Bilingualism: Language and Cognition

cambridge.org/bil

Emotional factors of early vocabulary in Spanish as a second language*

Natividad Hernández Muñoz[1] and Ana Blanco Canales[2]

[1]University of Salamanca, Salamanca, Spain and [2]University of Alcalá, Madrid, Spain

Research Article

*The article has been updated since original publication. A notice detailing the change has also been published.

Cite this article: Hernández Muñoz N, Blanco Canales A (2023). Emotional factors of early vocabulary in Spanish as a second language. *Bilingualism: Language and Cognition* **26**, 476–489. https://doi.org/10.1017/S136672892200061X

Received: 11 January 2022
Revised: 31 July 2022
Accepted: 26 August 2022
First published online: 24 November 2022

Keywords:
emotional factors; bilingual lexicon; Spanish vocabulary; valence; arousal

Address for correspondence:
Natividad Hernández Muñoz,
Facultad de Filología, Plaza de Anaya, 1,
37008 Salamanca, Spain
E-mail: natih@usal.es

Abstract

Previous studies analysing the differences in emotionality in first and second language suggest that affective content of lexical items is modulated in certain contexts. This paper investigates the differences in valence and arousal ratings for 300 early words, in both oral and written modalities, through speakers' subjective appraisal of words given by two immersion groups of Spanish late bilinguals (Chinese and European) compared with a group of native speakers. The main goal of our study is to identify the lexical areas where variability occurs, regarding to a set of affective (emotional charge and intensity), grammatical (nouns, adjectives and verbs) and semantic (concreteness) features of words. Our results show that valence is the dimension where the greatest variability is observed between native and bilinguals, although the influence of the independent factors differs considerably. Besides, arousal yields illuminating data regarding the grammatical category of words and differentiation between the groups of participants.

1. Introduction

In recent decades, one of the most frequently repeated claims in scientific discourse has been that cognition and emotion interact in the complex processes carried out by humans, including communication and learning (Damasio, 1994; Dolan, 2002). In both of these, language emerges as the creator and bearer of the emotional component that accompanies human experiences (Lindquist, MacCormack & Shablack, 2015). Hence, research interest in the relationship between language and emotion has grown exponentially in disciplines such as linguistics, anthropology, cognitive psychology and neuroscience (see reviews by Citron, 2012; Hinojosa, Moreno & Ferré, 2019; Pavlenko, 2012; Robinson & Altarriba, 2014). Likewise, there is also a need in bilingualism and language learning to understand how cognitive processes interact with the emotional content of the verbal elements that new speakers incorporate into their linguistic system. Thus, further research on how and to what extent new words in the bilingual lexicon incorporate affective meaning, especially in the early stages, can shed light on how vocabulary is integrated into learners' linguistic knowledge, how this differs from their native language system and how learning contexts affect the development of lexical competence.

Studies on emotional words in different languages indicate that the cultural context in which speakers are immersed determines not only the breadth of emotional vocabulary, but also how emotions are perceived, understood and described, so that each language possesses a unique emotional space (Dewaele, 2010; Robinson & Altarriba, 2014). Numerous studies are devoted to determine how emotionality varies in a speaker's first language (L1) and second or foreign languages (L2). Although these have sometimes differed widely in approach and methodology and have obtained diverse results, they all tend to agree that the effects of emotion are weaker in second languages, especially as regards the perception of emotions.

One of the models for exploring the emotional component of vocabulary is the explicit and subjective appraisal of the emotional connotations of words. Thus, words such as *party* or *gift* elicit high emotional arousal in the speaker-listener and carry a positive emotion while *illness* or *mistake* also bear high arousal but negative emotion.

Research on emotionality and second languages has examined emotional words broadly, those that refer directly to an emotion or feeling (*anger, joy*). These appear to have a different status within the mental lexicon (Pavlenko, 2008) and their semantic space is constructed from componential features of emotional experience (Soriano, 2013). In addition to emotional words, our word corpus included emotion-laden words, which are words that do not explicitly refer to emotions but nevertheless carry them (*party* or *illness*). In native speakers, these two categories often elicit behavioural differences (Pavlenko, 2008), but Kazanas and Altarriba (2016) indicate that the distinction between them is less robust in a second language. In addition, we included words considered neutral or less-emotional-charged words in studies on emotion (*book* or *mouth*), in the belief that such words possess traces of emotional charge derived from their contexts of use, especially in relation to positive or negative connotations.

© The Author(s), 2022. Published by Cambridge University Press

Since the measurement instruments we used enabled speakers to rate emotional variables on a scale, our pool of vocabulary incorporates the three types of words mentioned, as an innovation from previous studies. This enabled us to construct a vocabulary to be tested as close as possible to the one learners are exposed to during the process of lexicon acquisition. We were also interested in assessing both oral and written word perception. Most cognitive and neurocognitive studies use only written stimuli because these avoid additional non-verbal emotional information (Citron, 2012), but in terms of second language speakers' communicative skills, we believe it is important to compare both modalities in order to determine any modulations in the emotional content.

Thus, the main goal of this study of applied psycholinguistics to bilingual vocabulary is to explore how two groups of immersion learners of Spanish as a foreign language, each from different cultural backgrounds and writing systems (Chinese and European), assessed two emotional dimensions of words (valence and arousal) in their L2 compared with native speakers. Our second group of research questions consider a series of different word characteristics, such as affective (positive, neutral and negative words), grammatical (nouns, verbs and adjectives), and semantic (concreteness) features in order to identify the lexical areas where emotional resonance differs between groups. In addition, oral and written verbal stimuli are considered to determine whether modality influences the transfer of emotionality. So, these are our research questions:

1. To what extent do the emotional dimensions of valence and arousal vary in two groups of immersion learners of Spanish (Chinese and European) compared with natives?
2. Do inner characteristics of words (affective, grammatical or semantic) influence the degree of variation? Does written or oral modality also affect the emotional rates? Are there interactions in variation between word features and modality?

2. Theoretical background

2.1. Words and emotion in L1 and L2

Words in a first language contain an emotionality that is constructed through speakers' lifelong experience of the concepts represented by linguistic elements (Pavlenko, 2008, p. 155). The construction of this emotionality involves a combination of factors that have to do with linguistic meaning, shared by all speakers of a language, and experiential characteristics that vary from one person to another. In an L2, acquisition of the emotional associations of words is influenced by a number of interrelated dimensions, such as importation of meaning from the L1, frequency of use, context, socialisation and new learning experiences. Furthermore, language learning is a highly emotional process, not only because it requires simultaneous neurological activation of cognition and emotion (Eder, Hommel & de Houwer, 2007), but also because it generates a rich context of social interaction and personal growth.

One of the major claims in research on emotion in bilingual speakers is that it can be activated through the L1 and L2, so that the effect of emotion on cognitive tasks can be found in both languages. However, the key question is whether the effect of emotion is equally strong in the two languages and if not, what factors determine these differences (Robinson & Altarriba, 2014). Level of proficiency, order of language acquisition, language dominance, age and learning context influence the degree of arousal of emotional resonance in the L2 (Caldwell-Harris, 2014; Degner, Doycheva & Wentura, 2012; Dewaele, 2004, 2010; Pavlenko, 2012;). Analysing differences in emotional permeability in the L2 is hampered by the fact that bilingual speakers and L2 or foreign language learners present widely diverse individual characteristics (De Groot, 2011). This situation, whereby particular characteristics determine emotionality, has been called the EMOTIONAL CONTEXT OF LEARNING (Harris, Gleason & Ayçiçeği, 2006). Hence, another key factor is whether learners have experienced secondary affective socialisation in parallel with language learning itself (Dewaele, 2006; Pavlenko, 2008).

When bilinguals are highly proficient in their L1 and L2, emotional arousal tends to be similar in both languages (Altarriba & Basnight-Brown, 2011; Ferré, García, Fraga, Sánchez-Casas & Molero, 2010), whereas if they are not equally proficient in both, the effect will be stronger in their dominant language (Robinson & Altarriba, 2014). Thus, previous studies indicate that the effect of emotion on second language words is reduced in sequential bilinguals (people who learn an L2 after mastering their L1), especially those who leant their L2 at a later age (Ayçiçeği-Dinn & Caldwell-Harris, 2009; Eilola & Havelka, 2011; Harris, 2004). It follows that less-proficient bilinguals may be unaffected by the emotional content conveyed, with less engagement in communication, and that certain language tasks may therefore be affected by this reduced emotional component. However, this reduction in automatic affective processes in the L2 can also be seen as an advantage. For example, "thinking in a second language" may reduce bias in decision-making because the second language generates greater emotional distance than the native language, an effect that Pavlenko (2012) has called *disembodied* cognition.

2.2. Measuring emotion in bilingual lexicon

Researchers have used varying approaches to study the emotion contained in lexical items. Some studies have been based on speakers' perceptions of the emotional expressive power of each language and the emotional use of language in communication (alternation and code-switching) (Dewaele, 2004, 2010), while others have focused on the processing of emotion-laden words and phrases. These latter have analysed the results of cognitive tasks which reveal the automatic effect of emotional content – for example, memory tasks (Ayçiçeği-Dinn & Caldwell-Harris, 2009; Baumeister, Foroni, Conrad, Rumiati & Winkielman, 2017), word recognition tasks (Altarriba & Bauer, 2004; Degner et al., 2012; Kazanas & Altarriba, 2016), and affective Stroop tasks (Eilola & Havelka, 2011; Eilola, Havelka & Sharma, 2007; Sutton, Altarriba, Gianico & Basnight-Brown, 2007); also via skin conductance (Caldwell-Harris & Ayçiçeği-Dinn, 2009; Harris, Ayçiçeği & Gleason, 2003; Harris, 2004) or neuroimaging techniques (Fan, Xu, Wang, Zhang, Yang & Liu, 2016; Hernández, 2009).

Within the word recognition tasks, one of the most productive methods has been to use an introspective approach, where speakers attribute words with a series of emotional characteristics considered universal. This so-called DIMENSIONAL APPROACH has the advantage of being relatively neutral in terms of language and culture specificities. Assessments are obtained through a series of subjective measurement questionnaires in which speakers are asked to rate words according to different dimensions on a scale. These affective measures have given rise to a comprehensive

https://doi.org/10.1017/S136672892200061X Published online by Cambridge University Press

set of emotional norms for native speakers, almost all derived from the norms of the ANEW questionnaire (Bradley & Lang, 1999) applied to different languages such as Spanish (Stadthagen-González, Imbault, Pérez Sánchez & Brysbaert, 2017). Comparisons between the norms of different native languages indicate that speakers understand the emotional charge of words in a fairly similar way (Soares, Comesaña, Pinheiro, Simões & Frade, 2011). However, norms for second or foreign language speakers are scarce (Garrido & Prada, 2018; Imbault, Titone, Warriner & Kuperman, 2020), probably because of the diversity of speaker or learner profiles. As a result, experimental designs with bilinguals have often been based on data obtained from native speaker norms, subsequently translated into the corresponding second language, which in some cases may bias the results, especially in less-proficient or late ones.

Among the different affective dimensions identified (see Brosch, Pourtois & Sander, 2010 for a review), the most frequent in literature are valence and arousal. Valence refers to the extent to which a word elicits a positive or negative emotion and prompts a reaction of attraction or defence (Colombetti, 2005), whereas arousal indicates the degree of emotional activation that a term elicits, how much calmness or excitement it transmits (Strauss & Allen, 2008). These two characteristics are independent within the two-dimensional structure of affect, both behaviourally and in terms of brain dissociation, although they often show patterns of interrelation (Citron, 2012; Kousta, Vinson & Vigliocco, 2009; Kuperman, Estes, Brysbaert & Warriner, 2014). In studies with monolinguals, valence appears to be more pronounced because it is related to the appraisal of an emotional situation, whereas arousal is associated with less well-defined, uncontrolled physiological reactions (Citron, 2012). Typically, words with more valence, whether positive or negative, tend to have higher arousal values, and words with a negative valence tend to be rated with higher arousal than positive words (Citron, Weekes & Ferstl, 2014). It appears this can be explained by the automatic vigilance hypothesis, which states that negative valence is more important for survival as it activates the defence mechanisms against threatening situations that emerged during evolution. However, some studies with natives have found no asymmetry between positive and negative valence (Kousta et al., 2009). In general, it has been widely demonstrated in psycholinguistics that both play a role in the mental lexicon organisation of native and bilingual speakers alike (Altarriba & Bauer, 2004). Moreover, these dimensions may be perceived differently in an L1 and an L2, especially for some specific types of term such as taboo words and swear words (Dewaele, 2010; Garrido & Prada, 2018; Imbault et al., 2020).

Studies have also been conducted of the relationship between these emotional attributes and other semantic-lexical factors that influence processing tasks, such as formal features (e.g., number of letters, frequency) and semantic features (e.g., concreteness, imaginability) (Altarriba, 2003; Kuperman et al., 2014). Regarding the latest, emotional words and abstract words have shown a special relationship. At first, emotional words were considered to form part of the abstract word set in terms of processing. However, Altarriba and Bauer (2004) demonstrated that they are not processed in the same way and should be considered a separate lexical group in the linguistic system. This may be because emotional words are learnt very early in L1: language development coincides with the development of emotion regulation systems, which encompass autobiographical memories and the sensory, visual, auditory, olfactory, tactile, kinaesthetic and visceral systems (Pavlenko, 2008, p. 156). Thus, emotional words in the L1 are more richly and deeply encoded in terms of their semantic components.

The four previous studies comparing subjective ratings of affective variables between L1 and L2 obtained mixed results. Winskel (2013) compared the subjective valence ratings given by 54 Thai (L1) and English (L2) speakers and 54 native English speakers, using 40 words (20 negative and 20 neutral), and found no differences between the bilinguals' ratings in both their languages, or between English (L2) and English (L1) scores. Meanwhile, Garrido and Prada (2018) compared the emotional ratings given by 230 Portuguese (L1) and English (highly proficient L2) speakers for valence, emotional intensity and familiarity, using a set of 320 words in Portuguese and English. Their results indicate that L1 stimuli obtain more extreme scores for valence (positive ones are more positive and negative ones more negative) and familiarity. However, for emotional intensity, only taboo words obtained higher scores in the L1. Vélez-Uribe and Rosselli (2019) assessed the valence of 120 oral and written words divided into positive, negative and neutral. Their 149 participants were Spanish (L1) and English (L2) speakers living in South Florida who made extensive use of English. In this study, the participants rated positive and taboo words as more emotional in their L2, while negative words were rated as more emotional in their L1. They also found that the predictor of higher valence was high proficiency in both languages. Lastly, in a very recent study, Imbault et al. (2020) collected subjective valence and arousal ratings of 2628 words given by undergraduate students of English (L2) with different native languages. Comparing the valence data with native speaker norm scores, they found, like Garrido and Prada (2018), that L2 responses were more moderate than L1 responses at the extremes of the scale: the more extreme scores were in L1, the more moderate they were in L2. In addition, a wide difference was observed in valence between native and non-native speakers for low-frequency words, those to which learners have been less exposed, been more neutral for the latter. The authors described individual patterns that predict similarity between the two sets: the higher the degree of proficiency and the longer the time spent living in immersion in Canada, the greater the degree of similarity between L1 and L2 ratings. Imbault et al. (2020) linked their results to speakers having a more incomplete mental representation of the word in the L2, especially in denotations and emotional content, as posited by the lexical quality hypothesis (Perfetti, 2007), as well as a lower register of sensorimotor and emotional interactions for words in the L2, as argued by theories stating that emotionality is not acquired through denotative meaning alone, but through interaction with the words (embodied theoretical approach to language – Barsalou, 1999).The differences between the results of the four studies described above may be due to methodological issues, primarily differences between participants and the sets of words used as stimuli.

In our study, we compared both valence and arousal measures, as did Garrido and Prada (2018), and predicted that the greatest differences would be observed in the valence domain, and that extreme scores would tend to be moderated, in line with the results reported by Garrido and Prada (2018) and Imbault et al. (2020). As a major novelty, our study is the first to group bilinguals by cultural background and native language (Chinese and European), since we hypothesised that cultural background would influence the construction of emotions, as would previous learning contexts in the participants' respective countries. Our second innovation was to analyse, for the first time in comparisons with bilinguals, a set of words taken from the *Cervantes*

https://doi.org/10.1017/S136672892200061X Published online by Cambridge University Press

Institute Curriculum Plan corpus of learner words (Instituto Cervantes, 2006), which ensured a diverse set of stimuli of varying emotional intensity that would make it possible to obtain more nuanced scores. In addition, all the selected words corresponded to level A (initial), thus ensuring that the participants – late bilinguals with an intermediate level – would have been exposed in educational contexts to these words and would thus have acquired the denotative meaning of the words. A holistic exploration of a learner's lexical input must include both oral and written stimulus presentation. Hence, as with Vélez-Uribe and Rosselli (2019), we assessed all words in both modalities. In order to refine our analysis focus on linguistic characteristics of words that identify lexical areas with an altered emotional content, our third innovation was to include stimuli features in three independent dimensions: traditional affective factors (positive, negative and neutral words; high- and low-intensity words), the grammatical category of the words and concreteness as a semantic feature. Before exploring personal and contextual factors that may influence emotional charge, it is essential to analyse how emotional space of words is constructed taking into account their grammatical and semantic nature.

3. Methods

3.1. Participants

A total of 149 individuals participated in this study, distributed as follows: 55 native Chinese speakers, 42 native European language speakers (70% German, English, Italian and French; 30% Serbian, Greek, Russian, Slovak and Belarusian) and 52 native Spanish speakers.

The non-Spanish-native participants (all Spanish learners) were divided into these two categories for two reasons: the typological and genetic similarity of the native languages (Sino-Tibetan and Indo-European) and the native language writing system (Chinese script and the alphabetic system). All of them were attending language immersion programmes as part of their undergraduate or graduate curriculums, either at the University of Alcalá or at the Rey Juan Carlos University, both in the Community of Madrid (Spain). Their global level of Spanish was intermediate, ranged between A2 (25%), B1 (50%) and B2 (25%), according to the CEFR. Of the Chinese students, 37 were men and 18 were women, with an age range of 18 to 25 years old. In the case of the European students, 36 were women and 6 were men, and they formed two age groups: 37 were aged between 18 and 25 years old, and 5 between 26 and 35 years old. Most of the Chinese participants were studying Arts and Humanities (42) or Social Sciences (9). The rest were divided among other areas. Similarly, most of the European students were studying Arts and Humanities or Social Sciences (15 in each case).

As regards the use of Spanish with family and friends, in both groups, 50% (27 in the Chinese group and 21 in the European group) reported using it frequently or sometimes and the other 50% (28 and 21), never or almost never. However, the two groups differed in their contact with Spanish language and culture, since 76.5% of the European students scored between medium and very high, whereas only 39.1% of the Chinese students did so.

With respect to the native speaker participants, all were undergraduate students at the University of Salamanca, studying in the Faculty of Philology; 12 were men and 40 women, and they all belonged to the first age group (18-25 years old). They participated on a voluntary basis and received academic credits for their collaboration.

3.2. Instruments

The valence and arousal data collection questionnaire consisted of 300 words taken from the A1 and A2 levels of the *Cervantes Institute Curriculum Plan* (PCIC) (2006). There were 142 nouns, 84 adjectives and 72 verbs for the factor grammatical category; 160 positive, 90 neutral and 50 negative for emotional charge; 167 low and 133 high for intensity; and 151 concrete and 149 abstract words for concreteness factor.

Both emotional charge and intensity, as independent variables, were calculated from valence and arousal data extracted from *EmoFinder* (Fraga, Guasch, Haro, Padrón & Ferré, 2018). In the first case, the total scale was divided into three levels of similar amplitude: 1 to 3 were negative words, 3.1 to 6 were neutral words and 6.1 to 9 were positive words. With regard to intensity, the mean score (4.25) was taken as a cut-off, dividing scores into high or low arousal according to whether they were above or below the mean.

In the case of the concreteness factor, the *EmoFinder* database only contained measures for about 25% of the words in our corpus. Consequently, we collected data on subjective measures of concreteness for our 300 items in which 150 native Spanish speakers studying at the universities of Salamanca, Alcalá and the Basque Country participated. As with intensity, the mean (4.50) was calculated to establish a cut-off between concrete words (above 4.50) and abstract words (below 4.50).

To select the 300 items all single words of levels A1 and A2 from PCIC were extracted (N = 1094 words). Compounds, adverbs, complex noun phases and repeated words were not taken into account. We also discarded 175 words that were not included in *EmoFinder* and included 11 emotion words that were not in PCIC A level (N = 926). Lexicon from initial levels shows unbalanced grammatical categories with a high presence of nouns (nouns = 641; adjectives = 173; verbs = 112) and valence groups with a high number of positive words (positive = 464; neutral = 393; and negative = 69). Our selection responds to the motivation of including the three groups of valence (positive, negative and neutral) in each grammatical category. The 300 items represent the 32.3% of initial corpora (22% nouns; 49.1% adjectives; 66% verbs). So, the difference in numbers through words factors (such as grammatical category or emotional charge) is due to the aim of working with a real learning second-language lexical corpus for basic levels of proficiency, not an artificial experimental set of stimuli. Corpora offer the advantage of providing a large amount of authentic linguistic material that can be representative of large populations, as well as allowing for a wide variety of methodologies and approaches (Soriano, 2013). The apparent unequal distribution of number of items in each category is assumed by statistical analysis power.

3.3. Data collection

For data collection, the total number of words was divided into two questionnaires (questionnaire A and questionnaire B), which were identical in terms of the number of words for each grammatical category (noun, adjective and verb) and for each emotional level (positive, neutral and negative), and almost identical for the factors of concreteness (concrete and abstract) and intensity (low and high).

https://doi.org/10.1017/S136672892200061X Published online by Cambridge University Press

The two groups of foreign participants (Chinese and European) were divided into two subgroups (subgroup A and subgroup B), and each subgroup was administered one of the questionnaires. Thus, each participant rated 300 words: 75 words for valence and 75 for arousal presented in oral and written modalities. Participants scored the items on a scale of 1 to 7, where 1 indicated negative or no arousal and 7 indicated positive or high arousal. An additional box was included in case they did not know the meaning of the word. The oral stimuli were recorded by a professional announcer, male of medium age, with a central-peninsular Spanish accent. He was instructed to speak neutrally, avoiding prosodic changes that could affect emotional interpretation.

The questionnaires – consisting of four sections or tasks – were administered in paper format and completed in class, within a similar time frame, in the same context and without distractions. A week beforehand, the project was briefly explained to all participants, their collaboration was requested and they were provided with an informed consent form. They were also asked to complete a sociodemographic questionnaire.

The group of native speaker participants was organised in a similar way. Even though measures were available for all 300 items in the *EmoFinder* database (Fraga et al., 2018), they were only for written modality. So, it was necessary to collect measures in both modalities for this group as well. All statistical analysis were done using SPSS.26 except mixed-effect models which were made with R.

4. Results

4.1. General analysis

For the first research question, general comparison of overall distributions of valence and arousal rating were conducted, following the model of statistical simple linear regression of Imbault et al. (2020) where the differential between native rates (L1) and each group of non-native rates (L1-L2) is the dependent variable and the L1 rates, the independent variable. A negative difference means a word that receive more positive response or higher arousal in native than in non-native rating, and a positive difference signifies a lower valence (less positive o more negative) or lower arousal in native than in non-native scores. Figure 1 shows the differences between values in both factors, valence and arousal, and in two groups of late bilinguals (Chinese and European). Words close to zero line represent similar rates in natives and no natives. As in Imbault et al. (2020) non-native values are consistently more attenuated than native values at extreme position of the scale: positive words are less positive (higher differential rates) and negative words are less negative (lower differential rates). The same pattern is seen in arousal. The best-fit regression line demonstrated a linear effect of native ratings on differential measurements (valence L1-Chinese L2 differential, $Y = -1{,}049$, $R^2 = 29\%$, correlation .542, $p < .01$; valence L1-European L2 differential, $Y = -1{,}265 + .290x$, $R^2 = 40\%$, correlation .636, $p < .01$; arousal L1-Chinese L2 differential $Y = -1{,}342 + .36x$, $R^2 = 16\%$, correlation .398, $p < .01$; arousal L1-European L2 differential $Y = -1{,}402 + .307x$, $R^2 = 26\%$, correlation .402, $p < .01$). In order to delve in these differences a set of more complex analysis with interactions between and within factors has been developed.

4.2. Mixed-effect analysis

In order to answer the second research question, to analyse how word features and their interactions affect valence and arousal, two adjusted mixed-effect regression analysis were developed

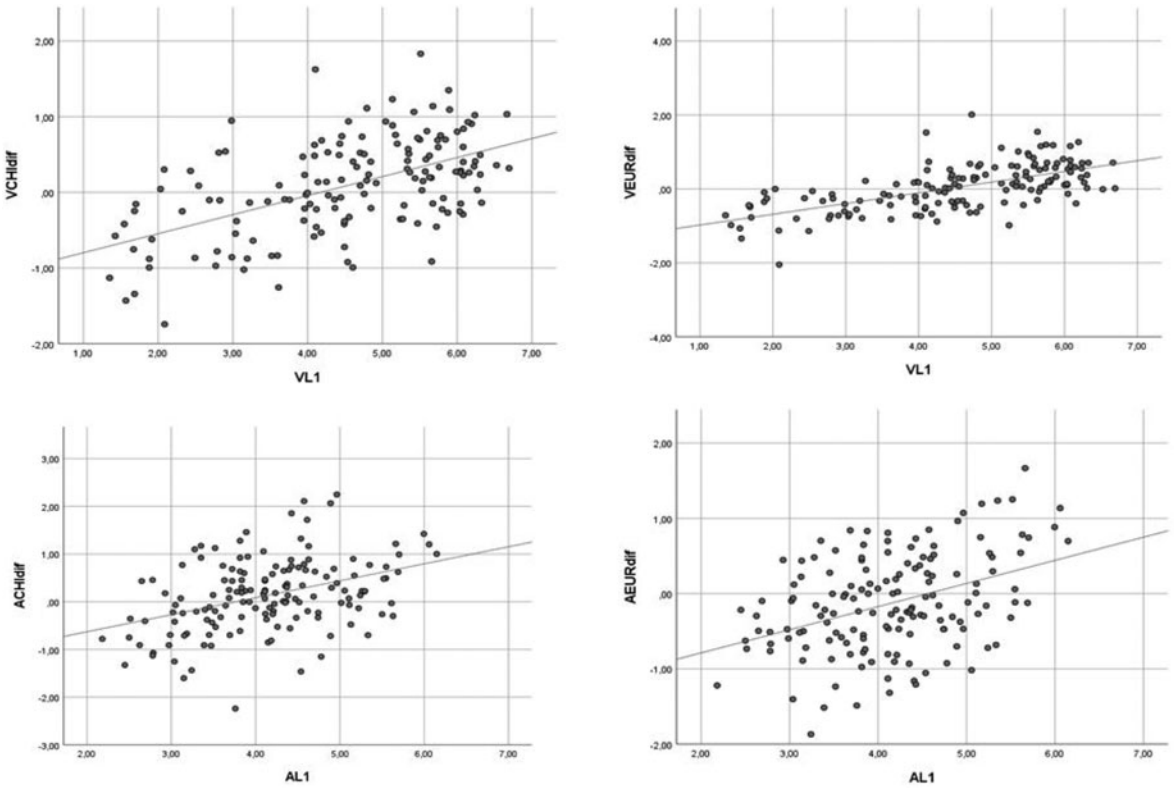


Fig. 1. Scatterplots of the differences between native (L1) and non-native valence (V) and arousal (A). Chinese differential (CHIdif) and European (EURdif).

https://doi.org/10.1017/S136672892200061X Published online by Cambridge University Press

with both dimensions rates (valence and arousal) as dependent variables. The models incorporate the main effects of six fixed factors (grammatical category, emotional charge, intensity, concreteness, modality and group) and their interactions as well as two random effects between participants and items with the differences on the slope of the explanatory effects. No interactions of random effects were included.

4.2.1. Valence

The model has been partially favourable. It presents as significant factors emotional charge (F (2, 130) = 2.20, $p < .001$), intensity (F (2, 129) = 10.57, $p = .001$) and modality (F (1, 21289) = 20.10, $p < .001$), and most of interactions where group and modality are included (emotional charge*group F (4, 21353) = 64.360, $p < .001$, intensity*group F (2, 21444) = 26.844, $p < .001$, concreteness*group F (2, 21391) = 8.447, $p < .001$, emotional charge*modality F (2, 21286) = 67.871, $p < .001$, concreteness*modality F (1, 21286) = 5.196, $p < .001$, group*modality F (2, 21289) = 2.809, $p = .06$). Variance explained by fixed effects is 35% (R^2 marginal = 0.35). If random factors are considered it grows until 47% (R^2 conditional = 0.47). Results suggest the convenience of the included factors. Coefficients with significance show that interactions with a more relevance are emotional charge*group (0.40, 0.71, 0.89), intensity*group (0.31) and emotional charge*modality (0.61). Even though category is not a significant factor, coefficients in the interactions indicate not to discard it in subsequent analysis.

4.2.2. Arousal

Arousal model is more favourable than the valence model because it shows significance values for more factors: emotional charge (F (2, 130) = 22.587, $p < .001$), intensity (F (1, 130) = 19.550, $p < .001$), concreteness (F (1, 131) = 6.445, $p = .012$) and group (F (2, 20833), $p < .001$). Grammatical category and modality are not significant as isolate factors. All interactions where group factor is involved are relevant (category*group F (4, 21280) = 8.998, $p < .001$, emotional charge*group F (4, 21883) = 23.862, $p < .001$, intensity*group F (2, 21321) = 5.988, $p = .003$, concreteness*group F (2, 21283) = 6.785, $p = .001$), and also emotional charge*modality (F (2, 21212) = 3.046, $p = .048$). However, the R_2 marginal is low (0.07), so the amount of variance explained by fixed factors is 7%. If random effects are considered, it reaches 27% (R_2 conditioned = 0.239). Coefficients show that the greatest effect on the dependent variable on the slope is caused by emotional charge (−0.824, −0.577), intensity (0.583), and interactions of emotional charge*group. (−0.394, −0.350, −0.811) and category*group (0.203, −0.275. −0.160).

To delve into the results of mixed analysis and, specially, about interaction between factors, a group of general linear model (GLM) of mixed repeated measures analysis has been developed. These models are designed from the factors that show more influence and theoretical relevance, groups and modality, and they will allow to see in detail the contexts and lexical spaces of affective discrepancies between L1 and L2.

4.3. Complementary Analysis

We performed various general linear model (GLM) analysis, mixed within-and-between repeated measures ANOVA, according to 3 × 2 × 3 or 3 × 2 × 2 models for each affective dimension (valence and arousal). Valence and arousal rates were dependent variables, groups was a between-subjects independent variable and modality (written, oral), emotional charge (positive, neutral, negative), intensity (low, high), concreteness (abstract, concrete) and grammatical category (noun, adjective, verb) were within-subjects independent factors. In all of them, group and modality were the constant independent variables, while the other four factors alternated. Despite the debate about considering affective values as categorical ones (Warriner, Shore, Schmidt, Imbault & Kuperman, 2017), we maintain the most frequent line of analysis of affective scores in bilingual approach to establish comparisons (Imbault et al., 2020; Vélez-Uribe & Rosselli, 2019). Tables 1 and 2 give the means and standard deviations for valence and arousal, respectively.

Analysis of valence

Factors: Groups, modality, and emotional charge

The data obtained from repeated measures analyses (Greenhouse-Geisser correction) indicate a significant interaction between the three factors (F (4,290) = 18.839, $p < .001$, *partial* $\eta^2 = .206$). This demonstrates that participants in the three groups rated the valence of positive, neutral and negative words differently, and that they also differed when rating them in the two modalities (written words scored higher than oral words for all three types of word and for the three groups). The partial η^2 effect size shows that the three-way interaction accounted for 20% of the variance in scores in this test, which is a large effect size (Richardson, 2011). Furthermore, the power analysis indicates

Table 1. Valence means and standard deviations.

Modality	Word feature	Chinese	European	Native
Written	Positive	5.35 (.58)	5.53 (.50)	5.64 (.52)
	Neutral	4.25 (.64)	4.17 (.50)	4.23 (.62)
	Negative	2.65 (.84)	2.43 (.64)	2.30 (.67)
	Low arousal	4.94 (.59)	4.89 (.49)	4.54 (.45)
	High arousal	3.94 (.61)	4.21 (.55)	4.82 (.50)
	Concrete	4.41 (.56)	4.54 (.39)	4.57 (.54)
	Abstract	4.70 (.48)	4.74 (.44)	4.81 (.45)
	Nouns	4.62 (.55)	4.64 (.43)	4.64 (.51)
	Adjectives	4.62 (.46)	4.67 (.37)	4.71 (.41)
	Verbs	4.48 (.66)	4.61 (.58)	4.67 (.72)
Oral	Positive	5.09 (.59)	4.83 (.62)	5.41 (.65)
	Neutral	4.10 (.58)	4.25 (.61)	3.93 (.59)
	Negative	2.71 (.84)	3.43 (1.24)	2.17 (.63)
	Low arousal	4.72 (.57)	4.58 (.53)	4.35 (.50)
	High arousal	3.87 (.49)	4.18 (.50)	4.51 (.55)
	Concrete	4.29 (.45)	4.44 (.44)	4.34 (.54)
	Abstract	4.50 (.47)	4.42 (.42)	4.51 (.55)
	Nouns	4.42 (.46)	4.45 (.41)	4.42 (.51)
	Adjectives	4.51 (.46)	4.42 (.43)	4.48 (.49)
	Verbs	4.20 (.51)	4.44 (.64)	4.31 (.81)

https://doi.org/10.1017/S136672892200061X Published online by Cambridge University Press

Table 2. Arousal means and standard deviations.

Modality	Word feature	Chinese	European	Native
Written	Positive	4.20 (.98)	4.69 (.89)	4.54 (.93)
	Neutral	3.38 (.92)	3.51 (.93)	3.66 (.79)
	Negative	4.45 (1.17)	4.32 (1.11)	4.80 (.91)
	Low arousal	3.78 (.88)	4.02 (.79)	3.94 (.74)
	High arousal	4.32 (.77)	4.62 (.82)	4.87 (.71)
	Concrete	3.90 (.90)	4.26 (.76)	4.13 (.71)
	Abstract	4.09 (.79)	4.30 (.87)	4.50 (.78)
	Nouns	4.03 (.82)	4.35 (.77)	4.29 (.66)
	Adjectives	4.03 (.82)	4.13 (.88)	4.09 (.69)
	Verbs	3.92 (.91)	4.35 (.98)	4.65 (1.02)
Oral	Positive	4.33 (.77)	4.68 (.92)	4.47 (.90)
	Neutral	3.58 (.87)	3.63 (.96)	3.74 (.80)
	Negative	4.36 (1.16)	4.14 (1.24)	4.93 (1)
	Low arousal	3.92 (.77)	4.02 (.8)	3.97 (.73)
	High arousal	4.39 (.73)	4.51 (.97)	4.87 (.7)
	Concrete	4.04 (.75)	4.24 (.81)	4.12 (.75)
	Abstract	4.17 (.72)	4.24 (.89)	4.55 (.72)
	Nouns	4.14 (.68)	4.27 (.83)	4.33 (.67)
	Adjectives	4.08 (.77)	4.12 (.91)	4.20 (.72)
	Verbs	4.04 (.84)	4.34 (1.02)	4.46 (.96)

that this interaction had great power (100%) to detect statistical differences.

Supplementary data from the pairwise comparison show that in written modality, there were statistically significant differences between the Chinese participants and native Spanish speakers in the case of positive and negative words (in neutral words, the three groups obtained very similar means). As can be seen in Table1, Figures 2 and 3, Chinese participants gave lower scores to positive words than native Spanish speakers ($p < .01$), but the opposite was true for negative words, suggesting that for the Chinese learners, the gap between positive and negative narrowed, rendering the extreme positions more moderate. This tendency is not significative for European bilinguals in written scores; however, for the oral rates positive and negative words are also different from natives ($p < .001$).

In oral modality, neutral words were perceived in the same way by all three groups. However, in the case of positive and negative words, the groups showed different behaviours, especially in the case of negative words. We found that native Spanish speakers gave significantly higher scores to positive words than the

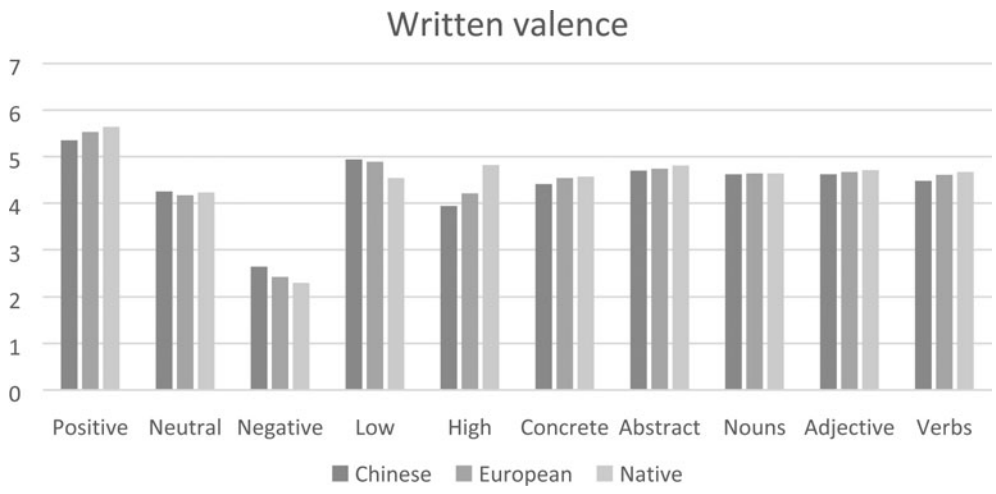


Fig. 2. Results of written rates of valence.

https://doi.org/10.1017/S136672892200061X Published online by Cambridge University Press

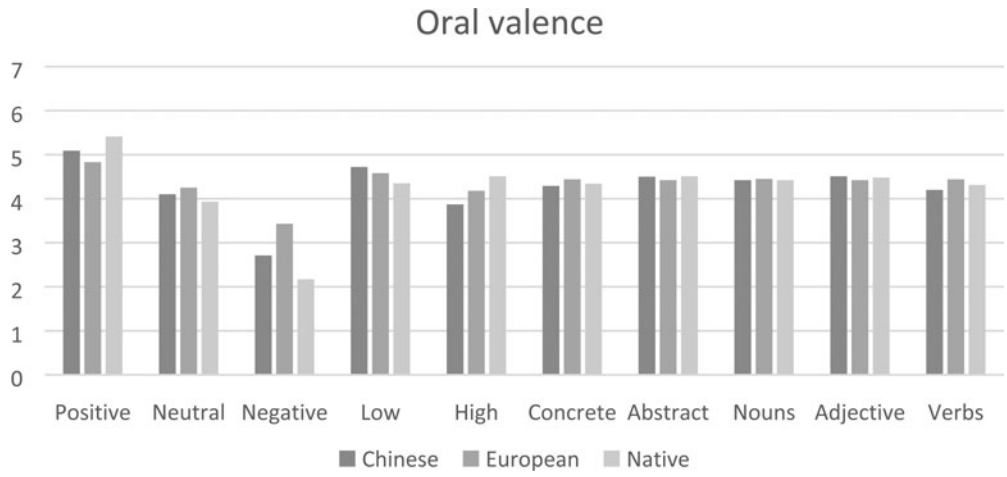


Fig. 3. Results of oral rates of valence.

European and Chinese participants ($p < .01$), (whereas there were no significant differences between these latter two). For negative words, significant differences were detected between all three groups in oral modality, but only between the Chinese and European participants in written modality. These findings therefore show that the largest differences between groups occurred with oral words, especially negative ones, and that the non-native speakers were less sensitive to the emotional charge of the words.

Factors: Groups, modality, and intensity

A repeated measures analysis indicated significant interaction between the three variables considered (F (2, 146) = 10.476, $p < .001$), with a medium effect size, showing that factor interaction accounted for 12% of the variance, and a power of the analysis very high (*partial* $\eta^2 = .125$, *power* = .987). Thus, the three groups of participants showed significant differences in the valence of low- and high-arousal words in both oral and written items (see Figures 2 and 3).

Pairwise comparisons revealed some important details that illuminated the results obtained. In the case of modality and groups, only one statistically significant difference ($p < .05$) was found, between Chinese participants and native Spanish speakers in written modality, whereby the former gave considerably lower scores to high-arousal words than native speakers, and slightly higher scores to low-arousal words.

As regards intensity and groups, the data indicate statistically significant differences between all three groups of participants in the case of high-intensity words, and between native and non-native speakers in the case of low-intensity words. In the first case (high intensity), native speakers gave the highest scores, followed by European and finally Chinese participants. In the second case (low intensity), the situation was reversed: native speakers obtained lower means than the other two groups (see Table 1). This indicates that the European and Chinese participants perceived high-arousal words as less positive than native speakers did and low-arousal words as more positive. These data therefore suggest that non-native speakers are less affected than native speakers by the intensity carried by words when assessing valence.

Factors: Groups, modality, and concreteness

Tests for within-subjects effects revealed that the interaction of all factors was not significant (F (2, 146) = 2.270, p > .05, partial $\eta^2 = .030$, *power* = .456). However, considered in isolation, modality (F (1, 146) = 40.501, $p < .001$, *partial* $\eta^2 = .217$, *power* = 1.000) and concreteness (F (1, 146) = 35.516, $p < .001$, *partial* $\eta^2 = .204$, *power* = 1.000) were significant, and the interaction between the two was also statistically significant (F (1, 146) = 14.404, $p < .001$, *partial* $\eta^2 = .090$, *power* = .965). In post-hoc tests, the Bonferroni statistic indicated that groups had no effect on the valence score ($p > .05$ at all levels). In the complementary analysis the most relevant result is that all three groups perceived abstract words more positively than concrete ones, in written and oral format alike.

Factors: Group, modality, and grammatical category

The Greenhouse-Geisser statistics indicated that there was no significant interaction between the factors considered (F (4,292) = 2.309, $p > .05$, *partial* $\eta^2 = .031$, *power* = .618). Analysed in isolation, both modality (F (1, 146) = 42.484, $p < .001$, *partial* $\eta^2 = .225$, *power* = 1.000) and grammatical category (F (2,292) = 4.694, $p < .05$, *partial* $\eta^2 = .031$, *power* = .679) showed significance levels. In the case of modality, this factor explained 22% of the variance (large effect size), whereas category only explained 3% (small effect size). Between-subjects tests and the Bonferroni statistic showed that differences according to groups were not significant at any of the possible levels.

In pairwise comparisons, the European participants were in full agreement with the native Spanish speakers, as both groups rated the valence of the three categories of words similarly. Chinese participants, however, presented statistically significant differences between adjectives and verbs ($p < .01$), with lower means for verbs (see Table 1 and Figure 2).

In the interaction between modality and category, we found no difference in written words by grammatical category (F (2, 145) = 1.91), but in the case of oral words (F (2, 145) = 5.63), we observed a difference between adjectives and verbs ($p < .01$), with slightly lower scores for verbs. This difference occurred in the Chinese group (F (2, 145) = 7.95) between nouns and verbs ($p < .05$) and between verbs and adjectives ($p < .01$).

Analysis of arousal

Factors: Groups, modality, and emotional charge

The Greenhouse-Geisser correction showed that there was no interaction between the three factors analysed (F (4,274) = 1.375, $p > .05$), but we did find statistically significant interaction effects

https://doi.org/10.1017/S136672892200061X Published online by Cambridge University Press

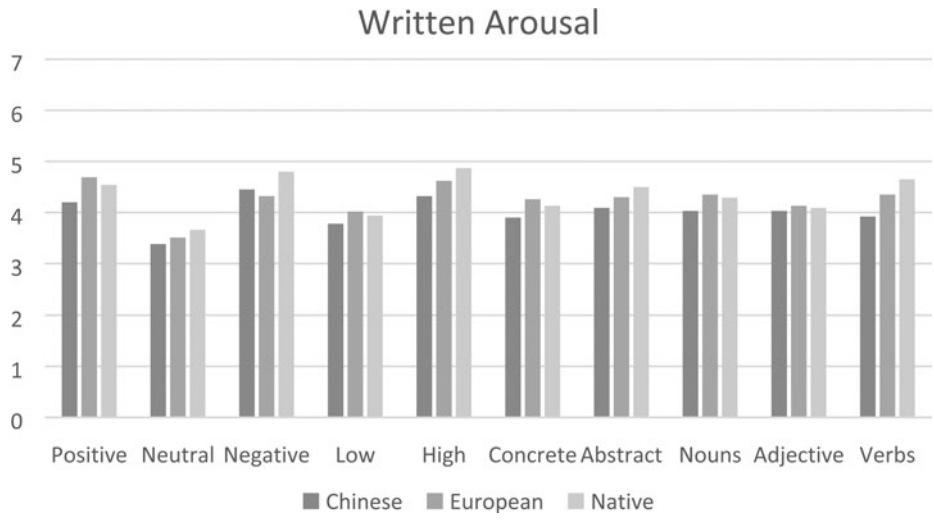


Fig. 4. Results of written arousal.

between place or origin and emotional charge (F (4,274) = 4.011, $p < .05$). The partial η^2 indicated that this interaction explained 6% of the variance, which is a medium effect size, with sufficient power of analysis (81%). Participants in the three groups behaved differently when measuring emotionally charged words. Modality, on the other hand, was not a statistically significant factor. Note also that both the between-subjects and Bonferroni post-hoc tests were close to statistical significance ($p = .06$).

Figures 4 and 5 show how the above-mentioned factors interacted. Pairwise comparisons revealed that neutral words had lower arousal than positive or negative words, and all three groups presented a statistically significant difference between neutral words and the other two sets of words ($p < .05$). The interaction between non-native and native speakers is clearly evident in the negative words. For native speakers, the words with the highest arousal were negative words in both modalities, but a statistically significant difference with respect to the non-native speaker groups was only detected in oral words. These results indicate that the arousal of negative words was very clearly lower in the L2 in oral modality. This reduction was particularly noticeable in the group of Europeans; in fact, the arousal of negative words was below that of positive ones. Meanwhile, in written items, we observed differences between the two non-native speaker groups: the Chinese participants rated positive words lower than the Europeans ($p < .05$).

Factors: Groups, modality, and intensity

There was no statistical interaction between the three factors (F (2, 138) = 0.192, $p > .05$), but there was an interaction between place or origin and intensity (F (2, 138) = 8.104, $p < .05$), which explained 10% of the variance (medium effect size), with sufficient statistical power (94%). This indicates that the three groups rated words of more or less intensity in significantly different ways. Native speakers tend to assign a higher rating to high-intensity words than the other two groups, in both modalities. In fact, in the between-subjects tests, the groups factor was very close to statistical significance ($F(2, 138) = 2.787$, $p = .06$), and the Bonferroni post-hoc tests indicated that the Chinese group differed significantly from the native speaker group for high-arousal words ($p = .05$). In this case, modality had no effect on the scores, and the means were remarkably similar for all the three groups.

Factors: Groups, modality, and concreteness

As in the previous analyses, we found no interaction between the three factors analysed (F (2, 138) = 0.982, $p > .05$), but we did observe an interaction between groups and concreteness (F (2,

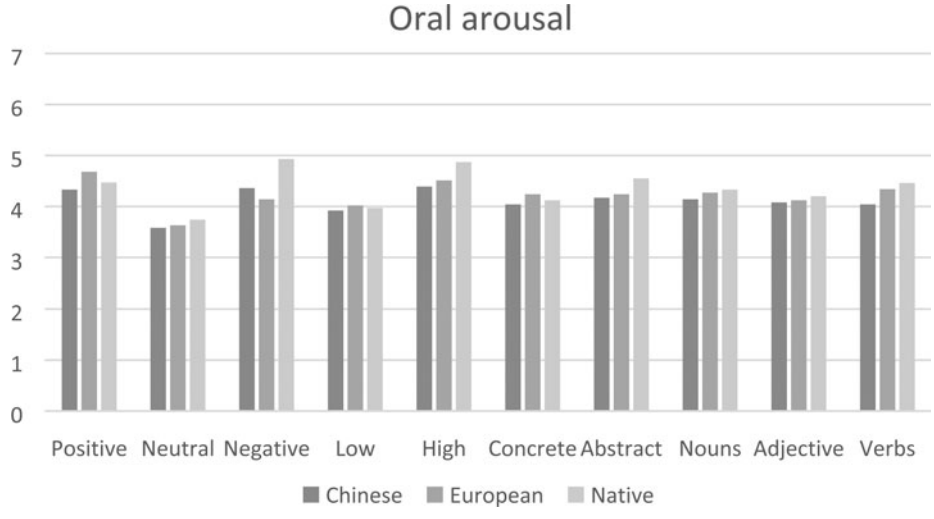


Fig. 5. Results of oral arousal.

Data availability: The data that support the findings of this study are openly available in EmoELE at https://grupoleide.com/herramienta-de-analisis/.

https://doi.org/10.1017/S136672892200061X Published online by Cambridge University Press

138) = 6.532, $p < .05$), with a partial η^2 indicating a medium effect size, explaining 8% of the variance and sufficient power (90%). Considered in isolation, the concreteness factor was significant (F (1, 138) = 2.775, $p < .01$) and explained 14% of the variance with a large effect size and 99% power: abstract words scored higher than concrete words. Pairwise comparisons showed that this latter statistical difference occurred between native Spanish speakers and Chinese participants in written items and only between native Spanish speakers in oral items. As for differences between groups, these only emerged in abstract words between the Chinese participants and native Spanish speakers, in both modalities.

Factors: Groups, modality and grammatical category

The repeated measures analysis revealed a clear interaction between the three factors (F (3.94, 270.46) = 3.571, $p < .05$). The partial η^2 showed that this interaction explained 5% of the variance, indicating a small effect size, with sufficient power of analysis (86%). In both modalities, there were significant differences between adjectives – the grammatical category with the lowest arousal – and nouns and verbs, which obtained the highest scores: written ($F = 7.979$) $p < .01$ and oral ($F = 4.793$) $p < .05$).

In pairwise comparisons, verbs obtained higher arousal scores in native Spanish speakers and European participants (see Table 2), but in the group of Chinese learners, verbs were the grammatical category that obtained the lowest arousal values in both modalities. Figure 4 illustrates descriptively the statistical difference between the native speaker and Chinese groups in written words. Likewise, the differences between grammatical categories were more extreme in the native speaker group, reaching statistical significance between the three categories in written modality and between adjectives and verbs in oral modality.

5. Discussion

Our main goal was to determine whether differences existed in subjective valence and arousal ratings between native Spanish speakers and two groups from different native languages (Chinese and European). The most widespread trend in previous research on behaviour and automatic emotion processing indicates a reduced emotionality in L2, but the results depend on the type of task performed and individual factors, such as the level of proficiency. Studies comparing conscious subjective appraisals of emotional factors have also obtained mixed results. Winskel (2013) found no difference between L1 and L2 values, whereas Garrido and Prada (2018) and Imbault et al. (2020) have reported neutralisation in L2 and more extreme L1 scores. In general, our hypothesis of lower emotional resonance in the L2 was confirmed because the results show that the scores obtained in the L2 tended to be less extreme than those obtained in the L1, especially for valence. However, we propose that such general statements are insufficient to describe what is actually happening in bilinguals' ratings and may even be misleading, because when a number of factors are taken into account, such as those we included (modality, grammatical category of words, concreteness), the variations are much more complex. Thus, some groups of words are more likely to obtain different emotional scores in the L2, both for valence and arousal, and in some cases the learners' L1 determines the type and direction of the variation. It has only been possible to reach this conclusion because a diverse but cohesive set of words drawn from the vocabulary to which learners are exposed in the early stages of language learning in formal education has been considered. Therefore, the results present a much more complex picture that requires a detailed analysis to truly understand what happens in a subjective appraisal of the affective content of words.

5.1. Valence and arousal in Spanish as L1 and L2

For valence, our results are consistent with the findings of previous research, whereby non-native speakers give fewer extreme scores than native speakers, with statistically significant differences for negative words and low-intensity words, especially in oral modality. Regarding positive words, our results do not completely coincide with those reported by Garrido and Prada (2018) or Imbault et al. (2020), because both studies show also a decrease in valence for positive words. Nevertheless, our results do agree with those obtained by Vélez-Uribe and Rosselli (2019), who only observed fewer extreme values in L1 for negative words but not for positive ones. So, negative words are consistently less negative in L2 but there are different patterns for positive words. Blanco Canales and Hernández Muñoz (2022) also find this last tendency for Brazilian and Greek Spanish students in non-immersion-context and relate this variability with socio-psychological factors like cultural context and attitudes and beliefs through L2. The better the attitudes and higher cultural-context, the higher the overestimation of positive words.

In terms of arousal, a tendency can also be observed in non-native speakers to assign lower arousal scores to negative words and high-intensity words, most evidently in oral stimuli. However, here we found one of the first differences between the Chinese and European groups. Whereas the Europeans' ratings of word intensity were similar to those of native speakers, Chinese learners assigned much lower ratings to high-intensity words than native speakers in both modalities.

In general, the group factor gave rise to two scenarios. On the one hand, significant differences often emerged between native speakers and both groups of non-native speakers (e.g., described valence scores). On the other, significant differences sometimes only occurred between the Chinese group and native speakers, with the Europeans falling somewhere in between and not reaching statistical significance (e.g., grammatical categories or intensity modulations).

In valence, moreover, all the factors observed in the cases of emotional charge and intensity interacted, revealing that the set of factors analysed exerted a strong effect on the participants' behaviour. The results for arousal were less salient, and in fact there were no broad interaction between the set of factors, but the more detailed analyses evidenced significant differences in the ratings. The disparity of our results for valence and arousal supports the notion that these two emotional variables assess different dimensions, in line with the findings reported by Citron (2012) and Citron et al. (2014). Whereas valence measures evaluative processes, arousal measures automatic processes. It is easier for speakers to assess the valence or affective charge of words, i.e., to judge whether an item is positive or negative, than it is to assess the degree of arousal, which refers to non-conscious processes that are much more difficult to conceptualise. Thus, the differences in valence are more marked and more extreme than in arousal for non-native speakers. In this respect, it should be recalled that Garrido and Prada (2018) only observed differences in arousal for taboo words.

One of the most significant contributions of our study is to have conducted a detailed analysis of arousal, the nature of which is

https://doi.org/10.1017/S136672892200061X Published online by Cambridge University Press

difficult to analyse as it does not tend to elicit extreme scores. Together with valence, this less-studied dimension shows internal variations that yield illuminating data in terms of understanding the development of the emotional component, such as the clear reduction in the emotional intensity of verbs and high-arousal negative words in the case of Chinese students. Arousal is also a key factor in determining differences between grammatical categories.

The fact that negative words show the greatest differences emphasises the importance of negative emotions for the individual. Our asymmetry between positive and negative valence supports the automatic vigilance hypothesis, which states that negative emotions are more salient because they ensure survival to a greater degree. This is not only evident in subjective rating tasks, but also in behavioural tasks: bilingual speakers tended to be more fluent when talking about positive rather than negative emotions, and in reading processes, L2 readers were more likely to neutralise negative but not positive words than L1 readers (Sheikh & Titone, 2016). These behaviours allow bilinguals to distance themselves from negative feelings and experiences (Dewaele & Pavlenko, 2002).

5.2. Written and oral modalities

Modality showed a markedly differentiated behaviour in valence and arousal. In the former, it proved a decisive factor in all analyses, revealing that written stimuli obtain higher scores than oral stimuli, rendering positive written words more positive and negative written words less negative. This was the same in all three groups, indicating that it is independent of whether the stimulus is in the participants' L1 or L2. These results differ from those reported by Vélez-Uribe and Rosselli (2019), who found no differences in valence between modalities neither L1 nor L2.

As it has been pointed out, valence, as a cognitive dimension, may be represented more consciously in our minds than arousal, a physiological dimension. It would be expected, therefore, that the valence would obtain higher scores in writing and the arousal in speaking. In addition, reading is the result of a formal learning process, whereas speaking is the result of a natural process. In this respect, it is reasonable to expect that valence (a cognitive process) is more closely aligned with writing than with speaking. Our results on valence in L1 and L2 support this view. However, we did not systematically obtain higher arousal in the oral items, perhaps because of the difficulty involved in consciously assessing this dimension (Citron et al., 2014).

Nevertheless, the specific cases in which we found significant differences between native and non-native speakers occurred in the oral modality. Although their results are not completely comparable, in the skin conductance study by Harris et al. (2003), native speakers showed higher arousal when the stimuli were auditory. The most important patterns indicated by our results show that bilinguals lose intensity in negative words and abstract words compared to native speakers (both in valence and arousal), and European learners and native speakers assign higher arousal to verbs than the Chinese group. This suggests that orality is the basis for different emotional perceptions between L1 and L2, and for the high resonance of negative words, abstract words and verbs.

5.3. Grammatical categories

It is well known that words belonging to different categories are processed differently and involve different brain regions (Citron, 2012). Emotion studies with native speakers using different grammatical categories have reported differences in processing time. Palazova, Mantwill, Sommer and Schacht (2001) have suggested that because nouns are easier to process and are acquired earlier, they may be processed more superficially without sustained attention to their emotional content, and consequently adjectives and verbs in their experiments were more affected by positive valence. Our results on arousal are only partially coincident with these findings. In valence, no differences emerged when rating nouns, adjectives or verbs. Nevertheless, our detailed analyses revealed that the Chinese group assigned lower valence scores than the European or native speakers. In the arousal results, grammatical categories emerged as a predictor of emotional intensity. Verbs were consistently rated with higher scores, nouns with medium scores and adjectives with the lowest scores in the European and native speaker groups in both modalities. This is an expected result if we consider that the grammatical category of verbs, especially action verbs, is the most embodied in sensorimotor processes and emotional interaction, according to the embodied approach to language (Barsalou, 1999). This notion is coincident with the findings reported by Bąk and Altarriba (2019), who also observed that verbs obtained higher arousal scores in native speakers of English and Polish.

It is worth noting that the Chinese learners did not show this grammatical category behaviour in arousal. In both modalities, the Chinese group's verb ratings fell far below that of nouns and adjectives. Again, our results suggest a reduced emotional charge of verbs in Chinese learners. This may be due to difficulties in acquiring the verbal paradigm, given that Chinese is an analytic language, whereas Europeans as a whole speak synthetic languages. So, in the study of grammatical categories the greatest differences between the non-native speaker groups emerged: whereas the Europeans behaved like native Spanish speakers, the Chinese demonstrated independent behaviour.

5.4. Concreteness

Concreteness was selected as a semantic measure for inclusion in the analyses because it has been widely demonstrated that the processing of concrete words differs from that of abstract words (Paivio, 1971). Furthermore, Citron (2012) highlights in neurocognitive studies that concreteness interacts with both valence and arousal. Our results in this respect were contrary to expectations, since we hypothesise that it might be easier for non-native speakers to attribute affective values to concrete words, which are more easily experienced by the senses. However, we found no interaction between concreteness and group, but we did observe a constant effect of concreteness on the independent variable, as abstract words were scored more positively. Similarly, non-concrete words obtained higher arousal scores than concrete words, a difference that was more pronounced in native speakers and in oral modality, suggesting that abstraction and orality exert an arousal effect in Spanish. Late bilinguals would be less sensitive to the arousal effect under these conditions. It would appear that emotional content is more salient in abstract words. This may be due to the clear interaction between emotional words and abstract words, which for a long time led to them being considered part of the same category (Pavlenko, 2008). This result requires further exploration in the future.

6. Conclusions

Research on the emotional resonance of words in bilingual speakers shows that this is not a simple question to answer and that

https://doi.org/10.1017/S136672892200061X Published online by Cambridge University Press

detailed analyses – rather than generalist approaches – are required that investigate the types of words in which emotional content is altered, as well as the different first languages, cultural backgrounds and learning contexts of second language speakers.

In this paper we have addressed at what extent new words in the bilingual lexicon incorporate affective meaning in the early stages of learning in immersion contexts. Explicit and subjective appraisal of the emotional connotations of words was analysed to contribute to the understanding of how vocabulary is integrated into learners' linguistic knowledge. We conclude that not only a general comparison between native and non-native rates is required, but also a depth consideration of semantic, grammatical, and emotional features of words, in both written and oral modalities. Our results show a decrease of emotional charge in non-natives, especially in negative words. Regarding the linguistic context of late bilinguals, differentials from L1 rates tend to be wider in the Chinese group than in the European one. Moreover, the divergence of our results for valence and arousal supports the notion that these two emotional variables assess different dimensions (e.g., grammatical category is only a predictor of variation in arousal).

Bilinguals' word ratings may depend on different emotional experiences in bilingual practices, such as socio-cultural or motivational elements. The processes we have described here are evidence of a dynamic interaction between lexical-semantic processes, affective processing, socio-cultural factors and educational practices (such as the mediation of literacy).

In addition, the alteration of a word's emotional content in late bilinguals may also be influenced by factors associated with incomplete vocabulary acquisition, a lower frequency of use because of a lower frequency of occurrence either in teaching materials or in everyday interactions, or a later age of acquisition (Imbault et al., 2020). This is particularly evident in research using words selected from native speakers' repertoires. In our case, we attempted to rectify this situation by using a corpus of beginner-level words, thus ensuring prior contact with the terms evaluated.

The present study is based on a componential and dimensional conceptualization of emotion (Scherer, 2001, 2013; Damasio, 1994). It assumes appraisal and embodiment as basic principles in the configuration and processing of emotionality in the lexicon of the languages. Its main aim is to explore how linguistic characteristics of words (semantic, grammatical and affective) also influence the variability of emotional dimension in the vocabulary of a second language, specially, early words. Although we have concluded that different mother tongues influence the change in affective load -Chinese and European students show clear linguistic typological differences, cultural contexts and discursive styles-, we cannot elucidate with our data how much depends on the mother tongue and its similarities with the L2 or on the sociocultural factors associated with each group. Investigating this issue exceeds the objectives of our study, but it should be noted that from studies on language and emotion in second languages from social anthropology, differentiated cultural patterns may contribute to creating particular emotional spaces (Wierzbicka, 1999). As a future researching line, it would be interesting to broaden the cultural factors considered, including a division of European participants as an unitary group, and analyse the results in the light of constructivist and anthropological theories of emotion, which consider the sociocultural context as determining (Averill, 1980; Karandashev, 2021; Ponsonnet, 2014). These approaches focus on the idea that differences in the structural linguistic organization of emotional space are due to different cultural patterns, which contribute to creating particular emotional spaces. In this way, it would be relevant to consider various aspects related to social characteristics, cultural representations, collective experiences, educational structures or attitudes and beliefs about the language (Blanco Canales & Hernández Muñoz, 2022; Caldwell-Harris, Staroselsky, Smashnaya & Vasilyeva, 2012; Caldwell-Harris, Tong, Lung & Poo, 2011).

To conclude, our results show that even in immersion contexts, late bilinguals reduce the emotional content of certain lexical items as regards both the affective charge and the intensity of words. In order for learners to incorporate affective, semantic and sensorimotor dimensions simultaneously, vocabulary learning must be reinforced with rich experiences that stimulate the creation of patterns in the autobiographical memory that involves emotion regulation systems. This occurs naturally in L1 acquisition, but not necessarily in an L2, which is often taught in the denaturalised L2 context of a classroom (Dewaele, 2004, 2010), particularly in the case of methodologies that place an emphasis on written presentation or learning mediated by writing. All this leads to "the development of "disembodied" words, used freely by speakers who do not experience their full impact" (Pavlenko, 2012, p. 421).

Acknowledgements. This research is part of the project "Comunicación, emoción e identidad en la adquisición y aprendizaje del español como segunda lengua", FFI2017-83166-C2-1-R, FEDER/Ministerio de Ciencia, Innovación y Universidades. Agencia Estatal de Investigación.

References

Altarriba, J (2003) Does cariño equal 'liking'? A theoretical approach to conceptual nonequivalence between languages. *International Journal of Bilingualism*, 305–322.

Altarriba, J and Basnight-Brown, DM (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. https://doi.org/10.1177%2F1367006910379261

Altarriba, J and Bauer, LM (2004) The distinctiveness of emotion concepts: A comparison between emotion, abstract, and concrete words. *American Journal of Psychology* **117**, 389–410. https://doi.org/10.2307/4149007

Averill, JR (1980) A constructivist view of emotion. In Plutchik, R and Kellerman, H (Eds.) *Emotion: Theory, research and experience*. New York: Academic Press, pp. 305–309.

Ayçiçeği-Dinn, A and Caldwell-Harris C L (2009) Emotion-memory effects in bilingual speakers: A levels-of-processing approach. *Bilingualism: Language and Cognition* **12**, 291–303. https://doi.org/10.1017/S1366728909990125

Bąk, H and Altarriba, J (2019) Feels different when s/he judges it – gender and grammar factors in emotion word evaluations. Poster presented at the International Society for Research on Emotions (ISRE), Amsterdam, The Netherlands, 12 July.

Barsalou, LW (1999) Perceptual symbol systems. *Behavioral and Brain Sciences* **22**, 577–660. https://doi.org/10.1017/s0140525×99002149.

Baumeister, J C, Foroni, F, Conrad, M, Rumiati, RI and Winkielman, P (2017). Embodiment and emotional memory in first vs. second language. *Frontiers in Psychology* **8**, Article 394. https://doi.org/10.3389/fpsyg.2017.00394

Blanco Canales A and Hernández Muñoz N (2022, forthcoming): The impact of language proficiency, cultural contact and attitudes on valence and arousal in Spanish as a second language. In Blanco A and Martín M (Eds.) *Emotion and identity in second language learning*. Oxford: Peter Lang.

Bradley, M M and Lang, PJ (1999) Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings. Technical Report C-1, The Center for Research in Psychophysiology, University of Florida.

https://doi.org/10.1017/S136672892200061X Published online by Cambridge University Press

Brosch, T, Pourtois, G and Sander, D (2010) The perception and categorization of emotional stimuli: A review. *Cognition and Emotion* **24**, 377–400. https://doi.org/10.1080/02699930902975754

Caldwell-Harris, CL (2014) Emotionality Differences Between a Native and Foreign Language: Theoretical Implications. *Frontiers in Psychology* **5**, Article 1055. https://doi.org/10.3389/fpsyg.2014.01055

Caldwell-Harris, CL and Ayçiçeği-Dinn, A (2009) Emotion and lying in a non-native language. *International Journal of Psychophysiology* **71**(3), 193–204. https://doi.org/10.1016/j.ijpsycho.2008.09.006

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** , 329–352.

Caldwell-Harris, CL, Staroselsky, M, Smashnaya, S and Vasilyeva, N (2012) Emotional resonances of bilinguals' two languages vary with age of arrival: The Russian-English bilingual experience in the U.S. In P Wilson (ed.), *Dynamicity in emotion concepts*. Frankfurt am Main, Germany: Peter Lang, 373–395.

Citron, FMM (2012) Neural correlates of written emotion word processing: A review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and Language* **122**, 211–226. https://doi.org/10.1016/j.bandl.2011.12.007

Citron, FMM, Weekes, BS and Ferstl, EC (2014) Arousal and emotional valence interact in written word recognition. *Language, Cognition and Neuroscience* **29**, 1257–1267. https://doi.org/10.1080/23273798.2014.897734

Colombetti, G (2005) Appraising Valence. *Journal of Consciousness Studies* **12**, 103–126.

Damasio, A (1994) *Descartes' Error: Emotion, Reason and the Human Brain*. New York: HarperCollins.

De Groot, AMB (2011) *Language and cognition in bilinguals and multilinguals: An introduction*. New York and Hove: Psychology Press.

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: Language and Cognition* **15**, 181–189. https://doi.org/10.1017/S1366728911000095

Dewaele, JM (2004) Perceived language dominance and language preference for emotional speech: The implications for attrition research. In Schmid, M, Köptke, B, Keijzer, M and Weilemar, L (eds.), *First language attrition: Interdisciplinary perspectives on methodological issues*. Amsterdam / Philadelphia: John Benjamins, pp. 81–104.

Dewaele, JM (2006) Expressing anger in multiple languages. In Pavlenko, A (ed.), *Bilingual minds: Emotional experience, expression, and representation*. Clevedon, UK: Multilingual Matters, pp. 118–151.

Dewaele, JM (2010) *Emotions in multiple languages*. Basingstoke, England: Palgrave Macmillan.

Dewaele, JM and Pavlenko, A (2002) Emotion vocabulary in interlanguage. *Language Learning* **52**, 265–324. https://doi.org/10.1111/0023-8333.00185

Dolan, RJ (2002) Emotion, Cognition, and Behavior. *Science* **298** , 1191–1194.

Eder, A, Hommel, B and de Houwer, J (2007) How distinctive is affective processing? On the implications of using cognitive paradigms to study affect and emotions. *Cognition & Emotion* **21**, 1137–1154. https://doi.org/10.1080/02699930701437386

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**, 353–369. https://doi.org/10.1177/1367006910379263

Eilola, TM, Havelka, J and Sharma, D (2007) Emotional activation in the first and second language. *Cognition and Emotion* **21**, 1064–1076. https://doi.org/10.1080/02699930601054109

Fan, L, Xu, Q, Wang, X, Zhang, F, Yang, Y and Liu, X (2016) Neural correlates of task-irrelevant first and second language emotion words: evidence from the emotional face–word Stroop task. *Frontiers in Psychology* **7**, 1672. https://doi.org/10.3389/fpsyg.2016.01672

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** , 760–785. https://doi.org/10.1080/02699930902985779

Fraga, I, Guasch, M, Haro, J, Padrón, I and Ferré, P (2018) EmoFinder: The meeting point for Spanish emotional words. *Behavior Research Methods* **50**, 84–93.

Garrido, MV and Prada, M (2018) Comparing the valence, emotionality and subjective familiarity of words in a first and a second language. *International Journal of Bilingual Education and Bilingualism* **24** , 275–291. https://doi.org/10.1080/13670050.2018.1456514

Harris, CL (2004) Bilingual speakers in the lab: Psychophysiological measures of emotional reactivity. *Journal of Multilingual and Multicultural Development* **25**, 223–47. https://doi.org/10.1080/01434630408666530

Harris, CL, Ayçiçeği, A and Gleason, JB (2003) Taboo Words and Reprimands Elicit Greater Autonomic Reactivity in a First Language Than in a Second Language. *Applied Psycholinguistics* **24**, 561–579. https://doi.org/10.1017/S0142716403000286

Harris, CL, Gleason, JB and Ayçiçeği, A (2006) When is a first language more emotional? Psychophysiological evidence from bilingual speakers. In Pavlenko, A (ed.) *Bilingual minds: Emotional experience, expression, and representation*. Clevedon, UK: Multilingual Matters, pp. 257–283.

Hernández, A (2009) Language switching in the bilingual brain: What's next? *Brain and Language* **109**, 133–140. https://doi.org/10.1016/j.bandl.2008.12.005

Hinojosa, JA, Moreno, EM and Ferré, P (2019) Affective neurolinguistics: Towards a framework for reconciling language and emotion. *Language, Cognition and Neurosicence* **49**, 1–27. https://doi.org/10.1080/23273798.2019.1620957 https://doi.org/10.3758/s13428-017-1006-3

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* **24** , 281–292. https://doi.org/10.1017/S1366728920000474

Instituto Cervantes (2006) *Plan curricular del Instituto Cervantes*. Madrid: Instituto Cervantes.

Karandashev, V (2021) *Cultural Models of Emotions*. Dordrecht: Springer.

Kazanas, SA and Altarriba, J (2016) Emotion Word Processing: Effects of Word Type and Valence in Spanish–English Bilinguals. *Journal of Psycholinguist Research* **45**, 395–406. https://doi.org/10.1007/s10936-015-9357-3

Kousta, ST, Vinson, DP and Vigliocco, G (2009) Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition* **112**, 473–481. https://doi.org/10.1016/j.cognition.2009.06.007

Kuperman, V, Estes, Z, Brysbaert, M and Warriner, AB (2014) Emotion and language: valence and arousal affect word recognition. *Journal of Experimental Psychology General* **143**, 1065–1081. https://doi.org/10.1037/a0035669

Lindquist, KA, MacCormack, JK and Shablack, H (2015) The role of language in emotion: Predictions from psychological constructionism. *Frontiers in Psychology* **6**, Article 444. https://doi.org/10.3389/fpsyg.2015.00444

Paivio, A (1971) *Imagery and verbal processes*. Hillsdale, N.J.: Erlbaum Associates.

Palazova, M, Mantwill, K, Sommer, W and Schacht, AK (2001) Are effects of emotion in single words non-lexical? Evidence from event-related brain potentials. *Neuropsychologia* **49**, 2766–75. https://doi.org/10.1016/j.neuropsychologia.2011.06.005.

Pavlenko, A (2008) Emotion and emotion-laden words in the bilingual brain. *Bilingualism: Language and Cognition* **11**, 147–164. https://doi.org/10.1017/S1366728908003283

Pavlenko, A (2012) Affective processing in bilingual speakers: Disembodied cognition? *International Journal of Psychology* **47**, 405–428. https://doi.org/10.1080/00207594.2012.743665.

Perfetti, C (2007) Reading Ability: Lexical Quality to Comprehension. *Scientific Studies of Reading* **11**, 357–383. https://doi.org/10.1080/10888430701530730

Ponsonnet, M (2014) *The Language of Emotions. The case of Dalabon (Australia)*. London: John Benjamins Publishing Company.

Richardson, JTE (2011) Eta squared and partial eta squared as measurements of effect size in educational research. *Educational Research Review* **6**, 135–147. https://doi.org/10.1016/j.edurev.2010.12.001

Robinson, C and Altarriba, J (2014) The interrelationship between emotion, cognition, and bilingualism. *Yearbook of the Poznan Linguistic Meeting* **1**, 103–117.

https://doi.org/10.1017/S136672892200061X Published online by Cambridge University Press

Scherer, KR (2001) Appraisal Considered as a Process of Multilevel Sequential Checking. In Scherer, KR, Schorr, A and Johnstone, T (Eds.) *Appraisal Processes in Emotion. Theory, Methods*. Oxford: Oxford University Press Research, pp. 92–120.

Scherer, KR (2013) Measuring the meaning of emotion words: A domain-specific componential approach. In Fontaine, JJR, Scherer K R y Soriano C (Eds.). *Components of emotional meaning*. Oxford: Oxford University Press, pp. 31–46.

Sheikh, NA and Titone, D (2016) The embodiment of emotional words in a second language: An eye-movement study. *Cognition and Emotion* **30**, 488–500. https://doi.org/10.3389/fpsyg.2017.00394

Soares, AP, Comesaña, M, Pinheiro, AP, Simões, A and Frade C (2011) The Adaptation of the Affective Norms for English Words (ANEW) for European Portuguese. *Behavior Research Methods* **44**, 256–269. https://doi.org/10.3758/s13428-011-0131-7

Soriano, C (2013) Linguistic theories of lexical meaning. In. Fontaine, JJR, Scherer, KR, y Soriano, C (eds.), *Components of emotional meaning: A sourcebook*. Oxford: Oxford University Press, pp. 63–82.

Stadthagen-González, H, Imbault, C, Pérez Sánchez, MA and Brysbaert, M (2017) Norms of Valence and Arousal for 14,031 Spanish Words. *Behavior Research Methods* **49**, 11–123. https://doi.org/10.3758/s13428-015-0700-2

Strauss, GP and Allen, DN (2008) Emotional Intensity and Categorisation Ratings for Emotional and Nonemotional Words. *Cognition and Emotion* **22**, 114–133. https://doi.org/10.1080/02699930701319154

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

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

Warriner, AB, Shore, DI, Schmidt, LA, Imbault, CL and Kuperman, V (2017) Sliding Into Happiness: A New Tool for Measuring Affective Responses to Words. *Canadian Journal of Experimental Psychology* **71**, 71–88 https://doi.org/10.1037/cep0000112.

Wierzbicka, A (1999) *Emotions Across Languages and Cultures: Diversity and Universals*. Cambridge: Cambridge University press.

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

https://doi.org/10.1017/S136672892200061X Published online by Cambridge University Press