languages

Article

How Do Foreign Language Learners Process L2 Emotion Words in Silent Reading? An Eye-Tracking Study

Dina Abdel Salam El-Dakhs [1,*], Suhad Sonbul [2] and Jeanette Altarriba [3]

1 Linguistics and Translation Department, College of Humanities and Sciences, Prince Sultan University, Riyadh 12435, Saudi Arabia
2 English Language Centre, Umm Al-Qura University, Mecca 24382, Saudi Arabia
3 Department of Psychology, University at Albany, State University of New York, Albany, NY 12222, USA
* Correspondence: ddakhs@psu.edu.sa

Abstract: The current study aimed to examine the processing of emotion words in L2 silent reading. We conducted two experiments in which Arab learners of English as a foreign language (EFL) read short English sentences in which target words were embedded. The participants' eye movements were recorded and analyzed. The results of Experiment 1, which compared the processing of emotionally positive versus neutral words by 44 participants, did not reveal any significant effect for word type. The results only showed a few instances of significant interactions between word type and word frequency (i.e., positive words were read faster than neutral words only in the case of high-frequency words) and arousal (i.e., positive words were recognized faster than neutral words only when the target words were low in arousal). The results of Experiment 2, which compared the processing of emotionally negative versus neutral words by 43 participants, only established one effect of word type on the skipping rate which was also modulated by length (i.e., negative words were less likely to be skipped, particularly shorter ones). Moreover, arousal interacted with word type (i.e., only the negative words with low arousal were read faster than neutral words in two eye-movement measures).

Keywords: emotion words; valence; arousal; word frequency; eye-movement

Citation: El-Dakhs, Dina Abdel Salam, Suhad Sonbul, and Jeanette Altarriba. 2023. How Do Foreign Language Learners Process L2 Emotion Words in Silent Reading? An Eye-Tracking Study. *Languages* 8: 125. https://doi.org/10.3390/languages8020125

Academic Editor: Jeanine Treffers-Daller

Received: 6 February 2023
Revised: 4 April 2023
Accepted: 25 April 2023
Published: 4 May 2023

1. Introduction

It is widely acknowledged that emotion words behave more distinctively than concrete and abstract words (e.g., Altarriba et al. 1999; Bauer et al. 2009; El-Dakhs and Altarriba 2018). It has also been recurrently reported that emotion words exhibit a processing advantage over neutral words across several cognitive tasks (Ferré 2002; Kazanas and Altarriba 2015a; Kousta et al. 2009). These findings have triggered further studies (Kuperman et al. 2014; Ponari et al. 2015; Sereno et al. 2015) that aimed to explore the potential modulating factors of this distinctiveness and processing advantage, such as word frequency, arousal and the participants' mood and L2 status.

Despite the intriguing findings of the above studies, they have been critiqued for largely focusing on isolated words. The tasks employed in these studies were designed to examine the behavior of individual words out of context. As the activation and integration of the meaning of words operate through different processes in normal silent reading, in cases of individual word recognition, there have been repeated calls to examine the behavior of emotion words under normal reading conditions. Thus, a few studies, e.g., (Knickerbocker et al. 2015; Scott et al. 2012), have recently attempted to explore the processing advantage of emotion words in contextualized tasks using eye-tracking. In these studies, the participants read sentences in which the target words were embedded, and the participants' eye movements were monitored against several predefined measures. However, these studies have mainly focused on native language speakers to the neglect of bilinguals who are, in fact, the majority of the world's residents.

The current study aims to address this gap through investigating the processing of L2 emotion words in normal silent reading by foreign language learners. Specifically, the current study will extend the work of Knickerbocker et al. (2015) on native English speakers to a population of Arab EFL learners. Through the use of 13 eye-movement measures, the current study will compare the processing of emotionally positive versus neutral words in Experiment 1 and of emotionally negative versus neutral words in Experiment 2.

The findings of the present study will be of great importance for this line of emotion research for three reasons. First, the study addresses a true gap in the literature. There is a dearth of research on the emotionality advantage among bilinguals in normal reading. This gap is particularly important because the native language is often described as being more emotionally advantageous than the L2 (Caldwell-Harris 2014) and as being emotionally close while subsequent languages are emotionally distant (Chen et al. 2015). Second, the findings will target a population that is relatively underrepresented in the psycholinguistic literature; namely, Arabic speakers of English. This particular population is important because Arabic and English are extremely distant languages that follow completely different linguistic systems. Third, the study will have theoretical implications in relation to the Revised Hierarchical Model (Kroll and Stewart 1994) and the Mental Lexicon Model (Jiang 2000). The study will also have practical implications in relation to language pedagogy.

To situate the present study, the following section will include a brief survey of earlier studies in this direction. This will be followed by listing our research questions and describing the study methodology. The results will then be presented and discussed in light of the literature survey. Finally, conclusions, including the relevant implications and suggested directions for future research, will be drawn.

2. Literature Review

2.1. The Emotionality Advantage in L1 Word Processing

Almost two decades ago, researchers started to consider distinguishing emotion words (e.g., *happy, sad*) from concrete words (e.g., *chair, table*) and abstract words (e.g., *freedom, jealousy*). The distinctiveness of emotion words was based on several reports that the three types of words behave differently. For example, Altarriba et al. (1999) compared the behavior of emotion words with concrete and abstract words in two experiments: one using rating scales that involved 78 English-speaking monolinguals and another one using the discrete word association task that involved 55 English-speaking monolinguals. The findings strongly suggested a more distinctive behavior for emotion words than abstract and concrete words in memory representation, recognition and retrieval. The distinctiveness of emotion words was similarly found in later studies using other tasks, including Algom and Bauer (2004) with the use of a free recall task, a rating task and a lexical decision task (LDT) and Bauer et al. (2009) with the use of a list-learning false memory paradigm.

Later reports also pointed out that emotion words have a processing advantage over neutral words in several cognitive tasks; that is, emotion words are processed with greater ease than neutral words in these tasks. For example, emotion words were better remembered than neutral ones in both immediate and delayed memory tests (Ferré 2002), were processed faster than neutral words in LDTs (Kousta et al. 2009), were recognized faster in naming tasks (Kuperman et al. 2014) and triggered shorter response times and greater priming effects in masked and unmasked priming tasks (e.g., Kazanas and Altarriba 2015b).

These findings concerning the special status of emotion words triggered the attention of scholars to explore the potential modulating factors of the emotional advantage. Kuperman et al. (2014) used a sample of 12,658 words and included many lexical and semantic control factors to determine the precise nature of the effects of arousal and valence on word recognition. The data were collected based on LD and naming tasks. The results showed that valence and arousal do not interact, but both interact with word frequency. Valence and arousal exerted clearer influence on low-frequency, rather than high-frequency, words.

The results also showed that negative words were recognized more slowly than positive words, and arousing words were recognized more slowly than calming words.

Similarly, Sereno et al. (2015) examined how word frequency and the participant's induced mood could influence the LDT's responses for emotionally positive, negative and neutral words. The results revealed an interaction between a word's emotionality and its frequency. Overall, positive and negative moods facilitated responses in comparison to neutral words. However, mood-congruent words, whether positive or negative, did not lead to additional facilitation. In terms of the emotion X frequency interaction, frequency influenced the behavior of negative words. A relative slowing of responses to negative words versus positive words was noted, especially among high-frequency negative words, which were recognized as slowly as neutral words.

In the same vein, Barriga-Paulino et al. (2022) investigated the modulating effect of word frequency and arousal on the processing advantage of valence in two LDT experiments involving 96 native Portuguese speakers. The results of the LDT showed that positive words were recognized faster and negative words were recognized slower than neutral words. Additionally, valence interacted with word frequency as it affected only low-frequency words. The effect of the valence X frequency interaction was reduced for high-arousal words when the pressure to respond was high.

In summary, the research on L1 word processing has revealed that emotional words are characteristically different from, and activate different relationships than, concrete and abstract words. Emotion words also exhibited a processing advantage over neutral words in several cognitive tasks. This processing advantage can, according to Kousta et al. (2009), be explained in terms of the motivated attention and affective states model (Lang et al. 1990, 1997), which postulates that attention is captured and sustained by motivationally significant stimuli rather than by neutral stimuli. Emotion words are viewed as motivationally relevant to attentional functions and thus capture attention differently than other words and lead to a rapid modification of behavior.

However, when examined separately, emotionally positive and negative words demonstrated different behaviors. Negative words often take a longer time to be recognized than positive words, particularly in LDTs (Barriga-Paulino et al. 2022; Kuperman et al. 2014; Sereno et al. 2015). These slower responses are often explained in terms of the automatic vigilance model of emotion (Pratto and John 1991), which assumes that negative words do not attract attention to a greater extent than either positive or neutral stimuli, but that they withhold attention longer (Fox et al. 2001) due to an innate defense mechanism that temporarily freezes all threatening activities (Algom et al. 2004). This leads to slower responses to negatively valenced stimuli in a variety of cognitive tasks (Algom et al. 2004; Estes and Verges 2008). In this context, it is expected that arousing words (Kuperman et al. 2014) and high frequency negative words (Sereno et al. 2015) will result in longer responses than calming words and low-frequency negative words, respectively, as arousing words and highly frequent negative words can similarly trigger more automatic vigilance.

2.2. *The Emotionality Advantage in L2 Word Processing*

Research on the emotionality advantage in the L2 lags behind research on the L1. One of the early studies in this regard was conducted by Algom and Basnight-Brown (2004). After training 60 English-speaking monolinguals on a set of Spanish words, Altarriba and Basnight-Brown requested their participants to complete a Stroop color-word task and a translation recognition task which included the target Spanish words. The participants produced faster naming times, longer recognition times and higher error rates when recognizing emotion vs. non-emotion words. Hence, the results supported the distinctiveness of emotion words compared to non-emotion words and a potential processing advantage for emotion words in naming tasks.

Another relevant study was reported by Ponari et al. (2015), who had 95 native English speakers and 156 English-as-a-second-language (ESL) speakers complete an LDT that included emotionally positive, negative and neutral words, in addition to non-words that served as control items. The ESL participants were classified as early L2 (mean age of English acquisition = 4.07) and late L2 (mean age of English acquisition = 11.95) bilinguals. Their native languages varied, but they were all highly proficient in English and residents of an English-speaking country. It was found that these highly proficient ESL learners exhibited the same facilitation in processing emotionally valanced words as native English speakers regardless of their L1, the age of English acquisition and the frequency and context of English use.

El-Dakhs and Altarriba (2018) investigated whether the distinctiveness of emotion words holds for foreign language learners. The results of a free recall task and a rating task that were completed by Arab EFL learners supported the emotionality advantage, particularly for highly proficient (n = 120) versus less proficient learners (n = 120) of English. More emotion words were recalled than non-emotion words. Additionally, emotion words were rated significantly differently than non-emotion words in terms of concreteness, imageability and context availability. However, no significant difference was noted between emotion and non-emotion words in a discrete word association task. This pattern of findings indicated that the distinctiveness of emotion words vs. non-emotion words can be task dependent. Emotion words behaved differently in this study in the free recall and rating tasks. In the word association task, no significant distinction emerged. This finding may have been caused by the different structures of the Arabic and the English mental lexicons. While the former is morphophonemically structured, the latter is more meaning-based.

In the same vein, Ferré et al. (2018) investigated whether the processing advantage for emotional words holds in the different languages of bilinguals. A group of 40 highly proficient Catalan-Spanish bilinguals, who learned both languages through immersion and remained in that context, performed an affective LDT and a normal LDT in both Catalan and Spanish. Additionally, a group of 51 Catalan-Spanish bilinguals who learned English at a later stage through an instructional setting and lived in a non-English speaking country, yet were proficient in English, performed an LDT. Both groups also completed an unexpected free recall task. The results showed that language status modulates the influence of emotional content on language processing. Although Experiments 1 and 2 did not show an interaction between language and emotionality, Experiment 3 revealed a clear interaction between the two variables. For example, Experiments 1 and 2 showed a facilitation effect for positive words and an interference effect for negative ones in both languages, similar to the findings in the L1 word processing research (see above). However, this pattern was only maintained in Spanish in Experiment 3, whereas English showed a facilitation effect for both types of words. This finding suggests that bilinguals who acquire their L2 in early childhood and through immersion can process the emotional content of both languages similarly. In contrast, those who learn their L2 at a later stage and in instructional settings seem to process emotionality differently across their languages.

The special influence of language status on the processing of emotion words can be interpreted in light of Kroll and Stewart's (1994) Revised Hierarchical Model (RHM). The RHM postulates that memory representation in bilinguals exists across three components, namely: (1) conceptual; (2) L1 lexicon; and (3) L2 lexicon. The first component resides within the conceptual level of processing, while the latter two are part of the lexical level of processing. Both lexical and conceptual connections hold across these levels. In the L1 lexicon, the conceptual connections among L1 words and their meanings are strong because these connections developed through extensive use of the native language. However, newly learned L2 words have stronger lexical than conceptual connections. They are often linked more strongly with their L1 translation counterparts. With increasing L2 exposure/proficiency, the conceptual connections for L2 words gradually strengthen, which enables L2 learners to access meaning directly without L1 mediation and can thus show similar language processing behavior to monolinguals. According to the RHM, Ferré et al.'s (2018) results fit perfectly with the model. Learning an L2 through immersion and at an early stage will strengthen the conceptual connections in the L2 lexicon much earlier and to a greater extent than learning an L2 through instructional performance and at an older age; however, see (Altarriba and Mathis 1997 for a slightly different view).

Another relevant model in this regard is Jiang's L2 Mental Lexicon Model (MLM), which postulates that L1 transfer is a central mechanism in the acquisition of L2 vocabulary. The L2 MLM divides lexical knowledge into: (1) lexeme (phonology, orthography and morphology) and (2) lemma (syntax and semantics), with the latter being amenable to L1 influence. The L2 MLM specifies three stages of acquiring L2 vocabulary. The first stage, known as word association, is characterized by connections between L2 words and their L1 translation equivalents without any connections between the L2 words and their concepts. In the second stage, called L2 mediation, L1-specific lemma information transfers into L2 words while the L2 words form weak connections with their concepts. With increasing L2 exposure, L2 words develop direct conceptual connections and the L2-specific lemma information emerges and replaces the L1-specific lemma information in the third developmental stage of word acquisition, which is known as full integration. According to Jiang (2000), only a few advanced L2 learners move to the third stage, while the majority remain stuck in the second stage, which is described as a fossilized L2 stage. The distinctive acquisition constraints that L2 words undergo can explain why language proficiency/exposure plays a clear role in processing emotion words. Less proficient learners do not process L2 emotion words based on direct conceptual connections, as L1 speakers do; hence, they are highly likely to exhibit a different pattern of behavior.

2.3. The Emotionality Advantage in Normal Reading

Despite the intriguing findings in the domain of word processing, this line of research has received some harsh criticism. For example, Scott et al. (2012) criticized the abundance of research on affective word processing because "words are presented in isolation and response time is relatively delayed", (p. 788). They thus called for an investigation of the emotionality advantage in normal reading, during which "individual word meanings are rapidly activated and integrated online into a developing discourse context" (p. 788). According to them, research on affective word processing alone can never suffice to establish a reliable effect for emotion in language processing. Following this plausible proposition, Scott et al. (2012) monitored the eye movements of 48 native speakers of English while reading sentences containing emotionally positive, negative and neutral words. Target word frequency was also considered. Measures of target fixation time demonstrated the significant effects of emotion and frequency, as well as an interaction between them. Emotionally positive and negative words demonstrated faster fixation times than their neutral counterparts, with the exception of high-frequency negative words, which were not fixated any faster than neutral words. The earliest eye-movement measures indicated that lexical processing is modulated by emotionality.

As a result of Scott et al. (2012) intermixing emotion-label words (e.g., happy and terrified/a state of mind) with emotion-laden words (e.g., marriage and debt/a concept that is associated with an emotional state) in their experimental stimuli, Knickerbocker et al. (2015) conducted another study to investigate the processing of emotional words in normal silent reading. Knickerbocker et al.'s (2015) study contained only emotion-label words. This was an important measure as a number of studies have pointed out that emotion-label and emotion-laden words behave differently (e.g., Kazanas and Altarriba 2015a, 2015b; Knickerbocker and Altarriba 2013; Zhang et al. 2017).

In their study, Knickerbocker et al. (2015) conducted two experiments. The first experiment was designed to compare the processing of positive words versus neutral words, while the second experiment compared the processing of negative words versus neutral words. The participants in the two experiments were native English speakers ($n = 64$ and $n = 78$, respectively). The participants completed the Beck Depression Inventory (BDI) (Beck et al. 1996) and the State-trait Anxiety Inventory (STAI) (Spielberger et al. 1983) surveys to assess their level of depression and anxiety, respectively. Then, they read the experimental sentences in which the target words were embedded. Several eye-movement measures were recorded and analyzed, including early (e.g., first fixation duration) and late (e.g., total time or post-target region) measures.

The results of Experiment 1 showed a consistent positive emotion processing advantage across multiple early, as well as late, measures. In other words, positive words were mostly read with greater ease than neutral words. Word type (i.e., positive vs. neutral) significantly interacted with arousal. The processing advantage of positive words was stronger in the higher arousal condition. Notably, there was no significant relation or interaction between the BDI or STAI results and the emotional advantage; that is, the facilitation effect of positive words was evident regardless of the participants' depressive or anxiety levels.

The results of Experiment (2) showed a similar facilitation effect for negative words over neutral words. However, the effect was found only in some early (e.g., first fixation and single fixation) and later (i.e., regression and total time) measures. Similarly to positive words, a significant interaction was observed between word type and arousal. Negative words with high arousal were processed more slowly. In contrast to Experiment 1, Experiment 2 demonstrated some interaction between negative words and the results of the BDI and STAI tests, which means that the participants were somehow influenced by their status of depression and/or anxiety while processing negative words.

3. The Current Study

Similar to Scott et al. (2012) and Knickerbocker et al. (2015), several studies have recently emerged to investigate word processing in silent reading using eye-trackers (Gerth and Festman 2021; Huang et al. 2022; Kuperman and Deutsch 2020). However, there is a paucity of research on the processing of emotion words among bilinguals. This is the gap that the current study aims to fill. This study, which extends the work of Knickerbocker et al. (2015), is designed to investigate how emotionally positive and negative words are processed in comparison with neutral words in normal silent reading across several eye-movement and eye-fixation measures. The focus here will be on foreign language learners, particularly Arab EFL learners, who learned English as late bilinguals in an instructional setting and live within an L1-dominant community. Will the processing advantage for emotion words cited above (Knickerbocker et al. 2015) be maintained in this context? Following Knickerbocker et al. (2015), we also aimed to examine any modulating effects that several item-related and participant-related factors might have on the potential advantage of emotion words (positive in Experiment 1 and negative in Experiment 2) over neutral words. More specifically, the current study addresses these research questions:

1. Is there an advantage for positive emotion words when compared to neutral words for EFL learners during natural reading? (Experiment 1)
2. What modulating effects do learner-related and item-related factors have on any potential advantage for positive emotion words over neutral words during natural reading? (Experiment 1)
3. Is there an advantage for negative emotion words when compared to neutral words for EFL learners during natural reading? (Experiment 2)
4. What modulating effects do learner-related and item-related factors have on any potential advantage for negative emotion words over neutral words during natural reading? (Experiment 2)

The modulating factors that we aim to examine in the present study include the following participant-related factors: measures of depression and anxiety (as represented by STAI and BDI scores); a rough proficiency measure (vocabulary test scores); and age, as our participants varied greatly in terms of their age. The item-related factors include: arousal; word length; and word frequency, which are well-established factors modulating the reading of emotion words (Angele et al. 2014; Barriga-Paulino et al. 2022; Brysbaert and Vitu 1998).

4. Experiment 1

4.1. Participants

A total of 44 participants took part in Experiment 1. These were L1 Arabic—L2 English speakers in Saudi Arabia (36 female and 8 male). They ranged in age between 18 and 38 (*M* = 23.52, *SD* = 5.29). Age was controlled for statistical significance in the analysis (see below). They started learning English at an average age of seven and a half (*M* = 7.48, *SD* = 4.50), mostly in schools. The participants' average self-reported proficiency across the four skills (reading, listening, reading and writing) on a score from 1 = very poor to 5 = excellent was 4.32 (*SD* = 0.52).

The participants completed the V_YesNo online vocabulary test (Meara and Miralpeix 2017; maximum score = 10,000) as a proxy for their L2 proficiency. Their scores ranged between 4159 and 8966 (M = 6543.75, SD = 1039.75), which indicates a good-to-high level of proficiency (Meara and Miralpeix 2017). We added the vocabulary test score as a main and interacting variable in the mixed-effects models (see below) in order to examine any potential effect of proficiency on reading times (RTs) for positive and neutral words.

The participants were presented with one of two versions of the eye-tracking experiment randomly (see *Design* below). Thus, nearly half of the participants took Version 1 of the Experiment (*n* = 23), and the other half took Version 2 (*n* = 21).

4.2. Design and Stimuli

The items were borrowed from Knickerbocker et al. (2015), who selected their target positive and neutral words from the ANEW database (Bradley and Lang 1999). Knickerbocker et al. (2015) included 36 pairs, but we opted to select the 21 pairs with the most frequent target words (see Appendix A). This was intended to ensure that our L2 participants were likely to know them. The target words were divided equally between three parts of speech (Verbs = 7, Nouns = 7, Adjective = 7). All target items belonged to the most frequent 4000-word families in English and were thus likely to be known by EFL speakers in the present context (see *Participants* above). Only four target words (*appliance, optimism, soothe* and *detest*) belonged to higher frequency levels. Thus, we included word frequency as a covariate in the analysis (see below) to control for its effect on emotion word processing. The items in each pair were matched for various factors. The characteristics of the target items are presented in Table 1. It is evident that the target items were similar along several variables, but they significantly differed in their valence and arousal scores. The former score is the variable under manipulation in the present study, and arousal will be controlled for in the statistical analysis (see below).

Table 1. Mean lexical characteristics for the target items used in Experiments 1 and 2.

	Neutral		*Positive*		*Negative*			
	M	***SD***	***M***	***SD***	***M***	***SD***	***F***	***p***
Valence	5.50	0.45	7.76	0.59	2.22	0.46	340.83	<0.001 ***
Arousal	4.34	0.76	5.63	1.34	5.91	1.13	6.72	0.004 **
Length (no. of letters)	5.90	1.45	6.48	1.94	6.10	2.10	0.43	0.66
Word frequency (Kučera–Francis)	24.81	22.61	37.05	51.60	30.11	28.03	1.05	0.36
Word frequency (log HAL)	8.79	1.09	8.43	1.85	8.75	1.31	0.09	0.91
Word frequency (SUBTLEX)	16.63	19.54	111.96	275.58	58.81	79.53	3.06	0.06
Orthographic neighborhood size	2.76	4.29	2.43	3.70	3.52	5.78	0.15	0.86
Mean naming RT (ms)	622.26	52.09	632.94	69.55	628.42	51.81	0.51	0.61
Mean LDT RT (ms)	633.87	50.21	642.44	89.52	636.23	70.26	0.26	0.77

** and ***: statistical significance.

The sentence contexts were also adopted from Knickerbocker et al. (2015). Each pair of items was presented in two sentence contexts (counter-balanced across lists). The sentences were already normed for understandability and predictability with 20 native speakers of English (see a study by Knickerbocker et al. (2015) for more details). The following is an example of sentence contexts for the target words (bake/love):

The Italian chef will probably bake/love pizza until the day he dies.

Erin's parents thought she would bake/love the food that was at the family reunion.

We checked for the lexical coverage of the sentence contexts and found it acceptable; 97.5% of the words in the context belonged to the most frequent 4000-word families (likely to be known by our participants, see above). Thus, no changes were made on the sentence contexts. One-third of the sentence contexts were presented with a comprehension question to ensure that participants were reading for meaning. (Please refer to Knickerbocker et al. (2015) for the full list of sentences.)

4.3. Procedures

After signing the consent form, each participant was asked to complete the BDI-II (Beck et al. 1996) and the State Version of the STAI (Spielberger et al. 1983). The items within each questionnaire were presented in random order. The BDI includes 21 items for which participants select the statement that indicates their feelings and behaviors during the past two weeks. The following item was excluded from the BDI-II as it was considered culturally sensitive:

0 I have not noticed any recent change in my interest in sex.

1 I am less interested in sex than I used to be.

2 I have almost no interest in sex.

3 I have lost interest in sex completely.

Thus, the possible scores on the BDI scale in the present study ranged between 0 and 60, with higher scores reflecting higher levels of depressive symptoms. The STAI includes 20 items that assess the feelings of the participants at that very moment. The possible scores ranged between 20 and 80, with higher scores indicating higher levels of anxiety. After completing the emotion questionnaires, the participant completed the vocabulary test, then the eye-tracking experiment began.

We used an SR Research EyeLink 1000+ eye-tracker. Eye movements were recorded monocularly and head movement was minimized using a desk-mounted chinrest. A 9-point grid calibration procedure was conducted before the experiment. Each screen was additionally preceded by a fixation point for drift correction. The eye-tracker was re-calibrated at least once during the experiment. The instructions stated that the sentences should be read as naturally as possible for comprehension and that the space bar should be pressed to go to the next screen. The sentences were presented across one line in random order, and the target positive and neutral words did not appear at the beginning or end of a line. A third of the sentences were followed by a comprehension question, and the scores indicated that the participants had attended to the text (average percentile score: 97.56%, SD = 4.61). Before leaving the lab, the participants completed a short language-background questionnaire.

4.4. Analysis

Trials were eliminated from the data analysis if the eye tracker lost track of the eye. Following Knickerbocker et al. (2015), we applied a two-step cleaning procedure in the DataViewer software. First, adjacent fixations (less than 0.5 degrees apart) were merged if one or both of them were short (less than 80 ms). Second, single fixations shorter than 100 ms and longer than 800 ms were removed (3.28% of all fixations).

The analysis was conducted in R version 4.0.5 (Team R Developement Core 2021). Separate models were constructed for 13 eye-movement measures examining early and late eye movements on the positive/neutral word (target region) and on the post-target region. Similar to Knickerbocker et al. (2015), the post-target region was operationalized as two words following the target word. Table 2 presents a list of the eye-movement measures and their specified features. Binary (0/1) outcome values were analyzed using a mixed-logit regression (or, mixed-effects logistic regression) analysis (*glmer* function in the *lme4* package). For continuous dependent variables, we employed linear mixed-effects (LME) models (*lmer* function in the *lme4* package).

Reading times were log-transformed to reduce skewness in the data. All 13 analyses (one for each eye-movement measure) adopted the maximal random-effects structure justified by the design. The main target variable was Word Type, with *neutral* as the reference level. We also included various variables that might modulate differences between positive and neutral words: Trial; Log vocabulary test score; STAI score; BDI score; Age; Arousal; Length; and Log word frequency. All models included random intercepts for subjects and items, as well as by-subject random slopes for Word Type. Seven 2-way interactions were then added stepwise, one-by-one (between Word Type and each of the other variables, except Trial), to examine any modulating effect of these variables on the difference between positive and neutral words. Log-likelihood tests and AIC values were used to compare the resulting model with the original model, and only interactions that significantly improved the model fit were retained[1].

All continuous variables were centered. Table 3 presents a summary of the continuous variables. The final models were checked for collinearity, and no issues were observed, with a variance inflation factor (VIF) of around 3.

Table 2. Summary of the eye-movement measures.

Measure	Variable Type	Description	Cleaning Details
		Early measures (target region)	
(a) First fixation duration	Log transformed continuous variable	The duration of the first fixation that readers made on the target word	Exclude skipped trials
(b) Single fixation duration	Log transformed continuous variable	The duration of the first fixation on the target word given that the reader made only one fixation on it during first pass reading	Exclude skipped trials
(c) First pass reading time	Log transformed continuous variable	The amount of time the reader spent on the target word before leaving it to the left or right	Exclude skipped trials
(d) Landing position	Log transformed continuous variable	The distance from the beginning of the word, in pixel positions, that the reader landed on the target word	Exclude skipped trials
(e) Skipping rate	Binary (0/1) variable, 1 = skipped	The percentage of trials in which the reader skipped the target word during first pass reading	None
		Late measures (target region)	
(f) Total time	Log transformed continuous variable	The total time spent on the target word, including any regressions back to the target word after leaving it	Exclude skipped trials
(g) Regressions in	Binary (0/1) variable, 1 = regression occurred	The percentage of trials in which the reader went back to the target word after leaving it	Exclude skipped trials
(h) Second pass time	Log transformed continuous variable (Log of reading time +1)	The amount of time spent on the second pass reading of the target word	Skipped trials coded as 0 ms
		Post-target region	
(i) Spillover	Log transformed continuous variable	The duration of the first fixation made after leaving the target region	Exclude skipped trials
(j) First fixation duration	Log transformed continuous variable	The duration of the first fixation that readers made on the post-target word	Exclude skipped trials
(k) First pass reading time	Log transformed continuous variable	The amount of time the reader spent on the post-target region before leaving it to the left or right	Exclude skipped trials
(l) Total time	Log transformed continuous variable	The total time spent on the target word, including any regressions back to the post-target region after leaving it	Exclude skipped trials
(m) Regressions out	Binary (0/1) variable, 1 = regression occurred	The percentage of regressions made out of the post-target region	Exclude skipped trials

Table 3. Summary of continuous variables (Experiment 1).

Variable	Range (Adjusted Range)	Mdn
Trial	28.00–83.00 (−1.68–1.71) trials	−0.02
Log vocabulary test score	4159–8966 (−2.69–2.00) points	0.19
STAI score	21.00–66.00 (−1.45–2.99) points	−0.17
BDI score	1.00–46.00 (−1.17–3.86) points	−0.16
Age	18.00–38.00 (−1.06–2.77) years	−0.48
Arousal	1.65–3.35 (−1.48–2.44) points	−0.1
Length (no. of letters)	3.00–10.00 (−1.88–2.25) characters	−0.11
Word frequency (log HAL)	88–165,830 (−2.77–2.28) points	0.005

Note. The second column shows the range of the variables. The adjusted range after transformation and/or centering, is presented in parentheses. Medians refer to the predictor values in the models. All variables are centered, and their means are zero.

4.5. Results and Discussion

Table 4 presents the means (continuous variables) and percentage (binary variables) for neutral and positive words. No evident direction is observed; while most of the measures exhibited the expected advantage for positive over neutral words (shorter RT for the former than the latter), the remaining measures showed a processing disadvantage.

Table 4. Mean of each eye-movement measure in Experiment 1.

	Neutral		Positive	
Measure	**Mean**	**SE**	**Mean**	**SE**
Early measures (target region)				
(a) First fixation duration	265.24	3.5	263.79	3.39
(b) Single fixation duration	272.82	4.33	266.21	4.16
(c) First pass reading time	360.21	7.38	376.15	7.82
(d) Landing position	39.17	0.78	41.2	0.79
(e) Skipping rate *	6.74%	—	8.02%	—
Late measures (target region)				
(f) Total time	515.27	12.37	511.8	12.13
(g) Regressions in *	17.10%	—	13.75%	—
(h) Second pass time	105.68	6.05	99.14	6.34
Post-target region				
(i) Spillover	264.02	3.49	264.56	3.56
(j) First fixation duration	260.96	3.36	261.94	3.42
(k) First pass reading time	517.96	10.28	514.83	9.64
(l) Total time	708.09	15.31	685.48	14.35
(m) Regressions out *	16.78%	—	13.47%	—

* Percentages are presented for binary data including skipping rate and regression.

The results of the mixed-effects modelling are presented in Table 5. It is evident that Word Type was not a significant main predictor for any of the 13 eye-movement measures (early, late, or post-target). However, Word Type interacted with several predictor variables. The significant interactions are presented in Figures 1–5.

Table 5. Results of mixed-effects modelling for each eye-movement measure (Experiment 1).

Early Measures (Target Region)																				
	(a) First Fixation Duration				**(b) Single Fixation Duration**				**(c) First Pass Reading Time**				**(d) Landing Position**				**(e) Skipping Rate**			
	Estimate	t value	*p*		Estimate	t value	*p*		Estimate	t value			Estimate	t value	*p*		Estimate	z value	*p*	
(Intercept)	5.50	216.77	<0.001	***	5.56	192.26	<0.001	***	5.75	153.87	<0.001	***	3.47	68.76	<0.001	***	−3.74	−10.70	<0.001	***
Trial	0.00	−0.25	0.80		0.00	0.17	0.86		0.00	−0.19	0.85		0.02	0.94	0.35		0.03	0.33	0.74	
Log vocabulary test score	−0.06	−3.44	0.001	**	−0.08	−3.15	0.003	**	−0.15	−4.22	<0.001	***	−0.01	−0.36	0.72		0.34	1.24	0.21	
STAI score	0.00	0.08	0.94		−0.03	−0.91	0.37		−0.06	−1.30	0.20		−0.04	−1.04	0.30		0.31	0.83	0.41	
BDI score	0.01	0.29	0.77		0.04	1.21	0.24		0.06	1.27	0.21		0.05	1.20	0.24		−0.07	−0.21	0.83	
Age	0.06	3.63	0.001	**	0.06	2.68	0.01	*	0.06	1.80	0.08		−0.01	−0.42	0.68		−0.44	−1.68	0.09	
Arousal	−0.02	−1.90	0.07		−0.02	−1.82	0.08		−0.04	−2.56	0.01	*	0.03	0.82	0.42		0.06	0.54	0.59	
Length (no. of letters)	0.02	2.26	0.03	*	0.04	2.81	0.007	**	0.12	7.91	<0.001	***	0.10	3.05	0.004	**	−0.64	−4.87	<0.001	***
Word frequency (log HAL)	0.00	−0.17	0.87		−0.05	−3.88	<0.001	***	−0.06	−4.27	<0.001	***	0.00	−0.12	0.91		0.09	0.81	0.42	
Word Type: Positive	0.01	0.24	0.81		0.00	−0.19	0.85		0.02	0.51	0.61		0.00	−0.05	0.96		0.52	1.60	0.11	
Word frequency (log HAL) × Word Type	−0.05	−2.34	0.02	*	—	—	—		—	—	—		—	—	—		—	—	—	—
Random effects:	Variance				Variance				Variance				Variance				Variance			
Subject	0.02				0.02				0.04				0.03				2.10			
Subject ǀ Word type	0.00				0.00				0.00				0.01				0.06			
Item	0.00				0.00				0.00				0.02				0.00			
Residual	0.11				0.11				0.16				0.66				—			

Late Measures (Target Region)												
	(f) Total Time				**(g) Regressions in**				**(h) Second Pass Time**			
	Estimate	t value	*p*		Estimate	z value	*p*		Estimate	t value	*p*	
(Intercept)	5.97	113.08	<0.001	***	−1.91	−10.08	<0.001	***	1.61	7.55	<0.001	***
Trial	0.00	−0.38	0.70		0.03	0.41	0.68		−0.05	−0.86	0.39	
Log vocabulary test score	−0.20	−3.97	<0.001	***	−0.04	−0.33	0.74		−0.32	−1.83	0.08	
STAI score	−0.02	−0.29	0.78		0.39	2.20	0.03	*	0.30	1.28	0.21	
BDI score	0.02	0.34	0.73		−0.29	−1.66	0.10		−0.31	−1.35	0.18	
Age	0.11	2.23	0.03	*	0.23	1.83	0.07		0.37	2.16	0.04	*
Arousal	−0.14	−4.40	<0.001	***	−0.12	−1.01	0.31		−0.47	−2.54	0.02	*
Length (no. of letters)	0.11	7.15	<0.001	***	−0.13	−1.11	0.27		−0.01	−0.14	0.89	
Word frequency (log HAL)	−0.10	−6.33	<0.001	***	−0.04	−0.39	0.70		−0.17	−1.99	0.054	
Word Type: Positive	0.03	0.76	0.45		−0.12	−0.49	0.63		0.06	0.33	0.75	
Arousal × Word Type	0.12	3.07	<0.001	***	—	—	—	—	0.47	2.19	0.04	*
BDI score × Word Type	−0.05	−1.98	0.047	*	—	—	—	—	—	—		

Table 5. *Cont.*

Late Measures (Target Region)			
	(f) Total Time	**(g) Regressions in**	**(h) Second Pass Time**
Random effects:	Variance	Variance	Variance
Subject	0.09	0.36	1.05
Subject∣Word type	0.00	0.00	0.00
Item	0.00	0.23	0.12
Residual	0.23	—	5.62

Post-Target Region																				
	(i) Spillover				**(j) First Fixation Duration**				**(k) First Pass Reading Time**				**(l) Total Time**				**(m) Regressions Out**			
	Estimate	t value	*p*		Estimate	t value	*p*		Estimate	t value	*p*		Estimate	t value	*p*		Estimate	z value	*p*	
(Intercept)	5.51	191.54	<0.001	***	5.5	194.2	<0.001	***	6.09	98.15	<0.001	***	6.35	83.50	<0.001	***	−1.94	−10.19	<0.001	***
Trial	0.00	−0.29	0.77		0.0	0.4	0.7		−0.01	−1.13	0.26		−0.01	−1.33	0.18		0.02	0.34	0.73	
Log vocabulary test score	−0.02	−0.99	0.33		0.0	−1.0	0.3		−0.12	−2.93	0.006	**	−0.14	−2.71	0.01	*	−0.06	−0.48	0.63	
STAI score	0.01	0.29	0.77		0.0	0.5	0.6		−0.06	−1.16	0.25		0.02	0.27	0.79		0.39	2.18	0.03	*
BDI score	0.00	0.02	0.98		0.0	0.0	1.0		0.01	0.12	0.91		−0.05	−0.68	0.50		−0.27	−1.57	0.12	
Age	0.04	1.76	0.09		0.0	1.9	0.1		0.07	1.84	0.07		0.12	2.38	0.02	*	0.24	1.90	0.06	
Arousal	−0.01	−0.38	0.71		0.0	−0.4	0.7		0.02	0.40	0.69		−0.01	−0.17	0.86		−0.13	−1.09	0.27	
Length (no. of letters)	0.01	0.78	0.44		0.0	1.0	0.3		0.00	0.07	0.95		−0.03	−0.70	0.49		−0.11	−0.96	0.34	
Word frequency (log HAL)	0.00	−0.14	0.89		0.0	−0.2	0.8		0.00	0.14	0.89		0.02	0.55	0.59		−0.06	−0.50	0.62	
Word Type: Positive	0.00	0.12	0.91		0.0	0.1	0.9		−0.01	−0.09	0.93		0.01	0.07	0.94		−0.11	−0.45	0.65	
BDI score x Word Type	—	—	—		—	—	—		0.04	2.07	0.04	*	—	—	—		—	—	—	
Random effects:	Variance				Variance				Variance				Variance				Variance			
Subject	0.01				0.02				0.06				0.10				0.39			
Subject∣Word type	0.00				0.00				0.00				0.00				0.01			
Item	0.01				0.01				0.04				0.06				0.22			
Residual	0.11				0.11				0.20				0.18				—			

*, ** and ***: statistical significance.

Figure 1 shows that the usual advantage for positive words over neutral words is only evident in the first fixation duration measure for higher-frequency words; that is, positive words were read faster than neutral words only when these were frequent in the language.

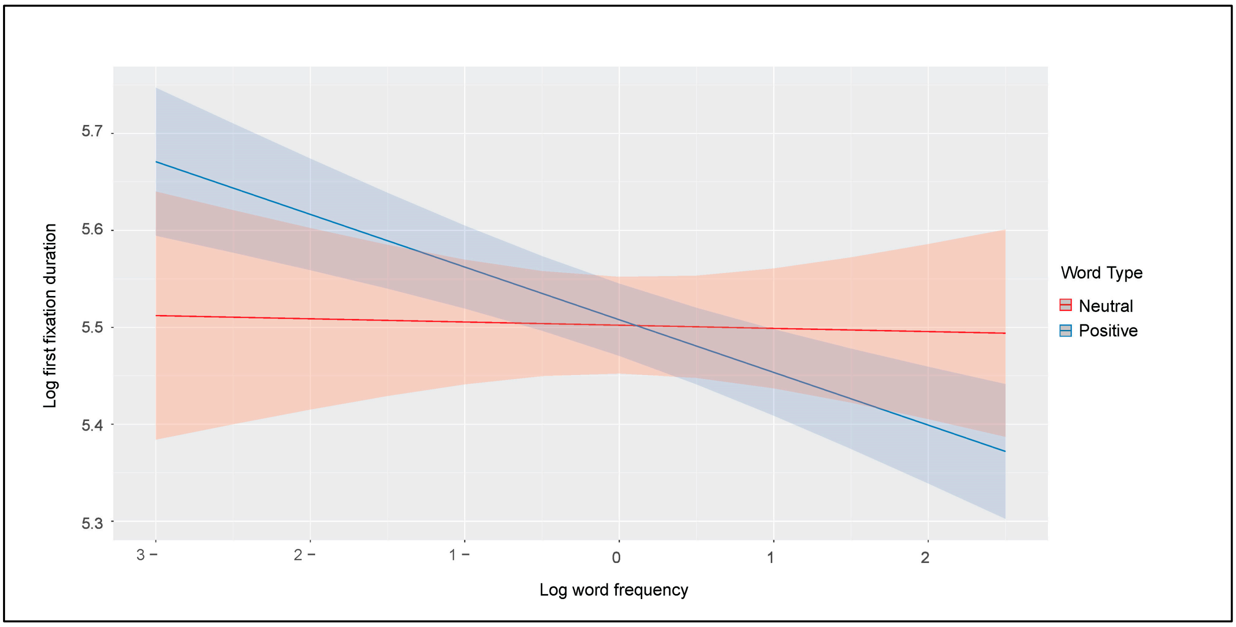

Figure 1. A graphical presentation of the interaction between word frequency and Word Type for the first fixation duration measure (target region, Experiment 1). Shading shows 95 per cent confidence intervals (CI).

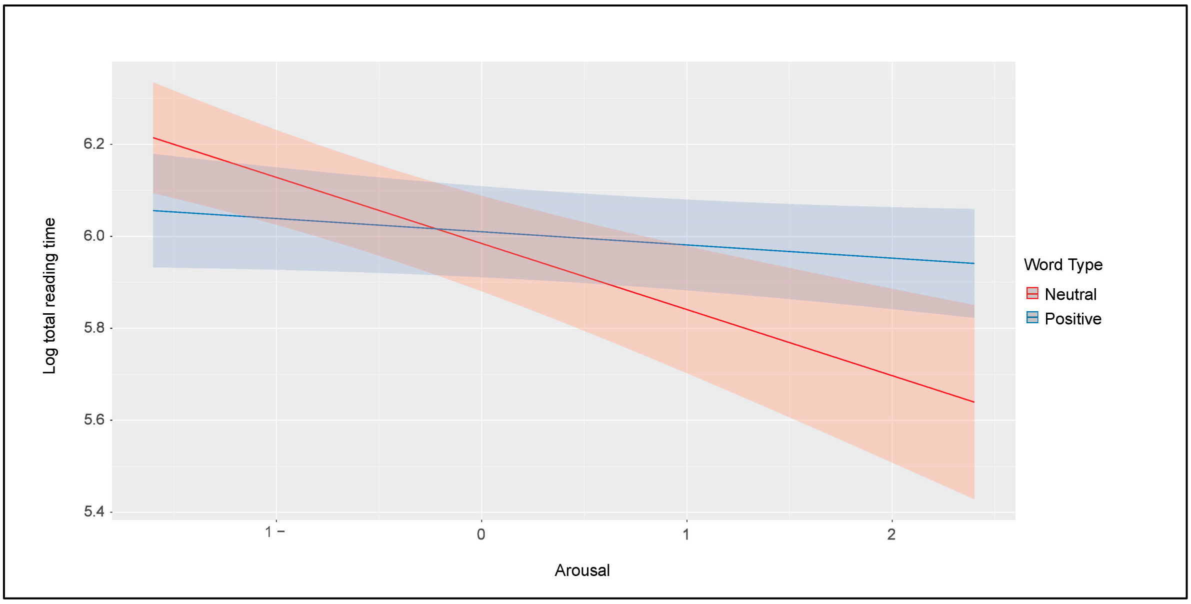

Figure 2. A graphical presentation of the interaction between arousal score and Word Type for the total reading time measure (target region, Experiment 1). Shading shows 95 per cent confidence intervals (CI).

The interactions depicted in Figures 2 and 3 seem to suggest that the advantage in the total reading time for positive words over neutral ones is evident for words with a low arousal score and for participants with a high BDI score, respectively. Arousal also seems to have an effect on the second pass reading time measure (Figure 4), again with an advantage for positive words with a low arousal score. This same effect is also evident for the first pass reading time in the post-target region (Figure 5). It is noteworthy that as the arousal scores increase for all three measures, positive words show the opposite effect; that is, a disadvantage in processing compared to neutral items.

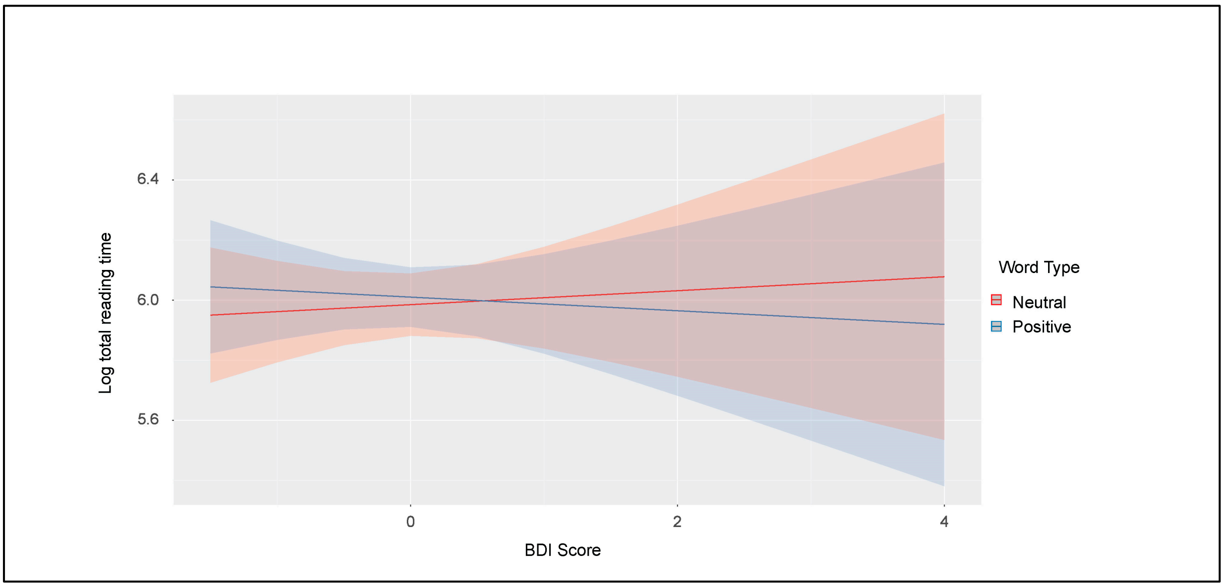

Figure 3. A graphical presentation of the interaction between the BDI score and Word Type for the total reading time measure (target region, Experiment 1). Shading shows 95 per cent confidence intervals (CI).

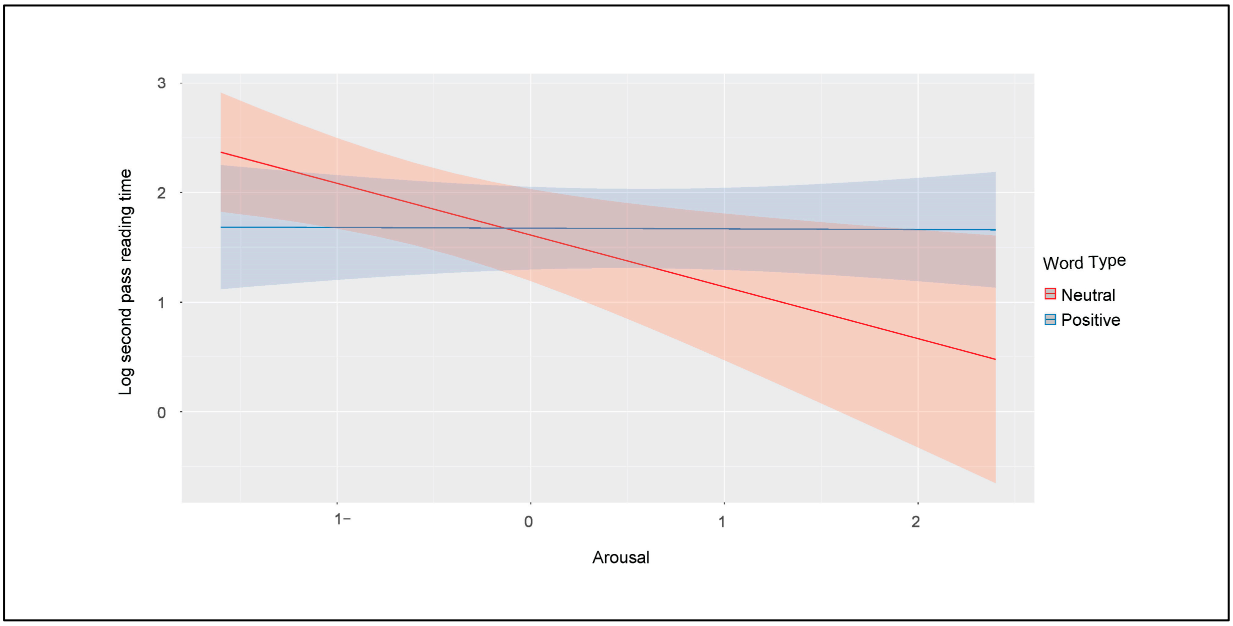

Figure 4. A graphical presentation of the interaction between arousal score and Word Type for the second pass reading time measure (target region, Experiment 1). Shading shows 95 per cent confidence intervals (CI).

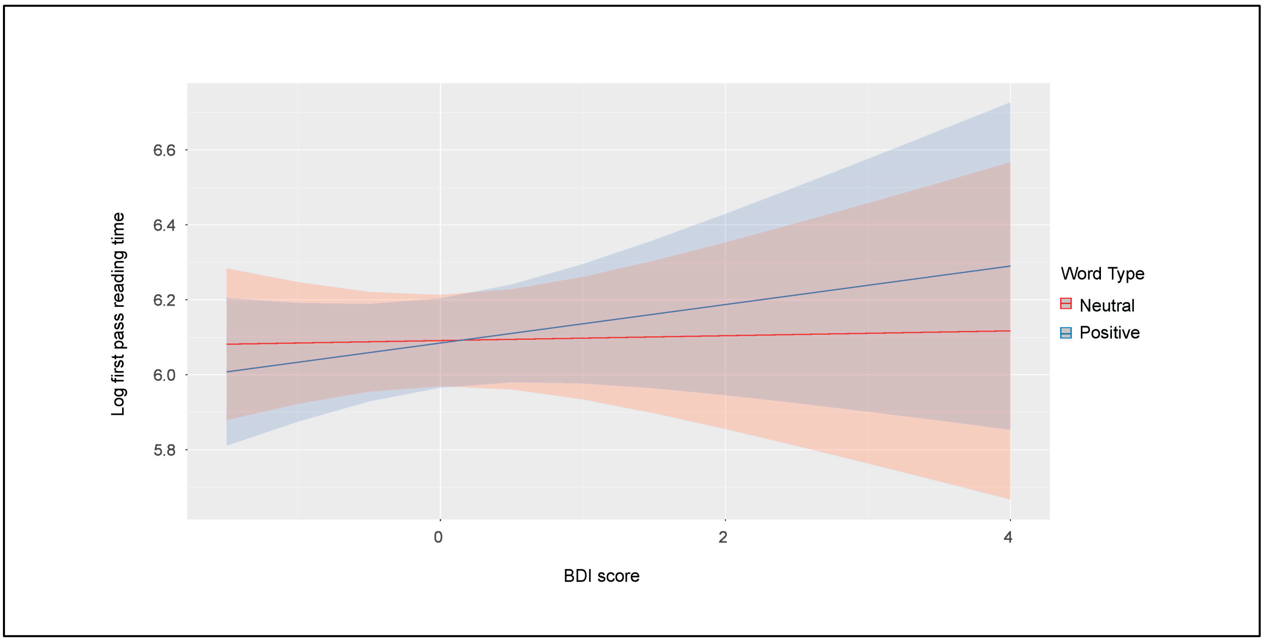

Figure 5. A graphical presentation of the interaction between the BDI score and Word Type for the first pass reading time measure (post-target region, Experiment 1). Shading shows 95 per cent confidence intervals (CI).

5. Experiment 2

5.1. Participants

Forty-three L1 Arabic—L2 English speakers participated in Experiment 2 (female = 29, male = 14). None of them had participated in Experiment 1. Their average age was 29.70 years (*Min* = 18, *Max* = 62, *SD* = 9.73). Similar to Experiment 1, age was added as a covariate in the analysis. They started learning English mostly in educational settings, at an average age of six (*M* = 6.30, *SD* = 4.00). The participants' average self-reported proficiency across the four skills was 4.34 (*SD* = 0.52).

The mean score achieved by the participants in the V_YesNo was 6492.37 (Min = 4608, Max = 8601, SD = 1147.28), indicating a good-to-high level of proficiency. The vocabulary scores were included as a factor in the analysis. The participants were randomly presented with one version of the experiment (Version 1, n = 21 and Version 2, n = 22).

5.2. Design and Stimuli

The items for Experiment 2 were borrowed from Knickerbocker et al. (2015), following the same procedures as Experiment 1. We selected the pairs with the same neutral words as the ones included in Experiment 1. The 21 negative/neutral pairs are presented in Appendix B. Here is an example for the pair (bake/hate):

I am pretty sure that Kevin will bake/hate the sugar cookie mix from the grocery store.

Every holiday, I return home and bake/hate the dinner that the family eats together.

5.3. Procedures

The procedures followed those in Experiment 1. The average percentile score in the comprehension question was high (*M* = 96.68%, *SD* = 4.23).

5.4. Analysis

The data cleaning and analysis followed the procedures described in Experiment 1. Only 2.49% of the data points were lost due to the cleaning procedure. Table 6 presents a summary of the continuous variables. The collinearity was checked for significant predictors in each model using the VIF. All values were around 3, indicating no collinearity issues.

Table 6. Summary of continuous variables (Experiment 2).

Variable	Range (Adjusted Range)	Mdn
Trial	28.00–83.00 (−1.71–1.71) trials	−0.03
Log vocabulary test score	4608–8601 (−1.89–1.71) points	−0.19
STAI score	20.00–61.00 (−1.49–2.55) points	−0.01
BDI score	0.00–33.00 (−1.43–3.07) points	−0.21
Age	18.0–62.0 (−1.00–3.57) years	−0.48
Arousal	3.15–8.17 (−1.61–2.50) points	−0.21
Length (no. of letters)	3.00–10.00 (−1.70–2.27) characters	−0.0003
Word frequency (log HAL)	473–44,285 (−2.21–1.64) points	0.15

Note. The second column shows the range of the variables. The adjusted range after transformation and/or centering, is presented in parentheses. Medians refer to the predictor values in the models. All variables are centered, and their means are zero.

5.5. Results and Discussion

The means and percentages for continuous and binary variables, respectively, are presented in Table 7. The emotion-word advantage is less clearly evident in Experiment 2 than in Experiment 1. Almost all of the eye-movement measures showed a disadvantage for negative words in comparison to neutral ones, with longer reading times in the former than the latter.

Table 7. Mean of each eye-movement measure in Experiment 2.

	Neutral		Negative	
Measure	**Mean**	**SE**	**Mean**	**SE**
Early measures (target region)				
(a) First fixation duration	286.94	3.84	288.64	3.96
(b) Single fixation duration	292.63	4.4	296.68	4.97
(c) First pass reading time	406.12	9.44	423.64	9.57
(d) Landing position	40.45	0.87	41.25	0.86
(e) Skipping rate *	6.01%	—	7.32%	—
Late measures (target region)				
(f) Total time	568.27	15.16	573.43	15.88
(g) Regressions in *	16.94%	—	15.85%	—
(h) Second pass time	114.66	7.03	106.71	7.4
Post-target region				
(i) Spillover	266.66	3.67	272.57	3.74
(j) First fixation duration	267.3	3.52	271.41	3.57
(k) First pass reading time	501.05	10.46	501.27	10.68
(l) Total time	678.06	15.1	659.51	15.29
(m) Regressions out *	15.90%	—	15.74%	—

* Percentages are presented for binary data including skipping rate and regression.

Looking now at the results of the mixed-effects modelling in Experiment 2 (Table 8), we find that Word Type was significant as a main variable for only one measure, which is the skipping-rate, with a medium effect (*odds ratio* = 0.29, $d = -0.67$). This result seems to suggest that the odds of skipping a neutral word were over three times (1/0.29) larger than those of skipping a negative word. It should be noted, however, that this effect is further modulated by word length (see Figure 6). Negative words are more likely to be skipped when they are short, but the difference seems to diminish for longer words.

Table 8. Results of mixed-effects modelling for each eye-movement measure (Experiment 2).

Early Measures (Target Region)																				
	(a) First Fixation Duration				**(b) Single Fixation Duration**				**(c) First Pass Reading Time**				**(d) Landing Position**				**(e) Skipping Rate**			
	Estimate	t value	*p*		Estimate	t value	*p*		Estimate	t value	*p*		Estimate	t value	*p*		Estimate	z value	*p*	
(Intercept)	5.59	289.76	<0.001	***	5.64	254.13	<0.001	***	5.80	138.89	<0.001	***	3.49	75.48	<0.001	***	−3.12	−11.51	<0.001	***
Trial	0.02	1.86	0.06		0.01	1.32	0.19		0.00	0.50	0.62		−0.01	−0.49	0.62		−0.11	−1.04	0.30	
Log vocabulary test score	−0.11	−6.37	<0.001	***	−0.13	−6.81	<0.001	***	−0.24	−8.08	<0.001	***	0.15	5.32	<0.001	***	0.69	3.85	<0.001	***
STAI score	0.01	0.42	0.68		0.00	0.20	0.84		0.03	1.06	0.29		−0.06	−1.85	0.07		0.13	0.80	0.43	
BDI score	0.00	0.08	0.94		0.00	0.15	0.88		0.00	−0.10	0.92		−0.01	−0.15	0.88		−0.28	−1.44	0.15	
Age	0.04	2.11	0.04	*	0.04	1.91	0.06		0.07	2.34	0.02	*	−0.07	−2.35	0.02	*	−0.34	−1.95	0.05	
Arousal	−0.01	−0.53	0.60		0.00	−0.14	0.89		−0.09	−2.25	0.03	*	0.02	0.72	0.48		0.44	2.79	0.005	**
Length (no. of letters)	0.00	−0.34	0.74		0.00	0.23	0.82		0.07	3.72	<0.001	***	0.10	3.20	0.003	**	−0.74	−2.78	0.005	**
Word frequency (log HAL)	0.00	−0.43	0.67		−0.01	−0.99	0.33		−0.07	−3.82	<0.001	***	0.01	0.34	0.74		0.07	0.43	0.67	
Word Type: Negative	0.01	0.37	0.71		0.00	0.14	0.89		0.06	1.35	0.19		−0.03	−0.38	0.71		−1.22	−2.43	0.02	*
Arousal × Word Type	—-	—-	—-		—-	—-	—-		0.12	2.47	0.02	*	—-	—-	—-		—-	—-	—-	—-
Length × Word Type	—-	—-	—-		—-	—-	—-		—-	—-	—-		—-	—-	—-		−0.93	−2.16	0.03	*
Random effects:	Variance				Variance				Variance				Variance				Variance			
Subject	0.00				0.01				0.02				0.01				0.44			
Subject ∣ Word type	0.00				0.00				0.00				0.00				0.00			
Item	0.00				0.00				0.01				0.01				0.14			
Residual	0.11				0.10				0.17				0.65				—-			
Late Measures (Target Region)																				
	(f) Total Time				**(g) Regressions in**				**(h) Second Pass Time**											
	Estimate	t value	*p*		Estimate	z value	*p*		Estimate	t value	*p*									
(Intercept)	6.05	97.71	<0.001	***	−1.90	−9.43	<0.001	***	1.82	8.65	<0.001	***								
Trial	−0.02	−1.68	0.09		0.08	1.18	0.24		−0.06	−1.10	0.27									
Log vocabulary test score	−0.26	−5.82	<0.001	***	−0.04	−0.33	0.74		−0.27	−1.61	0.12									
STAI score	0.04	0.91	0.37		0.18	1.43	0.15		0.14	0.76	0.45									
BDI score	−0.09	−1.64	0.11		−0.42	−3.03	0.002	**	−0.50	−2.48	0.02	*								
Age	0.10	2.21	0.03	*	0.14	1.21	0.23		0.21	1.18	0.24									
Arousal	−0.12	−2.25	0.03	*	0.01	0.05	0.96		−0.18	−1.40	0.17									
Length (no. of letters)	0.04	1.51	0.14		−0.52	−4.01	<0.001	***	−0.15	−1.32	0.19									
Word frequency (log HAL)	−0.13	−4.69	<0.001	***	−0.42	−3.35	<0.001	***	−0.29	−2.70	0.01	*								
Word Type: Negative	0.05	0.89	0.38		−0.06	−0.19	0.85		0.07	0.28	0.78									
Arousal × Word Type	0.14	2.11	0.04	*	—-	—-	—-	—-	—-	—-	—-									

Table 8. *Cont.*

Late Measures (Target Region)																				
	(f) Total Time				**(g) Regressions in**				**(h) Second Pass Time**											
Random effects:	Variance				Variance				Variance											
Subject	0.07				0.27				0.86											
Subject ∣ Word type	0.00				0.04				0.01											
Item	0.02				0.26				0.23											
Residual	0.19				—				5.59											
Post-Target Region																				
	(i) Spillover				**(j) First Fixation Duration**				**(k) First Pass Reading Time**				**(l) Total Time**				**(m) Regressions Out**			
	Estimate	t value	*p*		Estimate	t value	*p*		Estimate	t value	*p*		Estimate	t value	*p*		Estimate	z value	*p*	
(Intercept)	5.51	209.57	<0.001	***	5.51	213.24	<0.001	***	6.00	104.62	<0.001	***	6.30	97.15	<0.001	***	−1.98	−9.47	<0.001	***
Trial	−0.01	−0.81	0.42		−0.01	−0.99	0.32		−0.02	−2.09	0.04	*	−0.02	−1.98	0.048	*	0.09	1.32	0.19	
Log vocabulary test score	−0.07	−3.36	0.002	**	−0.07	−3.78	<0.001	***	−0.26	−7.23	<0.001	***	−0.25	−5.89	<0.001	***	−0.04	−0.31	0.76	
STAI score	0.02	0.87	0.39		0.02	1.11	0.27		0.03	0.79	0.43		0.03	0.72	0.48		0.18	1.44	0.15	
BDI score	0.01	0.56	0.58		0.01	0.54	0.59		0.02	0.58	0.57		−0.05	−1.06	0.30		−0.42	−2.98	0.003	**
Age	0.04	2.18	0.04	*	0.05	2.52	0.02	*	0.09	2.50	0.02	*	0.13	2.89	0.006	**	0.15	1.28	0.20	
Arousal	−0.01	−0.84	0.41		−0.01	−0.71	0.49		−0.06	−1.57	0.13		−0.05	−1.18	0.25		0.01	0.06	0.95	
Length (no. of letters)	0.01	0.91	0.37		0.02	1.12	0.27		−0.01	−0.38	0.71		−0.06	−1.54	0.13		−0.52	−3.83	<0.001	***
Word frequency (log HAL)	0.01	0.67	0.51		0.01	0.84	0.40		0.03	1.00	0.33		−0.01	−0.21	0.83		−0.47	−3.62	<0.001	***
Word Type: Negative	0.04	1.13	0.27		0.03	1.00	0.32		0.09	1.08	0.29		0.04	0.48	0.64		−0.02	−0.06	0.95	
Random effects:	Variance				Variance				Variance				Variance				Variance			
Subject	0.01				0.01				0.04				0.06				0.29			
Subject ∣ Word type	0.00				0.00				0.00				0.00				0.13			
Item	0.00				0.00				0.03				0.04				0.31			
Residual	0.11				0.11				0.21				0.18				—-			

*, ** and ***: statistical significance.

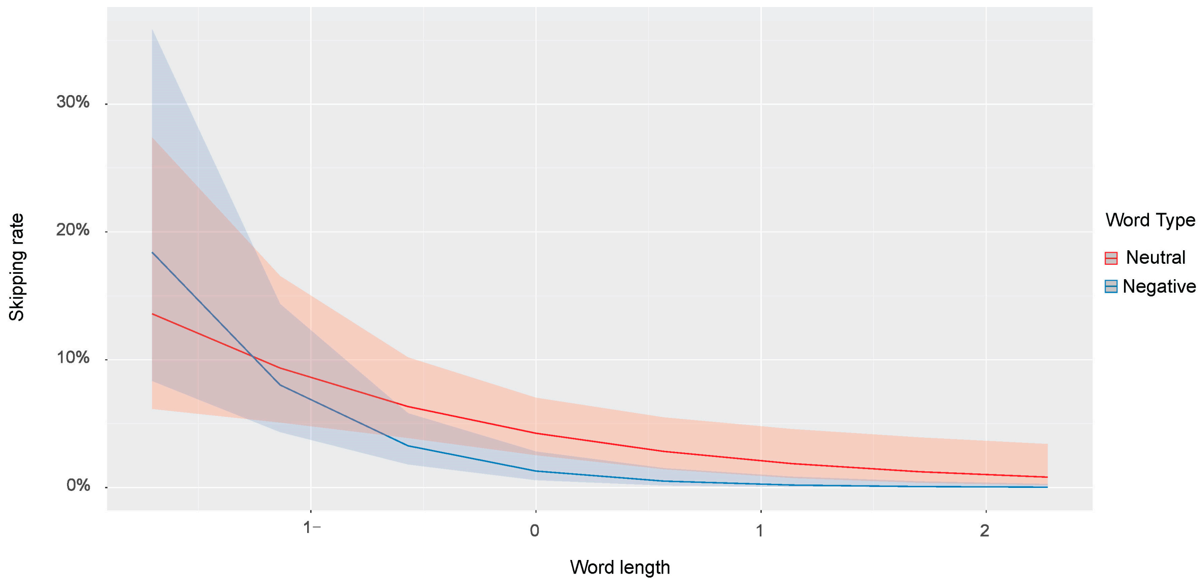

Figure 6. A graphical presentation of the interaction between word length and Word Type for the skipping rate measure (target region, Experiment 2). Shading shows 95 per cent confidence intervals (CI).

Finally, the effect of arousal, which was evident in Experiment 1 (whereby the advantage for emotion words was evident only for words with lower arousal score), was also demonstrated in one early (first pass reading time) and one late (total reading time) measure on the target region (see Figures 7 and 8). The BDI and STAI had no significant main or interreacting effects on the difference between negative and neutral words.

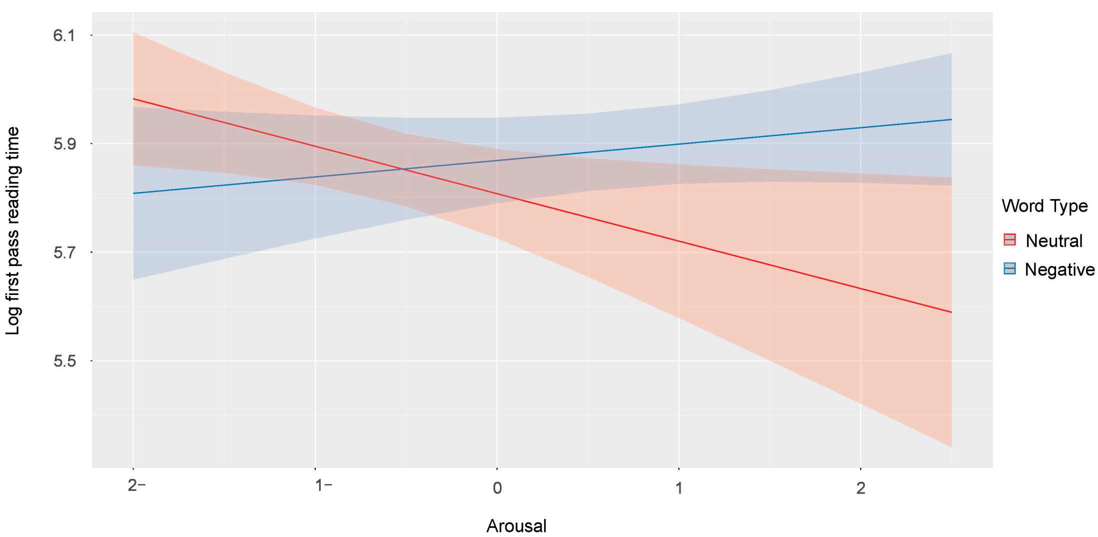

Figure 7. A graphical presentation of the interaction between arousal score and Word Type for the first pass reading time measure (target region, Experiment 2). Shading shows 95 per cent confidence intervals (CI).

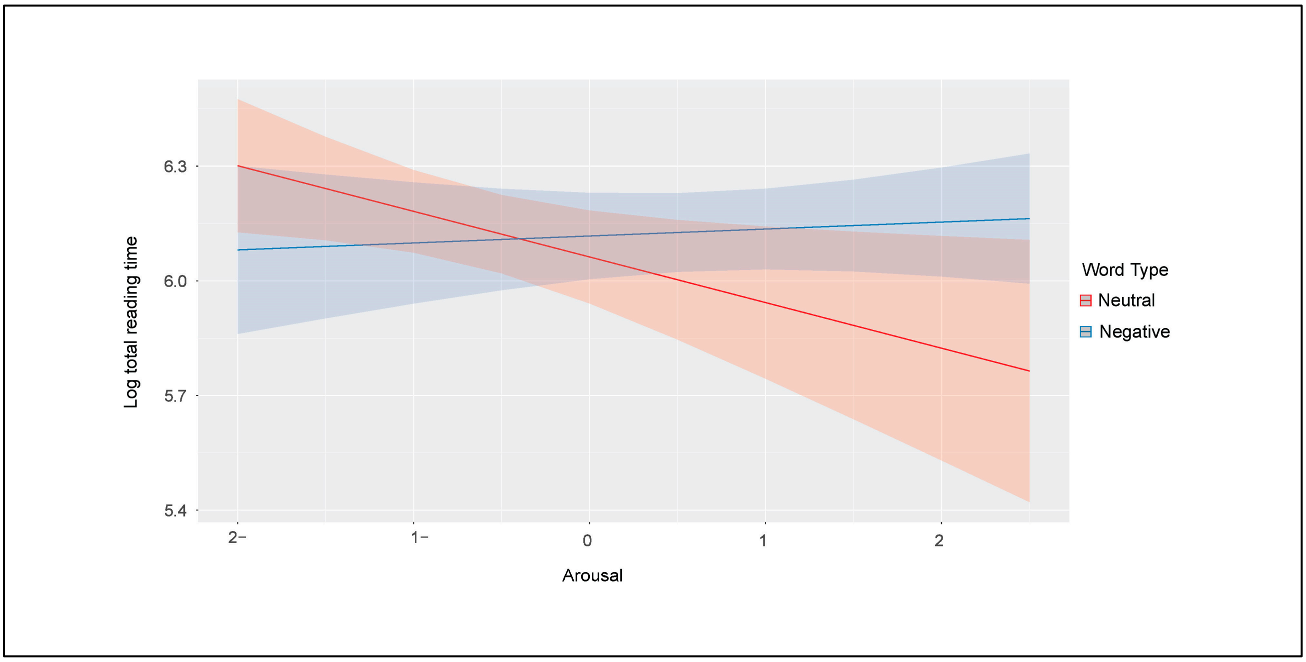

Figure 8. A graphical presentation of the interaction between arousal score and Word Type for the total reading time measure (target region, Experiment 2). Shading shows 95 per cent confidence intervals (CI).

6. Discussion

The current study aimed to examine the processing of emotionally positive and negative words versus neutral words in normal silent reading by Arab EFL learners. To this end, two experiments were conducted in which the participants read English sentences with embedded target words while their eye movements were monitored and recorded against 13 measures. Experiment 1 was designed to address Research Question 1 (RQ1), concerning whether positive emotion words display an advantage over neutral words, and Research Question 2 (RQ2), regarding the modulating effects of item-related and participant-related features; on the other hand, Experiment 2 was designed to respond to RQ3, concerning the advantage of negative emotion words when compared with neutral words and, again, the modulating effects of item-related and participant-related features. The results of the two experiments failed to reveal any significant advantage for word type (but see Experiment 2 for a main effect of word type on the skipping rate of negative words) against the 13 eye-movement measures.

This finding is in contrast to the confirmed processing advantage for emotion words in earlier studies in word processing (Ferré 2002; Kousta et al. 2009; Kazanas and Altarriba 2015b) and in normal reading (Knickerbocker et al. 2015; Scott et al. 2012). However, this should not come as a surprise as emotion words can behave differently in bilinguals, particularly for late bilinguals who learned their L2 in instructional settings, as is the case in the present study (Ferré et al. 2018). According to Kroll and Stewart's (1994) RHM, the strong conceptual connections between L2 words and their meanings may require additional meaningful exposure before L2 learners can access meaning directly, and thus behave similar to L1 speakers. Similarly, Jiang (2000) predicted that very few advanced language learners can escape the fossilization stage and deal with L2 vocabulary as naturally as native speakers. Additionally, the L2, particularly in foreign language learning contexts, is known to be emotionally distant (Caldwell-Harris 2014; Chen et al. 2015).

Despite failing to show any significant main effects for word type, Experiment 1 showed an interaction between word type and frequency. Positive words were only read faster than neutral words in the case of highly frequent words. This finding can again be interpreted in light of the RHM. Highly frequent L2 words are likely to access meaning through direct conceptual links due to their recurrent use. Hence, highly frequent words may connect with their concepts directly without L1 mediation and can thus mirror the behavior of L1 emotion words.

Similarly, Experiment 2 revealed a significant main effect of word type on the skipping rate, which was further modulated by length. Negative words were less likely to be skipped than neutral words (with a medium effect), particularly when they were shorter. It is well-known that adult readers skip approximately one-third of the words in a text as they read (Rayner 1998, 2009). It is also widely acknowledged that shorter words are more likely to be skipped than longer ones (Brysbaert and Vitu 1998; Rayner and McConkie 1976; Vitu et al. 1995). This can be related to *parafoveal lexical processing*, whereby shorter words are preprocessed while they are still at the parafoveal view, leading to a higher skipping rate (see Schotter 2018 for an overview of the evidence). In addition to word length, a number of other contributing factors have also been investigated, such as predictability from context (Drieghe et al. 2005; Gollan et al. 2011) and high word frequency (Angele et al. 2014; Rayner and Fischer 1996). The finding that negative words may be less likely to be skipped than neutral words is novel. Could this be a result of automatic vigilance (Pratto and John 1991)? That is, could negative words be more resistant to skipping because they require prolonged attention? Further research is needed to investigate this finding in depth.

In addition to word frequency and skipping, Experiments 1 and 2 showed a negative correlation between emotion words and arousal. Emotionally positive and negative words were only read faster than neutral words when they were low in arousal. In contrast, emotion words suffered from a processing disadvantage when the words were high in arousal. This finding is similar to previous studies in word recognition in which arousing words were recognized more slowly than calming words (Barriga-Paulino et al. 2022; Kuperman et al. 2014). This could again be interpreted in light of the automatic vigilance model of emotion (Pratto and John 1991) as highly aroused words may hold attention longer than words with low arousal. It must be noted, however, that Knickerbocker et al. (2015) also found a modulating effect for arousal on the processing advantage of emotion words, but the effect was facilitative for positive words and inhibitory in negative words.

Concerning the scores of the DBI and STAI, minimal main and interaction effects were attached to them in the present study, as was the case in Knickerbocker et al.'s (2015) study with native speakers of English. In Knickerbocker et al.'s (2015) study, this indicated that the processing advantage of emotion words can emerge in L1 silent reading regardless of the participants' status of anxiety or depression. However, this conclusion will not apply to the present study as no processing advantage for emotion words was noted.

7. Conclusions

The current study aimed to examine the processing of emotionally positive/negative words versus neutral words in L2 normal silent reading. Two experiments were conducted that included Arab EFL learners read English sentences in which positive versus neutral words (Experiment 1) or negative versus neutral words (Experiment 2) were embedded. The participants' eye-movements were monitored against 13 measures. In contrast to earlier L1 studies (Knickerbocker et al. 2015; Scott et al. 2012), the present study failed (overall) to find any main effects for word type in both experiments. However, an interaction was found between word type and word frequency/arousal in Experiment 1 and between word type and word length/arousal in Experiment 2. Only Experiment 1 showed a few instances of interaction between word type and the BDI inventory, which assesses depression. The results seem to support the hypothesis that L1 is emotionally close while L2 is emotionally distant to bilinguals, especially the late ones who learned the L2 in instructional settings (Caldwell-Harris 2014; Chen et al. 2015).

In terms of theoretical models/hypotheses, the current study partially supports Kroll and Stewart (1994) and Jiang (2000). L2 emotion words could have behaved similarly to their L1 counterparts if the participants had longer natural exposure to the L2. In this case, the direct connections between L2 words and their concepts could have been strengthened enough to allow L2 emotion words to behave similarly to their L1 counterparts. It must be noted, however, that the results of the current study call for the explicit integration of the affective factor in mental lexicon models. The current study also casts doubt on the applicability of two models on the processing of emotion words by bilinguals: the motivated attention and affective states model (Lang et al. 1990, 1997) and the automatic vigilance model of emotion (Pratto and John 1991). A revision of the models in light of bilingualism research could help provide a comprehensive understanding of the behavior of emotion words.

In terms of pedagogy, the results of the current study add further emphasis to the need to pay special attention to the teaching of emotion words in the L2 classroom and engage learners in experiences that would strengthen their understanding and appreciation of these words. The fact that the L2 is emotionally distant and that EFL learners struggle to express their true feelings are well-documented in the literature (Caldwell-Harris 2014; Chen et al. 2015; Pratto and John 1991). This difficulty emerges in the present study, although the participants' proficiency level was good-to-high based on the vocabulary test they completed prior to the experiments.

Based on the present study, we can suggest some directions for future research. As noted previously, work on emotion words (e.g., *happy*, *sad*, *joy*) abounds in the literature, but far fewer studies appear to focus on the reading of emotion-laden words (e.g., *cancer*, *butterfly*, *knife*). The question remains in terms of how eye movements are modulated by the type of emotion word that is read or viewed, and whether or not words of higher or lower arousal and positive or negative valence influence measures of eye movements and reading in significant and systematic ways. A true comparison would involve creating sets of these materials that control for various characteristics or features (e.g., length, frequency, orthographic neighbors, valence, arousal, etc.) and differ only on the basis of whether they directly label an emotional state or represent an emotion, in the case of emotion-laden words. Work on both early and late measures of processing for both of these word classes will further illuminate our understanding of emotion word processing on the whole.

Author Contributions: Conceptualization, D.A.S.E.-D., S.S. and J.A.; methodology, D.A.S.E.-D., S.S. and J.A.; software, S.S.; validation, S.S.; formal analysis, S.S.; investigation, D.A.S.E.-D. and S.S.; resources, D.A.S.E.-D.; data curation, D.A.S.E.-D.; writing—original draft preparation, D.A.S.E.-D. and S.S.; writing—review and editing, J.A.; visualization, S.S.; supervision, J.A.; project administration, S.S.; funding acquisition, D.A.S.E.-D. All authors have read and agreed to the published version of the manuscript.

Funding: This research is funded by Prince Sultan University through the research lab [Applied Linguistics Research Lab- RL-CH-2019/9/1].

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of Prince Sultan University (IRB-2021-03-0076, 24 March 2021).

Informed Consent Statement: All participants gave informed consent to participate in the study.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: The researchers thank Prince Sultan University for funding this research project through the research lab [Applied Linguistics Research Lab- RL-CH-2019/9/1].

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. List of Target Words (Experiment 1)

Word	PoS	Word Type	Length (No. of Letters)	Valence (Mean)	Valence (SD)	Arousal (Mean)	Arousal (SD)	Word Frequency (HAL)	Orthographic Neighborhood Size	Mean LDT RT (ms)	Mean Naming RT (ms)
joy	N	Positive	3	8.6	0.71	7.22	2.13	11,703	10	516.912	588.481
joyful	A	Positive	6	8.22	1.22	5.98	2.54	569	0	604.848	601.889
kindness	N	Positive	8	7.82	1.39	4.3	2.62	2295	0	674.516	704.333
love	V	Positive	4	8.72	0.7	6.44	3.35	165,830	12	492.806	528.25
merry	N	Positive	5	7.9	1.49	5.9	2.42	1879	6	759.176	663.407
nice	A	Positive	4	6.55	2.44	4.38	2.69	113,119	8	561.727	552.038
optimism	N	Positive	8	6.95	2.24	5.34	2.58	882	1	686.062	621.481
passion	N	Positive	7	8.03	1.27	7.26	2.57	6347	0	589.545	583.107
secure	V	Positive	6	7.57	1.76	3.14	2.47	14,198	0	657.438	731.654
thankful	A	Positive	8	6.89	2.29	4.34	2.31	2279	0	701.471	567
thrill	V	Positive	6	8.05	1.48	8.02	1.65	2569	1	660.333	670.107
delight	V	Positive	7	8.26	1.04	5.44	2.88	3207	0	649.938	615.786
brave	A	Positive	5	7.15	1.64	6.15	2.45	5524	5	573.094	560.704
soothe	V	Positive	6	7.42	1.12	4.53	2.91	286	0	663.414	714.36
acceptance	N	Positive	10	7.98	1.42	5.4	2.7	8774	0	660.75	720.308
cheer	V	Positive	5	8.1	1.17	6.12	2.45	1904	3	603.471	627.818
excitement	N	Positive	10	7.5	2.2	7.67	1.91	5307	0	734.094	732.519
fascinate	V	Positive	9	7.34	1.68	5.83	2.73	88	0	893.032	751
friendly	A	Positive	8	8.43	1.08	5.11	2.96	17,486	0	612.758	617.036
gentle	A	Positive	6	7.31	1.3	3.21	2.57	6694	1	659.676	584.429
happy	A	Positive	5	8.21	1.82	6.49	2.77	70,881	4	536.273	555.929
invest	V	Neutral	6	5.93	2.1	5.12	2.42	7965	5	685.25	693.222
museum	N	Neutral	6	5.54	1.86	3.6	2.13	10,047	0	701.438	643.167
noisy	A	Neutral	5	5.02	2.02	6.38	1.78	2746	1	554.091	589.423
iron	A	Neutral	4	4.9	1.02	3.76	2.06	19,187	2	580.743	565.222
privacy	N	Neutral	7	5.88	1.5	4.12	1.83	13,645	1	606.152	675.519
rock	V	Neutral	4	5.56	1.38	4.52	2.37	44,285	12	543.879	554.464
sheltered	A	Neutral	9	5.75	1.92	4.28	1.77	654	0	688.303	692.357
spray	V	Neutral	5	5.45	1.63	4.14	2.28	5153	2	679.412	697.577
statue	N	Neutral	6	5.17	0.7	3.46	1.72	7801	1	695	674
stiff	A	Neutral	5	4.68	1.97	4.02	2.41	6210	3	639.455	659.852
yellow	A	Neutral	6	5.61	1.94	4.43	2.05	19,319	3	618.909	560.821
activate	V	Neutral	8	5.46	0.98	4.86	2.56	3486	0	657.353	688.25
appliance	N	Neutral	9	5.1	1.21	4.05	2.06	816	0	697.382	623.519
bake	V	Neutral	4	6.17	1.71	5.1	2.3	2528	17	673.267	610.643

Word	PoS	Word Type	Length (No. of Letters)	Valence (Mean)	Valence (SD)	Arousal (Mean)	Arousal (SD)	Word Frequency (HAL)	Orthographic Neighborhood Size	Mean LDT RT (ms)	Mean Naming RT (ms)
basket	N	Neutral	6	5.45	1.15	3.63	2.02	2815	4	649.882	584.071
blond	A	Neutral	5	6.43	2.04	5.07	2.7	3228	4	569.7	545.667
chair	N	Neutral	5	5.08	0.98	3.15	1.77	18,589	3	607.03	606.571
circle	V	Neutral	6	5.67	1.26	3.86	2.13	19,887	0	634.129	635.143
elbow	V	Neutral	5	4.9	0.45	3.7	2	3401	0	571.529	544.185
custom	A	Neutral	6	5.85	1.53	4.66	2.12	18,290	0	607.103	616
highway	N	Neutral	7	5.92	1.72	5.16	2.44	11,286	0	651.364	607.889

Appendix B. List of Target Words (Experiment 2)

Word	PoS	Word Type	Length (No. of Letters)	Valence (Mean)	Valence (SD)	Arousal (Mean)	Arousal (SD)	Word Frequency (HAL)	Orthographic Neighborhood Size	Mean LDT RT (ms)	Mean Naming RT (ms)
irritate	V	Negative	8	3.11	1.67	5.76	2.15	642	1	752.387	738.391
loneliness	N	Negative	10	1.61	1.02	4.56	2.97	880	1	679.406	685.963
lonely	A	Negative	6	2.17	1.76	4.51	2.68	4872	1	624.97	569.185
nervous	A	Negative	7	3.29	1.47	6.59	2.07	6719	0	587.515	684.296
hurt	V	Negative	4	1.9	1.26	5.85	2.49	28,018	5	562.656	529.852
rage	N	Negative	4	2.41	1.86	8.17	1.4	6289	13	584.343	587.5
sad	A	Negative	3	1.61	0.95	4.13	2.38	17,183	18	574.031	628.333
terrible	A	Negative	8	1.93	1.44	6.27	2.44	12,323	1	625.5	639.111
upset	V	Negative	5	2	1.18	5.86	2.4	11,355	0	706.545	582.5
useless	A	Negative	7	2.13	1.42	4.87	2.58	13,510	0	684.581	679.96
afraid	A	Negative	6	2	1.28	6.67	2.54	23,303	0	599.594	653.759
agony	N	Negative	5	2.43	2.17	6.06	2.67	2245	0	657.879	633.259
anger	N	Negative	5	2.34	1.32	7.63	1.91	8608	2	604.242	603.577
cruel	A	Negative	5	1.97	1.67	5.68	2.65	5177	1	611.065	624.5
depression	N	Negative	10	1.85	1.67	4.54	3.19	7691	1	680.152	627.148
disappoint	V	Negative	10	2.39	1.44	4.92	2.64	975	0	823.5	632.833
fear	V	Negative	4	2.76	2.12	6.96	2.17	34,537	12	588.469	593.852
grief	N	Negative	5	1.69	1.04	4.78	2.84	3903	1	611.147	577.519
hate	V	Negative	4	2.12	1.72	6.95	2.56	44,130	16	524.324	568.423
horror	N	Negative	6	2.76	2.25	7.21	2.14	12,795	0	586.424	649.464
detest	V	Negative	6	2.17	1.3	6.06	2.39	473	1	692.033	707.393
invest	V	Neutral	6	5.93	2.1	5.12	2.42	7965	5	685.25	693.222
museum	N	Neutral	6	5.54	1.86	3.6	2.13	10,047	0	701.438	643.167
noisy	A	Neutral	5	5.02	2.02	6.38	1.78	2746	1	554.091	589.423

Word	PoS	Word Type	Length (No. of Letters)	Valence (Mean)	Valence (SD)	Arousal (Mean)	Arousal (SD)	Word Frequency (HAL)	Orthographic Neighborhood Size	Mean LDT RT (ms)	Mean Naming RT (ms)
iron	A	Neutral	4	4.9	1.02	3.76	2.06	19,187	2	580.743	565.222
privacy	N	Neutral	7	5.88	1.5	4.12	1.83	13,645	1	606.152	675.519
rock	V	Neutral	4	5.56	1.38	4.52	2.37	44,285	12	543.879	554.464
sheltered	A	Neutral	9	5.75	1.92	4.28	1.77	654	0	688.303	692.357
spray	V	Neutral	5	5.45	1.63	4.14	2.28	5153	2	679.412	697.577
statue	N	Neutral	6	5.17	0.7	3.46	1.72	7801	1	695	674
stiff	A	Neutral	5	4.68	1.97	4.02	2.41	6210	3	639.455	659.852
yellow	A	Neutral	6	5.61	1.94	4.43	2.05	19,319	3	618.909	560.821
activate	V	Neutral	8	5.46	0.98	4.86	2.56	3486	0	657.353	688.25
appliance	N	Neutral	9	5.1	1.21	4.05	2.06	816	0	697.382	623.519
bake	V	Neutral	4	6.17	1.71	5.1	2.3	2528	17	673.267	610.643
basket	N	Neutral	6	5.45	1.15	3.63	2.02	2815	4	649.882	584.071
blond	A	Neutral	5	6.43	2.04	5.07	2.7	3228	4	569.7	545.667
chair	N	Neutral	5	5.08	0.98	3.15	1.77	18,589	3	607.03	606.571
circle	V	Neutral	6	5.67	1.26	3.86	2.13	19,887	0	634.129	635.143
elbow	V	Neutral	5	4.9	0.45	3.7	2	3401	0	571.529	544.185
custom	A	Neutral	6	5.85	1.53	4.66	2.12	18,290	0	607.103	616
highway	N	Neutral	7	5.92	1.72	5.16	2.44	11,286	0	651.364	607.889

Note

[1] We followed the forward step-by-step method in evaluating the contribution of interactions to the fit of the models to avoid convergence issues.

References

Algom, Daniel, Eran Chajut, and Shlomo Lev. 2004. A rational look at the emotional Stroop phenomenon: A generic slowdown, not a Stroop Effect. *Journal of Experimental Psychology: General* 133: 323–38. [CrossRef]

Altarriba, Jeantte, and Dana M. Basnight-Brown. 2011. The acquisition of concrete, abstract and emotion words in a second language. *International Journal of Bilingualism* 16: 446–52. [CrossRef]

Altarriba, Jeanette, and Katherine M. Mathis. 1997. Conceptual and lexical development in second language acquisition. *Journal of Memory and Language* 36: 550–568. [CrossRef]

Altarriba, Jeanette, and Lisa M. Bauer. 2004. The distinctiveness of emotion concepts: A comparison between emotion, abstract and concrete words. *The American Journal of Psychology* 117: 389–410. [CrossRef]

Altarriba, Jeanette, Lisa M. Bauer, and Claudia Benvenuto. 1999. Concreteness, context-availability, and imageability ratings and word associations for abstract, concrete, and emotion words. *Behavior Research Methods, Instruments, & Computers* 31: 578–602.

Angele, Bernhard, Abby E. Laishley, Keith Rayner, and Simon P. Liversedge. 2014. The effect of high- and low-frequency previews and sentential fit on word skipping during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 40: 1181–203. [CrossRef]

Barriga-Paulino, Catarina I., Milene Guerreiro, Luís Faísca, and Alexandra Reis. 2022. Does emotional valence modulate word recognition? A behavioral study manipulating frequency and arousal. *Acta Psychologica* 223: 103484. [CrossRef]

Bauer, Lisa M., Erik L. Olheiser, Jeanette Altarriba, and Nicole Landi. 2009. Word type effects in false recall: Concrete, abstract, and emotion critical lures. *American Journal of Pshychology* 122: 469–481. [CrossRef]

Beck, Aaron T., Robert A. Steer, and Gregory K. Brown. 1996. *Manual for the Beck Depression Inventory-II*. Boston: ScienceOpen, Inc.

Bradley, Margaret M., and Peter J. Lang. 1999. *Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings, Technical Report C-1*. Gainesville: The Center for Research in Psychophysiology, University of Florida.

Brysbaert, Marc, and Françoise Vitu. 1998. Word skipping: Implications for theories of eye movement control in reading. In *Eye Guidance in Reading and Scene Perception*. Edited by Geoffrey Underwood. Oxford: Elsevier Science Ltd., pp. 125–47.

Caldwell-Harris, Catherine L. 2014. Emotionality differences between a native and foreign language: Theoretical implications. *Frontiers in Psychology* 5: 1055. [CrossRef]

Chen, Peiyao, Jie Lin, Bingle Chen, Chunming Lu, and Taomei Guo. 2015. Processing emotional words in two languages with one brain: ERP and fMRI evidence from Chinese–English bilinguals. *Cortex* 71: 34–48. [CrossRef]

Drieghe, Denis, Keith Rayner, and Alexander Pollatsek. 2005. Eye movements and word skipping during reading revisited. *Journal of Experimental Psychology: Human Perception and Performance* 31: 954. [CrossRef]

El-Dakhs, Dina A. S., and Jeanette Altarriba. 2018. The distinctiveness of emotions words: Does it hold for foreign language learners? The case of Arab EFL learners. *Journal of Psycholinguistic Research* 47: 1133–1149. [CrossRef]

Estes, Zachary, and Michelle Verges. 2008. Freeze or flee? Negative stimuli elicit selective responding. *Cognition* 108: 557–65. [CrossRef]

Ferré, Pilar. 2002. Advantage for emotional words in immediate and delayed memory tasks: Could it be explained in terms of processing capacity? *The Spanish Journal of Psychology* 5: 78–89. [CrossRef]

Ferré, Pilar, Manuel Anglada-Tort, and Marc Guasch. 2018. Processing of emotional words in bilinguals: Testing the effects of word concreteness, task type and language status. *Second Language Research* 34: 371–94. [CrossRef]

Fox, Elaine, Riccardo Russo, Robert Bowles, and Kevin Dutton. 2001. Do threatening stimuli draw or hold visual attention in subclinical anxiety? *Journal of Experimental Psychology: General* 130: 681. [CrossRef]

Gerth, Sabrina, and Julia Festman. 2021. Reading development, word length and frequency effects: An eye-tracking study with slow and fast readers. *Frontiers in Communication* 6: 743113. [CrossRef]

Gollan, Tamar H., Timothy J. Slattery, Diane Goldenberg, Eva Van Assche, Wouter Duyck, and Keith Rayner. 2011. Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General* 140: 186. [CrossRef]

Huang, Lingshan, Jinghui Ouyang, and Jingyang Jiang. 2022. The relationship of word processing with L2 reading comprehension and working memory: Insights from eye-tracking. *Learning and Individual Differences* 95: 102143. [CrossRef]

Jiang, Nan. 2000. Lexical representation and development in a second language. *Applied Linguistics* 21: 47–77. [CrossRef]

Kazanas, Stephanie A., and Jeanette Altarriba. 2015a. The automatic activation of emotion and emotion-laden words: Evidence from a masked and unmasked priming paradigm. *The American Journal of Psychology* 128: 323–336. [CrossRef]

Kazanas, Stephanie A., and Jeanette Altarriba. 2015b. Emotion word type and affective valence priming at a long stimulus onset asynchrony. *Language & Speech* 59: 339–352.

Knickerbocker, Hugh, and Jeanette Altarriba. 2013. Differential repetition blindness with emotion and emotion-laden word types. *Visual Cognition* 21: 599–627. [CrossRef]

Knickerbocker, Hugh, Rebecca L. Johnson, and Jeanette Altarriba. 2015. Emotion effects during reading: Influence of an emotion target word on eye movements and processing. *Cognition and Emotion* 29: 784–806. [CrossRef] [PubMed]

Kousta, Stavroula-Thaleia, David P. Vinson, and Gabriella Vigliocco. 2009. Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition* 112: 473–81. [CrossRef] [PubMed]

Kroll, Judith F., and Erika Stewart. 1994. Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language* 33: 149–74. [CrossRef]

Kuperman, Victor, and Avital Deutsch. 2020. Morphological and visual cues in compound word reading: Eye-tracking evidence from Hebrew. *Quarterly Journal of Experimental Psychology* 73: 2177–87. [CrossRef] [PubMed]

Kuperman, Victor, Zachary Estes, Marc Brysbaert, and Amy Beth Warriner. 2014. Emotion and language: Valence and arousal affect word recognition. *Journal of Experimental Psychology: General* 143: 1065–81. [CrossRef] [PubMed]

Lang, Peter J., Margaret M. Bradley, and Bruce N. Cuthbert. 1990. Emotion, attention, and the startle reflex. *Psychological Review* 97: 377. [CrossRef]

Lang, Peter J., Margaret M. Bradley, and Bruce N. Cuthbert. 1997. Motivated attention: Affect, activation, and action. In *Attention and Orienting: Sensory and Motivational Processes*. Mahwah: Lawrence Erlbaum Associates Publishers, pp. 97–135.

Meara, Paul, and Imma Miralpeix. 2017. *Tools for Researching Vocabulary*. Clevedon: Multilingual Matters, vol. 105.

Ponari, Marta, Sara Rodríguez-Cuadrado, David Vinson, Neil Fox, Albert Costa, and Gabriella Vigliocco. 2015. Processing advantage for emotional words in bilingual speakers. *Emotion* 15: 644–52. [CrossRef]

Pratto, Felicia, and Oliver P. John. 1991. Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology* 61: 380–91. [CrossRef]

Rayner, Keith, and George W. McConkie. 1976. What guides a reader's eye movements? *Vision Research* 16: 829–37. [CrossRef]

Rayner, Keith, and Martin H. Fischer. 1996. Mindless reading revisited: Eye movements during reading and scanning are different. *Perception & Psychophysics* 58: 734–47. [CrossRef]

Rayner, Keith. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* 124: 372. [CrossRef] [PubMed]

Rayner, Keith. 2009. Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology* 62: 1457–506. [CrossRef] [PubMed]

Schotter, Elizabeth R. 2018. Reading ahead by hedging our bets on seeing the future: Eye tracking and electrophysiology evidence for parafoveal lexical processing and saccadic control by partial word recognition. In *Psychology of Learning and Motivation*. Edited by Kara D. Federmeier and Duane G. Watson. Cambridge: Academic Press, vol. 68, pp. 263–98.

Scott, Graham G., Patrick J. O'Donnell, and Sara C. Sereno. 2012. Emotion words affect eye fixations during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 38: 783–92. [CrossRef] [PubMed]

Sereno, Sara C., Graham G. Scott, Bo Yao, Elske J. Thaden, and Patrick J. O'Donnell. 2015. Emotion word processing: Does mood make a difference? *Frontiers in Psychology* 6: 1191. [CrossRef] [PubMed]

Spielberger, Chales D., Richard L. Gorsuch, Robert E. Lushene, P. R. Vagg, and Gerard A. Jacobs. 1983. *Manual for the State-Trait Anxiety Inventory (STAI)*. Palo Alto: Consulting Psychologists Press, Inc.

Team R Developement Core. 2021. *A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Vitu, Françoise, John Kevin O'Regan, Albrecht W. Inhoff, and Richard Topolski. 1995. Mindless reading: Eye-movement characteristics are similar in scanning letter strings and reading texts. *Perception & Psychophysics* 57: 352–64. [CrossRef]

Zhang, Juan, Chenggang Wu, Yaxuan Meng, and Zhen Yuan. 2017. Different neural correlates of emotion-label words and emotion-laden words: An ERP study. *Frontiers in Human Neuroscience* 11: 455. [CrossRef]