Bilingualism: Language and Cognition 22 (1), 2019, 30–46 © Cambridge University Press 2017 doi:10.1017/S1366728917000517

The auditory and visual appraisal of emotion-related words in Spanish–English bilinguals*

IDALY VÉLEZ-URIBE
Department of Psychology, Florida Atlantic University, Davie, FL, USA
MÓNICA ROSSELLI
Department of Psychology, Florida Atlantic University, Davie, FL, USA

(Received: September 8, 2015; final revision received: August 1, 2017; accepted: August 22, 2017; first published online 5 October 2017)

Bilinguals experience emotions differently depending on which language they are speaking. Emotionally loaded words were expected to be appraised differently in the first versus the second language in Spanish–English bilinguals. Three categories of words (positive, negative, and taboo) were appraised in both languages in the visual and auditory sensory modalities. Positive word ratings were more positive in English than in Spanish. Negative words were judged as more negative in English than in Spanish. Taboo words were rated as more negative in Spanish than in English. Significant regression models were obtained for the visual and auditory positive words and auditory negative words with English and Spanish proficiency as the most significant predictors. Results support the view that there are differences in the appraisal of emotions in the two languages spoken by bilinguals; the direction of the difference depends on the emotion category of words, and it is influenced by language proficiency.

Keywords: emotion, bilinguals, Spanish, English, auditory, visual words

Introduction

Bilinguals often report a dissociation in the intensity of emotional verbal content between their first (L1) and second (L2) languages. Evidence indicates differences in the way speakers of two languages process emotional information in each of their languages. For instance, sequential bilinguals, who have learned two languages at different ages, self-report that their L2 is less emotional than their L1 (Dewaele, 2004, 2008, 2011, 2013). Similarly, the second language has been considered to provide bilinguals with more emotional distance (Bond & Lai, 1986; Dewaele & Costa, 2013).

Accordingly, emotion words may be processed differently than other types of words in L1 and L2 in bilingual speakers (Altarriba, 2008). Emotion words could refer to explicit affective states (i.e., happy), be associated with emotions (i.e., laughter) or induce an emotion (i.e., war) (Pavlenko, 2008, 2012); these words can be classified by their perceived emotional intensity and rated in terms of valence and arousal. Word valence ranges within a continuum from negative to positive (passing through neutral) depending on the level of pleasantness or unpleasantness attributed to the word, while arousal ranges from calm to excited (Bradley, Greenwald, Petry & Lang, 1992). There are databases of emotion words including norms of valence, arousal, and dominance in different languages such as in English (Bradley & Lang, 1999) and in Spanish (Redondo, Fraga, Padrón & Comesaña, 2007).

Two levels of processing of emotion words have been analyzed in bilingual speakers: conscious and unconscious. The conscious processing of emotion words has been studied by requiring overt analysis of the emotional strength of certain word categories. Dewaele's (2004) study collected information from 1,039 multilinguals on the perception of swear and taboo (S-T) words. Participants were asked in an online questionnaire whether these categories of words had the same emotional weight in each language. Results showed that the S-T words in the L1 are perceived to have much more emotional force than the same words in the L2. More recently, Dewaele (2016) found results in a different direction; multilingual speakers overestimated the offensiveness of swear words in their non-native languages. The author attributed this effect to the fear of inappropriate use of offensive words by language learners, who would, in turn, avoid using them as a result of insufficient experience with the newly learned language. These findings point to a particularly complex relationship between taboo words and their pragmatic application in second language learners. Dewaele (2004) collected information about an overall word category (swear and taboo words) whereas Dewaele (2016) included a set

* Our most sincere gratitude to Valeria Torres, Deven Christopher and AmberRose Reale for their editorial support.

Address for correspondence:
Mónica Rosselli, Department of Psychology, Charles E. Schmidt College of Science, Florida Atlantic University, 3200 College Ave, Davie, FL 33314, USA
mrossell@fau.edu

Supplementary material can be found online at https://doi.org/ 10.1017/S1366728917000517

of 30 words for which the participants indicated how offensive they found them on a 1–5 scale; which may have influenced the different direction of the result, due to the specificity of the stimuli.

To study the automatic or unconscious affective processing of words, a variety of research paradigms have included emotion words with different valence (positive, negative, neutral) and levels of arousal (high, moderate, low). Typically, the words are presented visually or aurally in each language with the assumption that the automatic processing will be reflected in emotional effects specific to each language. These effects include changes in memory recall (Anooshian & Hertel, 1994; Ayçiçeği & Harris, 2004; Ayçiçegi-Dinn & Caldwell-Harris, 2009; Ferré, García, Fraga, Sanchez-Casas & Molero, 2010), lexical decision tasks (Ponari, Rodríguez-Cuadrado, Vinson, Fox, Costa & Vigliocco, 2015), and facilitation or interference through congruency (Eilola, Havelka & Sharma, 2007; Sutton, Altarriba, Gianico & Basnight-Brown, 2007; Colbeck & Bowers, 2012; Altarriba & Basnight-Brown, 2010).

The differences in memory recall known as "emotion-memory effect" are reflected in the superior recall of emotion words compared to neutral words. Participants perform a task (such as rating the emotional intensity of words or making a lexical decision) and are later asked to recall as many words as possible from the task (Rubin & Friendly, 1986). In bilinguals, any difference in the processing of their two languages would be reflected in superior memory recall for emotion words in one language compared to the other.

Anooshian and Hertel (1994) studied the free recall of Spanish and English emotional and neutral words in 36 late Spanish–English bilinguals who had acquired fluency in their second language after eight years of age. Emotional words presented in the native language resulted in higher recall rates than neutral words. Ayçiçeği-Dinn and Harris (2004) evaluated differences between auditory and visual processing of positive, negative, neutral, and taboo words, and childhood reprimands; the resulting free recall and recognition rates in L1 and L2 were also assessed. Participants were Turkish–English bilinguals who had learned English after 12 years of age. Outcomes demonstrated no difference in the processing of visual words compared to auditory words. In L2, the recall of emotion words in all categories, except for the negative words, was significantly higher than the recall of neutral words. In L1, this effect was only observed for the recall of taboo and negative words. In the recognition condition, the L2 advantage was evident in positive, negative, and taboo words, whereas for L1 the advantage was only seen in taboo words.

Ayçiçegi-Dinn and Caldwell-Harris (2009) tested 59 Turkish native speakers with high proficiency in English, and implemented four tasks varying in level of processing: (a) rating the items for emotional intensity; (b) a shallow processing task (counting the number of letters containing a closed circle); (c) a translation task; and, (d) a word association task (provide as many word associates as possible in 10 seconds). The emotion-memory effect was equally strong for both languages when averaged over the four tasks, but it was restricted to the native language in the emotion intensity-rating task. On the letter-counting task, the emotion-memory effect was very similar in both languages and was the strongest of all four tasks, suggesting that automatic processing of emotion stimuli varies depending on the type of task.

In support of the equivalent emotion-memory effects in L1 and L2 are findings by Ferré et al. (2010). They tested memory for positive, negative, and neutral words in two groups of proficient Spanish–Catalan bilinguals. Participants had acquired their second language early in life, in an immersion context, and differed in their language dominance; 42 were dominant in Catalan and 33 were dominant in Spanish. These researchers also tested a group of 35 proficient Spanish–English bilinguals (late unbalanced bilinguals) who had learned the second language later in life but, before age 12, in an instructional setting. Results did not demonstrate differences in the emotion-memory effect between the two languages in early balanced bilinguals or late unbalanced bilinguals. The null differential effects of emotions in L1 and L2 in the Spanish–Catalan bilinguals could be interpreted as the result of the linguistic similarities between L1 and L2 (Catalan and Spanish). However, the fact that similar results were found for the Spanish–English late bilinguals might indicate that the emotion-memory effect is independent of this linguistic factor.

There seem to be two major implications for the results on the automatic emotion-memory effect of emotion words. First, in most cases, the emotion-memory effect in bilinguals seems to be more related to the characteristics of the stimuli (emotionality or novelty) than linguistic factors, including the age of acquisition of the languages in which it is presented. Second, the inconsistencies, such as the larger effect in L1 (Anooshian & Hertel, 1994), larger effect in L2 (Ayçiçeği & Harris, 2004), and equally strong effect (Ferré et al., 2010), may suggest that recall tasks are not the best-suited method to test the differences in the emotional experience of bilinguals in L1 and L2.

In addition to memory recall, automatic processing of emotion words in bilinguals has been evaluated using a simple lexical decision task including positive, negative, and neutral words in addition to pseudowords (Ponari et al., 2015). Highly proficient bilinguals have shown the same facilitation in processing emotionally valent words as native English speakers, regardless of their L1, age of English acquisition, or frequency and context of English use. Another method used to evaluate the automatic effect of emotion words in bilinguals is the emotional Stroop

https://doi.org/10.1017/S1366728917000517 Published online by Cambridge University Press

task which measures the interference elicited by emotional versus nonemotional stimuli. In bilingual speakers, it has been predicted that higher automaticity in L1 compared to L2 would result in greater interference in the L1 emotional Stroop. This interference would be reflected in higher RTs (reaction times) and therefore a larger emotional Stroop effect in L1 (Altarriba, 2008; Winskel, 2013).

Eilola et al. (2007) evaluated Finnish–English proficient bilinguals (who began acquiring English between 7 and 13 years of age) living in their native language environment, and found that negative and taboo words produced significantly longer RTs compared to neutral words when presented in the native language. However, the effect for negative and taboo words was not exclusive to L1, and extended to L2. Eilola and Havelka (2011) reported similar levels of emotional Stroop interference in both languages in early proficient Greek–English bilinguals who, differently from the bilingual sample in Eilola et al. (2007), were immersed in the English-speaking environment (L2) at the time of testing.

Similary, Sutton et al. (2007) found faster RTs in English (L2) than in Spanish (L1) in Spanish–English bilinguals who were highly proficient in both languages. However, though the emotional Stroop effect was present in both languages, it was greater in L2 than L1, possibly resulting from a shift in language dominance from L1 to L2. Similarly, Altarriba and Basnight-Brown (2010) applied the Affective Simon Task (a modification of the classic Simon task, including nouns and adjectives with emotional connotation) in two experiments, one with emotion-laden words and another with emotion words. In the bilingual group, the level of interference was larger in L2 than in L1, which the authors also attributed to a language dominance shift from L1 to L2.

Overall, the results from the affective Stroop (Sutton et al., 2007) and Simon task (Altarriba & Basnight-Brown, 2010) indicate that interference (larger emotional effect) occurs in both L1 and L2. The larger interference in L2 compared to L1 could be related to an increased automaticity in the second language, and a reflection of changes in proficiency levels and language dominance. Thus, if there is a shift in dominance from L1 to L2, the automaticity of emotion word processing in L2 increases, approaching the level of automaticity in L1. Other paradigms included in studies assessing "automatic affective processing" are rapid serial visual presentation (RSVP), and affective priming (Colbeck & Bowers, 2012; Kazanas & Altarriba, 2016).

Additionally, evidence of the differences in which bilinguals experience emotional content in L1 and L2 has been obtained through psychophysiological measures such as skin conductance responses (SCR). Harris, Ayçiçeği, and Gleason (2003) found that taboo words elicited the strongest SCRs in both languages among word categories. In a further study, Harris (2004) found that in both languages stimuli presented through the auditory modality elicited higher SCRs than stimuli presented through the visual channel for early but not for late learners. Late learners showed higher SCRs in L1 when exposed to childhood reprimands, but early learners did not show significant differences. These results seem to point to age of acquisition and perhaps proficiency in L2 as the most influential factors. Furthermore, Caldwell-Harris and Ayçiçeği-Dinn (2009) found stronger SCRs for items presented in the first language with the largest difference observed in the childhood reprimands category. Moreover, Eilola and Havelka (2011) reported greater reactivity to negative and taboo words in L1. Overall, autonomic reactivity seems to be greater for stimuli presented in L1 than in L2, although it appears to depend on the type of word, with taboo words provoking the strongest responses. The auditory modality seems to elicit higher reactivity. As emotional words are first acquired through the auditory modality while written emotional words (visual modality) are learned later in life, these results are consistent with the early learning hypothesis proposed by Caldwell-Harris and Ayçiçeği-Dinn (2009), which suggests that words or expressions learned early in life correspond with a stronger emotional association.

Other studies dealing with emotional language processing in bilinguals have used electrophysiological measures such as Event Related Potentials (ERPs) (Conrad, Recio & Jacobs, 2011; Opitz & Degner, 2012; Chen, Lin, Chen, Lu & Guo, 2015). The ERP components of interest include: the early posterior negativity (EPN), which seems to be enhanced by emotional valence particularly for positive words when compared to neutral words, and appears to indicate an attention shift toward words with emotional relevance at early processing stages; and the late positive complex (LPC), which seems to reflect a higher level semantic evaluation of words. Conrad et al. (2011) found that EPN and LPC effects can be elicited in both L1 and L2 in positive words in German–Spanish bilinguals. The differential effect was present for L1 versus L2 for negative words only in German native speakers, whereas native Spanish speakers showed no such effect; this could reflect a positivity bias induced by the native Spanish speakers being immersed in the second language culture. Pursuing this further, Opitz and Degner (2012) found no differences between L1 and L2 in the ERP processing of negative words in French–German bilinguals; participants showed greater EPNs for positive and negative words in both languages, with longer ERP latencies in L2, indicating delayed processing of L2 content. In a more recent study, Chen et al. (2015) studied Chinese–English bilinguals and found that L1 emotion words elicited larger negative deflections than neutral words in the 250–400 ms window at posterior electrode sites, resembling the classical EPN. Additionally in L1, emotion words generated smaller

https://doi.org/10.1017/S1366728917000517 Published online by Cambridge University Press

positive waves than neutral words starting at 500–800 ms window at centroparietal electrodes, consistent with the LPC. In L2, neutral words presented lower positivity than emotion words at the parietal sites during the 400–500 ms window.

Additional methods allowing for the examination of bilinguals during tasks involving emotion words have also been used. For example, functional neuroimaging studies indicate some differences in brain activation between L1 and L2 (Hsu, Jacobs & Conrad, 2015; Chen et al., 2015). Using functional magnetic resonance (fMRI), Hsu et al. (2015) observed the neural substrates associated with reading literary passages with emotional content in the first or the second language in late German–English bilinguals. Results showed that bilateral visual cortices were more active when reading in L1 (German) than in L2 (English) irrespective of emotion conditions. The authors found that the activation of the amygdala and the hippocampal and parahippocampal cortices were associated to emotional stimuli. In contrasting these activations in L1 and L2, the authors observed that activation in the amygdala was significant in the contrast [L1 > L2], while activation of the anterior insula was significant in the contrast [L2 > L1]. The authors concluded that reading literature in one's first language could induce a more intense emotional experience than reading it in one's second language. Furthermore, Chen et al. (2015) recorded fMRI during a lexical decision task in Chinese–English bilinguals. The task included words in the positive, neutral, and negative categories, as well as pseudowords. Participant classified these stimuli as words or non-words, resorting to the automatic processing of emotion content without incurring semantic processing costs. During this experiment, the mid-occipital gyrus showed decreased activity in response to positive words when compared to negative and neutral words in L1, but was inactive in response to emotional content in L2.

In summary, the connection between emotion processing and bilingualism has been analyzed using different research methodologies. Few studies have analyzed conscious affective processing, whereas the majority has included automatic affective processing of emotion words. Results from these multiple studies are not always consistent; while many have shown higher emotional responses for words in L1 compared to the same words in L2, others showed the opposite pattern or no emotion dissociations between the two languages. The inconsistent results across studies may be explained by variations in the experimental method and the characteristics of the sample's bilingualism.

The aim of the current study was to analyze the direct assessment of words through conscious appraisal of emotional valence in bilinguals. This matter has been reported by only one previous study to the authors' knowledge. Winskel (2013) tested the conscious affective processing of words, using emotionality ratings of 20 negative and 20 neutral Thai and English words in 57 Thai–English bilinguals; L1 was Thai and L2 English, which was learned as a foreign language in an instructional setting. Results showed that neutral words were rated lower in valence than negative words independent of the language in which the words were presented. Therefore, no differences between L1 (Thai) and L2 (English) were found in the conscious level of processing of negative emotion words. Winskel's (2013) findings can only generalize to late bilinguals who learned the second language in an academic framework, with very little exposure to the L2 culture, and who were not immersed in a bilingual environment; therefore, their underlying learning of emotion words in L2 took place almost exclusively in the educational context. These factors could all potentially impact the results. Moreover, participants in this study were all majoring in Languages or Linguistics, and therefore it is likely that they had a metalinguistic understanding of the foreign language that may have resulted in biased emotional consciousness with overestimation of emotional valence in L2.

Other studies have evaluated the perception of the emotional intensity of the expression "I love you" (Dewaele, 2008), and of swear and taboo words (Dewaele, 2004), as well as the overall appraisal of the emotional intensity of both languages, or more languages in the case of multilinguals (Dewaele, 2004, 2008, 2011, 2013).

The current study expanded on previous research by examining, in a sample of active Spanish–English bilinguals, the differences in L1 and L2 in the conscious appraisal of valence of 120 emotion-related words in three different emotion categories (positive, negative, and taboo) and in two sensory modalities (visual and auditory). In addition, the contribution of language proficiency and level of acculturation to the differences in appraisal between L1 and L2 were analyzed. Any differences in the appraisal of these word categories may reflect differences in the way emotions are experienced in each language.

The first hypothesis proposed a main effect of word category (positive, negative, and taboo) and of sensory modality (auditory and visual) on the differential L1-L2 valence scores. Overall, in all three valence categories, words were expected to be rated as more intense in L1 than in L2, independent of sensory modality. In addition, since aural words have shown greater emotional reactivity than written words in bilinguals (Harris et al., 2003; Harris, 2004), any differences in word valence scores in Spanish (L1) and English (L2) were expected to be significantly higher when presented aurally than when presented visually. Moreover, because taboo words have been shown to elicit higher emotional responses (Harris, 2004), they were expected to yield larger differential scores than negative or positive words.

https://doi.org/10.1017/S1366728917000517 Published online by Cambridge University Press

Additionally, the study aimed to test a multiple regression model including a set of relevant variables, which have been shown to influence the emotional experience of language in bilinguals. Based on this idea, our second hypothesis proposed that the difference in the emotional strength of words between L1 and L2 could be predicted by a regression model with seven independent variables: gender, age of acquisition of English, percent of life lived in the United States, English proficiency, Spanish proficiency, and Latino cultural and US cultural identity scores. More specifically, it was hypothesized that age of acquisition of L2 (English) would be one of the highest contributing variables to the regression models, as a lower age of acquisition might reflect a higher emotional competence in this language. It was expected that cultural identity and language proficiency would be significant predictors as well.

Dewaele and Pavlenko (2002) defined sociocultural competence as the ability to identify, categorize, perceive and engage in verbal and non-verbal behaviors similarly to other members of a particular language community. They analyzed the influence of this variable on the proportion of emotion words produced in narrative elicited by a three-minute film. Participants were two groups of Russian L1 speakers with English as their L2. One group had been exposed to the American culture (socioculturally competent L2 learners), and the other had never been exposed to the American culture. Results showed, despite the fact that participants with high sociocultural competence did not use more emotion vocabulary than Russians with low sociocultural competence, qualitative differences in the use of emotion language were seen in the elicited narratives.

De Leersnyder, Mesquita, and Kim (2011) analyzed "emotional acculturation" by measuring the similarity between the emotional patterns of an immigrant's culture and those of the host culture. They found that emotional concordance in Korean immigrants residing in the United States was highly correlated to the percentage of life spent in this country. These results showed that emotional acculturation is a dynamic process. Thus, it was expected in the current study that the level of acculturation of bilinguals would contribute to the variability of the emotional experience in their languages. As a consequence, it was predicted that Latino and US cultural identity scores would be significant contributors to the variance of valence scores when words were presented in the participant's native language (Spanish) and host language (English).

Another important predictor of disparity in emotionality between languages of bilinguals is the level of language proficiency (Dewaele, 2004). Dewaele and Pavlenko (2002) demonstrated that level of proficiency influenced emotion vocabulary in L2. In Belgian speakers of Dutch (L1) and French (L2) they found the level of proficiency in L2 influenced the proportion of emotion words produced in spontaneous conversation, with highly proficient speakers of French using more emotion word tokens. Also, Dewaele (2004) found age of acquisition (AoA), self-reported language proficiency, and frequency of use of the second language were significant predictors of the perception of emotional weight of swear and taboo (S-T) words in L2, and were significantly correlated with the emotional weight of affection expressions in L2 (Dewaele, 2008). Earlier AoA, higher level of proficiency, and higher frequency of use of L2 significantly predicted a stronger perception of S-T words in the second language. Dewaele's studies (2004, 2008) have a strong statistical power due to the large sample size, but they have the limitation of measuring the dependent variable with one response per participant in each language; having only one response per scale makes it difficult to assess the internal consistency of this scale. Although the independent measures of the current study are similar to those in Dewaele's (2004) research, the present study adds to his findings by including a highly reliable dependent measure (see reliability analyses reported in the results section): each participants' rating of 40 positive words, 40 negative words, and 40 taboo words in English and Spanish. Based on Dewaele's (2004, 2008) evidence, it was expected that higher language proficiency would be a significant predictor of intensity ratings of emotion words in the corresponding language.

Experiment I

Method

Participants

Participants were college students and members of the South Florida community. The initial sample included 149 Spanish–English bilingual participants; however, 48 were excluded from the analyses (11 reported other native languages, one had a data collection error, and 37 were simultaneous bilinguals). The final sample of 101 participants was composed of 32 males (32%) and 69 females (68%). Information on the characteristics of the sample is presented in Table 1. All participants (M_{age} = 21.82, SD = 5.60) acquired Spanish as their native language (that is, being exposed to it since birth) with 93% reporting it as the language used by parents to instruct and guide. English was acquired after Spanish (See table 1). Participants lived an average of 68.93% of their lives in the United States and had a high level of self-rated proficiency in English and Spanish. The sample was overall highly bicultural, with high scores in both Latino and US cultural identity (Table 2). All participants were active bilinguals, using Spanish and English on a daily basis (see Table 1). In

https://doi.org/10.1017/S1366728917000517 Published online by Cambridge University Press

Table 1. *Sample Characteristics.*

$N = 101$	Minimum	Maximum	M	SD
Age	18.00	44.00	21.82	5.60
Years of Education	12.00	27.00	14.89	2.77
Age of arrival to the US (n = 74)*	1.00	37.00	9.78	6.59
Total years in the US	0.50	32.00	14.69	5.23
Percent of life spent in the US	2.63	100.00	68.93	25.60
Age of acquisition of English	3.00	30.00	6.92	4.29

*Participants born in the US not included

Table 2. *Self-Reported Levels of Proficiency (LEAP-Q), Cultural Identity Subscales (AMAS-ZABB), and frequency of use.*

	Minimum	Maximum	M	SD
Level of proficiency				
Speaking English	4.00	10.00	9.39	1.08
Understanding English	6.00	10.00	9.55	0.80
Reading English	5.00	10.00	9.40	1.01
English Proficiency	5.33	10.00	9.45	0.86
Speaking Spanish	2.00	10.00	8.30	1.74
Understanding Spanish	4.00	10.00	9.02	1.26
Reading Spanish	2.00	10.00	7.94	1.80
Spanish Proficiency	3.33	10.00	8.39	1.46
Cultural Identity Subscales				
Latino Cultural Identity	10.00	24.00	20.73	3.34
US Cultural Identity	6.00	24.00	19.08	4.19
	Minimum	Maximum	M	SD
Daily Extent of Exposure to English				
When interacting with friends	3.00	10.00	9.23	1.38
When interacting with family	0.00	10.00	4.92	3.08
When watching TV	0.00	10.00	8.89	1.88
When listening to radio/music	1.00	10.00	8.23	1.98
When reading	0.00	10.00	9.11	1.79
In language-lab/self-instruction	0.00	10.00	8.38	3.05
Daily Extent of Exposure to Spanish				
When interacting with friends	0.00	10.00	3.80	2.83
When interacting with family	1.00	10.00	8.10	2.29
When watching TV	0.00	10.00	3.18	2.96
When listening to radio/music	0.00	10.00	4.70	3.04
When reading	0.00	10.00	2.58	2.51
In language-lab/self-instruction	0.00	10.00	2.30	2.70

regard to education, 62.4% of participants highest level of education was a high school degree, 30.7% had received an associate's degree, and 6.9% had achieved a bachelor's degree. Possibly due to an overall early age of acquisition and high level of schooling completed in English, it is reasonable that a large percentage of participants (80%) considered English as their dominant language. Additional information about the participants' reported experience of emotion-related terms in both languages, their language preferences for performing mathematical

https://doi.org/10.1017/S1366728917000517 Published online by Cambridge University Press

Table 3. *Frequency of Language Use for Mathematical Calculations*

Frequency of language use for Mathematical Calculations	Count	Percent
Native language		
Never	27	26.7%
Rarely	22	21.8%
Sometimes	25	24.8%
Frequently	13	12.9%
All the time	14	13.9%
Second language		
Never	2	2%
Rarely	5	5%
Sometimes	10	9.9%
Frequently	26	25.5%
All the time	58	57.4%

operations, and swearing preferences can be found in Tables 3 and 4.

Materials

Word rating task

The word rating task included 120 words rated on a valence scale, ranging from 1 to 9, (1 = most negative and 9 = most positive), replicating Bradley and Lang's (1994) procedure for rating words in terms of valence using the Self-Assessment Manikin (SAM, see Appendix 2, Supplementary Materials) as a rating system**.** Ratings obtained using SAM have shown to be valid in assessing emotions as they relate to stimuli, and are considered relatively culture and language bias free (Morris, 1995).

The task included four blocks of emotion words presented in two languages (English and Spanish) and two sensory modalities (visual and auditory). The words were selected from the ANEW Spanish and English databases to fit three categories, with 40 words per category: (a) Negative (valence $M = 2$; Range: 1–3); (b) Positive (valence $M = 8$; Range: 7–9), and (c) Taboo (20 words) (arousal $M = 7$; Range: 6–9). For the remaining 20 words of the taboo category, swear and offensive words were collected through an anonymous internet survey. Subsequently, they were screened and selected with the help of lab personnel of Hispanic origin from different nationalities, aiming to include terms that would apply to international standards of the English and Spanish languages (all words are depicted in Appendix 1, Supplementary Materials). Words in Spanish within positive and negative ranges were evaluated by a team of bilingual lab members for accuracy of translation and international applicability.

Table 4. *Emotion Language Use Preferences and Intensity of Languages*

Swearing language preference		
Native language		
Never	19	18.8%
Rarely	23	22.8%
Sometimes	29	28.7%
Frequently	13	12.9%
All the time	17	16.8%
Second language		
Never	9	8.9%
Rarely	6	5.9%
Sometimes	15	14.9%
Frequently	30	29.7%
All the time	41	40.6%
Emotional weight of swear and taboo words		
In native language		
Not strong	4	4%
Little	10	9.9%
Fairly	27	26.7%
Strong	23	20.8%
Very strong	35	34.7%
Not applicable	2	2%
In second language		
Not strong	3	3%
Little	12	11.9%
Fairly	47	46.5%
Strong	19	18.8%
Very strong	18	17.8%
Not applicable	2	2%
Is emotional weight of "I love you" different in the different languages		
Yes	40	39.6%
No	61	60.4%
Language in which it feels stronger ($n = 75$)		
English	16	21.3%
Spanish	59	78.7%
Difference in emotional significance for languages ($n = 97$)		
Yes	50	51.5%
No	47	48.5%

Language Experience and Proficiency Questionnaire (LEAP-Q)

The LEAP-Q assesses self-rated measures of proficiency (Marian, Blumenfeld & Kaushanskaya, 2007). LEAP-Q proficiency scores are divided in three components

https://doi.org/10.1017/S1366728917000517 Published online by Cambridge University Press

(speaking, understanding spoken language, and reading) on a 0 to 10 scale as follows: 0 = none, 1 = very low, 2 = low, 3 = fair, 4= slightly less than adequate, 5 = adequate, 6 = slightly more than adequate, 7 = good, 8 = very good, 9 = excellent, 10 = perfect. From this questionnaire, the average scores of the two subscales related to English and Spanish proficiency were used as a measure of self-reported proficiency for both languages.

Bilingual Emotional Experience Questionnaire Modified (BEQ-M)

The BEQ-M questionnaire was developed by integrating elements from the bilingualism and emotion questionnaire or BEQ (Dewaele & Pavlenko, 2001), as well as additional questions added to expand on the bilingual experience and the usage of emotion language by our participants. The final version included a total of 29 items and was named accordingly BEQ-M (Bilingualism and Emotion Questionnaire Modified); its objective is to collect information about how the participants experience emotions and emotional situations in their first and second languages.

Abbreviated Multidimensional Acculturation Scale (AMAS-ZABB)

The AMAS-ZABB is a 42-item scale shown to have good internal reliability and adequate concurrent validity on a sample of 246 participants composed of Central American immigrants and native Spanish speakers living in the United States. The analysis yielded three distinct factors: cultural identity, language competence, and cultural competence in 6 differentiated subscales (U.S. cultural identity - 6 items, Latino cultural identity - 6 items, English language - 9 items, Spanish language - 9 items, US cultural competence - 6 items, Latino cultural competence - 6 items). Scores are given on a 1–4 scale and calculated individually for each subscale (Zea, Asner-Self, Birman & Buki, 2003). For the current study, two subscales were considered: US cultural identity and Latino cultural identity.

Reliability of instruments

Reliability analyses for internal consistency by obtaining Cronbach's alpha (α) indices (using IBM SPSS Version 24, IBM Corp, 2014) were conducted for each scale. The English proficiency scale included three items and was highly reliable ($\alpha = .87$). The Spanish proficiency scale included three items and was highly reliable as well ($\alpha = .83$). Additionally, the cultural identity subscales included six items each and yielded high reliability indexes (US cultural identity, $\alpha = .87$; Latino cultural identity, $\alpha = .86$). The word ratings tasks were also analyzed for internal consistency and yielded high reliability scores for positive ($\alpha = .98$), negative ($\alpha = .98$), and taboo ($\alpha = .98$) words, including 160 items per category. For a detailed account of all reliability analyses, refer to Table 5.

Procedure

After giving informed consent, participants provided the information related to demographic and educational factors as well as other background information. Immediately following, the participants completed the computerized Word-Rating Task (WRT) designed on E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA) and presented on a 14-inch laptop. The task was divided into four counterbalanced sections: visual-English, visual-Spanish, auditory-English, and auditory-Spanish, with differently valenced items presented randomly across tasks. The experimenters instructed the participants in the language corresponding to each section. All visual stimuli were presented in Courier New, 18-point font in black with a white background for a duration of 2000 ms, and were followed immediately by the rating slide. All participants entered their ratings using the numeric pad on the laptop keyboard upon presentation of SAM (see Appendix 2, Supplementary Materials). All auditory stimuli were recorded in the voice of a female native speaker of the corresponding language using a neutral/natural tone of voice following the procedures in Harris et al. (2003). Words presented through the auditory modality were judged by a panel of three bilingual judges regarding prosody with the purpose of achieving a neutral tone across categories and between languages. Stimuli were presented while participants fixated on a cross on the center of the screen followed by the rating slide. Subsequently, participants were asked to complete the LEAP-Q, followed by the AMAS-ZABB and the BEQ-M. Lastly, a debriefing session was conducted in which the participant's questions were answered.

Results

The difference in valence scores between L1 and L2 was calculated by subtracting the valence ratings in English from those in Spanish. These scores are referred to as the differential valence scores (negative values indicate higher scores in English and positive values indicate higher scores in Spanish) (Table 6).

Differential valence scores were entered into a 2 (sensory modality: visual and auditory) x 3 (word category: positive, negative, and taboo) general linear model (GLM) analysis. Results revealed a significant main effect of word type, $F(1.79, 179.68) = 35.11$, $p < .001$, $p\eta^2 = .26$. Mauchly's test of sphericity was significant, $X^2(2) = 16.66$, $p < .001$, indicating a violation to

https://doi.org/10.1017/S1366728917000517 Published online by Cambridge University Press

Table 5. *Reliability Analyses for Word Rating Tasks, Acculturation, and Language Proficiency Subscale*

All items in English			α	*M*	*SD*	Items
English	Visual	Positive	0.94	322.93	30.99	40
		Negative	0.96	94.44	41.67	40
		Taboo	0.93	158.88	41.47	40
	Auditory	Positive	0.93	320.15	30.30	40
		Negative	0.95	99.08	39.45	40
		Taboo	0.94	151.12	41.59	40
All items in Spanish			α	*M*	*SD*	Items
Spanish	Visual	Positive	0.88	312.31	29.37	40
		Negative	0.93	105.83	36.80	40
		Taboo	0.93	150.06	43.77	40
	Auditory	Positive	0.92	314.83	32.04	40
		Negative	0.93	108.83	37.01	40
		Taboo	0.92	150.75	40.76	40
By language and word category			α	*M*	*SD*	Items
	English	Positive	0.97	642.85	60.31	80
		Negative	0.98	194.08	80.87	80
		Taboo	0.97	318.94	82.09	80
	Spanish	Positive	0.95	627.38	62.18	80
		Negative	0.96	215.50	71.48	80
		Taboo	0.96	305.34	81.63	80
By word category			α	*M*	*SD*	Items
		Positive	0.98	1267.26	119.60	160
		Negative	0.98	432.79	148.29	160
		Taboo	0.98	657.80	173.77	160
Acculturation and language proficiency subscales			α	*M*	*SD*	Items
		US Cultural Identity	0.87	19.08	4.19	6
		Latino Cultural Identity	0.86	20.73	3.34	6
		English Proficiency	0.87	28.32	2.60	3
		Spanish Proficiency	0.83	25.31	4.21	3

the sphericity assumption; therefore, Greenhouse-Geisser ($\xi = .89$) correction is reported using adjusted degrees of freedom. Pairwise comparisons showed significant differences between the differential valence scores for positive and negative words ($p < .001$) and for negative and taboo words ($p < .001$). While positive words and taboo words had higher values in English, negative words had higher values in Spanish (less negative). Contrary to what was expected, no main effect of sensory modality, $F(1, 89, 179.67) = 0.89$, $p = .35$, $p\eta^2 = .01$, was found, showing that the total differential valence sore for auditory stimuli ($M = -.13$; $SD = .028$) was very similar to that same total score in the visual modality ($M = -.11$; $SD = .031$). The interaction between sensory modality and word category was not significant, $F(1.86, 185.9) = 2.19$, $p = .117$, $p\eta^2 = .031$.

Due to the unbalanced gender distribution in the sample, the analyses were repeated entering gender as a covariate. It is important to mention that previous studies present a similar uneven gender distribution. For example: Ayçiçeği-Dinn and Caldwell-Harris (2009, 2009) – 88% females; Ferré et al. (2010) - 81% females; Ponari et al. (2015) – 76% females; Eilola et al. (2007) - 85% females; Eilola and Havelka (2011) – 71% females; Colbeck and Bowers (2012) – 75% females. Results did not change and no significant gender effect, $F(1,99) = 2.98$,

https://doi.org/10.1017/S1366728917000517 Published online by Cambridge University Press

Table 6. *Means and Standard Deviations for Valence Ratings: Raw and Subtracted (differential valence scores)*

	Raw Scores					
	Positive		Negative		Taboo	
	M	*SD*	*M*	*SD*	*M*	*SD*
English						
Visual	8.08	.76	2.31	.99	3.86	1.10
Auditory	8.03	.75	2.38	.97	3.85	1.08
Total	8.05	.75	2.34	.98	3.86	1.09
Spanish						
Visual	7.77	.77	2.60	.90	3.57	.98
Auditory	7.78	.77	2.59	.91	3.49	1.01
Total	7.78	.77	2.60	.91	3.53	.99
	Subtracted Scores					
	Positive		Negative		Taboo	
	M	*SD*	*M*	*SD*	*M*	*SD*
Visual	−.31	.64	.29	.52	−.30	.59
Auditory	−.25	.55	.21	.48	−.36	.59
Total	−.28	.60	.25	.50	−.33	.59

$p = .09$, $p\eta^2 = .03$, or interaction between gender and word category, $F(2, 98) = .18$, $p = .184$, $p\eta^2 = .004$, was found.

To further analyze the relationship of language (English and Spanish) and sensory modality (visual and auditory) within each word type (positive, negative, and taboo), three 2x2 GLM analyses were conducted, utilizing raw valence ratings as dependent variables on each word category. Results revealed a significant main effect of language on positive words, $F(1, 100) = 23.81$, $p < .001$, $p\eta^2 = .19$, with English words in both visual and auditory modalities being evaluated higher (more positive) than in Spanish words in both modalities. However, no significant main effect of sensory channel, $F(1, 100) = .916$, $p = .341$, $p\eta^2 = .009$, nor interaction between language and sensory channel, $F(1, 100) = 1.56$, $p = .214$, $p\eta^2 = .015$, were detected. When gender was entered into the analysis, no significant main effect was found, $F(1, 99) = 2.33$, $p = .130$, $p\eta^2 = .023$, nor interaction between gender and language, $F(1, 99) = .340$, $p = .561$, $p\eta^2 = .003$.

The 2x2 GLM analysis of negative words revealed a significant main effect of language, $F(1, 100) = 36.33$, $p < .001$, $p\eta^2 = .267$, with no significant effect of sensory channel, $F(1, 100) = 2.87$, $p = .093$, $p\eta^2 = .028$, and a nonsignificant interaction between language and sensory channel, $F(1, 100) = 2.17$, $p = .144$, $p\eta^2 = .021$, in which the English means were overall lower (more negative) than the Spanish means. However, when gender was entered as a covariate, a significant effect of gender was found, $F(1, 99) = 9.65$, $p = .002$, $p\eta^2 = .089$ with the disappearance of the language effect on negative words, $F(1,99) = .561$, $p = .456$, $p\eta^2 = .006$.

To have a better understanding of the gender effect over negative words, the male and female means were compared using a univariate GLM. Total English means (including visual and auditory) were significantly higher, $F(1, 99) = 9.57$; $p = .003$, $p\eta^2 = .088$ for males ($M = 2.76$, $SD = .16$) compared to females ($M = 2.14$, $SD = 11$). In addition, the total Spanish means, including visual and auditory, were significantly higher in males ($M = 2.95$; $SD = .15$) than in females ($M = 2.42$; $SD = .18$), $F(1,99) = 8.60$; $p = .004$, $p\eta^2 = .08$. However, the differential valence score between Spanish and English negative words (visual and auditory combined) was not significantly different between males ($M = .19$ $SD = .76$) and females ($M = .27$ $SD = .05$), $F(1, 99) = .687$; $p = .409$, $p\eta^2 = .007$.

Moreover, the 2x2 GLM analysis of the taboo word category yielded a significant effect of language, $F(1, 100) = 36.33$, $p < .001$, $p\eta^2 = .267$, with no significant effect of sensory modality, $F(1,100) = 2.87$, $p = .093$, $p\eta^2 = .028$, nor significant interaction between language and sensory modality, $F(1, 100) = 2.17$, $p = .144$, $p\eta^2 = .021$. The overall means in Spanish (visual and auditory) were lower (more negative) than those visual and auditory words in English. When gender was entered as a covariate, a significant effect was found, $F(1, 99) = 12.36$, $p = .001$, $p\eta^2 = .11$, with no significant interaction between gender

https://doi.org/10.1017/S1366728917000517 Published online by Cambridge University Press

and word category, $F(1, 99) = 1.50$, $p = .222$, $p\eta^2 = .015$, and with the main word category effects remaining significant over taboo words, $F(1, 99) = 7.80$, $p = .006$, $p\eta^2 = .073$.

To have a better understanding of the gender effect over taboo words, the male and female means were compared using a univariate GLM. Total English means (including visual and auditory) were significantly higher, $F(1, 99) = 12.66$; $p = .001$, $p\eta^2 = .113$ for males ($M = 4.38$, $SD = .87$) compared to females ($M = 3.61$, $SD = 1.08$). In addition, the total Spanish means, including visual and auditory, were significantly higher in males ($M = 3.96$; $SD = .16$) than in females ($M = 3.32$; $SD = .11$), $F(1, 99) = 10.15$; $p = .002$, $p\eta^2 = .093$. However, the differential valence score between Spanish and English taboo words (visual and auditory together) was not significantly different between males ($M = -.42$ $SD = .64$) and females ($M = -.28$ $SD = .49$), $F(1, 99) = 1.50$; $p = .222$, $p\eta^2 = .001$.

The second hypothesis regarding the association of demographic, bilingual, and cultural variables with emotion words in the sample was tested through six multiple regression analyses, which were conducted using a set of variables as predictors of the differential valence scores. The following predictors were included: (1) Age of acquisition of English; (2) Gender; (3) Percent of life lived in the United States; (4) Self-rated English proficiency; (5) Self-rated Spanish proficiency; (6) Latino cultural identity score; and, (7) US cultural identity score. The dependent variables were the differences in word ratings (Spanish minus English) in both sensory modalities (visual and auditory) for each of the three word categories (positive, negative, and taboo) which resulted in six dependent variables (Table 7). Bivariate analyses of the relationships between regression variables (regression results are reported further in this section) resulted in several significant correlations (two-tailed) (Table 8). The alpha level was adjusted for multiple comparisons to .008, corresponding to 66 correlations obtained in the analyses.

Gender was not a significant predictor in any of the regression analyses. The first regression model for visual positive words was significant, $F(7, 92) = 3.50$, $p = .002$, $f^2 = .21$, accounting for 21% of the variance in scores. One of the predictors had significant weight over the model: Spanish proficiency, $\beta = .28$, $t(100) = 2.60$, $p = .01$. In the visual negative category, the model was not significant, $F(7, 92) = 1.44$, $p = .197$, $f^2 = .09$, accounting for 9% of the variance in the visual negative differential scores. One significant predictor resulted from the analysis: English proficiency, $\beta = .27$, $t(100) = 2.16$, $p = .03$. The model was unsuccessful in predicting the differential scores for the visual taboo category, $F(7, 92) = .90$, $p = .50$, $f^2 = .06$.

In the auditory modality, the regression model was significant for the positive category, $F(7, 92) = 4.38$, $p < .001$, $f^2 = .25$, accounting for 25% of the variance

Table 7. *Summary of Multiple Regression Analyses on Differential Scores*

Differential valence scores	Visual Positive			Visual Negative			Visual Taboo			Auditory Positive			Auditory Negative			Auditory Taboo		
Predictors	*B*	*SE B*	*β*	*B*	*SE B*	*β*	*B*	*SE B*	*β*	*B*	*SE B*	*β*	*B*	*SE B*	*β*	*B*	*SE B*	*β*
Gender	.00	.14	.00	.00	.12	.00	.10	.14	.08	−.04	.11	−.04	.18	.10	.18	.12	.14	.09
AoA of English	.00	.02	.00	.08	.02	.07	.02	.02	.12	.00	.02	.01	−.03	.01	−.24	.00	.02	.01
Percent of Life in US	−.01	.00	−.19	.00	.00	.09	−.00	.00	−.12	−.00	.00	−.13	.00	.00	.23	−.01	.00	−.03
English Proficiency	−.14	.09	−.19	.16	.08	.27*	−.00	.09	.00	−.15	.07	−.24*	.04	.07	.07	.00	.09	.01
Spanish Proficiency	.12	.05	.28*	−.01	.04	−.02	−.02	.05	−.05	.13	.04	.36**	.01	.04	.04	.03	.05	−.08
Latino Cultural Identity	−.02	.02	−.12	.02	.02	.15	.02	.02	.13	−.02	.02	−.10	−.01	.02	−.010	−.01	.02	−.06
US Cultural Identity	−.00	.01	−.01	.00	.01	.02	.00	.02	−.01	−.01	.01	.98	−.00	.01	−.03	−.01	.02	−.07
R^2		.21			.1			.06			.25			.19			.03	
F		3.51 **			1.44			.85			4.38**			3.13*			.33	

https://doi.org/10.1017/S1366728917000517 Published online by Cambridge University Press

Table 8. *Correlation Matrix: Correlations between Regression Variables*

	Age of Acquisition of English	Percent of life spent in US	English Proficiency	Spanish Proficiency	Latino Cultural Identity	US Cultural Identity	Visual Positive	Visual Negative	Visual Taboo	Auditory Positive	Auditory Negative	Auditory Taboo
Age of Acquisition of English	1	–.49**	–.55**	.21*	–.11	–.08	.27**	–.14	–.16	.28**	–.36**	.05
Percent of life spent in US		1	.44**	–.28**	.04	.18	–.35**	.19	–19	–.29**	.32**	–.08
English Proficiency			1	–.04	–.02	.15	–.28**	.27*	–.13	–.30**	.29**	–.02
Spanish Proficiency				1	.32**	–.03	.30**	.01	.06	.36**	–.07	.08
Latino Cultural Identity					1	–.06	–.04	.13	.10	–.01	–.01	–.02
US Cultural Identity						1	–.06	.07	–.04	.05	.05	–.07
Visual Positive							1	–.44**	.05	.82**	–.53**	.06
Visual Negative								1	.2	–.34**	.43**	.02
Visual Taboo									1	–.01	.01	.72**
Auditory Positive										1	–.49**	–.04
Auditory Negative											1	.18
Auditory Taboo												1

*. P < .008, Adjusted for multiple comparisons

in scores. Two of the predictors had significant weight over the model: English proficiency, $\beta = -.24$, $t(100) = -2.11$, $p = .037$, and Spanish proficiency, $\beta = .36$, $t(100) = 3.46$, $p = .001$, suggesting that lower English proficiency scores and higher Spanish proficiency scores predict larger valence differences between L1 and L2 in auditory positive words. On the auditory negative category, the model was also significant, $F(7, 92) = 3.13$, $p = .005$, $f^2 = .19$, accounting for 19% of the model variance. However, no significant predictors emerged from the analysis, with only age of acquisition of English approaching significance, $\beta = -.24$, $t(100) = -1.91$, $p = .058$. These results showed that the older the age in which participants began acquiring English, the larger the differences between L1 and L2 in the perception of auditory negative words. On the auditory taboo category, $F(7, 92) = .334$, $p = .937$, $f^2 = .02$, the results were not significant. These results indicated that the included variables were unable to predict differential scores between L1 and L2 in the perception of taboo words. Table 8 summarizes the results of all the regression analyses.

Experiment II

Since Experiment I did not include neutral words, the question of whether the language effect is restricted to emotional words or whether it extends to neutral words remained to be explored. To overcome this limitation, an additional experiment was performed using a small bilingual sample ($N = 24$) with similar demographic characteristics to the sample from Experiment I.

Method

Participants

The initial sample consisted of 24 participants (18 females) recruited following the same procedures used in the previous experiment. This sample had been exposed to Spanish since birth and had a mean age of acquisition of English of 5.17 ($SD = 4.12$), having learned English either simultaneously ($N = 7$) or later in life ($N = 17$). To equate the characteristics of this sample to those of the sample in Experiment I, the seven simultaneous bilinguals were excluded from the analyses. The final sample included 17 bilinguals (12 females), with a mean age of 25.56 years ($SD = 5.15$) and mean years of education of 17.50 ($SD = 5.34$). All participants identified Spanish as their native language and had a high level of proficiency (as measured by the LEAP-Q) in both English ($M = 9.62$, $SD = .54$) and Spanish ($M = 8.89$, $SD = 1.20$), with 76.5% of participants selecting English as the dominant language. Participants' scores on the Latino cultural identity scale ($M = 21.56$, $SD = 4.76$) and the US cultural identity

https://doi.org/10.1017/S1366728917000517 Published online by Cambridge University Press

Table 9. *Experiment II: Means and Standard Deviations for Valence Ratings: Raw and Subtracted*

	Positive		Negative		Neutral		Taboo	
English	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Visual	8.01	.64	2.44	.91	5.00	.14	3.60	.89
Auditory	7.95	.75	2.35	.90	4.94	.19	3.60	1.00
Total	7.98	.69	2.39	.90	4.96	.22	3.60	.94
Spanish								
Visual	7.85	.85	2.53	.82	5.02	.18	3.30	.78
Auditory	7.52	.72	2.64	.74	5.01	.22	3.45	.67
Total	7.68	.78	2.58	.78	5.02	.28	3.37	.72
Subtracted Scores								
	Positive		Negative		Neutral		Taboo	
	M	*SD*	*M*	*SD*	*M*	*SD*	M	SD
Visual	−.16	.72	.09	.23	.08	.21	−30	.54
Auditory	−.43	.50	.28	.37	.07	.15	−15	.62
Total	−.29	.61	.19	.30	.07	.18	−.22	.58

($M = 19.87$, $SD = 5.17$) were similar to scores in the sample from Experiment I.

Materials

The materials included a word rating task that replicated the one used on Experiment I with the addition of a set of 40 neutral words (valence $M = 5$; Range: 4–6) selected from ANEW (refer to Appendix 1, Supplementary Materials) and following the procedures described in Experiment I to evaluate translation accuracy and international generalizability. Neutral words in the auditory modality were recorded and evaluated following the procedures outlined in Experiment I. Additional information about the participants' linguistic and cultural background was collected through the same instruments described in Experiment I: LEAP-Q, BEQ-M and AMAS-ZABB (for a detailed description refer to the materials section under Experiment I).

Procedure

The procedure replicated the methodology from Experiment I.

Results

To study the relationship of language (English and Spanish) and sensory modality (visual and auditory) within each word type (positive, negative, taboo, and neutral), a 2x2 (language X sensory modality) GLM analysis was performed on each word category (positive, negative, taboo, and neutral) including raw valence ratings as dependent variables. Table 9 depicts both raw valence ratings and subtracted scores, which were obtained following the procedure described in Experiment I.

The findings were consistent with those of Experiment I, with significant differences between ratings in both languages for the positive, negative, and taboo word categories but not for neutral words, supporting the indication that differences between languages are restricted to emotion words.

In the positive word category, results revealed a non-significant main effect of sensory modality, $F(1, 16) = 4.24$, $p = .06$, $p\eta^2 = .21$, but a significant main effect of language, $F(1, 16) = 8.55$, $p = .010$, $p\eta^2 = .349$, indicating that English ratings for visual and auditory positive words ($M = 7.98$, $SD = .69$) were higher than the visual and auditory ratings for the corresponding Spanish words ($M = 7.62$, $SD = .69$); this main effect was independent of the sensory modality, as indicated by the non-significant interaction between language and sensory modality, $F(1, 16) = 1.427$, $p = .25$, $p\eta^2 = .082$.

Furthermore, the 2x2 GLM analysis of negative words revealed a significant main effect of language, $F(1, 16) = 8.48$, $p = .010$, $p\eta^2 = .346$, with no significant effect of sensory modality, $F(1, 16) = .45$, $p = .834$, $p\eta^2 = .003$. Higher valences for negative words were observed in Spanish ($M = 2.58$, $SD = .78$) than in English ($M = 2.39$, $SD = .90$. Consistent with the results of Experiment I, the English visual ($M = 2.58$, $SD = .83$) and English auditory ($M = 2.35$, $SD = .90$) words were lower (more negative)

https://doi.org/10.1017/S1366728917000517 Published online by Cambridge University Press

than those in the Spanish visual and the Spanish auditory word categories. There was a significant interaction between language and sensory modality, $F(1,16) = 6.49$, $p = .021$, $p\eta^2 = .289$, showing that higher negative valence for negative words in Spanish compared to negative words in English was greater for words presented aurally. This interaction was not significant in Experiment 1.

Similarly, the analysis of taboo words revealed a non-significant main effect of sensory modality, $F(1, 16) = .378$, $p = .070$, $p\eta^2 = .19$, and a significant main effect of language, $F(1, 16) = 4.57$, $p = .048$, $p\eta^2 = .222$ with English taboo words rated higher in valence ($M = 3.60$, $SD = .94$) than in Spanish ($M = 3.37$; $SD = .17$). No significant interaction between language and sensory modality were found, $F(1, 16) = .938$ $p = .347$, $p\eta^2 = .057$.

The analysis of neutral words revealed non-significant main effects for sensory modality, $F(1, 16) = 1.22$, $p = .286$, $p\eta^2 = .071$, or language, $F(1, 16) = 2.04$, $p = .172$, $p\eta^2 = .113$, with a non-significant interaction between language and sensory modality, $F(1, 16) = 3.796$ $p = .070$, $p\eta^2 = .192$.

Findings from this experiment confirmed differences between the two languages in the processing of emotion words, and the absence of this effect in neutral words, which contributes to the validation of the results obtained in Experiment 1.

General discussion

The main aim of the current study was to compare the way in which bilinguals appraise emotion words in both languages and in two sensory modalities with three word categories. In support of the first hypothesis, the results show the conscious perception of emotion words in active Spanish–English bilinguals is different for Spanish (L1) and English (L2) depending on the word category (positive, negative, or taboo). The magnitude of the difference in valence between L1 and L2 was significant for the three emotion categories but in different directions; while positive and taboo words had higher values in English (L2), negative words had higher values in Spanish (L1). These results indicate that positive words were perceived as more pleasant in English than in Spanish, and taboo words were perceived to be less negative in English than in Spanish. Negative words, on the other hand, were perceived as less negative in Spanish than in English.

Our results support previous findings indicating that emotionality effects of language are not restricted to L1, but are also observed in L2 (Altarriba & Basnight-Brown, 2010; Ayçiçeği & Harris, 2004; Ayçiçeği-Dinn & Caldwell-Harris, 2009). In addition, they support the finding that bilinguals perceive positive and negative words more strongly in L2 compared to L1 (Ayçiçeği & Harris, 2004). However, Ayçiçeği and Harris (2004) found a stronger emotion-memory effect in L2 for words carrying negative associations. Similarly, our participants perceived negative words as more negative in English (L2) than in Spanish (L1). These findings can be related, even though the samples of the two studies had different characteristics. Our sample was very proficient in both languages, lived in a bicultural environment, and actively used both languages on a day-to-day basis. In contrast, Ayçiçeği-Dinn and Caldwell-Harris' (2009) sample included Turkish–English bilinguals who had learned English (L2) in an instructional setting with little exposure to emotion words in a natural environment. It could be possible that the novelty of the stimuli during the experimental condition in Ayçiçeği-Dinn and Caldwell-Harris' (2009) study played a confounding role in the recall process, resulting in stronger emotion-memory effect while in the present study, a language dominance effect could have taken place. These comparisons should be made cautiously, given the different methodologies under comparison. Our study used conscious affective assessment of words whereas Ayçiçeği-Dinn and Caldwell-Harris' (2009) resorted to unconscious processing of emotional stimuli.

Our findings are inconsistent with those in Winskel's (2013) study, who did not find differences in the emotional ratings of negative words between L1 and L2. There are methodological differences between the two studies. Winskel's (2013) experiment included bilinguals with limited exposure to the L2 culture and who had learned L2 almost exclusively in an academic environment. The present study analyzed active bilinguals who had learned L2 both in an educational setting and in daily interactions with peers. These differences in the characteristics of the sample could have potentially influenced the results.

Our results propose that the bilingual participants in our sample felt the emotional force of positive and negative words more strongly in English (L2) than in Spanish (L1), which could be due to frequency of use of the corresponding language (Dewaele, 2004, 2008), and could be related to the overall English dominance in the sample (Altarriba & Basnight-Brown, 2010; Dewaele, 2008). In the opposite direction, overall ratings of taboo words were lower (more negative) in Spanish (L1) than in English (L2). These results are consistent with previous findings of higher reactivity in SCRs to taboo words in L1 than in L2 (Eilola & Havelka, 2011; Harris et al., 2003). The greater sensitivity to taboo words in L1 could be related to the frequency of use, with most participants reporting the use of swear or taboo words in English (62.3%) rather than Spanish (38.6%), possibly increasing the sensitivity to those words in Spanish. In addition, considering that taboo words are more frequently used when interacting with peers than with family, and that our participants reported more frequent use of English with

https://doi.org/10.1017/S1366728917000517 Published online by Cambridge University Press

their peers and Spanish with their families, the higher sensitivity to taboo words in Spanish might be influenced by this interaction.

Our findings are in partial disagreement with studies which did not find differences between the automatic perception of emotion in L1 and L2 in early proficient bilinguals (Ferré et al., 2010; Ponari et al., 2015). Studies on language interference such as Altarriba and Basnight-Brown (2010) and Colbeck and Bowers (2012) have found that bilinguals with higher dominance in a second language show fewer differences between the way emotions are processed in the two languages. These individuals demonstrate an increased automaticity of emotion word processing in L2, making it very similar or even higher to the level of automaticity in L1, which seems to be consistent with our results in the positive and negative categories. Even though all participants in the current study had acquired Spanish as a first language, most were dominant in English and received most of their schooling in English. It is possible this aspect resulted in stronger ratings in English than in Spanish.

Overall, there seem to be differences in the findings of emotional processing of words in bilinguals depending on the specific methodologies used: automatic processing paradigms versus conscious processing. This raises questions about which methodology is better suited to reflect the differences in which bilingual individuals process emotion content in both languages. While self-report and single question accounts (Dewaele, 2004, 2008, 2011, 2013) and SCRs (Harris et al., 2003; Caldwell-Harris & Ayçiçeği-Dinn, 2009) are quite consistent in indicating decreased emotional sensitivity in L2, other findings using automatic processing paradigms are not as consistent (Ponari et al., 2015; Eilola et al., 2007; Sutton et al., 2007; Colbeck & Bowers, 2012; Altarriba & Basnight-Brown, 2010) and often vary across studies in the directionality of the effect (L2 or L1 as being more emotional), even between studies utilizing similar tasks (Anooshian & Hertel, 1994; Ayçiçeği & Harris, 2004; Ayçiçegi-Dinn & Caldwell-Harris, 2009; Ferré et al., 2010). Results from the present study showed that the conscious appraisal of emotions in active proficient bilinguals is different in L1 than L2, even though both languages are sensitive to the perception of emotions and are influenced by level of proficiency. The current findings are not in line with results obtained through the indirect assessment of emotions using memory recall or interference paradigms. Future research needs to compare the concurrent validity of these different methodologies.

Contrary to expectations and consistent with the findings by Ayçiçeği and Harris (2004), this study found neither a sensory modality main effect, nor a significant sensory modality x word category interaction. These nonsignificant findings suggest that the language effect over the emotional perception of words extends equally to visual and auditory stimuli when conscious affective paradigms are used in early and active bilinguals. It adds to the same findings provided by Ayçiçeği and Harris (2004) using automatic affective processing in late bilinguals.

The second hypothesis proposed the selected set of independent variables would significantly predict the differential valence scores in two sensory modalities and three word categories, and expected the age of acquisition of English, and both Latino and US cultural identity scores to be the highest contributing variables to the variance in scores in all categories. The regression model was fit for predicting scores for the positive word category in both the visual and auditory channels and the negative word category in the auditory modality. The significant contributors differed across these analyses. In the visual positive category, Spanish proficiency contributed significantly to the model, whereas in the auditory positive category, both English and Spanish proficiency had significant weights. Although the influence of level of proficiency is consistent with previous findings (Dewaele, 2004; Dewaele & Pavlenko, 2002), it must be interpreted with caution. The significant predictive value of Spanish and English proficiency over the visual positive word category indicates a complex relationship between the level of proficiency in both languages and the differences in valence scores. The bivariate negative correlation seems to indicate that a higher level of proficiency in English had an attenuating effect in the differences between languages in the positive category, whereas a higher level of proficiency in Spanish seemed to have the opposite effect, as indicated by the positive direction of the correlation. Participants with a higher percent of life in the US tended to have higher proficiency in English and lower proficiency in Spanish. As expected, participants who acquired English later in life presented lower levels of English proficiency and higher levels of Spanish proficiency. There was an unusual significant correlation in these results; participants with higher self-reported levels of Spanish proficiency tended to have higher scores in the US cultural identity scores. It is possible that those individuals in our sample who were highly proficient in Spanish were more assimilated to the mainstream culture than to their ethnic culture. However, future research should focus on testing this finding using a linguistic measure of proficiency instead of self-reports.

In conclusion, our results support the view that there are differences in the appraisal of emotions in the two languages spoken by proficient and active bilinguals. However, the direction of the difference in emotion valence between L1 and L2 depends on the emotion category of words and is influenced by language proficiency, while the receptive sensory modality does not seem to have a significant influence. Results from the current study are supported by strong statistical power due to a large homogenous sample and the use of instruments

https://doi.org/10.1017/S1366728917000517 Published online by Cambridge University Press

with strong internal consistency. Nevertheless, the main limitation of the study is the unequal gender distribution in the sample and replication of our findings is needed with a sample with equal number of females and males.

Supplementary material

To view supplementary material for this article, please visit https://doi.org/10.1017/S1366728917000517

References

Altarriba, J. (2008). Expressions of emotions as mediated by context. *Bilingualism: Language and Cognition, 11,* 165–167. doi:10.1017/S1366728908003295

Altarriba, J., & Basnight-Brown, D. M. (2010). The representation of emotion vs. emotion-laden words in english and spanish in the affective simon task. *International Journal of Bilingualism,* doi:10.1177/1367006910379261

Anooshian, L. J., & Hertel, P. T. (1994). Emotionality in free recall: Language specificity in bilingual memory. *Cognition and Emotion, 8,* 503–514. doi:10.1080/02699939408408956

Ayçiçeği, A., & Harris, C. L. (2004). Brief report: Bilinguals' recall and recognition of emotion words. *Cognition and Emotion, 18,* 977–987. doi:10.1080/02699930341000301

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

Bond, M. H., & Lai, T. (1986). Embarrassment and code-switching into a second language. *The Journal of Social Psychology, 126,* 179–186.

Bradley, M. M., & Lang, P. J. (1999). Affective norms for English words (ANEW): Instruction manual and affective ratings. (No. C-1). University of Florida: The Center for Research in Psychophysiology.

Bradley, M. M., Greenwald, M. K., Petry, M. C., & Lang, P. J. (1992). Remembering pictures: Pleasure and arousal in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 379–390.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25,* 49–59. doi:https://doi.org.ezproxy.fau.edu/10.1016/0005-7916(94)90063-9

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

Chen, P., Lin, J., Chen, B., Lu, C., & Guo, T. (2015). Processing emotional words in two languages with one brain: ERP and fMRI evidence from Chinese–English bilinguals. *Cortex, 71,* 34–48.

Colbeck, K. L., & Bowers, J. S. (2012). Blinded by taboo words in L1 but not L2. *Emotion, 12,* 217–222. doi:10.1037/a0026387

Conrad, M., Recio, G., & Jacobs, A. M. (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. doi:10.3389/fpsyg.2011.00351; 10.3389/fpsyg.2011.00351

De Leersnyder, J., Mesquita, B., & Kim, H. S. (2011). Where do my emotions belong? A study of immigrants' emotional acculturation. *Personality and Social Psychology Bulletin, 37(4),* 451–463. doi:10.1177/0146167211399103

Dewaele, J.M. (2004). The emotional force of swearwords and taboo words in the speech of multilinguals. *Journal of Multilingual and Multicultural Development, 25,* 204–222.

Dewaele, J.M. (2008). The emotional weight of "I love you" in multilinguals' languages. *Journal of Pragmatics, 40,* 1753–1780.

Dewaele, J. M. (2011). Self-reported use and perception of the L1 and L2 among maximally proficient bi-and multilinguals: a quantitative and qualitative investigation. *International journal of the sociology of language, 2011(208),* 25–51.

Dewaele, J. M. (2013). Multilingualism and Emotions. In C. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 1–7). Oxford: Wiley-Blackwell.

Dewaele, J. M. (2016). Thirty shades of offensiveness: L1 and LX English users' understanding, perception and self-reported use of negative emotion-laden words. *Journal of Pragmatics, 94,* 112–127.

Dewaele, J. M., & Costa, B. (2013). Multilingual clients' experience of psychotherapy. *Language and Psychoanalysis, 2,* 31–50.

Dewaele, J.M., & Pavlenko, A. (2001). Web questionnaire bilingualism and emotions. London: University of London.

Dewaele, J.M., & Pavlenko, A. (2002). Emotion Vocabulary in Interlanguage. *Language Learning 52 (2)*, 263–322.

Eilola, T. M., & 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.

Eilola, T. M., Havelka, J., & Sharma, D. (2007). Emotional activation in the first and second language. *Cognition and Emotion, 2*1, 1064–1076. doi:10.1080/02699930601054109

Ferré, P., García, T., Fraga, I., Sànchez-Casas, R., & 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. doi:10.1080/02699930902985779

Harris, C. L. (2004). Bilingual speakers in the lab: Psychophysiological measures of emotional reactivity. *Journal of Multilingual and Multicultural Development, 25,* 223–247.

Harris, C. L., Ayçiçegi, A., & Gleason, J. B. (2003). Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language. *Applied Psycholinguistics, 24(04),* 561. doi:10.1017/S0142716403000286

Hsu, C. T., Jacobs, A. M., & 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.

IBM Corp. (Released, 2014). IBM SPSS Statistics for Windows, Version 23.0. Armonk, NY: IBM Corp.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language and Hearing Research, 50,* 940.

Morris, J. D. (1995). Observations: SAM: The self-assessment manikin; an efficient cross-cultural measurement of emotional response. *Journal of Advertising Research, 35,* 63–68.

Opitz, B., & Degner, J. (2012). Emotionality in a second language: It's a matter of time. *Neuropsychologia, 50(8),* 1961–1967. doi:10.1016/j.neuropsychologia.2012.04.021

Pavlenko, A. (2008). Emotion and emotion-laden words in the bilingual lexicon. *Bilingualism: Language and Cognition, 11,* 147–164.

Pavlenko, A. (2012). Affective processing in bilingual speakers: Disembodied cognition? *International Journal of Psychology, 47(6),* 405–428. doi:10.1080/00207594.2012.743665

Ponari, M., Rodríguez-Cuadrado, S., Vinson, D., Fox, N., Costa, A., & Vigliocco, G. (2015). Processing advantage for emotional words in bilingual speakers. *Emotion, 15(5),* 644.

Redondo, J., Fraga, I., Padrón, I., & Comesaña, M. (2007). The spanish adaptation of ANEW (affective norms for english words). *Behavior Research Methods, 39,* 600–605. doi:10.3758/BF03193031

Rubin, D. C., & Friendly, M. (1986). Predicting which words get recalled: Measures of free recall, availability, goodness, emotionality, and pronunciability for 925 nouns. *Memory & Cognition, 14(1),* 79–94.

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

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(6),* 1090–1098.

Zea, M. C., Asner-Self, K. K., Birman, D., & Buki, L. P. (2003). The abbreviated multidimentional acculturation scale: Empirical validation with two Latino/Latina samples. *Cultural Diversity and Ethnic Minority Psychology, 9,* 107.