CAMBRIDGE JOURNALS

Proof Delivery Form

Journal of Public Policy

Date of delivery:

Journal and vol/article ref: APS 1900050

Number of pages (not including this page): 25

PROOFREADING INSTRUCTIONS

Dear Author,

Here is a proof of your article for publication in *Applied Psycholinguistics*. Please check the proofs carefully, make any corrections necessary on the PDF, and answer the queries on the e-proofs.

Please mark any necessary corrections on the page proofs using Adobe's comment and mark-up tools to convey edits. If the corrections can be explained clearly in a text message, please list the corrections in an email, citing page number, paragraph number, and line number. You may also list the corrections in a doc file. Send the corrections to Brian Mazeski (apsproduction@cambridge.org).

You are responsible for correcting your proofs. Errors not found may appear in the published journal.

The proof is sent to you for correction of typographical errors only. Revision of the substance of the text is not permitted, unless discussed with the editor of the journal.

To order reprints or offprints of your article or a printed copy of the issue, please visit the Cambridge University Reprint Order Center online at: www.sheridan.com/cup/eoc.

Please answer carefully any queries raised from the typesetter.

A new copy of a figure must be provided if correction of anything other than a typographical error introduced by the typesetter is required—please provide this in electronic file form.

Thank you in advance.

QUERIES

AQ1: The distinction between surnames can be ambiguous, therefore to ensure accurate tagging for indexing purposes online (e.g. for PubMed entries), please check that the highlighted surnames have been correctly identified, that all names are in the correct order and spelt correctly.

AQ2: Please check that affiliations of all the authors and the corresponding author details are correctly set.

Applied Psycholinguistics (2019), 1–25
doi:10.1017/S014271641900050X

CAMBRIDGE UNIVERSITY PRESS

ORIGINAL ARTICLE

HEROÍNA: Drug or hero? Meaning-dependent valence norms for ambiguous Spanish words

Daniel Huete-Pérez[1], Juan Haro[1], Isabel Fraga[2] and Pilar Ferré[1,*]

[1]Universitat Rovira i Virgili and [2]Universidade de Santiago de Compostela
*Corresponding author. Email: mariadelpilar.ferre@urv.cat

(Received 31 January 2019; revised 26 July 2019; accepted 22 September 2019)

Abstract

Semantically ambiguous and emotional words occur frequently in language, and the different meanings of ambiguous words can sometimes have different emotional loads. For example, the Spanish word *heroína* (heroin/heroine) can refer to a drug or to a woman who performs a heroic act. Because both ambiguity and emotionality affect word processing, there is a need for normative databases that include data on the emotionality of the different meanings of such words. Thus far, no bases of this type are available in Spanish. With this in mind, the current study will present meaning-dependent affective (valence) ratings for 252 Spanish ambiguous words. The analyses performed show that (a) among ambiguous words, those words with meanings that have distinct affective valence are quite frequent, (b) ambiguous words rated as neutral in isolation can have meanings of opposite valence (i.e., negative-positive or positive-negative), and (c) the valence estimated for ambiguous words in isolation is better explained by the weighted average of the valence of their meanings by dominance. A database of this kind can be useful both for basic research (e.g., relationship between emotion and language and ambiguity processing) and for applied research (e.g., cognitive and emotional biases in emotional disorders and second language learning).

Keywords: ambiguous words; norms; valence

Interest in the interplay between emotion and language has seen rapid growth in recent years. Psycholinguistic studies in this field have mostly focused on the modulation of word processing by emotional content, which is commonly defined in terms of valence and arousal. Valence indicates the hedonic value of an experience, ranging from negative to positive, whereas arousal indicates the degree of activation that it produces, ranging from calming to exciting (Bradley & Lang, 1999). Studies have consistently shown that emotional content affects word processing (e.g., Kousta, Vinson, & Vigliocco, 2009; Kuperman, 2015; Kuperman, Estes, Brysbaert, & Warriner, 2014; Rodríguez-Ferreiro & Davies, 2019; Vinson, Ponari, & Vigliocco, 2014; see Hinojosa, Moreno, & Ferré, 2019, for an overview), although findings concerning the specific effects of positive and negative valence

are mixed, with some reports of facilitative effects restricted to positive words (e.g., Kissler & Koessler, 2011; Kuperman et al., 2014), whereas others extend this to negative words (e.g., Kousta et al., 2009; Vinson et al., 2014). Similarly, arousal has been shown to interact with valence in some studies, but in others has been seen to be independent of it (see Citron, Weekes, & Ferstl, 2014, for an overview). Such discrepancies in the literature might be explained in part by differences in the lexical and semantic variables taken into account for the selection (and matching) of experimental materials.

To obtain experimental stimuli, researchers exploring the effect of emotion on language typically use normative studies in which large numbers of words are characterized in terms of valence and arousal. A variety of such studies have appeared over the last decade, providing affective ratings for different languages (e.g., Eilola & Havelka, 2010; Ferré, Guasch, Moldovan, & Sánchez-Casas, 2012; Guasch, Ferré, & Fraga, 2016; Hinojosa, Martínez-García, et al., 2016; Moors et al., 2013; Redondo, Fraga, Padrón, & Comesaña, 2007; Soares, Comesaña, Pinheiro, Simões, & Frade, 2012; Stadthagen-Gonzalez, Imbault, Pérez-Sánchez, & Brysbaert, 2017; Võ, Jacobs, & Conrad, 2006; Warriner, Kuperman, & Brysbaert, 2013). However, none of these studies has taken into account the fact that words can have more than one meaning, that is, that they can be semantically ambiguous (e.g., *cataract* is an ambiguous word, in that it can mean both "eye disease" and "waterfall"). More specifically, there are words that have distinct, unrelated meanings (so-called homonymous words; e.g., *bat*) and words that have different but interrelated meanings (so-called polysemous words; e.g., *twist*). Normative affective studies have not taken this fact into account. The usual procedure is to collect ratings for large sets of words by asking participants to estimate their valence or arousal without making the words' meanings explicit. Thus, for the valence dimension, participants would be presented with the word *cataract* and asked to rate it on a scale ranging from 1 (*highly unpleasant word*) to 9 (*highly pleasant word*). Ratings would then be averaged across participants to obtain the normative value for that word in terms of valence. This may not be the most appropriate approach for ambiguous words, in that it is unclear what participants do when asked to rate the affective properties of a word with more than one meaning. They might take into account each meaning and produce a rating that is an average value of the properties of the distinct meanings, or, by contrast, they might consider only the properties of the more common or dominant meaning. In any case, both possibilities would lead to inaccurate estimates, because it is unlikely that all the meanings of an ambiguous word have the same affective value. Moreover, in cases such as *cataract*, the different meanings have opposite affective properties (the meaning "eye disorder" has a negative value whereas that of "waterfall" has a positive one). It should be noted that this is not a phenomenon restricted to affective properties: the potential ambiguity of words has been also overlooked in normative studies on semantic properties such as concreteness (e.g., Brysbaert, Warriner, & Kuperman, 2014), imageability (e.g., Davies, Izura, Socas, & Domínguez, 2016), familiarity (e.g., Guasch, et al., 2016), and sensory experience (e.g., Hinojosa, Rincón-Pérez, et al., 2016).

Such neglect is striking if we consider the large number of ambiguous words that exist in any language (e.g., it is estimated that approximately 44% of English words are ambiguous; Britton, 1978) and also the interest that these words have had for psycholinguists over decades (e.g., Adriaens, Small, Cottrell, & Tanenhaus, 1988; Eddington & Tokowicz, 2015; Simpson, 1984). The reason for such interest is that semantic ambiguity depicts a one-to-many mapping between orthography and semantics (e.g., Hino & Lupker, 1996), and hence the study of the phenomenon has contributed to a better understanding of language processing. For instance, it has helped to elucidate how orthography and semantics interact during word processing (Balota, Ferraro, & Connor, 1991; Hino & Lupker, 1996; Pecher, 2001) and how contextual information guides the retrieval of semantic information (Klepousniotou, 2002; Onifer & Swinney, 1981; Swinney, 1979).

It is worth mentioning here that the scientific study of semantic ambiguity involves some methodological challenges, and several of these have still not been fully addressed. Perhaps the most crucial of these is how to characterize ambiguous words correctly. Determining the ambiguity status of a word is not a simple task. In order to decide whether a word is ambiguous or not, as well as what kind of ambiguity it carries (i.e., ambiguity between unrelated [*homonymy*] or related meanings [*polysemy*]), first we need to estimate the number of meanings (NOM) that it has and, in case there is more than one meaning, the relatedness of meanings (ROM). Although there are several different methods to do so, all have some weaknesses and limitations (see Haro, Ferré, Boada, & Demestre, 2017). In brief, there are two main approaches to estimate NOM and ROM. One is based on the linguistic knowledge of the speakers (i.e., subjective methods; e.g., Kellas, Ferraro, & Simpson, 1988; Nelson, McEvoy, Walling, & Wheeler, 1980), and the other is based on the information provided by dictionaries (i.e., objective methods, e.g., Jastrzembski, 1981; Rodd, Gaskell, & Marslen-Wilson, 2002). It should be noted that previous research suggests that using one or the other of these approaches may result in different NOM values for the same word (Ferraro & Kellas, 1990; Gernsbacher, 1984), and lead to opposite experimental results (Haro & Ferré, 2018). In particular, the studies that used dictionary measures to calculate NOM values found a disadvantage for ambiguous words over unambiguous words in the lexical decision task (e.g., Armstrong & Plaut, 2008; Rodd et al., 2002); instead, those studies that used NOM ratings based on participants' knowledge reported the opposite finding (e.g., Haro & Ferré, 2018; Hino, Kusunose, & Lupker, 2010; Pexman, Hino, & Lupker, 2004).

Apart from correctly estimating NOM and ROM, researchers interested in ambiguity have to match the distinct conditions of their experiments in many lexical, semantic, and affective variables that affect word processing. This is particularly important in that semantic ambiguity interacts with some of them (Jager & Cleland, 2016; Syssau & Laxén, 2012; Tokowicz & Kroll, 2007). To obtain the values for those variables, researchers rely on the normative studies mentioned above. However, as we have noted, such databases only contain a single value for each word in the reference variable, without distinguishing between different meanings.

In the present study we aim to fill at least part of this gap by collecting meaning-dependent affective (valence) ratings for a set of Spanish ambiguous words. There

are only a few existing norms that have used a similar approach, by disambiguating the meaning of ambiguous words to obtain ratings for each meaning in different lexicosemantic variables. All these studies have focused on English. For instance, Bird, Franklin, and Howard (2001) collected imageability and age of acquisition ratings for a set of 110 noun–verb homographs (disambiguated by the word "a" or "to" preceding them). Gilhooly and Logie (1980) gathered ratings for imagery, age of acquisition, familiarity, and concreteness for 387 ambiguous words (disambiguated by dictionary-based definitions of each meaning). In turn, Khanna and Cortese (2011) obtained age of acquisition estimates for a large set of ambiguous words ($n = 1{,}208$), including both homographs and homophones (disambiguated by definitions based on free association norms). Finally, Scott, Keitel, Becirspahic, Yao, and Sereno (2019), who collected ratings for a large set of English words on nine dimensions (i.e., arousal, valence, dominance, concreteness, imageability, familiarity, age of acquisition, semantic size, and gender association), identified a set of 379 ambiguous words within the data set for which they obtained meaning-dependent estimates (the words were presented together with another word referring to one of their meanings).

The above studies represent a step toward a better characterization of the semantic–affective properties of ambiguous words, although significant issues remain to be addressed. On the one hand, such studies have not examined the relationship between affective–semantic variables and ambiguity variables, such as meaning dominance. On the other hand, they have not explored in depth the relationship between the ratings of the disambiguated forms of the words (i.e., the words were presented together with information concerning the meaning to which they refer) and the non-disambiguated forms (i.e., the words were presented in isolation, without any meaning cue). One exception is Gilhooly and Logie (1980), who found that the correlation with isolated word ratings was higher for the rating of the dominant meaning than for the ratings of the nondominant (subordinate) meanings. Scott et al. (2019), in turn, carried out an informal examination of the data and observed that although isolated word ratings tended to be closer to the ones of dominant meanings than subordinate meanings, this was not true in all the cases. However, the authors could not draw strong conclusions because they lacked any independent measure of dominance.

In the present study, we aimed to address the above issues by presenting meaning-dependent valence ratings for a set of 252 Spanish ambiguous words. This is the first normative study in Spanish where word meaning is disambiguated. We chose an affective variable due to our interest in the study of the interplay between emotion and language. Furthermore, we focused on valence because of the inconsistencies in the literature with respect to this variable (Kousta et al., 2009; Kuperman et al., 2014; Larsen, Mercer, Balota, & Strube, 2008; Vinson et al., 2014). It might be that ambiguous words with different valence values for each meaning have been included in previous studies, making it difficult to know in each case whether participants were processing the intended meaning. This might have contributed to the discrepancies noted above.

An additional aim of the study was to explore in depth the relationship between affective and ambiguity variables and to examine which is the best estimate of the valence value of isolated words. To that end, we collected valence ratings for

the words presented in isolation (i.e., non-disambiguated) and included information about several relevant ambiguity variables. It should be noted that a large number of words and ambiguity measures of the present database come from the study of Haro et al. (2017), which provides both subjective and objective ambiguity measures for a sample of 530 ambiguous and unambiguous words. The study includes two subjective NOM measures: NOM ratings obtained by asking participants to choose the appropriate number of meanings for a word using a numerical scale (0 = *no meaning*, 1 = *one meaning*, 2 = *more than one meaning*), and an ambiguous/unambiguous categorization made by a group of judges on the basis of the word associates generated by participants. With regard to objective NOM measures, the study provides the number of dictionary definitions and the number of dictionary senses for each word. It also contains two subjective ROM measures, each one obtained by using a different approach. One of them was obtained by asking participants to choose the appropriate ROM value for an ambiguous word by using a numerical scale, ranging from unrelated meanings to related meanings (i.e., the same procedure employed by Hino, Pexman, and Lupker, 2006; ROM_2). The other (novel) measure was obtained by presenting each ambiguous word paired with a lexical associate related to one of its meanings (e.g., SIREN-sea and SIREN-ambulance) and asking participants to estimate the degree of relatedness between those two meanings (ROM_1). Therefore, some of the above ambiguity measures were incorporated into the present database (i.e., number of meanings and relatedness of meanings) as well as a new measure, which was a dominance index. Finally, a set of 252 unambiguous words (i.e., words with a single meaning) were also included. We collected ratings of the same variables as for ambiguous words (except for cases in which this did not make sense, such as relatedness of meanings). The aim was to provide researchers with a large set of well-characterized unambiguous words to which ambiguous words could be compared in future research.

The present database is of potential use to researchers interested in the relationship between emotion and language or in ambiguity processing. For instance, it will be extremely useful in studies where ambiguous words are presented in context, and hence their meaning has to be disambiguated. It will also benefit those working on natural language processing, discourse analysis, or sentiment analysis. Likewise, it will be available for studies on bilingualism and second language learning, as well as cross-language studies. Beyond its application in these areas of psycholinguistic research, the database might also serve as a useful new resource for applied work. Ambiguous words have a role in several contexts such as clinical (e.g., Angwin, Dissanayaka, McMahon, Silburn, & Copland, 2017; Dearing & Gotlib, 2009; Mogg, Bradbury, & Bradley, 2006; Taler, Klepousniotou, & Phillips, 2009), educational (e.g., Degani & Tokowicz, 2010; Kaplan, Fisher, & Rogness, 2009), legal (e.g., Schane, 2002) and biomedical fields (e.g., Stevenson & Guo, 2010). In the clinical context, for example, ambiguous emotional words could be used as an indirect measure to differentiate between populations, such as people with/without depression (Mogg et al., 2006) or people with/without dementia (Taler et al., 2009). Turning to the educational context, we note that ambiguity (e.g., Degani & Tokowicz, 2010; Fang & Perfetti, 2019) and emotionality (e.g., Altarriba & Basnight-Brown, 2011; Ferré, Ventura, Comesaña, & Fraga, 2015) are two variables

that influence learning. For instance, Degani and Tokowicz (2010) found that ambiguous words are harder to learn than unambiguous words. Furthermore, in a series of follow-up studies, they also found that such difficulty was modulated by the degree of relatedness between meanings (Bracken, Degani, Eddington, & Tokowicz, 2017) and the way in which both meanings were introduced to the learner (together or apart; Degani, Tseng, & Tokowicz, 2014). In contrast, Ferré et al. (2015) demonstrated that emotional content facilitates the acquisition of words in a foreign language, especially when those words were abstract, abolishing the typical disadvantage found for abstract words. Hence, characterizing the interaction between ambiguity and emotionality is clearly relevant for the field of novel vocabulary learning.

Method

Participants

A total of 634 Spanish native undergraduate students ($M = 21.00$ years; $SD = 4.25$; 487 women) from the Universitat Rovira i Virgili (URV, Tarragona, Spain) and the Universidade de Santiago de Compostela (USC, Santiago de Compostela, Spain) participated in the study. They were obtained mostly from the same population as the participants of Haro et al. (2017). Participants were recruited using convenience-volunteer sampling. They gave informed written consent and received academic credits for their participation.

Procedure and measures

Overview of the procedure

The database was developed following a series of steps. First, a pool of potential Spanish ambiguous words were identified by consulting previous studies (Armstrong, Tokowicz, & Plaut, 2012; Armstrong, Zugarramurdi, Cabana, Lisboa, & Plaut, 2016; Clifton, 2015; Gawlick-Grendell & Woltz, 1994; Gorfein, Viviani, & Leddo, 1982; Haro et al., 2017). Second, only ambiguous words that seemed to have two different meanings were selected, relying on definitions drawn from the electronic version of the *Diccionario de la Lengua Española* (Real Academia Española, 2014; http://dle.rae.es/). Third, ambiguous words in which the second meaning was probably unknown, according to the authors' intuitions, were discarded. Following these three steps, 410 Spanish ambiguous words remained. Then, to discover whether these words were psychologically ambiguous (i.e., people know that they have more than one meaning),[1] two subjective ambiguity measures were collected (lexical associates and NOM), with 158 words being consequently discarded. For the final set of 252 ambiguous words, ROM and valence ratings were collected. Regarding valence, three different ratings were considered for each word: (a) valence of the word presented in isolation, (b) valence of the word presented together with the definition of one of their meanings, and (c) valence of the word presented together with the definition of the other meaning.

Furthermore, in order to facilitate the comparison of ambiguous and unambiguous words in future research, we included an equal number of unambiguous words

in the database. To that end, we initially incorporated 136 unambiguous words from the database of Haro et al. (2017) and 150 potentially unambiguous words identified from the whole corpus of the emoFinder online search engine (Fraga, Guasch, Haro, Padrón, & Ferré, 2018). After NOM ratings were collected, 34 unambiguous words were discarded in order to have the same number of unambiguous and ambiguous words in the data set (i.e., 252).

As stated above, a large number of the words and measures of the present database come from the study of Haro et al. (2017). In order to collect the data for the new words included in our study, we strictly followed the same procedure and used the same instructions for each variable as Haro et al. (2017). Taking this into account, together with the fact that participants were obtained virtually from the same population, we were able to include all those words in the same data set.

In all the cases in which new ratings were collected, it was done online, by using TestMaker (Haro, 2012). For these new ratings, the following three-step trimming procedure was applied to each questionnaire, resulting in the removal of some participants from the analyses (the *N* reported in each variable is the number of valid responses after the trimming procedure). First, the questionnaire was checked, looking for participants in which data recording problems occurred (e.g., duplicated responses or missing data), 13 participants being eliminated for that reason. Second, the data were graphically analyzed in order to detect aberrant or random response patterns (e.g., giving the same response to all the stimuli), 1 participant being eliminated for that reason. Third, the Pearson correlation between the data of each participant and the mean of all participants was calculated. As a result, 21 participants who had low indexes were eliminated ($r < .40$): values close to zero were interpreted as idiosyncratic response patterns that do not represent the general trend, while negative values were attributed to an interpretation of the scale in the opposite direction.

Overview of the measures

The main measures of the present norms are number of dictionary entries and senses, number of lexical associates, NOM, ROM, valence of words presented in isolation, and valence of each meaning. Secondary measures derived from the principal ones are emotionality, a dominance-subordination index (DSI), and an index of the relative emotionality of meanings (IREM). A summary of all these measures can be seen in Table 1.

First measure of ambiguity: Number of dictionary entries and senses. To provide an objective measure of ambiguity, the number of dictionary entries and senses (irrespective of whether all the senses were listed in a single dictionary entry or not) of the electronic version of the *Diccionario de la Lengua Española* (Real Academia Española, 2014; http://dle.rae.es/) were included for each word in the data set.

Second measure of ambiguity: Associates. The associates of 221 of the 410 Spanish ambiguous words were directly obtained from the raw data of Haro et al. (2017). For the remaining 189 words, new data were collected. These 189 words were

Table 1. Description and source of each measure included in the database

Variable	Description	Source
Dictionary entries	Number of dictionary entries of the word	RAE (2014)
Dictionary senses	Number of dictionary senses of the word	RAE (2014)
Associates	Number of words related to each meaning generated in a free-association task	Haro et al. (2017) + New ratings
DSI	Ratio of associates of the first meaning in relation to associates of the second meaning ($Associates_{M1}$ / $Associates_{M2}$)	Calculated from associates
NOM	Subjective number of meanings of the word (0 = *no meaning*, 1 = *one meaning*, 2 = *more than one meaning*)	Haro et al. (2017) + New ratings
ROM	Subjective relatedness of meanings of the word (from 1 = *unrelated meanings* to 9 = *same meaning*)	Haro et al. (2017) + New ratings
Valence (isolation)	Subjective hedonic value (positive-negative) of the word presented in isolation (from 1 = *completely sad* to 9 = *completely happy*)	Haro et al. (2017) + Fraga et al. (2018) + New ratings
Valence (meaning-dependent)	Subjective hedonic value (positive-negative) of the word presented together with the definition of either the first meaning or the second meaning (from 1 = *completely sad* to 9 = *completely happy*)	New ratings
Emotionality	Absolute value of the emotionality of the word, regardless of if it is positive or negative, computed with respect to the central point of the scale (|5 – *valence*|)	Calculated from valence
IREM	Difference between the emotionality of the two meanings of the word ($Emotionality_{M1}$– $Emotionality_{M2}$)	Calculated from emotionality

Note: DSI, dominance-subordination index. M1, first/dominant meaning. M2, second/subordinate meaning. NOM, number of meanings. ROM, relatedness of meanings. IREM, index of the relative emotionality of meaning.

randomly distributed in two questionnaires, and 50 participants ($M = 22.50$ years; $SD = 4.54$; 45 women) were asked to write the first word that came to mind after reading each word. Two of the authors categorized each associate produced by the participants as (a) related to one of the two meanings previously considered, (b) related to another meaning not considered in the selection, or (c) not clearly related to one or the other meaning. This task was first performed individually, and the two judges met later to compare their results and discuss discrepancies (three associates was the maximum accepted discrepancy). The values included in the database were obtained by recounting the number of produced associates that could be associated to each meaning. For example, considering the Spanish word *arca* (ark/coffer) 23 associates were generated for one of the meanings

(17 participants produced *Noé* [Noah], 3 participants produced *barco/a* [boat], 1 participant produced *animales* [animals], 1 participant produced *mar* [sea], and 1 participant produced *biblia* [Bible]), whereas only 1 associate was generated for the other meaning (i.e., 1 participant produced *arcón* [chest]). Hence, for the Spanish word *arca*, we can say that *ark* is the dominant meaning (i.e., the one that has more associates) and *coffer* is the subordinate meaning (i.e., the one that has fewer associates).[2] As a result of the overall process, 24 words were removed because the associates generated were not clearly related to one or the other meaning (e.g., for the Spanish word *basurero*, most of the associates did not permit a distinction between the two meanings, *dustman* and *dumpster*). In addition, 26 words were eliminated because the associates revealed a third meaning that was not taken into consideration during the selection of the words (e.g., for the Spanish word *granito*, apart from the expected meanings of *granite* and *acne*, some participants also generated associates for the unexpected meaning of *small particle*).

An DSI of meanings was calculated through the division of the number of associates related to the first or dominant meaning by the number of associates related to the second or subordinate meaning[3]: an index of 1 would mean that the two meanings have the same number of associates, hence both meanings are balanced (e.g., as is the case with the two meanings of the Spanish word *genio* [gifted/magician]); an index of 2 would mean that the dominant meaning has twice as many associates as the subordinate meaning; an index of 3 would mean that the dominant meaning has three times as many associates as the subordinated meaning; and so on.

Third measure of ambiguity: NOM. The NOM of 221 of the 410 Spanish ambiguous words was directly obtained from the database of Haro et al. (2017), and for the remaining 189 words new data were collected. These 189 words were randomly distributed in three questionnaires, including the same number of fillers (half of them were unambiguous words [e.g., *geología* {geology}] and the other half were nonwords [e.g., *frubdión*]). A total of 67 participants ($M = 23.28$ years; $SD = 7.38$; 49 women) were asked to decide if each string of letters had *no meaning* (coded as 0), *one meaning* (coded as 1), or *more than one meaning* (coded as 2). At a later time, this procedure was repeated for the unambiguous words included in the database in order to check that they had a single meaning: the NOM of 136 of the 286 unambiguous words was taken directly from the database of Haro et al. (2017), and NOM ratings for the remaining 150 potentially unambiguous words were obtained through questionnaires ($N = 20$ participants), in which we included nonword fillers.

In order to decide which of the 410 Spanish ambiguous words were psychologically ambiguous (i.e., they had more than one meaning for the participants), two criteria had to be met: (a) at least one associate had to be generated for the subordinate meaning (note that cases with 0.5 associates in the subordinate meaning were included, because one of the judges, but not the other, considered that there was one associate related to that meaning), and (b) the NOM had to be equal to or higher than 1.40. This was the same NOM criterion to distinguish between

ambiguous and unambiguous words as that employed in previous studies (e.g., Haro, Comesaña, & Ferré, 2019; Haro & Ferré, 2018; Hino et al., 2006). A total of 108 words were rejected because they did not meet both criteria. Taking those words, together with the 24 words eliminated because they had several associates that could not be clearly related to one of the meanings, and also the 26 words that were eliminated because they had associates related to a third meaning, the final number of ambiguous words included in the data set was 252. Regarding unambiguous words, we removed from the data set a word whose NOM was higher than 1.40 (*tabique* [septum/partition]). Apart from that word, 33 out of 286 potentially unambiguous words were discarded in order to have the same number of unambiguous and ambiguous words in the data set (i.e., 252). To do that, we followed rational criteria (e.g., words that had a NOM lower than 1, suggesting that these words are relatively unknown or infrequent).

Measure of ROM. The ROM of 183 of the 252 Spanish ambiguous words were taken directly from the ROM_1 data of Haro et al. (2017), and for the remaining 69 new data were collected. All the words were presented in a single questionnaire. Each word was presented together with the most frequent associate generated for each meaning (e.g., the Spanish word *órgano* [organ] was presented with the associates *cuerpo* [body] and *música* [music]). Twenty-six participants ($M = 20.62$ years; $SD = 1.86$; 23 women) were asked to decide the degree of relationship between the two meanings of the ambiguous word indicated in the word pair on a 9-point scale (ranging from 1 = *unrelated meanings* to 9 = *same meaning*).

Measure of the valence of words in isolation. Following the same instructions and procedures of the published databases (e.g., Guasch et al., 2016), the Spanish adaptation of the Self-Assessment Manikin was used in order to obtain the valence rating of each ambiguous word when presented in isolation. The 252 words were randomly distributed in eight questionnaires, including the same number of unambiguous words (e.g., *ajedrez* [chess]) in each questionnaire as fillers. A total of 200 participants ($M = 21.09$ years; $SD = 4.07$; 161 women) rated the valence of the words on a 9-point scale (ranging from 1 = *completely sad* to 9 = *completely happy*).

Valence ratings for unambiguous words were directly taken from the following sources with the help of the emoFinder online search engine (Fraga et al., 2018): Ferré et al. (2012), Guasch et al. (2016), Haro et al. (2017), Hinojosa, Martínez-García, et al. (2016), and Stadthagen-González et al. (2017).

Measure of the valence of each meaning (meaning-dependent valence). In order to obtain the valence of each meaning separately, the Self-Assessment Manikin was again used. However, in this case, words were not presented in isolation. Instead, they were presented together with the definition of either the dominant meaning or the subordinate meaning. The definitions were adapted from the electronic version of the *Diccionario de la Lengua Española* (Real Academia Española, 2014; http://dle.rae.es/) and the electronic version of the dictionary of the *El Mundo* newspaper (http://www.elmundo.es/diccionarios/, no longer available), by keeping the length and difficulty of the definitions as uniform as possible (although a rigorous match was not performed). Participants were instructed that some of

the words could have more than one meaning, and that they had to evaluate the valence of each word considering only the meaning that was indicated by the definition (see Appendix A).

The 252 words were randomly distributed in eight questionnaires, each word appearing in two of these: in one questionnaire, the word appeared with the definition of the dominant meaning, and in the other questionnaire, it appeared with the definition of the subordinate meaning. A total of 236 participants ($M = 20.16$ years; $SD = 2.98$; 173 women) rated the valence of the words on a 9-point scale (ranging from 1 = *completely sad* to 9 = *completely happy*).

The emotionality for each meaning was calculated by subtracting the valence of that meaning to the middle point of the valence scale (i.e., 5), and was reported as the absolute value. Values in this variable range from 0 = *the meaning is completely neutral* (e.g., the dominant meaning of *segundo* [second], which refers to the "time unit" meaning) to 4 = *the meaning is extremely emotional* (e.g., the subordinate meaning of *viola* [viola/rapes], which refers to the present tense of the verb "to rape" and has an emotionality value of 3.86). For each word, an IREM[4] was calculated through the subtraction of the emotionality of the subordinate meaning to the emotionality of the dominant meaning: an index of 0 means that both meanings are equally emotional (e.g., *cinturón* [belt/seatbelt]), while positive values indicate that the dominant meaning is more emotional than the subordinate meaning (e.g., *mata* [kills/bush]) and negative values indicate that the subordinate meaning is more emotional than the dominant meaning (e.g., *terminal* [airport terminal/ terminal illness]).

Availability of the norms

The database can be downloaded from http://psico.fcep.urv.cat/exp/files/huete_et_al_emo_amb_database.xlsx as an Excel file, in which the following columns can be found: *Word* (word in Spanish), *Ambiguity_status* (whether the word is ambiguous or unambiguous), *Dictionary_entries* (number of dictionary entries), *Dictionary_senses* (number of dictionary senses), *NOM* (mean number of meanings), *ROM* (mean relatedness of meanings), *Associates_M1* (number of associates generated for the dominant meaning), *Associates_M2* (number of associates generated for the subordinate meaning), *DSI* (dominance-subordination index), *Valence_isolation_m* (mean valence of the word presented in isolation), *Valence_isolation_sd* (standard deviation of the valence of the word presented in isolation), *Valence_M1_m* (mean valence of the word presented together with the definition of the dominant meaning), *Valence_M1_sd* (standard deviation of the valence of the word presented together with the definition of the dominant meaning), *Emotionality_M1* (emotionality of the dominant meaning), *Valence_M2_m* (mean valence of the word presented together with the definition of the subordinate meaning), *Valence_M2_sd* (standard deviation of the valence of the word presented together with the definition of the subordinate meaning), *Emotionality_M2* (emotionality of the subordinate meaning), *IREM* (index of the relative emotionality of meanings), *Definition_M1* (Spanish definition of the dominant meaning), *Definition_M2* (Spanish definition of the subordinate meaning), and *Source* (indicates whether the word was taken from the database of Haro et al., 2017, or it was a new word).

Table 2. Descriptive statistics of the 252 ambiguous words and the 252 unambiguous words included in the database

	Val	Ent	Sen	NOM	ROM	DSI
Ambiguous words						
Mean	5.25	1.52	8.65	1.73	2.49	8.28
SD	1.14	0.70	5.61	0.15	1.09	9.40
MNVR	25	–	–	21	26	25
Unambiguous words						
Mean	5.00	1.02	2.42	1.09	–	–
SD	1.60	0.19	1.99	0.08	–	–
MNVR	–	–	–	20	–	–

Note: All values are rounded to the last decimal. Val, valence in isolation. Ent, number of dictionary entries. Sen, number of dictionary senses. NOM, number of meanings. ROM, relatedness of meanings. DSI, dominance-subordination index. MNVR, minimum number of valid responses when questionnaires were used.

Results

All analyses were conducted with IBM SPSS Statistics (version 25) and JASP (version 0.9.2.0). Where applicable, the critical level of significance was $\alpha = .05$. In cases where the data seriously violate the assumptions of the parametric tests (e.g., a distribution that does not meet the normality criteria), their nonparametric equivalents were used (e.g., the Mann–Whitney *U* test instead of the independent samples *t* test, or Kendall's τ correlation instead of the Pearson correlation). In the Bayesian[5] analyses, the default options in JASP were left unchanged.

Reliability

Following Guasch et al. (2016), the intraclass correlation coefficient (ICC; for an overview of this coefficient, see Koo & Li, 2016) was computed for each questionnaire in order to obtain an index of interrater reliability of each measure. More specifically, the type of ICC used was the two-way random effects based on the absolute agreement of multiple raters. The ICCs were all significant (all $p < .001$), supporting the reliability of the data: ICC = .93–.95 for valence in isolation ($M = .94$, $SD = .01$), ICC = .96–.98 for valence with the definition of each meaning ($M = .97$, $SD = .01$), ICC = .98–.99 for NOM ($M = .98$, $SD = .01$), ICC = .93 for ROM.

Ambiguity measures

The descriptive statistics of ambiguous and unambiguous words are represented in Table 2.

Comparison between ambiguous and unambiguous words

The validity of the a priori arbitrary cutoff point established to differentiate between ambiguous and unambiguous words (i.e., a NOM value of 1.40) was reinforced by

Table 3. Kendall's tau correlation coefficients (first value) and the respective BF_{10} (in parentheses) between objective and subjective measures of ambiguity

	Ent	Sen	NOM	ROM[a]
Ent	–			
Sen	.34*** (3.87 e+27)	–		
NOM	.33*** (3.15 e+25)	.56*** (8.05 e+74)	–	
ROM[a]	−.22*** (3.27 e+4)	−.01 (0.08)	−.12** (4.70)	–

Note: All values are rounded to the last decimal. Ent, number of dictionary entries. Sen, number of dictionary senses. NOM, number of meanings. ROM, relatedness of meanings. [a]The correlations with ROM were only conducted with the ambiguous words ($N = 252$), but the correlation between the other variables were conducted with all the words ($N = 504$). $^{**}p < .01$. $^{***}p < .001$.

the significant differences observed in objective ambiguity measures, such as the number of dictionary entries, $U = 19216.00$, $z = -10.81$, $p < .001$, $r_{rb} = .39$, $BF_{10} = 6.68$ e+8, and the number of dictionary senses, $U = 5009.50$, $z = -16.49$, $p < .001$, $r_{rb} = .84$, $BF_{10} = 2.52$ e+11.

Correlation between objective and subjective measures of ambiguity

The correlations between the different measures of ambiguity are presented in Table 3. The pattern of correlations found reveals that most of the ambiguity variables are significantly correlated, supporting the validity of those variables as indices of ambiguity. This pattern is similar to that found by Haro et al. (2017). The only difference is that number of senses and ROM are not significantly correlated in the present database. Such discrepancy may be explained by a methodological difference: while in the study of Haro et al. (2017) ROM was collected for both ambiguous and unambiguous words, in the present study it was collected for ambiguous words only. If we take into account that excluding unambiguous words restricts the range/variability in both number of senses and ROM values (unambiguous words tend to have lower values of number of senses and higher values of ROM than ambiguous words), it is not surprising that the correlation diminishes (for a reminder of how range restriction decreases correlation coefficients, see, for example, Goodwin, & Leech, 2006). If we remove all the unambiguous words from the data of Haro et al. (2017), the correlation between number of senses and ROM drops from $r\ (530) = -.39$, $p < .001$ to $r\ (386) = -.08$, $p = .111$.

DSI

The values of the DSI range from 1 (i.e., the two meanings of the word have the same number of associates) to 48 (i.e., the dominant meaning of the word has 48 times more associates than the subordinate meaning). As shown in Table 2, the mean DSI for the ambiguous words included in the database is $M = 8.28$. However, because the distribution of this variable is far enough from normality

Table 4. Descriptive statistics of the 252 unambiguous words by emotional categories

	n	Example	Val	Ent	Sen	NOM
Neg	66	*demencia* (dementia)	2.79 (0.76)	1.00 (0.00)	1.88 (1.00)	1.07 (0.08)
Neu	122	*milímetro* (millimeter)	5.22 (0.51)	1.03 (0.26)	2.52 (1.68)	1.10 (0.08)
Pos	64	*juerga* (revelry)	6.86 (0.64)	1.02 (0.13)	2.78 (2.97)	1.10 (0.09)

Note: The value indicated is the mean of all the stimuli in each condition and the standard deviations are in parentheses. Neg, negative. Neu, neutral. Pos, positive. Val, valence in isolation. Ent, number of dictionary entries. Sen, number of dictionary senses. NOM, number of meanings.

($p < .05$ in Kolmogorov–Smirnov test and visual inspection of histogram and Q-Q plots), the mean is not a good measure of central tendency in this case (see Field, 2017, Chapter 1). Consequently, the median and the other quartiles are chosen to describe that variable: 1.84 as first quartile, 3.95 as second quartile, and 11.95 as third quartile. Ambiguous words of the first rank (i.e., DSI $\leq$ 1.84) can be considered as being balanced in terms of ambiguity. Ambiguous words of the second (i.e., 1.84 $<$ DSI $\leq$ 3.95), the third (i.e., 3.95 $<$ DSI $\leq$ 11.95), and the fourth ranks (i.e., DSI $>$ 11.95) can be considered as being slightly, moderately, and extremely unbalanced, respectively. It is important to note that this categorization is arbitrary and is only used for descriptive purposes.

Valence measures

Emotional categorization of the words in the database

Following previous studies (e.g., Ferré et al., 2012), the 9-point valence dimension was divided to form three categories: negative (i.e., valence $<$ 4), neutral (i.e., 4 $\leq$ valence $\leq$ 6), and positive (i.e., valence $>$ 6). Concerning unambiguous words, three types are distinguished according to that division: negative, neutral, and positive words. Their descriptive statistics are represented in Table 4. Regarding ambiguous words, nine types are distinguished as a result of the combination of the three valence categories for each meaning: negative-negative, negative-neutral, negative-positive, neutral-negative, neutral-neutral, neutral-positive, positive-negative, positive-neutral, and positive-positive words (see Figure 1). The descriptive statistics of these words are represented in Table 5.

The distribution of unambiguous words in the distinct emotional categories reveals that most of the words belong to the neutral category (Table 4). The same pattern is observed if we look at the valence of the dominant meaning of ambiguous words (Table 5). However, if we consider both the dominant and the subordinate meaning and the distribution of ambiguous words into the nine emotional categories, some interesting patterns emerge. On the one hand, quite a lot of ambiguous words have meanings with different valence. In our classification, more than half the words (54.37%) have incongruent valence between their meanings, while the other words have congruent valence (i.e., negative-negative,

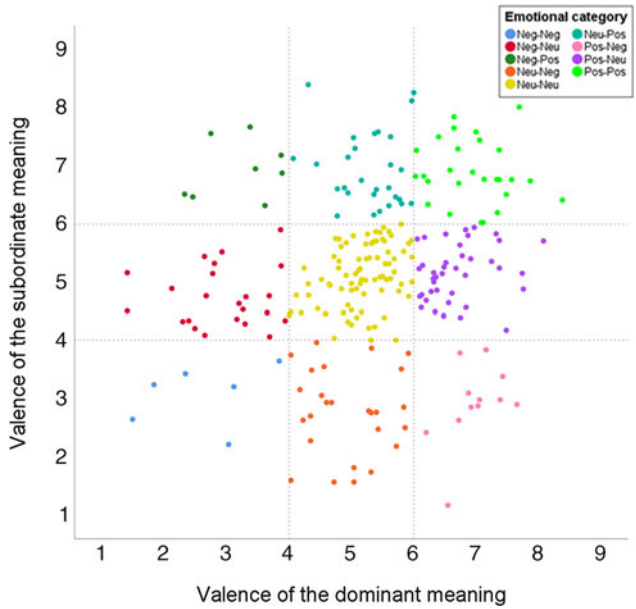

Figure 1. Distribution of ambiguous words in the affective space defined by the valences of the dominant meaning and the subordinate meaning. Emotional categories are based on the division of the 9-point valence dimension as follows: negative (i.e., valence < 4), neutral (i.e., $4 \leq$ valence ≤ 6), and positive (i.e., valence > 6).

neutral-neutral, positive-positive). On the other hand, it should be noted that among the 252 words, 115 of them have at least one positive meaning, while only 75 of them have at least one negative meaning. These results are consistent with a positive bias in language, as reported in several studies (see Dodds et al., 2015, for instance).

Prediction of valence in isolation

Four linear regression models were created to explore which of the following variables best predicted the values of valence in isolation: (a) the valence of the subordinate meaning alone, (b) the valence of the dominant meaning alone, (c) the arithmetic mean of the valences of both meanings, and (d) the weighted mean of the valences of both meanings in function of the number of associates (i.e., dominance). The results of these models are presented in Table 6.

As can be seen in Table 6, the four models significantly predicted the values of valence in isolation, although there were small differences between the last three models in the amount of explained variance (R^2). To examine if those differences were significant, the Pearson correlation coefficient of each model was compared to the others by using the web version of cocor (Diedenhofen & Musch, 2015). The results revealed that the Pearson correlation coefficient of Model 1 was significantly smaller than that of Model 2 (difference of 0.31, 95% confidence interval; CI [0.20, 0.42]), $z = 5.62$, $p < .001$, Model 3 (difference of 0.34, 95% CI [0.27, 0.41]), $z = 9.07$, $p < .001$, and Model 4 (difference of 0.38, 95% CI [0.29, 0.48]), $z = 7.98$,

Table 5. Descriptive statistics of the 252 ambiguous words by emotional categories

	n	Example	Val	Val-M1	Val-M2	Ent	Sen	NOM	ROM	DSI	IREM
Neg-neg	6	*Cardenal* (cardinal/bruise)	3.16 (0.50)	2.60 (0.88)	3.05 (0.54)	1.33 (0.52)	8.67 (6.35)	1.70 (0.14)	2.64 (0.86)	6.33 (10.64)	0.45 (0.89)
Neg-neu	23	*Mata* (kills/bush)	3.81 (0.93)	2.93 (0.72)	4.74 (0.51)	1.70 (0.77)	8.39 (4.94)	1.70 (0.17)	2.25 (0.69)	8.06 (9.35)	1.57 (0.82)
Neg-pos	8	*Éxtasis* (ecstasy/bliss)	4.57 (0.64)	3.21 (0.62)	6.95 (0.51)	1.38 (0.52)	6.50 (3.67)	1.74 (0.20)	2.40 (1.18)	9.77 (13.24)	–0.17 (0.88)
Neu-neg	26	*Terminal* (airport terminal/terminal illness)	4.37 (0.87)	4.93 (0.62)	2.76 (0.74)	1.58 (0.81)	8.38 (5.89)	1.76 (0.13)	2.65 (1.19)	6.49 (6.93)	–1.69 (0.87)
Neu-neu	82	*Gato* (cat/jack)	5.08 (0.65)	5.18 (0.48)	5.11 (0.52)	1.54 (0.67)	9.61 (6.17)	1.74 (0.14)	2.33 (0.99)	8.67 (9.87)	–0.01 (0.41)
Neu-pos	28	*Lista* (list/clever)	5.71 (0.75)	5.29 (0.51)	6.94 (0.65)	1.39 (0.57)	7.61 (5.14)	1.74 (0.16)	2.49 (1.35)	5.79 (6.87)	–1.44 (0.65)
Pos-neg	12	*Sangria* (sangria/bloodletting)	6.01 (0.90)	6.98 (0.40)	2.90 (0.70)	1.58 (0.79)	7.00 (4.31)	1.78 (0.16)	2.48 (0.86)	12.19 (15.04)	–0.13 (0.95)
Pos-neu	40	*Bonito* (beautiful/tuna)	5.91 (0.72)	6.69 (0.52)	5.15 (0.48)	1.68 (0.86)	8.60 (5.22)	1.70 (0.14)	2.44 (1.20)	10.07 (10.05)	1.27 (0.52)
Pos-pos	27	*bombón* (bonbon/hottie)	6.75 (0.67)	6.95 (0.59)	6.91 (0.55)	1.19 (0.40)	8.74 (6.10)	1.72 (0.16)	3.06 (1.15)	7.16 (6.17)	0.05 (0.86)

Note: The value indicated is the mean of all the stimuli in each condition and the standard deviations are in parentheses. Neg-neg, negative-negative. Neg-neu, negative-neutral. Neg-pos, negative-postive. Neu-neg, neutral-negative. Neu-neu, neutral-neutral. Neu-pos, neutral-positive. Pos-neg, positive-negative. Pos-neu, positive-neutral. Pos-pos, positive-positive. Val, valence in isolation, Val-M1, valence of the dominant meaning. Val-M2, valence of the subordinate meaning. Ent, number of dictionary entries. Sen, number of dictionary senses. NOM, number of meanings. ROM, relatedness of meanings. DSI, dominance-subordination index. IREM, index of the relative emotionality of meanings.

Table 6. Linear regression models for the prediction of valence in isolation

	R^2	*b*	95% CI	β	*p*	BF_{10}
Model 1	.22				<.001	2.28 e+12
Constant		3.34	[2.87, 3.80]		<.001	
Valence of the subordinate meaning		0.37	[0.29, 0.46]	.47	<.001	
Model 2	.61				<.001	6.33 e+48
Constant		1.80	[1.44, 2.16]		<.001	
Valence of the dominant meaning		0.65	[0.58, 0.71]	.78	<.001	
Model 3	.65				<.001	2.04 e+55
Constant		0.81	[0.39, 1.22]		<.001	
Arithmetic mean of both valences		0.85	[0.77, 0.93]	.81	<.001	
Model 4	.73				<.001	6.19 e+68
Constant		0.98	[0.65, 1.31]		<.001	
Weighted mean of both valences (by associates)		0.80	[0.74, 0.87]	.85	<.001	

Note: All values are rounded to the last decimal. In all models, $N = 252$.

$p < .001$. Focusing on the other three models, the Pearson correlation coefficient of Model 4 was significantly larger than that of both Model 2 (difference of 0.08, 95% CI [0.05, 0.11]), $z = 5.31$, $p < .001$, and Model 3 (difference of 0.05, 95% CI [0.02, 0.08]), $z = 2.87$, $p = .004$; meanwhile, no significant differences were observed between Models 2 and 3 (difference of 0.03, 95% CI [–0.02, 0.08]), $z = 1.21$, $p = .226$. Taking all this into account, it seems that both the valence of the dominant meaning alone and the arithmetic mean of the valences of both meanings explain quite well the valence of the word in isolation (60.60% and 65.08% of the variance, respectively), and they do it much better than the valence of the subordinated meaning alone (22.24% of the variance). However, the best predictor of the valence ratings in isolation is the weighted mean of the valence of both meanings considering the number of associates (explaining the 72.81% of the variance).

In order to illustrate the superior predictive capacity of the weighted valence of both meanings on valence values in isolation, the valences of the word in isolation, the dominant meaning, and the subordinate meaning of some ambiguous words are represented in Figure 2 (inspired by Figure 2b of Scott et al., 2019).

The above findings show that, in order to predict the values of valence in isolation, it is not enough to consider only the valence of the dominant meaning: both meanings should be considered, taking into account their dominance. Unfortunately, our analyses do not allow us to establish which of the following possibilities has been the responsible of this result: (a) participants have taken the two valences of the meanings of ambiguous words into account and made mental averages of them, or (b) there were two groups of participants with extreme opposite values (i.e., a group of participants focused on one meaning and another

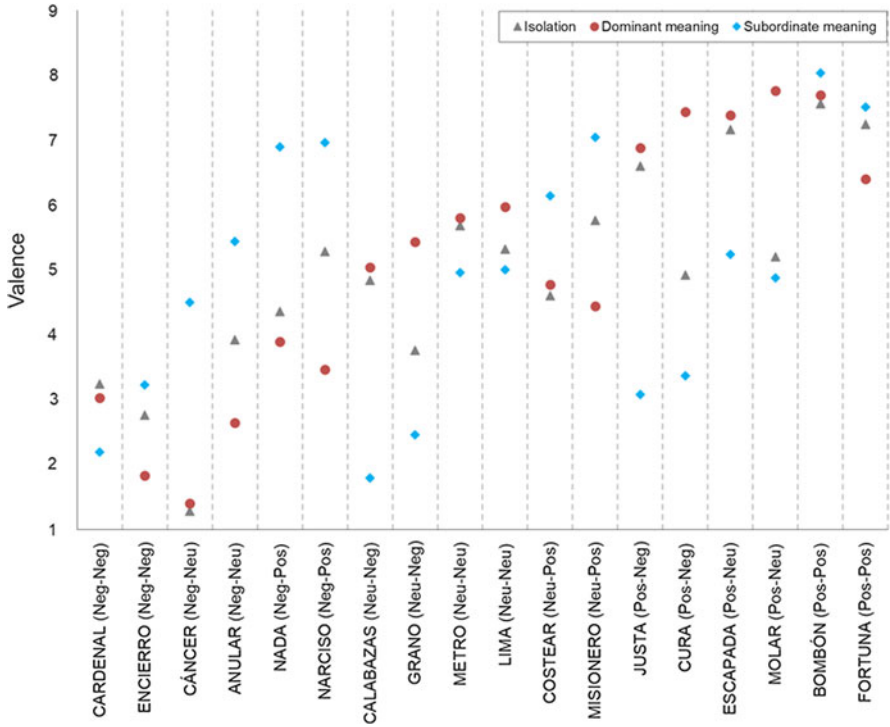

Figure 2. Distribution of the three valences (valence of the word in isolation, of the dominant meaning, and of the subordinate meaning) of some ambiguous words. Two example ambiguous words are presented for each category (e.g., *metro* and *lima* for the neutral-neutral category). In each category, the first example is a case where the valence of the word presented in isolation is very similar to the valence of the dominant meaning, while this is not the case for the second example.

group focused on the other meaning). The only way to examine this issue is to ask the participants directly at the end of the rating task what has been the strategy followed for each word. Such an approach may be very informative for our understanding of the processing of language ambiguity. If future studies reveal that people take both meanings into account (Possibility a), this would give support to the idea that the different meanings of ambiguous words are activated when they are presented in isolation (in line with the exhaustive access/search model; see, for instance, Cai & Vigliocco, 2018; Lin & Chen, 2015). Due to the characteristics of the task, we could not know if such activation was automatic or rather due to a conscious search process. Further research with other speeded tasks may shed light on this issue.

Limitations and conclusions

The present database includes meaning-dependent affective (valence) ratings for 252 Spanish ambiguous words. Before highlighting the main implications of this study, we would like to mention a possible limitation, which is related to the use

of definitions for the meaning-dependent valence rating task.[6] On the one hand, it may be that participants rely on the definition more than on the feeling produced by the word itself (especially for the subordinate meanings, which could be unknown by some of the speakers). On the other hand, the definitions themselves have an affective content, which can have affected ratings. As for the first issue, there is no way to know with the present design the number of cases where participants did not know the subordinate meaning. However, researchers interested in the study of emotion and ambiguity may use the dominance index provided in the database to avoid selecting highly unbalanced words for their experiments. In this way, they would be sure that the valence ratings for each meaning refer to something previously known by the speaker. As for the second issue, the effect of the affective content of definitions on word valence ratings cannot be avoided. A negative word has a negative meaning. Consequently, its definition has a negative valence. Another possibility would have been to use an associate to each meaning instead of its definition. However, most associates can be emotionally charged too. Another problem with this approach is that the type of associates and their associative strength can greatly differ between the two meanings of ambiguous words.

Apart from the above limitation, the present study has produced several relevant findings that have clear implications for research on both ambiguity and emotionality. First, the analysis carried out has shown that a high amount of ambiguous words have meanings with different valence. This fact is especially relevant for studies on affective processing, which include ambiguous words, as it could be the case that some of them have two meanings with different valence values, and this could affect results (e.g., a facilitative effect of positive valence on word processing might not be found if there are several words in the experimental set in which the content of both meanings is not positive). Second, many of the ambiguous words that are highly ambivalent in terms of the valence of their meanings are rated as neutral when presented in isolation. Thus, of the 20 words with meanings of opposite affective valence (i.e., negative-positive and positive-negative), 14 have a neutral range (4–6) in the isolated estimate. It is important to bear this in mind, as there is a risk that these words might be included as neutral in studies if only valence in isolation is considered. Third, and probably the most relevant contribution of the present work, the study demonstrated that the valence rating of an ambiguous word in isolation is better explained by a weighted average of the valence of its two meanings in terms of their dominance (i.e., number of lexical associates). It should be noted, however, that the valence of the dominant meaning alone predicts also quite well the valence of the isolated ambiguous words. Overall, the present findings have consequences both for basic studies (e.g., relationship between emotion and language, ambiguity processing, etc.) and applied work (e.g., cognitive and emotional biases in emotional disorders, second language learning, etc.).

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/S014271641900050X

Acknowledgments. This work was supported by Grants PSI2015-63525-P (MINECO/FEDER) and Ref 2017PFR-URV-B2-32 from the Universitat Rovira i Virgili.

Notes

1. Our aim was to develop a database of Spanish words that, regardless of the number of different meanings included in the dictionary, had only two subjective or psychologically different meanings for the general population. For example, the word *despecho* has two different meanings according to the *Diccionario de la Lengua Española* (RAE, 2014; http://dle.rae.es/): a first one related to a negative affective state, and a second one related to stop breastfeeding. However, after the subjective ambiguity variables were assessed, *despecho* had a NOM of 1.10 and no associates were generated for the second dictionary meaning. Hence, despite having two different meanings linguistically, the word *despecho* can be said to have only the psychological meaning of the negative affective state. For further discussion of the pros and cons of the dictionary and subjective approaches, see Fraga, Padrón, Perea, and Comesaña (2017), Haro and Ferré (2018), and Haro et al. (2017).

2. When there was a discrepancy between the two judges that could not be resolved and that involved 3 or fewer associates, the average number of associates was computed, and this was the value included in the database. For example, for the subordinate meaning of the Spanish word *super* (supermarket/super), one judge counted 10 associates while the other judge recounted 8: the reported value of associates was the average of these two values, which is 9. Taking into account this averaging procedure, it is not surprising that there are words with decimal associates (e.g., the reported value of associates for the dominant meaning of the Spanish word *agarrado* [grasped/stingy] is 11.5, because one judge recounted 12 associates while the other judge recounted 11).

3. It has to be noted that five of the ambiguous words were wholly balanced (i.e., DSI = 1), and for descriptive purposes which was the first meaning and which one the second meaning was established at random as follows: *carga* (burden/charge), *celo* (heat/cellophane), *claustro* (senate/cloister), *genio* (gifted/magician), and *tira* (strip/throws). Consequently, the Spanish word *carga* (burden/charge) was emotionally categorized as negative-neutral, but it could have been categorized as neutral-negative too.

4. Caution is required when using the IREM: this index provides information on the relative emotionality of meanings, but not about whether the first meaning is more or less positive or negative than the second.

5. For an introduction to the theory and practice of Bayesian analysis, see the papers in the Special Issue edited by Vandekerckhove, Rouder, & Kruschke (2018). Specifically, we recommend Wagenmakers, Marsman, et al. (2018) in order to understand the rationale behind using Bayesian analyses. We also recommend Wagenmakers, Love, et al., (2018) for a practical introduction to Bayesian analyses using JASP (and their Table 1 for a guide to the interpretation of Bayes factors BF_{10}, p. 67).

6. We thank the reviewer Constance Imbault for this suggestion.

References

Adriaens, G., Small, S. L., Cottrell, G. W., & Tanenhaus, M. K. (1988). *Lexical ambiguity resolution: Perspectives from psycholinguistics, neuropsychology, and artificial intelligence*. San Mateo, CA: Kaufmann.

Altarriba, J., & Basnight-Brown, D. M. (2011). The representation of emotion vs. emotion-laden words in English and Spanish in the Affective Simon Task. *International Journal of Bilingualism*, **15**, 310–328. doi: 10.1177/1367006910379261

Angwin, A. J., Dissanayaka, N. N., McMahon, K. L., Silburn, P. A., & Copland, D. A. (2017). Lexical ambiguity resolution during sentence processing in Parkinson's disease: An event-related potential study. *PLOS ONE*, **12**, 1–14. doi: 10.1371/journal.pone.0176281

Armstrong, B. C., & Plaut, D. C. (2008). Settling dynamics in distributed networks explain task differences in semantic ambiguity effects: Computational and behavioral evidence. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 273–278). Austin, TX: Cognitive Science Society.

Armstrong, B. C., Tokowicz, N., & Plaut, D. C. (2012). eDom: Norming software and relative meaning frequencies for 544 English homonyms. *Behavior Research Methods*, **44**, 1015–1027. doi: 10.3758/s13428-012-0199-8

Armstrong, B. C., Zugarramurdi, C., Cabana, Á., Lisboa, J. V., & Plaut, D. C. (2016). Relative meaning frequencies for 578 homonyms in two Spanish dialects: A cross-linguistic extension of the English eDom norms. *Behavior Research Methods*, **48**, 950–962. doi: 10.3758/s13428-015-0639-3

695 **Balota, D. A., Ferraro, F. R., & Connor, L. T.** (1991). On the early influence of meaning in word
696 recognition: A review of the literature. In P. J. Schwanenflugel (Ed.), *The psychology of word meanings*
697 (pp. 187–222). Hillsdale, NJ: Erlbaum.
698 **Bird, H., Franklin, S., & Howard, D.** (2001). Age of acquisition and imageability ratings for a large set of
699 words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers*, **33**,
700 73–79. doi: 10.3758/BF03195349
701 **Bracken, J., Degani, T., Eddington, C., & Tokowicz, N.** (2017). Translation semantic variability: How
702 semantic relatedness affects learning of translation-ambiguous words. *Bilingualism: Language and*
703 *Cognition*, **20**, 783–794. doi: 10.1017/S1366728916000274
704 **Bradley, M. M., & Lang, P. J.** (1999). *Affective Norms for English Words (ANEW): Instruction manual and*
705 *affective ratings (Technical report C-1)*. Gainesville, FL: University of Florida, Center for Research in
706 Psychophysiology.
707 **Britton, B. K.** (1978). Lexical ambiguity of words used in English text. *Behavior Research Methods &*
708 *Instrumentation*, **10**, 1–7. doi: 10.3758/BF03205079
709 **Brysbaert, M., Warriner, A. B., & Kuperman, V.** (2014). Concreteness ratings for 40 thousand
710 generally known English word lemmas. *Behavior Research Methods*, **46**, 904–911. doi: 10.3758/
711 s13428-013-0403-5
712 **Cai, Z. G., & Vigliocco, G.** (2018). Word processing. In S. L. Thompson-Schill & J. T. Wixted (Eds.),
713 *Stevens' handbook of experimental psychology and cognitive neuroscience: Vol. 3. Language & thought.*
714 Hoboken, NJ: Wiley.
715 **Citron, F. M., Weekes, B. S., & Ferstl, E. C.** (2014). Arousal and emotional valence interact in written word
716 recognition. *Language, Cognition and Neuroscience*, **29**, 1257–1267. doi: 10.1080/23273798.2014.897734
717 **Clifton, J.** (2015). *Training the interpretation of ambiguity: A thesis submitted to the Victoria University of*
718 *Wellington in fulfilment of the requirements for the degree of doctor of philosophy* (Unpublished doctoral
719 dissertation, Victoria University of Wellington).
720 **Davies, S. K., Izura, C., Socas, R., & Dominguez, A.** (2016). Age of acquisition and imageability norms for
721 base and morphologically complex words in English and in Spanish. *Behavior Research Methods*, **48**,
722 349–365. doi: 10.3758/s13428-015-0579-y
723 **Dearing, K. F., & Gotlib, I. H.** (2009). Interpretation of ambiguous information in girls at risk for
724 depression. *Journal of Abnormal Child Psychology*, **37**, 79–91. doi: 10.1007/s10802-008-9259-z
725 **Degani, T., & Tokowicz, N.** (2010). Ambiguous words are harder to learn. *Bilingualism: Language and*
726 *Cognition*, **13**, 299–314. doi: 10.1017/S1366728909990411
727 **Degani, T., Tseng, A. M., & Tokowicz, N.** (2014). Together or apart: Learning of translation-ambiguous
728 words. *Bilingualism: Language and Cognition*, **17**, 749–765. doi: 10.1017/S1366728913000837
729 **Diedenhofen, B., & Musch, J.** (2015). cocor: A comprehensive solution for the statistical comparison of
730 correlations. *PLOS ONE*, **10**, e0121945. doi: 10.1371/journal.pone.0121945
731 **Dodds, P. S., Clark, E. M., Desu, S., Frank, M. R., Reagan, A. J., Williams, J. R., . . . Megerdoomian, K.**
732 (2015). Human language reveals a universal positivity bias. *Proceedings of the National Academy of*
733 *Sciences*, **112**, 2389–2394. doi: 10.1073/pnas.1411678112
734 **Eddington, C. M., & Tokowicz, N.** (2015). How meaning similarity influences ambiguous word processing:
735 The current state of the literature. *Psychonomic Bulletin & Review*, **22**, 13–37. doi: 10.3758/s13423-014-
736 0665-7
737 **Eilola, T. M., & Havelka, J.** (2010). Affective norms for 210 British English and Finnish nouns. *Behavior*
738 *Research Methods*, **42**, 134–140. doi: 10.3758/BRM.42.1.134
739 **Fang, X., & Perfetti, C. A.** (2019). Learning new meanings for known words: Perturbation of original
740 meanings and retention of new meanings. *Memory & Cognition*, **47**, 130–144. doi: 10.3758/s13421-
741 018-0855-z
742 **Ferraro, F. R., & Kellas, G.** (1990). Normative data for number of word meanings. *Behavior Research*
743 *Methods, Instruments, & Computers*, **22**, 491–498. doi: 10.3758/BF03204432
744 **Ferré, P., Guasch, M., Moldovan, C., & Sánchez-Casas, R.** (2012). Affective norms for 380 Spanish words
745 belonging to three different semantic categories. *Behavior Research Methods*, **44**, 395–403. doi: 10.3758/
746 s13428-011-0165-x
747 **Ferré, P., Ventura, D., Comesaña, M., & Fraga, I.** (2015). The role of emotionality in the acquisition of new
748 concrete and abstract words. *Frontiers in Psychology*, **6**, 1–10. doi: 10.3389/fpsyg.2015.00976
749 **Field, A.** (2017). *Discovering statistics using IBM SPSS statistics* (5th ed.). Thousand Oaks, CA: Sage.

Fraga, I., Guasch, M., Haro, J., Padrón, I., & Ferré, P. (2018). EmoFinder: The meeting point for Spanish emotional words. *Behavior Research Methods*, **50**, 1–10. doi: 10.3758/s13428-017-1006-3

Fraga, I., Padrón, I., Perea, M., & Comesaña, M. (2017). I saw this somewhere else: The Spanish Ambiguous Words (SAW) database. *Lingua*, **185**, 1–10. doi: 10.1016/j.lingua.2016.07.002

Gawlick-Grendell, L. A., & Woltz, D. J. (1994). Meaning dominance norms for 120 homographs. *Behavior Research Methods, Instruments, & Computers*, **26**, 5–25. doi: 10.3758/BF03204557

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, **113**, 256–281. doi: 10.1037/0096-3445.113.2.256

Gilhooly, K. J., & Logie, R. H. (1980). Meaning-dependent ratings of imagery, age of acquisition, familiarity, and concreteness for 387 ambiguous words. *Behavior Research Methods & Instrumentation*, **12**, 428–450. doi: 10.3758/BF03201694

Goodwin, L. D., & Leech, N. L. (2006). Understanding correlation: Factors that affect the size of r. *Journal of Experimental Education*, **74**, 249–266. doi: 10.3200/JEXE.74.3.249-266

Gorfein, D. S., Viviani, J. M., & Leddo, J. (1982). Norms as a tool for the study of homography. *Memory & Cognition*, **10**, 503–509. doi: 10.3758/BF03197654

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

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

Haro, J., Comesaña, M., & Ferré, P. (2019). Is there an orthographic boost for ambiguous words during their processing? *Journal of Psycholinguistic Research*, **48**, 519–534. doi: 10.1007/s10936-018-9616-1

Haro, J., & Ferré, P. (2018). Semantic ambiguity: Do multiple meanings inhibit or facilitate word recognition? *Journal of Psycholinguistic Research*, **47**, 679–698. doi: 10.1007/s10936-017-9554-3

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

Hino, Y., Kusunose, Y., & Lupker, S. J. (2010). The relatedness-of-meaning effect for ambiguous words in lexical-decision tasks: When does relatedness matter? *Canadian Journal of Experimental Psychology*, **64**, 180–196. doi: 10.1037/a0020475

Hino, Y., & Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance*, **22**, 1331–1356. doi: 10.1037/0096-1523.22.6.1331

Hino, Y., Pexman, P. M., & Lupker, S. J. (2006). Ambiguity and relatedness effects in semantic tasks: Are they due to semantic coding? *Journal of Memory and Language*, **55**, 247–273. doi: 10.1016/j.jml.2006.04.001

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

Hinojosa, J. A., Moreno, E. M., & Ferré, P. (2019). Affective neurolinguistics: Towards a framework for reconciling language and emotion. *Language, Cognition and Neuroscience*. Advance online publication. doi: 10.1080/23273798.2019.1620957

Hinojosa, J. A., Rincón-Pérez, I., Romero-Ferreiro, M. V., Martínez-García, N., Villalba-García, C., Montoro, P. R., & Pozo, M. A. (2016). The Madrid Affective Database for Spanish (MADS): Ratings of dominance, familiarity, subjective age of acquisition and sensory experience. *PLOS ONE*, **11**, e0155866. doi: 10.1371/journal.pone.0155866

Jager, B., & Cleland, A. A. (2016). Polysemy advantage with abstract but not concrete words. *Journal of Psycholinguistic Research*, **45**, 143–156. doi: 10.1007/s10936-014-9337-z

Jastrzembski, J. E. (1981). Multiple meanings, number of related meanings, frequency of occurrence, and the lexicon. *Cognitive Psychology*, **13**, 278–305. doi: 10.1016/0010-0285(81)90011-6

Kaplan, J. J., Fisher, D. G., & Rogness, N. T. (2009). Lexical ambiguity in statistics: What do students know about the words association, average, confidence, random and spread? *Journal of Statistics Education*, **17**. doi: 10.1080/10691898.2009.11889535

Kellas, G., Ferraro, F. R., & Simpson, G. B. (1988). Lexical ambiguity and the timecourse of attentional allocation in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, **14**, 601–609. doi: 10.1037/0096-1523.14.4.601

Khanna, M. M., & Cortese, M. J. (2011). Age of acquisition estimates for 1,208 ambiguous and polysemous words. *Behavior Research Methods*, **43**, 89–96. doi: 10.3758/s13428-010-0027-y

Kissler, J., & Koessler, S. (2011). Emotionally positive stimuli facilitate lexical decisions—An ERP study. *Biological Psychology*, **86**, 254–264. doi: 10.1016/j.biopsycho.2010.12.006

Klepousniotou, E. (2002). The processing of lexical ambiguity: Homonymy and polysemy in the mental lexicon. *Brain and Language*, **81**, 205–223. doi: 10.1006/brln.2001.2518

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

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

Kuperman, V. (2015). Virtual experiments in megastudies: A case study of language and emotion. *Quarterly Journal of Experimental Psychology*, **68**, 1693–1710. doi: 10.1080/17470218.2014.989865

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. doi: 10.1037/a0035669

Larsen, R. J., Mercer, K. A., Balota, D. A., & Strube, M. J. (2008). Not all negative words slow down lexical decision and naming speed: Importance of word arousal. *Emotion*, **8**, 445–452. doi: 10.1037/1528-3542.8.4.445

Lin, C. J. C., & Chen, Y. R. (2015). Exhaustive semantic activation for reading ambiguous verbs in Chinese sentences. *Lingua Sinica*, **1**, 1–23. doi: 10.1186/s40655-015-0008-2

Mogg, K., Bradbury, K. E., & Bradley, B. P. (2006). Interpretation of ambiguous information in clinical depression. *Behaviour Research and Therapy*, **44**, 1411–1419. doi: 10.1016/j.brat.2005.10.008

Moors, A., De Houwer, J., Hermans, D., Wanmaker, S., Van Schie, K., Van Harmelen, A. L., . . . 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

Nelson, D. L., McEvoy, C. L., Walling, J. R., & Wheeler, J. W. (1980). The University of South Florida homograph norms. *Behavior Research Methods & Instrumentation*, **12**, 16–37. doi: 10.3758/BF03208320

Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory & Cognition*, **9**, 225–236. doi: 10.3758/BF03196957

Pecher, D. (2001). Perception is a two-way junction: Feedback semantics in word recognition. *Psychonomic Bulletin & Review*, **8**, 545–551. doi: 10.3758/BF03196190

Pexman, P. M., Hino, Y., & Lupker, S. J. (2004). Semantic ambiguity and the process of generating meaning from print. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **30**, 1252–1270. doi: 10.1037/0278-7393.30.6.1252

Real Academia Española. (2014). *23ª edición del Diccionario de la lengua española*. Madrid: Author.

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

Rodd, J. M., Gaskell, M. G., & Marslen-Wilson, W. D. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language*, **46**, 245–266. doi: 10.1006/jmla.2001.2810

Rodríguez-Ferreiro, J., & Davies, R. (2019). The graded effect of valence on word recognition in Spanish. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **45**, 851–868. doi: 10.1037/xlm0000616

Schane, S. (2002). Ambiguity and misunderstanding in the law. *Thomas Jefferson Law Review*, **25**, 167–193.

Scott, G. G., Keitel, A., Becirspahic, M., Yao, B., & Sereno, S. C. (2019). The Glasgow Norms: Ratings of 5,500 words on nine scales. *Behavior Research Methods*, **51**, 1258–1270. doi: 10.3758/s13428-018-1099-3

Simpson, G. B. (1984). Lexical ambiguity and its role in models of word recognition. *Psychological Bulletin*, **96**, 316–340. doi: 10.1037/0033-2909.96.2.316

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

Stadthagen-Gonzalez, H., Imbault, C., Pérez-Sánchez, M. A., & Brysbaert, M. (2017). Norms of valence and arousal for 14,031 Spanish words. *Behavior Research Methods*, **49**, 111–123. doi: 10.3758/s13428-015-0700-2

Stevenson, M., & Guo, Y. (2010). Disambiguation in the biomedical domain: The role of ambiguity type. *Journal of Biomedical Informatics*, **43**, 972–981. doi: 10.1016/j.jbi.2010.08.009

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re) consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, **18**, 645–659. doi: 10.1016/S0022-5371(79)90355-4

Syssau, A., & Laxén, J. (2012). The influence of semantic richness on the visual recognition of emotional words. *Canadian Journal of Experimental Psychology*, **66**, 70–78. doi: 10.1037/a0027083

Taler, V., Klepousniotou, E., & Phillips, N. A. (2009). Comprehension of lexical ambiguity in healthy aging, mild cognitive impairment, and mild Alzheimer's disease. *Neuropsychologia*, **47**, 1332–1343. doi: 10.1016/j.neuropsychologia.2009.01.028

Tokowicz, N., & Kroll, J. F. (2007). Number of meanings and concreteness: Consequences of ambiguity within and across languages. *Language and Cognitive Processes*, **22**, 727–779. doi: 10.1080/01690960601057068

Vandekerckhove, J., Rouder, J. N., & Kruschke, J. K. (Eds.) (2018). Bayesian methods for advancing psychological science [Special Issue]. *Psychonomic Bulletin & Review*, **25**.

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

Võ, 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

Wagenmakers, E J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, A. J., . . . Morey, R. D. (2018). Bayesian statistical inference for psychological science: Part II. Example applications with JASP. *Psychonomic Bulletin & Review*, **25**, 58–76. doi: 10.3758/s13423-017-1323-7

Wagenmakers, E. J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., . . . Matzke, D. (2018). Bayesian inference for psychology: Part I. Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, **25**, 35–57. doi: 10.3758/s13423-017-1343-3

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

Appendix A. Instructions used for valence ratings of words with the definition of one of the two meanings

Spanish (original)

Page 1

A continuación se te presenta un conjunto de palabras que tendrás que evaluar mediante el Maniquí de Auto-Evaluación (MAE).

Observa que el MAE dispone de una escala que va desde 1 hasta 9. Utiliza estos valores para evaluar cada palabra en la dimensión AGRADO de acuerdo con los siguientes criterios:

- Si la palabra te hace sentir completamente triste, la evaluarás con un 1.
- Si te hace sentir completamente alegre, la evaluarás con un 9.
- Si no te hace sentir ni alegre ni triste, sino neutral, la evaluarás con un 5.

Puedes utilizar otros valores si la palabra te hace sentir un poco triste (3) o un poco alegre (7). Observa que también puedes evaluar tu nivel de alegría o tristeza utilizando otros valores (2, 4, 6, 8) situados entre las figuras.

Page 2

Algunas palabras pueden hacer referencia a más de un concepto. Por ello, junto a cada palabra se presenta una definición para acotar con cuál de sus significados debes trabajar. Debes valorar el agrado de la palabra solo en relación al concepto/significado que se indica en la definición. Imagina el siguiente ejemplo con la palabra caña:

> CAÑA – Instrumento utilizado para pescar que lleva en el extremo más delgado una cuerda de la que pende el sedal con el anzuelo.

La mayoría de la gente evaluaría esta palabra en valores medios de la escala, ya que no es ni agradable ni desagradable, sino más bien neutra. Otros significados posibles para esta palabra serían "caña de cerveza," "caña de azúcar," o "caña de chocolate." No debes tener en cuenta estos significados para responder: debes evaluar el agrado de la palabra solo pensando en el concepto/significado de la definición.

English (translated, example changed for illustrative purposes: We did not use English ambiguous words in any questionnaire)

Page 1

You will be presented with a set of words. You are asked to rate them using the Self-Assessment Manikin (SAM).

Note that the SAM has a scale that ranges from 1 to 9. Use these values to rate each word in the PLEASANTNESS dimension according to the following criteria:

- If the word makes you feel completely sad, then you should rate it with a 1.
- If the word makes you feel completely happy, then you should rate it with a 9.
- If it makes you feel neither happy nor sad, but neutral, you should rate it with a 5.

You can use other values if the word makes you feel a little sad (3) or a little happy (7). Note that you can also rate your level of happiness or sadness by using other values (2, 4, 6, 8) between these figures.

Page 2

Note that some words can refer to more than one concept. For this reason, a definition is presented together with each word to indicate which of its meanings you should rate. You have to evaluate the word's pleasantness only in relation to the meaning that is indicated by the definition. Imagine the following example with the word "bar":

BAR – Long straight piece of metal or other materials.

Most people would rate this word on the values in the middle of the scale, as it is neither pleasant nor unpleasant, but rather neutral. Other possible meanings for this word would be "pub" and "musical compass." You do not have to take these meanings into account in order to answer: you have to rate the word's pleasantness only in thinking about the meaning/concept indicated by the definition.

Cite this article: Huete-Pérez, D., Haro, J., Fraga, I., and Ferré, P. HEROÍNA: Drug or hero? Meaning-dependent valence norms for ambiguous Spanish words. *Applied Psycholinguistics*. https://doi.org/10.1017/S014271641900050X