The time course of emotion effects in first and second language processing: a cross cultural ERP study with German–Spanish bilinguals

Markus Conrad[1]*, Guillermo Recio[2] and Arthur M. Jacobs[1]

[1] *Department of General and Neurocognitive Psychology, Freie Universität Berlin, Berlin, Germany*
[2] *Department of Biological Psychology and Psychophysiology, Humboldt Universität zu Berlin, Berlin, Germany*

To investigate whether second language processing is characterized by the same sensitivity to the emotional content of language – as compared to native language processing – we conducted an EEG study manipulating word emotional valence in a visual lexical decision task. Two groups of late bilinguals – native speakers of German and Spanish with sufficient proficiency in their respective second language – performed each a German and a Spanish version of the task containing identical semantic material: translations of words in the two languages. In contrast to theoretical proposals assuming attenuated emotionality of second language processing, a highly similar pattern of results was obtained across L1 and L2 processing: event related potential waves generally reflected an early posterior negativity plus a late positive complex for words with positive or negative valence compared to neutral words regardless of the respective test language and its L1 or L2 status. These results suggest that the coupling between cognition and emotion does not qualitatively differ between L1 and L2 although latencies of respective effects differed about 50–100 ms. Only Spanish native speakers currently living in the L2 country showed no effects for negative as compared to neutral words presented in L2 – potentially reflecting a predominant positivity bias in second language processing when currently being exposed to a new culture.

Keywords: emotion, visual word recognition, bilinguals, second language processing, ERPs

INTRODUCTION

Emotions are a basic element of human communication which is mostly operated through language. From early on, linguistic theory assigned a central role to emotion as evident from Bühler's (1992) and Jakobson and Halle's (1969) definitions of language as comprising descriptive, expressive and appellative functions. The intimate link between cognition and emotion in general (see Dolan, 2002, for a review) and language and emotion in particular is drawing increasing attention in the fields of cognitive psychology and cognitive neuroscience. The role of emotion for language perception is being investigated in a range of psycholinguistic domains and methodologies: While it may appear intuitively evident that larger units of text or speech describing emotional events (e.g., Altmann et al., submitted; Bohrn et al., submitted), or the emotional prosody of speech (e.g., Kotz and Paulmann, 2007; Pell et al., 2009) would trigger emotional processing in the reader or listener, also for the processing of more fine grained units of language, emotion effects have been observed. This holds true for – superficially – purely cognitive tasks, and even when just single words are visually presented: Emotion-laden words are recalled better than neutral words [Rubin and Friendly, 1986; see Dietrich et al., 2001, for an event related potential (ERP) study] and they also seem to possess a processing advantage in terms of speed of lexical access, e.g., when the task consists of deciding whether a given letter string is a word or not in lexical decision (Eviatar and Zaidel, 1991;

Kuchinke et al., 2005; Kousta et al., 2009; Schacht and Sommer, 2009a, among others).

Rather recently, this apparent emotion–cognition coupling for visually presented single words has been studied using neuroscientific methods trying to explore physiological (pupil dilatation, Võ et al., 2008), electrophysiological (ERPs, Schacht and Sommer, 2009a,b; Hofmann et al., 2009, see also Junghöfer et al., 2006; Kissler et al., 2006, for reviews), and neuronal (fMRI, Kuchinke et al., 2005, 2006) correlates of emotional word content during visual word recognition in native language processing.

On the other hand, when we learn a second language, the question arises of whether the same emotion–cognition coupling as in native language processing can be assumed given that second language processing is generally less efficient and automatic (see van Heuven and Dijkstra, 2010, for a review), and emotional connotations for words might be established especially during childhood.

The way bilinguals or second language learners deal with their two languages has been the topic of a large number of empirical studies in the field of cognitive psychology, mainly investigating the issues of simultaneous activation of the two language systems or the cognitive costs of language switch (for ERP studies, see, e.g., Chauncey et al., 2008; Midgley et al., 2008, 2011; Gillon-Dowens et al., 2010; van der Meij et al., 2011).

Rather few empirical studies are available to date addressing directly the issue of second language learners' sensitivity to

emotional content in L2, and their results are heterogeneous: From early on, research seemed to be driven by the general assumption that second language processing would be characterized by increased emotional distance compared to native language processing. Bilingual speakers were reported to feel more free to talk about potentially embarrassing topics in their L2 (Bond and Lai, 1986), possibly because their emotional involvement was stronger for native language than for L2 emotion representations. Quoting Sutton et al. (2007): "When emotions words are learned in the first language they are more deeply coded, experienced in more contexts, and applied in various ways as compared to their L2 equivalents." Linking such qualitative findings or theoretical considerations to emotion theories including physiological correlates of emotion (see Schachter and Singer, 1962), taboo words – especially prone to provoke physiological arousal- were often used as stimulus material to assess emotion sensitivity in L2 (see Dewaele, 2004). Decreased physiological responses to such words when presented in L2 as compared to L1 (Gonzales-Reigosa, 1976; Harris et al., 2003, see also Harris, 2004) seemed to support the view of attenuated emotionality of L2. Also concerning word recall, an emotion advantage was initially reported to be limited to L1 (Anooshian and Hertel, 1994), but Ayçiçegi and Harris (2004) more recently provided evidence for the contrary.

Generally, recent research rather points to the conclusion that second language processing should not be understood as *a priori* or always less emotional than native language processing, but that L2 emotionality is modulated by factors such as age of acquisition, proficiency, and exposure. For instance, Eilola et al., 2007, see also Sutton et al., 2007) obtained comparable effects in L1 and L2 presenting negative and taboo words in a Stroop paradigm, concluding that "for late bilinguals with good knowledge of their second language, the first (L1) and second (L2) language are equally capable of activating the emotional response to word stimuli representing threat." Harris et al. (2006) could partly extend this claim to physiological correlates showing that bilinguals' skin conductance response to emotion-laden words in L2 varied as a function of age of L2 acquisition.

In sum, there is a growing body of evidence that emotion plays an important role not only in native but also in second language processing, but clearly more data is needed to understand how mandatory and how early or automatically an access to emotional content of L2 words would occur. A promising method to study the time course of processes at the interplay of emotion and cognition is EEG registration with its high temporal resolution. Surprisingly, to our knowledge, only one study has intended to contrast ERPs for emotion-laden words presented in L1 and L2 (Kim, unpublished doctoral dissertation). Unfortunately, no conclusive results were obtained either regarding general ERP effects of emotional valence or differential ones as a function of language status. This failure to obtain reliable effects led Harris et al. (2003) to propose that "words' emotionality may be a case where electrodermal recording is preferred to ERPs, despite the latter's superior temporal resolution."

However, to date, this claim can no longer be supported given that a number of recent studies have shown ERPs' sensitivity to both emotional valence and arousal (Junghöfer et al., 2006; Kissler et al., 2006; Schupp et al., 2006; Schacht and Sommer, 2009a,b;

Hofmann et al., 2009) in native language processing. In particular, two components of the ERP signal have been reported to be modulated by emotional valence in a visual lexical decision task: especially positive words – as compared to neutral ones – produced increased early posterior negativity (EPN) at occipito-temporal electrode sites (EPN, starting, for instance, at about 370 ms, immediately (50 ms) after the onset of the Lexicality effect between word and non-words in Schacht and Sommer, 2004, see also Kissler et al., 2007) and a late positive complex (LPC) with increased positivity at centro-parietal electrodes (LPC, peaking at around 600 ms, e.g., Schacht and Sommer, 2009a).

This specific pattern of results provides a promising framework for the investigation of the nature of emotion processing in L2, because these two distinct components reflect different processing levels, at which emotion sensitivity between L1 and L2 might prove to be differential: the EPN for words with emotional valence is interpreted as an attention shift toward words with apparent emotional relevance at early processing stages, whereas the LPC is supposed to reflect higher level semantic evaluation (see, e.g., Kissler et al., 2006). Even assuming some general sensitivity to emotional word content in L2 – as suggested by Eilola et al. (2007), Harris et al. (2006), and Sutton et al. (2007) – it is an open question whether sensitivity to emotional content during L2 processing would be given already during early stages of visual word recognition or whether emotional relevance given to words in L2 would rather result from a re-translation of these words to L1 – based on the emotional context of these words in the native language, which is assumed to be central for emotional relevance (see Sutton et al., 2007). Accordingly, a specific hypothesis for an ERP study investigating emotion effects in L2 could be that no modulation of the EPN component – intimately bound to lexical access in the presented language – might be observed when words are presented in L2 – due to a potentially attenuated early emotion–cognition coupling in L2. Instead, possibly only during more elaborated processing stages – as reflected by the LPC – emotion sensitivity in L2 might be observed, potentially arising from a re-translation of words to L1 resulting in the activation of L1 emotion representations.

On the other hand, if both ERP components proved to be affected by emotional valence in both L1 and L2, such a result would indicate comparable – early and direct– processing of emotional content in L2 and L1.

THE PRESENT STUDY

We designed a visual lexical decision task presenting words with either positive, neutral, or negative valence to two groups of German–Spanish late bilinguals – native speakers of each language with sufficient proficiency in the respective other language. Identical or at least closely comparable semantics for words presented in L1 and L2 are, of course, a necessary condition for the comparison of emotion processing across the two languages, but it is not trivial to obtain such stimulus material, first, because emotion concepts may vary across languages (see Altarriba, 2003; Pavlenko, 2008), and, second, because a number of other factors have to be controlled across experimental conditions (see Materials and Design for details). Due to the increasing interest in emotion-related processes in the field of cognitive psychology, a growing number of

normative databases with rating values on emotional dimensions for different languages is becoming available, but very few databases have been published so far containing rating data for a shared set of words from different languages (e.g., Redondo et al., 2007, for a Spanish version of the ANEW corpus of Bradley and Lang, 1999, for the English language, see also Schmidtke et al., in preparation). Before conducting this study, a normative database allowing for the comparison of German and Spanish word emotion ratings was not available. Several databases have been published for each the German (Võ et al., 2006, 2009) and the Spanish (Redondo et al., 2005; Redondo et al., 2007) language, but their semantic overlap was limited. Within a current research project entitled "bilingualism and affectivity in reading" at the Freie Universität Berlin, we have been collecting emotion rating data for about 6,000 words in each the English, Spanish, and German language. These three corpora (Conrad et al., in preparation) provide sufficient semantic overlap to tackle the question of whether identical words with comparable emotional valence across languages would produce different emotion effects in L2 and L1. Presenting the stimuli to two different participant groups – with different L1 – furthermore allows testing whether eventual effects would generalize over different cultural/linguistic backgrounds, and each L1 participant group would serve as control group for the other's L2 performance.

PARTICIPANTS

Forty German native speakers participated in the experiment. All were late Spanish–German bilinguals having acquired the respective other language after the age of 12.

Note that the two L1 groups are not strictly comparable in terms of L2 exposure, because all Spanish participants were living and being tested in the country of their L2, whereas German participants were presently living in their native country – though most of them had reported past stays in Spanish speaking countries for several months.

Only individuals who had proven capable of fluent L2 conversation over the telephone (during participant acquisition) were invited to participate. Prior to the experiment, second language linguistic profile and proficiency was assessed via self-report questionnaires (i.e., Language Experience and Proficiency Questionnaire, LEAP-Q, Marian et al., 2007) administered in their native

language and via an academic proficiency test (DIALANG, Huhta et al., 2002). Only individuals reaching at least an A2[1] level of second language proficiency in the DIALANG test participated in the study. Characteristics of L2 proficiency for the two participant groups are presented in **Table 1**. Note that Spanish native speakers had started to learn their L2 later than participants in the German group. Accordingly they had less years of L2 learning experience. On the other hand – related to the afore-mentioned systematic difference between the two groups, this latter group had been living in the L2 country for more months than the German participants. Direct measures of L2 proficiency suggested that L2 proficiency was higher for the Spanish than for the German group in terms of scores in an L2 vocabulary test (although not significant) and standardized levels of L2 proficiency (ranging from A2 to C1 in the present sample, but note the more detailed discussion of this issue later on where we attempt to use individual L1 and L2 RT data from this study as an alternative and potentially more reliable measure of L2 proficiency). Prior to the experimental session all participants read and signed consent to the study. All of them were students and obtained a small amount of money for their participation in the experiment. All participants were right handed as evaluated with Oldfield (1971) and none reported neurological, or language problems.

MATERIALS AND DESIGN

From our databases comprising emotion ratings for about 6,000 Spanish and German words (Conrad et al., in preparation), 3×80 word pairs (Spanish–German translations) were selected for the three cells with positive, neutral, or negative valence. Words were entered in the "positive" condition when their corresponding mean valence ratings in both native language contexts were higher than $+1.2$ on a bipolar seven point scale ranging from -3 ("very negative") to $+3$ ("very positive") used for normative data collection in the two databases. They were entered in the "negative" condition when corresponding mean ratings in both languages were

---

[1]According to the European Reference Frame for Language Competence, A1–A2 correspond to an "elementary level," B1–B2 to an "independent level," and C1–C2 to a "competent level" of second language use.

Table 1 | Means (ranges, or SD) for measures of second language proficiency for the group of native Spanish speakers (SP = 26), for the complete sample of German native speakers (GE = 40) and for a matched Sub-sample of German native speakers (GE = 26).

	♂ – ♀*	Age*[2] *	AP*[2] *	Vocabulary	StAge L2**[2] ***	Years L2*	Months LC*[2] **	RT L1/L2*	ERRL1/L2
GE = 40	12–28	26.20 (20–33)	8–17–5–3	442 (201)	18.65 (14–26)	7.55 (1–14)	22.63 (0–224)	0.84 (0,14)	0.30 (0,26)
GE = 26	10–16	26.31 (21–32)	6–14–3–2	460 (192)	19.04 (14–24)	7.27 (1–14)	25.23 (0–224)	0.81 (0,11)	0.27 (0,24)
SP = 26	14–12	28.54 (20–38)	1–9–11–3	497 (313)	23.15 (12–33)	5.38 (1–15)	50.35 (4–180)	0.76 (0,15)	0.39 (0,72)

∗ p < 0.05 only between GE = 40 and SP = 26.

∗[2] ∗ p < 0.05 between GE = 26 and SP = 26 and between GE = 40 and SP = 26.

∗[2] ∗∗ p < 0.05 between GE = 26 and SP = 26 and p < 0.01 between GE = 40 and SP = 26.

∗∗[2] ∗∗∗ p < 0.01 between GE = 26 and SP = 26 and p < 0.001 between GE = 40 and SP = 26.

values are presented only for participants entering the ERP analyses.

LEAP-Q-measures: Startage of L2 learning (StAge L2), years of L2 learning experience (years L2), months of living experience in the L2 country (months LC).

DIALANG-measures: levels of academic proficiency (in increasing order: a2-b1-b2-c1; AP), mean vocabulary test scores (vocabulary).

Lexical decision task measures: the ratios between individual response latencies to words in L1 and L2 (RTL1/L2) and between the respective error rates (ERRL1/L2).

lower than −1.2. Words in the "neutral" condition had mean valence ratings ranging between −0.5 and +0.5 and mean arousal ratings lower than 3.3 in both languages (rated on a five point scale from 1 "very calm" to 5 "very exciting"). In addition, in all three categories, words were only entered when mean valence ratings for specific pairs did not display differences across languages greater than 0.675[2] in order to assure that no cultural/linguistic differences in emotional concepts or connotations were given for the selected word material. Furthermore, only word pairs were used, for which translations seemed unambiguous (for many word pairs this would not be the case, e.g., the German word "Eis" refers to both "ice" and "ice-cream," corresponding to the separate Spanish lexical entries "hielo" and "helado"). Examples for positive German–Spanish word pairs from the stimulus material are: "Kuβ/beso" (kiss) or "Sommer/verano" (summer); for negative pairs: "Krieg/guerra" (war) or "Alptraum/pesadilla" (nightmare); and finally, "Wand/pared" (wall) or "Inhalt/contenido" (content) for neutral words.

Words across the three categories were closely matched in both languages on a number of variables known to or plausible to influence speed of lexical access or ERPs: word length (in terms of letters and phonemes), word frequency (as number of occurrences per 1 Million and Log frequency), regular vs. irregular stress, word class (in all cells about 65% of stimuli were nouns, about 20% were adjectives, about 15% were verbs, lemmata in all cases), density, and frequency of orthographic neighborhood (see Holcomb et al., 2002; Molinaro et al., 2010), initial syllable frequency (Barber et al., 2004; Hutzler et al., 2004). Word statistics were derived from CELEX (Baayen et al., 1993) for the German and from LEXESP (Sebastián-Gallés et al., 2000) for the Spanish language. German words comprised between three and nine letters and between one

---

[2]Note that this cut-off seems liberal enough to account for the general pattern that valence ratings in the Spanish database were on average about 0.15 points more positive than in the German database, which probably rather reflects more general culture specific attitudes than different emotional concepts at the level of single lexical items.

and four syllables, Spanish words between four and nine letters and two and five syllables. Characteristics for word stimuli are presented in **Table 2**. Non-words in both languages were constructed as pronounceable letter strings matched to words in each language on letter and syllable length. The complete material in each language contained 545 items: 80 words for each condition of positive, neutral, or negative valence, 240 non-words (involving a manipulation beyond the focus of this paper) and 65 filler pseudowords (pseudohomophones or non-word items with a high number of orthographic neighbors) that were added to increase general task difficulty. None of the non-words could form a word in the respective L2.

PROCEDURE
Participants were seated in a quiet, dimly illuminated room in a comfortable chair. Stimuli were presented on a 17″ color monitor (70 Hz refresh rate) placed at 80 cm from participants' eyes. Participants were instructed to read the combination of letters appearing in the screen and respond as fast and as accurately as possible whether the stimulus was a word or not in the presented language by pressing one of two buttons in a game device with their index fingers. The assignment of word and non-word responses to the left or the right button was balanced across participants. Two separate blocks for each language contained exclusively either Spanish or German words and non-words. Order of blocks (Spanish or German) was counterbalanced across participants so that within each L1 group half of the participants performed the task first in L1 and then in L2, and vice versa. There was a practice block consisting of 10 randomized items (5 words and 5 non-words) before each block. Each trial started with a fixation cross at the center of the screen, replaced after 300 ms by the stimulus, which was present until participants responded. After responses, a blank screen was presented for 2000 ms before the next trial started. Participants were instructed to blink – if necessary – only during this blank screen. The whole experimental session consisted of 1110 trials (including practice items) presented in randomized order in white letters (font Times New Roman, size 20) over black

Table 2 | Means for independent variables (valence) and control variables, i.e., word length (syllables, S; letters, L; and phonemes, PH) word frequency (per 1 million of occurrences: F/1Mio and log (10) per 1 million of occurrences: LogF), orthographic neighborhood density (N) and frequency (number of higher frequency orth. neighbors: HFN) and initial syllable frequency [log (10) per 1 million of occurrences: LogFS1] for German and Spanish word stimuli.

	VAL	F/1Mio	LogF	S	L	PH	N	HFN	LogFS1
GERMAN WORDS									
Negative	−2.09	36.88	1.29	2.05	6.55	5.67	1.53	0.38	2.61
Neutral	0.06	37.04	1.27	2.08	6.44	5.46	1.71	0.61	2.43
Positive	1.99	31.33	1.33	2.08	6.55	5.61	1.92	0.45	2.43
p>	0.0001	0.4	0.4	0.8	0.6	0.3	0.3	0.1	0.3
SPANISH WORDS									
Negative	−2.18	34.08	1.29	2.84	6.75	6.59	1.49	0.19	3.38
Neutral	0.07	32.03	1.23	2.79	6.73	6.58	1.55	0.34	3.40
Positive	2.04	30.08	1.25	2.94	6.89	6.81	1.31	0.26	3.45
p>	0.0001	0.5	0.4	0.2	0.4	0.3	0.5	0.1	0.4

p>, Minimum p-values for tests of significant mean differences between any two conditions.

background. There were short pauses after 250 items within each block and a longer pause between the two blocks.

BEHAVIORAL RESULTS

To examine effects of emotional valence on response latencies and recognition rates in first vs. second language processing, we analyzed the data from both blocks of German and Spanish words separately using valence as within-subject and participants' first language as between-subjects factors.

Mean correct response latencies and percent of errors were submitted to separate analyses of variance (ANOVAs) for participants and items (F1 and F2 respectively).

Response latencies deviating more than two SD from the mean for participant and condition were considered outliers and removed from analyses (4.52% of the data). Items with corresponding error rates equal to or higher than 50% in any of the two participant groups' responses were excluded from all analyses (six German and seven Spanish words). Only participants, whose ERP data will be presented later on (two participants from the Spanish group had to be removed due to exceeding number of artifacts) entered the analyses (see **Table 3** for means of dependent variables in all conditions).

LEXICAL DECISION WITH GERMAN WORDS

Responses to words were 317 ms slower for Spanish than for German native participants, $F1\ (1,64) = 17.59$, $p < 0.0001$, $\eta_p^2 = 0.216$; $F2\ (1,231) = 903.18$, $p < 0.0001$, $\eta_p^2 = 0.796$, this slowdown of visual word recognition in L2 as compared to L1 was mirrored by an effect on error rates where Spanish participants committed more errors (9.65 vs. 1.45%) than German native speakers, $F1\ (1,64) = 33.65$, $p < 0.0001$, $\eta_p^2 = 0.345$; $F2\ (1,231) = 131.99$, $p < 0.0001$, $\eta_p^2 = 0.364$.

Significant effects of Lexicality (comparing the 240 word and the 240 non-word stimuli ignoring fuller items) were obtained in both participant groups: German native speakers responded 46 ms more quickly to words than to non-words, $F1\ (1,39) = 11.62$, $p < 0.003$, $\eta_p^2 = 0.229$; $F2\ (1,472) = 75.75$, $p < 0.0001$, $\eta_p^2 =$

0.138, whereas the respective difference for the Spanish native speakers was 228 ms, $F1\ (1,25) = 16.34$, $p < 0.0001$, $\eta_p^2 = 0.395$; $F2\ (1,472) = 133.12$, $p < 0.0001$, $\eta_p^2 = 0.220$.

Rather surprisingly, only a marginally significant effect for valence on response latencies was obtained over participants, with positive words yielding the fastest (791 ms) and negative words the slowest (828 ms) responses (802 ms for neutral words), $F1\ (2,128) = 2.59$, $p < 0.08$, $\eta_p^2 = 0.039$; $F2\ (2,231) = 2.13$, $p < 0.13$, $\eta_p^2 = 0.018$. No interaction valence × participant group was given, $p > 0.1$. No significant effect of valence on response latencies did arise in either of the two participant groups in separate analyses. In contrast, highly significant effects of word valence were present in the error data: positive words provoked the least number of errors (3.62%), negative words were the most probables not to be recognized (7.34%), followed by neutral words (5.70%), $F1\ (2,128) = 24.90$, $\eta_p^2 = 0.315$, $p < 0.0001$; $F2\ (1,231) = 8.01$, $p < 0.0005$, $\eta_p^2 = 0.065$. But note that this effect was modulated by participant group, $F1\ (2,128) = 28.03$, $p < 0.0001$, $\eta_p^2 = 0.305$; $F2\ (2,231) = 9.22$, $p < 0.0001$, $\eta_p^2 = 0.074$, and separate analyses revealed that no valence effect on error rates was present for the German native participants', $F < 0.1$. Rather, valence caused a significant effect in the Spanish participants' L2 data alone, $F1\ (2,50) = 20.91$, $p < 0.0001$, $\eta_p^2 = 0.455$; $F2\ (2,231) = 8.87$, $p < 0.0002$, $\eta_p^2 = 0.071$. All single tests between error rates for positive (5.86%), neutral (9.91%) and negative words (13.18%) were significant (all $p < 0.05$ over participants and items).

LEXICAL DECISION WITH SPANISH WORDS

Again, as in the German data, word recognition was more efficient in L1 as compared to L2: Native speakers' correct responses to words were 101 ms faster than second language learners', $F1\ (1,64) = 3.84$, $p < 0.06$, $\eta_p^2 = 0.057$; $F2\ (1,230) = 10.44$, $p < 0.0001$, $\eta_p^2 = 0.599$, accompanied by a consistent effect on error rates (1.59 vs. 6.81%), $F1\ (1,64) = 10.88$, $p < 0.002$, $\eta_p^2 = 0.145$; $F2\ (1,231) = 92.50$, $p < 0.0001$, $\eta_p^2 = 0.287$, but notice

Table 3 | Means and SD of response latencies (RT) and error percentages (%ERR) for words and non-words in the German and Spanish lexical decision task (LDT).

	German participants (N = 40)				German participants (N = 26)				Spanish participants (N = 26)			
	RT	SD	%ERR	SD	RT	SD	%ERR	SD	RT	SD	%ERR	SD
GERMAN LDT												
Pseudowords	864	203	8.11	5.73	864	154	7.45	5.00	1443	659	21.15	11.49
Non-words	695	142	0.96	1.72	700	137	0.78	1.26	1184	537	9.85	11.53
Negative words	654	132	1.47	1.78	657	116	1.19	1.17	1003	449	13.61	12.30
Neutral words	654	142	1.48	1.80	663	152	1.36	1.65	950	392	9.92	9.40
Positive words	639	131	1.38	2.06	648	130	1.05	1.58	943	546	5.86	6.03
SPANISH LDT												
Pseudowords	1042	306	15.80	17.50	1084	278	14.34	17.84	1006	449	8.18	8.47
Non-words	870	230	4.87	7.89	926	230	4.80	8.94	836	328	2.10	2.75
Negative words	795	224	6.19	9.10	821	135	5.05	5.58	690	188	1.34	1.82
Neutral words	829	255	8.87	9.50	847	141	7.92	6.80	712	204	1.84	1.89
Positive words	754	167	5.39	6.05	778	107	4.81	4.45	674	185	1.60	1.77

that the size of the respective group effects is much attenuated relative to the German lexical decision data suggesting that L2 proficiency – as assessed by speed of lexical decision – was higher in the German than in the Spanish native speakers' group: Spanish participant's L2 responses differed more from native responses than Germans' L2 responses did. Significant effects of Lexicality were obtained in both participant groups: Spanish native speakers responded 139 ms more quickly to words than to non-words, $F1\ (1,25) = 16.80$, $p < 0.0001$, $\eta_p^2 = 0.387$; $F2\ (1,471) = 411.16$, $p < 0.0001$, $\eta_p^2 = 0.466$, whereas the respective difference for the German native speakers was 79 ms, $F1\ (1,39) = 5.67$, $p < 0.03$, $\eta_p^2 = 0.127$; $F2\ (1,471) = 70.63$, $p < 0.0001$, $\eta_p^2 = 0.130$.

For the factor emotional valence, this time, highly significant effects emerged also in the RT data: Responses were fastest for positive words, followed by negative ones, and slowest for neutral words, $F1\ (2,128) = 17.84$, $p < 0.0001$, $\eta_p^2 = 0.218$; $F2\ (2,230) = 10.44$, $p < 0.0001$, $\eta_p^2 = 0.083$. No interaction valence × participant group was given, $p > 0.1$, suggesting that emotional valence of spanish words influenced processing speed in L2 and L1 processing in similar ways. In particular, the effect of valence within the Spanish native speakers' data, $F1\ (2,50) = 17.44$, $p < 0.0001$, $\eta_p^2 = 0.411$; $F2\ (2,230) = 10.11$, $p < 0.0001$, $\eta_p^2 = 0.081$, was characterized by significant differences between positive (674 ms) and neutral (712 ms; $p < 0.0001$) as well as between negative (690 ms) and neutral conditions ($p < 0.03$). Differences between positive and negative words were only marginally significant over participants ($p < 0.06$; $p > 0.2$ in the item analysis)[3]. Most importantly, also for second language reading, valence caused a significant effect on RTs, $F1\ (2,78) = 13.12$, $p < 0.0001$, $\eta_p^2 = 0.252$; $F2\ (2,230) = 7.80$, $p < 0.001$, $\eta_p^2 = 0.063$, the general direction of which was the same as in the native speakers' data: responses to positive words (754 ms) were faster than to neutral (829 ms; $p < 0.0001$) or negative ones (795 ms; $p < 0.03$ in the participant, $p > 0.1$ in the item analyses), differences between the latter two being only marginally significant ($p < 0.06$) in the participant, but not ($p > 0.1$) in the item analyses[4].

Also on error rates a significant valence effect was present, with positive words provoking the smallest and negative words the greatest number of errors, $F1\ (2,128) = 9.73$, $p < 0.0001$, $\eta_p^2 = 0.132$; $F2\ (2,230) = 2.56$, $p < 0.09$, $\eta_p^2 = 0.022$. This effect was modulated by participant group, $F1\ (2,128) = 9.56$, $p < 0.003$, $\eta_p^2 = 0.093$; $F2\ (2,230) = 2.78$, $p < 0.07$, $\eta_p^2 = 0.024$. Separate comparison showed that whereas valence did not affect Spanish native speakers' error rates, $F < 1$, it did so for second language performance, $F1\ (2,78) = 13.85$, $p < 0.0001$, $\eta_p^2 = 0.262$; $F2\ (2,230) = 2.96$, $p < 0.06$, $\eta_p^2 = 0.025$. *Post hoc* test revealed that German native speakers' error rates were highest in the neutral condition (8.87%) differing significantly – at least in the participant analyses from both positive (5.39%; $p < 0.0001$; $p < 0.06$ in the item analysis) and negative words (6.19%; $p < 0.0003$; $p > 0.1$ in the item analysis).

---

[3]Always the greatest p-value after Bonferroni correction corresponding to either the participant or the item analysis is given.
[4]Note that item analyses between RTs for negative words and those in the other two conditions would be significant without Bonferroni correction.

DISCUSSION OF BEHAVIORAL DATA

At least for the Spanish lexical decision task, effects in the RT data corroborate previous reports of a processing advantage for emotional over neutral words that is especially robust for positive, but also observable for negative stimuli (e.g., Kousta et al., 2009). The important novel finding here is that the same pattern can be observed in L2 processing, which, in turn, seems not to be characterized by emotional distance – although respective effect sizes in the present data suggest that this emotional processing advantage is more pronounced in L1 than in L2. But note that generally increased noise in L2 RT data might also hold responsible for this attenuation of effect sizes.

The absence of valence effects in the German lexical decision RT data was clearly unexpected, but we believe that specific characteristics of non-words used might account for this. A specific manipulation within the German non-words (otherwise irrelevant for the topic of this study) involving many low-frequency letter clusters within these stimuli might have made them less word-like than the Spanish non-words (Note that unlike in German, constraints of Spanish orthography make almost any Spanish letter combination either illegal or relatively high-frequent). For the German lexical decision, a resulting modulation of task context via non-word characteristics might have induced a shift in participants' response criteria for giving a "yes" response toward a more liberal fast-guess strategy (see Grainger and Jacobs, 1996). The differential pattern of lexicality effects – being relatively weak especially in the German natives' L1 data as compared to the Spanish natives' L1 data –supports the assumption that potential RT differences between different German word conditions have been reduced due to the application of a more liberal response criterion all German words would benefit from. Of course, this does not imply that no lexical access or semantic processing would have occurred for German words (which would make it improbable to observe any emotional effects in the ERP data), rather it means that lexical access or word identification would no longer have been alone responsible for triggering correct responses to German words. Clearly, ERPs are more sensitive to the multifaceted character of the time course of visual word recognition than response latencies – reflecting mainly the final point of a decision process. The absence of a significant effect of emotional valence in the German LDT data should, therefore, not represent a problem for the interpretation of potential ERP effects of valence for the same word material. On the contrary, the absence of significant RT differences between conditions of positive, negative, and neutral German words in both participant groups has the advantage that any ERP differences between these conditions are unlikely to be contaminated by response processes.

Analyses of error rates interestingly add to the picture of emotion effects across L1 and L2 word processing: Whereas word valence did not affect error rates in either of the two L1 data sets, for both cases of processing L2 words error rates differed significantly between conditions of emotional valence: Whereas Spanish native speakers' recognition of German words displayed a clear positivity bias with positive words provoking the least and negative words the highest amount of errors, German native speakers recognized both types of emotional L2 words more often than neutral ones. We assume that the clear distinction of respective effects between

L1 and L2 data can best be interpreted as a shift of the locus of effect: Recognition errors in L1 probably only represent a failure to retrieve a word, but these type of errors in the L2 data might also reflect an emotion bias effective during the learning process of a second language where emotional words (especially positive ones) might be more efficiently encoded and memorized than neutral ones.

Finally, this data shed an interesting light on the problem of assessing L2 proficiency: As mentioned above, our participants' scores on standardized academic proficiency tests predicted higher L2 proficiency within the Spanish native group. The RT data of the present experiment, on the other hand, suggest that German participants' L2 processing was more native-like than the Spanish group's and it should be noted that our subjective impression during data acquisition was also that the German group was more fluent in Spanish than the inverse. This contradictory picture might best be explained by a general problem with comparing standard L2 proficiency tests across different languages, because their specific choice of items within every language might strongly constrain the validity of such comparisons. Moreover, the relatively high importance these tests assign to the comprehension of rather sophisticated grammar problems (especially evident in the Spanish test) might not perfectly reflect proficiency in everyday L2 use.

The present experiment providing individual response latencies for words in L1 and L2, in turn, could itself be considered an assessment of L2 proficiency. We suggest that calculating the individual difference of processing speed for identical material across the two languages might provide an adequate estimate of individual L2 proficiency. Dividing individual mean response latencies in L1 by those in L2, the resulting ratio indeed differed significantly (see **Table 1**) between the two participant groups suggesting higher L2 proficiency within the German group, matching the pattern of main effects for participant group within the experimental subsets for each language. We will again refer to this measure when trying to establish two subsamples of participants matched on L2 proficiency for the analyses of ERP data.

PSYCHOPHYSIOLOGICAL RECORDINGS

Electroencephalogram (EEG) was recorded from 60 electrodes (Fp1,Fpz,Fp2,AF3,AF4,F5,F3,F1,Fz,F2,F4,F6,FT7,FC3,FC1,FCz, FC2,FC4,FT8,T7,C5,C3,C1,Cz,C2,C4,C6,T8,TP7,CP5,CP3,CP1, CPz,CP2,CP4,CP6,TP8,P9,P7,P5,P3,P1,Pz,P2,P4,P6,P8,P10,PO9, PO7,PO3,POz,PO4,PO8,PO10,O1,Oz,O2,Iz,M1,M2) using two 32 channel amplifiers (Brainamp, Brain Products, Germany). Electrodes were mounted in a cap and referenced to the right mastoid. Four additional electrodes were attached under, above, and next to the eyes to record vertical and horizontal electro-oculogram. Electrode impedance was kept below 5 kΩ. EEG was recorded at 500 Hz sampling rate. All channels were offline filtered with a band pass filter 0.1–20 Hz and notch filter of 50 Hz, and recalculated to average reference meeting the recommendation to calculate EPN and LPC components (Junghöfer et al., 2006; Schacht and Sommer, 2009a). EEG continuous signal was segmented in 1.8 s[5] epochs starting

---

[5]This segment length was chosen, because it was unclear for how much potential emotional effects in L2 would be delayed compared to L1.

200 before stimulus onset, which served as reference point for the pre-stimulus baseline correction. Ocular artifacts were corrected using independent component analyses (ICA) with Analyzer 2.0 software (Brain Products, Germany). Only segments with correct responses and free of artifacts were averaged over experimental conditions. Differences in values >80 μV in intervals of 70 ms as well as amplitudes >50 or <−50 μV were considered artifacts. Data from two participants with Spanish L1 were excluded from the analyses due to excessive numbers of artifacts.

Numbers of valid trials entering the analyses were very similar across conditions and participants: German participants: M neutral = 74.4 (SD = 5.8), M positive = 74.4 (SD = 5.2), and M negative = 74.7 (SD = 4.7) in L1 and M neutral = 70 (SD = 8), M positive = 72 (SD = 6.4), and M negative = 71.5 (SD = 8.6) in L2. Spanish participants M neutral = 74.7 (SD = 5.0), M positive = 74.1 (SD = 6.4), and M negative = 74.9 (SD = 5.8) in L1 and M neutral = 69.9 (SD = 8.1), M positive = 71.5 (SD = 8.3), and M negative = 66.4 (SD = 9.9) in L2.

STATISTICAL ANALYSES

Given general processing differences between L1 and L2 with processing of the latter one to be expected more time-consuming, different onsets for potential ERP components of emotional valence in L1 and L2 could be expected. Visual inspection of the data confirmed this assumption and thus, averaged ERP mean amplitudes were segmented into 25 ms epochs from stimulus onset (0 ms) to 700 ms onward. This exploratory method is also suitable to compare the latencies of consecutive components with different onsets as the EPN and the LPC (see Recio et al., 2011). Repeated measure ANOVAs were calculated for each time window of ERP data including the factors electrode site (60 levels) and valence (3 levels). By definition, the average reference method sets the mean activity across all electrodes to zero, hence, in ANOVAs including all electrodes only interactions of experimental factors with electrode site are meaningful. As we are interested in two components (EPN and LPC) with distinctive topography distribution, we considered more appropriate to include the activity across all electrodes rather than defining regions of interest for our analyses. For sake of brevity, in the results section, reports of main effects for the factors electrode site or valence will be omitted, only interactions of the two effects will be reported as "valence effects." p-Values were Huynh–Feldt adjusted. *Post hoc* – Bonferroni corrected – comparisons between different cells of the factor valence (valence x electrode) were conducted when significant interactions between the three level factor valence and electrode site were observed. To meet the problem of multiple testing for consecutive time windows, only significant effects appearing in at least three consecutive time windows were considered meaningful and are reported.

RESULTS AND DISCUSSION

Results from ANOVAs on the ERP data involving effects of valence in 25 ms time windows between 200 (no significant effects had appeared previously) and 700 ms are presented in **Figure 1** (involving each 26 German and Spanish participants – see paragraph "*Reanalyses matching participant groups on second language proficiency*"). Mean activity on representative electrodes F3, F4, C1, C2, PO9, and PO10 is depicted in **Figure 2**.
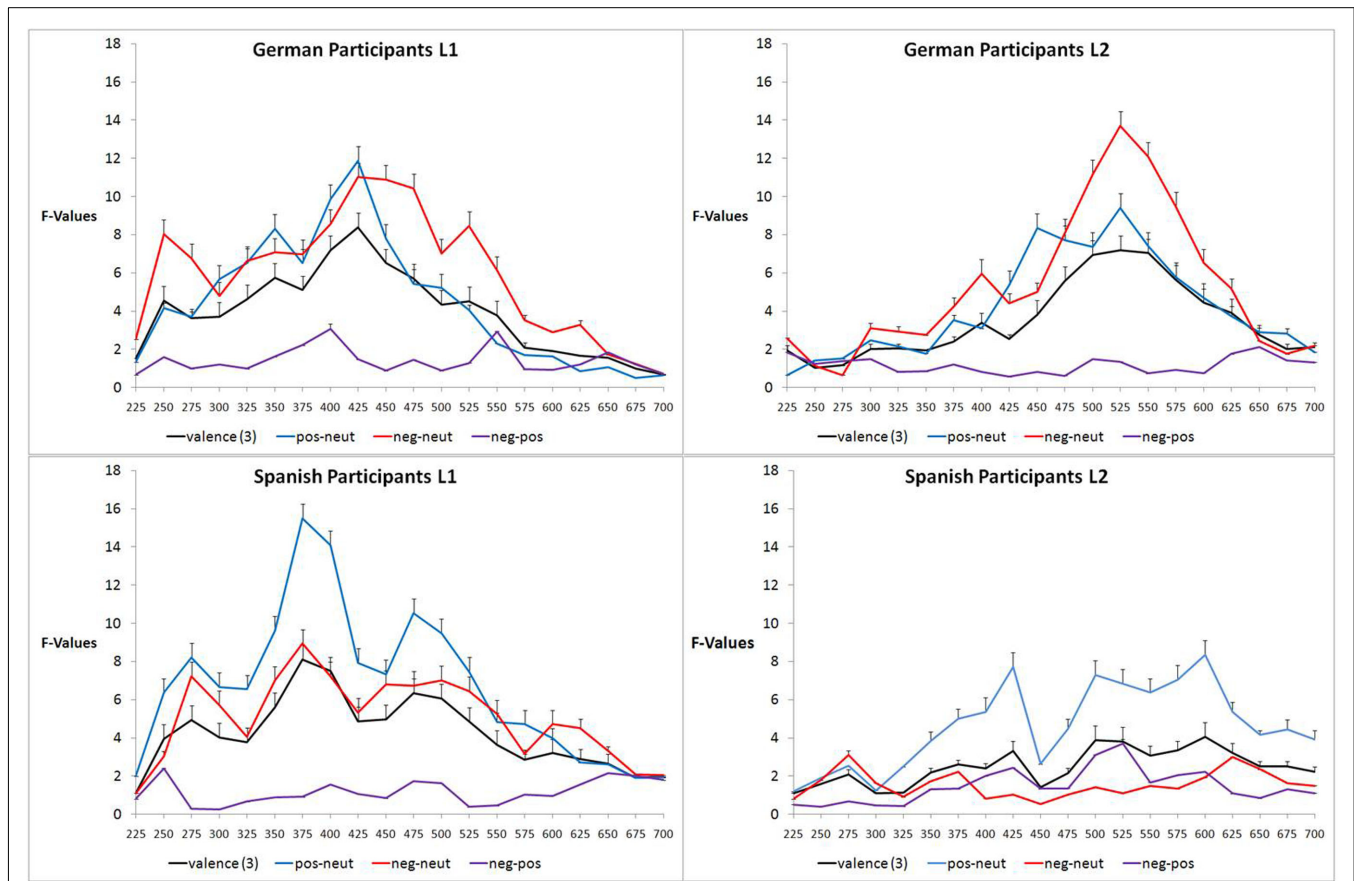
FIGURE 1 | *F*-values from ANOVAs for 26 Spanish and 26 German participants over mean ERP amplitudes in 25 ms time epochs from 200 ms until 700 ms (numbers represent the end of respective time windows) after stimulus onset. Reported *F*-values refer to interactions of the factors electrode and valence. Error bars give levels of significance ($p < 0.0001$, $p < 0.001$, and $p < 0.05$ are represented as numerical values 0.75, 0.5, and 0.25). *p*-Values are Bonferroni corrected for *post hoc* comparisons between pos–neut, neg–neut, and neg–pos.

German participants (N = 40) L1

For the group of German speakers reading words in L1 (German) valence yielded significant effects in all intervals between 200 and 675 ms, all Fs $(118,4602) > 2.2$; all $ps < 0.018$; $\varepsilon s = 0.059$–0.087. *Post hoc* comparisons revealed significant differences in ERP amplitudes between positive and neutral words in all time intervals between 225 and 525 ms, Fs $(59,2301) > 3.2$; $ps < 0.05$; $\varepsilon s = 0.057$–0.082. **Figure 3** depicts global field power (GFP) of ERP amplitudes for all participant groups and task languages. As evident from **Figure 3**, GFP increased for positive compared to neutral words and topographic maps (based on the time window between 200 and 425 ms) of difference wave comparisons (positive minus neutral) reveal enhanced negativity in posterior electrode sites for positive as compared to neutral words, representing the typical topography for the EPN. The same comparison in a later interval (425–600 ms) reveals enhanced centro-parietal positivity for positive words – typical for the LPC. Thus, our data replicates previous accounts for ERP effects of emotional valence in German native language processing (e.g., Schacht and Sommer, 2009a).

Interestingly, closely comparable effects were present for comparisons between negative and neutral words: ANOVAs provided significant effects for all intervals between 200 and 625 ms, Fs $(59,2301) > 3.1$; $ps < 0.045$; $\varepsilon s = 0.057$–0.102. Topographic maps corresponding to the same windows as used for positive–neutral comparisons reveal that also negative words produced EPN and LPC components when compared to neutral ones. Note also that even between negative and positive words significant effects emerged in time windows between 450 and 675 ms Fs $(59,2301) > 3.1$; $ps < 0.05$; $\varepsilon s = 0.061$–0.072, with increased GFP (see **Figure 3**) for negative relative to positive words.

German participants (N = 40) L2

Analyses of variances revealed significant effects of valence in all time epochs from 250 to 700 ms, Fs $(118,4602) > 2.1$; $ps < 0.043$; $\varepsilon s = 0.056$–0.079. According to single comparisons, effects held true for both positive, Fs $(59,2301) > 3.0$; $ps < 0.036$; $\varepsilon s = 0.056$–0.086, and negative words, Fs $(59,2301) > 3.6$; $ps < 0.036$; $\varepsilon s = 0.051$–0.082, presented in L2 compared to neutral ones in consecutive time intervals 300–700 ms and 275–700 ms respectively. Most interestingly, topographic maps for differences for either positive or negative minus neutral words in time intervals 300–450 ms and 450–700 ms in L2 reveal most similar EPN and LPC components as the ones obtained for the same participants processing the same words in L1. But note also that the onset of valence effects in
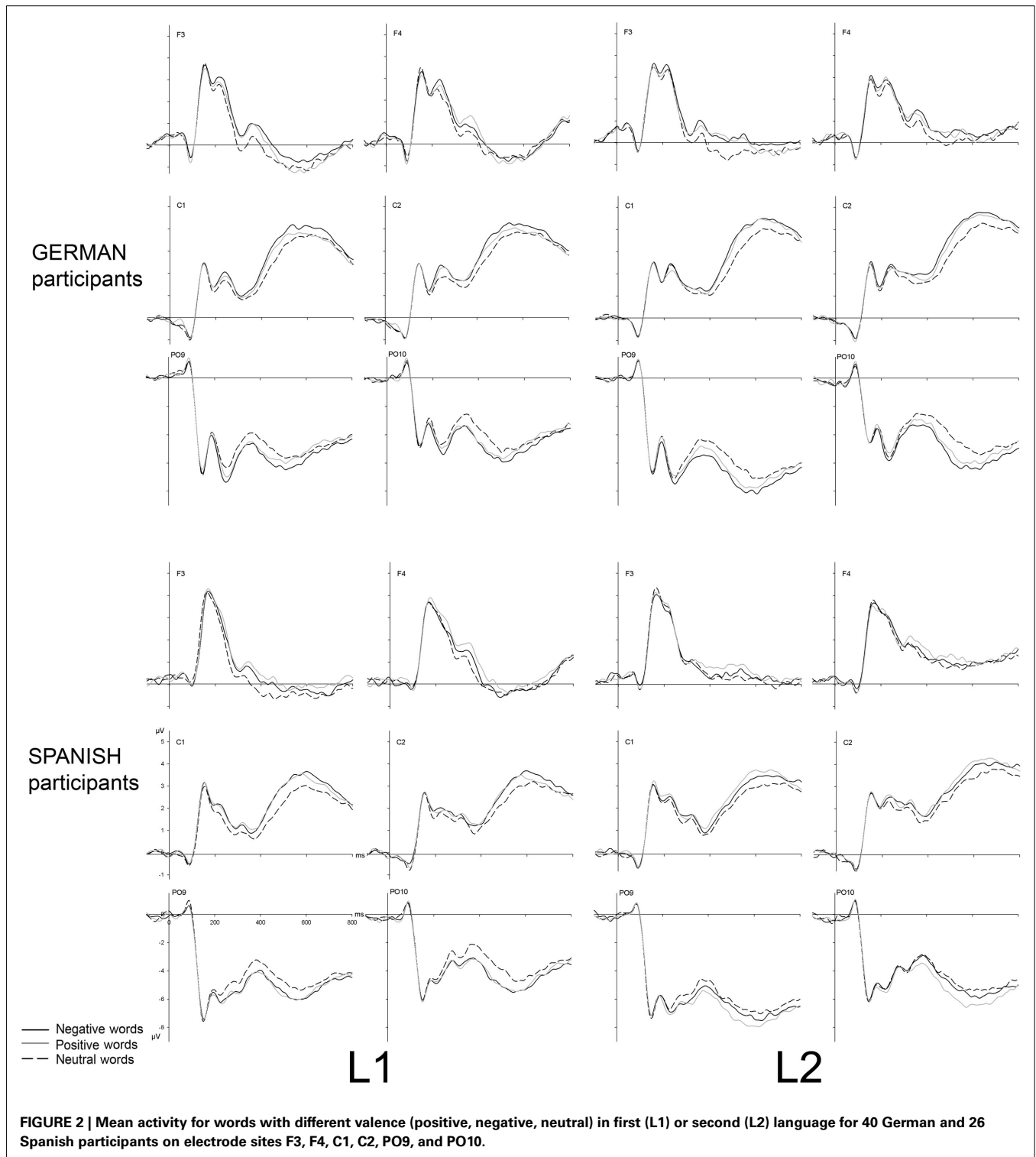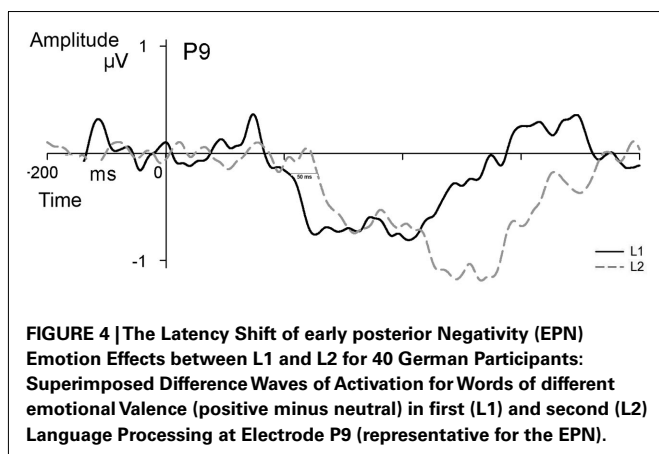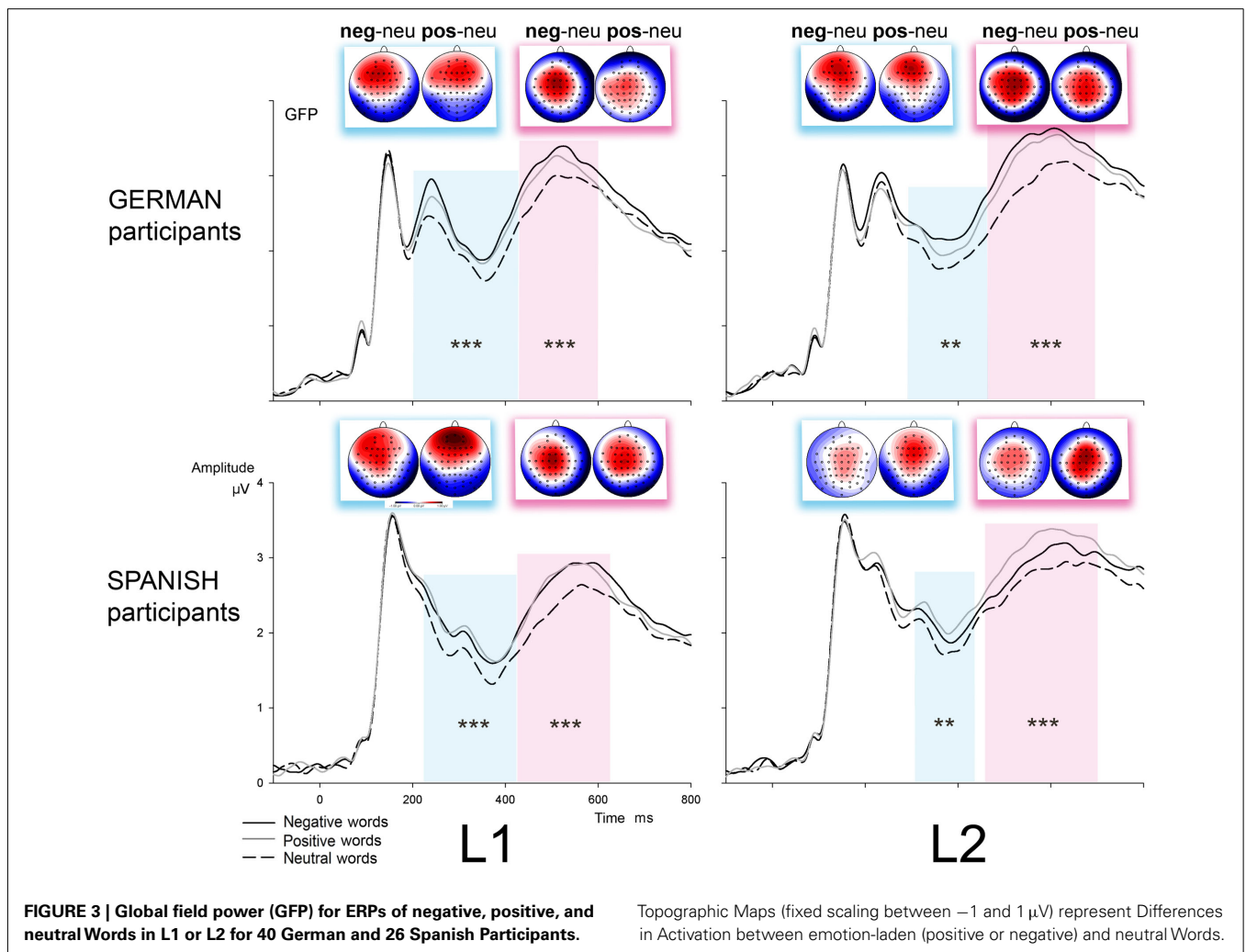
FIGURE 2 | Mean activity for words with different valence (positive, negative, neutral) in first (L1) or second (L2) language for 40 German and 26 Spanish participants on electrode sites F3, F4, C1, C2, PO9, and PO10.

the L2 data for German participants was 50 ms later as compared to L1 (in the respective analyses involving valence as three levels factor). **Figure 4** represents this latency shift superimposing difference waves for respective valence effects (positive − neutral words) in L1 and L2 for the 40 German participants at an electrode site representative for the EPN (P9).

Spanish participants (N = 26) L1

For the group of Spanish speakers, ANOVAs revealed effects of valence in L1 (Spanish) in all intervals between 225 and 650 ms, Fs (118,2950) > 2.6; $ps < 0.01$; εs = 0.069–0.093. Similarly to the German group in L1, for both positive, Fs (59,1475) > 4.7; $ps < 0.001$; εs = 0.064–0.095, and negative words,

FIGURE 3 | Global field power (GFP) for ERPs of negative, positive, and neutral Words in L1 or L2 for 40 German and 26 Spanish Participants. Topographic Maps (fixed scaling between −1 and 1 μV) represent Differences in Activation between emotion-laden (positive or negative) and neutral Words.

FIGURE 4 | The Latency Shift of early posterior Negativity (EPN) Emotion Effects between L1 and L2 for 40 German Participants: Superimposed Difference Waves of Activation for Words of different emotional Valence (positive minus neutral) in first (L1) and second (L2) Language Processing at Electrode P9 (representative for the EPN).

Fs (59,1475) > 3.0; ps < 0.05; εs = 0.066–0.094, ERP amplitudes differed from the ones for neutral words between 225 and 600 ms and 225–650 (with the exception of 550–575 ms with p-corrected = 0.054) respectively. Unlike in the analyses with 40 German participants' L1 data, no consecutive significant effects

between positive and negative words were found, topographic maps for comparisons of emotional (both positive or negative) minus neutral words for Spanish speakers in L1 also show the posterior negativity typical of the EPN component in the interval 225–425 ms, and enhanced centro-parietal positivity for the LPC between 425 and 625 ms. Thus, our data corroborates that these effects are robust and mostly comparable across different languages – obtained for both positive and negative words compared to neutral ones in both cases of native language processing in the present study.

Spanish participants (N = 26) L2

Effects of valence in the group of Spanish speakers in L2, again, started later than in L1: They were present in all time windows between 325 and 700 ms (except 425–450 ms; ps < 0.04) Fs (118,2950) > 2.0; ps < 0.04; εs = 0.057–0.079. This time, effects were restricted to the comparison between positive and neutral words, Fs (59,1475) > 3.8; ps < 0.015; εs = 0.055–0.107, with positive words eliciting larger ERP amplitudes than neutral ones (see **Figures 2** and **3**). No other comparisons were significant after Bonferroni corrections. Again, topographic maps (based on the same intervals as for the German participants' L2 data, except

425–450 ms being removed from the present interval for the EPN) show similar distributions for the EPN and LPC components as the ones obtained for the same participants reading words in L1- at least when comparing positive to neutral words. The same holds true comparing German and Spanish participants' L2 data. Note that also in the topographic maps differences between negative and neutral words appear clearly attenuated for Spanish participants' L2 data.

REANALYSES FOR A REDUCED SAMPLE OF GERMAN PARTICIPANTS (N = 26)

The fact that our participant groups differed not only in numbers, but also – to some degree – in second language proficiency represents a potential problem for the interpretation of effects differing across the two participant groups. In particular, the question arises of whether the absence of an effect between negative and neutral words in the Spanish L2 data (present in both the L1 and L2 data of German participants) might be due to decreased statistical power in the Spanish sample or might be a consequence of the distinct linguistic L2 profiles of the two groups. To meet this problem we selected 26 out of our German native speakers matched to Spanish native speakers on second language proficiency for new analyses of the German native speakers' ERP data. We had mentioned above that the best measure of individual L2 proficiency might be obtained from the ratio of individual response latencies to words in L1 and L2. We thus selected 26 out of the 40 German native participants controlling for the mismatch on this and other variables present between the two complete samples (40 German vs. 26 Spanish participants) used for the analyses presented above. No more significant differences between these subsamples of 26 each individuals in the German and the Spanish group were found on the following variables: RT L1/RT L2, sex, years of learning contact with L2. Respective differences were at least strongly reduced concerning months of living experience in the L2 country (see **Table 1**). The same analyses of ERP data as presented above were conducted for this reduced German participant sample representing a better match of L2 linguistic profile and proficiency to the Spanish one.

German participants (N = 26) L1

Effects closely resembled the ones obtained for the complete participant set (at least concerning the comparisons between either positive or negative and neutral words): ANOVAs revealed significant effects of valence in all time epochs from 225 to 575 ms, Fs $(118, 2950) > 2.09$; $ps < 0.034$; $\varepsilon s = 0.062$–0.084. According to single comparisons, effects held true for both positive, Fs $(59, 1475) > 3.7$; $ps < 0.042$; $\varepsilon s = 0.053$–0.085, and negative words, Fs $(59, 2301) > 3.5$; $ps < 0.024$; $\varepsilon s = 0.003$–0.072, presented in L1 compared to neutral ones in consecutive time intervals 225–525 ms and 225–575 ms respectively. In contrast to the previous analyses, no significant effects in consecutive time windows were present between conditions of positive and negative words for this reduced participant sample.

German participants (N = 26) L2

Also for L2 processing, effects for this reduced set of participants resembled the ones obtained for the complete participant set:

ANOVAs revealed significant effects of valence in all time epochs from 275 to 700 ms, Fs $(118, 2950) > 2.04$; $ps < 0.040$; $\varepsilon s = 0.058$– 0.089. According to single comparisons, effects held true for both positive, Fs $(59, 1475) > 3.9$; $ps < 0.045$; $\varepsilon s = 0.068$–0.097, and negative words, Fs $(59,2301) > 2.7$; $ps < 0.048$; $\varepsilon s = 0.066$–0.093, presented in L2 compared to neutral ones in consecutive time intervals 350–675 (with the exception of 375–400 ms with p-corrected $= 0.057$) ms and 275–625 ms (with the exception of 325–350 ms with p-corrected $= 0.087$) respectively.

RESULTS FOR ANOVAS ON EXTENDED TIME WINDOWS

To further explore our data, and to provide a comprehensive overview allowing for direct comparisons of effect sizes across participant groups and L1 vs. L2 performance we conducted additional ANOVAs collapsing time windows where significant EPN and LPC effects had appeared in the explorative analyses for each 26 Spanish and German participants presented above (see **Figure 1**).

L1 Data

Concerning the EPN, ANOVAs based on a time window from 225 to 425 ms yielded significant effects for both the German and the Spanish native data.

For positive words (compared to neutral words) respective effects were F $(59,1475) = 14.98$; $p < 0.0001$; $\eta_p^2 = 0.373$ in the Spanish, and F $(59,1475) = 10.51$; $p < 0.0001$; $\eta_p^2 = 0.296$ in the German native data.

When comparing negative to neutral words, respective effects were F $(59,1475) = 9.47$; $p < 0.0001$; $\eta_p^2 = 0.275$ in the Spanish and F $(59,1475) = 12.07$; $p < 0.0001$; $\eta_p^2 = 0.326$ in the German native data.

Unlike for these EPN effects – according to the explorative analyses – onsets and duration of LPC effects slightly differed for either negative or positive as compared to neutral words and also between the Spanish and German native data. We, therefore, selected specific time windows for the ANOVAs that best reflected the specific LPC effects. Effects were analyzed and resulted significant for positive as compared to neutral words between 425 and 600 ms for the Spanish, F $(59,1475) = 9.53$; $p < 0.0001$; $\eta_p^2 = 0.276$, and between 425 and 525 ms for the German native data, F $(59,1475) = 6.59$; $p < 0.0001$; $\eta_p^2 = 0.209$. Analyses of time windows between 425 and 625 ms yielded LPC effects between negative and neutral words in both the Spanish, F $(59,1475) = 7.89$; $p < 0.0001$; $\eta_p^2 = 0.240$, and the German native data, F $(59,1475) = 8.96$; $p < 0.0001$; $\eta_p^2 = 0.264$.

L2 Data

Early posterior negativity effects for positive words (as compared to neutral ones) when reading L2 were analyzed and proved significant between 325 and 425 ms for Spanish participants, F $(59,1475) = 7.34$; $p < 0.0001$; $\eta_p^2 = 0.227$, and between 350 and 450 ms for German participants, F $(59,1475) = 4.18$; $p < 0.003$; $\eta_p^2 = 0.143$. For negative words, respective analyses were based on the time window 275–450 ms and resulted in a significant effect only for the German, F $(59,1475) = 6.26$; $p < 0.0001$; $\eta_p^2 = 0.200$, but not for the Spanish participants, F $(59,1475) = 1.43$; $p > 0.2$; $\eta_p^2 = 0.054$, when reading L2 words.

Late positive complex effects were analyzed for positive words as compared to neutral ones between 450 and 675 ms resulting in significant effects for both Spanish, F (59,1475) = 8.07; $p < 0.0001$; $\eta_p^2 = 0.244$, and German participants, F (59,1475) = 8.08; $p < 0.0001$; $\eta_p^2 = 0.244$.

Concerning LPC effects between negative and neutral words in L2 processing, analyses were based on a time window between 450 and 625 ms and resulted significant only for the German, F (59,1475) = 13.56; $p < 0.0001$; $\eta_p^2 = 0.352$, but, again, not for the Spanish participants, F (59,1475) = 1.90; $p > 0.1$; $\eta_p^2 = 0.071$.

The typical morphology of EPN and LPC effects – as evident in **Figure 3** – obtained in these analyses on extended time windows was supported by additional analyses testing for significance of respective effects at single electrode sites. It appears that EPN effects are generally most robust at frontal and posterior electrode sites (but mostly absent at central electrodes) whereas LPC effects are strongest at central electrode positions, but either attenuated or absent at frontal and posterior electrode sites.

Looking at the global picture of all ERP effects presented above, the most important finding seems that EPN and LPC effects for emotional words can be obtained in both L1 and L2 processing for both participant groups– at least concerning positive words. Only with regard to negative words our results display one unpredicted differential pattern concerning effects across L1 and L2 reading performance and participant groups: for German native speakers significative differences between negative and neutral words when processing Spanish words were found, but such differences were absent when Spanish native speakers processed German words. Note that these differences hold true for the two participant sets matched on numbers and – as closely as possible – second language proficiency. To test for statistical significance of these apparent group differences with regard to sensitivity to negative emotional word content in L2 processing we conducted multifactorial ANOVAs with valence (negative vs. neutral) as within-subject and participant group (Native German vs. native Spanish) as between-subjects factor for the two groups' respective L2 ERP data. ANOVAs revealed a tendency for an interaction between effects of emotional valence and participant group for the mean activity across the EPN time window where significant effects had been obtained for the German participants (275–450 ms), F (2,2950) = 1.75; $p = 0.141$; $\eta_p^2 = 0.034$. A significant result for the same interaction was obtained for the respective LPC time window (450–625 ms), F (2,2950) = 2.71; $p = 0.029$; $\eta_p^2 = 0.051$. No such interaction between group and valence effects concerning negative and neutral words, on the other hand, was obtained running the same analyses for the two groups' L1 data (using 225–425 ms for EPN and 450–625 ms for LPC effects), $Fs < 1$, suggesting that the two groups differed especially in their L2 sensitivity to negative stimuli. But note also that concerning positive–neutral comparisons, effect sizes where always greater in the Spanish than in the German group in both L1 and L2 data, whereas the inverse was the case for comparisons between negative and neutral words (see also **Figures 1**–**3**).

GENERAL DISCUSSION
The central topic of the present study is the question of emotionality of language processing in L2. Previous studies mainly used either behavioral data from word recall (Anooshian and Hertel, 1994; Ayçiçegi and Harris, 2004), priming (Altarriba and Canary, 2004, see also Altarriba, 2006), naming latencies in Stroop paradigms (Eilola et al., 2007; Sutton et al., 2007), or physiological parameters as skin conductance responses (Harris et al., 2003, 2006; Harris, 2004) to emotional words to compare emotional word processing in L2 and L1. The present study, documenting response latencies and ERPs for emotion-laden words presented to two participant groups (German and Spanish native speakers) in both L1 and L2 context of a visual lexical decision directly taps into the time course of the effects in question.

EMOTION EFFECTS IN THE BEHAVIORAL DATA
Already the behavioral data of the present lexical decision task provides the novel finding that a processing speed advantage for emotional words is not limited to L1: Both positive and negative words in the Spanish lexical decision task were responded to faster than neutral words not only by Spanish native speakers, but also by German participants reading them as L2 words (see method section for a detailed discussion of behavioral data from the German lexical decision task). From a more general perspective, these findings corroborate recent proposals (Kousta et al., 2009) that a processing advantage for emotional words would generalize over both the positive and negative domain of emotional valence. Yet, it should be noted that respective RT effects were more pronounced for positive than for negative words and that RT differences between negative and neutral Spanish words were only marginally significant in the L2 data. Concerning the emotionality of L2 processing, it should also be noted that effect sizes of respective valence effects in the RT data decreased from L1 to L2 processing, but it seems hard to tell whether this indicates an attenuation of emotion sensitivity in L2 or rather results from general noise in the RT data also increasing from L1 to L2. Whereas, thus, generally comparable effects for emotional valence influencing lexical decision response latencies were found across L1 and L2, the error data displayed a different pattern: Here, significant effects were present for both German and Spanish words only when read as L2 words. We propose that word recognition errors in L2 might reflect a special role of emotion during second language acquisition: Words with emotional content might preferentially attract learners' attention and seem to be memorized more efficiently – leading to increased recognition rates in the present experiment. Most interestingly, this holds true for both types of emotional words (positive and negative ones) for the German participants when presented with L2 material, whereas Spanish participants' recognition rates were worst for negative words. This might indicate a specific positivity bias in this group's second language acquisition, an issue we will refer to again in the discussion of the ERP data.

ERP EMOTION EFFECTS IN L1
For ERP analyses, we focused on two components of the ERP signal previously described (e.g., Junghöfer et al., 2006; Kissler et al., 2006; Schacht and Sommer, 2009a) as characteristic to emotional word valence in native language processing: an EPN and a LPC. In the present study, both components were found to be affected by emotional valence when words had to be read in L1 in consistent

ways for both languages used in the present experiment: Significant effects representing an EPN topography were found in the comparison between positive and neutral words between 225 and 425 ms in the German and Spanish L1 data, and the same pattern was present for the comparison between negative and neutral words in both cases (with an even earlier onset at 200 ms in the 40 participants' German L1 data). Accordingly, in both languages – when read as L1 – for both negative and positive words (as compared to neutral ones) the typical morphology of LPC effects was present in the time window between 425 and 625 ms. Thus, our data not only show that ERP effects of emotional valence can be obtained for different languages, respective effects also appear not to be restricted to positive words, but to be of comparable strength and nature for negative stimuli.

Although also ERP effects between negative and neutral words have been previously reported, they used to be less pronounced than effects for positive words (see Junghöfer et al., 2006; Kissler et al., 2006), especially concerning the earlier EPN component (EPN effects had been restricted to positive words, e.g., in the data of Schacht and Sommer, 2009a, but see also Palazova et al., 2011). Our data, therefore, make a strong case suggesting valence effects to generalize across the positive and negative domain (see Kuchinke et al., 2005; Kousta et al., 2009).

The EPN for emotion-laden words is generally interpreted as an early attention shift to words with emotional content – intimately bound to lexical access (see Schacht and Sommer, 2004, 2009a; Kissler et al., 2007).

In particular, the respective onsets of effects of Lexicality and EPN for emotional words in referential studies investigating valence effects for German words (Schacht and Sommer, 2004, 2009a) followed each other with a relatively small time lapse of 50 ms. Emotion effects starting around 380 ms lead these authors to interpret valence effects as post-lexical in nature. This claim has been maintained by Palazova et al. (2011) although respective onsets of Lexicality and Valence effects in their study coincided at 300 ms. In our data, EPN effects in L1 appeared even earlier – at around 200 or 225 ms – (see Kissler et al., 2007; Hofmann et al., 2009, for reports of comparably early ERP effects for emotional words) and they either coincided or even slightly preceded the moment where respective ERP curves for words and non-words started to distinguish (normally understood as the moment when lexical access has occurred and when word semantics, in consequence, should get available). In this context, it seems important to mention that the earliest visually observable distinction in GFP-difference waves between words and non-words in our data for 40 German participants appeared at 230 ms in L1 and at 270 ms in the L2 data (260 and 280 ms respectively for the 26 Spanish participants). We do not propose that such a pattern of results would imply that effects of emotional valence were of prelexical nature – clearly, emotional word content is a semantic propriety and should therefore only influence word processing after some information of word meaning is available. But, on the other hand, the labeling of emotional effects as "post-lexical" seems incompatible with the repeatedly obtained finding of speeded lexical decision responses to emotional words. If response latencies are understood as reflecting the moment when lexical access to a word occurs, then any process speeding response latencies can not be purely post-lexical.

An analogy to computational models of visual word recognition might help to reconcile these different views: We propose that the appearance of differences in ERP waves for words and non-words might not precisely reflect the first moment when some semantic information for words becomes available. Both theoretical accounts and computational models of visual word recognition (e.g., Grainger and Jacobs, 1996) assume (or implement) that non-words presented in a lexical decision task would activate word representations. In consequence, they should – in the same way as word stimuli – also activate semantic information. In the MROM of Grainger and Jacobs (1996) activation levels for word representations continuously increase for a number of processing cycles before a lexical decision corresponding to word identification is made when activation reaches a threshold of 90% of the corresponding asymptotic function. They also do so when a non-word stimulus is presented – they only rarely reach the threshold corresponding to word identification in this case. Accordingly, some activation of a word stimulus' meaning might have occurred before ERP waves for words and non-words start to distinguish. More precisely, this specific point might rather reflect the moment when searching for a stimulus word's meaning comes to an end – whereas it continues in the case of non-words, for which no exact match in the mental lexicon can be found. The comparable latencies for EPN effects of emotional valence and Lexicality in our data might best be understood assuming that words' emotional valence helps to achieve lexical access – representing rather the end of a decision process than the very beginning of semantic activation – due to an early shift of attention toward emotional stimuli. Already during early processing stages an emotional word's semantics might have triggered this attention shift before the reading system is sure whether the stimulus represents a word or not. From a general perspective, such an attention shift – occurring before lexical access is completed – seems plausible even when the potentially rewarding or, in particular, aversive character of a stimulus is yet just a vague idea rather than an established certainty. Otherwise, the consistently obtained speeded response latencies to emotion-laden words in lexical decision (see Kousta et al., 2009) would be hard to explain if semantic information including emotional valence would only be available after lexical access was completed. The general view that emotional valence would speed the process of lexical access, instead, is well in line with the finding of ERP effects of emotional valence (EPN) and Lexicality arising at the same time in our data (see also Palazova et al., 2011).

ERP EMOTION EFFECTS IN L2

As to the comparability of emotion effects across word L1 and L2 processing, the specific question at hand was whether comparable EPN and LPC effects as in L1 processing would also be observed in L2.

Previous empirical studies reporting emotional effects in L2 could not answer the question of whether an assignment of emotional relevance to words in L2 would be immediate – occurring during early phases of word perception or visual word recognition – or would rather result from an internal re-translation of respective words to L1- whereby words might gain emotional relevance only via the emotional context of their L1 translations. In particular, such an indirect access to L2 words' emotional content

might prevent early EPN effects of valence in the L2 ERP data from arising, because of the relation of this component to early attention and lexical access processes – presumably preceding any re-translation from L2 to L1.

In our data, emotional valence yielded significant EPN effects for the comparisons between positive and neutral words for both groups of Spanish (325–425 ms) and German participants (300–450 ms; 26 participants: 350–450 ms) when presented with words in their respective L2. A similar effect was present between negative and neutral words, but, only for German participants reading Spanish words (yet, again, starting especially early: at 275 ms).

These findings suggest that emotion sensitivity must be understood as a mandatory feature occurring at early stages of L2 processing, respective onsets of EPN effects shifted only about 50 ms (in the case of earliest L2 EPN effects) or 100 ms between L1 and L2 data. Beside the striking similarity in the morphologies of the respective effects, these rather small time delays suggest that no qualitative differences seem to be given concerning early processing of emotional word content in L1 and L2. In other words, the 50- to 100-ms time shift between L1 and L2 EPN latencies in our data seems to reflect rather generally delayed visual word recognition processes (see van Heuven and Dijkstra, 2010, for a review), but not specifically delayed emotion processing in L2.

And, unless for lexical access in L2 activation of L1 word representations would be mandatory (see Kroll and Stewart, 1994, for the proposal of early L2–L1 translation in bilingual word recognition, see also Midgley et al., 2008, 2011; Geyer et al., 2011, for ERP studies on bilinguals' simultaneous processing of two languages), it appears implausible that – if emotional relevance was assigned to words only via re-translation to L1 – such a process would coincide with lexical access in L2.

Also for the LPC -generally interpreted as reflecting higher level semantic processing of stimuli (see Kissler et al., 2006; Schacht and Sommer, 2009a) – our data show that it can consistently be obtained for emotion-laden words in second language processing.

Significant effects with LPC morphology were obtained for both positive and negative words when compared to neutral ones in time windows between 450 and 700 ms when German participants read Spanish L2 words. Interestingly, the same type of effect was obtained in the case of Spanish participants reading German L2 words only concerning positive, but not when comparing negative to neutral words (though significant EPN and LPC effects for both negative and positive words were present in L1 for the same participants).

This represents an interesting case of a dissociation of effects between the positive and the negative domain across L1 and L2 on the one hand, and between our two different participant groups on the other. Note that this pattern of effects was replicated in additional analyses for a reduced sample of participants with identical group sizes and improved matching of L2 proficiency between the two groups of native German and native Spanish speakers. Furthermore, these differences resulted in a significant interaction between effects of valence and participant group, which should no longer have arisen via phenomena of differing group size or second language proficiency.

We suggest that this pattern of results might best be explained as a consequence of the following systematic difference between our two participant groups that remained, even when respective L2 proficiency was controlled for: All Spanish participants – being tested in Berlin – were actually living in the L2 country and, therefore, currently involved in an intense ongoing process of both improving their academic L2 knowledge and collecting life experiences in a foreign country and with a new culture. We propose that these specific conditions – normally described as emotionally highly stimulating and enriching – might have formed the ground for the specific observed positivity bias observed in Spanish participants' L2 processing (suggested by both the absence of ERP effects and increased error rates for negative L2 words). Our Spanish participants' L2 data might, therefore, partially reflect a "euphoric" stage of second language acquisition: When coinciding with the positive life experience that a student's stay in a foreign country usually is, the general emotional attitude vs. an L2 and vs. the culture it represents might be characterized by a positivity bias that became observable in our ERP data (see Schumann, 1998, for the claim that L2 proficiency is strongly related to motivational aspects of second language learning where positive emotions determine the base of interest in a second language). Although most of our German participants had also reported stays in a Spanish speaking country and probably undergone the same positive emotions, they were not currently exposed to the L2 culture, and the specific positivity bias presumably accompanying L2 acquisition and perception might already have faded in their case.

Clearly, this remains a speculative argument, and at least one alternative explanation for this specific aspect of the present results should be mentioned: Both visual inspection of ERP curves and comparison of effect sizes for statistical analyses reported above converge on the following: ERP effects were always stronger for negative than for positive words in the German participants' L1 and L2 data, but always stronger for positive than for negative words in the Spanish participants' L1 and L2 data. Although these apparent differences did result in significant interactions reflecting differential patterns of positivity/negativity bias between our two participant groups only for the L2, but not for the L1 data, it can not be excluded that native German participants' general language processing – in contrast to native Spanish language processing would be characterized by a more pronounced negativity bias. Therefore, the fact that effects for negative words in the L2 data were present only for the German participants might possibly reflect an export of a culture specific negativity bias into their L2 processing.

CONCLUSION

Our ERP data, in the first place, corroborate previous reports of electrophysiological correlates (EPN and LPC) of emotional word valence during native language processing showing, in particular, that effects are generally comparable across domains of positive and negative valence. Latencies of EPN effects, in particular in the data from 40 German native speakers were even earlier for negative words (200 ms) than for positive words (225 ms) potentially reflecting a bias to process stimuli representing threat as early as possible. In general, latencies for EPN emotion effects in our ERP data mostly coincided or even slightly preceded effects of Lexicality suggesting that emotional valence helps to achieve lexical access,

an argument that is supported also by speeded response latencies to emotional words.

With regard to second language processing, we could show that it also seems characterized by an early and direct sensitivity to emotional content – as evident from EPN and LPC effects that were closely comparable to L1 processing. Respective onsets of EPN emotion effects shifted across L1 and L2 processing about 50–100 ms suggesting only generally delayed L2 processing rather than qualitative differences with regard to the processing of emotion across L1 and L2.

REFERENCES

Altarriba, J. (2003). Does cariño equal "liking?" A theoretical approach to conceptual nonequivalence between languages. *Int. J. Billing.* 7, 305–322.

Altarriba, J. (2006). "Cognitive approaches to the study of emotion-laden and emotion words in monolingual and bilingual memory," in *Bilingual Minds: Emotional Experience, Expression, and Representation*, ed. A. Pavlenko (Clevedon: Multilingual Matters), 232–256.

Altarriba, J., and Canary, T. (2004). The influence of emotional arousal on affective priming in monolingual and bilingual speakers. *J. Multiling. Multicult. Dev.* 25, 248–265.

Anooshian, J. L., and Hertel, P. T. (1994). Emotionality in free recall: language specificity in bilingual memory. *Cogn. Emot.* 8, 503–514.

Ayçiçegi, A., and Harris, C. (2004). Bilinguals' recall and recognition of emotion words. *Cogn. Emot.* 18, 977–987.

Baayen, R. H., Piepenbrock, R., and van Rijn, H. (1993). *The CELEX Lexical Database (CD-ROM)*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Barber, H., Vergara, M., and Carreiras, M. (2004). Syllable-frequency effects in visual word recognition: evidence from ERPs. *Neuroreport* 15, 545–548.

Bond, M. H., and Lai, T. (1986). Embarrassment and code-switching into a second language. *J. Soc. Psychol.* 126, 179–186.

Bradley, M. M., and Lang, P. J. (1999). *Affective Norms for English Words (ANEW)*. Gainesville, FL: University of Florida; NIMH Center for the Study of Emotion and Attention.

Bühler, K. (1992). *Sprachtheorie. Die Darstellungsfunktion der Sprache.* Stuttgart: Fischer.

Chauncey, K., Grainger, J., and Holcomb, P. J. (2008). Code-switching effects in bilingual word recognition: a masked priming study with ERPs. *Brain Lang.* 105, 161–174.

Dewaele, J.-M. (2004). The emotional force of swearwords and taboo words in the speech of multilinguals. *J. Multiling. Multicult. Dev.* 25, 204–222.

Dietrich, D., Waller, C., Johannes, S., Wieringa, B. M., Emrich, H. M., and Münte, T. F. (2001). Differential effects of emotional content on event-related potentials in word recognition memory. *Neuropsychobiology* 43, 96–101.

Dolan, R. J. (2002). Emotion, cognition, and behavior. *Science* 298, 1191–1194.

Eilola, T. M., Havelka, J., and Sharma, D. (2007). Emotional activation in the first and second language. *Cogn. Emot.* 21, 1064–1076.

Eviatar, Z., and Zaidel, E. (1991). The effects of word length and emotionality on hemispheric contribution to lexical decision. *Neuropsychologia* 29, 415–428.

Geyer, A., Holcomb, P. J., Midgley, K. J., and Grainger, J. (2011). Processing words in two languages: an event-related brain potential study of proficient bilinguals. *J. Neurolinguistics* 24, 338–351.

Gillon-Dowens, M., Vergara, M., Barber, H. A., and Carreiras, M. (2010). Morpho-syntactic processing in late L2 learners. *J. Cogn. Neurosci.* 22, 1870–1887.

Gonzales-Reigosa, F. (1976). "The anxiety-arousing effect of taboo words in bilinguals," in *Cross-Cultural Anxiety*, eds C. D. Spielberger and R. Diaz-Guerrero (Washington, DC: Hemisphere), 89–105.

Grainger, J., and Jacobs, A. M. (1996). Orthographic processing in visual word recognition: a multiple read-out model. *Psychol. Rev.* 103, 518–565.

Harris, C. (2004). Bilingual speakers in the lab: psychophysiological measures of emotional reactivity. *J. Multiling. Multicult. Dev.* 25, 223–247.

Harris, C., Ayçiçegi, A., and Gleason, J. (2003). Taboo words and reprimands elicit greater autonomic reactivity in a first than in a second language. *Appl. Psycholinguist.* 4, 561–578.

Harris, C., Gleason, J., and Ayçiçegi, A. (2006). "When is a first language more emotional? Psychophysiological evidence from bilingual speakers," in *Bilingual Minds: Emotional Experience, Expression, and Representation*, ed. A. Pavlenko (Clevedon: Multilingual Matters), 257–283.

Hofmann, M. J., Kuchinke, L., Tamm, S., Võ, M. L.-H., and Jacobs, A. M. (2009). Affective processing within 1/10th of a second: high arousal is necessary for early facilitative processing of negative but not positive words. *Cogn. Affect. Behav. Neurosci.* 9, 389–397.

Holcomb, P. J., Grainger, J., and O'Rourke, T. (2002). An electrophysiological study of the effects of orthographic neighborhood size on printed word perception. *J. Cogn. Neurosci.* 14, 938–950.

Huhta, A., Luoma, S., Oscarson, M., Sajavaara, K., Takala, S., and Teasdale, A. (2002). "DIALANG – a diagnostic language assessment system for learners," in *Case Studies of the Use of the Common European Framework*, ed. J. C. Alderson (Strasbourg: Council of Europe), 130–145.

Hutzler, F., Bergmann, J., Conrad, M., Kronbichler, M., Stenneken, P., and Jacobs, A. M. (2004). Inhibitory effects of first syllable-frequency in lexical decision: an event related potential study. *Neurosci. Lett.* 372, 179–184.

Jakobson, R., and Halle, M. (1969). *Die Grundlagen der Sprache. Schriften zur Phonetik, Sprachwissenschaft und Kommunikationsforschung (1).* Berlin: Akademie Verlag.

Junghöfer, M., Peyk, P., Flaisch, T., and Schupp, H. (2006). "Neuroimaging methods in affective neuroscience: selected methodological issues," in *Understanding Emotions*, Vol. 156, eds G. Anders, G. Ende, M. Junghöfer, J. Kissler, and D. Wildgruber (Amsterdam: Elsevier), 123–143.

Kissler, J., Assadollahi, R., and Herbert, C. (2006). Emotional and semantic networks in visual word processing: insights from ERP studies. *Prog. Brain Res.* 156, 147–183.

Kissler, J., Herbert, C., Peyk, P., and Junghöfer, M. (2007). Buzzwords: early cortical responses to emotional words during reading. *Psychol. Sci.* 18, 475–480.

Kotz, S. A., and Paulmann, S. (2007). When emotional prosody and semantics dance cheek to cheek: ERP evidence. *Brain Res.* 1151, 107–118.

Kousta, S. T., Vinson, D. P., and Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition* 112, 473–481.

Kroll, J. F., and Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *J. Mem. Lang.* 33, 149–174.

Kuchinke, L., Jacobs, A. M., Grubich, C., Võ, M. L.-H., Conrad, M., and Herrmann, M. (2005). Incidental effects of emotional valence in single word processing: an fMRI study. *Neuroimage* 28, 1022–1032.

Kuchinke, L., Jacobs, A. M., Võ, M. L.-H., Conrad, M., Grubich, C., and Herrmann, M. (2006). Modulation of prefrontal cortex activation by emotional words in recognition memory. *Neuroreport* 17, 1037–1041.

Marian, V., Blumenfeld, H. K., and Kaushanskaya, M. (2007). the language experience and proficiency questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *J. Speech Lang. Hear. Res.* 50, 940–967.

Midgley, K., Holcomb, P. J., and Grainger, J. (2011). Effects of cognate status on word comprehension in second language learners: an ERP investigation. *J. Cogn. Neurosci.* 23, 1634–1647.

Midgley, K. J., Holcomb, P. J., van-Heuven, W. J. B., and Grainger, J. (2008). An electrophysiological investigation of cross-language effects of orthographic neighborhood. *Brain Res.* 1246, 123–135.

Molinaro, N., Conrad, M., Barber, H., and Carreiras, M. (2010). On the functional nature of the N400: contrasting effects related to visual word recognition and contextual semantic integration. *Cogn. Neurosci.* 1, 1–7.

Oldfield, R. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.

Palazova, M., Mantwill, K., Sommer, W., and Schacht, A. (2011). Are effects of emotion in single words non-lexical? Evidence from event-related brain potentials. *Neuropsychologia* 49, 2766–2775.

Pavlenko, A. (2008). Emotion and emotion-laden words in the bilingual lexicon. *Biling. Lang. Cogn.* 11, 147–164.

Pell, M. D., Monetta, L., Paulmann, S., and Kotz, S. A. (2009). Recognizing emotions in a foreign language. *J. Nonverbal Behav.* 33, 107–120.

Recio, G., Sommer, W., and Schacht, A. (2011). Electrophysiological correlates of perceiving and evaluating static and dynamic facial emotional expressions. *Brain Res.* 1376, 66–75.

Redondo, J., Fraga, I., Padron, I., and Comesana, M. (2007). The Spanish adaptation of ANEW (affective norms for English words). *Behav. Res. Methods* 39, 600–605.

Redondo, J., Fraga, I., Comesaña, M., and Perea, M. (2005). Estudio normativo del valor afectivo de 478 palabras españolas. *Psicológica* 26, 317–326.

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

Schacht, A., and Sommer, W. (2004). "Emotional valence accelerates lexical decisions: functional localization with ERPs," in *Evoked Potentials International Conference XIV*, eds A. Widmann, E. Schröger, T. Jacobsen, T. Gruber, M. M. Müller, J. Jescheniak, A. D. Friederici, T. C. Gunter, and C. S. Herrmann (Leipzig: Leipziger Universitätsverlag), 43.

Schacht, A., and Sommer, W. (2009a). Emotions in word and face processing: early and late cortical responses. *Brain Cogn.* 69, 538–550.

Schacht, A., and Sommer, W. (2009b). Time course and task dependence of emotion effects in word processing. *Cogn. Affect. Behav. Neurosci.* 9, 28–43.

Schachter, S., and Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychol. Rev.* 69, 379–399.

Schumann, J. H. (1998). *The Neurobiology of Affect in Language*. Oxford: Blackwell.

Schupp, H. T., Flaisch, T., Stockburger, J., and Junghöfer, M. (2006). "Emotion and attention: event-related brain potential studies," in *Understanding Emotions*, eds G. E. S. Anders, M. Junghöfer, J. Kissler, and D. Wildgruber (Amsterdam: Elsevier), 31–51.

Sebastián-Gallés, N., Martí, M. A., Carreiras, M., and Cuetos, F. (2000). *LEXESP: Una base de datos informatizada del español*. Barcelona: Servicio de Publicaciones de la Universitat de Barcelona.

Sutton, T. M., Altarriba, J., Gianico, J. L., and Basnight-Brown, D. M. (2007). The automatic access of emotion: emotional stroop effects in Spanish-English bilingual speakers. *Cogn. Emot.* 21, 1077–1090.

van der Meij, M., Cuetos, F., Carreiras, M., and Barber, H. (2011). Electrophysiological correlates of language switching in second language learners. *Psychophysiology* 48, 44–54.

van Heuven, W. J. B., and Dijkstra, T. (2010). Language comprehension in the bilingual brain: fMRI and ERP support for psycholinguistic models. *Brain Res. Rev.* 64, 104–122.

Võ, M.-L., Jacobs, A. M., and Conrad, M. (2006). Crossvalidating the Berlin affective word list (BAWL). *Behav. Res. Methods* 38, 606–609.

Võ, M. L.-H., Conrad, M., Kuchinke, K., Urton, K., Hofmann, M. J., and Jacobs, A. M. (2009). The Berlin affective word list reloaded. *Behav. Res. Methods* 41, 534–538.

Võ, M. L.-H., Jacobs, A. M., Kuchinke, L., Hofmann, M., Conrad, M., Schacht, A., and Hutzler, F. (2008). The coupling of emotion and cognition in the eye: introducing the pupil old/new effect. *Psychophysiology* 45, 130–140.