The embodiment of emotional words in a second language: An eye-movement study

Naveed A. Sheikh and Debra Titone

Department of Psychology, Centre for Research on Brain, Language and Music, McGill University, Montreal, QC, Canada

(Received 14 July 2014; accepted 7 February 2015)

The hypothesis that word representations are emotionally impoverished in a second language (L2) has variable support. However, this hypothesis has only been tested using tasks that present words in isolation or that require laboratory-specific decisions. Here, we recorded eye movements for 34 bilinguals who read sentences in their L2 with no goal other than comprehension, and compared them to 43 first language readers taken from our prior study. Positive words were read more quickly than neutral words in the L2 across first-pass reading time measures. However, this emotional advantage was absent for negative words for the earliest measures. Moreover, negative words but not positive words were influenced by concreteness, frequency and L2 proficiency in a manner similar to neutral words. Taken together, the findings suggest that only negative words are at risk of emotional disembodiment during L2 reading, perhaps because a positivity bias in L2 experiences ensures that positive words are emotionally grounded.

Keywords: Emotion; Second language; Embodied cognition; Proficiency; Eye-movement measures of reading.

Emotion is thought to play a foundational role in grounding semantic representations during first language (L1) processing (Kousta, Vigliocco, Vinson, Andrews, & Del Campo, 2011). However, the role of emotion in grounding second language (L2) semantic representations remains an open question: Some studies find that bilinguals process emotional and neutral words (e.g., sex vs. pin, respectively) differently in their L2, similar to what native speakers do in their L1 (e.g., Sutton, Altarriba, Gianico, & Basnight-Brown, 2007), whereas other studies do not (e.g., Degner, Doycheva, & Wentura, 2012). Such discrepancies may arise for a variety of potential reasons, which include whether bilinguals process emotional words differently as a function of L2 proficiency, emotional polarity or comprehension demands. Here, we test these alternatives, with the ultimate goal of determining whether an embodied theoretical approach to language (e.g., Barsalou, 1999) can explain when bilinguals process L2 emotional words like L1 emotional words. We first review the bilingual literature on L2 emotional word processing, and then describe the embodied approach that comprises the theoretical basis of this study.

One source of complexity regarding L2 emotional word processing concerns L2 proficiency. For example, bilinguals who are as proficient in
their L2 as their L1 show intact emotional word processing in emotional Stroop tasks (Eilola, Havelka, & Sharma, 2007; Sutton et al., 2007). Conversely, bilinguals who are less proficient in their L2 than their L1 show reduced or no emotional effects in other tasks. For example, bilinguals show reduced skin conductance responses to L2 childhood reprimands and taboo words in emotion rating tasks (Harris, Ayçiçeği, & Gleason, 2003; see also, Segalowitz, Trofimovich, Gatbonton, & Sokolovskaya, 2008; Degner et al., 2012, for implicit affect association and affective priming tasks, respectively). The fact that low L2 proficient bilinguals show impaired emotional processing effects implies that bilinguals require additional, and likely direct experience with L2 words in order to treat them emotionally. It also implies that automatic cross-language activation that routinely occurs during bilingual language processing, which is highly probable for low L2 proficient bilinguals (e.g., Thierry & Wu, 2007; reviewed in Baum & Titone, 2014; Kroll, Bobb, Misra, & Guo, 2008), is insufficient for leading to emotional word processing effects in the L2.

Another source of complexity with respect to L2 emotional word processing is the potential effect of emotional polarity (i.e., negative vs. positive emotionality). Most work focuses exclusively on negative words—thus, few have compared negative and positive words to each other, while simultaneously taking into account the other linguistic ways that words vary (e.g., frequency). A recent exception is Conrad, Recio, and Jacobs (2011), who compared emotional and non-emotional words using event-related potentials (ERP) gathered during lexical decisions in two groups of bilinguals differing in L2 proficiency. Here, more proficient bilinguals showed an enhanced early posterior negativity and late positive complex in L1 for negative and positive words vs. neutral words, with delayed but similar effects in the L2. Less proficient bilinguals, however, showed ERP modulations in the L2 for positive but not negative words. Thus, negative but not positive emotional words may be treated in an unemotional manner in the L2. In contrast with this work, Degner et al. (2012) and Segalowitz et al. (2008) found impoverished emotional word processing in an implicit affect association task and an affective priming task, respectively, which did not interact with continuous measures of L2 proficiency. Interestingly, however, Degner et al. found that bilinguals who reported frequently using their L2 processed emotional words in their L2 more like they did in their L1, even after controlling for L2 proficiency.

Thus, an open question concerns the amount and kind of L2 experience required for intact emotional word processing. One potentially useful approach for addressing this question is to view emotional effects on word processing as part of a larger class of language embodiment phenomena, which considers how real-world experiences ground semantic representations generally. Consider the words *apple* and *grief*, for which people have very different bodily experiences. Apples are tangible objects that may be grasped, eaten or physically manipulated in any number of ways (e.g., sliced, cooked, stewed, etc.), which are all typically classed as sensorimotor experiences. Grief is a less tangible concept that may have indirect physical manifestations that we typically class as emotional experiences (e.g., a lump in one’s throat). In embodied approaches to language, words are grounded in whichever type of experience they tend to co-occur with during language use (Zwaan, 2008). Vigliocco and colleagues recently proposed that all words are represented by linguistic information (e.g., word associations), emotional information and sensorimotor information (Kousta et al., 2011). They further stipulated that emotional and sensorimotor information primarily represent abstract and concrete semantics, respectively. With respect to L2 word processing in bilinguals, an embodied approach to language forces us to consider more deeply the behavioural ecology of bilinguals (Green, 2011), and whether this ecology would afford the embodiment of emotional words in one’s L2. Accordingly, bilinguals’ L2 words may be “disembodied” (Pavlenko, 2012) if they do not use their L2 in social contexts that provide opportunities for co-occurrence with emotional experiences. This idea is consistent with work on language choice.
For example, bilinguals curse to express anger using their L1 more than their L2 (Dewaele, 2004), and switch to their L1 for emotionally charged interactions with their romantic partners, even when their partners have limited knowledge of the language (Pavlenko, 2005). Thus, bilinguals may experimentally lock their L2 words out of emotional contexts, thus rendering them less salient than L1 words during on-going language processing.

Bilinguals might also show differences between negative and positive words to the extent that they experientially lock their L2 out of negative but not positive emotional contexts. For example, Conrad et al. (2011) suggested that bilinguals showed no emotional effects for negative words but not positive words because a positivity bias ensures that bilinguals use their L2 in emotionally positive contexts. The notion of a positivity bias is supported by work on the Pollyanna hypothesis showing that human communication is centred on emotionally positive exchanges (Boucher & Osgood, 1969). It is also supported by work on emotion regulation showing that positive emotion is more often up-regulated and less often down-regulated than negative emotion (Matsumoto, Yoo, Hirayama, & Petrova, 2005), particularly with respect to interactions with colleagues and strangers vs. family and friends. An embodied perspective also forces us to consider differences between emotional and sensorimotor embodiment in bilinguals—which are reflected in emotional and concreteness advantages, respectively (Kousta et al., 2011). Accordingly, bilinguals may have difficulty grounding L2 words in emotional experiences specifically but not sensorimotor experiences more generally. To test this idea, we can capitalise on interactions found in the L1 literature between word emotionality and concreteness, which shows that emotion is more likely to facilitate semantic categorisation accuracy for abstract but not concrete words (Newcombe, Campbell, Siakaluk, & Pexman, 2012). Similarly, in eye-movement paradigms, emotion facilitates first-pass reading times for abstract words and not concrete words, and conversely, concreteness facilitates neutral words but not emotional words (Sheikh & Titone, 2013). Thus, if bilinguals have disembodied negative L2 words, those negative words should be facilitated by concreteness, like neutral words, and not by emotionality.

Differences among bilinguals in L2 proficiency may not have predicted emotional effects in previous studies because proficiency does not reflect the kinds of contexts in which bilinguals use their L2. However, L2 proficiency might modulate facilitation by concreteness to the extent that sensorimotor referents of words are independent of context. For example, there is no evidence suggesting that sensorimotor experiences vary as a function of social context, as has been shown for emotional experiences (Matsumoto et al., 2005). Thus, L2 proficiency should modulate concreteness advantages for negative and neutral words. Moreover, L2 proficiency should modulate word frequency effects, which reflect how often words occur in language, to the extent that frequency effects also do not depend on context-specific L2 experiences. Consistent with this conjecture, differences in language use measures among bilinguals predict L2 and L1 frequency effects (e.g., Whitford & Titone, 2012) even though the usage measures do not reflect the context of use. Bilinguals generally also show larger L2 than L1 frequency effects, presumably because bilinguals experience words less frequently in their L2 (e.g., Whitford & Titone, 2012). Thus, bilinguals should show less dependence on word frequency for modulating embodiment effects, in contrast to native speakers (Juhasz, Yap, Dicke, Taylor, & Gullick, 2011; Sheikh & Titone, 2013).

The present study

The work just reviewed leads to the following open questions about bilinguals reading in their L2: (1) Do bilinguals show reduced or eliminated emotional facilitation for negative words, but not positive words, relative to neutral words? (2) Do bilinguals show facilitation of negative words by concreteness, like neutral words? (3) Do differences in L2 proficiency among bilinguals predict facilitation by concreteness and frequency, but not emotionality? We address these questions using eye-movement measures of natural sentence reading (Rayner, 2009). We used semantically neutral sentences in which we
embedded English target words that varied on emotional valence, frequency and concreteness.

French–English bilinguals (all French L1) read English sentences presented in their entirety, with no goal other than comprehension. Since embodiment theories are based mostly on data from single word paradigms, they do not help target particular reading time measures for the disembodiment predictions. Thus, we analysed all first-pass fixation time measures from first fixation duration (FFD) to go-past time (GPT) to identify the point in time at which emotional disembodiment manifests in bilinguals for negative words. To obtain more evidence that our bilinguals have disembodied negative words but not positive words, we compared the bilinguals with native English speakers from a previous study that used the same materials (Sheikh & Titone, 2013).

METHOD

Participants

We recruited 34 bilinguals (French L1, English L2, mean age = 24.97, standard deviation [SD] = 5.16) at McGill University that were less proficient in their L2 than their L1 and primarily used their L2 in formal environments rather than informal social environments (see supplementary materials). We report below how we determined sample size, all data exclusions, manipulations and measures in the study.

Materials and design

We used 156 target words grouped into 52 triplets, each consisting of a negative, neutral and positive word. Frequency and concreteness were manipulated, but balanced across emotional categories (Table 1). Length was longer for low-compared with high-frequency words, which was statistically controlled in all analyses. We created three sentences for each triplet which were well-formed when combined with any of the triplet members. Differences in sentential fit across emotional categories were avoided, and any potential effects were statistically controlled (see Sheikh & Titone, 2013, where we used the same stimuli and describe their development in depth). Targets and sentence frames were presented once in a given experimental list and the combinations were counterbalanced across participants (Table 2).

A modified Language Experience and Proficiency Questionnaire (LEAP-Q; Marian, Blumenfeld, & Kaushanskaya, 2007) was also administered. We used the LEAP-Q to verify that all participants were native speakers of French and dominant in their native language, that English was their L2, and to measure L2 proficiency using a 7-point scale that ranged from 1 (beginner) to 7 (near-native).

Apparatus and procedure

Eye movements were recorded using an Eyelink 1000 that sampled eye position every millisecond. The right eye was recorded but viewing was

Table 1. Mean (SDs) for lexical attributes for target words and their latent semantic analysis (LSA) values within sentence frames

|                  | Low frequency |                      |                  | High frequency |                      |
|------------------|---------------|----------------------|------------------|----------------|----------------------|
|                  | Negative      | Neutral              | Positive         | Negative       | Neutral              | Positive         |
| Log frequency\(^a\) | 7.58 (1)      | 7.18 (1.27)          | 7.86 (1.03)      | 9.89 (0.77)    | 10.31 (0.91)         | 10.11 (0.91)    |
| Valence\(^b\)    | 3.15 (0.58)   | 5.28 (0.51)          | 7.11 (0.42)      | 3.05 (0.85)    | 5.39 (0.54)          | 7.08 (0.45)     |
| Arousal\(^b\)    | 5.13 (0.8)    | 4.36 (0.87)          | 5.42 (1.07)      | 5.25 (0.65)    | 4.54 (0.73)          | 5.54 (0.88)     |
| Concreteness\(^c\) | 429.62 (106.9) | 463.6 (138.2)       | 459.14 (125.8)  | 445.4 (120.85) | 471.93 (116.24)      | 468.42 (129.05) |
| Length           | 6.31 (2.32)   | 7.2 (3.52)           | 7 (2.14)         | 5.9 (2.34)     | 4.81 (2.37)          | 5.39 (1.75)     |
| LSA value        | −0.005 (0.06) | 0.013 (0.06)         | 0.005 (0.06)     | 0.004 (0.05)   | 0.009 (0.04)         | 0.010 (0.06)    |

\(^a\)The English Lexicon Project (Balota et al., 2007); \(^b\)Kousta, Vinson, and Vigliocco (2009); \(^c\)MRC Psycholinguistic Database (Coltheart, 1981).
binocular. Sentences were displayed on a 21-in. Viewsonic CRT monitor with 144 Hz refresh rate, 1024 × 768 resolution, yellow 10-point Monaco font on black background, with three characters subtending approximately 1° of visual angle. Yes or no questions followed 21% of trials to ensure participants read for comprehension. The LEAP-Q was administered after the reading task.

RESULTS

We tested our hypotheses using first-pass fixation duration measures, which reflect early stages of lexical processing (rather than later measures that reflect the integration of word meaning into the sentential context) (Rayner, 2009). We analysed FFD (the duration of the first fixation on a word), single fixation duration (SFD; fixation time in cases where the word was fixated exactly once), gaze duration (GD; the sum of the durations of all fixations made during the first pass, before the eyes left the target region) and GPT (the sum of the durations of all fixations on the word from the point when the word is first fixated up until the eyes move past the word to the right). We also analysed the probability of fixating and regressing to a target, which are presented as supplementary material.

The eye-movement measures were analysed in R (R Development Core Team, 2010) with linear mixed models (LMMs), and generalised LMMs for the binary data, using the lme4 package (Bates, Mächler, & Dai, 2009). To test our hypotheses regarding negative and positive valence, we specified predictors for valence (negative vs. neutral vs. positive), frequency (continuous), concreteness (continuous), L2 proficiency (continuous) and their interactions in the fixed-effect structure. The model baseline was set to neutral valence because our hypotheses specifically concern differences between negative and positive words vs. neutral words. All continuous predictors were standardised. We also included a covariate for word length (continuous) in all analyses.

We maximised the random-effect structure to the extent possible (Barr, Levy, Scheepers, & Tily, 2013). We included random intercepts for participant (subject), word and triplet; and by-subject random slopes for valence (including the slope–intercept correlation) and frequency (including the slope–intercept correlation for SFD, but not the other measures).

We also used likelihood ratio tests to confirm that the highest-order significant effect in each model was justified to ensure that the data were not overfitted, though the full models are presented in Table 4 for the sake of comparison. The models were fit to log-transformed observations to meet model assumptions, and non-log predicted values were calculated for plotting partial effects. We report the estimated coefficient ($b$), standard

| Counterbalance | Sentence                                                                 |
|----------------|--------------------------------------------------------------------------|
| List 1         | The art teacher presented the *smoke* that the students were going to paint.  
The news report mentioned the *hill* that was discussed at work.  
The school paper reviewed the *drink* that had caused concern among students. |
| List 2         | The art teacher presented the *drink* that the students were going to paint.  
The news report mentioned the *smoke* that was discussed at work.  
The school paper reviewed the *hill* that had caused concern among students. |
| List 3         | The art teacher presented the *hill* that the students were going to paint.  
The news report mentioned the *drink* that was discussed at work.  
The school paper reviewed the *smoke* that had caused concern among students. |

Note: Targets were not presented with asterisks during the experiment.
error (SE) and $t$ value. An absolute value of $|t| \geq 1.96$ falls within the 95% confidence interval and is significant. For convenience, we calculated $p$ values using an upper bound for the degrees of freedom (number of observations – the number of fixed effects in the model), which should not lead to anti-conservative values according to Baayen (2008), since we have well over the minimum 100 observations that he indicates should suffice for calculating $p$ values in this manner. As seen below, there was no discrepancy between these $p$ values and the $|t| \geq 1.96$ criterion.

Means and SDs for the eye-movement measures are presented in Table 3. Mean accuracy on the comprehension questions was 81.83% (SD = 4.85). Twenty-six per cent of the trials were discarded because of track losses, lack of fixation on the target words, blinks, or when observations were shorter than 80 ms. Visually identified outliers were eliminated from the remaining trials (FFD > 800 [0.57%], SFD > 800 [0.71%], GD > 1500 [0.22%], GPT > 1500 [0.32%]).

Setting aside the effect of L2 proficiency, we can see in the LMM outputs (Table 4) that negative and neutral words did not differ for FFD, SFD or GD. However, the GPT model showed a single main effect for negative valence, suggesting that negative words (with sufficient processing time) eventually became faster than neutral words. In contrast to negative words, significant main effects for positive valence across all measures indicate that positive words were faster than neutral words. Moreover, in the FFD data, there was an unexpected two-way interaction between positive valence and frequency indicating that the emotional advantage for positive words was even larger for high-frequency words compared with low-frequency words. Additional models with negative valence as the baseline (as opposed to neutral described above) showed that positive words were faster than negative words across all measures ($p < .05$). Thus, the emotional advantage is reduced for negative words but not positive words, and emotional advantages occur for both high-frequency and low-frequency words.

Regarding the effect of L2 proficiency, there were no interactions for FFD or SFD. However, there were L2 proficiency interactions for GD and GPT. In both cases, this manifested as a four-way interaction between positive valence, frequency, concreteness and proficiency. L2 proficiency never interacted with negative valence. Thus, the effect of proficiency on GD and GPT was identical for negative and neutral words, which manifested in

| Low concreteness | High concreteness |
|------------------|-------------------|
| **Measures**     | **Negative**      | **Neutral**     | **Positive**     | **Negative** | **Neutral** | **Positive** |
|                   | Low frequency     |                   |                   | High frequency |
|                   | First fixation    | 280 (107)         | 279 (105)         | 272 (102)     | 285 (116)   | 273 (100)   | 279 (107) |
|                   | Single fixation   | 293 (113)         | 310 (118)         | 282 (103)     | 294 (115)   | 280 (101)   | 279 (104) |
|                   | Gaze duration     | 378 (224)         | 435 (247)         | 371 (213)     | 362 (198)   | 372 (231)   | 330 (155) |
|                   | Go-past time      | 418 (243)         | 485 (271)         | 397 (224)     | 374 (202)   | 394 (258)   | 342 (169) |
|                   | Fixation probability | 0.91         | 0.92              | 0.97          | 0.87        | 0.88        | 0.89 |
|                   | Regression probability | 0.15         | 0.18              | 0.16          | 0.15        | 0.16        | 0.09 |
|                   | High frequency    | 272 (108)         | 276 (108)         | 255 (89)      | 269 (93)    | 279 (109)   | 255 (88) |
|                   | First fixation    | 275 (106)         | 284 (114)         | 261 (92)      | 275 (97)    | 287 (112)   | 257 (88) |
|                   | Single fixation   | 321 (155)         | 323 (147)         | 297 (137)     | 309 (133)   | 306 (133)   | 284 (122) |
|                   | Gaze duration     | 333 (166)         | 339 (160)         | 311 (148)     | 319 (146)   | 327 (168)   | 297 (136) |
|                   | Go-past time      | 0.82              | 0.82              | 0.82          | 0.83        | 0.78        | 0.81 |
|                   | Regression probability | 0.13         | 0.14              | 0.15          | 0.15        | 0.16        | 0.14 |

Note: Frequency and concreteness were continuous variables in the analyses.
Table 4. Test of interaction between valence, frequency, concreteness and proficiency, showing emotional advantage for positive valence across all measures, which is absent for negative valence except for GPT, and a frequency × concreteness × proficiency interaction that is different for negative and neutral words vs. positive words

| Fixed effect                  | First fixation |          |          |          |          |          |          |          |          |          |          |
|-------------------------------|---------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
|                               | b  | SE  | t       | b  | SE  | t       | b  | SE  | t       | b  | SE  | t       |
| (Intercept)                   | 5.56 | 0.03 | 185.02*** | 5.61 | 0.03 | 172.27*** | 5.76 | 0.04 | 158.94*** | 5.82 | 0.04 | 153.52*** |
| Neg                           | 0   | 0.02 | -0.05   | -0.01 | 0.02 | -0.48   | -0.04 | 0.02 | -1.83   | -0.05 | 0.02 | -2.4*   |
| Pos                           | -0.03 | 0.01 | -2.22*  | -0.06 | 0.02 | -3.63*** | -0.09 | 0.02 | -4.38*** | -0.11 | 0.02 | -5.1***  |
| Freq                          | -0.01 | 0.01 | -0.78   | -0.02 | 0.01 | -1.23   | -0.06 | 0.02 | -3.2**  | -0.05 | 0.02 | -2.81**  |
| Conc                          | -0.01 | 0.01 | -0.6    | -0.01 | 0.01 | -0.49   | 0.01  | 0.02 | 0.37    | 0.01  | 0.02 | 0.8     |
| Prof                          | -0.02 | 0.03 | -0.73   | -0.03 | 0.03 | -0.95   | -0.07 | 0.03 | -2.06*  | -0.06 | 0.04 | -1.82    |
| Len                           | -0.01 | 0.01 | -0.74   | 0.02  | 0.01 | 1.53    | 0.1   | 0.01 | 6.73*** | 0.14  | 0.02 | 9.12***  |
| Neg × Freq                    | -0.01 | 0.02 | -0.38   | -0.01 | 0.02 | -0.53   | 0     | 0.02 | 0.04    | -0.03 | 0.02 | -1.09    |
| Pos × Freq                    | -0.03 | 0.01 | -2.15*  | -0.02 | 0.02 | -1.14   | -0.01 | 0.02 | -0.37   | 0     | 0.02 | 0.04    |
| Neg × Conc                    | 0   | 0.01 | -0.24   | 0     | 0.02 | -0.22   | -0.01 | 0.02 | -0.57   | -0.02 | 0.02 | -0.81    |
| Pos × Conc                    | 0.02 | 0.01 | 1.58    | 0.03  | 0.02 | 1.46    | 0.02  | 0.02 | 1.19    | 0.02  | 0.02 | 0.77     |
| Freq × Conc                   | -0.01 | 0.01 | -0.85   | 0     | 0.01 | -0.17   | -0.01 | 0.01 | -0.95   | 0     | 0.01 | -0.2     |
| Neg × Prof                    | 0.01 | 0.01 | 0.49    | -0.01 | 0.02 | -0.31   | -0.01 | 0.02 | -0.34   | -0.01 | 0.02 | -0.71    |
| Pos × Prof                    | 0.02 | 0.01 | 1.4     | 0.02  | 0.02 | 0.98    | 0     | 0.02 | 0.21    | -0.01 | 0.02 | -0.49    |
| Freq × Prof                   | 0   | 0.01 | 0.07    | 0.01  | 0.01 | 0.65    | 0.04  | 0.01 | 2.63**  | 0.03  | 0.02 | 2.02*    |
| Conc × Prof                   | -0.01 | 0.01 | -0.72   | -0.01 | 0.01 | -0.52   | 0     | 0.01 | -0.24   | -0.01 | 0.01 | -0.55    |
| Neg × Freq × Conc             | 0.01 | 0.02 | 0.48    | 0     | 0.02 | 0.09    | 0.01  | 0.02 | 0.54    | 0     | 0.03 | 0.07     |
| Pos × Freq × Conc             | -0.01 | 0.01 | -0.85   | -0.01 | 0.02 | -0.75   | 0     | 0.02 | -0.19   | -0.01 | 0.02 | -0.45    |
| Neg × Freq × Prof             | -0.02 | 0.01 | -1.18   | -0.03 | 0.02 | -1.39   | -0.02 | 0.02 | -0.83   | -0.01 | 0.02 | -0.74    |
| Pos × Freq × Prof             | 0   | 0.01 | 0.16    | 0     | 0.02 | -0.19   | -0.03 | 0.02 | -1.96   | -0.02 | 0.02 | -1.32    |
| Neg × Conc × Prof             | 0.01 | 0.01 | 1.08    | 0.02  | 0.02 | 1.04    | 0.02  | 0.02 | 0.91    | 0.03  | 0.02 | 1.45     |
| Pos × Conc × Prof             | -0.01 | 0.01 | -0.41   | 0     | 0.02 | 0.05    | 0     | 0.02 | -0.12   | 0     | 0.02 | -0.22    |
| Freq × Conc × Prof            | 0.01 | 0.01 | 0.76    | 0.01  | 0.01 | 0.52    | 0.03  | 0.01 | 2.7**   | 0.03  | 0.01 | 2.95**   |
| Neg × Freq × Conc × Prof      | 0.01 | 0.01 | 0.74    | 0.01  | 0.02 | 0.74    | -0.03 | 0.02 | -1.41   | -0.03 | 0.02 | -1.59    |
| Pos × Freq × Conc × Prof      | -0.01 | 0.01 | -0.91   | -0.02 | 0.02 | -1      | -0.04 | 0.02 | -2.21*  | -0.04 | 0.02 | -2.6**   |

Note: Neg = negative; Pos = positive; Freq = frequency; Conc = concreteness; Prof = second language proficiency.
*p < .050, **p < .010, ***p < .001.
both cases as a three-way interaction between frequency, concreteness and proficiency. We visualised the pattern of effects from the model for GD using the coefficients for these significant interactions in a partial effects plot (Figure 1).

We fit separate models to the GD data, split by median concreteness, to follow-up on the four-way interaction in Figure 1, and to test our hypothesis that L2 proficiency predicts the concreteness advantage but not the emotional advantage. For words low in concreteness (i.e., abstract words), there was a main effect for positive valence \((b = -0.1, \ SE = 0.03, t = -3.56, p < .001)\) and no interactions, which indicates that for abstract words, processing was faster for positive compared with neutral words irrespective of frequency and proficiency. This emotional advantage for abstract positive words can be seen in all of the panels in Figure 1 at low concreteness values.

For words high in concreteness (i.e., concrete words), there was an interaction between positive valence, frequency and proficiency \((b = -0.08,\)

![Figure 1. Frequency, concreteness and proficiency produced identical effects for negative and neutral words, which differed from positive words.](image-url)
SE = 0.03, \( t = -2.86, p < .01 \). This interaction indicates that frequency and proficiency produced identical effects for concrete negative and neutral words, but different effects for positive words. Specifically, concreteness facilitated processing for negative and neutral low-frequency words, but only at high levels of proficiency. And as words became more concrete, this reduced the emotional advantage for positive words. This pattern can be seen in Figure 1 in the bottom right panel at high concreteness values. We also found this pattern of effects in separate models fit to the GPT data split by concreteness. Thus, the results show that concreteness advantages, but not emotional advantages, depend on proficiency, and that concreteness produces identical effects for negative and neutral words, but different effects for positive words.

L2 readers vs. L1 readers

Next, we test differences between the bilinguals reading in their L2 (L2 readers) in this study and 43 native speakers of English\(^1\) (L1 readers) that read the same materials in a previous study (Sheikh & Titone, 2013). To compare L2 vs. L1 readers, we tested whether language group interacts with valence, frequency and concreteness. For a complete exposition of the L1 data, see Sheikh and Titone (2013). The most relevant finding in the previous study is that the concreteness advantage was limited to low-frequency neutral words in L1 readers, producing three-way interactions between valence, frequency and concreteness for negative and positive words for the first-pass fixation time measures. Thus, if emotional disembodiment is specific to negative words in L2 readers, a comparison with L1 readers should produce four-way interactions between language group, valence, frequency and concreteness for negative words, but not positive words. As expected, this four-way interaction was significant for negative words for FFD (\( b = 0.04, SE = 0.02, t = 2.09, p < .05 \)), SFD (\( b = 0.04, SE = 0.02, t = 1.93, p = .05 \)), GD (\( b = 0.09, SE = 0.02, t = 4.01, p < .001 \)) and GPT (\( b = 0.06, SE = 0.02, t = 2.53, p < .05 \)). There were no four-way interactions for positive words (\( ps > .12 \)).

DISCUSSION

The purpose of the study was to investigate whether bilingual readers exhibit L2 emotional word processing effects, guided by predictions from an embodied approach to L2 word representation. We found that bilinguals reading in their L2 showed word processing facilitation by embodied knowledge only for some types of words, presumably because they capitalise on some but not all sources of experiential information to ground L2 semantics. Specifically, bilinguals processed positive words faster than neutral words, suggesting that they capitalise on emotionally positive experiences. Moreover, bilinguals with high L2 proficiency processed negative and neutral words faster when they were concrete compared to abstract, suggesting they also capitalise on sensorimotor experiences. However, bilinguals processed negative words faster than neutral words only for the latest first-pass measure, suggesting that they do not as readily capitalise on emotionally negative experiences.

Previous work on L1 embodiment indicates that a concreteness advantage, where observed, is diagnostic of emotional neutrality because it does not occur for emotionally charged words (Sheikh & Titone, 2013). Thus, the concreteness advantage for negative words, coupled with the absence of an early emotional advantage for negative words suggests that only negative words were emotionally disembodied in our bilingual participants. The finding that negative words alone are emotionally disembodied is consistent with research showing that bilinguals prefer to use their L2 less often than their L1 in emotional contexts (Dewaele, 2004;\(^1\) Average age for the L1 readers (\( M = 24.14, SD = 5.14 \)) did not differ from L2 readers. We also verified that there were no differences in total alexithymia scores (Bagby, Parker, & Taylor, 1994) between L1 readers (\( M = 43, SD = 11.26 \)) and L2 readers (\( M = 40, SD = 9.73 \)), because emotional word processing by the L1 readers was modulated by individual differences in alexithymia (Sheikh & Titone, 2013).
An L1 preference for emotional contexts would reduce co-occurrence between words and emotional experiences, leaving words emotionally disembodied in the L2. We see evidence of this in our data given that increased L2 proficiency among our bilingual participants correlated with increased L2 use in formal contexts like the workplace but not informal social exchanges. However, the same is not true of positive words, possibly because a positivity bias ensures that L2 use co-occurs with emotionally positive experiences. We also found that the emotional advantage for positive words was larger for high-frequency items compared with low-frequency items for first fixation duration only, which was unexpected, suggesting that bilinguals more easily retrieve positive emotional semantics for high-frequency words than low-frequency words at the earliest stage of processing, though this needs to be confirmed in future work.

The selective disembodiment for negative words observed here is consistent with Conrad et al. (2011), who found that bilinguals who were less proficient in their L2 than L1 showed emotional effects only for positive and not negative words. Crucially, we extend their findings by showing that this valence asymmetry occurs during natural reading, at the earliest stages of comprehension. Bilinguals are able to eventually compute the negative valence of disembodied words, which emerged as a relatively late emotional advantage for negative words. The late advantage is consistent with Harris et al. (2003) who observed attenuated skin conductance responses for taboo words and childhood reprimands even though bilinguals were ultimately able to identify their negative valence on a rating task. Interestingly, bilinguals continue to process negative and neutral words similarly in terms of how negative and neutral words are influenced by frequency, concreteness and proficiency, even after the emergence of the late emotional advantage. Thus, the late emotional advantage for negative words and the emotional advantage for positive words do not differ solely in terms of time course.

The embodiment approach also suggests a potential explanation for why studies vary in whether reduced or eliminated emotional effects are observed during L2 language processing. When the L2 is at least equal in proficiency to the L1, there is no L1 preference that locks the L2 out from emotional contexts, as that preference is partly driven by greater L1 proficiency (Dewaele, 2004). Thus, bilinguals can ground words in emotional experiences (Eilola et al., 2007; Sutton et al., 2007). In contrast, when bilinguals are less proficient in their L2, the L2 gets locked out of those contexts and bilinguals end up with emotionally disembodied words (Harris et al., 2003; Segalowitz et al., 2008). Of course, bilinguals could acquire words in emotional contexts if forced into emotional contexts irrespective of proficiency—the present results do not preclude that possibility, and indeed, suggest a future test of the embodiment hypothesis.

The findings also clarify the role of L2 proficiency. Specifically, L2 proficiency predicted concreteness advantages but not emotional advantages presumably because sensorimotor experiences are not context-specific the way emotional experiences are (Matsumoto et al., 2005). Moreover, concreteness advantages were limited to low-frequency words at high levels of L2 proficiency, similar to L1 readers (Sheikh & Titone, 2013). Thus, the bilingual experiences that underlie word processing facilitation by frequency and concreteness seem less dependent on context than the experiences that underlie facilitation by emotion. Interestingly, facilitation by concreteness at high levels of L2 proficiency did not occur for first fixation duration or single fixation duration. Thus, the representational changes correlated with L2 proficiency do not appear to be activated on the very first fixation on target words.

The present findings also add to the small number of studies for emotional target words embedded in sentences using eye-movement measures of reading, though these previous studies all examined L1 processing. Some work on L1 processing using isolated words found that valence had a monotonic effect on word recognition (i.e., negative words were slower than neutral words, which were slower than positive words; Kuperman, Estes,
Brysbaert, & Warriner, 2014). However, other studies found that people process negative and positive words faster than neutral words (e.g., Vinson, Ponari, & Vigliocco, 2014), which converges with recent eye-movement results for low-frequency words (Scott, O’Donnell, & Sereno, 2012; Sheikh & Titone, 2013). There is also an earlier eye-movement study by Hyönä and Häikö (2005), but emotionally charged words in that study were only presented parafoveally and never directly fixated by readers, as in our study. This methodological difference makes it difficult to directly compare the results across the studies. In contrast, Scott et al. (2012) used a natural sentence reading paradigm, and the present findings converge with their results for high-frequency words. Scott et al. found that for low-frequency words, people processed negative and positive words (which did not differ) faster than neutral words. For high-frequency words, positive words were faster than neutral words, but negative words did not differ from neutral words. This is precisely the pattern that we observed here for bilinguals. One possible explanation that Scott et al. (2012) posited to explain their findings for high-frequency items involves the selective attenuation of emotional charge for negative words. Specifically, they proposed that negative emotionality may be reduced by high frequency of exposure, which they compared to desensitisation in psychotherapy. Thus, although the specifics differ, our interpretation converges with their suggestion in terms of negative words not being as emotional as positive words.

To conclude, our findings demonstrate that bilinguals have emotionally disembodied negative words during L2 reading, and that these words are instead grounded in sensorimotor experiences like neutral words. Our study also shows that L2 proficiency predicts concreteness advantages but not emotional advantages during natural reading. Thus, sensorimotor experiences are more readily available than emotionally negative experiences for grounding L2 words. Similarly, emotionally positive experiences are more readily available for grounding L2 words than emotionally negative experiences.

Acknowledgements
Special thanks to Dr. Gabriella Vigliocco for providing us with stimuli normed by her lab.

Disclosure statement
No potential conflict of interest was reported by the authors.

Funding
This work was supported by Natural Sciences and Engineering Research Council of Canada [grant number 261769], Social Sciences and Humanities Research Council of Canada [grant number 410-2011-1806] and a doctoral scholarship from Le Fonds de recherche du Québec-Nature et technologies.

Supplemental data
Supplemental data for this article can be accessed here.

REFERENCES
Baayen, R. H. (2008). Analyzing linguistic data: A practical introduction to statistics using R. Cambridge: Cambridge University Press.
Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia scale—I. Item selection and cross-validation of the factor structure. Journal of Psychosomatic Research, 38, 23–32. doi:10.1016/0022-3999(94)90005-1
Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., … Treiman, R. (2007). The English lexicon project. Behavior Research Methods, 39, 445–459.
Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language, 68, 255–278.
Barsalou, L. W. (1999). Perceptual symbol systems. Behavioral and Brain Sciences, 22, 577–660.
Bates, D., Maechler, M., & Dai, B. (2009). Lme 4: Linear mixed-effects models using S4 classes. R package version 0.999375-27. Retrieved from http://lme4.r-forge.r-project.org
Baum, S., & Titone, D. (2014). Moving toward a neuroplasticity view of bilingualism, executive control, and aging. Applied Psycholinguistics, 5, 857–894.
SECOND LANGUAGE EMOTIONAL WORDS

Boucher, J., & Osgood, C. E. (1969). The Pollyanna hypothesis. Journal of Verbal Learning and Verbal Behavior, 8, 1–8. doi:10.1016/S0022-5371(69)80002-2

Coltheart, M. (1981). The MRC psycholinguistic database. The Quarterly Journal of Experimental Psychology Section A, 33, 497–505. doi:10.1080/14640748108400805

Conrad, M., Recio, G., & Jacobs, A. M. (2011). The time course of emotion effects in first and second language processing: A cross cultural ERP study with German–Spanish bilinguals. Frontiers in Psychology: Language Sciences, 2, 351. doi:10.3389/fpsyg.2011.00351

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. doi:10.1017/S1366728911000095

Dewaele, J.-M. (2004). Blistering barnacles! What language do multilinguals swear in? Estudios de Sociolinguistica, 5(1), 83–105.

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

Green, D. W. (2011). Language control in different contexts: The behavioral ecology of bilingual speakers. Frontiers in Psychology: Language Sciences, 2, 103. doi:10.3389/fpsyg.2011.00103

Harris, C. L., Ayçiçeği, A., & Gleason, J. B. (2003). Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language. Applied Psycholinguistics, 24, 561–579.

Hyöna, J., & Häikiö, T. (2005). Is emotional content obtained from parafoveal words during reading? An eye movement analysis. Scandinavian Journal of Psychology, 46, 475–483. doi:10.1111/j.1467-9450.2005.00479.x

Juhasz, B. J., Yap, M. J., Dicke, J., Taylor, S. C., & Gullick, M. M. (2011). Tangible words are recognized faster: The grounding of meaning in sensory and perceptual systems. Quarterly Journal of Experimental Psychology, 64, 1683–1691. doi:10.1080/17470218.2011.605150

Kousta, S.-T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. Journal of Experimental Psychology: General, 140, 14–34. doi:10.1037/a0021446

Kousta, S.-T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. Cognition, 112, 473–481. doi:10.1016/j.cognition.2009.06.007

Kroll, J. F., Bobb, S. C., Misra, M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. Acta Psychologica, 128, 416–430. doi:10.1016/j.actpsy.2008.02.001

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.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. Journal of Speech, Language, and Hearing Research, 50, 940–967. doi:10.1044/1092-4388(2007/067)

Matsumoto, D., Yoo, S. H., Hirayama, S., & Petrova, G. (2005). Development and validation of a measure of display rule knowledge: The display rule assessment inventory. Emotion, 5, 23–40. doi:10.1037/1528-3542.5.1.23

Newcombe, P. I., Campbell, C., Siakaluk, P. D., & Pexman, P. M. (2012). Effects of emotional and sensorimotor knowledge in semantic processing of concrete and abstract nouns. Frontiers in Human Neuroscience, 6, 275. doi:10.3389/fnhum.2012.00275

Pavlenko, A. (2005). Emotions and multilingualism. Cambridge: Cambridge University Press.

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

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. Quarterly Journal of Experimental Psychology, 62, 1457–1506. doi:10.1080/17470210902816461

R Development Core Team. (2010). R: A Language and environment for statistical computing. Retrieved from http://www.R-project.org

Scott, G. G., O’Donnell, P. J., & Sereno, S. C. (2012). Emotion words affect eye fixations during reading. Journal of Experimental Psychology: Learning, Memory, and Cognition, 38, 783–792. doi:10.1037/a0027209

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. doi:10.1075/ml.3.1.05seg

Sheikh, N. A., & Titone, D. A. (2013). Sensorimotor and linguistic information attenuate emotional word processing benefits: An eye-movement study. Emotion, 13, 1107–1121. doi:10.1037/a0032417

Sutton, T. M., Altarriba, J., Gianico, J. L., & Basnight-Brown, D. M. (2007). The automatic access of COGNITION AND EMOTION, 2016 499
emotion: Emotional Stroop effects in Spanish–English bilingual speakers. *Cognition & Emotion, 21*, 1077–1090. doi:10.1080/02699930601054133

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *PNAS, 104*, 12530–12535.

Vinson, D., Ponari, M., & Vigliocco, G. (2014). How does emotional content affect lexical processing? *Cognition & Emotion, 28*, 737–746.

Whitford, V., & Titone, D. (2012). Second-language experience modulates first- and second-language word frequency effects: Evidence from eye movement measures of natural paragraph reading. *Psychonomic Bulletin & Review, 19*, 73–80. doi:10.3758/s13423-011-0179-5

Zwaan, R. A. (2008). Experiential traces and mental simulations in language comprehension. In M. DeVega, A. M. Glenberg, & A. C. Graesser (Eds.), *Symbols, embodiment, and meaning* (pp. 165–180). Oxford: Oxford University Press.