



Love me in L1, but hate me in L2: How native speakers and bilinguals rate the affectivity of words when feeling or thinking about them

Pilar Ferré¹ , Marc Guasch¹, Hans Stadthagen-Gonzalez²
and Montserrat Comesaña^{3,4} 

Research Article

Cite this article: Ferré P, Guasch M, Stadthagen-Gonzalez H, Comesaña M (2022). Love me in L1, but hate me in L2: How native speakers and bilinguals rate the affectivity of words when feeling or thinking about them. *Bilingualism: Language and Cognition* 1–15. <https://doi.org/10.1017/S1366728922000189>

Received: 26 July 2021
Revised: 17 February 2022
Accepted: 28 February 2022

Address for correspondence:
Pilar Ferré, Departament de Psicologia, URV.
Ctra. de Valls, s/n, 43007 – Tarragona (Spain)
Email: mariaadelpilar.ferre@urv.cat

¹Universitat Rovira i Virgili, Department of Psychology and CRAMC, Tarragona, Spain; ²School of Psychology, The University of Southern Mississippi, Long Beach, MS, USA; ³Research Unit in Human Cognition, CIPsi, School of Psychology, University of Minho, Braga, Portugal and ⁴Centro de Investigación Nebrija en Cognición (CINC), Nebrija University, Madrid, Spain

Abstract

This study examines the distinction between knowing the meaning of a word and experiencing the feelings associated with it. We collected affective ratings for a set of emotional and neutral English words from a group of English native speakers and a group of European Portuguese–English bilinguals. Half of the emotional words named emotions (emotion words) and the other half did not name emotions but could provoke them (emotion-laden words). Some participants were asked to focus on the meaning of words while others were asked to focus on the feeling produced by the words. Native speakers of English produced more intense affective ratings than Portuguese–English bilinguals. Such difference was larger when participants focused on their feelings than when they focused on the words' meaning. Accordingly, such distinction should be considered in the study of bilingual affective language processing. Finally, the type of emotional word (emotion vs. emotion-laden) had only modest effects.

Introduction

The processing of affective language in bilinguals has become a topic of interest in psycholinguistics during the last decade (for reviews, see Caldwell-Harris, 2014; Jończyk, 2016; Pavlenko, 2012). The main objective in that line of research has been to elucidate if the native language (L1) has a stronger emotional resonance than the non-native language (L2).

Different methodological approaches have been followed in the attempt to answer that question: i) introspective studies that assess bilinguals' perception of emotionality in their respective languages (e.g., Dewaele, 2004, 2008; Dewaele, Lorette, Rolland & Mavrou, 2021); ii) cognitively oriented studies where the effects of the emotional content of words (defined in relation to the dimensions of valence-ranging from pleasant to unpleasant, and arousal-ranging from exciting to calming) in the two languages is examined with different experimental paradigms, such as the lexical decision task (e.g., Ferré, Anglada-Tort & Guasch, 2018; Ponari, Rodríguez-Cuadrado, Vinson, Fox, Costa & Vigliocco, 2015), the emotional Stroop task (e.g., Eilola & Havelka, 2011; Sutton, Altarriba, Gianico & Basnight-Brown, 2007), the affective priming paradigm (e.g., Degner, Doycheva & Wentura, 2012) or memory tasks (e.g., Ayçiçeği-Dinn & Caldwell-Harris, 2009; Ferré, García, Fraga, Sánchez-Casas & Molero, 2010); iii) psychophysiological studies, where physiological markers of arousal (mostly the skin conductance response, SCR) are recorded when words that differ in emotional content (including, for instance, swearwords or taboo words) are presented in L1 and L2 (e.g., Baumeister, Foroni, Conrad, Rumiati & Winkielman, 2017; Caldwell-Harris, Tong, Lung & Poo, 2011; Eilola & Havelka, 2011; see also Iacozza, Costa & Duñabeitia, 2017 and Toivo & Scheepers, 2019, for studies recording the pupillary response); iv) and electrophysiological and neuroimaging studies, where the time course of emotional processing is recorded (e.g., Conrad, Recio & Jacobs, 2011; Jończyk, Boutonnet, Musiał, Hoemann & Thierry, 2016; Wu & Thierry, 2012), or the neural areas involved in such processing are identified (e.g., Hsu, Jacobs & Conrad, 2015) while bilingual speakers are presented with L1 and L2 words.

Overall, previous research in this area has documented a reduced effect of emotional content when bilinguals process linguistic information in their L2. Concretely, the effects of affective content are either observed only in L1 (e.g., Degner et al., 2012) or, more commonly, they are also found in the L2, but with a smaller magnitude (e.g., Baumeister et al., 2017; Colbeck & Bowers, 2012; García-Palacios, Costa, Castilla, Del Río, Casaponsa & Duñabeitia, 2018; Ivaz, Costa & Duñabeitia, 2016; Toivo & Scheepers, 2019), or the effects appear later than in L1 (Conrad et al., 2011; Opitz & Degner, 2012), with some studies indicating that this reduced

emotional effect is especially true for negative content (e.g., Jończyk et al., 2016; Wu & Thierry, 2012). There are, however, other studies that report an effect of the same magnitude in both languages (e.g., Eilola, Havelka & Sharma, 2007; Eilola & Havelka, 2011; Ferré et al., 2010; Ferré, Sánchez-Casas & Fraga, 2013; Ponari et al., 2015), and even others that have found a larger emotional effect in L2 than in L1 (Ayçiçeği-Dinn & Caldwell-Harris, 2009; Caldwell-Harris et al., 2011). Such discrepancies might be explained by methodological differences between the studies, such as the type of bilinguals tested, or variables related with the L2 like age of acquisition, proficiency level or the type of acquisition context (Caldwell-Harris, 2014; Pavlenko, 2012). This last variable is very relevant if one considers that the acquisition of the L1 is closely linked to sensory and affective experiences while the acquisition of the L2 commonly takes place in a classroom environment, devoid of such experiences. When L2 words are learned exclusively in an academic setting, they are often disembodied (Pavlenko, 2012): although the speakers know that the meaning of those L2 words is emotional, they may not experience an affective reaction in response to them (i.e., bilinguals KNOW the words, but do not FEEL them).

The emotional (positive and negative) words and neutral words included in previous research on this area have been commonly selected from normative studies, where native speakers rate a large set of words in terms of their affective properties, most often valence and arousal. Subjective norms are available for a variety of languages, such as English (e.g., Warriner, Kuperman & Brysbaert, 2013); Spanish (e.g., Guasch, Ferré & Fraga, 2016; Stadthagen-Gonzalez, Imbault, Perez Sanchez & Brysbaert, 2016); European Portuguese -EP- (e.g., Soares, Comesaña, Pinheiro, Simões & Frade, 2012); French (e.g., Monnier & Syssau, 2014); German (e.g., Vö, Conrad, Kuchinke, Urton, Hofmann & Jacobs, 2009); Polish (e.g., Imbir, 2015); Croatian (e.g., Čoso, Guasch, Ferré & Hinojosa, 2019); Finnish (e.g., Eilola & Havelka, 2010); Italian (e.g., Montefinese, Ambrosini, Fairfield & Mammarella, 2014); Dutch (e.g., Moors, De Houwer, Hermans, Wanmaker, Van Schie, Van Harmelen, De Schryver, De Winne & Brysbaert, 2013); or Chinese (e.g., Lin & Yao, 2016). When bilingual experiments are designed, those words are commonly translated to the other language involved in the study and their valence and arousal levels are assumed to be the same in both languages. In order to know if this assumption is correct, it is necessary to examine if the subjective perception of the affective properties of words varies between the L1 and the L2. This is the main goal of the present study, where we use, for the first time, an approach that allows us to distinguish between the meaning of the words and the feelings they produce. This approach also enables us to test Pavlenko's (2012) proposal, according to which bilinguals might know the meaning of L2 emotional words, but not feel them.

Only a few studies to date have compared L1 and L2 affective ratings. Some of them included such ratings as a complement of an experimental task; for instance, Harris (2004) recorded the SCR of Spanish-English bilinguals while they rated the unpleasantness of a set of L1 and L2 words and expressions (taboo words, neutral words, sexual words, childhood reprimands, endearments and insults). There were differences between L1 and L2 in the SCR and in the unpleasantness ratings elicited by only one type of expressions: namely, childhood reprimands. Similar results regarding affective ratings were obtained by Caldwell-Harris et al. (2011) with a group of Mandarin-English bilinguals, who were asked to rate the emotional intensity of the

same types of words as in Harris (2004). Of note, ratings of emotional intensity of endearments were similar in Mandarin and English, despite the fact that SCR was larger for the former. Winskel (2013), in turn, asked a group of Thai-English bilinguals to provide pleasantness (valence) ratings for a small set of Thai and English words (20 negative and 20 neutral words) and found no difference between L1 and L2 ratings. Interestingly, the same participants performed an emotional Stroop task that yielded an emotional effect but restricted to L1 words. The studies by Caldwell-Harris et al. (2011) and Winskel (2013) suggest that self-reported affective ratings do not always match with other behavioral or physiological measures.

Other studies have focused on affectivity ratings per se, without including any additional task. For instance, Vélez-Urbe and Rosselli (2019) collected valence ratings for positive, negative and neutral words, as well as for taboo words. Spanish-English bilinguals performed the ratings in both languages. The results revealed that valence ratings in the more dominant language (English) were more positive for positive and taboo words and more negative for negative words when compared to valence ratings in the less dominant language (Spanish). In a similarly oriented study, Garrido and Prada (2021) collected ratings for positive, negative and neutral words, as well as for taboo words from European Portuguese (EP)-English bilinguals. The dimensions rated in this study were valence, emotional intensity, and familiarity. The results showed that positive words were rated as more positive in L1 (EP) than in L2 (English), negative words were rated as more negative in L1 than in L2, and taboo words were rated as more negative and more emotionally intense in L1 than in L2.

Harris (2004), Caldwell-Harris et al. (2011), Winskel (2013), Vélez-Urbe and Rosselli (2019), and Garrido and Prada (2021) all used a within-participants design; the same bilinguals rated the affective properties of words in both languages. This is the most suitable design for reducing the effects of individual differences, but it may be that presenting words in the two languages also affects the participants' ratings. Concretely, ratings may have been influenced by bilinguals' belief about differences in emotional intensity between the two languages (i.e., they may have produced lower affective ratings for L2 words than for L1 words because they consider that the non-native language is less emotional than the native language). This would increase the difference between L1 and L2 ratings. Conversely, due to the bilingual characteristics of the task, participants might be on a bilingual mode (Grosjean, 2010), which might enhance the activation of the L1 translation equivalents when performing the L2 ratings, and reduce the difference between L1 and L2 ratings.

In order to avoid these possible confounds, it would be better to ask bilinguals to perform ratings only in their L2, and compare them to ratings from native speakers for those same words. This was the approach followed by Imbault, Titone, Warriner, and Kuperman (2020), who collected valence ratings for a large set of English words from L2 speakers living in Canada and compared those ratings with those of native speakers (Warriner et al., 2013). In line with the results of Garrido and Prada (2021), L2 speakers produced attenuated ratings (i.e., less positive and negative values) in comparison to native speakers of English. Such reduction in emotionality in L2 was observed for both positive and negative words, but it was more salient for negative words. They also found that the differences in affective ratings between L1 and L2 words were smaller for highly frequent

words, and for bilinguals who had spent more time living in Canada and were more proficient in English.

The overall conclusion from the reviewed rating studies seems to be that L2 words are perceived as less emotionally intense than L1 words. No conclusions can be drawn, however, regarding the proposal of Pavlenko (2012), who suggested that bilinguals might know the emotional meaning of L2 words, but not feel it. The reason for this is that the criterion used by participants to do their ratings is not clear. Focusing on valence, which is the variable of interest here, and depending on the instructions used, participants may rely either on their semantic knowledge about the affective content of the stimulus (i.e., “murder” refers to something bad), or on the affective response elicited by the stimulus (i.e., the word “murder” evokes a negative feeling). Even a mix of both approaches is possible. The studies reviewed previously did not use identical instructions. Imbault et al. (2020) and Vélez-Urbe and Rosselli (2019) instructed participants to rate the feeling the words produced in them. Caldwell-Harris et al. (2011) asked participants to think of a real situation where the words/expressions were used and to rate how emotional that use was for them. Garrido and Prada (2021), in turn, instructed participants to rate the extent to which they thought the word generated positive or negative feelings. Finally, participants in Harris (2004) and Winskel’s (2013) studies were asked to rate the words for unpleasantness/pleasantness.

Pavlenko’s proposal (2012) may be tested by instructing participants explicitly, when they perform their ratings, to focus either on the meaning of words or on the feelings they evoke. Although some previous studies have asked participants to focus on feelings (Caldwell-Harris et al., 2011; Imbault et al., 2020; Vélez-Urbe & Rosselli, 2019), the type of instructions used in others (Garrido & Prada, 2021; Harris, 2004; Winskel, 2013) does not allow us to know if participants relied on meanings or on feelings. A systematic study comparing these two types of instructions has never been carried out with words. This approach, which has been recently used with pictures (e.g., Itkes, Kimchi, Haj-Ali, Shapiro & Kron, 2017), is the one followed in the present study.

Itkes and Kron (2019) recently pointed out the confusion that can be created by using the term *VALENCE* to refer to two different things. Researchers have used *VALENCE* to refer to both the emotional response produced by the stimuli (affective valence) as well as the semantic knowledge about their affective content (semantic valence). However, they are not the same. Affective response refers to a physiological (i.e., autonomous nervous system activity), body (facial expression and bodily posture), and subjective (feelings) change in response to a particular event (Itkes & Kron, 2019). Semantic valence, by contrast, refers to general conceptual knowledge about the affective properties of objects or events (e.g., prison or violence refer to negative things). Ratings collected in normative studies may reflect either the affective response or the semantic valence, depending on how participants understand the task. If they understand that the task is about how positive or negative the stimulus is, they will probably rely mostly on their semantic knowledge. By contrast, when they understand that they have to report their feelings, it is more likely that their ratings will be based on the feelings provoked by the stimulus. Simply asking participants about stimulus or feelings may not be sufficient to make clear the difference between them. For this reason, Itkes and Kron (2019) developed a set of instruction protocols where, after explaining the difference between affective valence and semantic valence to the participants, they were

asked to focus on one of these two aspects to perform their ratings. They also provided evidence of the validity of these ratings. In particular, they found that ratings obtained with feeling-focused instructions predicted better the physiological response of participants (facial electromyography, heart rate, and electrodermal change) than those obtained with knowledge-focused instructions (Hamzani, Mazar, Itkes, Petranker & Kron, 2020). Furthermore, whereas affective valence ratings were attenuated with the repeated presentation of the stimuli (i.e., there was habituation), semantic ratings for valence were not attenuated (Itkes et al., 2017). These findings reveal that affective and semantic valence can be dissociated empirically and that, when participants are instructed to focus on one of these two aspects, they can do so.

In this study, we adopted the procedure of Itkes and colleagues with the aim of examining the impact of language in the evaluation of the affective and semantic valence of words. To that end, we collected ratings for a set of English words from native speakers of English as well as from EP-English bilinguals. We asked participants to provide positivity and negativity ratings for those words, either focusing on the feelings the words produced to them or on their meaning. Considering the results of previous rating studies, we expected differences between native speakers of English and bilinguals, and, in particular, we expected valence ratings to be more extreme for native speakers. Furthermore, in accordance with Pavlenko (2012), we expected these differences to be larger for feeling-focused ratings than for knowledge-focused ratings. Considering that, in some cases, experimental differences in emotional effects between L1 and L2 have been restricted to negative words (Jończyk et al., 2016; Wu & Thierry, 2012), we expected such differences in ratings to be more pronounced for negative words than for positive words.

A second goal of our study was to explore whether, if there were indeed a difference, it would depend on the type of emotional word involved. Emotional words are not a homogeneous set. There are emotion terms (e.g., *happiness* or *fear*) that refer directly to emotions (i.e., they denote an emotion) which we may call “emotion words”. Words can also provoke intense emotional reactions even though they do not denote an emotion (e.g., *murder*), by virtue of their connotative meaning. These may be called “emotion-laden words” (Pavlenko, 2008), which are perceived by the speakers as positive or negative (e.g., *party* or *murder*, respectively). Psycholinguistic research on emotional word processing on both native speakers and bilinguals has not commonly distinguished between these two types of words (see Hinojosa, Moreno & Ferré, 2020, for a review), although recent attempts have been made to explore this distinction (e.g., Martin & Altarriba, 2017; Wang, Shanguan & Lu, 2019). The few rating studies conducted in bilinguals have not distinguished between emotion words and emotion-laden words (Garrido & Prada, 2021; Imbault et al., 2020; Vélez-Urbe & Rosselli, 2019; Winskel, 2013). This can be, however, a relevant distinction for the purposes of this work. Emotion words directly refer to emotion terms, while emotion-laden words become affectively loaded probably through associative mechanisms (e.g., emotion-laden words could be associated with an emotion as a result of the direct experience of the speaker, and/or through his/her indirect exposure to information about the world). In consequence, the affective content of emotion-laden words may be more prone to individual differences and culture-specific effects than that of emotion words, depending more on the affective experiences linked to the acquisition and use of those words. If that is the case, the

differences between L1 and L2 ratings explored here may be larger for emotion-laden words than for emotion words. This is the first time that this issue is investigated.

Method

Participants

A total sample of 194 participants took part in this study. Among them, 117 participants were native speakers of English and 77 participants were EP-English bilinguals. The native speakers of English were students at the University of Southern Mississippi (United States of America). They were raised in the USA (with two exceptions who were raised in other English-speaking countries), had English as their mother tongue (two of them had an additional mother tongue), and were currently living in the USA. The bilingual participants were recruited at the Universidade do Minho (Portugal). They were all raised in Portugal (with four exceptions, raised in Brazil), had EP as their mother tongue, and were currently living in Portugal. Most of them (88.31%) learned English at school. All participants electronically signed an informed consent form before starting the experiment and received academic credit for their participation. The study was conducted with the approval of the Ethics Committee for Human Research of the University of Minho (Ref: CEICSH 052/2019) and of the University of Southern Mississippi's Institutional Review Board (Ref: IRB-20-383).

Participants from the two groups were randomly assigned to one of the instruction protocols: feeling-focused or knowledge-focused ratings; thus generating four groups of participants. A group of 62 native speakers of English (Mean age = 20.65 years, SD = 6.27, 83.87% females) were assigned to the knowledge-focused instructions, and another group of 55 native speakers of English (Mean age = 21.16 years, SD = 6.09, 80.00% females) were assigned to the feeling-focused instructions. Similarly, a group of 42 EP-English bilinguals (Mean age = 21.60 years, SD = 4.97, 85.71% females) were assigned to the knowledge-focused instructions, and another group of 35 EP-English bilinguals (Mean age = 21.86 years, SD = 6.43, 91.43% females) were assigned to the feeling-focused instructions.

All participants started the study by taking the LexTALE English test, developed by Lemhöfer and Broersma (2012; <http://www.lextale.com/takethetest.html>) to measure English proficiency. This test consists of a lexical decision task where participants have to indicate whether or not a string of letters is an existing English word. LexTALE provides a proficiency score that is interpreted as follows, in relation to the Common European Framework of Reference for Languages proficiency levels: participants with scores between 80 and 100 are considered advanced users (C1-C2), those with scores between 60 and 79 are upper intermediate users (B2), and those with scores of 59 and below are lower intermediate (B1) or basic users (A2), (Lemhöfer & Broersma, 2012, Table 9).

Participants also completed a language background questionnaire after entering their ratings. Table 1 shows the mean scores obtained by the four groups of participants in the LexTALE test as well as the data collected from the language background questionnaire. Participants estimated their age of acquisition of English in years, and their proficiency in English (listening, reading, speaking, and writing) in a 7-point scale (1 = *very poor* to 7 = *native-like*). They also estimated their daily use of English (EP-English bilinguals also estimated their use of EP).

Table 1 shows that native speakers of English acquired English early in life, and self-rated their proficiency in English as almost perfect. Importantly, the two groups of native speakers did not differ significantly in terms of age or any of the variables presented in Table 1 (all $ps > .136$). EP-English bilinguals acquired English during childhood, studied English for about ten years, their self-rated English proficiency in the four skills was above the middle point of the scale, and they scored in the upper intermediate range as measured by LexTALE. All bilingual participants declared using EP more than English in their daily life, but their percentage of use of English throughout the day was not negligible. Importantly, the two groups of bilinguals did not differ significantly in age or any of the aforementioned variables (all $ps > .150$).

Statistical comparisons revealed that native English speakers and EP-English bilinguals did not differ in age, $t(192) = 0.94$, $p = .347$, but, as intended, they differed in all the other variables examined (all $ps < .001$): native speakers obtained better scores in the LexTALE test, acquired English earlier, perceived themselves as more proficient, and used English to a greater extent than bilinguals.

Participants were also asked in the background questionnaire about their knowledge of other languages. 31.62% of the native speakers of English and 68.83% of the EP-English bilinguals indicated that they were familiar with other languages. However, their self-rated proficiency in those languages was low ($M = 2.79$ and $M = 3.19$ in a 1-to-7 scale for native speakers of English and EP-English bilinguals, respectively). None of the native speakers of English reported having knowledge of Portuguese.

Materials

The critical stimuli set consisted of 160 English words (see Appendix A), distributed as follows: half of them were emotional words and the other half were neutral words. The emotional words were divided into emotion words and emotion-laden words. Furthermore, half of them had a positive valence whereas the other half had a negative valence. Therefore, there were 20 positive emotion words (e.g., *love*), 20 negative emotion words (e.g., *envy*), 20 positive emotion-laden words (e.g., *dream*), 20 negative emotion-laden words (e.g., *exile*) and 80 neutral words (e.g., *notion*).

Words were selected from the English affective norms of Warriner et al. (2013). We first identified a potential set of emotion words in that dataset, and, to be sure that they were perceived as words referring to emotions, relied on the emotional prototypicality ratings of their translation equivalents in a Spanish database (Pérez-Sánchez, Stadthagen-Gonzalez, Guasch, Hinojosa, Fraga, Marín & Ferré, 2021). We used this dataset because it is the most recent and largest for this type of ratings. We focused on emotion words (nouns and adjectives) with prototypicality ratings higher than 3 (in a 5-point scale). From this pool, we selected 20 positive and 20 negative words, according to the valence values in the Warriner et al. (2013) database. A word was considered to be positive if it had a valence rating higher than 6, and it was considered to be negative if it had a valence rating lower than 4 (in a 9-point scale). Positive and negative emotion words differed in valence, $t(38) = 29.36$, $p < .001$, but they were matched in a set of affective and psycholinguistic variables: arousal (Warriner et al., 2013), concreteness (Brysbaert, Warriner & Kuperman, 2014), Zipf value (a measure of word frequency, Van Heuven, Mandera, Keuleers & Brysbaert, 2014), and length taken as number of letters (all $ps > .164$), (see Table 2). In addition, as one set of

Table 1. Mean and standard deviation (in parentheses) of the LexTALE score and of the language background questionnaire data in the four groups of participants

Group	Bilingualism	Native speakers		Bilinguals	
	Type of rating	Knowledge-focused	Feeling-focused	Knowledge-focused	Feeling-focused
LexTALE score (%)		85.08 (8.62)	86.82 (7.45)	72.40 (10.59)	71.64 (10.34)
Years studying English		-	-	10.05 (2.91)	10.20 (3.45)
Age of acquisition of English					
Listening		1.27 (1.61)	1.07 (1.16)	8.29 (3.12)	7.57 (2.98)
Reading		3.65 (1.71)	3.52 (1.86)	8.71 (2.24)	8.91 (2.61)
Speaking		1.24 (1.07)	1.46 (1.19)	8.86 (3.15)	8.66 (3.28)
Writing		4.02 (1.95)	3.74 (1.87)	9.10 (2.76)	9.11 (2.74)
Mean		2.54 (2.07)	2.45 (1.95)	8.74 (2.83)	8.56 (2.94)
Self-rated English proficiency (7-point scale)					
Listening		6.85 (0.57)	6.85 (0.45)	5.57 (0.91)	5.63 (1.00)
Reading		6.82 (0.46)	6.84 (0.54)	5.45 (0.97)	5.40 (1.19)
Speaking		6.90 (0.30)	6.91 (0.29)	5.14 (1.03)	4.77 (1.21)
Writing		6.82 (0.53)	6.89 (0.31)	4.86 (1.07)	4.66 (1.41)
Mean		6.85 (0.47)	6.87 (0.41)	5.26 (1.03)	5.11 (1.27)
Daily use (%)					
English		98.24 (7.88)	99.76 (0.84)	19.02 (11.97)	16.46 (10.86)
EP		-	-	79.24 (13.63)	80.54 (13.42)

participants consisted of EP–English bilinguals, the orthographic similarity between the English words and their Portuguese translations was also controlled. For this purpose, we relied on the Normalized Levenshtein Distance (NLD), calculated with NIM (Guasch, Boada, Ferré & Sánchez-Casas, 2013). Positive and negative emotion words were matched in NLD, $t(38) = 0.48$, $p = .635$. Finally, both conditions consisted of 75% nouns and 25% adjectives.

Words from the emotion-laden and neutral conditions were also taken from the norms of Warriner et al. (2013) and consisted of 75% nouns and 25% adjectives in each condition as well. The valence criteria used to select emotion-laden positive and negative words were the same as for the emotion words. Words were considered as neutral if they had a valence rating between 4 and 6 (see Table 2). As intended, emotion-laden positive and negative words differed in valence, $t(38) = 28.49$, $p < .001$, but not in the other variables (all $ps > .279$). Furthermore, positive emotion-laden words were matched in all the variables with positive emotion

words (all $ps > .120$), and negative emotion-laden words were matched in all the variables with negative emotion words (all $ps > .326$).

Finally, the set of 80 neutral words differed from emotional words (emotion words and emotion-laden words considered together) only in arousal, $t(158) = 9.18$, $p < .001$. There were no significant differences for the other variables (all $ps > .198$). The emotional words included both positive and negative words, so the average value of their valences was similar to that of neutral words. However, when comparing the valence of positive and negative words separately against the valence of neutral words, significant differences were observed in both comparisons (both $ps < .001$).

Procedure

The instructions for the ratings were adapted from Itkes et al. (2017; see Appendices B and C for the complete instructions).

Table 2. Mean and standard deviation (in parentheses) of the affective and psycholinguistic characteristics of the stimuli

Condition	Valence	Arousal	Concreteness	Zipf	Letters	NLD _{port}
Emotion_{positive}	7.54 (0.55)	5.10 (1.23)	2.01 (0.31)	4.15 (0.85)	7.85 (2.39)	0.41 (0.21)
Emotion_{negative}	2.75 (0.48)	5.06 (0.74)	2.14 (0.27)	3.90 (0.62)	7.40 (2.87)	0.38 (0.22)
Emotion-laden_{positive}	7.27 (0.52)	4.73 (0.66)	2.08 (0.33)	3.87 (0.65)	7.70 (2.11)	0.41 (0.20)
Emotion-laden_{negative}	2.90 (0.45)	4.94 (0.53)	2.14 (0.24)	3.77 (0.73)	8.05 (2.26)	0.41 (0.21)
Neutral	5.24 (0.55)	3.93 (0.56)	2.15 (0.29)	3.82 (0.60)	7.96 (2.25)	0.44 (0.20)

Note. Valence and arousal rated in a 9-point scale. Concreteness rated in a 5-point scale. NLD_{port} refers to the proportion of orthographic similarity between the English words and their Portuguese translation equivalents.

Participants were first presented with an explanation of the distinction between KNOWING and FEELING regarding words. Then, the particular instructions for each version of the ratings were provided. For the knowledge-focused version, participants were instructed to focus on word meaning to rate how positive/negative the word was. It was explained to participants that they should provide ratings based on the valence of the word, not on their internal feelings (see Appendix B). Conversely, in the feeling-focused version, participants were instructed to focus on the internal feelings evoked by the words in order to rate them in terms of positivity and negativity. It was explained to participants that they were to provide ratings based on their internal feelings, not on their knowledge of the affective meaning of the words (see Appendix C). All participants completed the questionnaire online.

The procedure was the same for all participant groups. First, participants read and electronically signed the informed consent form. Then, they were prompted to take the online version of the English LexTALE test (Lemhöfer & Broersma, 2012), and to report the score obtained. Each participant was then randomly assigned to one of the conditions and the corresponding set of instructions was presented on the screen (see Appendices B and C). The following screen explained the use of the scales to provide ratings. Following the procedure of Itkes et al. (2017), we used two sliding scales for each word in which participants could indicate values ranging from 0 to 10. The first one was a positivity scale, and the second one was a negativity scale. In the knowledge-focused version, 0 meant that the word had *no positive meaning*

at all or *no negative meaning at all*, while 10 meant that the word had *an extremely positive meaning* or *an extremely negative meaning* (for the positivity and negativity scales, respectively). In the feeling-focused version, 0 meant that the word elicited *no positive feeling at all* or *no negative feeling at all* while 10 meant that the word produced *an extremely positive feeling* or *an extremely negative feeling* (for the positivity and negativity scales, respectively). Participants could indicate that they didn't know the word.

After the instructions, the 160 words were presented in a different random order for each participant, one word per screen, together with the positive and the negative scales (see Figure 1).

In order to progress to the next item, participants had to provide a response on each scale or indicate that they did not know the meaning of the word. After rating all the words on the list, participants completed the language background questionnaire described in the Participants section.

Results

Three sets of analyses were performed: 1) an analysis that included all the words, where emotional words (including emotion words and emotion-laden words) were compared to neutral words; 2) an analysis restricted to emotional words, that focused on the comparisons between emotion words with emotion-laden words; and 3) an analysis restricted to emotional words, that focused on the comparison between positive and negative words. In all the cases, we carried out two analyses of variance

Fig. 1. Layout of the rating screen.

(ANOVA), one by participants and the other by items. Significant interactions were subjected to post-hoc Bonferroni tests. When needed, the effect size was calculated using Cohen’s *d* (Cohen, 1988). When the interactions were significant both by participants and by items (or when they were close to significance by participants, and significant by items), the statistical data provided with respect to Bonferroni tests (i.e., *p* and *d* values) correspond to the analysis by participants. In the few cases where the interactions reached statistical significance only in the analysis by items, the statistical data reported correspond to that analysis. Finally, all the means and standard deviations that appear in the tables, in the figures, as well as in the main text refer to the analyses by participants.

1- Analyses including all the words

We carried out two analyses of variance (ANOVA by participants and by items) on positivity ratings. The factors included in the analyses were emotionality (emotional words vs. neutral words), group (native speakers of English vs. EP-English bilinguals) and type of rating (knowledge-focused vs. feeling-focused: *know* vs. *feel* hereafter). Emotionality was treated as a within-group factor in the analysis by participants and as a between-groups factor in the analysis by items. Group and type of rating were treated as between-groups factors in the analysis by participants and as within-group factors in the analysis by items. The same two ANOVAs were carried out on negativity ratings. Table 3 shows the mean and standard deviation of the affective ratings across conditions.

Positivity ratings

The analyses of positivity ratings revealed a main effect of the three factors involved: positivity ratings were higher in native speakers ($M = 3.77, SD = 1.48$) than in bilinguals ($M = 3.33, SD = 1.72$), $F_1(1, 190) = 6.44, p = .012, \eta_p^2 = 0.03, F_2(1, 158) = 98.93, p < .001, \eta_p^2 = 0.39$, and they were higher in the *know* condition ($M = 4.18, SD = 1.41$) than in the *feel* condition ($M = 2.92, SD = 1.53$), $F_1(1, 190) = 49.20, p < .001, \eta_p^2 = 0.21, F_2(1, 158) = 776.05, p < .001, \eta_p^2 = 0.83$. The ratings were higher in emotional words ($M = 3.93, SD = 1.17$) than in neutral words ($M = 3.26, SD = 1.86$), although in this case the difference was only significant in the analysis by participants, $F_1(1, 190) = 75.81, p < .001, \eta_p^2 = 0.29, F_2(1, 158) = 3.39, p = .067, \eta_p^2 = 0.02$. The interaction

between group and type of rating was close to significance in the analysis by participants and reached statistical significance in the analysis by items, $F_1(1, 190) = 3.86, p = .051, \eta_p^2 = 0.02, F_2(1, 158) = 206.84, p < .001, \eta_p^2 = 0.57$. This interaction, which is displayed in Figure 2a, indicates that the difference between native speakers and bilinguals was significant in the *feel* condition ($p = .002$), but not in the *know* condition ($p = .674$). Furthermore, the difference in positivity ratings between the *feel* condition and the *know* condition was significant in both groups of participants (both $ps < .001$), but the effect size calculated using Cohen’s *d* (Cohen, 1988) was larger for bilinguals ($d = 1.14$) than for native speakers of English ($d = 0.69$). Finally, the interaction between emotionality and type of rating was also significant, $F_1(1, 190) = 11.52, p < .001, \eta_p^2 = 0.06, F_2(1, 158) = 28.78, p < .001, \eta_p^2 = 0.15$. That interaction indicates that both emotional and neutral words were rated as more positive in the *know* condition than in the *feel* condition (both $ps < .001$), the effect size being larger for emotional words ($d = 0.93$) than for neutral words ($d = 0.88$). It also indicates that emotional words were rated as more positive than neutral words in both the *know* and the *feel* condition (both $ps < .001$), being the effect size larger in the *feel* condition ($d = 0.65$) than in the *know* condition ($d = 0.31$).

Negativity ratings

The analyses of negativity ratings revealed a similar pattern of findings. The three factors involved had a significant effect. Indeed, negativity ratings were higher in native speakers of English ($M = 3.47, SD = 1.63$) than in bilinguals, ($M = 2.85, SD = 1.59$), $F_1(1, 190) = 13.98, p < .001, \eta_p^2 = 0.07, F_2(1, 158) = 175.46, p < .001, \eta_p^2 = 0.53$; they were higher in the *know* condition ($M = 3.68, SD = 1.47$) than in the *feel* condition ($M = 2.70, SD = 1.67$), $F_1(1, 190) = 35.03, p < .001, \eta_p^2 = 0.16, F_2(1, 158) = 323.80, p < .001, \eta_p^2 = 0.67$, and they were higher in emotional words ($M = 4.12, SD = 1.26$) than in neutral words ($M = 2.34, SD = 1.48$), $F_1(1, 190) = 826.07, p < .001, \eta_p^2 = 0.81, F_2(1, 158) = 25.39, p < .001, \eta_p^2 = 0.14$. The interaction between group and type of rating was also significant, although only in the analysis by items, $F_1(1, 190) = 1.15, p = .285, \eta_p^2 = 0.01, F_2(1, 158) = 56.48, p < .001, \eta_p^2 = 0.26$. This interaction, which is displayed in Figure 2b, indicates that, although negativity ratings were higher in native speakers than in bilinguals, and in the *know* condition than in the *feel* condition (all $ps < .001$), the effect size of the difference between

Table 3. Mean and standard deviation (in parentheses) of the affective ratings across conditions

Participants	Native speakers of English				EP-English Bilinguals			
	Positivity		Negativity		Positivity		Negativity	
Ratings	Know	Feel	Know	Feel	Know	Feel	Know	Feel
Instructions								
Emotion _{positive}	8.28 (1.09)	7.07 (1.75)	1.31 (1.04)	1.16 (1.05)	8.07 (1.42)	5.84 (2.66)	0.89 (0.69)	0.58 (0.52)
Emotion _{negative}	1.01 (1.16)	0.69 (0.69)	8.31 (1.23)	6.66 (2.13)	0.85 (0.60)	0.29 (0.40)	7.86 (1.19)	5.87 (2.54)
Emotion-laden _{positive}	7.54 (1.33)	6.42 (1.95)	1.56 (1.25)	1.28 (1.13)	7.12 (1.62)	5.06 (2.63)	1.13 (0.85)	0.70 (0.64)
Emotion-laden _{negative}	1.02 (1.12)	0.70 (0.74)	7.97 (1.30)	6.48 (2.02)	0.96 (0.76)	0.28 (0.43)	7.41 (1.38)	5.29 (2.58)
Neutral	3.96 (1.80)	2.79 (1.53)	2.97 (1.48)	2.16 (1.46)	3.96 (1.76)	1.92 (1.65)	2.50 (1.26)	1.29 (1.12)

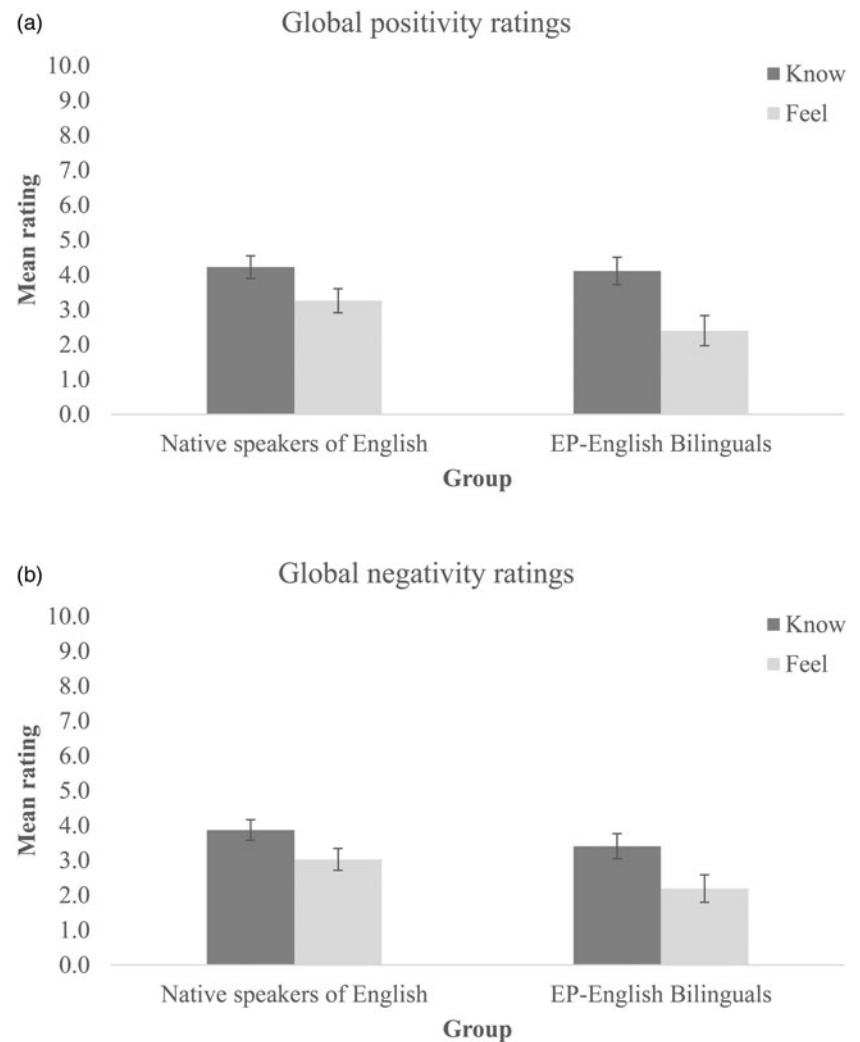


Fig. 2a and 2b. Analysis with all the words: Interaction between group and type of rating in positivity ratings (2a) and negativity ratings (2b).

native speakers and bilinguals was larger in the *feel* condition ($d = 0.39$) than in the *know* condition ($d = 0.16$). On the other hand, the effect size of the difference between the *feel* condition and the *know* condition was larger for bilinguals ($d = 0.51$) than for native speakers of English ($d = 0.33$).

2- Analyses restricted to emotional words: Comparison between Emotion words and Emotion-laden words

We carried out two ANOVAs (by participants and by items) on positivity ratings, restricted to positive words. The factors were as follows: type of emotional word (emotion words vs. emotion-laden words), group (native speakers of English vs. EP-English bilinguals) and type of rating (*know* vs. *feel*). Type of emotional word was treated as a within-group factor in the analysis by participants and as a between-groups factor in the analysis by items. Group and type of rating were treated as between-groups factors in the analysis by participants and as within-group factors in the analysis by items. The same two ANOVAs were carried out on negativity ratings, focusing now on negative words.

Positivity ratings

There was a main effect of the three factors involved: positivity ratings for positive words were higher in native speakers ($M =$

7.36, $SD = 1.68$) than in bilinguals ($M = 6.62$, $SD = 2.39$), $F_1(1, 190) = 9.80$, $p = .002$, $\eta_p^2 = 0.05$, $F_2(1, 38) = 50.30$, $p < .001$, $\eta_p^2 = 0.57$; they were higher in the *know* condition ($M = 7.78$, $SD = 1.41$) than in the *feel* condition ($M = 6.24$, $SD = 2.29$), $F_1(1, 190) = 41.88$, $p < .001$, $\eta_p^2 = 0.18$, $F_2(1, 38) = 724.06$, $p < .001$, $\eta_p^2 = 0.95$, and they were higher in emotion positive words ($M = 7.45$, $SD = 1.93$) than in emotion-laden positive words ($M = 6.68$, $SD = 2.04$), $F_1(1, 190) = 159.62$, $p < .001$, $\eta_p^2 = 0.46$, $F_2(1, 38) = 4.83$, $p = .034$, $\eta_p^2 = 0.11$. The interaction between group and type of rating was close to significance in the analysis by participants and reached statistical significance in the analysis by items, $F_1(1, 190) = 3.66$, $p = .057$, $\eta_p^2 = 0.02$, $F_2(1, 38) = 111.01$, $p < .001$, $\eta_p^2 = 0.74$. Such interaction, which is displayed in Figure 3a, indicates that the difference between native speakers of English and bilinguals was significant in the *feel* condition ($p = .001$), but not in the *know* condition ($p = .371$). It also indicates that, although positivity ratings were higher in the *know* condition than in the *feel* condition for both native speakers of English and for bilinguals (both $ps < .001$), the effect size was larger for bilinguals ($d = 1.01$) than for native speakers ($d = 0.74$).

Negativity ratings

There was a main effect of group: Negativity ratings were higher in native speakers of English ($M = 7.40$, $SD = 1.86$) than in

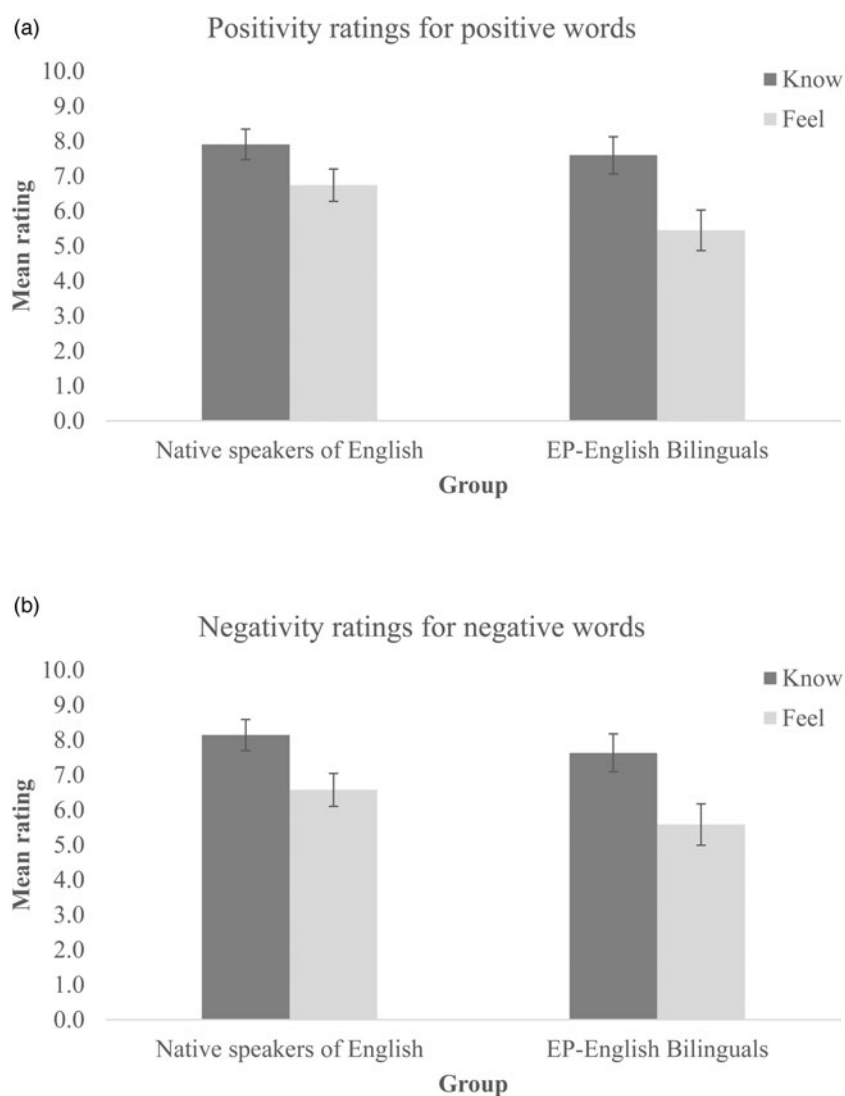


Fig. 3a and 3b. Analysis with emotional words: Interaction between group and type of rating in positivity ratings for positive words (3a) and in negativity ratings for negative words (3b).

bilinguals, ($M = 6.70$, $SD = 2.22$), $F_1(1, 190) = 8.24$, $p = .005$, $\eta_p^2 = 0.04$, $F_2(1, 38) = 64.38$, $p < .001$, $\eta_p^2 = 0.63$. The effect of type of rating was also significant: ratings were higher in the *know* condition ($M = 7.94$, $SD = 1.30$) than in the *feel* condition ($M = 6.19$, $SD = 2.32$), $F_1(1, 190) = 48.18$, $p < .001$, $\eta_p^2 = 0.20$, $F_2(1, 38) = 351.79$, $p < .001$, $\eta_p^2 = 0.90$. The effect of type of emotional word reached statistical significance only in the analysis by participants, $F_1(1, 190) = 56.19$, $p < .001$, $\eta_p^2 = 0.23$, $F_2(1, 38) = 1.92$, $p = .174$, $\eta_p^2 = 0.05$, indicating that participants rated emotion words as more negative ($M = 7.31$, $SD = 2.02$) than emotion-laden words ($M = 6.94$, $SD = 2.05$). The interaction between group and type of rating was significant too, although in this case only in the analysis by items, $F_1(1, 190) = 0.86$, $p = .356$, $\eta_p^2 = 0.004$, $F_2(1, 38) = 17.11$, $p < .001$, $\eta_p^2 = 0.31$. This interaction, which is displayed in Figure 3b, indicates that negative words were perceived as more negative by native speakers of English than by bilinguals in both the *know* condition and the *feel* condition (both $ps < .001$), but the effect size was larger in the *feel* condition ($d = 0.99$) than in the *know* condition ($d = 0.54$). On the other hand, both native speakers of English and bilinguals rated negative words as more negative in the *know* condition than in the *feel* condition (both $ps < .001$), being the effect size larger in bilinguals ($d = 2.00$) than in native speakers ($d = 1.75$).

Finally, the interaction between group and type of emotional word was significant in the analysis by participants, $F_1(1, 190) = 6.07$, $p = .015$, $\eta_p^2 = 0.03$, $F_2(1, 38) = 1.54$, $p = .223$, $\eta_p^2 = 0.04$. This interaction, which is displayed in Figure 4, indicates that negativity ratings were higher for native speakers of English than for bilinguals, both when considering emotion words ($p = .020$) and emotion-laden words ($p = .001$), the effect size being larger in emotion-laden words ($d = 0.41$) than in emotion words ($d = 0.29$). It also indicates that negativity ratings were higher for emotion words than for emotion-laden words, both in native speakers of English and in bilinguals (both $ps < .001$), with a higher effect size in bilinguals ($d = 0.55$) than in native speakers of English ($d = 0.14$).

3- Analyses restricted to emotional words: Comparison between positive words and negative words

The above analyses reveal a similar pattern of findings with positive words (i.e. those with positivity ratings) and negative words (i.e. those with negativity ratings). A direct comparison between them is needed, however, to examine directly the effects of valence (i.e., positive vs. negative). To that end, a single bipolar valence score (positive rating minus negative rating) for each word was

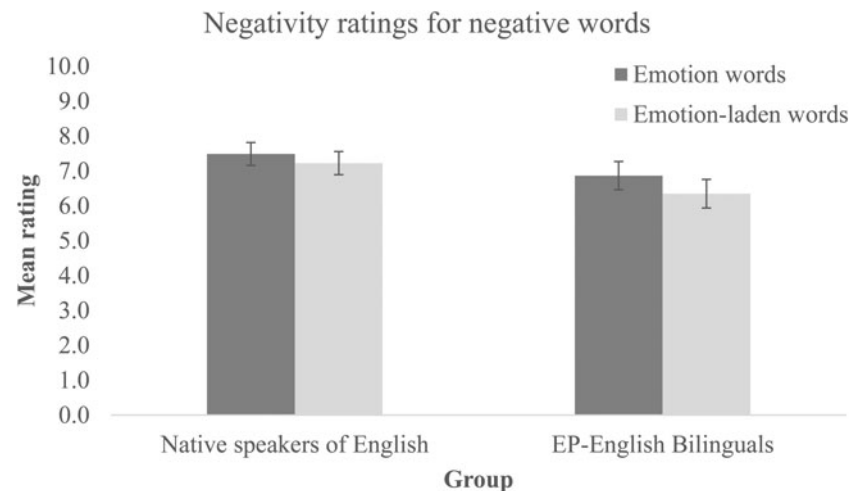


Fig. 4. Analysis with emotional words: Interaction between group and type of emotional word in negativity ratings for negative words.

computed, following the procedure of Itkes et al. (2017). A positive bipolar valence score for a given word indicates that the positivity rating is higher than the negativity rating for that word. Conversely, a negative bipolar valence score indicates that the positivity rating for that word is lower than the negativity rating. We carried out an ANOVA on this variable including these factors: valence (positive word vs. negative word), group (native speakers of English vs. EP-English bilinguals) and type of rating (*know* vs. *feel*). Valence was treated as a within-groups factor in the analysis by participants and as a between-groups factor in the analysis by items, while group and type of rating received the same treatment as in the above analyses. The results revealed a main effect of valence, $F_1(1, 190) = 2333.45$, $p < .001$, $\eta_p^2 = 0.92$, $F_2(1, 78) = 1461.78$, $p < .001$, $\eta_p^2 = 0.95$, showing, as expected, that the bipolar valence score was higher for positive words ($M = 5.93$,

$SD = 1.87$) than for negative words ($M = -6.35$, $SD = 1.99$). The three two-way interactions reached statistical significance: The interaction between valence and type of rating, $F_1(1, 190) = 29.17$, $p < .001$, $\eta_p^2 = 0.13$, $F_2(1, 78) = 346.11$, $p < .001$, $\eta_p^2 = 0.82$; between group and type of rating, $F_1(1, 190) = 2.94$, $p = .088$, $\eta_p^2 = 0.02$, $F_2(1, 78) = 6.47$, $p = .013$, $\eta_p^2 = 0.08$, and the interaction between group and valence, $F_1(1, 190) = 2.49$, $p = .116$, $\eta_p^2 = 0.01$, $F_2(1, 78) = 15.32$, $p < .001$, $\eta_p^2 = 0.16$, although in the last two cases, the interaction only reached statistical significance in the analysis by items. Interestingly, these two-way interactions were qualified by a three-way interaction between group, type of rating, and valence, which was significant in the analysis by items, $F_1(1, 190) = 0.99$, $p = .321$, $\eta_p^2 = 0.01$, $F_2(1, 78) = 20.10$, $p < .001$, $\eta_p^2 = 0.20$ (see Figure 5). In order to interpret this interaction, we analyzed separately the bipolar valence scores for *know* and *feel*

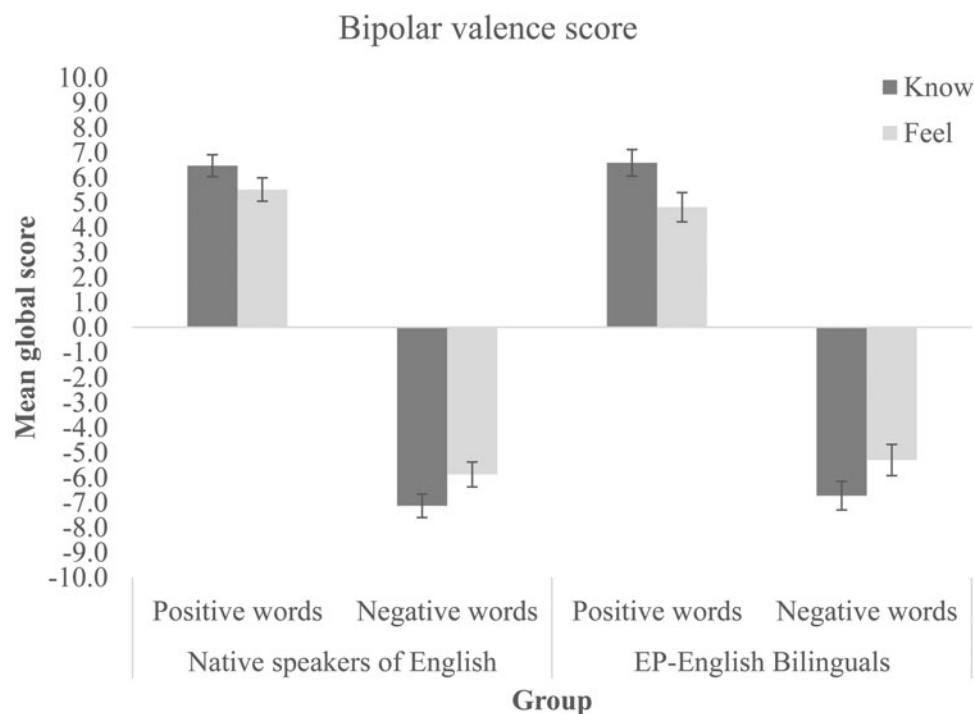


Fig. 5. Analysis with emotional words: Interaction between group, type of rating and valence in the bipolar valence score.

focused instructions. The analysis focused on the *know* condition revealed a main effect of valence, $F_2(1, 78) = 1536.31, p < .001, \eta_p^2 = 0.95$, indicating that the overall bipolar valence score was larger for positive words ($M = 6.52, SD = 1.81$) than for negative words ($M = -6.92, SD = 1.42$). The analysis focused on the *feel* condition revealed again a significant valence effect, $F_2(1, 78) = 1235.75, p < .001, \eta_p^2 = 0.94$, with higher values for positive words ($M = 5.14, SD = 1.70$) than for negative words ($M = -5.56, SD = 1.25$). A significant interaction between group and valence was also present in the *feel* condition, $F_2(1, 78) = 33.01, p < .001, \eta_p^2 = 0.30$. Such interaction indicates that the bipolar valence score was higher for positive words than for negative words in both groups of participants (both $ps < .001$), being the effect size larger in native speakers of English ($d = 7.99$) than in bilinguals ($d = 6.86$). It also indicates that native speakers of English produced more extreme scores for both positive and negative words (i.e., a larger bipolar valence score indicates that a positive word is considered as more positive, and that a negative word is considered as more negative) than bilinguals (both $ps < .001$). In this case, the effect size was larger for negative words ($d = 0.50$) than for positive words ($d = 0.46$).

Discussion

In this study we examined whether there is a difference between native and second-language speakers regarding affective ratings of words. Native speakers of English and EP-English bilinguals provided positivity and negativity ratings for a set of emotional and neutral English words. Among the emotional words, there were emotion words and emotion-laden words. Half of the participants were asked to focus on the feelings elicited by each word to do the ratings, while the other half were asked to focus on their knowledge of the word's meaning. The results were clear-cut: native speakers produced more intense affective ratings than bilinguals, and that difference was larger when participants focused on their feelings than when they focused on the meaning of words. Regarding the comparison between emotion words and emotion-laden words, the effects were rather modest and restricted to negative words, where the difference in ratings between native speakers of English and EP-English bilinguals was larger for emotion-laden words than for emotion words.

In a review of the literature on affective language processing in bilinguals, Pavlenko (2012) suggested that bilinguals might know the emotional meaning of L2 words, but not feel it. The results of the few affective rating studies conducted so far largely support that proposal (Garrido & Prada, 2021; Imbault et al., 2020; Vélez-Urbe & Rosselli, 2019; but see Winkler, 2013, for contrasting findings and Harris, 2004, and Caldwell-Harris et al., 2011, for effects restricted to a particular type of words – namely, childhood reprimands). However, as stated in the introduction, the procedure used in those studies does not allow us to conclude whether the differences between native speakers and bilinguals' ratings are mostly related to the meaning of the words or to the feelings they evoke. Winkler (2013) and Harris (2004) asked participants to rate words' pleasantness and unpleasantness, respectively. With these ambiguous instructions, the aspect on which participants focused to provide their ratings cannot be known. Similarly, the instructions by Garrido and Prada (2021) are somewhat ambiguous because they asked participants to rate the extent to which they thought the word generated positive or negative feelings. It is not clear from these instructions if participants actually focused on the feelings the words produced in them (feeling-focused

ratings) or on their knowledge about the capacity of the words to evoke some feelings (knowledge-focused ratings). The studies by Caldwell-Harris et al. (2011), Imbault et al. (2020), and Vélez-Urbe and Rosselli (2019), by contrast, had less ambiguous instructions, asking participants to pay attention to the feelings that the words produced in them.

It is not necessary to explicitly focus on one's feelings to be affected by the emotional content of words. In fact, there is ample evidence of physiological, electrophysiological, and neuroimaging effects of emotional words in tasks where participants' attention was focused on aspects other than the feelings evoked by words, such as in word recognition tasks (e.g., Conrad et al., 2011; Degner et al., 2012; Sulpizio, Toti, Del Maschio, Costa, Fedeli, Job & Abutalebi, 2019; Toivo & Scheepers, 2019). However, as pointed out by Caldwell-Harris et al. (2011), when it comes to bilingual emotional processing, asking bilinguals to rate words or expressions for pleasantness/valence or emotional intensity may not be the most suitable approach, because they can rely on rote cultural knowledge of the meaning of the word. The ratings most sensitive to language-learning history seem, by contrast, to be those obtained by asking participants about the personal emotional feelings evoked by the stimuli (i.e., feeling-focused ratings). However, this is not enough to test the proposal made by Pavlenko (2012). To that end, it is necessary to compare these feeling-focused ratings with ratings focused on the meaning of words. This, which has not been done before, is what we have done in the present study. Importantly, the few studies conducted with images using such distinction (Hamzani et al., 2020; Itkes & Kron, 2019) show that ratings obtained with feeling-focused instructions are better predictors for the physiological response of participants than those obtained with knowledge-focused instructions. Taking this into consideration, applying that distinction to the field of emotional word processing in the native and the non-native language is a promising methodological approach to better characterize such processing.

The most relevant finding in this study is the interaction between group and type of rating. This interaction indicates that the difference between bilinguals' and native speakers' ratings is more pronounced in the feeling-focused condition than in the knowledge-focused condition. It also indicates that the difference between the feeling-focused condition and the knowledge-focused condition is larger for bilinguals than for native speakers. This result goes beyond what previous research has indicated (Caldwell-Harris et al., 2011; Garrido & Prada, 2021; Harris, 2004; Imbault et al., 2020; Vélez-Urbe & Rosselli, 2019), because it demonstrates that, although differences between native speakers and bilinguals in affective ratings are observed both when they focus on their knowledge of the words or on the feelings they produce, they are larger in the latter. These findings are in line with Pavlenko's (2012) proposal, suggesting that, although bilinguals know the emotional meaning of words in the second language, they may not FEEL them with the same intensity as native speakers. This is the case at least for bilinguals who have acquired their L2 mostly in instructional settings, like the participants of this study. Such difference in the feelings produced by L2 words might be reduced or even disappear when bilinguals have learnt the L2 by immersion, in which case they may have the chance to associate the foreign words with sensory and affective experiences.

Apart from the main result found here, we also expected some differences between the various types of emotional words involved in the study. Although the literature about emotional word processing has not commonly distinguished between emotion

words and emotion-laden words, there is evidence of differences in their processing, both in the native language (e.g., Kazanas & Altarriba, 2015) and in the non-native language of bilingual speakers (Zhang, Wu, Yuan & Meng, 2020). The findings of these studies seem to suggest that emotion words have an advantage in activating emotions over emotion-laden words (see also Wu & Zhang, 2020, for an overview). In this study, we predicted that the difference in affective ratings between native speakers of English and EP-English bilinguals might be larger for emotion-laden words than for emotion words because the affective content of emotion-laden words probably comes from their association with affective experiences, being more prone to individual and culture differences.

Our results supported that prediction in relation to negative words, but not to positive words. In this last case, only a main effect of type of emotional word emerged, which was of the same magnitude for native speakers as for bilinguals. Therefore, these results suggest that, when it comes to positive words, the fact that their affective load comes from their denotative meaning (i.e., emotion words) or from associative mechanisms (i.e., emotion-laden words) does not modulate the difference in affective ratings between native speakers and bilinguals. By contrast, when it comes to negative words, the distinction between emotion words and emotion-laden words matters: the difference in ratings' intensity between native speakers and bilinguals is larger for the latter, as revealed by the interaction between group and type of emotional word. This suggests that the overlap between the meanings of (and the feelings produced by) negative emotion words across languages is larger than that of negative emotion-laden words.

There is no clear explanation at present for the different pattern of findings between positive and negative words in relation to the emotion word/emotion-laden word distinction. It is worth noting that the three-way interaction obtained in the direct comparison between positive and negative words (i.e., the analyses on the bipolar valence score) also revealed a slight difference between these two types of words in feeling-focused ratings. Concretely, native speakers produced a more differentiated bipolar valence score (i.e., more positive ratings for positive words and more negative ratings for negative words, indicating more extreme scores) than bilingual speakers. Importantly, although this was true for all the words, the difference between native speakers and bilingual speakers was larger for negative words than for positive words. This result is in line with previous findings showing larger differences in affective processing between L1 and L2 with negative words in comparison to positive words (Jończyk et al., 2016; Wu & Thierry, 2012). Our results also suggest that this is especially evident when participants focus on the feelings produced by the words (i.e., the difference between positive and negative words was not observed with knowledge-focused ratings). However, considering the subtle effects found here in relation to this issue, more research is needed to examine the effects of valence on bilinguals' affective ratings and their possible interaction with the type of emotional word (emotion words vs. emotion-laden words).

To sum up, we have collected affective ratings for words and conducted the first systematic comparison between feeling-focused instructions vs. knowledge-focused instructions. The results show that bilinguals produce attenuated affective ratings in their L2 in comparison to native speakers. Such difference is more pronounced when participants focus on the feelings produced by the words than when they focus on their meaning.

These findings provide both theoretical and methodological contributions to the literature on the subject. From a theoretical point of view, they inform us about the perception of affective language in bilingual speakers, highlighting the relevance of the distinction between feelings and meanings. From a methodological point of view, they show the need for incorporating such distinction in normative studies from which researchers obtain their experimental stimuli. An important direction for future studies would be to provide additional validity to the distinction between feeling-focused ratings and knowledge-focused ratings by using psychophysiological measures, as has been done with images. Future research should also examine in more depth the relevance of the distinction between different types of emotional words (i.e., emotion words vs. emotion-laden words; positive vs. negative words) in relation to affective processing in bilinguals.

Competing interests. The authors declare none.

Data availability. The data that support the findings of this study are openly available in figshare at <https://doi.org/10.6084/m9.figshare.19182248.v1>.

Acknowledgements. This study was supported by the Ministerio de Ciencia, Innovación y Universidades of Spain (PID2019-107206GB-I00, RED2018-102615-T), and by the Universitat Rovira i Virgili (2019PFR-URV-B2-32). This has been also supported by the Foundation for Science and Technology (FCT) through the Portuguese State Budget (UIDB/01662/2020).

References

- Ayçiçeği-Dinn A and Caldwell-Harris CL (2009). Emotion-memory effects in bilingual speakers: A levels-of-processing approach. *Bilingualism* 12(3), 291–303. <https://doi.org/10.1017/S1366728909990125>
- Baumeister JC, Foroni F, Conrad M, Rumiati RI and Winkielman P (2017). Embodiment and emotional memory in first vs. second language. *Frontiers in Psychology* 8:394. <https://doi.org/10.3389/fpsyg.2017.00394>
- Brybaert M, Warriner AB and Kuperman V (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods* 46, 904–911. <https://doi.org/10.3758/s13428-013-0403-5>
- Caldwell-Harris CL (2014). Emotionality differences between a native and foreign language: Theoretical implications. *Frontiers in psychology* 5:1055. <https://doi.org/10.3389/fpsyg.2014.01055>
- Caldwell-Harris CL, Tong J, Lung W and Poo S (2011). Physiological reactivity to emotional phrases in Mandarin-English bilinguals. *International Journal of Bilingualism* 15(3), 329–352. <https://doi.org/10.1177/1367006910379262>
- Cohen J (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Colbeck KL and Bowers JS (2012). Blinded by taboo words in L1 but not L2. *Emotion* 12(2), 217–222. <https://psycnet.apa.org/doi/10.1037/a0026387>
- Conrad M, Recio G and Jacobs AM (2011). The time course of emotion effects in first and second language processing: a cross cultural ERP study with German-Spanish bilinguals. *Frontiers in psychology* 2:351. <https://doi.org/10.3389/fpsyg.2011.00351>
- Ćoso B, Guasch M, Ferré P and Hinojosa JA (2019). Affective and concreteness norms for 3,022 Croatian words. *Quarterly Journal of Experimental Psychology* 72(9), 2302–2312. <https://doi.org/10.1177/0271747021819834226>
- Degner J, Doycheva C and Wentura D (2012). It matters how much you talk: On the automaticity of affective connotations of first and second language words. *Bilingualism* 15(1), 181–189. <https://doi.org/10.1017/S1366728911000095>
- Dewaele JM (2004). The emotional force of swearwords and taboo words in the speech of multilinguals. *Journal of multilingual and multicultural development* 25(2–3), 204–222. <https://doi.org/10.1080/01434630408666529>
- Dewaele JM (2008). The emotional weight of I love you in multilinguals' languages. *Journal of Pragmatics* 40(10), 1753–1780. <https://doi.org/10.1016/j.pragma.2008.03.002>

- Dewaele JM, Lorette P, Rolland L and Mavrou I (2021). Differences in emotional reactions of Greek, Hungarian and British users of English when watching television in English. *International Journal of Applied Linguistics* 31(3), 345–361. <https://doi.org/10.1111/ijal.12333>
- Eilola TM and Havelka J (2010). Affective norms for 210 British English and Finnish nouns. *Behavior Research Methods* 42(1), 134–140. <https://doi.org/10.3758/BRM.42.1.134>
- Eilola TM and Havelka J (2011). Behavioural and physiological responses to the emotional and taboo Stroop tasks in native and non-native speakers of English. *International Journal of Bilingualism* 15(3), 353–369. <https://doi.org/10.1177%2F1367006910379263>
- Eilola TM, Havelka J and Sharma D (2007). Emotional activation in the first and second language. *Cognition and Emotion* 21(5), 1064–1076. <https://doi.org/10.1080/02699930601054109>
- Ferré P, Anglada-Tort M and Guasch M (2018). Processing of emotional words in bilinguals: Testing the effects of word concreteness, task type and language status. *Second Language Research* 34(3), 371–394. <https://doi.org/10.1177%2F0267658317744008>
- Ferré P, García T, Fraga I, Sánchez-Casas R and Molero M (2010). Memory for emotional words in bilinguals: Do words have the same emotional intensity in the first and in the second language? *Cognition and Emotion* 24(5), 760–785. <https://doi.org/10.1080/02699930902985779>
- Ferré P, Sánchez-Casas R and Fraga I (2013). Memory for emotional words in the first and the second language: Effects of the encoding task. *Bilingualism: Language and Cognition* 16(3), 495–507. <https://doi.org/10.1017/S1366728912000314>
- García-Palacios A, Costa A, Castilla D, Del Río E, Casaponsa A and Duñabeitia JA (2018). The effect of foreign language in fear acquisition. *Scientific Reports* 8, Article 1157. <https://doi.org/10.1038/s41598-018-19352-8>
- Garrido MV and Prada M (2021). Comparing the valence, emotionality and subjective familiarity of words in a first and a second language. *International Journal of Bilingual Education and Bilingualism* 24(2), 275–291. <https://doi.org/10.1080/13670050.2018.1456514>
- Grosjean F (2010). *Bilingual: Life and reality*. Harvard University Press. <https://doi.org/10.4159/9780674056459>
- Guasch M, Boada R, Ferré P and Sánchez-Casas R (2013). NIM: A Web-based Swiss Army knife to select stimuli for psycholinguistic studies. *Behavior Research Methods* 45, 765–771. <https://doi.org/10.3758/s13428-012-0296-8>
- Guasch M, Ferré P and Fraga I (2016). Spanish norms for Affective and Lexico-Semantic variables for 1,400 words. *Behavior Research Methods* 48, 1358–1369. <https://doi.org/10.3758/s13428-015-0684-y>
- Hamzani O, Mazar T, Itkes O, Petrunker R and Kron A (2020). Semantic and affective representations of valence: Prediction of autonomic and facial responses from feelings-focused and knowledge-focused self-reports. *Emotion* 20(3), 486–500. <https://doi.org/10.1037/emo0000567>
- Harris CL (2004). Bilingual speakers in the lab: Psychophysiological measures of emotional reactivity. *Journal of multilingual and multicultural development* 25(2-3), 223–247. <https://doi.org/10.1080/01434630408666530>
- Hinojosa JA, Moreno EM and Ferré P (2020). Affective neurolinguistics: towards a framework for reconciling language and emotion. *Language, Cognition and Neuroscience* 35:7, 813–839. <https://doi.org/10.1080/23273798.2019.1620957>
- Hsu CT, Jacobs AM and Conrad M (2015). Can Harry Potter still put a spell on us in a second language? An fMRI study on reading emotion-laden literature in late bilinguals. *Cortex* 63, 282–295. <https://doi.org/10.1016/j.cortex.2014.09.002>
- Iacoza S, Costa A and Duñabeitia JA (2017). What do your eyes reveal about your foreign language? Reading emotional sentences in a native and foreign language. *PLoS one* 12(10), Article e0186027. <https://doi.org/10.1371/journal.pone.0186027>
- Imbault C, Titone D, Warriner AB and Kuperman V (2020). How are words felt in a second language: Norms for 2,628 English words for valence and arousal by L2 speakers. *Bilingualism: Language and Cognition*. Advance online publication. <https://doi.org/10.1017/S1366728920000474>
- Imbir KK (2015). Affective norms for 1,586 Polish words (ANPW): Duality-of-mind approach. *Behavior research methods* 47(3), 860–870. <https://doi.org/10.3758/s13428-014-0509-4>
- Itkes O, Kimchi R, Haj-Ali H, Shapiro A and Kron A (2017). Dissociating affective and semantic valence. *Journal of Experimental Psychology: General* 146(7), 924–942. <https://psycnet.apa.org/doi/10.1037/xge0000291>
- Itkes O and Kron A (2019). Affective and semantic representations of valence: A conceptual framework. *Emotion Review* 11(4), 283–293. <https://doi.org/10.1177%2F1754073919868759>
- Ivaz L, Costa A and Duñabeitia JA (2016). The emotional impact of being myself: Emotions and foreign-language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 42(3), 489–496. <https://psycnet.apa.org/doi/10.1037/xlm0000179>
- Jończyk R (2016). *Affect-language interactions in native and non-native English speakers*. Springer International Publishing AG. <https://doi.org/10.1007/978-3-319-47635-3>
- Jończyk R, Boutonnet B, Musiał K, Hoemann K and Thierry G (2016). The bilingual brain turns a blind eye to negative statements in the second language. *Cognitive, Affective, & Behavioral Neuroscience* 16(3), 527–540. <https://doi.org/10.3758/s13415-016-0411-x>
- Kazanas SA and Altarriba J (2015). The automatic activation of emotion and emotion-laden words: Evidence from a masked and unmasked priming paradigm. *The American journal of psychology* 128(3), 323–336. <https://doi.org/10.5406/amerjpsyc.128.3.0323>
- Lemhöfer K and Broersma M (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior research methods* 44(2), 325–343. <https://doi.org/10.3758/s13428-011-0146-0>
- Lin J and Yao Y (2016). Encoding emotion in Chinese: a database of Chinese emotion words with information of emotion type, intensity, and valence. *Lingua Sinica* 2(1), 1–22. <https://doi.org/10.1186/s40655-016-0015-y>
- Martin JM and Altarriba J (2017). Effects of valence on hemispheric specialization for emotion word processing. *Language and speech* 60(4), 597–613. <https://doi.org/10.1177/0023830916686128>
- Monnier C and Syssau A (2014). Affective norms for French words (FAN). *Behavior research methods* 46(4), 1128–1137. <https://doi.org/10.3758/s13428-013-0431-1>
- Montefinese M, Ambrosini E, Fairfield B and Mammarella N (2014). The adaptation of the affective norms for English words (ANEW) for Italian. *Behavior research methods* 46(3), 887–903. <https://doi.org/10.3758/s13428-013-0405-3>
- Moors A, De Houwer J, Hermans D, Wanmaker S, van Schie K, Van Harmelen AL, De Schryver M, De Winne J and Brysbaert M (2013). Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. *Behavior research methods*, 45(1), 169–177. <https://doi.org/10.3758/s13428-012-0243-8>
- Opitz B and Degner J (2012). Emotionality in a second language: It's a matter of time. *Neuropsychologia* 50(8), 1961–1967. <https://doi.org/10.1016/j.neuropsychologia.2012.04.021>
- Pavlenko A (2008). Emotion and emotion-laden words in the bilingual lexicon. *Bilingualism* 11(2), 147–164. <https://doi.org/10.1017/S1366728908003283>
- Pavlenko A (2012). Affective processing in bilingual speakers: Disembodied cognition? *International Journal of Psychology* 47(6), 405–428. <https://doi.org/10.1080/00207594.2012.743665>
- Pérez-Sánchez MA, Stadthagen-Gonzalez H, Guasch M, Hinojosa JA, Fraga I, Marin J and Ferré P (2021). EmoPro – Emotional Prototypicality for 1,286 Spanish Words: Relationships with Affective and Psycholinguistic Variables. *Behavior research methods*. Advance online publication. <https://doi.org/10.3758/s13428-020-01519-9>
- Ponari M, Rodríguez-Cuadrado S, Vinson D, Fox N, Costa A and Vigliocco G (2015). Processing advantage for emotional words in bilingual speakers. *Emotion* 15(5), 644–652. <https://doi.org/10.1037/emo0000061>
- Soares AP, Comesana M, Pinheiro AP, Simões A and Frade CS (2012). The adaptation of the Affective Norms for English words (ANEW) for European Portuguese. *Behavior research methods* 44(1), 256–269. <https://doi.org/10.3758/s13428-011-0131-7>
- Stadthagen-Gonzalez H, Imbault C, Perez Sanchez MA and Brysbaert M (2016). Norms of Valence and Arousal for 14,031 Spanish Words. *Behavior Research Methods* 49, 111–123. <https://doi.org/10.3758/s13428-015-0700-2>
- Sulpizio S, Toti M, Del Maschio N, Costa A, Fedeli D, Job R and Abutaleb J (2019). Are you really cursing? Neural processing of taboo words in native

and foreign language. *Brain and language* **194**, 84–92. <https://doi.org/10.1016/j.bandl.2019.05.003>

Sutton TM, Altarriba J, Gianico JL and Basnight-Brown DM (2007). The automatic access of emotion: Emotional Stroop effects in Spanish–English bilingual speakers. *Cognition and Emotion* **21**(5), 1077–1090. <https://doi.org/10.1080/02699930601054133>

Toivo W and Scheepers C (2019). Pupillary responses to affective words in bilinguals’ first versus second language. *Plos one* **14**(4), Article e0210450. <https://doi.org/10.1371/journal.pone.0210450>

Van Heuven WJB, Mandera P, Keuleers E and Brysbaert M (2014). Subtlex-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology* **67**(6), 1176–1190. <https://doi.org/10.1080%2F17470218.2013.850521>

Vélez-Urbe I and Rosselli M (2019). The auditory and visual appraisal of emotion-related words in Spanish–English bilinguals. *Bilingualism* **22**(1), 30–46. <https://doi.org/10.1017/S1366728917000517>

Võ ML, Conrad M, Kuchinke L, Urton K, Hofmann MJ and Jacobs AM (2009). The Berlin affective word list reloaded (BAWL-R). *Behavior research methods* **41**(2), 534–538. <https://doi.org/10.3758/BRM.41.2.534>

Wang X, Shangguan C and Lu J (2019). Time course of emotion effects during emotion-label and emotion-laden word processing. *Neuroscience letters* **699**, 1–7. <https://doi.org/10.1016/j.neulet.2019.01.028>

Warriner AB, Kuperman V and Brysbaert M (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods* **45**, 1191–1207. <https://doi.org/10.3758/s13428-012-0314-x>

Winkel H (2013). The emotional Stroop task and emotionality rating of negative and neutral words in late Thai–English bilinguals. *International Journal of Psychology* **48**(6), 1090–1098. <https://doi.org/10.1080/00207594.2013.793800>

Wu C and Zhang J (2020). Emotion type should be incorporated in affective neurolinguistics: a commentary on Hinojosa, Moreno and Ferré (2019). *Language, Cognition and Neuroscience* **35**(7), 840–843. <https://doi.org/10.1080/23273798.2019.1696979>

Wu YJ and Thierry G (2012). How reading in a second language protects your heart. *Journal of Neuroscience* **32**(19), 6485–6489. <https://doi.org/10.1523/JNEUROSCI.6119-11.2012>

Zhang J, Wu C, Yuan Z and Meng Y (2020). Different early and late processing of emotion-label words and emotion-laden words in a second language: An ERP study. *Second Language Research* **36**(3), 399–412. <https://doi.org/10.1177/0267658318804850>

Appendix A. List of critical stimuli used in the study, divided by condition (sorted by alphabetical order):

Emotion words		Emotion-laden words	
Positive	Negative	Positive	Negative
admiration	agony	authenticity	blasphemy
affection	anguish	creation	censorship
astonished	anxiety	dream	cheat
compassion	depressed	excellence	disadvantage
desire	despair	free	disgrace
euphoric	desperate	freshness	doubt
excitement	disappointment	fruitful	exile
fun	enraged	heavenly	extinction
grateful	envy	honesty	failure
happiness	fear	masterful	loss
hope	guilt	memorable	negativity

(Continued)

Appendix A. (Continued.)

Emotion words		Emotion-laden words	
Positive	Negative	Positive	Negative
love	hate	parenting	poverty
passion	humiliation	phenomenon	punishment
peaceful	melancholy	sincerity	rejection
pleasure	misfortune	soul	sadistic
proud	nervousness	survival	stereotype
satisfaction	pity	triumph	uncertain
serenity	scared	unity	unfaithful
sympathy	shame	victory	unstable
tenderness	terrified	wisdom	wrong
Neutral words			
act	hypothesis	palpable	reminder
anonymous	identity	past	renounce
attempt	indigenous	phase	repertoire
average	input	philosophy	resemblance
busy	jurisdiction	plausible	resonance
conjunction	law	positioning	resource
continuity	leadership	posterity	restriction
debut	legitimacy	practical	rhetoric
deduction	manageable	presumption	selection
description	mentality	previous	similar
development	method	principle	sovereign
dissent	minimum	prophecy	speculation
distinction	mobility	psychic	suitable
emphasis	motto	rampant	tactical
fade	naive	random	theory
feasible	navigation	rate	trend
fixation	negotiable	rationing	underlying
formality	notion	reaction	variable
genre	odd	reign	viable
guidance	opposite	relativity	withdrawal

Appendix B. Instructions for the knowledge-focused questionnaires:

“In this questionnaire you will be presented with a set of words. We are asking you to report what you think about the emotional meaning of these words. Some words have an emotional meaning (i.e., they refer to something positive or negative) while other words do not. We want to make a distinction between what you think and what you feel: what you think about the meaning of a word and how that word makes you feel inside. For instance, imagine you saw on the screen the word GUN. It is possible that you read that word, and know that it has a negative emotional meaning (i.e., it means something negative), but you don’t really feel anything inside. On the other hand, you may read the word GUN and feel strong emotions inside.

In this questionnaire we are asking you to report **what you think** about the **meaning** of the word, and **not about your feelings**. Please report what you think about the emotional meaning of the word, the degree to which it is positive and/or negative.”

Appendix C. Instructions for the feeling-focused questionnaires:

“In this questionnaire you will be presented with a set of words. We are asking you to report about your feelings; about what you feel inside while reading each word. Reporting about your feelings is not always easy. Sometimes it is not exactly clear what to report when asked to report about your feelings. In these instructions we will explain precisely what we mean when we ask you to report about your feelings.

Let's begin with an example. Imagine you read in the screen the word GUN. We want to make a distinction between what you FEEL and what you THINK: how the word makes you FEEL INSIDE and what you THINK ABOUT the meaning of the word. For example, you may read the word GUN and feel strong emotions inside. However, you may know that a word

has a negative meaning, but you don't really feel anything inside when reading it. That is to say, you may know that guns are negative but you yourself don't feel anything when reading that word.

In this questionnaire we are asking you to report how the word **actually made you feel** -- if **your feelings** when reading the word were positive and/or negative.

We are **not** asking what you think about the meaning of the word, your opinion about the word, how you expected yourself to feel, or what you thought we expected you to feel; just the **actual feelings** that you experienced while reading the word – how you actually felt.

Compare yourself to a smoke detector that lights up when it detects smoke. In the same way think of yourself as a feelings detector. Similar to a smoke detector that lights up when it detects smoke, the feelings detector lights up when it detects feelings.”