, a finding confirmed by the strong correlation between our Spanish norms and Warriner et al.'s English norms ($r=0.79$, $p<0.001$ and $r=0.54$, $p<0.001$ for valence and arousal, respectively).

Conjugated verbs

In Spanish, syntactic information about a word is often expressed in inflectional morphemes added to the word stem (similarly to the suffix –s in the English verb *walks* indicating the third person singular). We examined several word forms of Spanish verbs to test whether syntactic characteristics of verbal tense, person, mode or aspect systematically affect emotional responses to verbal word forms. As can be seen in Table 3, there were no differences between first and third person ratings for either Valence or Arousal. This seems to indicate that the differences between first and second person found on self-referential processing of emotional words using other paradigms (e.g. Northoff et al., 2006; Fossati et al., 2003) are not captured by subjective ratings such as the ones presented here. Furthermore, despite the fact that most of the other differences between conjugated verb forms were significant, the size of such

differences was in all cases very small (a maximum of 4.7% of the 9-point scale for Valence and of 4.0% for Arousal). Tables 4 and 5 show the words with the largest difference between the different conjugations for Valence and Arousal, respectively.

[Table 3 about here]

[Table 4 about here]

[Table 5 about here]

It is possible that differences between emotional ratings of different verb conjugations are more marked for irregular verbs, since the differences between word forms are larger.

However, when we compared ratings for just the irregular verbs in our set ($n=58$), we obtained similar results as before (see Table 6): there are several significant differences, but they are very small in magnitude (less than 6.6% of the scale for valence and less than 6.3% for arousal).

Taken together, and given the small variations between the different conjugations, these results indicate that it is generally safe to assume that emotional ratings for infinitive verbs can be used for their inflected forms.

[Table 6 about here]

Availability of the Norms

The full set of norms is available in a comma-separated (.csv) file as supplementary materials to this article. The words are organized alphabetically and the headings for the table are as follows: ValenceMean = Valence mean value for all valid responses; ValenceSD = Standard Deviation of Valence ratings; %ValenceRaters = number of participants that knew and rated the word for valence; ArousalMean = Arousal mean value for all valid responses; ArousalSD = Standard Deviation of Arousal ratings; %ArousalRaters = number of participants that knew and

rated the word for arousal. %ValenceRaters and %ArousalRaters are included as measures of word prevalence, which has been shown to be a strong predictor of reaction times in lexical decision (Keuleers, Stevens, Mandera, & Brysbaert, 2015). We also indicate each word's dominant part of speech (POS) according to EsPal (Duchon et al., 2013). Ninety five words did not have an entry in EsPal so their POS was determined manually. Entries with less than 10 raters in either or both variables are marked with an asterisk.

General Discussion

In summary, the present paper expands on the work of Redondo et al. (2007) by providing valence and arousal ratings for over 14 thousand Spanish words, greatly increasing the availability of items normed on those variables. We have been able to collect such a large set of norms by using a modular approach, and we have shown that our norms are reliable and compatible with previous, similar, studies. We have also shown that there is little variation in ratings for different conjugations of the same verb. The availability of such a large set of norms will not only be useful in opening up research opportunities in Spanish by helping in the creation of stimuli sets, correlational studies, and text analyses, but its compatibility with the Warriner et al. (2013) norms and others will also be of value for research on emotional aspects of bilingualism, translation studies, and cross-language research.

References

- Ayçiçeği, A. & Harris, C. (2004) Bilinguals' recall and recognition of emotion words. *Cognition and Emotion*, *18*, 977-987. doi: 10.1080/02699930341000301
- Alonso, M.A., Fernandez, A., & Diez, E. (2015). Subjective age-of-acquisition norms for 7,039 Spanish words. *Behavior Research Methods*, *47*, 268-274. doi: 10.3758/s13428-014-0454-2
- Anooshian, L.J. & Hertel, P.T. (1994) Emotionality in free recall: Language specificity in bilingual memory. *Cognition and Emotion*, *8*, 503-514. doi: 10.1080/02699939408408956
- Altarriba, J. & Canary, T.M. (2004). The Influence of Emotional Arousal on Affective Priming in Monolingual and Bilingual Speakers. *Journal of Multilingual and Multicultural Development*, *25*, 248-265. doi: 10.1080/01434630408666531
- Barrett, L. F., & Russell, J. A. (1999). The structure of current affect: Controversies and emerging consensus. *Current Directions in Psychological Science*, *15*, 79-85. doi: 10.1111/1467-8721.00003
- Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut norms. *Journal of Experimental Psychology*, *80*, 1-46. doi: 10.1037/h0027577
- Bonin, P., Méot, A., Aubert, L., Malardier, N., Niedenthal, P., & Capelle-Toczek, M.-C. (2003). Normes de concrétude, de valeur d'imagerie, de fréquence subjective et de valence émotionnelle pour 866 mots. *L'Année Psychologique*, *104*, 655-694. doi: 10.3406/psy.2003.29658

- Bornstein, R.F. & Craver-Lemley, C. (2004). Mere exposure effect. In Pohl, Rüdiger F. (ed.), *Cognitive Illusions: A Handbook on Fallacies and Biases in Thinking, Judgement and Memory* (pp. 215–234). Hove, UK: Psychology Press.
- Bornstein, R.F. (1989) Exposure and affect: overview and meta-analysis of research, 1968–1987. *Psychological Bulletin*, *106*, 265–289. doi: 10.1037/0033-2909.106.2.265
- Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Stimuli, instruction manual and affective ratings* (Technical Report No. C-1). Gainesville, FL: University of Florida, NIMH Center for Research in Psychophysiology.
- 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
- 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. doi: 10.1016/j.actpsy.2014.04.010
- Campos, A., & Astorga, V. M. (1988). Abstractness and emotional values for French and Spanish words. *Perceptual & Motor Skills*, *66*, 649-650. doi: 10.2466/pms.1988.66.2.649
- Colbeck, K. L. & Bowers J. S. (2012). Blinded by taboo words in L1 but not L2. *Emotion*, *12*, 217–222. doi: 10.1037/a0026387

- Degner, J., C. Doycheva and D. Wentura (2011). It matters how much you talk: On the automaticity of affective connotations of first and second language words. *Bilingualism: Language and Cognition*, 15, 181–189. doi: 10.1017/S1366728911000095
- Dewaele, J.M. (2004) The Emotional Force of Swearwords and Taboo Words in the Speech of Multilinguals. *Journal of Multilingual and Multicultural Development*, 25, 204-222. doi: 10.1080/01434630408666529
- Dewhurst, S. A., & Parry, L. A. (2000). Emotionality, distinctiveness and recollective experience. *European Journal of Cognitive Psychology*, 12, 541–551. doi: 10.1080/095414400750050222
- Díaz-Lago, M., Fraga, I., & Acuña-Fariña, C. (2015) Time course of gender agreement violations containing emotional words. *Journal of Neurolinguistics*, 36, 79-93. doi: 10.1016/j.jneuroling.2015.07.001
- 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
- Feldman Barrett, L. (1998). Discrete Emotions or Dimensions? The Role of Valence Focus and Arousal Focus. *Cognition and Emotion*, 12, 579-599. doi: 10.1080/026999398379574
- 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

- Fontaine, J. R. J., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. *Psychological Science, 18*, 1050-1057. doi: 10.1111/j.1467-9280.2007.02024.x
- Fossati, P., Hevenor, S.J., Graham, S.J., Grady, C. et al (2003). In Search of the Emotional Self: An fMRI Study Using Positive and Negative Emotional Words, *The American Journal of Psychiatry, 160*, 1938-1945. doi: 10.1176/appi.ajp.160.11.1938
- Gilet, A. L., Grühn, D. D., Studer, J. J., & Labouvie-Vief, G. G. (2012). Valence, arousal, and imagery ratings for 835 French attributes by young, middle-aged, and older adults: The French Emotional Evaluation List (FEEL). *European Review of Applied Psychology, 62*, 173–181. doi: 10.1016/j.erap.2012.03.003
- Harris, C., L., Ayçiçeği, A., & Gleason, J. B. (2003). Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language. *Applied Psycholinguistics, 24*, 561-579. doi: 10.1017/S0142716403000286
- 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, Epub ahead of print*. doi: 10.3758/s13428-015-0572-5
- Imbir, K.K. (2015). Affective norms for 1,586 polish words (ANPW): Duality-of-mind approach. *Behavior Research Methods, 47*, 860-870. doi: 10.3758/s13428-014-0509-4
- 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

- 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
- Kazanas, S. A., & Altarriba, J. (in press). Emotion Word Processing: Effects of Word Type and Valence in Spanish–English Bilinguals. *Journal of Psycholinguistic Research*. doi: 10.1007/s10936-015-9357-3
- Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition, 31*, 1169–1180. doi: 10.3758/BF03195800
- Keuleers, M., Stevens, M., Mandera, P., & Brysbaert, M. (2015). Word knowledge in the crowd: Measuring vocabulary size and word prevalence in a massive online experiment. *Quarterly Journal of Experimental Psychology, 68*, 1665–1692. doi: 10.1080/17470218.2015.1022560
- Kousta, S.T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition, 112*, 473–481. doi: 10.1016/j.cognition.2009.06.007
- Kristensen, C. H., de Azevedo Gomes, C. F., Justo, A. R., & Vieira, K. (2011). Normas brasileiras para o Affective Norms for English Words. *Trends in Psychiatry and Psychotherapy, 33*, 135–146. doi: 10.1590/S2237-60892011000300003
- Kuchinke, L., Võ, M. H., Hofmann, M., & Jacobs, A.M. (2007). Pupillary responses during lexical decisions vary with word frequency but not emotional valence. *International Journal of Psychophysiology, 65*, 132–140. doi: 10.1016/j.ijpsycho.2007.04.004

- Kuperman, V., Estes, Z., Brysbaert, M., and Warriner, A.B. (2014). Emotion and language: Arousal and valence affect word recognition. *Journal of Experimental Psychology: General*, *143*, 1065-1081. doi: 10.1037/a0035669
- Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30 thousand English words. *Behavior Research Methods*, *44*, 978-990. doi: 10.3758/s13428-012-0210-4
- Kuppens P., Tuerlinckx, F., Russell, J.A., & Barrett L.F. (2012). The relation between valence and arousal in subjective experience. *Psychological Bulletin*, *139*, 917-940. doi: 10.1037/a0030811
- LaBar, K., & Phelps, E. (1998). Arousal-mediated memory consolidation: Role of the medial temporal lobe in humans. *Psychological Science*, *9*, 490–493. doi: 10.1111/1467-9280.00090
- Lahl, O., Göritz, A. S., Pietrowsky, R., & Rosenberg, J. (2009). Using the World-WideWeb to obtain large-scale word norms: 190,212 ratings on a set of 2,654 German nouns. *Behavior Research Methods*, *41*, 13–19. doi: 10.3758/BRM.41.1.13
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International Affective Picture System (IAPS): Affective ratings of pictures and instruction manual (Technical Report A-8)*. Gainesville, FL: University of Florida, Center for Research in Psychophysiology.
- Lang, P. J. (1995). The emotion probe: Studies of motivation and emotion. *American Psychologist*, *50*, 372-385. doi: 10.1037/0003-066X.50.5.372
- Larsen, R.J., & Diener, E. (1992). Promises and problems with the circumplex model of emotion. In M. Clark (Ed.), *Review of Personality and Social Psychology*, *13*, 25- 59.

- Leveau, N., Jhean-Larose, S., Denhière, G., & Nguyen, B. L. (2012). Validating an interlingual metanorm for emotional analysis of texts. *Behavior Research Methods*, *44*, 1007–1014. doi: 10.3758/s13428-012-0208-y
- MacKay, D. G., & Ahmetzanov, M. V. (2005). Emotion, memory, and attention in the taboo Stroop paradigm: An experimental analog of flashbulb memories. *Psychological Science*, *16*, 25–32. doi: 10.1111/j.0956-7976.2005.00776.x
- 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
- Mammarella, N., Borella, E., Carretti, B., Leonardi, G., & Fairfield, B. (2013). Examining an emotion enhancement effect in working memory: Evidence from age-related differences. *Neuropsychological Rehabilitation*, *23*, 416–428. doi: 10.1080/09602011.2013.775065
- Marful, A., Díez, E., & Fernandez, A. (2014). Normative data for the 56 categories of Battig and Montague (1969) in Spanish. *Behavior Research Methods*, [Epub ahead of print]. doi: 10.3758/s13428-014-0513-8
- Martín-Loeches, M., Fernández, A., Schacht, A., Sommer, W., Casado, P., Jiménez-Ortega, L., & Fondevila, S. (2012). The influence of emotional words on sentence processing: Electrophysiological and behavioral evidence. *Neuropsychologia*, *50*, 3262–3272. doi: 10.1016/j.neuropsychologia.2012.09.010

- Mathewson, K. J., Arnell, K. M., & Mansfield, C. A. (2008). Capturing and holding attention: The impact of emotional words in rapid serial visual presentation. *Memory & Cognition, 36*, 182–200. doi: 10.3758/MC.36.1.182
- Monnier, C., & Syssau, A. (2013). Affective Norms for French Words (FAN). *Behavior Research Methods*. Advance online publication. doi: 10.3758/s13428-013-0431-1
- Monnier, C., & Syssau, A. (2008). Semantic contribution to verbal short term memory: Are pleasant words easier to remember than neutral words in serial recall and serial recognition? *Memory & Cognition, 36*, 35–42. doi: 10.3758/MC.36.1.35
- 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
- Moors, A. , De Houwer, J., Hermans, D., Wanmaker, S., van Schie, K., Van Harmelen, A.L., De Schryver, M., De Winne, J., & Brysbaert, M. (2013). Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. *Behavior Research Methods, 45*, 169-177. doi: 10.3758/s13428-012-0243-8
- Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with suboptimal and optimal stimulus. *Journal of Personality and Social Psychology, 64*, 723–739. doi: 10.1037/0022-3514.64.5.723
- Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., & Panksepp, J. (2006). Self-referential processing in our brain--a meta-analysis of imaging studies on the self. *Neuroimage, 31*, 440-457. doi: 10.1016/j.neuroimage.2005.12.002

Opitz, B., & Degner, J. (2012). Emotionality in a second language: It's a matter of time.

Neuropsychologia, 50, 1961-1967. doi: 10.1016/j.neuropsychologia.2012.04.021

Ortigue, S., Michel, C. M., Murray, M. M., Mohr, C., Carbonnel, S., & Landis, T. (2004). Electrical

neuroimaging reveals early generator modulation to emotional words, *NeuroImage*, 21,

1242–1251. doi: 10.1016/j.neuroimage.2003.11.007

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). The measurement of meaning. Urbana:

University of Illinois Press.

Pavlenko, A. (2012). Affective processing in bilingual speakers: disembodied cognition?

International Journal of Psychology, 47, 405–428. doi: 10.1080/00207594.2012.743665

Robinson, C. J., & Altarriba, J. (2015). The interrelationship between emotion, cognition, and

bilingualism. In K. Dziubalska-Kořaczyk, J. Weckwerth, M. Marecka, & M. Gruszecka

(Eds.), *Yearbook of the Poznan linguistic meeting: Vol. 1*. Berlin, Germany: DeGruyter

Open.

Redondo, J., Fraga, I., Padr3n, I., & Comesaña, M. (2007). The Spanish adaptation of

ANEW(Affective Norms for EnglishWords). *Behavior Research Methods*, 39, 600–605.

doi: 10.1037//0022-3514.67.3.525

Reisenzein, R. (1994). Pleasure-arousal theory and the intensity of emotions. *Journal of*

Personality and Social Psychology, 67, 525-539. doi: 10.1037/0022-3514.67.3.525

Rivera, S. M., Bates, E.A., Orozco-Figueroa, A., & Wicha, N.Y.Y. (2010). Spoken verb processing

in Spanish: An analysis using a new online resource. *Applied Psycholinguist*, 31, 29–57.

doi: 10.1017/S0142716409990154

- Russell, J.A. (1991). Culture and the categorization of emotions. *Psychological Bulletin*, *110*, 426- 450. doi: 10.1037/0033-2909.110.3.426
- Russell, J.A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*, 1161-1178. doi: 10.1037/h0077714
- Scott, G. G., O'Donnell, P. J., Leuthold, H., & Sereno, S. C. (2009). Early emotion word processing: Evidence from event-related potentials, *Biological Psychology*, *80*, 95–104. doi: 10.1016/j.biopsycho.2008.03.010
- 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*, e72859. doi: 10.1371/journal.pone.0072859
- Stormark, K.M., Nordby, H., & Hugdahl, K. (1995). Attentional shifts to emotionally charged cues—Behavioral and ERP data. *Cognition & Emotion*, *9*, 507–523. doi: 10.1080/02699939508408978
- Syssau, A., & Laxén, J. (2012). L'influence de la richesse sémantique dans la reconnaissance visuelle des mots émotionnels. *Canadian Journal of Experimental Psychology*, *66*, 70–78. doi: 10.1037/a0027083
- Talmi, D., & Moscovitch, M. (2004). Can semantic relatedness explain the enhancement of memory for emotional words? *Memory & Cognition*, *32*, 742–751. doi: 10.3758/BF03195864

- Thayer, R.E. (1989). *The biopsychology of mood and arousal*. New York: Oxford University Press.
- Van Heuven, W.J.B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, *67*, 1176-1190. doi: 10.1080/17470218.2013.850521
- Vö, M. L., 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
- Vö, M. L., Jacobs, A. M., & Conrad, M. (2006). Cross-validating the Berlin affective word list. *Behavior Research Methods*, *38*, 606–609. doi: 10.3758/BF03193892
- Watson, D. & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, *98*, 219-235. doi: 10.1037/0033-2909.98.2.219
- Warriner, A.B., & Kuperman, V. (2015). Affective biases in English are bi-dimensional. *Cognition and Emotion*, *29*, 1147-1167. doi: 10.1080/02699931.2014.968098
- 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
- Wundt, W. M. (1924). *An introduction to psychology* (R. Pintner, Trans.). London: Allen & Unwin. (original work published 1912).
- Zajonc, R.B. (2001). Mere Exposure: A Gateway to the Subliminal. *Current Directions in Psychological Science*, *10*, 224-228. doi: 10.1111/1467-8721.00154
- Zajonc, Robert B. (1968). Attitudinal Effects Of Mere Exposure. *Journal of Personality and Social Psychology*, *9*, 1–27. doi: 10.1037/h0025848

Notes

1. A comparison between the genders was not indicated for our study, as some ratings were provided by 4 male participants only.
2. A first, rough, translation was done using Google Translate, and then checked word by word by hand for accuracy.

Tables

Table 1: Words with the highest and lowest valence and arousal ratings

Spanish Word	English Translation	Valence Rating	Arousal Rating
<i>Relajado</i>	Relaxed	7.48	1.40
<i>Atentado</i>	Attack	1.85	8.45
<i>Entierro</i>	Burial	1.15	7.35
<i>Libre</i>	Free	8.85	3.73

Table 2: Correlations between word characteristics and their arousal and valence

	Arousal	Valence
Frequency (Zipf value; n=13932)	-0.042, $p < 0.001$	0.163, $p < 0.001$
Familiarity (n=5280)	-0.029, $p = 0.032$	0.176, $p < 0.001$
Imageability (n=5118)	-0.084, $p < 0.001$	0.062, $p < 0.001$
Concreteness (n=5296)	-0.127, $p < 0.001$	0.020, $p = 0.142$
Age of Acquisition (n=5081)	0.184, $p < 0.001$	-0.189, $p < 0.001$

Note: The word frequencies used to calculate the Zipf value, as well as the subjective ratings for Familiarity, Imageability, and Concreteness were obtained from EsPal (Duchon et al., 2013), while Age of Acquisition ratings came from Alonso, Fernandez, & Diez (2015).

Table 3: Comparisons between each tense (reporting mean \pm SE and paired two-sample t-test)

Comparison	Valence	Arousal
1 st person present vs. 3 rd person present	5.33 \pm 0.136 and 5.38 \pm 0.144; n=101, t=-0.281, df=100, p=0.7791	5.32 \pm 0.080 and 5.38 \pm 0.10; n=101, t=-1.11, df=100, p=0.385
Future tense vs. conditional tense	5.33 \pm 0.15 and 5.09 \pm 0.13; n=98, t=-4.11, df=97, p<0.001, Cohen's d=0.172	5.72 \pm 0.07 and 5.54 \pm 0.074; n=98, t=2.52, df=97, p<0.001, Cohen's d=0.149
Present (1 st and 3 rd) vs. future tense	5.34 \pm 0.14 and 5.33 \pm 0.15; n=98, t=0.131, df=97, p=0.896	5.36 \pm 0.086 and 5.72 \pm 0.07; n=98, t=-6.78, df=97, p<0.001, Cohen's d=0.451
Present (1 st and 3 rd) vs. conditional tense	5.34 \pm 0.14 and 5.12 \pm 0.13; n=98, t=3.08, df=97, p=0.00281, Cohen's d=0.167	5.36 \pm 0.086 and 5.54 \pm 0.074; n=98, t=-3.78, df=97, p<0.001, Cohen's d=0.223
1 st person vs. infinitive form	5.33 \pm 0.13 and 5.52 \pm 0.17; n=102, t=-2.96, df=101, p=0.00386, Cohen's d=0.121	5.32 \pm 0.08 and 5.55 \pm 0.10; n=102, t=-4.26, df=101, p<0.001, Cohen's d=0.25
3 rd person vs. infinitive form	5.38 \pm 0.14 and 5.52 \pm 0.17; n=101, t=-2.12, df=100, p=0.0364, Cohen's d=0.0857	5.38 \pm 0.10 and 5.55 \pm 0.10; n=101, t=-3.03, df=100, p=0.00312, Cohen's d=0.171
Future vs. infinitive form	5.34 \pm 0.15 and 5.52 \pm 0.17; n=101, t=-2.85, df=100, p=0.00526, Cohen's d=0.113	5.72 \pm 0.07 and 5.56 \pm 0.10; n=101, t=2.62, df=100, p=0.0101, Cohen's d=0.180
Conditional vs. infinitive form	5.09 \pm 0.13 and 5.51 \pm 0.18; ; n=98, t=-5.22, df=97, p<0.001, Cohen's d=0.274	5.54 \pm 0.07 and 5.56 \pm 0.10; n=98, t=-0.353, df=97, p=0.725

Table 4: Words with the biggest difference in Valence between conjugations

	Spanish Lemma	English Lemma	Spanish Conjugation	English Conjugation
1 st person present vs. 3 rd person present	Leer	To Read	Leo (5.55) Lee (7.50)	I read He/She reads
Future tense vs. conditional tense	Alcanzar	To Reach	Alcanzaré (7.55) Alcanzaría (5.65)	I will reach I would reach
Present (1 st and 3 rd) vs. future tense	Conseguir	To get	Consigo (6.50) Consigue (6.65) Consiguiré (8.00)	I get He/She gets I will get
Present (1 st and 3 rd) vs. conditional tense	Ofender	To offend	Ofendo (2.63) Ofende (2.05) Ofendería (3.90)	I offend He/She offends I would offend
1 st person vs. infinitive form	Comer	To eat	Como (5.75) Comer (7.44)	I eat To eat
3 rd person vs. infinitive form	Resentir	To resent	Resiente (4.47) Resentir (2.90)	He/She resents To resent
Future vs. infinitive form	Pensar	To think	Pensaré (5.65) Pensar (7.30)	I will think To think
Conditional vs. infinitive form	Sentir	To feel	Sentiría (5.80) Sentir (7.95)	I would feel To feel

Table 5: Words with the biggest difference in Arousal between conjugations

	Spanish Lemma	English Lemma	Spanish Conjugation	English Conjugation
1 st person present vs. 3 rd person present	Creer	To Believe	Creo (5.20) Cree (3.85)	I believe He/She believes
Future tense vs. conditional tense	Imaginar	To Imagine	Imaginaré (6.10) Imaginaría (4.60)	I will imagine I would imagine
Present (1 st and 3 rd) vs. future tense	Sentir	To know	Siento (4.80) Siente (3.70) Sentiré (6.10)	I know He/She knows I would know
Present (1 st and 3 rd) vs. conditional tense	Escribir	To write	Escribo (4.40) Escribe (4.10) Escribiría(5.500)	I write He/She writes I would write
1 st person vs. infinitive form	Conocer	To know	Conozco (4.25) Conocer (6.00)	I know To know
3 rd person vs. infinitive form	Sentir	To feel	Siente (3.70) Sentir (5.45)	He/She feels To feel
Future vs. infinitive form	Inspirar	To inspire	Inspiraré (5.80) Inspirar (4.125)	I will inspire To inspire
Conditional vs. infinitive form	Saber	To know	Sabría (5.84) Saber (4.25)	I would know To know

Table 6: Comparisons between each tense of irregular verbs (reporting mean \pm SE and paired two-sample t-test)

Comparison	Valence	Arousal
1 st person present vs. 3 rd person present	5.57 \pm 0.15 and 5.57 \pm 0.15; n=58, t=0.155, df=57, p=0.877	5.07 \pm 0.094 and 5.13 \pm 0.12; n=58, t=-0.9127, df=57, p=0.420
Future tense vs. conditional tense	5.48 \pm 0.18 and 5.23 \pm 0.13; n=56, t=3.06, df=55, p<0.01, Cohen's d=0.218	5.58 \pm 0.095 and 5.35 \pm 0.087; n=56, t=3.62, df=55, p<0.001, Cohen's d=0.346
Present (1 st and 3 rd) vs. future tense	5.57 \pm 0.15 and 5.46 \pm 0.17; n=58, t=1.66, df=57, p=0.102	5.01 \pm 0.099 and 5.58 \pm 0.091; n=58, t=-7.19, df=57, p<0.001, Cohen's d=0.663
Present (1 st and 3 rd) vs. conditional tense	5.59 \pm 0.15 and 5.23 \pm 0.13; n=58, t=5.66, df=57, p<0.001, Cohen's d=0.346	5.10 \pm 0.10 and 5.35 \pm 0.087; n=56, t=-4.52, df=57, p<0.001, Cohen's d=0.348
1 st person vs. infinitive form	5.57 \pm 0.15 and 5.81 \pm 0.20; n=58, t=-3.10, df=57, p<0.01, Cohen's d=0.191	5.07 \pm 0.094 and 5.28 \pm 0.12; ; n=58, t=-2.83, df=57, p<0.01, Cohen's d=0.26
3 rd person vs. infinitive form	5.58 \pm 0.15 and 5.82 \pm 0.19; n=58, t=-2.58, df=57, p<0.05, Cohen's d=0.187	5.13 \pm 0.12 and 5.28 \pm 0.12; n=58, t=-1.78, df=57, p=0.080
Future vs. infinitive form	5.46 \pm 0.17 and 5.81 \pm 0.19; n=58, t=-4.13, df=57, p<0.001, Cohen's d=0.260	5.58 \pm 0.095 and 5.28 \pm 0.12; n=58, t=4.101, df=57, p<0.001, Cohen's d=0.373
Conditional vs. infinitive form	5.23 \pm 0.13 and 5.82 \pm 0.20; n=58, t=-5.47, df=57, p<0.001, Cohen's d=0.480	5.35 \pm 0.22 and 5.28 \pm 0.087; n=56, t=0.886, df=57, p=0.378

Figures

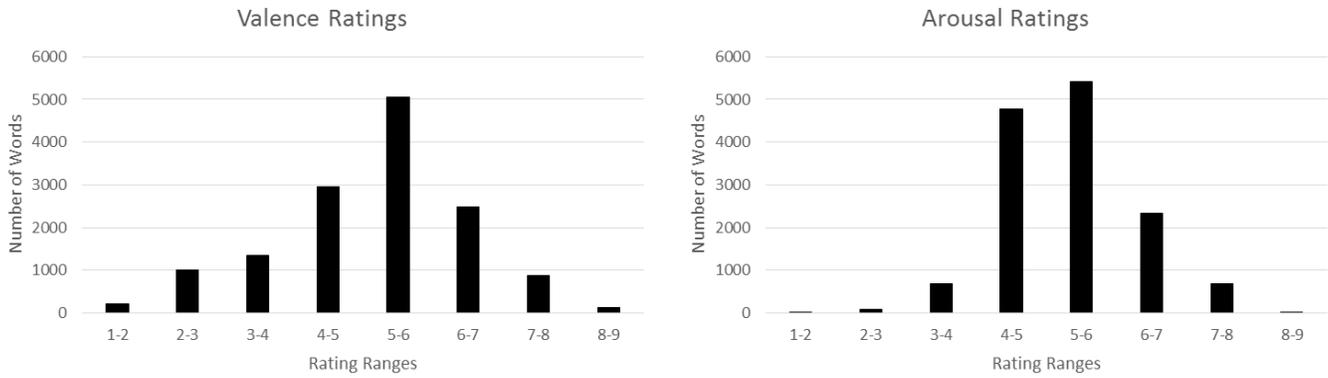


Figure 1: Distribution of ratings for valence and arousal.

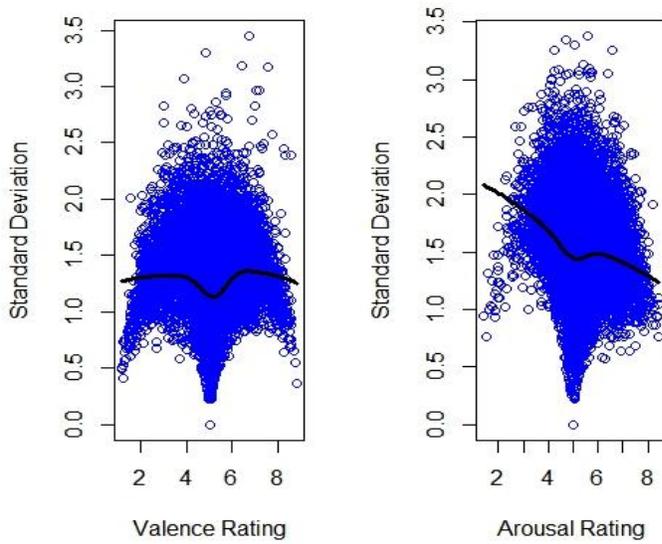


Figure 2: Average standard deviation (variance among responders) across valence and arousal range.

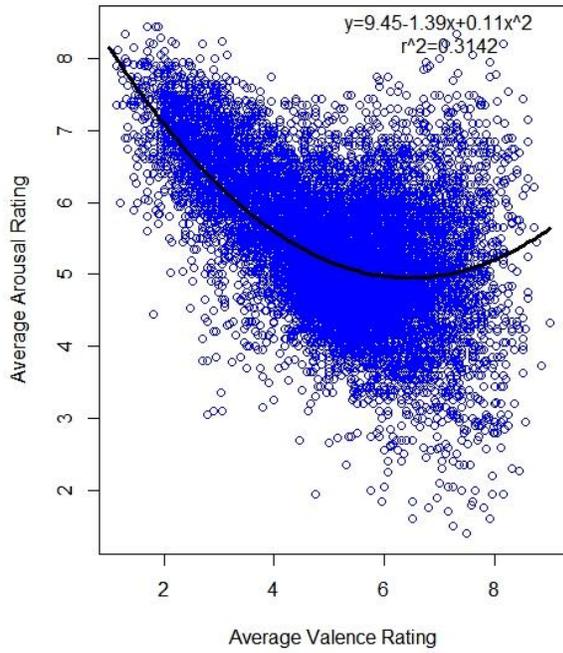


Figure 3: The average valence and arousal rating per word plotted with the quadratic best fit.

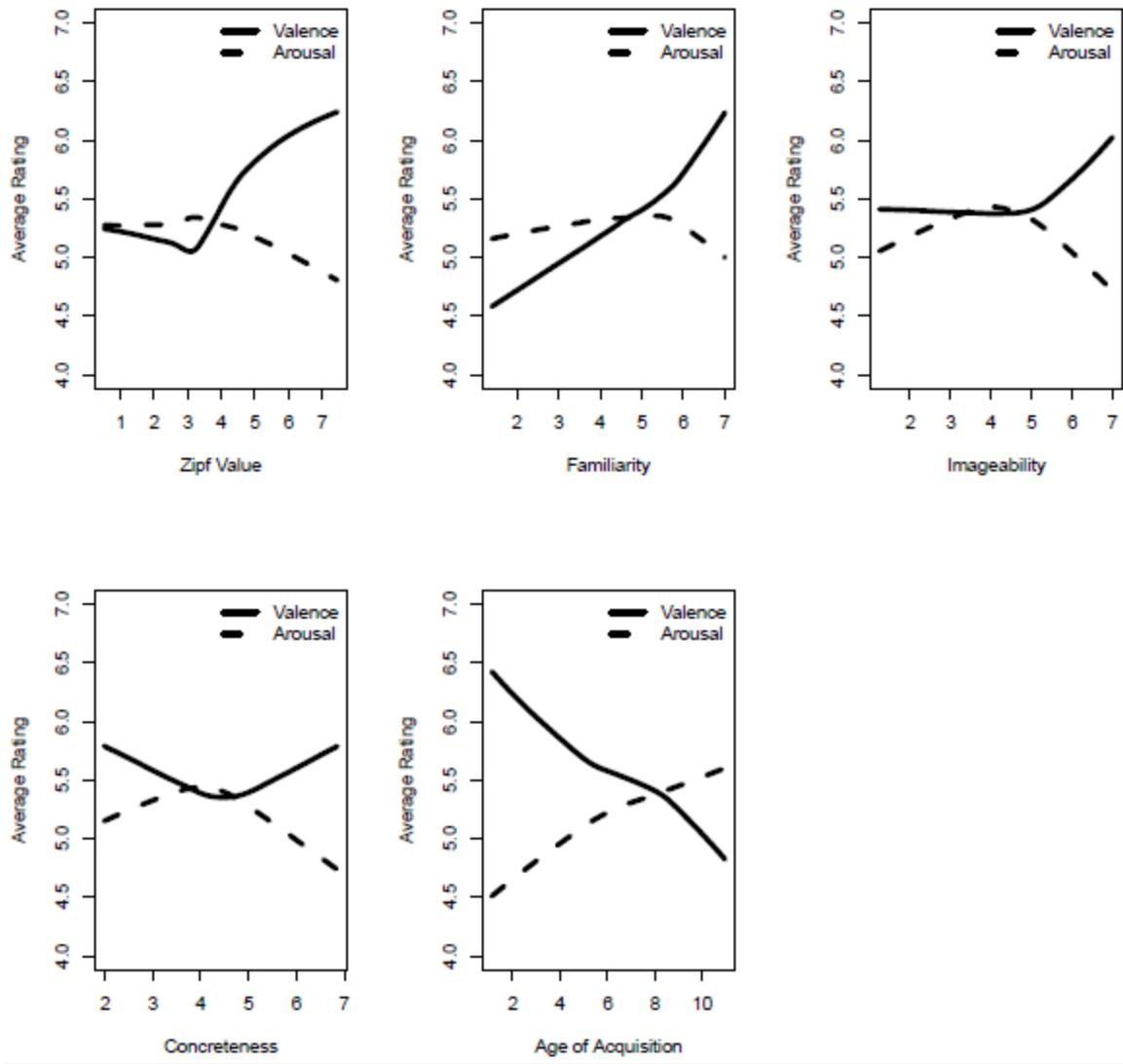


Figure 4: Average frequency (n=13932), familiarity (n=5280), imageability (n=5118), concreteness (n=5296), and age of acquisition (n=5081) across the valence and arousal spectrum.

Note: The word frequencies used to calculate the Zipf value, as well as the subjective ratings for Familiarity, Imageability, and Concreteness were obtained from EsPal (Duchon et al., 2013), while Age of Acquisition ratings came from Alonso et al. (2015).

Appendix

Instructions for VALENCE in Spanish (Adapted from Warriner et al., 2013):

AGRADABILIDAD:

Te invitamos a participar en este estudio sobre las emociones y cómo la gente responde a diferentes tipos palabras. Usarás una escala para indicar cómo te sientes mientras lees cada palabra. La escala que usarás va desde

1 = "Infeliz" hasta 9 = "Feliz"

A un extremo de la escala estás completamente infeliz, molesto(a), insatisfecho(a), melancólico(a), o desesperado(a). Cuando una palabra te haga sentir infeliz debes indicarlo seleccionando el 1.

El otro extremo de la escala es cuando te sientes feliz, alegre, satisfecho(a), contento(a), o esperanzado(a). Puedes indicar el sentirte completamente feliz al leer una palabra seleccionando el 9. Los valores en la escala también te permiten describir sentimientos intermedios de bienestar seleccionando cualquier otro sentimiento. Si te sientes completamente neutral al leer una palabra, ni feliz ni triste, escoge el medio de la escala (el número 5).

Por favor responde rápidamente y no pases mucho tiempo pensando en cada palabra.

Es mejor que hagas tus valoraciones basándote en tu primera reacción inmediata mientras lees cada palabra.

English Translation of the Instructions for VALENCE:

VALENCE:

We invite you to take part in the study about emotions and how people how people respond to different types of words.

You will use a scale to rate how you felt while reading each word. The scale ranges from

1 "Unhappy" to 9 "Happy".

At one extreme of the scale, you are completely unhappy, annoyed, unsatisfied, melancholic, despaired, or bored. When a word makes you feel unhappy you should indicate it by selecting 1.

The other end of the scale is for when you feel happy, pleased, satisfied, contented, or hopeful. You can indicate feeling completely happy when you read a word by selecting 9.

The numbers on the scale also allow you to describe intermediate feelings of pleasure by selecting any other feeling.

If you feel completely neutral when you read a word, neither happy nor sad, select the middle of the scale (number 5).

Please work at a rapid pace and don't spend too much time thinking about each word.

It's better if you base your ratings on your immediate reaction while you read each word.

Instructions for AROUSAL in Spanish:

ACTIVACIÓN:

Te invitamos a participar en este estudio sobre las emociones y cómo la gente responde a diferentes tipos palabras.

Usarás una escala para indicar cómo te sientes mientras lees cada palabra.

La escala que usarás va desde

1 = "tranquilo(a)" hasta 9 = "excitado(a)"

A un extremo de la escala estás relajado(a), tranquilo(a), aletargado(a), aburrido(a), soñoliento(a), o desactivado(a). Cuando una palabra te haga sentir completamente tranquilo(a) debes indicarlo seleccionando el 1.

El otro extremo de la escala es cuando te sientes completamente estimulado(a), excitado(a), frenético(a), nervioso(a), completamente despierto(a) o activado(a). Puedes indicar el sentirte excitado(a) al leer una palabra seleccionando el 9.

Los valores en la escala también te permiten describir sentimientos intermedios de tranquilidad/excitación seleccionando cualquier otro sentimiento.

Si te sientes completamente neutral al leer una palabra, ni excitado(a) ni completamente tranquilo(a), escoge el medio de la escala (el número 5)

Por favor responde rápidamente y no pases mucho tiempo pensando en cada palabra.

Es mejor que hagas tus valoraciones basándote en tu primera reacción inmediata mientras lees cada palabra.

English Translation of the Instructions for AROUSAL:

AROUSAL:

We invite you to take part in the study about emotions and how people how people respond to different types of words.

You will use a scale to rate how you felt while reading each word. The scale ranges from

1 "Calm" to 9 "Excited".

At one extreme of the scale, you are completely relaxed, calm, sluggish, dull, sleepy, or unaroused. When a word makes you feel totally calm you should indicate it by selecting 1.

The other end of the scale is for when you feel stimulated, excited, frenzied, jittery, wide-awake, or aroused. You can indicate feeling excited when you read a word by selecting 9.

The numbers on the scale also allow you to describe intermediate feelings of calmness/arousal by selecting any other feeling.

If you feel completely neutral when you read a word, neither excited nor totally calm, select the middle of the scale (number 5).

Please work at a rapid pace and don't spend too much time thinking about each word.

It's better if you base your ratings on your immediate reaction while you read each word.