Bilingualism: Language and Cognition

cambridge.org/bil

Emotion effects in second language processing: Evidence from eye movements in natural sentence reading

Enze Tang and Hongwei Ding

Speech-Language-Hearing Center, School of Foreign Languages, Shanghai Jiao Tong University, Shanghai, China

Research Article

Cite this article: Tang, E., & Ding, H. (2024). Emotion effects in second language processing: Evidence from eye movements in natural sentence reading. *Bilingualism: Language and Cognition*, **27**, 460–479. https://doi.org/10.1017/S1366728923000718

Received: 10 March 2023
Revised: 18 September 2023
Accepted: 20 September 2023
First published online: 18 October 2023

Keywords:
emotion; eye movement; second language; sentence reading

Corresponding author:
Hongwei Ding
Email: hwding@sjtu.edu.cn

Abstract

There exists insufficient eye-tracking evidence on the differences in emotional word processing between the first language (L1) and second language (L2) readers. This study conducted an eye-tracking experiment to investigate the emotional effects in L2 sentence reading, and to explore the modulation of L2 proficiency and individual emotional states. Adapted from Knickerbocker et al. (2015), the current study recorded eye movements at both early and late processing stages when late Chinese–English bilinguals read emotion-label and neutral target words in natural L2 sentences. Results indicated that L2 readers did not show the facilitation effects of lexical affective connotations during sentence reading, and they even demonstrated processing disadvantages for L2 emotional words. Additionally, the interaction effect between L2 proficiency and emotion was consistently significant for the measure of total reading time in positive words. Measurements of participants' depressive and anxious states were not robustly correlated with eye movement measures. Our findings supplemented new evidence to existing sparse eye-tracking experiments on L2 emotion processing, and lent support to several theoretical frameworks in the bilingual research field, including the Emotional Contexts of Learning Theory, Lexical Quality Hypothesis and Revised Hierarchical Model.

1. Introduction

The multilingual community has been enlarged over the years, and it is necessary to explore whether language representation is identical or different when the language is used as the first or a foreign language. In every language, emotion plays an essential role in grounding the semantic connotations (Kousta et al., 2011) and functions indispensably in human communication. Therefore, an increasing body of research has investigated how emotional stimuli modulate the processing of the non-dominant language, especially among late bilinguals. A late bilingual is defined as one who was not immersed in a bilingual environment since birth and did not acquire a second language until later stages of language development (Toivo & Scheepers, 2019), and the boundary age is normally around six years old (Brothers et al., 2021; X. Liu et al., 2017). This late bilingual population has been found to feel the emotional meanings of words in a second language (L2) less intensely (Ferré et al., 2022) and to use fewer high-arousal emotional words facing moral dilemmas in the L2 context (Kyriakou et al., 2022). These characteristics of L2 emotion processing have certain implications in real-life settings, such as the Foreign Language Effect in decision making (Keysar et al., 2012; L. Liu et al., 2021), the clinical intervention and rehabilitation of language use (Monaco et al., 2019), and the pedagogy for foreign language teaching. As such, there exists practical and theoretical significance to investigate emotion processing among the bilingual population in different scenarios.

Some consensus has been reached in this research field, but disagreements still exist. Many studies have investigated the differences in emotional density between the dominant and non-dominant languages by presenting L1 and L2 readers with the same set of emotional words and directly comparing their responsive activities. A line of studies has collected the affective ratings of a set of emotional words from the L1 and L2 populations, and has mostly reported attenuated or less extreme emotional feelings in the L2 context (Ferré et al., 2022; Garrido & Prada, 2021). Similarly, previous studies that have recorded participants' physiological responses also seemed to have identified weaker responses to emotional contents in the L2 group, including weaker pupillary effects (Toivo & Scheepers, 2019) as well as decreased facial motor resonance and skin conductance responses (Baumeister et al., 2017; Jankowiak & Korpal, 2018). However, there have been conflicting findings in studies that have examined how participants responded to the emotional content in their L1 or L2 using psychological

© The Author(s), 2023. Published by Cambridge University Press

This article has earned badges for transparent research practices: Open Data and Open Materials. For details see the Data Availability Statement.

and cognitive tasks. Some studies have reported weaker memory facilitation effects elicited by the emotional connotations in the L2 group (Baumeister et al., 2017), but others have controversially observed comparable affective priming effects (Degner et al., 2012), emotion-memory effects (Ayçiçeği-Dinn & Caldwell-Harris, 2009), emotional Stroop effects (Ahn & Jiang, 2022) and response speed of lexical decision (Ponari et al., 2015) in the L2 population. According to the Emotional Contexts of Learning Theory (Harris et al., 2006), late bilinguals usually acquire their L2 without sufficient involvement of emotional experiences, so that their emotional embodiment tends to be less grounded and they may process the L2 words more semantically with reduced automaticity of affective activation (Pavlenko, 2012). However, considering the inconsistent findings in previous studies, further investigations are still needed to examine the similarities and differences between the L1 and L2 populations.

Despite the intense interest in the comparison of L1-L2 emotional processing, few studies, however, have touched on the nuance caused by L2 proficiency within the bilingual population. As proposed by the Revised Hierarchical Model (Kroll & Stewart, 1994), higher L2 proficiency or exposure is accompanied by stronger conceptual connections of L2 words, so the bilinguals with higher L2 proficiency could directly access the meaning of L2 words without the mediation of L1 translation. This model suggests that less proficient bilinguals cannot directly access the L2 lexical nodes to the conceptual nodes, and they have to indirectly refer to the lexical-level links from L2 to L1 translation equivalents. As for proficient bilinguals, they can establish direct connections between L2 lexical nodes and conceptual nodes, which parallel the semantic representations of L1 speakers. Based on this model, it can thus be inferred that highly proficient bilinguals would show more native-like behaviours and responses while processing L2 words, compared to the less proficient L2 learners who primarily rely on the lexical connections of L1 counterparts. In a similar vein, the Lexical Quality Hypothesis (Perfetti, 2007) indicates that higher quality representations of a word's identity are more fully-specified and stable, including the orthographic, phonological and semantic constituents. As such, this hypothesis supports that word identification and processing would be more efficient and less cognitively demanding if the language proficiency is high. Hence, as suggested by Imbault et al. (2021), a prediction pertaining to the proficiency effects on L2 processing could be made–that is, less proficient L2 learners would demonstrate stronger neutralization while processing the emotional connotations of L2 words, compared to the more proficient L2 speakers.

Based on these theoretical frameworks, bilinguals with different proficiency levels could be predicted to show dissimilar behaviours or responses during L2 word processing, due to different semantic representations. Some previous studies have explored the proficiency effects on L2 emotion processing, but they have reported inconsistent findings on this issue. Imbault et al. (2021) collected the valence and arousal ratings for 2,628 L2 English words by a large-scale affective rating experiment, and they compared the valence scores across five L2 proficiency levels, finding that more proficient bilinguals tended to report native-like affective ratings. Caldwell-Harris et al. (2011) recorded the skin conductance responses of bilinguals during lexical processing, and they observed that participants with higher language proficiency showed elevated physiological responses to English endearments. However, a few prior studies using behavioural tasks did not seem to identify such proficiency effects. For example, Degner et al. (2012) reported similar affective priming effects between the L1 and L2 populations, and they did not observe significant differences caused by self-rated language proficiency. Similarly, Ahn and Jiang (2022) adopted a classic Stroop paradigm to investigate the automaticity of emotion activation in L2 readers, but they did not observe the main effect of L2 proficiency on the emotional Stroop effect. As such, L2 proficiency might not be predictive of different processing patterns for emotional connotations among L2 readers. Considering the highly discrepant but sparse empirical evidence on the association between L2 proficiency and L2 emotion processing, it is urgent and necessary to use a few more psychophysiological and neuroimaging paradigms to investigate the differences in affective processing among bilinguals with different proficiency levels.

The effects of a few influential factors on L2 emotion processing have been explored. In terms of the lexical materials, emotion polarity has shown certain modulating effects. Specifically, positive words seemed to elicit stronger processing advantages among L2 readers (Shenaut & Ober, 2021), while negative words, including swear and taboo words, in turn, might be processed similarly to neutral words (Arriagada-Mödinger & Ferreira, 2022) or even demonstrate processing disadvantages (Ferré et al., 2018) in the L2 context. In addition to the behavioural observations, some neuroimaging studies have also identified differences in processing positive and negative L2 words. Jończyk et al. (2016) enrolled Polish–English bilinguals to read sentences ending with positive, negative and neutral words, which manifested an incomplete and suppressed semantic access to L2 negative words based on the reduced N400 amplitudes. Such an early suppression and later stronger reevaluation of negative semantics, rather than positive or neutral semantics, corroborated previous results (Y. J. Wu & Thierry, 2012), indicating an involvement of additional brain structures and more efforts in negative word processing (Sulpizio et al., 2019). Therefore, it seems that positive words might show processing advantages over negative words in the L2 population. However, prior studies have mainly observed the influence of emotional valence in L2 with single-word stimuli, and such an effect may also need to be examined at the above-word levels.

Another concern is the use of emotion-label and emotion-laden words as the experimental materials, which have been reported to possibly involve different processing procedures over the recent years. Pavlenko (2008) has provided specific definitions for these two word types. Emotion-label words refer to the words that straightforwardly describe particular affective states (e.g., "happy", "sad") or processes (e.g., "to surprise", "to worry"), while emotion-laden words denote the words that indirectly express (e.g., "loser", "champion") or elicit emotions from the interlocutors (e.g., "birthday", "funeral"). By definition, compared to emotion-laden words, emotion-label words should normally have more salient emotional content and direct conceptual links to the core emotions they represent (Haro et al., 2023; Knickerbocker & Altarriba, 2013). Previous studies have found facilitated processing of emotion-label words in a few behavioural tasks, such as the lexical decision task (Kazanas & Altarriba, 2015). Other research using event-related potentials also observed that emotion-label words elicited enhanced N170 compared to emotion-laden words (Zhang et al., 2017).

In the L2 population, similar processing facilitation for emotion-label words has also been reported. Specifically, bilinguals showed higher accuracy and shorter response time while processing L2 emotion-label words in emotion categorization tasks (D. Tang et al., 2023) and priming tasks (C. Wu et al.,

https://doi.org/10.1017/S1366728923000718 Published online by Cambridge University Press

2022), and they needed smaller brain activation during the emotion perception primed by L2 emotion-label words (C. Wu et al., 2022). According to the "mediated account" (Altarriba & Basnight-Brown, 2011), the processing of emotional connotations in emotion-laden words needs to be mediated by the labelled conceptual meanings and associated affective experiences. However, since emotion-label words directly and explicitly represent emotions, the processing of their emotional meanings would be easier and more automatic. Similarly, the "emotion duality model" also suggests that emotional responses activated by emotion-label words are automatic and biologically rooted, while the processing of emotion-laden words requires more cognitive efforts through a reflective system (Imbir et al., 2019; D. Tang et al., 2023). Besides the difficulty and higher cognitive demands in processing emotion-laden words, the establishment of emotional connections to the emotion-laden words also seems to be more complex with larger individual and cross-cultural variations. In an extreme case illustrated by Pavlenko (2008), some emotion-laden words that are commonly perceived as insults, such as swearwords, may appear as friendly terms of affection in other languages, while words like "liberal" or "elite" may be perceived as insults in certain cultures. In fact, while constructing the cross-culturally universal "colexification" networks of emotional expressions (Jackson et al., 2019), researchers only considered words that could be directly felt as efficient emotion concepts. Therefore, considering the less demanding processing requirements, higher emotional salience and the more cross-culturally common emotional connotations for emotion-label words, we thus used emotion-label words as our experimental materials for this preliminary eye-tracking exploration of L2 emotion processing.

In addition to the factors related to lexical materials, L2 emotion processing might also be influenced by individual emotional states. Previous studies have mostly acknowledged that the processing of emotional information is abnormal in mood disorders (Panchal et al., 2019; E. Tang et al., 2022). Even in the healthy population, individuals' emotional states have also been reported to cause differences in word learning and emotional word processing. In a novel word learning paradigm (Guo et al., 2018), converging results were obtained from a semantic category judgment task and a word-picture semantic consistency judgment task, suggesting that negative emotional states could lead to worse performance in word learning. However, during the processing of emotional words, negative mood could narrow participants' attention and enhance their distinction of words within the positive, negative or neutral clusters (Sereno et al., 2015). In a recent review article (Naranowicz, 2022), it has been summarized that positive moods could activate global processing while negative moods strengthen detailed and local processing during language comprehension. In addition to the effects of individual mood on language processing, the anxiety states also seemed to influence people's comprehension of context information. Using the spelling task and lexical decision task, Blanchette and Richards (2003) found that, compared to the control group, participants with state anxiety were more likely to adopt an emotional interpretation of the ambiguous homophones under the emotional context and a neutral interpretation under the neutral context. Hence, these previous studies have demonstrated the influence of individuals' mood and anxiety state on the processing of emotion-related information. Knickerbocker et al. (2015) have analyzed the associations between participants' eye movement measures in L1 processing and the emotional measures of their depression and anxiety traits, and reported significant interactions between participants' anxiety and the emotion processing effects for negative words. However, few studies have explored how individual emotional states influence emotional word processing in the L2 context. Therefore, the current study was also intended to evaluate the depression and anxiety states of the L2 readers, and to associate these measures with their eye movements in L2 word processing.

Despite substantial explorations on L2 emotion processing, the eye-tracking method has not yet been widely used in this field to investigate the facilitation effects of L2 emotional connotations using passive sentence reading tasks. Earlier research (Scott et al., 2012) carried out among native speakers showed that emotion qualities of words can accelerate the reading time during natural sentence reading, which justified the effectiveness of eye-tracking experiments for the investigation of reading patterns as a function of different emotional meanings of words. Sheikh and Titone (2013, 2016) selected 52 triplets of English words and compiled them into sentence frames, and recorded the eye movements of native and bilingual readers during their reading procedure. Results indicated that emotion facilitation effects occurred for both positive and negative words with low frequency among L1 readers (Sheikh & Titone, 2013), whereas, among L2 readers, benefits were only identified for positive words, and their negative word processing was modulated by concreteness, frequency and L2 proficiency similarly to neutral words (Sheikh & Titone, 2016). These two experiments that compared the reading behaviours of L1 and L2 readers with the same set of materials suggested a joint influence of linguistic, emotional and sensorimotor information on the early stages of word processing in natural sentence reading. Moreover, they seemed to indicate that negative words might not be grounded in emotional experiences in the L2 context (Monaco et al., 2019). Hence, positive words could be predicted to elicit more advantages over negative words in L2 word processing. Knickerbocker et al. (2015) respectively tracked the reading behaviours for positive and negative words in native readers, and they corroborated that both positive and negative emotion-label words can facilitate the L1 reading at both early and late processing stages, compared to neutral words. The materials in this study have not yet been used in bilinguals to examine whether similar effects also stood in the non-dominant language environment. A few extant bilingual studies recorded participants' pupillary responses to emotional and neutral words, rather than words fitted to sentence frames, and they consistently reported attenuated pupil responses for emotional words in L2 (Toivo & Scheepers, 2019; Yao et al., 2023). Nevertheless, there is still a dearth of direct eye-tracking evidence to support the claim that bilinguals processed L2 emotional words differently from L1 readers, especially at the sentence level.

The present study primarily intended to investigate whether the emotion facilitation effects in L1 also occur during L2 sentence reading, and thus supplement new evidence to limited eye-tracking experiments on L2 emotion word processing. Taking existing evidence into consideration, we tried to address the following research questions, including (1) whether and how L2 emotion-label words influence the reading behaviours of bilinguals; (2) whether bilinguals with high and low L2 proficiency show similar or different patterns of L2 emotional word processing; and (3) whether depressive and anxious symptoms are associated with bilinguals' eye movements for emotional word processing. Based on previous findings and the theoretical frameworks, we predicted that bilinguals would not demonstrate the facilitation effects of emotional connotations in L2 processing,

https://doi.org/10.1017/S1366728923000718 Published online by Cambridge University Press

but positive words might show processing ease compared to negative words. In the current study, we would use the term "emotion processing effect" hereafter to represent the ease of processing emotional words compared to neutral words. Meanwhile, highly proficient bilinguals should process L2 words faster than less proficient ones in general, and they might also show stronger emotion processing effects considering their advancement in semantic representations and lexical quality for L2 words. Last, since the associations between individual emotional states and the eye movements of L1 processing have been relatively unstable (Knickerbocker et al., 2015), such associations could be even weaker in L2 processing considering the increased cognitive burdens. However, as there only exists little evidence concerning how individual depression and anxiety states influence the processing of L2 emotional words, we could thus tentatively predict that few eye movement measures would be correlated with individuals' emotional states, and any established associations should be interpreted with cautions and require further validations.

A few novelties and contributions of our current study are demonstrated here. Methodologically, tracking eye movements during L2 reading can appropriately observe bilinguals' automatic physiological responses to emotional words under involuntary control, and the recorded data can objectively reflect the emotional effects elicited by the emotional connotations of words. Moreover, due to factors such as language acquisition contexts and L1-L2 interaction, the L2 population tends to show prominent individual variations (Fricke et al., 2019), even for L2 readers with similar proficiency levels. As such, the comparability of emotional processing effects obtained in separate L2 groups might be decreased in a traditional between-subject design, even when the demographic background and language use experience were matched. Therefore, we adopted a within-subject design to minimize the potentially unexpected individual variances, so that the emotion processing effects of L2 positive and negative words could be more reasonably compared together. As for the selected materials, we used sentences rather than isolated words as experimental stimuli, which were more ecologically valid and could restore the emotion processing effects in the authentic scenes. It has been suggested that contextual information was more accurately predictive of lexical representation than isolated words (Garrido & Prada, 2021; Snefjella & Kuperman, 2016), and the use of sentence materials could be more sensitive to demonstrate the differences in context interpretation and lexical processing across the L1 and L2 populations. In terms of statistical analyses, we intended to analyze as many as 13 eye-tracking measures for all emotional words and examine the influences of L2 proficiency and individual emotional states on L2 emotion processing. Hence, our current study would not only supplement limited eye-tracking evidence in the field of L2 emotion processing, but also provide new statistics for some factorial effects that have been insufficiently investigated in previous studies. Last, we have drawn on some theoretical frameworks of bilingual research, including the Emotional Contexts of Learning Theory, Lexical Quality Hypothesis and Revised Hierarchical Model to establish hypotheses for our findings, and would in turn validate relevant theoretical premises with our empirical data.

2. Methods

2.1. Participants

Forty-two Chinese–English bilinguals from Shanghai Jiao Tong University participated in this experiment. All participants were native speakers of Mandarin Chinese and reported English as their second language. They all had normal or corrected-to-normal vision, with no history of language impairments, reading difficulty, or learning disorders. Unqualified participation was excluded from later analyses. Three participants wore thick glasses and their eye movements could not be successfully tracked. Two participants did not finish all trials due to fatigue or emergent personal issues and chose to withdraw from the experiment. One participant had an operation of crystalline lens resection, and failed to pass the calibration. As a result, the eye-tracking data from 36 participants (18 males) aging between 18 and 28 ($M = 22.14$, $SD = 2.45$) were analyzed. The final set of bilingual participants were all born and raised in mainland China with Chinese being the dominant language in daily use, and they were all late L2 learners who were not exposed to English until the age of six (mean age of first exposure = 8.89, $SD = 2.49$, range = 6-13). Based on the self-rated English proficiency on the 1-7 scale, these included bilinguals had an averagely intermediate proficiency in reading ($M = 5.51$, $SD = 1.04$), writing ($M = 4.92$, $SD = 0.98$), listening ($M = 4.78$, $SD = 1.48$) and speaking ($M = 4.68$, $SD = 1.40$) capacities. Only one participant once lived in an English-speaking country (the UK) for 10 months during the undergraduate period, while others did not claim overseas living experiences for more than six months. As for the language history, 10 participants reported knowing a third language but with limited proficiency (Japanese: n = 6; French: n = 3; Russian: n = 1). All participants signed a written informed consent before the formal experiment, and received monetary compensation for their participation.

2.2. Materials

The present study constructed the experiment lexicons and sentence frames based on Knickerbocker et al. (2015), where the effects of affective meanings on sentence reading have been examined in native speakers of English. In total, 36 positive emotion-label words, 36 negative emotion-label words, and 36 neutral contrasts were selected. For both the emotional and neutral words, there were 12 nouns, 12 verbs, and 12 adjectives. To ensure that all selected words were emotionally appropriate for the current study, their emotional characteristics were re-examined in the L2 population before our formal experiment.

The influence of emotional connotations on L2 reading was evaluated in two separate blocks, with positive or negative emotional words compared against the neutral words in each block. Within each block, 72 sentence frames were designed to locate the emotional and neutral words in pairs so that each sentence frame could grammatically and semantically fit one emotional word and one paired neutral word. The target words never appeared as the first three or the last three words of the sentence. As illustrated in Table 1, there were two sentence frames designed for each word pair, and each word would randomly appear in either of the sentence frames for each participant. Moreover, each target word and sentence frame would only appear once within one block, and the word-sentence matching was counterbalanced across all participants. As a result, each participant read 72 sentences in each block. All experimental sentences can be accessed in the two Appendixes. The mean sentence length was 11.36 words ($SD = 1.50$) in the block containing positive words, and it was 13.63 words ($SD = 2.04$) in the block containing negative words. To ensure that the sentences were appropriate for L2 reading, the difficulty, understandability and predictability were also re-assessed in the L2 population.

https://doi.org/10.1017/S1366728923000718 Published online by Cambridge University Press

Table 1. Examples of experimental stimuli.

Sentence frame	Word pair
Neutral/positive words	
Frame one	patient/jolly
Even on busy days, the **patient/jolly** conductor waited for everyone to board the train.	
Frame two	
During the holidays, the **patient/jolly** elf handed out toys to every child.	
Neutral/negative words	
Frame one	elbow/hurt
Helen watched the older student **elbow/hurt** the other child during an argument.	
Frame two	
At the game, the boy did not mean to **elbow/hurt** the player on the other team.	

2.3. Normative examination

First, the affective properties of the selected lexical items were re-evaluated in L2 to ensure that these experimental words could effectively introduce the positive, negative or neutral meanings as intended. Twenty Chinese–English bilinguals who did not participate in the formal eye-tracking experiment were enrolled to rate the familiarity, valence and arousal parameters of all included words. On the seven-point scale, *1* referred to *the least familiar*, *the most negative* and *the least arousing*, while *7* denoted *the most familiar*, *positive* and *arousing*. As shown in Table 2, all included words were highly familiar to the L2 population, and there was no significant difference between neutral words and either positive ($t(70) = 0.24$, $p = 0.81$) or negative ($t(70) = 0.24$, $p = 0.81$) words. Compared to the neutral words, positive words received significantly higher valence ($t(70) = 9.07$, $p < 0.001$) and arousal ($t(70) = 8.85$, $p < 0.001$) scores, while negative words received significantly lower valence ($t(70) = -13.52$, $p < 0.001$) and higher arousal ($t(70) = 6.51$, $p < 0.001$) scores. In summary, these selected emotional words contained prominent emotional meanings with moderately high arousing effects, and the neutral words were emotionally neutral with low arousing effects in L2. Hence, these words were suitable for the present experiment. Other parameters, including the word length, frequency and orthographic neighbourhood size, have been matched in Knickerbocker et al. (2015), and these objective measures were regarded as similar in the L1 and L2 conditions.

As the experiment materials of our current study followed closely to those of Knickerbocker et al. (2015), we have further examined whether the distribution of valence and arousal ratings of these experimental words was matched between the L1 and L2 populations. Since the seven-point and nine-point scales were differently adopted by our study and that previous study, the average raw scores were standardized into the percentages they took into account of the full scales for each affective metric, which ensured the cross-study comparability. To clearly observe how much difference there was between the actual percentages and the midpoint of the scales, we subtracted 50% from the actual percentage statistics at each metric. As depicted in Fig. 1, the valence ratings of the three word types had the same distribution in both populations, and those of the L1 population seemed to be more extreme than the L2 population. As for the arousal dimension, although the overall distribution was roughly the same across the two populations, these selected negative words seemed to be more arousing in the L1 context but less arousing in the L2 context, compared to the positive words. Despite these minor differences, it has been guaranteed that the emotional words could be clearly distinguished from the neutral words in both populations, so the selected words could satisfy our current research objectives. However, we still acknowledge that the observed differences here might impose certain limitations on the interpretation of our findings.

Second, properties at the sentence level, especially difficulty, understandability, and predictability, were also assessed on the

Table 2. Properties of experimental lexical items presented by mean (M) and standard deviation (SD).

	Neutral		Positive		Negative	
Parameters	*M*	*SD*	*M*	*SD*	*M*	*SD*
Familiarity	5.89	0.99	5.96	1.28	5.94	0.88
Valence	4.11	0.57	5.54	0.76	2.40	0.49
Arousal	3.08	0.69	4.66	0.82	4.13	0.68
Length	5.94	1.57	6.42	1.76	6.25	1.93
Word Frequency (Kučera-Francis)	29.39	27.13	35.44	49.08	23.97	26.15
Word Frequency (log HAL)	8.85	1.32	8.33	1.82	8.43	1.34
Orthographic neighbourhood Size	2.83	3.74	2.28	3.61	2.94	5.15

https://doi.org/10.1017/S1366728923000718 Published online by Cambridge University Press

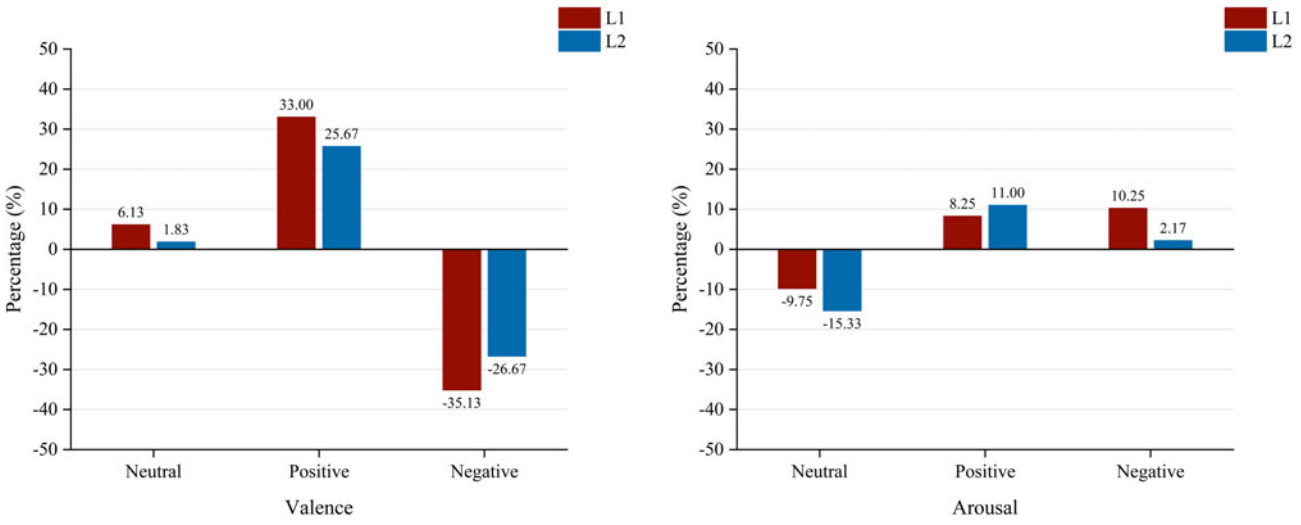


Figure 1. Comparison of valence and arousal distributions of experimental words between the L1 and L2 populations.

seven-point scale in the L2 population. Specifically, difficulty was defined as how difficult the sentences were for L2 reading, which focused on the use of unfamiliar words and the complexity of grammatical structure. Understandability was intended to measure whether the sentence meanings were clear and acceptable, as a reflection of how well the target words fit the sentence structures. Predictability measured whether the target words could be predicted by the preceding sentence frames. Another group of 20 participants who did not participate in the formal eye-tracking experiment provided the difficulty and understandability norms for the experiment sentences, with score *1* representing *extremely difficult/not understandable* and *7* representing *extremely easy/very understandable*. Sentences were counterbalanced so that each rater evaluated each sentence frame and each target word only once. As shown in Table 3, all experimental sentences were easy and natural to read, and there were no significant differences between the sentences containing emotional or neutral words under each condition. Another 15 bilinguals were given the incomplete sentence frames preceding the target words, and were asked to write down one English word for each sentence that was the most likely to occur at the next position. Results showed that the predictability was low for positive ($M = 0.00\%$), negative ($M = 0.83\%$), and neutral ($M = 0.00\%$) target words in these sentence contexts.

2.4. Apparatus and procedure

Eye-movements were recorded by the SR Eyelink 1000 eye-tracker (SR Research, ON, Canada) at a sampling rate of 1000 Hz per second. Experiment sentences were displayed 73 cm away from participants' eyes on a 24-inch CRT screen (resolution: 1,024*768; refresh rate: 100 Hz) of a DELL computer, and each sentence was presented in a single line in 12-point Courier New font where two characters equalled 1° of visual angle. Although the sentence reading was binocular, all calibrations and recordings of the eye movements were based on the right eye.

On arrival, participants registered their language history with an adapted version of the LEAP-Q questionnaire (Marian et al., 2007). Participants were then seated comfortably in front of the eye-tracker, with their heads positioned securely on the chin and forehead rest. Before recording the eye movements, the nine-point grid calibration was carried out, which was accepted if an average error was < 0.50° of visual angle (approximately 1 character) with a maximum error < 1°. With an acceptable calibration, a fixation cross would occur before each trial at the position where the first word of the following sentence was located, and drift calibration was conducted. Re-calibration was performed when necessary. To become familiarized with the experiment tasks, participants firstly conducted 10 practice trials, all of which did not appear again in the formal experiment, followed by 72 experiment trials presented in randomized order. During the experiment, participants were instructed to read the single-line sentences presented on the screen silently in a natural manner and to press the space bar on the keyboard if finished. Following a third of the sentences appeared a yes-or-no comprehension question that was related to the meaning of the preceding sentence but was irrelevant to the target words and the post-target regions (two-word region after the target words). Participants were asked to press one of the two buttons representing "yes" or "no" as a response, and the left-right corresponding relationships

Table 3. Properties of experimental sentences presented by mean (standard deviation).

Block	Parameter	SE	SN	AS	Statistics
Positive	DI	6.28 (0.35)	6.28 (0.43)	6.28 (0.39)	$t(142) = 0.03$, $p = 0.98$
	UN	6.18 (0.41)	6.12 (0.47)	6.15 (0.44)	$t(142) = 0.73$, $p = 0.47$
Negative	DI	6.21 (0.30)	6.26 (0.29)	6.24 (0.29)	$t(142) = 1.11$, $p = 0.27$
	UN	5.92 (0.43)	6.01 (0.43)	5.97 (0.43)	$t(142) = 1.23$, p = 0.22

Note: SE: sentences containing emotional target words; SN: sentences containing neutral target words; AS: all sentences; DI: difficulty; UN: understandability

https://doi.org/10.1017/S1366728923000718 Published online by Cambridge University Press

between the answers and buttons were counterbalanced. Half of the questions appeared after the sentences containing emotional words and the other half after the neutral sentences, and the number of questions requiring a "yes" answer or a "no" answer was also evened. Only the data of participants whose average accuracy for comprehension questions was at least 1.5 times the chance level (75%; Z. Wu & Wang, 2022) were valid for further analyses. The formal experiment session had two blocks, with one comparing positive and neutral words and the other negative and neutral words. Two blocks were counterbalanced in presentation sequence, and the whole experiment lasted around 40 minutes.

After the eye-tracking experiment, participants continued to complete a set of measures for affective state and L2 proficiency. The symptoms of depression were assessed by the Beck Depression Inventory Short Form (BDI-SF; Beck & Steer, 1993), the 13-item cognitive-affective subscale of BDI, which has been found effective and valid in depression detection (Furlanetto et al., 2005; Stukenberg et al., 1990). The Spielberger State Trait Anxiety Inventory (STAI; Spielberger, 1983) was used to measure the state anxiety and trait anxiety of participants, which consists of the STAI-S and STAI-T subscales to respectively measure the transient status and general propensity of the anxious disposition.

In the last step, English proficiency was measured by the LexTALE and C-test. The LexTALE is a quick and valid predictor of English vocabulary knowledge (Lemhöfer & Broersma, 2012), and has been used for L2 vocabulary assessment (Zhang et al., 2020). As an alternative to the traditional cloze test, the C-test (Raatz & Klein-Braley, 1981) normally consists of five independent short texts, and the second half of every second word in each text is deleted so that 20 blanks need to be completed. It outperforms some vocabulary tests as a valid and reliable predictor of receptive language skills, L2 academic achievements, or even general language proficiency (Daller et al., 2020; Harsch & Hartig, 2015). Following previous L2 research that used the C-test to discriminate L2 proficiency groups (Qiu, 2022), we compiled five excerpts from prior studies (Babaii & Ansary, 2001; Dörnyei & Katona, 1992), whose reliability, validity and sensitivity were satisfactory in the L2 population. Based on the median split of average language proficiency test scores (Qian et al., 2018), L2 participants were divided into the higher proficiency group (higher than 65) and the lower proficiency group (lower than 65).

Table 4 reported the statistics of affective and L2 proficiency assessments in our current experiment. There were significant differences in LexTALE ($t(34) = 4.21$, $p < 0.001$), C-test ($t(34) = 8.70$, $p < 0.001$), and the general L2 proficiency ($t(34) = 7.78$, $p < 0.001$) between two groups, while no between-group significant difference was found for STAI ($t(34) = 1.36$, $p = 0.18$) and BDI ($t(34) = 0.93$, $p = 0.36$).

Table 4. L2 proficiency and affective state measures of participants.

Parameters	All	Higher prof. group	Lower prof. group
LexTALE	63.44 (12.58)	70.69 (11.59)	56.18 (8.95)
C-test	63.06 (16.07)	76.22 (6.62)	49.89 (11.00)
EngProf	63.25 (12.94)	73.46 (7.54)	53.04 (8.20)
STAI	41.47 (10.20)	39.19 (9.71)	43.75 (10.44)
BDI-SF	4.44 (4.31)	3.78 (4.26)	5.11 (4.36)

Note: EngProf = English proficiency reflected by the average score of LexTALE and C-test.

2.5. Data processing and statistical analysis

The eye-tracking data were automatically cleaned. Short fixations (< 80 ms) were combined with nearby fixations within one character space. Extremely short (< 80 ms) and long (> 1000 ms) isolated fixations were removed from further analyses. Trials with a blink or pupil absence in the pre-target, target, or post-target (two-word region following the target word) regions were deleted. In addition, trials without any fixations were also deleted. In total, 3.04% of data on the target region and 2.90% of data on the post-target region were trimmed in the positive emotion block, while 2.44% of data on the target region and 2.34% of data on the post-target region were trimmed in the negative emotion block.

Following Knickerbocker et al. (2015, 2019), the current study analyzed a few early and late measures in the target and post-target regions. Early measures included: (1) first fixation duration (the duration of the first fixation on the target word); (2) single fixation duration (the duration of the first fixation on the target word when there was only one fixation on it); (3) gaze duration (or first-pass fixation duration, the total duration of all first-pass fixations on the target word before exiting it); (4) landing position (the number of characters from the horizontal position of the first fixation landing on the target word to its left edge); (5) skipping rate (the percentage of trials where the target word had no fixation during first-pass reading). Late measures included: (1) total dwell time (the total duration of all fixations on the target word); (2) regressions in (the percentage of trials where the participant returned to the target word from the regions to its right side); (3) second pass time (the total duration of all second-pass fixations on the target word). Besides the target words, five measures were also analyzed in the post-target regions, including (1) spillover (the duration of the first fixation after leaving the target word); (2) first fixation duration; (3) gaze duration; (4) total dwell time; (5) regressions out (the percentage of trials where regressions were made from the post-target regions to earlier interest areas before leaving the post-target regions in the forward direction).

Eye-tracking statistics were analyzed by R version 4.0.3 (R Core Team, 2020) in the linear mixed effect models (LMMs) for continuous dependent measures and generalized linear mixed effect models (GLMMs) for categorical metrics. These two models were respectively fitted using the *lmer()* function and the *glmer ()* function of the *lme4 1.1-26* package (Bates et al., 2015). The *p*-values and pairwise comparisons for statistical examinations were calculated by the *lmerTest 3.1-3* package (Kuznetsova et al., 2017) and the *lsmeans 2.30-0* package (Lenth, 2016). The remaining data were processed before fitting the statistical models. Metrics measuring fixation duration were lg-transformed to ensure a normal distribution, while other binomial metrics were analyzed by means of proportions. For each dependent variable, two contrast-coded factors and their interactions were coded as fixed effects, including the target word type (neutral: -0.5 vs. positive: +0.5 or neutral: -0.5 vs. negative: +0.5) and English proficiency (lower proficiency group: -0.5 vs. higher proficiency group: +0.5). Here, as English proficiency is fundamentally a continuous variable, it was also coded as a continuous dependent variable in additional models to reexamine the observed effects of L2 proficiency. As we also intended to investigate whether depression and anxiety states affected the emotion-related sentence reading, either BDI or STAI score was also coded as an additional fixed effect in two separate formulas for each dependent variable. Random effects were constructed with by-subject and

https://doi.org/10.1017/S1366728923000718 Published online by Cambridge University Press

by-item random intercepts. Starting with the full model, the backward model selection procedure followed the Akaike information criterion (AIC; Akaike, 1974) and significance tests (Matuschek et al., 2017). Models that reported convergence failure were excluded. To assess the construct of each model, the *r.squaredGLMM()* function of the *MuMIn 1.43.17* package (Bartoń, 2020; Nakagawa et al., 2017) was adopted to summarize the explanatory power of fixed effects ($R^2_{marginal}$) and the combined power of fixed and random effects ($R^2_{conditional}$). The partial eta-squared effect size was also intended for significant findings using the *eta_squared()* function of the *effectsize 0.6.0.1* package (Ben-Shachar et al., 2020). According to Cohen (1988), the benchmarks for small ($\eta^2 = 0.01$), medium ($\eta^2 = 0.06$) and large ($\eta^2 = 0.14$) effects were defined.

3. Results

3.1. Effects of presentation sequence on neutral word processing

Before conducting the formal statistical analyses, we were encouraged to examine the effects of presentation sequence on the processing of neutral words. In our current within-subject experiment design, the same set of neutral words was presented twice, but the emotional words only appeared once in the two separate blocks. Although possible familiar effects brought by the repeated presentation have been addressed by our counterbalance manipulation, there still existed an issue of averaging the eye-tracking data of neutral word processing obtained from the first and second presentations. Therefore, for each of the 13 eye-tracking measures in neutral word processing, we set the presentation sequence as the fixed effect with the by-subject and by-item random effects in LMMs or GLMMs models. As shown in Table 5, no significant differences were identified in any of our involved eye-tracking measures (all *ps* > 0.32). Results here indicated that no prominent differences existed in the neutral word processing between first-time and second-time presentations. In other words, the repeated presentation did not significantly influence the processing of neutral words in our current L2 sample. Hence, our following statistical analyses were modestly justified.

Table 5. Comparisons of neutral word processing between first and second presentations.

Measures	χ^2	*p*-value
Target region		
First fixation	0.38	0.54
Single fixation	0.00	0.97
First pass (gaze)	0.08	0.77
Landing position	0.19	0.66
Skipping rate	0.37	0.54
Total time	0.20	0.66
Regressions In	0.55	0.46
Second pass time	0.26	0.61
Post-target region		
Spillover	0.06	0.81
First fixation	0.02	0.89
First pass (gaze)	0.46	0.50
Total time	0.99	0.32
Regressions Out	0.01	0.91

3.2. Positive emotional words vs. neutral words

Overall, the comprehension accuracy for the positive block was acceptable for all participants ($M = 87.15\%$, $SD = 8.23\%$, range = 75% to 100%), indicating that they were attentive to the reading task. Table 6 presents the means and standard deviations of reading times concerning the eye movements at both target and post-target regions. Details of the LMM and GLMM models were provided in the supplementary data.

In the early stage of L2 sentence reading, an extra processing burden was observed for positive words and such a tendency was more obvious in the lower proficiency group. In general, when participants read positive emotional words, they showed a tendency to spend longer first-pass gaze duration ($\beta = 0.03$, $SE = 0.02$, $t = 1.58$, $\chi^2(1) = 2.50$, $p = 0.11$) and lower skipping rate ($\beta = -0.16$, $SE = 0.39$, $t = -0.42$, $\chi^2(1) = 0.00$, $p = 0.95$). The landing position ($\beta = -0.01$, $SE = 0.03$, $z = -0.39$, $\chi^2(1) = 0.16$, $p = 0.69$) was comparably similar in neutral and positive conditions. Nevertheless, there were no significant main effects of emotion in any of these early-stage measures. As for the comparison between the two proficiency groups, the higher proficiency group was overall faster to read the L2 words than the lower proficiency group, but the interaction between English proficiency and target word type reached significance only in the eye-tracking measure of landing position ($\beta = 0.07$, $SE = 0.03$, $t = 1.94$, $\chi^2(1) = 3.75$, $p = 0.05$, $R^2_{Marginal} = 0.006$, $R^2_{Conditional} = 0.102$). However, with an additional model where the L2 proficiency was coded as a continuous variable, the significant interaction in landing position disappeared ($\chi^2(1) = 0.27$, $p = 0.60$), suggesting that the observed L2 proficiency effects in the early processing stage were not robust enough.

With respect to the late measure, prolonged reading time was also observed for positive emotional words, and the burden of reading was more severe in the lower proficiency group. Participants spent longer total time ($\beta = 0.02$, $SE = 0.02$, $t = 0.90$, $\chi^2(1) = 0.82$, $p = 0.37$) and second pass time ($\beta = 0.01$, $SE = 0.02$, $t = 0.85$, $\chi^2(1) = 0.73$, $p = 0.39$) while reading positive words, and they had fewer regressions back to the target word region ($\beta = -0.05$, $SE = 0.14$, $z = -0.33$, $\chi^2(1) = 0.14$, $p = 0.71$) under this condition. The interaction between English proficiency and target word type was significant in the measure of total time ($\beta = 0.04$, $SE = 0.02$, $t = 2.06$, $\chi^2(1) = 4.25$, $p = 0.04$, $R^2_{Marginal} = 0.033$, $R^2_{Conditional} = 0.377$). Post-hoc analyses showed that compared to L2 readers in the lower proficiency group, those in the higher proficiency group spent significantly less total time reading positive words ($\beta = -0.12$, $SE = 0.05$, $t = -2.38$, $p = 0.02$), but their reading time for neutral words were comparable ($\beta = -0.08$, $SE = 0.05$, $t = -1.64$, $p = 0.11$). When coding the L2 proficiency as a continuous variable, this significant interaction was also identified ($\beta = 0.001$, $SE = 0.001$, $t = 2.02$, $\chi^2(1) = 4.09$, $p = 0.04$, $R^2_{Marginal} = 0.045$, $R^2_{Conditional} = 0.377$). As such, it could be cautiously inferred that higher L2 proficiency might be associated with stronger emotion processing ease for positive words at the late processing stage than less proficient bilinguals (Fig. 2).

In the post-target region, neither the main effect of emotion nor the interaction effects between English proficiency and emotion reached significance. Interestingly, the measure of spillover

https://doi.org/10.1017/S1366728923000718 Published online by Cambridge University Press

Table 6. Raw eye-tracking data for the positive block.

	Overall		Higher prof. group		Lower prof. group	
Measures	Neutral	Positive	Neutral	Positive	Neutral	Positive
Target region						
First pass measures						
First fixation	289.75 (54.12)	293.61 (49.15)	273.65 (40.47)	277.94 (42.00)	305.86 (61.99)	309.27 (51.86)
Single fixation	318.67 (67.10)	321.38 (68.03)	299.98 (57.84)	296.17 (42.40)	337.36 (72.00)	346.59 (79.90)
First pass (gaze)	420.34 (95.52)	461.02 (115.78)	383.17 (72.77)	409.32 (70.14)	457.51 (102.78)	512.73 (130.45)
Landing position	1.80 (0.12)	1.80 (0.12)	1.83 (0.09)	1.84 (0.10)	1.78 (0.14)	1.75 (0.12)
Skipping rate	3.4% (4.1%)	3.3% (3.1%)	4.3% (4.6%)	3.9% (3.2%)	2.5% (3.4%)	2.8% (3.0%)
Late measures						
Total time	840.47 (306.50)	899.46 (356.43)	745.26 (247.08)	767.22 (269.07)	935.68 (336.40)	1031.71 (390.03)
Regressions In	40.5% (18.4%)	39.4% (20.7%)	40.4% (15.2%)	38.0% (16.7%)	40.6% (21.6%)	40.7% (24.5%)
Second pass time	352.15 (76.36)	359.56 (79.58)	313.42 (39.59)	324.39 (55.28)	390.33 (85.17)	394.74 (85.81)
Post-target region						
Spillover	282.17 (51.65)	281.16 (60.29)	266.83 (36.73)	261.57 (29.94)	297.50 (60.38)	300.75 (75.98)
First fixation	275.95 (45.03)	278.42 (50.73)	257.55 (32.10)	263.07 (29.05)	294.35 (49.27)	293.76 (62.90)
First pass (gaze)	598.33 (106.74)	600.86 (129.90)	555.96 (78.99)	551.19 (89.70)	640.71 (115.82)	650.53 (146.52)
Total time	1064.74 (345.77)	1017.87 (316.81)	974.12 (255.67)	924.30 (262.18)	1155.35 (404.21)	1111.45 (345.49)
Regressions Out	44.4% (19.6%)	41.9% (17.6%)	46.0% (18.1%)	40.7% (15.1%)	42.9% (21.5%)	43.1% (20.2%)

Note. Reading times are in milliseconds.

did not manifest a significant interaction effect in the categorical model ($\chi^2(1) = 1.30$, $p = 0.26$), but this interaction effect has reached significance in the continuous model ($\beta = 0.001$, $SE = 0.001$, $t = 1.98$, $\chi^2(1) = 3.90$, $p = 0.05$, $R^2_{Marginal} = 0.014$, $R^2_{Conditional} = 0.185$).

In separate runs of the models examining the effects of individuals' emotional states on L2 reading, BDI scores did not show significant main effects (all $ps > 0.24$) or interaction effects (all $ps > 0.15$) in any of the eye movement measures. As for the STAI scores, although the main effects were not significant for all eye-tracking measures (all $ps > 0.28$), significant interactions between STAI scores and emotion were recognized in the early measures of positive word processing, including the first fixation ($\chi^2(1) = 5.89$, $p = 0.02$, $R^2_{Marginal} = 0.018$, $R^2_{Conditional} = 0.136$) and

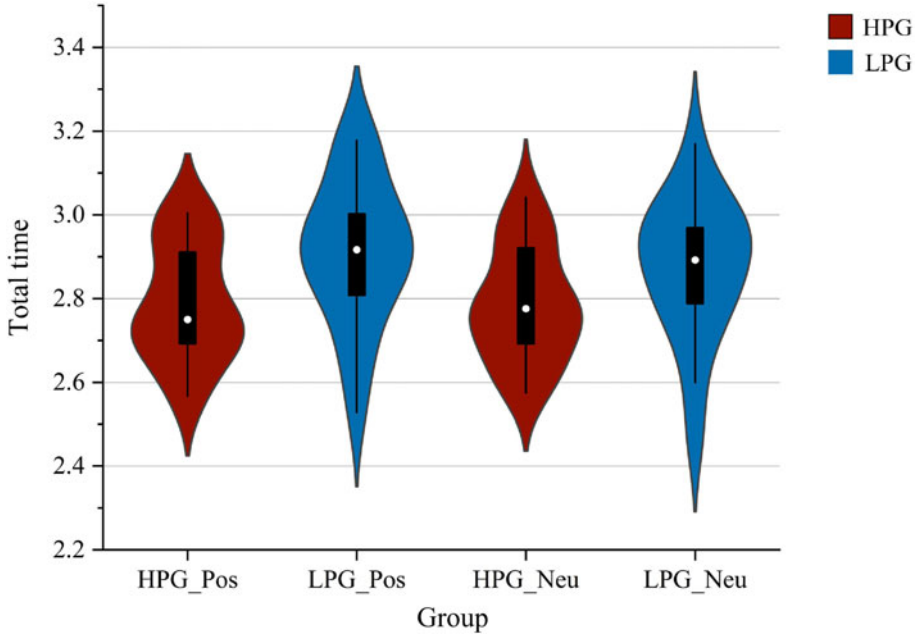


Figure 2. Significant interaction effects of total reading time for positive words. (HPG: higher proficiency group; LPG: lower proficiency group; Pos: positive word reading; Neu: neutral word reading. Reading time is lg-transformed.)

https://doi.org/10.1017/S1366728923000718 Published online by Cambridge University Press

single fixation ($\chi^2(1) = 3.92$, $p = 0.05$, $R^2_{Marginal} = 0.021$, $R^2_{Conditional} = 0.211$). Hence, the anxious mood, but not the depressive states, was correlated with the L2 processing of positive emotional words in the early stage.

In summary, L2 readers showed no significant differences in eye movements when reading the neutral and positive words. However, bilinguals with higher L2 proficiency showed significantly stronger emotion processing effects for L2 positive words at the late processing stage, as reflected by the robust interaction effects of total reading time. Other eye-tracking measures, such as landing position and spillover, might also display significant interaction effects, but cautions were needed for relevant interpretations due to the low robustness. Additionally, only the measures of first fixation and single fixation demonstrated significant interaction effects of STAI scores in the positive block.

3.3. Negative emotional words vs. neutral words

Overall, the comprehension accuracy for the negative block was acceptable for all participants ($M = 86.92\%$, $SD = 6.84\%$, range = 75% to 100%), suggesting that they read the sentences carefully and followed our instructions. Table 7 presents the means and standard deviations of reading times. Details of the LMM and GLMM models were provided in the supplementary data.

Of all early-stage measures, the reading time of negative emotional words was generally longer than that of neutral words. When participants read negative emotional words, they tended to spend longer first fixation duration ($\beta = 0.02$, $SE = 0.01$, $t = 1.93$, $\chi^2(1) = 3.68$, $p = 0.06$), single fixation duration ($\beta = 0.01$, $SE = 0.01$, $t = 1.16$, $\chi^2(1) = 1.36$, $p = 0.24$), first-pass gaze duration ($\beta = 0.02$, $SE = 0.02$, $t = 0.87$, $\chi^2(1) = 0.76$, $p = 0.38$). On the other hand, the landing position ($\beta = 0.02$, $SE = 0.03$, $t = 0.48$, $\chi^2(1) = 0.23$, $p = 0.63$) and skipping rate ($\beta = 0.03$, $SE = 0.29$, $z = 0.11$, $\chi^2(1) = 0.06$, $p = 0.80$) were also higher for the negative words than the neutral words. As for the proficiency effects on L2 emotional reading, a significant interaction effect of single fixation was identified ($\beta = 0.04$, $SE = 0.02$, $t = 2.04$, $\chi^2(1) = 4.17$, $p = 0.04$, $R^2_{Marginal} = 0.008$, $R^2_{Conditional} = 0.166$). However, this interaction effect did not reach a significant magnitude in the models where L2 proficiency was coded as a continuous variable ($\chi^2(1) = 2.52$, $p = 0.11$), suggesting that the proficiency effects identified in the categorical models might not be robust.

As for late measures, L2 participants also spent longer reading time for negative words than neutral words. In specific, the total time ($\beta = 0.01$, $SE = 0.03$, $t = 0.53$, $\chi^2(1) = 0.28$, $p = 0.59$), second pass time ($\beta = 0.00$, $SE = 0.02$, $t = 0.18$, $\chi^2(1) = 0.03$, $p = 0.85$), and regressions into the target regions ($\beta = 0.02$, $SE = 0.16$, $z = 0.10$, $\chi^2(1) = 0.02$, $p = 0.90$) were all slightly higher for negative words. No significant interactions between L2 proficiency and emotion were found in any of the eye-tracking measures at the late stage of negative word processing.

In the post-target region, significant main effects of emotion were only observed in the first fixation measure, while no other significant findings were identified. When reading the post-region words following the negative targets, L2 participants spent significantly longer first fixation duration ($\beta = 0.02$, $SE = 0.01$, $t = 2.55$, $\chi^2(1) = 6.50$, $p = 0.01$, $R^2_{Marginal} = 0.009$, $R^2_{Conditional} = 0.114$). The emotion effects on spillover also reached marginal significance ($\beta = 0.02$, $SE = 0.01$, $t = 1.83$, $\chi^2(1) = 3.32$, $p = 0.07$, $R^2_{Marginal} = 0.010$, $R^2_{Conditional} = 0.149$). These two eye-tracking measures seemed to suggest that the early processing of what closely follows negative words might require a longer time in the L2 population.

Table 7. Raw eye-tracking data for the negative block.

	Overall		Higher prof. group		Lower prof. group	
Measures	Neutral	Negative	Neutral	Negative	Neutral	Negative
Target region						
First pass measures						
First fixation	285.95 (43.86)	297.37 (46.23)	275.08 (35.79)	294.02 (44.14)	296.83 (49.29)	300.72 (49.28)
Single fixation	313.64 (61.33)	320.63 (64.68)	295.75 (45.63)	320.35 (68.90)	331.53 (70.59)	320.92 (62.18)
First pass (gaze)	423.44 (91.15)	446.20 (104.35)	389.56 (74.46)	413.68 (84.31)	457.33 (95.54)	478.72 (114.33)
Landing position	1.78 (0.15)	1.81 (0.16)	1.84 (0.12)	1.87 (0.10)	1.73 (0.16)	1.74 (0.18)
Skipping rate	4.9% (6.6%)	5.5% (4.3%)	4.3% (5.4%)	5.6% (3.7%)	5.6% (7.7%)	5.4% (5.0%)
Late measures						
Total time	842.41 (372.07)	876.94 (398.55)	746.36 (314.13)	756.82 (310.26)	938.47 (408.42)	997.07 (447.44)
Regressions In	37.2% (19.8%)	37.4% (19.4%)	35.2% (14.1%)	34.3% (12.2%)	39.2% (24.4%)	40.6% (24.5%)
Second pass time	352.95 (85.12)	372.73 (101.70)	316.04 (76.78)	334.82 (70.36)	392.03 (77.32)	410.65 (115.33)
Post-target region						
Spillover	255.93 (36.11)	274.83 (51.47)	251.71 (34.18)	261.95 (36.45)	260.16 (38.45)	287.71 (61.43)
First fixation	262.90 (33.81)	275.39 (44.65)	253.85 (26.65)	264.92 (30.93)	271.95 (38.33)	285.85 (54.00)
First pass (gaze)	510.25 (98.35)	520.53 (104.58)	471.45 (89.58)	487.46 (85.41)	549.05 (93.28)	553.60 (113.61)
Total time	870.52 (351.91)	859.60 (311.00)	795.15 (350.17)	773.75 (264.92)	945.89 (346.87)	945.46 (336.66)
Regressions Out	41.9% (19.6%)	40.1% (20.3%)	39.7% (15.4%)	36.6% (13.6%)	44.1% (23.2%)	43.7% (25.3%)

Note. Reading times are in milliseconds.

https://doi.org/10.1017/S1366728923000718 Published online by Cambridge University Press

The English proficiency, however, was not predictive of any emotion processing effects here.

Two separate models were operated to examine the effects of participants' emotional states, as reflected by the BDI and STAI scores. We observed no significant main effects of BDI scores (all $ps > 0.15$) or STAI scores (all $ps > 0.06$) in any of the eye movement measures. As for the interaction effects, only one significant interaction effect was identified between BDI scores and emotion in the late measure of regressions out ($\chi^2(1) = 4.08$, $p = 0.04$, $R^2_{Marginal} = 0.022$, $R^2_{Conditional} = 0.265$), whereas all other interaction effects of BDI were not significant (all $ps > 0.18$). No significant interaction effects were recognized for STAI scores (all $ps > 0.23$).

In summary, the emotion processing effects of negative words shortly occurred in the early processing stage at the post-target region among L2 readers. However, as all eye-tracking measures at the target region and most measures at the post-target region did not demonstrate significant main effects of emotion, the emotion processing effects might not truly exist for negative words. As for the interaction effects between L2 proficiency and emotion, only the single fixation measure manifested a significant result when dividing all participants into two proficiency groups. Additionally, only one interaction effect between BDI scores and emotion was found for the measure of regressions out. However, these interaction effects need to be interpreted with caution due to the low robustness.

4. Discussion

The present study investigated whether Chinese–English bilinguals benefit from the emotional information while processing L2 emotion-label words, and measured participants' eye movements at the early and late processing stages in the passive sentence reading tasks. Results suggested that emotion processing effects in L1 did not occur in our L2 population, and these emotional words even possibly exerted burdens on L2 processing, as reflected by the overall longer reading time for both positive and negative emotional words. The interaction effects between L2 proficiency and emotion were stably significant for positive words at the late processing stage. Moreover, since significant interaction effects of depressive and anxious states were only identified in few eye-tracking measures, they might not robustly correlate with eye movement measures in L2 emotional word processing.

Emotional words have been claimed to facilitate lexical processing and sentence reading in L1, but such an emotional advantage was not found in our L2 sample. In fact, L2 readers in our study even showed a reversed tendency of longer reading time for L2 emotional word processing, indicating that there might be disadvantages in this process. Similar to our eye movement measures, a few studies have presented bilinguals with emotional and neutral words in their L1 and L2 conditions, and recorded their pupillary responses. These experiments reported less pronounced contrasts of pupil size in the L2 condition only (Toivo & Scheepers, 2019), and observed similar weaker pupil responses for emotional words in L2 even across two cognate languages (Yao et al., 2023), thus reflecting reduced emotional responses for L2 words. Some studies using other physiological measures have also reported similar findings. For instance, bilinguals showed significantly decreased activities of the corrugator muscle for L2 emotional words (Baumeister et al., 2017), and their skin conductance levels did not differ significantly between emotionally charged and neutral words (Eilola & Havelka, 2010). Our eye movement data supplemented existing evidence and demonstrated that bilinguals might not be sensitive to the emotional messages in L2 so that they read the L2 emotional and neutral words in similar ways.

One possible explanation for the missing of emotion processing effects in L2 reading is that bilinguals had less emotional engagement in the L2 context (Harris et al., 2006), with longer emotional distance and reduced emotional resonance (Costa et al., 2014; Degner et al., 2012) in L2 processing. Several factors, such as age of acquisition, context of acquisition, language dominance, L2 proficiency, and cross-language emotional memory (Caldwell-Harris, 2015; Degner et al., 2012; Yao et al., 2023), have been argued to influence the formation of emotionality in L2. In support of the Emotional Contexts of Learning Theory (Harris et al., 2006), late bilinguals usually acquire their second language through formal courses in the classroom, and they may not be able to integrate the L2 lexical meanings with emotionally relevant autobiographical experiences and everyday interactions (Pavlenko, 2017). Due to the lack of sensory involvement, emotional connotations of L2 words and the ensuing sensorimotor activation are not as strong as their first languages (Dudschig et al., 2014; Eilola & Havelka, 2010). Affective rating experiments have proved that bilinguals were inclined to evaluate L2 emotional words less intensely (Ferré et al., 2022; Imbault et al., 2021), and they subjectively confessed to feeling less emotional and natural in L2 despite high proficiency (Dewaele & Nakano, 2013). In other words, late bilinguals may clearly know the emotional meaning of L2 words, but they either find it difficult to feel the embedded emotion in person or only perceive the L2 emotional words as insincere and artificial (Pavlenko, 2012). As a result, their reading behaviours for emotional and neutral words did not manifest significant differences overall, and the facilitation effects of emotional connotations were thus not successfully allowed.

Nevertheless, our eye-tracking statistics from the L2 readers manifested longer reading time for emotional words, which seemed to indicate disadvantages for L2 emotion processing rather than merely an attenuation of the advantageous effects in L1. The overall longer reading time in the L2 population might be caused by the extra cognitive load required for L2 processing. For bilinguals, the limited attentional and cognitive resources are preferentially occupied by prioritized orthographic, semantic, grammatical and phonetic processing in L2 (Franconeri et al., 2013; Hinojosa et al., 2010; Jiang, 2021; Lieder & Griffiths, 2019), so the additional processing of emotional connotations may lead to cognitive overload and working memory burden. Neuroimaging evidence has also corroborated that L2 brains favoured semantic processing over emotion decoding and more cognitive resources were needed to process the lexical-semantic components of L2 words (L. Liu et al., 2021), which hindered or at least attenuated the access to the emotional aspects of L2 words and led the bilinguals to overlook emotional information (L. Liu et al., 2022). As a result, when our bilingual participants read the target words and sentences in the cognitively demanding L2 context, they had to use longer time to comprehend the semantic meanings, and spent even longer duration decoding the emotional connotations associated with the emotional words in a suppressed or delayed manner.

The valence effects were also observed in our current study. Previous studies have mostly revealed processing disadvantages for negative words and processing advantages for positive

https://doi.org/10.1017/S1366728923000718 Published online by Cambridge University Press

words, compared to neutral words (Barriga-Paulino et al., 2022; Bromberek-Dyzman et al., 2021; Unkelbach et al., 2010). In our study, similar disadvantages in negative word processing have also been reflected by bilinguals' eye movements in the L2 context. This could be explained by the Automatic Vigilance Theory (Estes & Adelman, 2008; Estes & Verges, 2008; Pratto & John, 1991), which ascribes such effects to the increased engagement and delayed disengagement of attention for negative stimuli. According to this theory, external inputs are automatically evaluated as positive or negative in the first place, and then extended attention and cognitive resources would be allocated to negative stimuli due to their danger and fatality indications, so that relevant responses to the negative stimuli would be prolonged compared to the neutral stimuli. A more pertinent interpretation for the negative disadvantages in our L2 sample might be the suppressed and reevaluated processing of L2 negative stimuli. Jończyk et al. (2016) observed reduced N400 amplitudes when bilinguals read L2 sentences ending with negative words only. It has thus been proposed that the semantic access to L2 negative information in bilinguals was incomplete and partially suppressed at the early processing stage, which might further trigger stronger reevaluation in later processing stages, as reflected by the greater amplitudes in late positive complex (LPC) range (Y. J. Wu & Thierry, 2012). Because of the early suppression and reevaluated processing, the reading time for L2 negative words was reasonably longer than the neutral words.

However, the disadvantages of positive word processing in our study were not expected, which seemed to be inconsistent with previous evidence (Sheikh & Titone, 2013, 2016). In this case, the effects of the arousal dimension, which have been found to function on the word recognition duration independently from the valence effects (Kever et al., 2019), might have influenced the processing of emotional words in our study. Kuperman et al. (2014) tried to determine the precise nature of the effects of valence and arousal on word recognition by analyzing 12,658 words, and they observed that words with higher arousal were recognized more slowly than calming (low-arousal) words. As illustrated in Fig. 1, our included positive words were perceived as the most arousing category in the L2 population, compared to the neutral and negative words. Therefore, the higher arousal traits of our positive words might elicit extra difficulties in word recognition, which further caused the processing disadvantages of longer reading time. If the valence-arousal interaction is taken into further consideration, our included lexical materials seem to have caused extra cognitive burdens in L2 processing. In an extended version of the Approach-Withdrawal Theory (Robinson et al., 2004), it is proposed that negative high-arousal stimuli suggest potential threat (withdrawal) while positive low-arousal stimuli represent safety (approach). In our study, however, the experimental materials instead encompassed positive high-arousal and negative low-arousal words. According to the Valence-Arousal Conflict Theory (Robinson et al., 2004), there exists a tendency that positive high-arousal and negative low-arousal stimuli are more difficult to process because of their breach of the traditional approach-withdrawal relationships. Therefore, due to the mismatched valence-arousal relationships of our emotional words, these positive words might have imposed more processing difficulties at both the pre-attentive level and the subsequent response stage, and the reading time for these words was thus prolonged. Another possible reason underlying this processing disadvantage might be the individual psychological states of our participants. Recent studies have claimed that reading behaviours for positive words were highly susceptible to individual differences, and found that individuals with higher levels of need for affect would read positive words more slowly (Lei et al., 2023). However, since the current study only collected the BDI and STAI scores as reflections of participants' emotional states, it could not be determined whether other factors related with individual differences (e.g., need for affect) have modulated the eye movements during the L2 reading.

The proficiency of L2 was also coded as an essential factor for our subgroup analyses, which distinguished different reading behaviours between the higher proficiency group and the lower proficiency group, especially for positive words. Based on the Revised Hierarchical Model and Lexical Quality Hypothesis, these two proficiency groups might have different semantic representations of L2 words. Hence, it was hypothesized in our current study that bilinguals with different L2 proficiency could show different eye movement patterns while processing the L2 words with various emotional connotations. Indeed, we have found one stable significant proficiency-emotion interaction effect of the total reading time at the late processing stage of positive words. As for other eye-tracking measures, we also identified some unstable interaction effects, which reached significance in only one of the models that coded L2 proficiency as either a categorical or a continuous variable. These measures included the early landing position and the late spillover metrics in the positive block, and the single fixation of target words in the negative block. Our findings suggested that the effect of proficiency, reflected by the faster reading speed of L2 readers with higher proficiency, was larger for positive words than neutral words in the late processing stage. In fact, semantic and emotional processing in L2 may take place via separate channels. Compared to the less proficient bilinguals, highly proficient bilinguals can automatically activate the L2 affective connotations (Degner et al., 2012) with a more integrated word processing procedure (Sianipar et al., 2015), and they tend to show prioritized emotion perception during L2 processing (Ponari et al., 2015). As for the lower proficiency bilinguals, they seem to have extra demanding cognitive costs (Hasegawa et al., 2002; Stiller & Schworm, 2019), which would further influence the emotional access (L. Liu et al., 2021) and lead to the down-regulation of emotional circuits (L. Liu et al., 2022; Van Dillen et al., 2013). However, a prior eye-tracking experiment (Sheikh & Titone, 2016) found that L2 proficiency was only predictive of the concreteness advantage in L2 processing, rather than the emotional advantage. As such, since the proficiency effects on L2 emotion processing have not yet been sufficiently investigated, our findings on the proficiency effects need to be interpreted with caution and still require further examination.

We have also investigated the associations between depressive or anxious symptoms and the reading behaviours for L2 emotional words. However, only a few measures were significantly correlated to our bilinguals' affective states, indicating that these findings might not be stable and robust enough to be generalized to a larger population. Previous evidence has displayed a facilitated processing of negative target words under a more anxious mood (Knickerbocker et al., 2015). Our statistics, on the other hand, manifested processing ease for positive words as a function of higher anxiety level, but also at the early processing course. In addition to the anxious states, we also found that trait depression was associated with the percentage of regressions out in the post-target regions for negative words. Recent research mostly inspected the relationship between anxiety and negative emotion

https://doi.org/10.1017/S1366728923000718 Published online by Cambridge University Press

processing, finding that anxiety symptoms were associated with higher avoidant attention allocation and greater pupil dilation (Shechner et al., 2017). Nevertheless, negative moods in general, such as anxiety and depression, were found to promote detail-oriented L2 lexical processing and temporarily suppress the complete L2 semantic integration (Naranowicz et al., 2022), so the seemingly different correlations here might just be representations of the same effect in two emotional word categories. After all, investigation of the influence of depression and anxiety symptoms on the emotional access of words and sentence reading, especially in the bilingual context, is still scant and remains open for discussion, so that the findings reported here also warrant further examination.

Several limitations are acknowledged in the present study. First, there might exist possible influences of the current within-subject design on the interpretation of our findings. For instance, certain unexpected effects might be introduced by the repeated reading of the neutral sentences. To address this problem, we have tried to minimize such effects by the counterbalance manipulation, and revealed the non-significant effects of presentation sequence on neutral word processing. As a result, the statistical analyses of the current study were credible. Future studies could adopt more sophisticated experiment materials and designs in the exploration of L2 emotion processing. Second, the affective perception of experimental materials showed some differences between the L1 and L2 populations. For instance, compared to the positive words, negative words were perceived as more arousing in the L1 readers but less arousing in the L2 readers. Although these differences did not influence the fact that both positive and negative words could be clearly distinguished from neutral words in both populations, they might still modulate our current eye-tracking observations and elicit more caution for the interpretation of our results. Third, we only enrolled late bilinguals with similar language dominance, and more diversified L2 populations need to be included to examine the emotion processing effects in bilingual and multilingual contexts. Previous research has indicated that the construction and development of semantic representations were largely different in early and late bilinguals (Gathercole & Moawad, 2010), whose language networks also manifested different patterns (X. Liu et al., 2017). It can be the case that two types of bilinguals may show different activation and responsive patterns when reading the same set of emotional words in the L2 context. Last, the emotional density of emotion-label and emotion-laden words can be different in bilinguals' lexicon (Pavlenko, 2008), and it is still unclear whether these two emotion-related word types cause identical or discrepant physiological responses in L2 reading. According to previous behavioural and neuroimaging studies, the processing of emotion-laden words required longer response time and higher cognitive efforts than emotion-label words (D. Tang et al., 2023; C. Wu et al., 2022). In our current study, we did not identify significant emotion processing effects by using the emotion-label words, so it might be predicted that such facilitation effects caused by emotional connotations might not occur in the reading of L2 emotion-laden words. Probably, due to the indirect emotional connections of emotion-laden words and their associations with various personal experiences, the processing of these words might display more prominent disadvantages and larger individual variation compared to emotion-label words. Hence, future studies that analyze the eye movements in L2 reading of emotion-laden words are needed, which can help illustrate the L2 emotion processing effects more comprehensively.

5. Conclusions

To conclude, the facilitation effects of affective connotations in emotional words did not exist in the L2 context. Instead, processing disadvantages for emotional words were observed in our L2 sample. Our findings implied that bilinguals activate fewer emotional connotations when reading non-dominant L2 words in natural sentences. Moreover, readers with higher L2 proficiency showed greater ease in processing positive words at the late stage compared to those with lower L2 proficiency. As only a few interaction effects of individuals' emotional states were identified in our study, it appeared that the depressive and anxious states of bilinguals might not robustly influence their processing of L2 emotional words. The current study provides new evidence to a limited number of eye-tracking experiments on L2 emotion processing, and supports several theoretical frameworks in the field of bilingual language processing.

Supplementary Material. For supplementary material accompanying this paper, visit https://doi.org/10.1017/S1366728923000718

Conflict of interest. We have no known conflict of interest to disclose.

Acknowledgements. This work was supported by grants from the Major Program of National Social Science Foundation of China (Grant number: No. 18ZDA293).

Competing interests. The authors declare none.

Data availability. The essential data of this research are accessible by the electronic supplementary file.

References

Ahn, S., & Jiang, N. (2022). Can adult learners sense L2 emotional words automatically? The role of L2 use on the emotional Stroop effect. *Second Language Research*, *0*(0). https://doi.org/10.1177/02676583221131256

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Altarriba, J., & Basnight-Brown, D. M. (2011). The representation of emotion vs. emotion-laden words in English and Spanish in the Affective Simon Task. *International Journal of Bilingualism*, *15*(3), 310–328. https://doi.org/10.1177/1367006910379261

Arriagada-Mödinger, F., & Ferreira, R. A. (2022). The effect of emotional valence on auditory word recognition memory in English as a foreign language. *Journal of Psycholinguistic Research*, *51*(2), 309–322. https://doi.org/10.1007/s10936-022-09841-3

Ayçiçeği-Dinn, A., & Caldwell-Harris, C. L. (2009). Emotion-memory effects in bilingual speakers: A levels-of-processing approach. *Bilingualism: Language and Cognition* *12*(3), 291–303. https://doi.org/10.1017/S1366728909990125

Babaii, E., & Ansary, H. (2001). The C-test: A valid operationalization of reduced redundancy principle? *System*, *29*(2), 209–219. https://doi.org/10.1016/S0346-251X(01)00012-4

Barriga-Paulino, C. I., Guerreiro, M., Faísca, L., & Reis, A. (2022). Does emotional valence modulate word recognition? A behavioral study manipulating frequency and arousal. *Acta Psychologica*, *223*, 103484. https://doi.org/10.1016/j.actpsy.2021.103484

Bartoń, K. (2020). *MuMIn: Multi-Model Inference*. (R package version 1.43.17.). https://cran.r-project.org/package=MuMIn.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

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

Beck, A. T., & Steer, R. A. (1993). Beck Depression Inventory Manual. Psychological Corporation, San Antonio.

Ben-Shachar, M. S., Lüdecke, D., & Makowski, D. (2020). Effectsize: Estimation of effect size indices and standardized parameters. *The Journal of Open Source Software*, *5*, 2815. https://doi.org/10.21105/joss.02815

Blanchette, I., & Richards, A. (2003). Anxiety and the interpretation of ambiguous information: Beyond the emotion-congruent effect. *Journal of Experimental Psychology: General*, *132*(2), 294–309. https://doi.org/10.1037/0096-3445.132.2.294

Bromberek-Dyzman, K., Jończyk, R., Vasileanu, M., Niculescu-Gorpin, A.-G., & Bąk, H. (2021). Cross-linguistic differences affect emotion and emotion-laden word processing: Evidence from Polish–English and Romanian–English bilinguals. *International Journal of Bilingualism*, *25*(5), 1161–1182. https://doi.org/10.1177/1367006920987306

Brothers, T., Hoversten, L. J., & Traxler, M. J. (2021). Bilinguals on the garden-path: Individual differences in syntactic ambiguity resolution. *Bilingualism: Language and Cognition*, *24*(4), 612–627. https://doi.org/10.1017/s1366728920000711

Caldwell-Harris, C. L. (2015). Emotionality differences between a native and foreign language: Implications for everyday life. *Current Directions in Psychological Science*, *24*(3), 214–219. https://doi.org/10.1177/0963721414566268

Caldwell-Harris, C. L., Tong, J., Lung, W., & Poo, S. (2011). Physiological reactivity to emotional phrases in Mandarin-English bilinguals. *International Journal of Bilingualism*, *15*(3), 329–352. https://doi.org/10.1177/1367006910379262

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New York, NY: Routledge Academic.

Costa, A., Foucart, A., Arnon, I., Aparici, M., & Apesteguia, J. (2014). "Piensa" twice: On the foreign language effect in decision making. *Cognition*, *142*, 236–254. https://doi.org/10.1016/j.cognition.2015.05.013

Daller, M., Müller, A., & Wang-Taylor, Y. (2020). The C-test as predictor of the academic success of international students. *International Journal of Bilingual Education and Bilingualism*, *24*(10), 1502–1511. https://doi.org/10.1080/13670050.2020.1747975

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*(1), 181–189. https://doi.org/10.1017/S1366728911000095

Dewaele, J.-M., & Nakano, S. (2013). Multilinguals' perceptions of feeling different when switching languages. *Journal of Multilingual and Multicultural Development*, *34*(2), 107–120. https://doi.org/10.1080/01434632.2012.712133

Dörnyei, Z., & Katona, L. (1992). Validation of the C-test amongst Hungarian EFL learners. *Language Testing*, *9*(2), 187–206. https://doi.org/10.1177/0265532292009002

Dudschig, C., de la Vega, I., & Kaup, B. (2014). Embodiment and second-language: Automatic activation of motor responses during processing spatially associated L2 words and emotion L2 words in a vertical Stroop paradigm. *Brain and Language*, *132*, 14–21. https://doi.org/10.1016/j.bandl.2014.02.002

Eilola, T. M., & Havelka, J. (2010). Behavioural and physiological responses to the emotional and taboo Stroop tasks in native and non-native speakers of English. *International Journal of Bilingualism*, *15*(3), 353–369. https://doi.org/10.1177/1367006910379263

Estes, Z., & Adelman, J. S. (2008). Automatic vigilance for negative words is categorical and general. *Emotion*, *8*(4), 453–457. https://doi.org/10.1037/a0012887

Estes, Z., & Verges, M. (2008). Freeze or flee? Negative stimuli elicit selective responding. *Cognition*, *108*, 557–565. https://doi.org/10.1016/j.cognition.2008.03.003

Ferré, P., Anglada-Tort, M., & Guasch, M. (2018). Processing of emotional words in bilinguals: Testing the effects of word concreteness, task type and language status. *Second Language Research*, *34*(3), 371–394. https://doi.org/10.1177/0267658317744008

Ferré, P., Guasch, M., Stadthagen-Gonzalez, H., & Comesaña, M. (2022). Love me in L1, but hate me in L2: How native speakers and bilinguals rate the affectivity of words when feeling or thinking about them. *Bilingualism: Language and Cognition*, *25*(5), 786–800. https://doi.org/10.1017/s1366728922000189

Franconeri, S. L., Alvarez, G. A., & Cavanagh, P. (2013). Flexible cognitive resources: Competitive content maps for attention and memory. *Trends in Cognitive Sciences*, *17*, 134–141. https://doi.org/10.1016/j.tics.2013.01.010

Fricke, M., Zirnstein, M., Navarro-Torres, C., & Kroll, J. F. (2019). Bilingualism reveals fundamental variation in language processing. *Bilingualism: Language and Cognition*, *22*(1), 200–207. https://doi.org/10.1017/S1366728918000482

Furlanetto, L. M., Mendlowicz, M. V., & Romildo Bueno, J. (2005). The validity of the Beck Depression Inventory-Short Form as a screening and diagnostic instrument for moderate and severe depression in medical inpatients. *Journal of Affective Disorders*, *86*(1), 87–91. https://doi.org/10.1016/j.jad.2004.12.011

Garrido, M. V., & Prada, M. (2021). Comparing the valence, emotionality and subjective familiarity of words in a first and a second language. *International Journal of Bilingual Education and Bilingualism*, *24*(2), 275–291. https://doi.org/10.1080/13670050.2018.1456514

Gathercole, V. C. M., & Moawad, R. A. (2010). Semantic interaction in early and late bilinguals: All words are not created equally. *Bilingualism: Language and Cognition*, *13*(4), 385–408. https://doi.org/10.1017/s1366728909990460

Guo, J., Zou, T., & Peng, D. (2018). Dynamic influence of emotional states on novel word learning. *Frontiers in Psychology*, *9*, 537. https://doi.org/10.3389/fpsyg.2018.00537

Haro, J., Calvillo, R., Poch, C., Hinojosa, J. A., & Ferré, P. (2023). Your words went straight to my heart: The role of emotional prototypicality in the recognition of emotion-label words. *Psychological Research*, *87*, 1075–1084. https://doi.org/10.1007/s00426-022-01723-6

Harris, C., Gleason, J., & 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). Bristol, Blue Ridge Summit: Multilingual Matters.

Harsch, C., & Hartig, J. (2015). Comparing C-tests and Yes/No vocabulary size tests as predictors of receptive language skills. *Language Testing*, *33*(4), 555–575. https://doi.org/10.1177/0265532215594642

Hasegawa, M., Carpenter, P. A., & Just, M. A. (2002). An fMRI study of bilingual sentence comprehension and workload. *NeuroImage*, *15*, 647–660. https://doi.org/10.1006/nimg.2001.1001

Hinojosa, J. A., Méndez-Bértolo, C., & Pozo, M. A. (2010). Looking at emotional words is not the same as reading emotional words: Behavioral and neural correlates. *Psychophysiology*, *47*(4), 748–757. https://doi.org/10.1111/j.1469-8986.2010.00982.x

Imbault, C., Titone, D., Warriner, A. B., & Kuperman, V. (2021). How are words felt in a second language: Norms for 2,628 English words for valence and arousal by L2 speakers. *Bilingualism: Language and Cognition*, *24*(2), 281–292. https://doi.org/10.1017/s1366728920000474

Imbir, K. K., Jurkiewicz, G., Duda-Goławska, J., & Żygierewicz, J. (2019). The role of valence and origin of emotions in emotional categorization task for words. *Journal of Neurolinguistics*, *52*, 100854. https://doi.org/10.1016/j.jneuroling.2019.100854

Jackson, J. C., Watts, J., Henry, T. R., List, J. M., Forkel, R., Mucha, P. J., Greenhill, S. J., Gray, R. D., & Lindquist, K. A. (2019). Emotion semantics show both cultural variation and universal structure. *Science*, *366*, 1517–1522. https://doi.org/10.1126/science.aaw8160

Jankowiak, K., & Korpal, P. (2018). On modality effects in bilingual emotional language processing: Evidence from galvanic skin response. *Journal of Psycholinguistic Research*, *47*(3), 663–677. https://doi.org/10.1007/s10936-017-9552-5

Jiang, N. (2021). Orthographic friends and lexical strangers in the L2 lexicon. *Journal of Second Language Studies*, *4*(2), 224–244. https://doi.org/10.1075/jsls.21012.jia

Jończyk, R., Boutonnet, B., Musial, K., Hoemann, K., & Thierry, G. (2016). The bilingual brain turns a blind eye to negative statements in the second language. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(3), 527–540. https://doi.org/10.3758/s13415-016-0411-x

Kazanas, S. A., & Altarriba, J. (2015). The automatic activation of emotion and emotion-laden words: Evidence from a masked and unmasked priming

https://doi.org/10.1017/S1366728923000718 Published online by Cambridge University Press

paradigm. *The American Journal of Psychology*, *128*, 323–336. https://doi.org/10.5406/amerjpsyc.128.3.0323

Kever, A., Grynberg, D., Szmalec, A., Smalle, E., & Vermeulen, N. (2019). "Passion" versus "patience": The effects of valence and arousal on constructive word recognition. *Cognition and Emotion*, *33*(6), 1302–1309. https://doi.org/10.1080/02699931.2018.1561419

Keysar, B., Hayakawa, S. L., & An, S. G. (2012). The foreign-language effect: Thinking in a foreign tongue reduces decision biases. *Psychological Science*, *23*(6), 661–668. https://doi.org/10.1177/0956797611432178

Knickerbocker, H., & Altarriba, J. (2013). Differential repetition blindness with emotion and emotion-laden word types. *Visual Cognition*, *21*(5), 599–627. https://doi.org/10.1080/13506285.2013.815297

Knickerbocker, H., Johnson, R. L., & Altarriba, J. (2015). Emotion effects during reading: Influence of an emotion target word on eye movements and processing. *Cognition and Emotion*, *29*(5), 784–806. https://doi.org/10.1080/02699931.2014.938023

Knickerbocker, H., 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. https://doi.org/10.1016/j.actpsy.2018.10.008

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. https://doi.org/10.1037/a0021446

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*(2), 149–174. https://doi.org/10.1006/jmla.1994.1008

Kuperman, V., Estes, Z., Brysbaert, M., & Warriner, A. (2014). Emotion and language: Valence and arousal affect word recognition. *Journal of Experimental Psychology: General*, *143*(3), 1065–1081. https://doi.org/10.1037/a0035669

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Kyriakou, A., Foucart, A., & Mavrou, I. (2022). Moral judgements in a foreign language: Expressing emotions and justifying decisions. *International Journal of Bilingualism*, *0*(0). https://doi.org/10.1177/13670069221134193

Lei, A., Willems, R. M., & Eekhof, L. S. (2023). Emotions, fast and slow: Processing of emotion words is affected by individual differences in need for affect and narrative absorption. *Cognition and Emotion*, *37*(5), 997–1005. https://doi.org/10.1080/02699931.2023.2216445

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, *44*(2), 325–343. https://doi.org/10.3758/s13428-011-0146-0

Lenth, R. V. (2016). Least-squares means: The R package lsmeans. *Journal of Statistical Software*, *69*(1), 1–33. https://doi.org/10.18637/jss.v069.i01

Lieder, F., & Griffiths, T. L. (2019). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*, e1. https://doi.org/10.1017/s0140525×1900061x

Liu, X., Tu, L., Wang, J., Jiang, B., Gao, W., Pan, X., Li, M., Zhong, M., Zhu, Z., Niu, M., Li, Y., Zhao, L., Chen, X., Liu, C., Lu, Z., & Huang, R. (2017). Onset age of L2 acquisition influences language network in early and late Cantonese-Mandarin bilinguals. *Brain & Language*, *174*, 16–28. https://doi.org/10.1016/j.bandl.2017.07.003

Liu, L., Margoni, F., He, Y., & Liu, H. (2021). Neural substrates of the interplay between cognitive load and emotional involvement in bilingual decision making. *Neuropsychologia*, *151*, 107721. https://doi.org/10.1016/j.neuropsychologia.2020.107721

Liu, L., Schwieter, J. W., Wang, F., & Liu, H. (2022). First and second languages differentially affect rationality when making decisions: An ERP study. *Biological Psychology*, *169*, 108265. https://doi.org/10.1016/j.biopsycho.2022.108265

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*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Monaco, E., Jost, L. B., Gygax, P. M., & Annoni, J. M. (2019). Embodied semantics in a second language: Critical review and clinical implications. *Frontiers in Human Neuroscience*, *13*, 110. https://doi.org/10.3389/fnhum.2019.00110

Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of The Royal Society Interface*, *14*(134), 20170213. https://doi.org/10.1098/rsif.2017.0213

Naranowicz, M. (2022). Mood effects on semantic processes: Behavioural and electrophysiological evidence. *Frontiers in Psychology*, *13*, 1014706. https://doi.org/10.3389/fpsyg.2022.1014706

Naranowicz, M., Jankowiak, K., Kakuba, P., Bromberek-Dyzman, K., & Thierry, G. (2022). In a bilingual mood: Mood affects lexico-semantic processing differently in native and non-native languages. *Brain Sciences*, *12*, 316. https://doi.org/10.3390/brainsci12030316

Panchal, P., Kaltenboeck, A., & Harmer, C. J. (2019). Cognitive emotional processing across mood disorders. *CNS Spectrums*, *24*, 54–63. https://doi.org/10.1017/S109285291800130X

Pavlenko, A. (2008). Emotion and emotion-laden words in the bilingual lexicon. *Bilingualism: Language and Cognition*, *11*(2), 147–164. https://doi.org/10.1017/s1366728908003283

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

Pavlenko, A. (2017). Do you wish to waive your rights? Affect and decision-making in multilingual speakers. *Current Opinion in Psychology*, *17*, 74–78. https://doi.org/10.1016/j.copsyc.2017.06.005

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

Ponari, M., Rodríguez-Cuadrado, S., Vinson, D., Fox, N., Costa, A., & Vigliocco, G. (2015). Processing advantage for emotional words in bilingual speakers. *Emotion*, *15*(5), 644–652. https://doi.org/10.1037/emo0000061

Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology*, *61*(3), 380–391. https://doi.org/10.1037/0022-3514.61.3.380

Qian, Z., Lee, E.-K., Lu, D. H.-Y., & Garnsey, S. M. (2018). Native and non-native (L1-Mandarin) speakers of English differ in online use of verb-based cues about sentence structure. *Bilingualism: Language and Cognition*, *22*(5), 897–911. https://doi.org/10.1017/s1366728918001001

Qiu, X. (2022). Picture or non-picture? The influence of narrative task types on lower- and higher-proficiency EFL learners' oral production. *International Review of Applied Linguistics in Language Teaching*, *60*(2), 383–409. https://doi.org/10.1515/iral-2017-0094

Raatz, U., & and Klein-Braley, C. (1981). The C-Test; A Modification of the Cloze Procedure, In T. Culhane, C. Klein-Braley, & D.K. Stevenson (eds.), *Practice and Problems in Language Testing 7*, Colchester, England: University of Essex, Dept. of Language & Linguistics.

R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Robinson, M. D., Storbeck, J., Meier, B. P., & Kirkeby, B. S. (2004). Watch out! That could be dangerous: Valence-arousal interactions in evaluative processing. *Personality and Social Psychology Bulletin*, *30*, 1472–1484. https://doi.org/10.1177/0146167204266647

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*(3), 783–792. https://doi.org/10.1037/a0027209

Sereno, S. C., Scott, G. G., Yao, B., Thaden, E. J., & O' Donnell, P. J. (2015). Emotion word processing: Does mood make a difference? *Frontiers in Psychology*, *6*, 1191. https://doi.org/10.3389/fpsyg.2015.01191

https://doi.org/10.1017/S1366728923000718 Published online by Cambridge University Press

Shechner, T., Jarcho, J. M., Wong, S., Leibenluft, E., Pine, D. S., & Nelson, E. E. (2017). Threats, rewards, and attention deployment in anxious youth and adults: An eye tracking study. *Biological Psychology*, *122*, 121–129. https://doi.org/10.1016/j.biopsycho.2015.10.004

Sheikh, N. A., & Titone, D. A. (2013). Sensorimotor and linguistic information attenuate emotional word processing benefits: An eye-movement study. *Emotion*, *13*(6), 1107–1121. https://doi.org/10.1037/a0032417

Sheikh, N. A., & Titone, D. A. (2016). The embodiment of emotional words in a second language: An eye-movement study. *Cognition and Emotion*, *30*(3), 488–500. https://doi.org/10.1080/02699931.2015.1018144

Shenaut, G. K., & Ober, B. A. (2021). Multi-trial episodic recall and recognition of emotion-laden words in first versus second language. *Journal of Psycholinguistic Research*, *50*(3), 623–643. https://doi.org/10.1007/s10936-020-09727-2

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. https://doi.org/10.1371/journal.pone.0144576

Snefjella, B., & Kuperman, V. (2016). It's all in the delivery: Effects of context valence, arousal, and concreteness on visual word processing. *Cognition*, *156*, 135–146. https://doi.org/10.1016/j.cognition.2016.07.010

Spielberger, C. D. (1983). *Manual for the State-Trait Anxiety Inventory* (*STAI*). PaloAlto, CA: Consulting Psychologists Press

Stiller, K. D., & Schworm, S. (2019). Game-based learning of the structure and functioning of body cells in a foreign language: Effects on motivation, cognitive load, and performance. *Frontiers in Education*, *4*, 18. https://doi.org/10.3389/feduc.2019.00018

Stukenberg, K. W., Dura, J. R., & Kiecolt-Glaser, J. K. (1990). Depression screening scale validation in an elderly, community-dwelling population. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *2*(2), 134–138. https://doi.org/10.1037/1040-3590.2.2.134

Sulpizio, S., Toti, M., Del Maschio, N., Costa, A., Fedeli, D., Job, R., & Abutalebi, J. (2019). Are you really cursing? Neural processing of taboo words in native and foreign language. *Brain and Language*, *194*, 84–92. https://doi.org/10.1016/j.bandl.2019.05.003

Tang, E., Zhang, M., Chen, Y., Lin, Y., & Ding, H. (2022). Recognition of affective prosody in bipolar and depressive conditions: A systematic review and meta-analysis. *Journal of Affective Disorders*, *313*, 126–136. https://doi.org/10.1016/j.jad.2022.06.065

Tang, D., Fu, Y., Wang, H., Liu, B., Zang, A., & Kärkkäinen, T. (2023). The embodiment of emotion-label words and emotion-laden words: Evidence from late Chinese–English bilinguals. *Frontiers in Psychology*, *14*, 1143064. https://doi.org/10.3389/fpsyg.2023.1143064

Toivo, W., & Scheepers, C. (2019). Pupillary responses to affective words in bilinguals' first versus second language. *PLoS One*, *14*(4), e0210450. https://doi.org/10.1371/journal.pone.0210450

Unkelbach, C., von Hippel, W., Forgas, J. P., Robinson, M. D., Shakarchi, R. J., & Hawkins, C. (2010). Good things come easy: Subjective exposure frequency and the faster processing of positive information. *Social Cognition*, *28*(4), 538–555. https://doi.org/10.1521/soco.2010.28.4.538

Van Dillen, L. F., Papies, E. K., & Hofmann, W. (2013). Turning a blind eye to temptation: How cognitive load can facilitate self-regulation. *Journal of Personality and Social Psychology*, *104*, 427–443. https://doi.org/10.1037/a0031262

Wu, Y. J., & Thierry, G. (2012). How reading in a second language protects your heart. *The Journal of Neuroscience*, *32*(19), 6485–6489. https://doi.org/10.1523/JNEUROSCI.6119-11.2012

Wu, Z., & Wang, Y. (2022). The roles of social status information in irony comprehension: An eye-tracking study. *Frontiers in Psychology*, *13*, 959397. https://doi.org/10.3389/fpsyg.2022.959397

Wu, C., Zhang, J., & Yuan, Z. (2022). An ERP investigation on the second language and emotion perception: The role of emotion word type. *International Journal of Bilingual Education and Bilingualism*, *25*(2), 539–551. https://doi.org/10.1080/13670050.2019.1703895

Yao, Y., Connell, K., & Politzer-Ahles, S. (2023). Hearing emotion in two languages: A pupillometry study of Cantonese-Mandarin bilinguals' perception of affective cognates in L1 and L2. *Bilingualism: Language and Cognition*, *26*(4), 795–808. https://doi.org/10.1017/s1366728922000931

Zhang, J., Wu, C., Meng, Y., & Yuan, Z. (2017). Different neural correlates of emotion-label words and emotion-laden words: An ERP study. *Frontiers in Human Neuroscience*, *11*, 455. https://doi.org/10.3389/fnhum.2017.00455

Zhang, J., Wu, C., Yuan, Z., & Meng, Y. (2020). Different early and late processing of emotion-label words and emotion-laden words in a second language: An ERP study. *Second Language Research*, *36*(3), 399–412. https://doi.org/10.1177/0267658318804850

Appendix 1. Experimental materials in the positive block.

ID	sentence frames	neutral/positive word
1	The science lecture will activate/delight the curious minds in the audience.	activate/delight
2	The strong aroma of perfume will activate/delight the shoppers' sense of smell.	activate/delight
3	The new building looked alien/terrific across the skyline.	alien/terrific
4	Modern art always seems so alien/terrific to me because of the free flowing forms.	alien/terrific
5	Laura wanted to have the appliance/joy she saw in the commercial on television.	appliance/joy
6	Michael believed that the new appliance/joy he found was exactly what he needed.	appliance/joy
7	Erin's parents thought she would bake/love the food that was at the family reunion.	bake/love
8	The Italian chef will probably bake/love pizza until the day he dies.	bake/love
9	My guests got a little basket/merry at the Christmas party this weekend.	basket/merry
10	We may get a little basket/merry at the holiday celebration.	basket/merry
11	Yesterday, there was a blond/nice doorman who helped me carry my things inside.	blond/nice
12	Clarissa watched the blond/nice teenager hand a toy to the crying child.	blond/nice
13	The chef changed the menu increasing the amount of butter/bliss in every dish.	butter/bliss
14	The farmer thought butter/bliss was necessary to get through the day.	butter/bliss
15	The couple found their old chair/passion as they visited their first home together.	chair/passion

(*Continued*)

https://doi.org/10.1017/S1366728923000718 Published online by Cambridge University Press

Appendix 1. (*Continued.*)

ID	sentence frames	neutral/positive word
16	Nora found her chair/passion about a year ago.	chair/passion
17	The sheriff is going to circle/secure the area around the crime scene.	circle/secure
18	The secret service will circle/secure the perimeter before the senator arrives.	circle/secure
19	The experienced nurse will coast/comfort patients through the pre-operation procedures.	coast/comfort
20	The new cruise ship will coast/comfort passengers through the waters of the Arctic.	coast/comfort
21	Tim always used to concentrate/hope to finish his work early every day.	concentrate/hope
22	The debate team will concentrate/hope to make it to the state championship.	concentrate/hope
23	Hank hired the custom/thankful painter who had lost his job.	custom/thankful
24	I helped the custom/thankful cabinet maker find several new clients.	custom/thankful
25	Jean wanted to increase the level of detail/ecstasy included in her paintings.	detail/ecstasy
26	Writers must work to inject detail/ecstasy into their literary works.	detail/ecstasy
27	The teacher watched the older student elbow/soothe the younger child during recess.	elbow/soothe
28	Regina tried to elbow/soothe her boyfriend during the fight.	elbow/soothe
29	Steven was trying to find some highway/acceptance in Los Angeles.	highway/acceptance
30	The artist could not find the highway/acceptance he was looking for in New York.	highway/acceptance
31	It is important to invest/cheer for your child's future goals.	invest/cheer
32	I prefer to invest/cheer with people that I know.	invest/cheer
33	The bronze statue of the iron/brave soldier stood in the park.	iron/brave
34	The book had a(n) iron/brave robot as a main character.	iron/brave
35	We clapped as the key/elated player made the winning shot.	key/elated
36	The audience listened as the key/elated soloist finished the song.	key/elated
37	Doug watched the limber/grateful acrobat thank the audience for attending.	limber/grateful
38	While at the Olympics, the limber/grateful gymnast used the new equipment.	limber/grateful
39	The family thinks the museum/excitement this afternoon will be too much for grandpa.	museum/excitement
40	Penny told me to check out the museum/excitement in the park this weekend.	museum/excitement
41	Allison thought the noisy/gentle animal was looking for food at the campsite.	noisy/gentle
42	Ben wanted to adopt the noisy/gentle puppy he saw at the pet store.	noisy/gentle
43	The clown gave the child paint/pleasure at his birthday party.	paint/pleasure
44	The art teacher provided a lot of paint/pleasure to her new students.	paint/pleasure
45	We have to get the patent/desire to move the company forward.	patent/desire
46	Someone has to patent/desire the new product before we can market it.	patent/desire
47	Even on busy days, the patient/jolly conductor waited for everyone to board the train.	patient/jolly
48	During the holidays, the patient/jolly elf handed out toys to every child.	patient/jolly
49	The tenured professor gave his poster/proud presentation at the conference.	poster/proud
50	The teacher thought that Josh was a poster/proud boy because of his community service.	poster/proud
51	My mother emphasized the importance of privacy/optimism in having a balanced life.	privacy/optimism
52	The teacher stated that some privacy/optimism was important for healthy living.	privacy/optimism
53	During the show, the quiet/awed crowd watched the lion tamer.	quiet/awed
54	From the capital building, the quiet/awed politician watched the sunrise.	quiet/awed
55	The huge fireworks display will rattle/dazzle the younger children in the audience.	rattle/dazzle
56	The explosions in the action movie will rattle/dazzle the viewers at the premier.	rattle/dazzle
57	The company's actions this week will rock/thrill their investors and creditors.	rock/thrill
58	The author's new novel will rock/thrill the critics and many readers.	rock/thrill

(*Continued*)

https://doi.org/10.1017/S1366728923000718 Published online by Cambridge University Press

Appendix 1. (*Continued.*)

ID	sentence frames	neutral/positive word
59	The doctor said I need more salad/enjoyment in my life to be healthier.	salad/enjoyment
60	You can never have too much salad/enjoyment at the dinner party.	salad/enjoyment
61	Despite being so young, the sheltered/friendly child went out of his way to make friends.	sheltered/friendly
62	At the rock concert, the sheltered/friendly teenager met diverse people.	sheltered/friendly
63	The birthday clown plans to spray/fascinate the children with his water gun.	spray/fascinate
64	The orca whales at Sea World always spray/fascinate the entire crowd.	spray/fascinate
65	Many tourists found that the statue/kindness in the city was amazing.	statue/kindness
66	Jake told me he did not expect the statue/kindness that he received.	statue/kindness
67	I listened to the stiff/happy mayor present his new policy.	stiff/happy
68	At his retirement, the stiff/happy general gave a speech.	stiff/happy
69	The news reported on the umbrella/hopeful company that opened in Seattle.	umbrella/hopeful
70	The rainy town wanted the umbrella/hopeful manufacturer to increase production.	umbrella/hopeful
71	The first grader drew a yellow/joyful bird sitting in a tree.	yellow/joyful
72	The child pets the yellow/joyful cat purring on her lap.	yellow/joyful

Appendix 2. Experimental materials in the negative block.

ID	sentence frames	neutral/negative words
1	I think the flower is going to activate/irritate my allergies and give me a runny nose.	activate/irritate
2	The strong aromas from the restaurant will probably activate/irritate my sense of smell.	activate/irritate
3	The story described a small, alien/helpless creature who had landed here from outer space.	alien/helpless
4	George was awoken by a small alien/helpless figure drowning in his backyard pool.	alien/helpless
5	Beth would soon forget about the appliance/grief her family had at the farm house.	appliance/grief
6	The realtor could actually view the appliance/grief as he peered into the vacant home.	appliance/grief
7	Every holiday, I return home and bake/hate the dinner that the family eats together.	bake/hate
8	I am pretty sure that Kevin will bake/hate the sugar cookie mix from the grocery store.	bake/hate
9	The little boy seemed to ignore the basket/loneliness that his new friend had.	basket/loneliness
10	The teacher commented on the basket/loneliness that the new student had.	basket/loneliness
11	Ashley tried to comfort her blond/sad friend, who was upset about his dead cat.	blond/sad
12	While in the store, I saw a blond/sad child crying because he was lost.	blond/sad
13	Alex typically had a lot of butter/pain in the morning at breakfast.	butter/pain
14	Many people have too much butter/pain in their daily lives to be healthy.	butter/pain
15	The family talked about the chair/anger that was a nuisance and always in the way.	chair/anger
16	The discussion focused on the chair/anger that Carol got during her divorce from Phil.	chair/anger
17	The new young street gang will circle/fear us if we face them in their territory.	circle/fear
18	The cowboys will likely circle/fear the Native Americans if they attack the fort.	circle/fear
19	The worker will coast/offend at work if he is not supervised.	coast/offend
20	My dad has a tendency to coast/offend while driving the car.	coast/offend
21	Darla wanted to concentrate/scorn our efforts to update the equipment.	concentrate/scorn
22	Bob's employer wanted to concentrate/scorn his efforts towards obtaining a promotion.	concentrate/scorn

(*Continued*)

https://doi.org/10.1017/S1366728923000718 Published online by Cambridge University Press

Appendix 2. (*Continued.*)

ID	sentence frames	neutral/negative words
23	Billy tried to sell the custom/useless toolbox to his nephew.	custom/useless
24	Jean saved money to buy the custom/useless bike that no one else at the store liked.	custom/useless
25	The buyer was not expecting the detail/gloom found in every room in the house.	detail/gloom
26	The art critic disliked the detail/gloom found in the paintings.	detail/gloom
27	Helen watched the older student elbow/hurt the other child during an argument.	elbow/hurt
28	At the game, the boy did not mean to elbow/hurt the player on the other team.	elbow/hurt
29	Paula always found that the highway/agony was the worst in the evening.	highway/agony
30	Will hoped that he could avoid the highway/agony for the rest of the work week.	highway/agony
31	Charlotte tried not to invest/detest the gifts that she inherited from her parents.	invest/detest
32	Jeremy thinks that very few employees invest/detest the money that they make.	invest/detest
33	Carrie was paid to paint the iron/terrible sculpture that is in front of the school.	iron/terrible
34	We decided to remove the iron/terrible gates that surrounded the entire property.	iron/terrible
35	At the basketball practice, the key/distressed player argued with the coach.	key/distressed
36	The boss just fired the key/distressed salesman for being late to work every day.	key/distressed
37	The coach called to see if the limber/sick athlete could play in the game on Saturday.	limber/sick
38	Zachary was watching the limber/sick child upset his parents by climbing the jungle gym.	limber/sick
39	The teacher did not like the children's museum/rage she saw on the field trip.	museum/rage
40	The mayor did not like the town's museum/rage and tried to fix it.	museum/rage
41	Ursula was forced to tell the noisy/cruel child that he could not yell on the bus.	noisy/cruel
42	The little boy was noisy/cruel and had to be sent to his room.	noisy/cruel
43	Lisa is not sure about the room and will paint/despise it if we decide to move into this house.	paint/despise
44	Ned informed me that he will paint/despise the coffee table even if I like it.	paint/despise
45	The art teacher tried to patent/humiliate my new design.	patent/humiliate
46	The professor openly attempted to patent/humiliate my ideas about time travel.	patent/humiliate
47	Even when trapped by the police, the patient/guilty robber was able to remain calm and escape.	patient/guilty
48	Jeff could not believe the patient/guilty parent was ignoring the crying child.	patient/guilty
49	Cathy immediately reported the poster/hatred and arguments she found on the message board.	poster/hatred
50	Groups of people displayed their poster/hatred as they shouted at the criminals.	poster/hatred
51	The lawyer fought hard for the privacy/depression of his clients to be noted by the court.	privacy/depression
52	Faith argued with Tom about privacy/depression because he was not respecting her feelings.	privacy/depression
53	The observant teacher told the quiet/unhappy toddler to play with the other children.	quiet/unhappy
54	Even at a young age, Todd was quiet/unhappy and had trouble meeting new people.	quiet/unhappy
55	I am warning you that Roxanne will rattle/betray your friendship and trust at some point.	rattle/betray
56	Tom has a dark past and will rattle/betray the ethical standards of this company.	rattle/betray
57	The heavy metal band will rock/disappoint the entire audience with their new songs.	rock/disappoint
58	These new findings are going to rock/disappoint the scientific and medical community.	rock/disappoint
59	Theresa did not like the salad/jealousy she had at her friend's house.	salad/jealousy
60	Though the restaurant was a success, the salad/jealousy still made the owner unhappy.	salad/jealousy
61	Sadly, Rachel always felt sheltered/afraid and was never an outgoing person.	sheltered/afraid
62	After growing up in a town, Sean was sheltered/afraid and did not think he could live in a city.	sheltered/afraid
63	Victor really did not mean to spray/upset the girls as he washed his car in their driveway.	spray/upset
64	The cook is planning to play a joke and spray/upset the troublesome busboy with flat soda.	spray/upset
65	The director was concerned about the statue/horror in the first scene of the scary movie.	statue/horror

(*Continued*)

https://doi.org/10.1017/S1366728923000718 Published online by Cambridge University Press

Appendix 2. (*Continued.*)

ID	sentence frames	neutral/negative words
66	I certainly did not expect the statue/horror that you made me see today.	statue/horror
67	Everyone knew that Craig was stiff/nervous when he spoke in front of the whole class.	stiff/nervous
68	Usually, Franklin was too shy and stiff/nervous to play guitar in front of other people.	stiff/nervous
69	I ignored Robert's lecture about my umbrella/misery because he was repeating himself.	umbrella/misery
70	A book about the history of the umbrella/misery in a rainy town would not sell very well.	umbrella/misery
71	When I looked up, the yellow/lonely sun was the only thing in the sky.	yellow/lonely
72	In art class, the child drew a yellow/lonely stick figure on a blank sheet of paper.	yellow/lonely

https://doi.org/10.1017/S1366728923000718 Published online by Cambridge University Press