Reduced cross-modal affective priming in the L2 of late bilinguals depends on L2 exposure

Miriam S. Tenderini

Queen Mary University of London

Esther de Leeuw

Queen Mary University of London

Tiina M. Eilola

University of Leeds

Marcus T. Pearce

Queen Mary University of London

Author Note

Miriam S. Tenderini, School of Biological and Chemical Sciences, Queen Mary University of London; Esther de Leeuw, School of Languages, Linguistics and Film, Queen Mary University of London; Tiina M. Eilola, School of Psychology, University of Leeds; Marcus T. Pearce, School of Electronic Engineering and Computer Science, Queen Mary University of London.

This research was supported by grants from the Centre for Mind in Society, the School of Biological and Chemical Sciences and the School of Languages Linguistics and Film at Queen Mary University of London. Correspondence concerning this article should be addressed to Miriam S. Tenderini, Department of Experimental Psychology, Queen Mary University of London, London, E1 4NS. Contact: m.s.tenderini@zoho.com

Abstract

Processing of emotional meaning is crucial in many areas of psychology, including language and music processing. This issue takes on particular significance in bilinguals because it has been suggested that bilinguals process affective words differently in their first (L1) and second, later acquired languages (L2). We undertook a series of five experiments examining affective priming between emotionally-valenced language and emotionally-valenced music. Adult English monolinguals and two groups of proficient adult late bilinguals (German-English and Italian-English) with recent L2 exposure were examined. Priming effects were investigated using music to prime word targets and words to prime music targets. For both groups of bilinguals, music showed equivalent affective priming of L1 and L2 words, suggesting no difference in deliberate processing of affective meaning. Conversely, when words primed music, L2 words lacked the affective priming strength of L1 words for both late bilingual groups. Among various language background factors, only greater length of residence in the L2 context was positively related to the affective priming strength of L2 words. These results show strong activation of emotional meaning in the L1 of late bilinguals but reduced activation in the L2, where level of activation depends on the duration of everyday exposure to the L2.

Keywords: bilinguals, emotion, music, affective priming

Understanding the affective meaning of language is a fundamental part of effective communication across a wide range of social situations. While this is true of both native and non-native languages, most research examining emotional processing of language has focused on monolinguals (e.g., Citron, 2012; Kissler, Assadollahi, & Herbert, 2006; Knickerbocker et al., 2019) and we know much less about how people who speak more than one language process emotions differently in their languages. However, this is an important area to address given that a majority of people in the world speak more than one language (Bialystok, 2017; Grosjean, 2010) and second languages are likely to be used increasingly often in both personal and professional contexts. Therefore, the present research investigates potential differences in psychological processing of affective meaning in the first and second languages (L1 and L2) of adult late bilinguals who acquired their L2 after childhood.

Theories of conceptual and affective processing in bilingualism

Most psychological theories of bilingualism posit a language-independent conceptual (or semantic) store containing mental representations of word meaning (Dijkstra & van Heuven, 2002; Kroll, Van Hell, Tokowicz, & Green, 2010; Paradis, 1993, 2004). According to these theories, the strength with which words activate conceptual meanings depends on language proficiency (Kroll et al., 2010) or frequency and recency of language use (Dijkstra & van Heuven, 2002; Paradis, 2004). These theories make no distinction between the lexical access of affective and non-affective meaning in the conceptual store. However, *the emotional contexts of learning theory* (Harris, Gleason and Aycicegi, 2006) proposes that differences in processing of affective and non-affective meaning arise as a result of the context in which the L2 is acquired. Early language acquisition (before 6 years in Harris et al., 2006) occurs in tandem with the development of the emotion regulation system (Bloom & Beckwith, 1989) and usually in an extremely emotional context in close social interaction with the caregiver(s).

In contrast, languages acquired later in life often (although not necessarily) lack the intensity of this emotional and social setting. Harris et al. (2006) propose that during early language learning, the context of language use is stored together with the language form instead of being "stripped away during learning" (Harris et al., 2006, p. 273). A word experienced in a distinctive emotional context becomes associated with both the corresponding emotion and concurrent autobiographical events. The more a word is used in such emotional contexts, the richer its conceptual associations become, strengthening its affective meaning, which is potentially less likely to occur in an L2 that is acquired after early childhood.

Empirical research on affective processing in bilinguals

Introspective studies have found that bilinguals generally experience greater self-reported emotionality in their L1 than the L2 and that the difference reduces in size with earlier arrival in the L2 context (Dewaele, 2008; Pavlenko, 2004; Puntoni, Langhe, & Van Osselaer, 2009) which is often tightly correlated with increased L2 exposure. While this result is consistent with the emotional contexts of learning theory, it relies on conscious interpretation rather than allowing examination of the underlying psychological processes involved in lexical activation of affective meaning, which is generally thought to occur at least partially automatically (Pavlenko, 2012, 2017).

Cognitive research that directly tests the more specific hypothesis that processing of affective meaning is less well developed in the L2 of late bilinguals relative to the L1 has tended to use one of three different experimental approaches focusing on memory, interference, and affective priming respectively (Pavlenko, 2012). Studies of differences in memory recall for L1 and L2 words are based on the hypothesis that if affective meaning is activated less strongly in the L2, then affective words should be remembered less well in the L2 than the L1.

However, these studies have produced conflicting findings, sometimes showing an L1 advantage in recall of emotional words (Anooshian & Hertel, 1994), sometimes an unexpected L2 advantage (Ayçiçeği-Dinn & Caldwell-Harris, 2009; Ayçiçeği & Harris, 2004) and sometimes no difference between the L1 and L2 (Ferré, García, Fraga, Sánchez-Casas, & Molero, 2010). The variability in these results may arise because memory recall is influenced substantially by a range of factors other than the emotional meaning of the words (Pavlenko, 2012), which also emphasises the fact that memory paradigms provide a rather indirect measure of the activation of affective meaning.

Research examining interference effects has adopted variants of the emotional Stroop task in which participants typically show slower RTs when reporting the colour of emotion words compared with emotionally neutral words. This effect is assumed to reflect interference resulting from the automatic activation of the word's emotional meaning, which is hypothesised to be less strongly activated in the L2 than the L1. There is some evidence that bilinguals show lower levels of interference in their L2 for taboo words (Colbeck & Bowers, 2012) and negative words (Winskel, 2013) relative to neutral words. However, other studies have shown equal interference effects for both negative and taboo words between the L1 and L2 of bilinguals (Eilola & Havelka, 2011; Eilola, Havelka, & Sharma, 2007). In these studies, Stroop effects have been reported for negatively-valenced and taboo words but not positively-valenced words (Eilola & Havelka, 2011; Eilola et al., 2007), raising an additional question about whether the effects reflect activation of affective meaning in general. Furthermore, the interpretation of the results is complicated by a lack of theoretical clarity over the psychological processes responsible for the emotional Stroop effect (MacLeod, 2005; Yiend, 2010), with many having been proposed, including for example increased attentional allocation for negative than neutral words (Pavlenko, 2012) or perceptual prioritisation of affective words

(Pavlenko, 2017). The point is made explicitly by Macleod (2005) who concludes his review by noting that: "interpreting the results of Stroop experiments as evidence for a particular type of processing or for a particular process is suspect" (p. 34).

Research on affective priming in bilinguals

The *affective priming paradigm* (Fazio, 2001; Fazio et al., 1986; Klauer & Musch, 2003) provides a more direct way of targeting the activation of affective meaning. The logic of this experimental paradigm is to present a priming stimulus (e.g., a positive or negative word), which the participants are instructed to ignore, followed rapidly by a target stimulus (e.g., another positive or negative word) for which the participants make an evaluative decision (e.g., is the word positive or negative?). Priming is said to have taken place if the nature of the prime influences the response to the target, typically involving faster reaction times when the target is affectively congruent with the prime (i.e., both positive or both negative) than when the prime and target are affectively incongruent (i.e., one positive, the other negative). This congruence effect is critically dependent on the stimulus-onset asynchrony (SOA, the difference between the onset time of the prime and the onset time of the target), emerging only with short SOAs of no more than 300ms (Fazio et al., 1986; Hermans, De Houwer, & Eelen, 2001; Klauer, Robnagel, & Musch, 1997), suggesting that the prime activates affective meaning automatically (Pavlenko, 2017).

Degner, Doycheva and Wentura (2012) employed an affective priming paradigm to compare activation of affective meaning in the L1 and L2 of two groups of proficient, late French-German and German-French bilinguals. The groups did not differ in terms of Age of Acquisition (AoA) of the L2 nor in terms of self-rated proficiency. However, the French-German group reported significantly more intense current use of their L2 (German) than the German-French group (whose L2 was French), reflecting the fact that both groups were living in Germany at the time the experiment was conducted. The stimuli were positive and negative nouns in French and German, matched according to word length and frequency of usage. While both bilingual groups showed an affective priming effect in their L1, only the French-German bilinguals, who had recent L2 exposure, showed a comparable affective priming effect in their L2. The same participants also undertook a semantic priming task with non-affective words, in which they made a lexical decision about the target. In contrast to the affective priming results, both groups of participants showed no difference in priming between their L1 and L2. Overall, these results suggest reduced activation of affective meaning in the L2 of late bilinguals that does not reflect lack of proficiency or general semantic processing ability but rather lack of recent L2 use and immersion in the L2 context. While at face value, this conclusion runs counter to the emotional contexts of learning theory of Harris et al. (2006), it remains possible that adult immersion in the L2 context may provide sufficient emotional context, given emotional connections to the people who speak the L2.

The affective priming task used by Degner et al. (2012) relies on using two groups of bilinguals with complementary L1 and L2 to compare the priming effect of L1 and L2 words on word targets in the same language. Cross-modal experimental designs provide a simpler and potentially more robustly controlled approach with the advantages of not requiring two matched, complementary bilingual groups and assessing congruence for both L1 and L2 words using a common affective stimulus. For example, Segalowitz, Trofimovich, Gatbonton, and Sokolovskaya (2008) developed an affective variant of the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) referred to as the Implicit Affect Association Task (IAAT). They asked English-French bilinguals to judge the valence of L1 and L2 stimuli in a sequence of alternating affective pictures (happy and sad faces, intact and broken objects) and noun phrases (e.g., “the gentle child”). On congruent trials of the IAAT, the valence of the noun phrase was the same as that of the preceding picture, while for incongruent trials, the noun phrase differed in valence from the preceding picture. On neutral trials, positive and negative noun phrases were preceded by pictures with neutral valence. In both languages, RTs in the congruent condition were faster than in the neutral condition while RTs in the incongruent condition were slower compared with the neutral condition. The latter effect was stronger in the L1 than in the L2, leading the authors to conclude that affect is processed more automatically in the L1 than in the L2.

However, in contrast to the affective priming paradigm, stimulus presentation in the IAAT is self-paced with no overlap between the visual (pictures) and linguistic stimuli (L1 and L2 words), a maximum of 3000 ms to respond, and average RTs ranging from 900 to 1500 ms. Therefore, it seems likely that the IAAT relies on distinct (less rapid and automatic) psychological processes than the unimodal affective priming task of Degner et al. (2012), which depends on an SOA of 300 ms or less (see De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009, for related discussion of the IAT).

In the present series of experiments, we used a cross-modal affective priming paradigm with an SOA of 200 ms to examine affective congruency between L1 and L2 words and music as the common affective stimulus for both languages. The cross-modal experimental design depends on monolingual demonstrations of conceptual priming of language by music (Daltrozzo & Schön, 2009; Koelsch et al., 2004), which has been extended to affective priming (Goerlich et al., 2012; Goerlich, Witteman, Aleman, & Martens, 2011; Sollberger, Reber, & Eckstein, 2003; Steinbeis & Koelsch, 2008, 2011). For example, Goerlich et al. (2012, 2011) presented 600 ms segments of music expressing either happy or sad emotions along with affective Dutch nouns, both with music priming words and vice-versa, using an SOA of 200 ms. Participants judged the valence of the target stimulus as positive or negative. In both studies, there was a significant priming effect of words on music but the priming effect of music on words was only marginally significant. The relatively weak priming effect of music may reflect differing amounts of information in the positive and negative musical stimuli, given that music expressing negative affect typically has a slower tempo than music expressing positive affect. For this reason, in the present research, we employ a set of controlled musical excerpts especially created for music emotion research by Vieillard et al. (2008), taking care to match tempo between stimuli with positive and negative valence. Importantly, these musical excerpts have been shown to convey emotional meaning consistently in several participant groups with different first languages (L1) (Castro & Lima, 2014; Hill & Palmer, 2014; Hunter, Schellenberg, & Stalinski, 2011; Vieillard et al., 2008; Vieillard, Roy, & Peretz, 2012).

It is important to examine the psychological processes underlying affective priming effects. It is thought that when the priming stimulus is evaluated, it activates a semantic representation corresponding to its affective meaning (i.e., positive vs negative), which in turn activates a preparatory response pathway based on its valence (i.e., for a negative or positive response) (De Houwer et al., 2009; Fazio, 2001; Klauer & Musch, 2003). If the subsequently presented target turns out to be congruent with the prime, this preparatory response activation has a facilitatory effect on the response to the target, reducing reaction times. On the other hand, if the target is incongruent, the response activation elicited by the prime has an inhibitory effect on the response to the target, increasing reaction times. The effect of the prime is thought to be automatic, especially because of the very short SOA required (see above), but also because the effect persists when the prime is presented subliminally and in the presence of mental load introduced by a concurrent task (see De Houwer et al., 2009, for a detailed discussion).

The present research

In the traditional affective priming paradigm used by Degner et al. (2012), both prime and target are words, whereas in the cross-modal affective priming paradigm adopted in the present research, musical excerpts are used to prime L1 and L2 words, and L1 and L2 words are used to prime musical targets. In both cases because L1 and L2 words have exactly the same musical counterparts, the cross-model affective priming paradigm provides a well-controlled experimental design for a single group of bilinguals. When L1 and L2 words are used as primes, we hypothesise that L2 words will activate affective meaning less strongly than L1 words, leading to weaker preparatory response activation for the music targets. When music is used as the prime, we hypothesise equivalent response activation for both L1 and L2 targets. Furthermore, we hypothesise that the deliberate processing of L1 and L2 words as targets will not differ (in contrast to our hypothesis for their automatic processing as primes).

What factors influence the weaker activation of affective meaning by L2 words? The analysis conducted by Degner et al. (2012) suggested that the affective priming effect depended on present-day frequency of L2 use in daily life, rather than self-rated proficiency. Frequency of L2 use was assessed by a combination of self-rated reported use in professional, private-personal and private-leisure contexts, all of which were highly correlated. Participants who used their L2 less in everyday life showed weaker affective priming. This finding is potentially consistent with the emotional contexts of learning theory (Caldwell-Harris, 2015; Harris et al., 2006) if we allow for the possibility that some of that greater usage included L2 use in emotional contexts (e.g., with partners, friends, family). However, the fact that professional and personal use were highly correlated means that frequency of L2 use may not be the best way to test the theory. In the present research, in addition to language use and proficiency, we included another potential context effect, Length of Residency (LoR), which may provide a better measure since longer residency in the L2 context is potentially associated with greater use of L2 in emotionally charged contexts.

To summarise, our objectives in the present research were to examine whether early, automatic activation of affective meaning is as robust in the L2 as the L1 of late, proficient bilinguals and whether its robustness depends on factors such as language use, proficiency, and LoR. We employed a cross-modal affective priming paradigm to compare affective processing in the L1 and L2 of bilinguals, the L1 being either German or Italian acquired from birth, and the L2 being English acquired later in life but with at least one year of recent immersion living in the L2 environment (UK). In this task, visually-presented L1 and L2 words were used as both primes and targets in different conditions, so as to investigate both deliberate affective processing of word targets, as well as automatic processing of the same words as non-task-relevant primes. Participants rated the valence of the targets presented and RTs were analysed with respect to congruence between the valence of the prime and target (congruent, incongruent) and language (L1, L2).

In Experiments 1 and 2, the cross-modal affective priming paradigm was tested and baselined in two separate groups of monolingual English speakers with positively- and negatively-valenced music priming positively- and negatively-valenced target words (Experiment 1) and with words priming target music (Experiment 2). In Experiment 3, late German-English bilinguals were presented with L1 and L2 words as targets primed by music. In Experiment 4, a separate group of late German-English bilinguals was presented with music targets primed by L1 and L2 words. Experiment 5 attempted to replicate the results of Experiment 4 with another group of late bilinguals having the same L2 (English) and the L1 from a distinct language family (Italian). Across Experiments 4 and 5, we also examined whether the L2 priming effect was predicted by variables such as L2 proficiency, age of acquisition (AoA) of English, length of residence in the UK (LoR) and amount of L2 use in everyday life (L2 use).

These five experiments were designed to test the following hypotheses:

1. In English monolinguals, music will prime words (Experiment 1) and words will prime music (Experiment 2) as suggested by previous studies with different stimuli (Goerlich et al., 2012, 2011);
2. In late bilinguals, music will prime words from the L1 (German) and the L2 (English) equally because the emotional meaning of music is perceived accurately and consistently across languages and participants will have sufficient L2 proficiency to deliberately evaluate the valence of the target words (Experiment 3);
3. In late bilinguals, L1 words (German in Experiment 4 and Italian in Experiment 5) will prime music to a greater extent than L2 words (English) because L1 words show stronger automatic activation of affective meaning (Harris et al., 2006);
4. Language exposure rather than proficiency will predict the strength of the priming effect in the L2 (Degner et al., 2012).

Experiment 1: Monolingual English speakers - music priming words

Before assessing affective priming in bilinguals' L1 and L2 (the latter being English), a baseline was established with monolingual English speakers to confirm that the stimuli show an affective priming effect of music on English words. As indicated by Hypothesis (1), we predicted an affective priming effect, reflected by shorter RTs when evaluating the valence of the English word targets preceded by affectively congruent compared with incongruent musical excerpts.

Method

Participants. In all five experiments, a minimum sample size of 25 was specified *a priori* based on the existing literature (Daltrozzo & Schön, 2009; Degner et al., 2012; Fazio et al., 1986; Goerlich et al., 2012, 2011; Segalowitz et al., 2008; Sollberger et al., 2003; Steinbeis & Koelsch, 2008, 2011) but we collected data from everyone who responded to our adverts to accommodate the possibility of having to exclude participants from the subsequent analyses (e.g., due to technical difficulties, a high number of errors or language background).

Thirty-three monolingual native English speakers participated in Experiment 1 (aged 18 - 27, M_{age} = 20.26 years, *SD* = 2.05 years; 28 females). Two participants were excluded due to too many errors (in both congruent and incongruent trials, in total > 30%) as it could not be guaranteed that these participants had understood the instructions properly (e.g. evaluating the prime instead of the target). As indicated by the language background questionnaire (modified version of the Language History Questionnaire by Li, Sepanski, & Zhao, 2006, Appendix A1), the participants considered English to be their native language and learned English from birth onwards. Due to foreign language education at school, some of the monolinguals learned other languages to a certain degree, but reported their proficiency to be negligible, and that they rarely used the other languages. These participants were therefore considered to be functional monolinguals in comparison to the bilinguals tested in the subsequent experiments. Some participants had some form of musical training and most reported emotional engagement with music and good perceptual musical skills, assessed using the Goldsmiths Musical Sophistication Index (Müllensiefen, Gingras, Musil, & Stewart, 2014) and shown in Table 1. Since musical training has not been shown to influence priming effects with musical stimuli in previous research (Steinbeis & Koelsch, 2011), participants were not required to fulfil special criteria in terms of their musical background. All participants reported normal hearing and were reimbursed for their participation with course credits or cash.

Table 1
Means and standard deviations for the factors Perceptual Abilities, Musical Training and Emotion from the Goldsmiths Musical Sophistication Index (Müllensiefen et al., 2014).

Expt	**Perceptual Abilities** ***M (SD)***	**Musical Training** ***M (SD)***	**Emotion** ***M (SD)***
1	5.24 (0.73)	2.72 (1.36)	5.38 (0.88)
2	5.41 (0.77)	3.09 (1.34)	5.80 (1.47)
3	4.88 (1.04)	3.07 (1.61)	5.45 (1.03)
4	5.10 (0.93)	3.30 (1.46)	5.74 (0.65)

Note. Scores range from 1 to 7 (7 = highest possible score)

Material. Forty-eight 1 s musical stimuli were taken from a pool of unfamiliar musical excerpts composed and validated by Vieillard et al. (2008). The musical excerpts were intended to convey four emotion categories covering the four quadrants of the circumplex model (Russell, 1980) extended along the two dimensions of valence and arousal: happiness (positive valence and high arousal), peacefulness (positive valence and low arousal), threat (negative valence and high arousal) and sadness (negative valence and low arousal). Only excerpts with high arousal and thus high tempo (≥ 110 bpm) were chosen for the present research to ensure a relatively large number of musical events in the relatively short segments of 1 s. Happy excerpts were selected to represent positive valence and scary excerpts to represent negative valence. The original midi files provided by Viellard et al., (2008) were used and recorded to audio in a piano timbre.

The target word stimuli consisted of 24 English nouns with positive valence (e.g. “friend” or “vacation”) and 24 with negative valence (e.g. “war” or “failure”, for the complete list of words see Appendix B1). These emotion words were retrieved from the database ANEW (Affective Norms for English Words, Bradley & Lang, 1999). Positive and negative words were matched regarding frequency of use, word length and arousal but differed in valence, t (46) = -28.047, $p < .001$. Words were not matched for concreteness but since the same words were presented in the congruent and incongruent condition they should have impacted the different conditions similarly.

Procedure. Participants were tested individually after giving informed consent, and completing the language background questionnaire and musical background survey (in English). The experiment consisted of two blocks with 48 test-trials each. The trials started with a black fixation cross in the middle of the screen (200 ms), followed by the musical stimulus presented for 1s (or less if presentation was terminated earlier by the participant’s response). The target word appeared with a stimulus onset asynchrony (SOA) of 200 ms after the music onset (see Figure 1). The target presentation was terminated by the participants’ response. Trials were followed by inter-stimulus-intervals of 1.5 s. Music-word pairs were either congruent in their valence (both stimuli had positive or both negative valence) or incongruent (a positive stimulus paired with a negative stimulus). The target stimuli appeared as single words on the screen which was approximately 60 cm away from the participant. Words were presented with a point size of 30 in Courier New. Music stimuli were presented over headphones. Before the test trials started, participants indicated to the experimenter their preferred sound level, which then remained the same for all trials.

Music stimuli and target words were presented twice, once with a congruent and once with an incongruent partner. Prime–target pairs were created and order of presentation was randomized individually for each participant. Each stimulus occurred only once per block. In one block, participants used the A-key to rate the target words as pleasant and the L-key respectively to rate a word as unpleasant while the response pattern was reversed in the other block to control for possible RT differences between the right and left hand. The order of blocks was counterbalanced across participants. To familiarize the participant with the task and facilitate the hand-switch between the blocks, both blocks were preceded by 24 training trials (not included in the analysis).

Participants were instructed that they would hear some sounds but should focus on the words on the screen and judge the word as pleasant or unpleasant. They were not informed of any potential link between the musical excerpts and the following words. Participants were

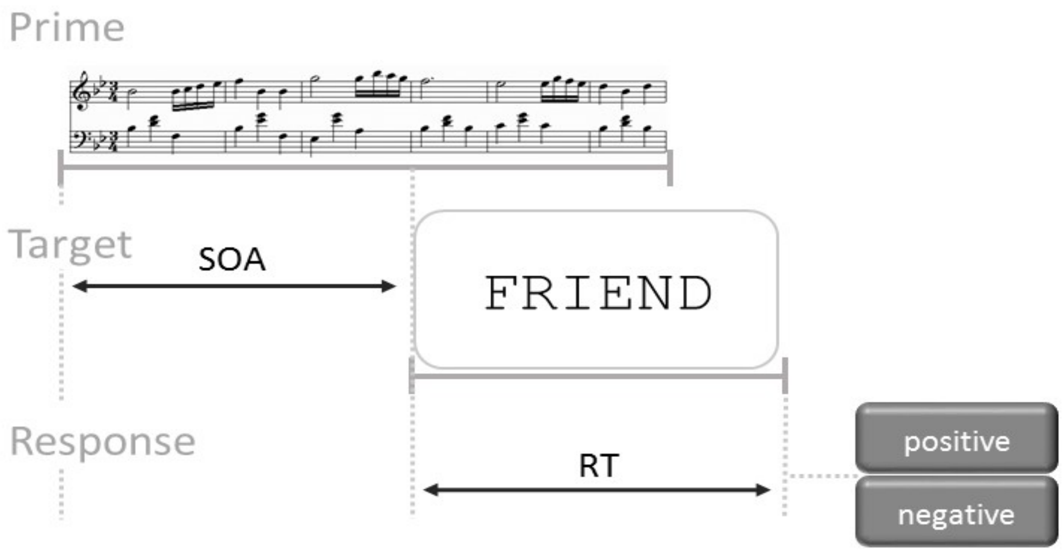

Figure 1. Affective priming paradigm with a music prime and a word target. The presentation of the target stimulus was terminated by the rating of the participant. SOA = stimulus onset asynchrony, excerpt from Vieillard et al. (2008). Note that the musical score is for illustrative purposes and does not show the actual number of musical events presented during the SOA.

asked to rate the words as accurately and quickly as possible using their index finger. RTs were recorded starting with the target-onset, i.e. with the start of the presentation of the word on the monitor. The stimulus presentation was controlled using E-Prime version 2.0 (Psychology Software Tools, Pittsburgh, PA).

Data analysis. The planned primary analysis focused on RTs, consistent with the existing literature on affective priming (De Houwer et al., 2009; Fazio, 2001; Klauer & Musch, 2003; Neely, 1977). Trials with RTs below 250 ms and more than three times the SD from the overall group mean (1768 ms) were excluded (see, e.g., Campbell & Xue, 2001) as outliers (0.94%) as well as trials with target ratings not agreeing with the predefined valence (see, e.g., Bradley & Lang, 1999) here referred to as *errors* (9.10%) presented in Table A1. Mean RTs were calculated for each participant for all four stimulus conditions: 1. positive primes with congruent positive targets; 2. positive primes with incongruent negative targets; 3. negative primes with congruent negative targets; 4. negative primes with incongruent positive targets. The RT data showed a skewed distribution and were therefore normalized by means of a log transformation for parametric analysis. For ease of interpretation, raw RTs (in ms) are reported in the text and figures. Appendix C contains an analysis of error rates for all five experiments, shown in Table A1, to exclude the possibility that a speed-accuracy trade-off influenced the primary analyses.

Results

The mean RTs for congruent and incongruent conditions are shown in Figure 2. Log-transformed RTs were submitted to a 2 x 2 repeated-measures analysis of variance (ANOVA) with the within-subject independent variables: *prime* (positive, negative) and *target* (positive, negative). There was no significant main effect of prime, $F(1,30) = 0.144$, $p = .707$, $\eta_p^2 = .005$; or target, $F(1,30) = 3.381$, $p = .076$, $\eta_p^2 = .101$ but there was a significant two-way Prime x

Target interaction, $F(1,30) = 11.577$, $p = .002$, $\eta_p^2 = .278$. A post-hoc analysis (Bonferroni corrected p-values, alpha = .05) showed target words were evaluated significantly faster if the preceding musical excerpt had affectively congruent rather than incongruent valence. This applied for negative targets, $F(1,30) = 5.307$, $p = .028$, $\eta_p^2 = .150$, and positive targets, $F(1,30) = 4.171$, $p = .050$, $\eta_p^2 = .122$.

Discussion

The purpose of this experiment was to establish a baseline for the priming effect with short musical excerpts priming English word targets in native English speakers. The results support our hypothesis that congruent music-word pairs would result in shorter RTs compared with incongruent ones. Neither the valence of the prime nor the target impacted the RT response.

This experiment is the first to show a clear priming effect of genuine musical excerpts, as opposed to isolated chords or chord sequences (Ignacio, 2016; Müller, Klein, & Jacobsen, 2011), on RTs to single written words. Previous studies presenting comparable short musical stimuli in a conceptual priming paradigm (Daltrozzo & Schön, 2009) or affective priming

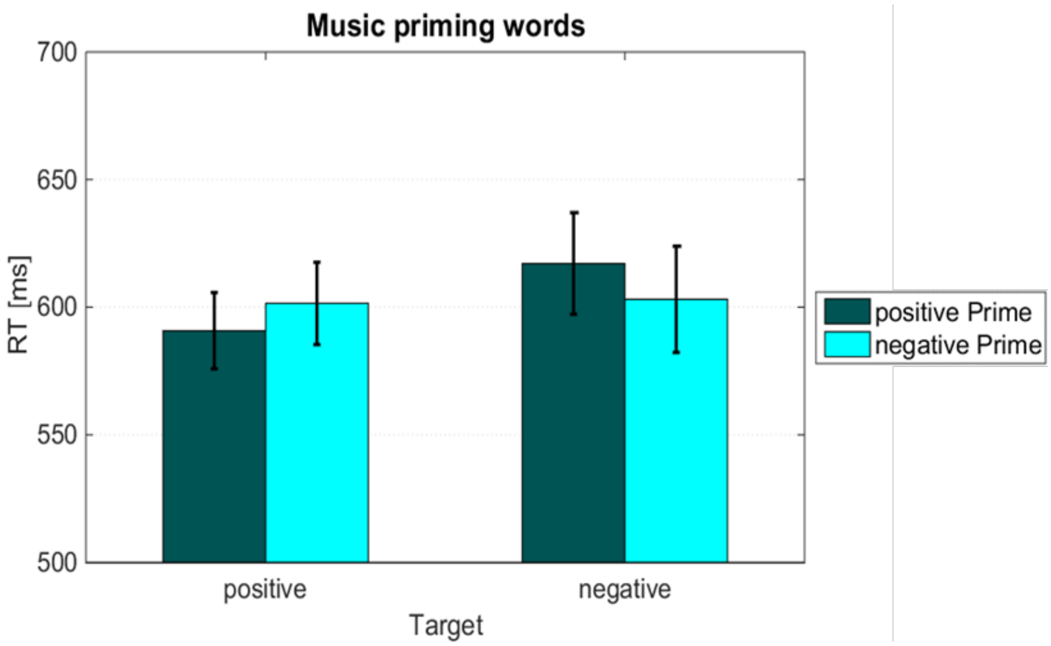

Figure 2. Mean reaction times (RT) for positive and negative word targets as a function of affectively congruent and incongruent music primes in native English speakers. Error bars represent standard errors.

paradigm (Goerlich et al., 2012, 2011) did not find significant priming effects in the RTs but only non-significant trends. The stronger priming effect in the present results may be due to the shorter SOA of 200 ms (Fazio, 2001) compared with 800 ms in Daltrozzo and Schön, and longer musical excerpts (1000 ms) than those presented in the studies of Goerlich et al. (600 ms). Furthermore, excerpts employed in the present experiment differed in valence but were better matched in terms of arousal (and consequently tempo, which determines the amount of musical information presented) than the sad and happy excerpts presented by Goerlich et al (2012, 2011).

The results of this experiment show that musical excerpts with a duration of 1000 ms are capable of activating psychological representations of affective meaning sufficiently to induce affective priming of subsequently presented affective words. In Experiment 2, we examined whether affective linguistic stimuli impact processing of the musical stimuli in a similar manner.

Experiment 2: Monolingual English speakers - words priming music

The purpose of Experiment 2 was to establish a baseline with English monolinguals for the priming effect of English words on congruent and incongruent music targets. The procedure, stimuli and participant criteria were the same as in Experiment 1 with the exception that words primed music. As indicated by Hypothesis (1), we predicted that words would prime music, reflected by shorter RTs when evaluating the valence of music targets preceded by affectively congruent compared with incongruent word primes.

Method

Participants. Thirty-seven monolingual native English speakers participated in Experiment 2 (aged 18 - 27, M_{age} = 19.54 years; *SD* = 1.69 years; 23 females). None had participated in Experiment 1. Four participants were excluded due to too many errors (in both congruent and incongruent trials, in total > 30%). As was the case in Experiment 1, the language background questionnaire (Appendix A1) indicated that all participants considered English to be their native language and had learned English from birth onwards. If other languages had been learned, for example at school, they were rarely used and proficiency was reported to be negligible. Information about the participants' music background is presented in Table 1. All other details were exactly as described for Experiment 1.

Materials and procedure. The music and word stimuli were the same as those used in Experiment 1 except that words were used as primes and the musical excerpts as targets. The forty-eight emotion words were presented visually as primes with a duration of 200 ms, each followed by one of the 48 musical excerpts with a SOA of 200 ms and a duration of 1 s. If the participant responded earlier than 1 s after target onset the presentation was terminated by the response. For this and all following experiments participants were asked to pay attention to the prime as well because this would be relevant for a subsequent task. This additional instruction was introduced to ensure participants would not close their eyes and only perceive the target music. Otherwise, the procedure was exactly the same as in Experiment 1.

Data analysis. The data were prepared for analysis in the same way as in Experiment 1. RTs below 250 ms and more than three times the SD from the overall mean (2318 ms) were excluded as outliers (1.94%) as were trials containing errors (15.86%) presented in Table A1.

Results

The mean RTs for congruent and incongruent stimulus pairs are shown in Figure 3. Log-transformed RTs were submitted to a 2 x 2 repeated-measures analysis of variance (ANOVA) with the within-subject independent variables: *prime* (positive, negative) and *target* (positive, negative). The only significant main effect was for target valence, $F(1,32) = 5.714$, $p = .023$,

$\eta_p^2 = .152$, with positive targets (*M* = 858.727 ms, *SD* = 223.538 ms) eliciting faster responses than negative targets (*M* = 879.914 ms, *SD* = 223.843 ms). Prime valence had no significant effect, $F(1,32) = 0.504$, $p = .483$, $\eta_p^2 = .016$. Furthermore, the analysis showed a significant Prime x Target interaction, $F(1,32) = 11.719$, $p = .002$, $\eta_p^2 = .268$. A post-hoc analysis (Bonferroni corrected p-values, alpha = .05) revealed significantly shorter RTs when music targets were primed by affectively congruent compared with incongruent words. This applied to music targets conveying positive, $F(1,32) = 9.771$, $p = .004$, $\eta_p^2 = .234$ and negative valence, $F(1,32) = 4.280$, $p = .047$, $\eta_p^2 = .118$.

Discussion

The purpose of this experiment was to establish a baseline for the priming effect with affective English words priming short musical excerpts in native English speakers. The results confirmed the hypothesis that congruent word-music pairs elicit faster responses than incongruent pairs.

The affective priming effect of single words on the response of music targets is consistent with the findings of Goerlich et al. (2012, 2011) and, furthermore, a faster response to positive

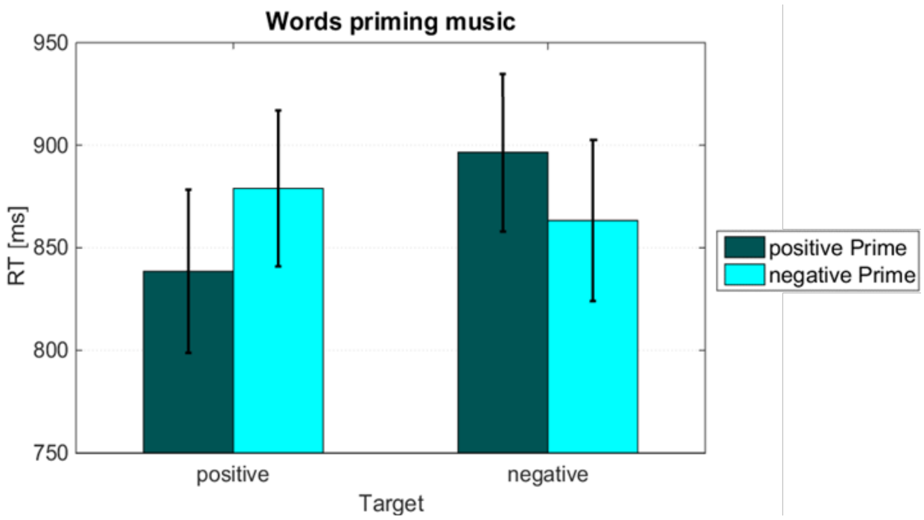

Figure 3. Mean reaction times (RT) for positive and negative music targets as a function of affectively congruent and incongruent word primes in native English monolinguals. Error bars represent standard errors.

target music (happy excerpts) has also been reported previously (Goerlich et al., 2011). The valence of happy music seems to be recognised more easily and with higher consistency than other categories of emotions (Hunter & Schellenberg, 2010; Vieillard et al., 2008).

Together with Experiment 1, these findings support the bi-directionality of the priming effect in monolingual English speakers and indicate that English words and music activate representations of emotional meaning capable of inducing cross-modal priming. In the remaining three experiments, the same design and stimuli were employed to test the processing of emotional information in bilinguals speaking English as their L2.

Experiment 3: German-English bilinguals - music priming words

This experiment was designed to test whether a group of German-English late bilinguals with English as their L2 would show a similar priming effect as monolingual English speakers using the same stimuli and paradigm as in Experiment 1. Furthermore, a new set of German words (the participants' L1) was introduced to compare the priming effects of music on L1 and L2 words. As indicated by Hypothesis (2), we predicted that music would prime words from the L1 and the L2 equally, reflected by shorter RTs in the congruent compared with the incongruent condition for both languages.

Method

Participants. Twenty-five German-English late bilinguals were newly recruited for Experiment 3 (aged 22-57, M_{age} = 36.08 years; *SD* = 10.71 years; 16 females). One participant was excluded due to too many errors (in both congruent and incongruent trials, in total > 30%) and another due to technical problems with the experimental interface. As revealed by the modified version of the Language History Questionnaire (Li et al., 2006, Appendix A2), these bilinguals started learning English after the age of eight years (AoA: *M* = 11.04, *SD* = 1.79

years). At the time of the experiment, they had all been living in the UK for at least one year (LoR: *M* = 7.30, *SD* = 6.74 years) and used English to different degrees in everyday life as shown in Table 2. Their self-assessed L2 proficiency indicated that they were highly proficient in English (Table 2). Most (*n* = 22) had learned additional languages, such as French or Spanish, to different degrees but not before the age of eight years. Some participants had some musical training, assessed using the Goldsmiths Musical Sophistication Index (Müllensiefen et al., 2014), while most reported emotionally engaging with music and having good perceptual musical skills as shown in Table 1.

Table 2
Means and standard deviations for the language background variables of the bilingual participants.

	English use (%)			English Proficiency			
Expt	work *M (SD)*	SoH *M (SD)*	SaH *M (SD)*	SC *M (SD)*	SP *M (SD)*	reading *M (SD)*	writing *M (SD)*
3	81.40 (19.52)	71.48 (23.60)	64.40 (32.67)	5.12 (.73)	5.08 (.76)	4.32 (.75)	4.92 (.99)
4	76.32 (26.77)	69.74 (22.77)	60.17 (32.72)	5.57 (.59)	5.49 (.51)	5.65 (.49)	4.39 (.66)
5	80.55 (20.49)	65.51 (19.70)	45.08 (39.60)	4.66 (.75)	4.47 (.89)	4.89 (.69)	4.50 (.67)

Note. *English use* was self-rated by the participant through a percentage of daily language use out of 100, *proficiency* was self-assessed on four subscales from 1 (*very poor*) to 7 (*native like*), *SoH* = socializing outside home, *SaH* = socializing at home, *SC* = speech comprehension, *SP* = speech production.

Materials. The music stimuli were the same as in the previous experiments. The target word stimuli consisted of the same 24 English words used in Experiments 1 and 2, as well as 24 German nouns with positive and negative valence. The German words were similar in meaning to the English words but not always the exact translation (for the complete list of words see Appendix B1). The German emotion words were taken from BAWL-R (The Berlin

Affective Word List Reloaded, Võ et al., 2009). The positive and negative word-stimuli were matched between and within languages regarding frequency of use and word length but differed in valence within the respective language: *English:* t (46) = -28.047, $p < .001$, German: t (46) = -24.24, $p < .001$. For the German words only, arousal was higher in negative (M = 3.92, SD = 0.53; 1 = *low arousing*, 5 = *high arousing*) than positive words (M = 2.84, SD = 0.66), t (46) = 6.25, $p < .001$. Since the databases of English and German word stimuli used different scales (ANEW: 9-point, BAWL-R: 5-point), the arousal ratings were not compared between languages. German words were matched in terms of imageability while these data were not available for the English words.

Procedure. Questionnaires were completed online in English prior to participation. The procedure of the experiment was generally the same as in Experiment 1, apart from adding a block with the German words. The order of the two language blocks was counterbalanced across participants. In order to prime the respective language frame before starting the L1 and L2 trials, participants were exposed to the language of the forthcoming block by listening to recorded instructions for approximately 2 minutes. The audio instructions were read by a native English and a native German speaker matched for age and sex. The instructions were followed by a brief written version of the instructions (the same as in the Experiments 1 and 2) in the respective language. Half of the participants rated the positive stimuli with the right hand and the negative stimuli with the left hand, whilst the other half used the converse hands to respond. Experiments 3 and 4 were conducted in a sound proof booth and music was presented via speakers instead of headphones. Again, the participant could indicate their preferred sound level before starting the experiment.

Data analysis. The data were prepared for analysis in the same way as in the previous experiments. RTs below 250 ms and more than three times the SD from the overall mean

(German: > 1771 ms = 1.92%, English: > 2226 ms = 1.88%) were excluded as were trials containing errors (German: 3.69%, English: 2.94%) presented in Table A1.

Results

The mean RTs for all four conditions are shown in Figure 4 separately for the L1 and L2. Log-transformed RTs were submitted to a 2 x 2 x 2 repeated measures ANOVA with the within-subject factors *language* (German, English), *prime* (positive, negative) and *target* (positive, negative). There was a significant main effect of target valence with positive target words eliciting faster responses (M = 683.507 ms, SD = 34.325 ms) than negative ones (M = 703.412 ms, SD = 34.504 ms), $F(1,22) = 11.954$, $p = .002$, $\eta_p^2 = .352$. However, the main effects of language, $F(1,22) = 0.273$, $p = .606$, $\eta_p^2 = .012$, and prime valence, $F(1,22) = 1.162$, $p = .293$, $\eta_p^2 = .050$, as well as the interactions between language and prime, $F(1,22) = 0.007$, $p = .935$, $\eta_p^2 = .000$, language and target, $F(1,22) = 2.977$, $p = .098$, $\eta_p^2 = .119$, and language, prime and target, $F(1,22) = 1.045$, $p = .322$, $\eta_p^2 = .045$, were not significant. The two-way interaction between prime and target showed a significant effect, $F(1,22) = 18.051$, $p < .001$, $\eta_p^2 = .451$. The post-hoc analysis (Bonferroni corrected p-values, alpha = .05) of the Prime x Target interaction across both languages revealed that negative targets were evaluated more quickly when primed by congruent negative music primes (M = 692.379 ms, SD = 32.564 ms)

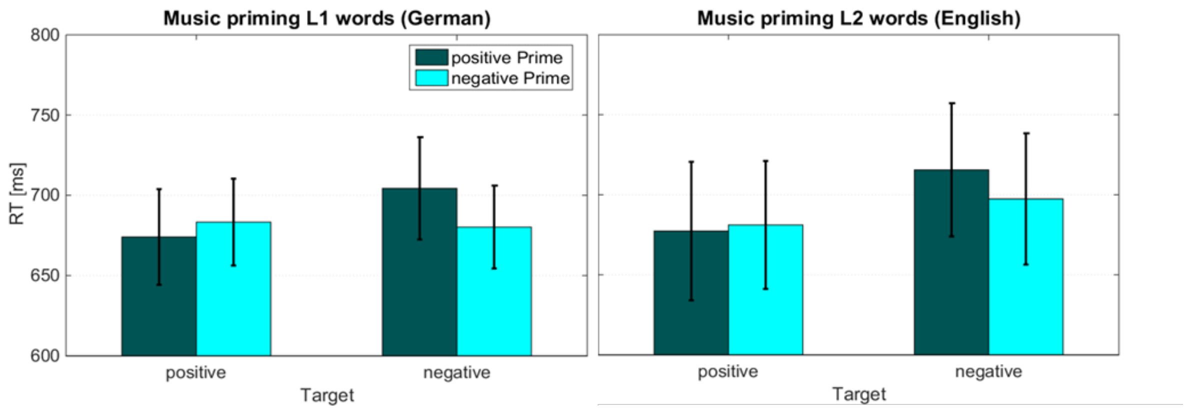

Figure 4. Mean reaction times (RT) for positive and negative word targets in the L1 (left panel) and the L2 (right panel) as a function of affectively congruent and incongruent music primes in German-English bilinguals. Error bars represent standard errors.

compared with incongruent positive music primes (*M* = 714.445 ms, *SD* = 36.674 ms), $F(1,22) = 9.601$, $p = .005$, $\eta_p^2 = .304$. Similarly, positive targets were evaluated more quickly when primed by congruent positive music (*M* = 679.729 ms, *SD* = 36.032 ms) compared with negative music (*M* = 714.445 ms, *SD* = 36.674 ms), $F(1,22) = 5.248$, $p < .032$, $\eta_p^2 = .193$.

Discussion

The purpose of this experiment was to test for a difference in the priming effect between bilinguals' L1 and L2 when primed by the same musical excerpts. The results corroborate the hypothesis of a comparable priming effect of music on word targets in both languages. Both positive and negative targets were evaluated more rapidly when preceded by congruent than incongruent music. Generally, positive target words elicited faster responses, consistent with literature showing more rapid processing of positive than negative words (e.g. Kuperman, Estes, Brysbaert, & Warriner, 2014).

The affective priming effect of music on the evaluation of single target words from Experiment 1 was extended in the present experiment to another language (German) and most importantly to bilinguals' L2. To the best of our knowledge this is the first experiment showing similar priming effects of music on emotion words presented in the L1 and L2 of bilinguals. In other words, there did not appear to be any L1 advantage in the affective priming of music on emotion words.

Experiment 4 examined whether the same L1 and L2 words (presented as targets in the present experiment) differ in automatic activation of affective meaning when presented as primes for music targets.

Experiment 4: German-English bilinguals - words priming music

The purpose of Experiment 4 was to examine differences in affective priming between bilinguals' L1 and L2 (as in Experiment 3) when late bilinguals responded to music targets primed by congruent and incongruent words (as in Experiment 2). Words from the L1 (German) and L2 (English) were presented as primes while participants evaluated the valence of the music targets. As indicated by Hypothesis (3), we predicted that L1 words would prime music to a greater extent than L2 words, reflected by shorter RTs in the congruent compared with the incongruent condition in the L1 and a weaker RT difference in the L2.

Method

Participants. Twenty-five German-English late bilinguals participated (aged 20 - 47, M_{age} = 31.61 years, *SD* = 7.52 years, 16 females) none of whom participated in Experiment 3. Two participants were excluded because of technical problems with the experimental interface. As revealed by a modified version of the Language History Questionnaire (Li et al., 2006, Appendix A2), participants started learning English after the age of eight years (AoA, *M* = 9.91, *SD* = 1.68 years). At the time of the experiment, they had all been living in the UK for at least one year (LoR: *M* = 6.43, *SD* = 4.87 years) and used English to different degrees in everyday life as shown in Table 2. The self-assessed language proficiency indicated that these late bilinguals were highly proficient in English. Most (*n* = 21) had learned a third language beside German and English to different degrees but not before the age of eight years.

Some participants had some musical training, assessed using the Goldsmiths Musical Sophistication Index (Müllensiefen et al., 2014), while most reported emotionally engaging with music and having good perceptual musical skills as shown in Table 1. All other details were exactly as described for Experiment 2.

Materials and procedure. Stimuli and procedure were the same as in Experiment 3 apart from participants evaluating the music stimuli primed by words (as in Experiment 2).

Data analysis. The data preparation for analysis was the same as in the previous experiments. RTs below 250 ms and more than three times the SD from the overall mean (German: > 2201 ms = 1.52%, English: > 2091 ms = 1.52%) were excluded as were trials containing errors, (German: 7.47%, English: 8.29%) presented in Table A1.

Results

Log-transformed RTs were submitted to a 2 x 2 x 2 repeated measures ANOVA with the within-subject factors *language* (German, English), *prime* (positive, negative) and *target* (positive, negative). The results showed a significant main effect of target valence but here positive targets elicited slower responses (M = 878.049 ms, SD = 47.914 ms) than negative ones (M = 819.881 ms, SD = 43.050 ms), $F(1,22) = 24.697$, $p < .001$, $\eta_p^2 = .529$. Neither the main effect of language, $F(1,22) = 0.536$, $p = .472$, $\eta_p^2 = .024$, nor prime valence, $F(1,22) =$ 658, $p = .658$, $\eta_p^2 = .009$, was significant. Furthermore, there was no significant interaction between language and prime, $F(1,22)=.018$, $p =.896$, $\eta_p^2=.001$, nor language and target $F(1,22)=.890$, $p =.356$, $\eta_p^2=.039$. The two-way interaction between prime and target was significant, $F(1,22) = 29.271$, $p < .001$, $\eta_p^2 = .571$, and this was qualified by a significant three-way Language x Prime x Target interaction, $F(1,22) = 5.214$, $p = .032$, $\eta_p^2 = .192$. Two additional 2 x 2 repeated measures ANOVAs on the RTs were conducted for the L1 (German) and L2 (English) words separately with the within-subject factors *prime* (positive, negative) and *target* (positive, negative).

German. The mean RTs for the prime target interaction are shown in Figure 5. The repeated measures ANOVA for German words showed a significant Prime x Target interaction, $F(1,22) = 33.900$, $p < .001$, $\eta_p^2 = .606$. The post-hoc analysis (Bonferroni corrected p-values, alpha = .05) revealed that positive music was responded to more quickly when primed by positive words than negative words, $F(1,22) = 30.634$, $p < .001$, $\eta_p^2 = .582$. Similarly, negative music was evaluated faster when primed by negative than positive words, $F(1,22) = 9.453$, $p = .006$, $\eta_p^2 = .301$.

English. The mean RTs for the prime target interaction are shown in Figure 5. The repeated measures ANOVA for English words showed a significant Prime x Target interaction, $F(1,22) = 5.819$, $p = .025$, $\eta_p^2 = .209$. The post-hoc analysis (Bonferroni corrected p-values, alpha = .05) revealed no significant differences between congruent and incongruent conditions, either for positive targets, $F(1,22) = 3.289$, $p = .083$, $\eta_p^2 = .130$, or negative targets, $F(1,22) = 3.009$, $p = .097$, $\eta_p^2 = .120$.

Discussion

The purpose of this experiment was to compare affective priming by L1 and L2 words in late, proficient bilinguals. As hypothesized, the results showed a difference in the priming effect between bilinguals' L1 and L2. While in the L1 there was a clear priming effect indicated by shorter RTs in response to congruent compared with incongruent stimulus pairs, the prime target interaction in the L2 differed significantly. Although there was a trend for congruent stimulus pairs to elicit faster responses than incongruent ones in the L2, the priming effect was

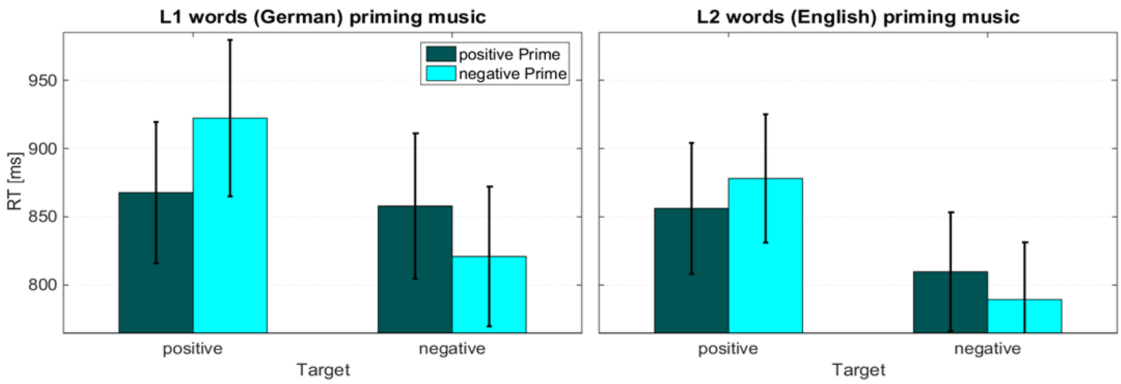

Figure 5. Mean reaction times (RT) for positive and negative music targets as a function of affectively congruent and incongruent word primes presented in the L1 (left panel) and the L2 (right panel) in German-English bilinguals. Error bars represent standard errors.

not statistically significant. To the best of our knowledge, this is the first experiment comparing the affective priming strength of L1 and L2 words on music targets. Crucially the results show that the same L1 and L2 word stimuli which were primed equally by music when presented as deliberately evaluated targets (Experiment 3) elicited different priming effects when presented as primes themselves. L1 words showed a clear priming effect while L2 words did not, suggesting reduced automatic activation of the affective meaning of L2 words than for L1 words.

Overall, responses were faster for negative than positive music targets. This differs from Experiment 2 where English monolinguals made faster responses to positive than negative music targets. A possible explanation may be the generally higher arousal in the negative compared with the positive words in the German stimuli. They may match better the high arousal in the music eliciting faster responses in line with research showing that music excerpts with fast tempo induce a greater selection of high arousal words (Tay & Ng, 2019). However, this cannot explain why the pattern also occurred in the L2 stimuli that were better matched in arousal and did not produce this pattern of results in the monolingual group.

The next experiment was conducted to test whether the difference in affective priming strength between L1 and L2 words generalizes to late bilinguals with a different L1 tested with exactly the same experimental design.

Experiment 5: Italian-English bilinguals - words priming music

The purpose of this experiment was to replicate the results of Experiment 4 (words affectively priming music in the L1 but not the L2 of German-English bilinguals) for bilinguals with the same L2 but a different L1. Therefore, a group of Italian-English late bilinguals was tested with the same cross-modal affective priming paradigm, with German stimuli replaced by Italian words. We hypothesized that the results of Experiment 4 would generalize to this group of late bilinguals with a similar L2 acquisition profile. Consequently, as indicated by Hypothesis (3), we predicted that L1 words would prime music to a greater extent than L2 words, reflected by shorter RTs in the congruent compared with the incongruent condition in the L1 and a weaker RT difference in the L2.

Method

Participants. Thirty-nine late Italian-English bilinguals (aged 20 - 39, M_{age} = 28.39 years; *SD* = 6.02 years; 25 females) participated. One participant was excluded due to too many errors (in both congruent and incongruent trials, in total > 30%) and another due to having had exposure to English in early childhood. As revealed by a modified version of the Language History Questionnaire (Li et al., 2006, Appendix A2), this group started learning English after the age of eight years (AoA: *M* = 10.11, *SD* = 5.92 years). At the time of the experiment, they had all been living in the UK for at least one year (LoR: *M* = 3.98 SD = 0.70 years) and used English to different degrees in everyday life as shown in Table 2. The self-assessed language proficiency indicated that these late bilinguals had high self-assessed proficiency in English as shown in Table 2. Most (*n* = 33) had learned languages other than Italian and English to different degrees but not before the age of eight years. Although the groups were recruited with the same criteria, the German-English bilinguals in Experiment 4 had lived longer in the UK (*M* = 6.32, *SD* = 0.89) than the Italian-English bilinguals (*M* = 3.98, *SD* = 0.70, $F(1,58) = 4.286$, $p = 0.043$, η_p^2 = 069). The groups did not differ significantly in regards of AoA, L2-use or self-rated proficiency ($p > .05$).

Materials and procedure. Stimuli and procedure were the same as in Experiment 4 except for replacing the German audio instructions with Italian ones (read by an Italian speaker matched in terms of age and sex with the reader of the English instructions) and by replacing the German words with Italian words (from the adaptation of ANEW for Italian, Montefinese, Ambrosini, Fairfield, & Mammarella, 2014). Due to orthographic similarities between some English words and their Italian translation - as well as the challenge of matching words as closely as possible for arousal, word length and frequency - some of the English words used in the previous experiments were changed (for the complete list of words see Appendix B2). However, the English word stimuli used in the present experiment did not differ in terms of valence, arousal, word length or frequency from those presented in the previous experiments ($p > .54$). The positive and negative word-stimuli were matched between and within languages regarding frequency of use and word length but differed in terms of valence within the respective language: English: t (46) = -29.122, $p < .001$, Italian: t (46) = -30.745, $p < .001$. While positive and negative English words were matched in terms of arousal, for the Italian stimuli negative words (M = 6.56, SD = 0.57; 1= *low-arousing*, 9 = *high-arousing*) were higher in arousal than positive words, t (46) = 4.099, $p < .001$. Italian words were matched in terms of imageability while these data were not available for the English words. Musical stimuli were presented via headphones as in Experiments 1 and 2 (rather than using loudspeakers as in Experiments 3 and 4). As before, the participants could indicate their preferred sound level before starting the experiment.

Data analysis. The data were prepared for analysis as in the previous experiments. RTs below 250 ms and more than three times the SD from the overall mean (Italian: > 2415 ms = 1.50%, English: > 2119 ms = 1.73%) were excluded as were trials containing errors (Italian: 10.42%, English: 10.44%) presented in Table A1.

Results

Log-transformed RTs were submitted to a 2 x 2 x 2 repeated measures ANOVA with the within-subject factors *language* (Italian, English), *prime* (positive, negative) and *target* (positive, negative). There was a significant main effect of target with positive music targets eliciting slower responses (*M* = 860.437 ms, *SD* = 213.380 ms) than negative targets (*M* = 839.240 ms, *SD* = 234.090 ms), $F(1,36) = 5.792$, $p = .021$, $\eta_p^2 = .135$. No other main effect was significant, language: $F(1,36) = 0.636$, $p = .430$, $\eta_p^2 = .017$, prime: $F(1,36) = 0.259$, $p = .614$, $\eta_p^2 = .007$. There was no significant interaction between language and prime valence, $F(1,36) = 0.004$, $p = .953$, $\eta_p^2 = .000$, nor between language and target valence, $F(1,36) = 0.407$, $p = .528$, $\eta_p^2 = .011$. However, the two-way interaction between prime and target valence was significant, $F(1,36) = 9.128$, $p = .005$, $\eta_p^2 = .202$. This was qualified by a significant three-way Language x Prime x Target interaction, $F(1,36) = 4.212$, $p = .047$, $\eta_p^2 = .105$. Two additional two-by-two repeated measures ANOVAs on the RTs were conducted for the L1 (Italian) and L2 (English) words separately with the within-subject factors *prime* (positive and negative) and *target* (positive and negative).

Italian. The mean RTs for the prime target interaction are shown in Figure 6. The repeated measures ANOVA for Italian words with the factors prime and target showed a significant Prime x Target interaction, $F(1,36) = 10.517$, $p = .003$, $\eta_p^2 = .226$. The post-hoc analysis (Bonferroni corrected p-values, alpha = .05) indicated that positive music targets elicited faster responses when primed by affectively congruent positive words than incongruent negative words, $F(1,36) = 4.716$, $p = .037$, $\eta_p^2 = .116$. Similarly, negative music targets elicited faster responses when primed by congruent negative words than incongruent positive words, $F(1,36) = 7.179$, $p = .011$, $\eta_p^2 = .166$.

English. The mean RTs for the prime target interaction are shown in Figure 6. The repeated measures ANOVA for English words with the factors prime and target showed, contrary to the analysis of the Italian words, no prime x target interaction, $F(1,36) = 1.692$, $p = .202$, $\eta_p^2 = .045$, but a main effect of target, $F(1,36) = 6.433$, $p = .016$, $\eta_p^2 = .152$, with faster responses to negative (M = 868.037 ms, SD = 277.653 ms) than positive targets (M = 891.202 ms, SD = 247.419 ms). There was no main effect of prime, $F(1,36) = 0.140$, $p = .711$, $\eta_p^2 = .004$.

Discussion

The results for the Italian-English bilinguals closely resemble those of the German-English bilinguals in Experiment 4. As hypothesised, we found a difference in affective priming between the bilinguals' L1 and L2. Participants made faster responses to positive and negative music targets primed by congruent than incongruent primes from the L1 (Italian) but there was no such effect for the L2 (English). As in Experiment 4, participants responded generally more rapidly to negative target music, especially when it was primed by L2 words.

The results of Experiments 4 and 5, obtained with two different groups of bilinguals, strongly indicate a difference in the rapid, automatic activation of affective meaning between the L1 and L2 of bilingual individuals. In the next section, we analyse potential influences of the participants' language background, particularly whether L2 exposure influenced the priming strength of L2 words.

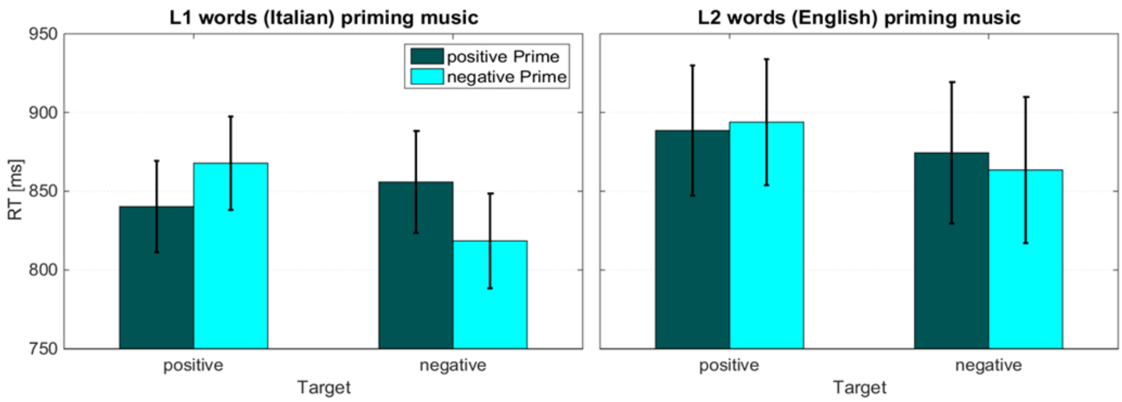

Figure 6. Mean reaction times (RT) for positive and negative music targets as a function of affectively congruent and incongruent word primes presented in the L1 (left panel) and the L2 (right panel) in Italian-English bilinguals. Error bars represent standard errors.

The effect of L2 exposure

Effects of the participants' language background were investigated by examining potential influences of L2 exposure and other L2 related factors on the affective priming strength of L2 words from Experiments 4 and 5. As indicated by Hypothesis (4), we predicted that L2 exposure – quantified as the two variables LoR in years in the UK and frequency of daily L2 use – would influence the priming strength (the difference in RT between the congruent and incongruent condition) of L2 words. In other words, the longer the late bilinguals had resided in the UK, and the more often they used English in their daily lives, the stronger the predicted priming strength of L2 words presented as primes.

Table 3 shows pairwise correlations between L2 priming strength (PS L2) derived from the pooled data from Experiments 4 and 5 (which used the same task and did not differ in priming strength, p = .245, N = 60) and two measures of L2 exposure (LoR , L2 use) as well as L2 proficiency and AoA. Since proficiency in speech comprehension, speech production,

Table 3: Spearman correlation coefficients between bilinguals' L2 (English) priming strength (Experiments 4 and 5) and L2 history factors (N = 60).

Variable	2	3	4	5
1. PS L2	.33**	.19	.09	.14
2. LoR	--	.31*	.14	.21
3. L2 use		--	.06	.07
4. AoA			--	-.06
5. Proficiency				--

Note. Significant correlations with $p < .05$ * and $p <.01$ **, PS L2 = L2 Priming Strength, LoR = Length of residency in the UK, L2 use = frequency of L2 use, AoA = Age when L2 acquisition started, Proficiency = overall L2 proficiency (including comprehension, speaking, writing and reading).

writing and reading were highly correlated (Cronbach's alpha = .86) proficiency was included as a single averaged variable. For the same reason, L2 use in different contexts (at work, for socializing inside and outside home; Cronbach's alpha = .74) was also included as a single averaged predictor. As shown in Table 3, the L2 priming strength was positively correlated with LoR in the UK.

Since LoR and L2 use were also positively correlated with one another, semi-partial Spearman correlations were run to assess the unique proportions of variance accounted for by these two variables. The results showed a significant relationship between LoR and L2 priming strength, while holding L2 use constant, $r = .3$, $p = .02$ but not between L2 use and L2 priming strength, while holding LoR constant, $r = .09$, $p = .48$. The proportion of variance in L2 priming strength accounted for by both LoR and L2 use (i.e., the variance in common between these variables) was 2.6%, which is larger than the 0.8% accounted for uniquely by L2 use but rather smaller than the 9% accounted for uniquely by LoR. These findings suggest that LoR was the strongest predictor of L2 priming strength when L2 words primed music.

In summary, these results demonstrate an influence of L2 exposure, but not proficiency or AoA, on L2 priming strength with LoR rather than L2 use being the critical measure of L2 exposure that influenced affective priming in the L2.

General Discussion

A number of interesting findings arose from this series of experiments. First, we observed, as hypothesized, consistent priming effects in all five experiments when L1 words primed music and when music primed L1 words. Second, we found, as hypothesised, that music primed language with equal strength in the L1 and L2 of late bilinguals (Experiment 3). Third, and importantly, the results showed consistently weaker affective priming of music by the L2 than the L1 in two groups of late bilinguals with different L1 (Experiments 4 and 5). Finally, we found that the strength of the L2 priming effect was related to measures of L2 exposure, in particular LoR in the L2 context (Experiments 4 and 5).

A novel aspect of this research was the use of a cross-modal affective priming paradigm using music and word stimuli to compare deliberate and automatic affect processing in bilinguals' L1 and L2. Music primed both L1 and L2 target words supporting our hypothesis that music would consistently prime the deliberate affective evaluation of L1 and L2 words. Crucially, however, when the same words were presented as primes, processed automatically, both groups of bilinguals (late German-English and late Italian-English) showed a clear priming effect in their L1 but not in their L2, supporting the hypothesis that the L1 and L2 differ in terms of rapid, automatic activation of affective meaning. The finding of a priming effect in English monolinguals shows that the lack of such an effect in English L2 speakers cannot be explained by the stimulus characteristics *per se* but rather indicates a processing difference between the bilinguals' L1 and L2.

These findings have implications for psychological models of the storage and retrieval of lexical semantics in bilinguals. The influence of language exposure on strength of affective priming in the L2 lends support to the *revised Bilingual Interactive Activation model* of Dijkstra and van Heuven (2002) and the Activation Threshold Hypothesis of Paradis (1993, 2004), which stipulate that increased activation of the L2 (i.e., greater exposure) will facilitate the accessibility of the L2 word meaning. It is not consistent with the *revised hierarchical model* (Kroll & Stewart, 1994; Kroll, Van Hell, Tokowicz, & Green, 2010), which posits increased proficiency as the driving force for the acquisition of L2 word meaning, strengthening the link between the L2 lexicon and the conceptual store (containing all representations of meaning in a language independent manner) and thus facilitating access to word meaning. The present results, like those of Degner et al. (2012), suggest that proficiency alone seems not to account for affective processing of L2 words; affective language processing appears to depend rather on increased exposure to the L2.

While these models do not explicitly distinguish processing of affective meaning from general processing of lexical semantics, the emotional contexts of learning theory (Harris et al., 2006) posits that the emotional understanding of language develops through use of language in emotional contexts in early childhood. Our results are partially consistent with this prediction, showing reduced affective priming in the L2 of individuals who acquired their L2 in adulthood. However, the influence of L2 exposure suggests that sensitivity to emotional meaning can also develop in adulthood and is not restricted to childhood. With increasing L2 exposure in a range of varied life situations, emotional meaning in language may become more strongly associated with extralinguistic emotional experience. Further research is required to examine exactly the conditions under which such sensitivity develops during L2 exposure and whether it is dependent on usage in emotional contexts, as predicted by Harris et al. (2006) for children, or merely greater L2 exposure and use (Dijkstra & van Heuven, 2002; Paradis, 1993, 2004).

The results are broadly consistent with those of Degner et al. (2012), who examined semantic and affective priming in French-German and German-French bilinguals living in Germany. They found differences in L2 affective priming for the German-French bilinguals but not for the French-German bilinguals, who were living in the L2 context. The German-French group had spent less time living abroad in their L2 context (French) and presumably also less recently since they were living in Germany at the time of the experiment, while they also used their L2 more and their L1 less than the French-German bilinguals. Our finding of a difference in affective priming between the L1 and L2 of late bilinguals is consistent with the results for the German-French group. It is reasonable to ask why our results differ from those for the French-German group, which is the most comparable to our German-English and Italian-English bilinguals, who were also currently living in the L2 context. It may be that the nature of L2 exposure differed between our late bilinguals and Degner et al.'s French-German group, in ways that are not captured by LoR or intensity of use. Another possibility is that differences arose from Degner et al.'s use of words to prime target words compared with our use of words to prime music targets. Although Degner et al. took pains to match the L1 and L2 words in terms of word length, valence and frequency of use, psychologically meaningful differences between the L1 and L2 target words cannot be completely ruled out whereas our use of the same musical targets for both L1 and L2 primes excludes any such differences, potentially making for a more precise and sensitive experimental comparison.

While semantic information is decoded very early in the process of L2 acquisition (already after a single word learning session, Altarriba & Mathis, 1997), there is evidence that semantic and affective meaning have distinct psychological representations during L2 learning (Sianipar, Middelburg, & Dijkstra, 2015) which persist even for highly proficient L2 speakers (Degner et al., 2012). The present results suggest that affective word processing is dependent on longer periods of exposure to the L2 context, ranging over years. Our finding of a modulatory effect of L2 exposure on affect processing in the L2 appeared to reflect primarily LoR in the L2 context rather than current frequency of L2 use. Of course, LoR cannot exert a direct influence since an individual could live in a foreign country without speaking the language. However, in our experiments, LoR was correlated with frequency of L2 use, suggesting that participants with longer residency in the L2 context used the L2 more frequently. What is striking is that it was LoR rather than L2 use per se that predicted L2 affective priming strength, suggesting an influence of some aspect of the L2 context *other than* increased use of the L2. This appears to extend the results of Degner et al. (2012), whose affective priming effect was influenced primarily by frequency of L2 use. Since L2 use was correlated with LoR in the present analysis and Degner et al. did not analyse LoR, the findings may be consistent if longer term L2 exposure was also an underlying influence in Degner et al.'s analysis.

Nevertheless, the present results and those of Degner et al., (2012) are both consistent with a prediction made by the emotional contexts of learning theory (Harris et al., 2006) that emotional understanding of language depends on experiencing language in emotional contexts. However, as noted above, both the present results and those of Degner et al., suggest that effects of exposure on emotional understanding of language extend beyond childhood into adulthood, with adults showing generally reduced affective priming in their second language but with affective priming becoming stronger with greater exposure. Regardless of age, future research should go beyond frequency of use and length of residency to examine whether affective processing in the L2 depends more specifically on language experience in emotional contexts, and the detailed type and quality of those experiences (Pavlenko, 2017). For example, bilinguals who have lived longer in an L2 environment may have a greater likelihood of establishing relationships with speakers of the L2 and thereby experiencing emotional situations in their L2. Conversely, it also remains to be confirmed whether lack of L1 use in emotional contexts can result in L1 attrition of affective processing (see Sutton et al., 2007) as observed in other language domains such as phonetic and phonological attrition of the L1 in bilinguals living in the L2 context (de Leeuw, 2019; de Leeuw, Schmid, & Mennen, 2010).

The present results are also significant in providing clear evidence for cross-modal affective priming between music and language. While conceptual priming of non-affective meaning between music and language is well established (Daltrozzo & Schön, 2009; Koelsch et al., 2004), support for affective priming was previously less clear, specifically in the case of music priming affectively-valenced words, where only marginally significant effects had been reported (Goerlich et al., 2012, 2011). Here we used longer musical excerpts (1000 ms instead of 600 ms) and took care to ensure that positive and negative excerpts were matched for tempo and arousal, resulting in clear and consistent cross-modal affective priming for music priming words in monolinguals as well as bilinguals' L1 and L2. The consistency of these priming effects provides strong evidence that music can automatically evoke representations of emotional meaning capable of priming responses to affective words. However, one potential limitation of our research is worth noting. The lack of any difference in affective priming of music on L1 and L2 words in Experiment 3 may reflect Type II error resulting from lack of power, estimated at 70% ($\alpha = .05$, N = 23, η^2 = 0.223, from Degner et al., 2012). Therefore, future research could attempt to replicate the results of Experiment 3 with larger sample sizes.

Several other directions remain open for future research. For example, we do not know whether similar cross-modal results would arise with other modalities, such as visual images varying in valence or aurally presented words. Furthermore, as already mentioned, although we did not find an effect of AoA nor of self-reported English proficiency on L2 affective priming (consistent with Degner et al., 2012), it would be useful to replicate these results with more detailed, objective instruments for measuring proficiency (e.g., Ross, 1998). Furthermore, given that Ponari et al., (2015) did not find an effect of L2 exposure on processing of emotion words using a lexical decision task, a comprehensive comparison of the lexical decision and affective priming tasks is warranted. It would also be interesting to observe whether similar findings are obtained with simultaneous bilinguals and whether, for this group, language exposure would also be the strongest predictor variable. Finally, applying conceptual priming paradigms with non-affective interaction of musical and linguistic meaning (e.g., Daltrozzo & Schön, 2009; Koelsch et al., 2004) in a bilingual context would shed further light on semantic processing more generally in the L2 as compared with the L1.

In summary, proficient late bilinguals demonstrated reduced automatic activation of emotional meaning in their L2 compared with the L1 when words were presented as non-task relevant, automatically processed primes. Furthermore, the strength of this effect was influenced by length of residency in the L2 environment. These findings could have important practical implications. Reduced emotional understanding of the L2 by individuals with little everyday L2 usage could lead to misunderstandings in cross-cultural interaction and impede acculturation in the L2 context, thereby disadvantaging non-native speakers. Such consequences could be far-reaching when one considers the increasing mobility of the global population and increasing use of non-primary languages in both professional and personal contexts.

References

Altarriba, J., & Mathis, K. M. (1997). Conceptual and lexical development in second language acquisition. *Journal of Memory and Language*, *36*, 550–568.

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

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

Ayçiçeği, A., & Harris, C. L. (2004). Bilinguals' recall and recognition of emotion words. *Cognition and Emotion*, *18*, 977–987.

Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin*, *143*, 233–262.

Bloom, L., & Beckwith, R. (1989). Talking with feeling: Integrating affective and linguistic expression in early language development. *Cognition and Emotion*, *3*, 313–342.

Bradley, M. M., & Lang, P. J. (1999). Affective norms for English words (ANEW): Instruction manual and affective ratings. *Technical Report C-1, The Center for Research in Psychophysiology, University of Florida.*

Caldwell-Harris, C. L. (2015). Emotionality differences between a native and foreign language: Implications for everyday life. *Current Directions in Psychological Science*, *24*, 214–219.

Campbell, J. I. D., & Xue, Q. (2001). Cognitive arithmetic across cultures. *Journal of Experimental Psychology: General*, *130*, 299–315.

Castro, S. L., & Lima, C. F. (2014). Age and musical expertise influence emotion recognition in music. *Music Perception*, *32*, 125–142.

Citron, F. M. M. (2012). Neural correlates of written emotion word processing: A review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and Language*, *122*, 211–226.

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

Daltrozzo, J., & Schön, D. (2009). Conceptual processing in music as revealed by N400 effects on words and musical targets. *Journal of Cognitive Neuroscience*, *21*, 1882–1892.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin*, *135*, 347–368.

de Leeuw, E. (2019). Phonetic attrition. In B. Schmid, B. Köpke, M. Cherciov, M. Keijzer, E. de Leeuw, T. Karayayla, & T. Mehotcheva (Eds.), *The Oxford Handbook of Language Attrition* (pp. 204–217).

de Leeuw, E., Schmid, M. S., & Mennen, I. (2010). The effects of contact on native language pronunciation in an L2 migrant setting. *Bilingualism: Language and Cognition*, *13*, 33–40.

Degner, J., Doycheva, C., & Wentura, D. (2012). It matters how much you talk: On the automaticity of affective connotations of first and second language words. *Bilingualism: Language and Cognition*, *15*, 181–189.

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

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*, 175–197.

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. *The International Journal of Bilingualism*, *15*, 353–369.

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

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion*, *15*, 115–141.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229–238.

Ferré, P., García, T., Fraga, I., Sánchez-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.

Goerlich, K. S., Witteman, J., Aleman, A., & Martens, S. (2011). Hearing feelings: Affective categorization of music and speech in alexithymia, an ERP study. *PLoS One*, *6*, e19501.

Goerlich, K. S., Witteman, J., Schiller, N. O., Heuven, V. J. Van, Aleman, A., & Martens, S. (2012). The nature of affective priming in music and speech. *Journal of Cognitive Neuroscience*, *24*, 1725–1741.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.

Grosjean, F. (2010). *Bilingual: Life and reality.* Cambridge, MA: Harvard University Press.

Harris, C. L., Gleason, J. B., & Ayçiçeği, A. (2006). When is a first language more emotional? Psychophysiological evidence from bilingual speakers. In A. Pavlenko (Ed.), *Bilingual minds: Emotional experience, expression and representation* (pp. 257–283). Clevedon, England: Multilingual Matters.

Hermans, D., De Houwer, J., & Eelen, P. (2001). A time course analysis of the affective priming effect. *Cognition and Emotion*, *15*, 143–165.

Hill, W. T., & Palmer, J. A. (2014). Affective responses to music without recognition: Beyond the cognitivist hypothesis. *The New School Psychology Bulletin*, *11*, 42–49.

Hunter, P. G., & Schellenberg, E. G. (2010). Music and emotion. In M. R. Jones, R. R. Fay, & A. N. Popper (Eds.), *Music Perception* (pp. 129–164). New York: Springer.

Hunter, P. G., Schellenberg, E. G., & Stalinski, S. M. (2011). Liking and identifying emotionally expressive music: Age and gender differences. *Journal of Experimental Child Psychology*, *110*, 80–93.

Ignacio, D. A. (2016). The cognitive effects of musical expectations on affective priming. *Psychomusicology*, *26*, 358–365.

Kissler, J., Assadollahi, R., & Herbert, C. (2006). Emotional and semantic networks in visual word processing: Insights from ERP studies. *Progress in Brain Research*, *156*, 147–83.

Klauer, K. C., & Musch, J. (2003). Affective priming: Findings and theories. In K. C. Klauer & J. Musch (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 9–50). London: Lawrence Erlbaum Associates.

Klauer, K. C., Robnagel, C., & Musch, J. (1997). List-context effects in evaluative priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 246–255.

Knickerbocker, F., Johnson, R. L., Starr, E. L., Hall, A. M., Preti, D. M., Slate, S. R., & Altarriba, J. (2019). The time course of processing emotion-laden words during sentence reading: Evidence from eye movements. *Acta Psychologica*, *192*, 1–10.

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T. C., & Friederici, A. D. (2004). Music, language and meaning: Brain signatures of semantic processing. *Nature Neuroscience*, *7*, 302–307.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and*

Language, *33*, 149–174.

Kroll, J. F., Van Hell, J., Tokowicz, N., & Green, D. W. (2010). The revised hierarchical model: A critical review and assessment. *Bilingualism: Language and Cognition*, *13*, 373–381.

Kuperman, V., Estes, Z., Brysbaert, M., & Warriner, A. B. (2014). Emotion and language: Valence and arousal affect word recognition. *Journal of Experimental Psychology: General*, *143*, 1065–1081.

Li, P., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: A web-based interface for bilingual research. *Behavior Research Methods*, *38*, 202–210.

MaCleod, C. M. (2005). The Stroop task in cognitive research. In A. Wenzel & D. C. Rubin (Eds.), *Cognitive methods and their application to clinical research* (pp. 17–40). Washington: APA.

Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2014). The adaptation of the Affective Norms for English Words (ANEW) for Italian. *Behavior Research Methods*, *46*, 887–903.

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PLoS ONE*, *9*, e89642.

Müller, M., Klein, J., & Jacobsen, T. (2011). Beyond demand: Investigating spontaneous evaluation of chord progressions with the affective priming paradigm. *Music Perception*, *29*, 93–108.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*, 226–254.

Paradis, M. (1993). Linguistic, psycholinguistic, and neurolinguistic aspects of "interference" in bilingual speakers: The activation threshold hypothesis. *International Journal of Psycholinguistics*, *9*, 133–145.

Paradis, M. (2004). *A Neurolinguistic Theory of Bilingualism*. Amsterdam: John Benjamins Publishing.

Pavlenko, A. (2004). “Stop doing that, Ia komu skazala!”: Language choice and emotions in parent—child communication. *Journal of Multilingual and Multicultural Development*, *25*, 179–203.

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

Pavlenko, A. (2017). Affect and decision-making in multilingual speakers. *Current Opinion in Psychology*, *17*, 1–5.

Ponari, M., Rodríguez-Cuadrado, S., Vinson, D., Fox, N., Costa, A., & Vigliocco, G. (2015). Processing advantage for emotional words in bilingual speakers. *Emotion*, *15*, 644–652.

Puntoni, S., Langhe, B. D. E., & Van Osselaer, S. M. J. (2009). Bilingualism and the emotional intensity of advertising language. *Journal of Consumer Research*, *35*, 1012–1025.

Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experiential factors. *Language Testing*, *15*, 1–20.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*, 1161–1178.

Segalowitz, N., Trofimovich, P., Gatbonton, E., & Sokolovskaya, A. (2008). Feeling affect in a second language: The role of word recognition automaticity. *The Mental Lexicon*, *3*, 47–71.

Sianipar, A., Middelburg, R., & Dijkstra, T. (2015). When feelings arise with meanings: How emotion and meaning of a native language affect second language processing in adult learners. *PLoS ONE*, *10*, e0144576.

Sollberger, B., Reber, R., & Eckstein, D. (2003). Musical chords as affective priming context in a word-evaluation task. *Music Perception*, *20*, 263–282.

Steinbeis, N., & Koelsch, S. (2008). Comparing the processing of music and language meaning using EEG and fMRI. *PLoS One*, *3*, e2226.

Steinbeis, N., & Koelsch, S. (2011). Affective priming effects of musical sounds on the processing of word meaning. *Journal of Cognitive Neuroscience*, *23*, 604–621.

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

Tay, R. Y. L., & Ng, B. C. (2019). Effects of affective priming through music on the use of emotion words. *PLoS ONE*, *14*, e0214482.

Vieillard, S., Peretz, I., Gosselin, N., Khalfa, S., Gagnon, L., & Bouchard, B. (2008). Happy, sad, scary and peaceful musical excerpts for research on emotions. *Cognition and Emotion*, *22*, 720–752.

Vieillard, S., Roy, M., & Peretz, I. (2012). Expressiveness in musical emotions. *Psychological Research*, *76*, 641–653.

Võ, M. L.-H., Conrad, M., Kuchinke, L., Urton, K., Hofmann, M. J., & Jacobs, A. M. (2009). The Berlin Affective Word List Reloaded (BAWL-R). *Behavior Research Methods*, *41*, 534–8.

Winskel, H. (2013). The emotional Stroop task and emotionality rating of negative and neutral words in late Thai–English bilinguals. *International Journal of Psychology*, *48*, 1090–1098.

Yiend, J. (2010). The effects of emotion on attention: A review of attentional processing of emotional information. *Cognition and Emotion*, *24*, 3–47.

Appendix A1

Language questionnaire of Experiments 1 and 2

1. At what age did you start learning English __________
2. Rate your proficiency in English according to the different language skills using the following scale (circle the number in the table):

	Very poor	Poor	Fair	Functional	Good	Very Good	Native like
Reading proficiency	○	○	○	○	○	○	○
Writing proficiency	○	○	○	○	○	○	○
Speaking fluently	○	○	○	○	○	○	○
Listening ability	○	○	○	○	○	○	○

3. Did you learn other languages than English?
   - yes
   - no

4. Which other languages do you speak? Please list these languages and indicate the age at which you started learning them. If you do not know any other languages a part from English, please go to Question 13.
5. Please rate your level of proficiency in the languages you identified in Section 4 *(for each language a table was presented like in Question 2)*

6. Estimate, in terms of percentages, how often you use English and any other languages per day (in all daily activities combined)
   English (%) _________, other languages (%) _________
7. Estimate in terms of percentages, how often you use English and any other languages at work/ school/ university.
   English (%) _________, other languages (%) _________
8. Estimate in terms of percentages, how often you use English and any other languages for socializing outside home.
   English (%) _________, other languages (%) _________
9. Estimate in terms of percentages, how often you use English and any other languages for socializing at home.
   English (%) _________, other languages (%) _________
10. Which language do you consider to be the most emotionally expressive to you (i.e. which language do you prefer to use when you express emotions such as anger or affection)? ___________________
11. If there is anything else that you feel is important about your language background or language use, please comment below. _____________________________________

Appendix A2

Country of origin

1. Which country where you born in?
2. How many years have you lived in this country?

Country of residence

3. Which country do you live in at the moment?
4. For how long have you lived in this country?
5. If relevant, please provide additional information

Language Background

6. What is your native language?
    - German
    - other
    - I grew up with more than one language (please specify)
7. How old have you been when you started learning German?
8. In which context have you learned German?
    - school
    - family
    - friends
    - work
    - living in a German speaking country
9. Rate your proficiency in German, please, according to the different language skills using the following scale (tick the appropriate category in the table)

	Very poor	Poor	Fair	Functional	Good	Very Good	Native like
Reading proficiency	○	○	○	○	○	○	○
Writing proficiency	○	○	○	○	○	○	○
Speaking fluently	○	○	○	○	○	○	○
Listening ability	○	○	○	○	○	○	○

10. How old have you been when you started learning English?
11. In which context have you learned English?
    - school
    - family
    - friends
    - work
    - living in an English speaking country

(continues)

Appendix A2 (continued)

12. Rate your proficiency in English, please, according to the different language skills using the following scale (tick the appropriate category in the table)

	Very poor	Poor	Fair	Functional	Good	Very Good	Native like
Reading proficiency	○	○	○	○	○	○	○
Writing proficiency	○	○	○	○	○	○	○
Speaking fluently	○	○	○	○	○	○	○
Listening ability	○	○	○	○	○	○	○

13. Did you learn other languages than German and English?
    - Yes
    - No

(if previous question was responded to with ''Yes'')

Which other languages do you speak? Please list these languages and indicate the age at which you started learning them.

(For each of the languages participants were asked to indicate the context of acquisition and the proficiency as shown above in question 11 and 12)

14. Estimate, in terms of percentages, how often you use German, English and any other languages per day (in all daily activities combined)
    - German (%)
    - English (%)
    - other languages (%)
15. Estimate in terms of percentages, how often you use German, English and any other languages at work/ school/ university.
    - German (%)
    - English (%)
    - other languages (%)
16. Estimate in terms of percentages, how often you use German, English and any other languages for socialising outside home.
    - German (%)
    - English (%)
    - other languages (%)

(continues)

Appendix A2 (continued)

17. Estimate in terms of percentages, how often you use German, English and any other languages for socialising at home.
    - German (%)
    - English (%)
    - other languages (%)

18. Which language do you consider to be the most emotionally expressive to you (i.e. which language do you prefer to use when you express emotions such as anger or affection)?
19. Which language do you use for mental arithmetic/ counting?
20. Which language do you dream in?
21. Do you feel different when using different languages? If so in what way?
22. If there is anything else that you feel is important about your language background or language use, please comment below.

Appendix B1

German and English word stimuli with positive and negative valence (Experiments 1 - 4).

German		**English**	
positive	*negative*	*positive*	*negative*
VORTEIL	UNFALL	ADVANTAGE	ACCIDENT
BRAUT	QUAL	BRIDE	AGONY
KUCHEN	LEICHE	CAKE	CORPSE
MEISTER	ABSTURZ	CHAMPION	CRASH
FREUDE	KRISE	DELIGHT	CRISIS
TRAUM	GEFAHR	DREAM	DANGER
RUHM	TEUFEL	FAME	DEVIL
BLÜTE	VERSAGEN	FLOWER	FAILURE
FREIHEIT	ANGST	FREEDOM	FEAR
URLAUB	FIEBER	HOLIDAY	FEVER
HOFFNUNG	ABFALL	HOPE	GARBAGE
SPASS	TRAUER	JOKE	GRIEF
WUNDER	LEID	MIRACLE	MISERY
PALAST	FEHLER	PALACE	MISTAKE
LIED	ALPTRAUM	SONG	NIGHTMARE*
ERFOLG	MITLEID	SUCCESS	PITY
ABENDROT	ARMUT	SUNSET	POVERTY
REISE	DRUCK	TRAVEL	PRESSURE
VERTRAUEN	GEFÄNGNIS	TRUST	PRISON
WAHRHEIT	STRAFE	TRUTH	PUNISHMENT
MELODIE	ZORN	TUNE	RAGE
FERIEN	AUFSTAND	VACATION	REVOLT
SIEG	NARBE	VICTORY	SCAR
HOCHZEIT	FOLTER	WEDDING	TORTURE

*In Experiment 1 ‘gossip’ was included instead of ‘nightmare’.

Appendix B2

Italian and English word stimuli with positive and negative valence (Experiment 5).

Italian		**English**	
positive	*negative*	*positive*	*negative*
ALBA	MORTO	SUNSET	CORPSE
MERAVIGLIA	ABUSO	MIRACLE	ASSAULT
SETA	CARCERE	SILK	PRISON
TRIONFO	PERICOLO	VICTORY	DANGER
MELODIA	URLO	TUNE	SCREAM
SALUTE	ERRORE	HEALTH	MISTAKE
NEVE	IRA	SNOW	RAGE
VACANZA	DEMONIO	HOLIDAY	DEVIL
VANTAGGIO	INCUBO	ADVANTAGE	NIGHTMARE
PIACERE	PRESSIONE	DELIGHT	PRESSURE
GIARDINO	COLLISIONE	GARDEN	CRASH
FERIE	PENA	VACATION	PAIN
SPOSA	LADRO	BRIDE	THIEF
RICOMPENSA	PIETÀ	REWARD	PITY
PARADISO	FALLIMENTO	HEAVEN	FAILURE
LIBERTÀ	PUNIZIONE	FREEDOM	PUNISHMENT
SCHERZO	INCIDENTE	JOKE	ACCIDENT
MIELE	LUTTO	HONEY	GRIEF
VERITÀ	POVERTÀ	TRUTH	POVERTY
VIAGGIARE	FEBBRE	TRAVEL	FEVER
VINCITORE	MACELLO	CHAMPION	SLAUGHTER
RISATA	PAURA	LAUGHTER	FEAR
TORTA	LITIGIO	CAKE	QUARREL
INCENTIVO	ARMA	INCENTIVE	WEAPON

Appendix C

Accuracy analysis for Experiments 1 to 5

The primary analyses focused on reaction times for trials in which participants correctly identified the valence of the target. In this appendix, we conduct a comparable analysis of the proportion of errors in the different conditions, to test for any influence of a speed-accuracy trade-off. Therefore, the total number of errors per condition was calculated for each participant. Since these were generally not normally distributed, analyses were conducted using Wilcoxon signed-rank tests, with a Bonferroni-corrected alpha level of .017 (to correct for three tests): (i) comparing error rates for positive and negative targets; (ii) comparing congruent and incongruent conditions for positive targets; and (iii) comparing congruent and incongruent conditions for negative targets. For Experiment 3, we also compared the overall error rate for L1 and L2 targets, producing a Bonferroni-corrected alpha level of .013 (to correct for four tests). Mean numbers of errors and standard deviations for Experiments 1 to 5 are presented in Table A1.

In most cases, the analyses below do not show significant effects of congruence between the prime and target on accuracy. When significant effects do arise, they are never in the opposite direction to the effects of congruence on RT reported in the main analyses. In three cases, the significant effects of congruence on accuracy are consistent with RT effects (Expt 1, positive targets; Expt 4, German, positive targets; Expt 5, Italian, positive and negative targets). In the two remaining cases, congruent stimuli show greater accuracy than incongruent stimuli in the absence of any significant effect of congruence on RT in the primary analyses (Expt 4, English, positive and negative targets; Expt 5, English, negative targets). Overall, these results speak against the possibility that the primary RT analyses were influenced by confounding effects of a speed-accuracy trade-off.

Experiment 1: native English speakers - music priming words

Accuracy did not differ for positive and negative targets, $Z = -.887$, $p = .375$. For negative targets, there was no difference in accuracy between congruent and incongruent stimuli, $Z = -.342$, $p = .733$. For positive targets, however, accuracy was greater for congruent than incongruent stimuli, $Z = -3.313$, $p = .001$.

Experiment 2: native English speakers - words priming music

Accuracy was greater for positive than negative targets, $Z = -3.436$, $p = .001$. Accuracy was also greater for congruent than incongruent stimuli both for positive targets, $Z = -2.762$, $p = .006$ and negative targets, $Z = -3.152$, $p = .002$.

Experiment 3: German-English bilinguals - music priming words

In trials with *German* words, accuracy was greater for positive than negative targets, $Z = -3.145$, $p = .002$. There was no difference between targets preceded by congruent or incongruent music primes either for positive, $Z = -1.116$, $p = .265$, or negative word targets, $Z = -.414$, $p = .679$. In trials with *English* words, there were no differences between targets differing in valence, $Z = -.632$, $p = .528$, or those preceded by congruent or incongruent primes; positive: $Z = -.769$, $p = .442$, negative: $Z = -.052$, $p = .959$. Finally, overall accuracy did not differ between the L1 and L2, $Z = -1.497$, $p = .134$.

Experiment 4: German-English bilinguals - words priming music

In trials with *German* words, there was no difference in accuracy between positive and negative targets, $Z = -.714$, $p = .475$. Accuracy was greater for congruent than incongruent stimuli for positive targets, $Z = -3.523$, $p < .001$, but not for negative targets, $Z = -2.051$, $p = .040$. In trials with *English* words there was no difference in accuracy between positive and negative targets, $Z = -.945$, $p = .345$. Accuracy was greater for congruent than incongruent stimuli, for both positive targets, $Z = -2.477$, $p = .013$, and negative targets, $Z = -3.431$, $p = .001$.

Experiment 5: Italian-English bilinguals - words priming music

In trials with *Italian* words, accuracy was greater for positive than negative targets, $Z = -3.066$, $p = .002$. Accuracy was also greater for congruent than incongruent stimuli, both for positive, $Z = -2.917$, $p = .004$, and negative targets, $Z = -3.148$, $p = .002$. In trials with *English* words there was no difference in accuracy between positive and negative music targets, $Z = -2.062$, $p = .039$. While there was no effect of congruence for positive targets, $Z = -1.755$, $p = .079$, for negative targets, accuracy was greater for congruent than incongruent stimuli, $Z = -3.338$, $p = .001$.

Table A1

Mean percentage of errors and standard deviation (SD) for words primed by congruent and incongruent musical excerpts in native English speakers' L1 (Experiment 1) and German-English bilinguals' L1 and L2 (Experiment 3) as well as musical excerpts primed by congruent and incongruent emotion words in native English speakers' L1 (Experiment 2), German-English (Experiment 4) and in Italian-English (Experiment 5) bilinguals' L1 and L2.

	L1		L2	
	Positive target	Negative target	Positive target	Negative target
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Experiment 1				
Positive prime	7.53 (7.01)	9.01 (8.27)		
Negative prime	11.83 (7.84)	8.20 (7.18)		
Experiment 2				
Positive prime	7.58 (7.76)	21.72 (14.80)		
Negative prime	13.26 (11.48)	14.65 (10.16)		
Experiment 3				
Positive prime	2.17 (2.21)	5.25 (4.34)	3.08 (4.57)	2.36 (3.46)
Negative prime	2.83 (2.94)	4.50 (3.49)	3.44 (5.57)	3.90 (11.28)
Experiment 4				
Positive prime	4.39 (6.29)	11.18 (7.74)	7.07 (5.13)	10.51 (7.47)
Negative prime	13.60 (9.73)	7.02 (5.65)	10.69 (7.25)	4.89 (4.24)
Experiment 5				
Positive prime	5.29 (5.95)	15.20 (10.31)	7.77 (7.23)	14.75 (10.92)
Negative prime	9.80 (9.48)	10.59 (9.35)	10.13 (8.70)	9.46 (7.96)