Psychonomic Bulletin & Review

Chinese Readers Utilize Emotion Information for Word Segmentation

Journal:	*Psychonomic Bulletin & Review*
Manuscript ID	PBR-BR-23-224.R1
Manuscript Type:	Brief Report
Date Submitted by the Author:	n/a
Complete List of Authors:	Huang, Linjieqiong; Institute of Psychology Chinese Academy of Sciences, Zhang, Xiangyang; Institute of Psychology Chinese Academy of Sciences Li, Xingshan; Institute of Psychology Chinese Academy of Sciences
Keywords:	reading, word perception, visual word recognition

Chinese Readers Utilize Emotion Information for Word Segmentation

Linjieqiong Huang[1], Xiangyang Zhang[2], and Xingshan Li[1, 3]

[1] CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

[2] CAS Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

[3] Department of Psychology, University of Chinese Academy of Sciences, Beijing, China

This research was supported by a grant from the National Natural Science Foundation of China (32371156, 31970992). This work was also jointly funded by the National Natural Science Foundation of China (NSFC) and the German Research Foundation (DFG) in Project Crossmodal Learning, NSFC 62061136001/DFG TRR-169. Linjieqiong Huang was supported by China Postdoctoral Science Foundation (2022M723362) and the Scientific Foundation of Institute of Psychology, Chinese Academy of Sciences (E2CX6625CX). The data, code and materials associated with this study are publicly available at Science Data Bank and can be accessed at https://doi.org/10.57760/sciencedb.psych.00130.

Correspondence should be addressed to Xingshan Li, 16 Lincui Road, Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China. Email: lixs@psych.ac.cn.

Word count: 2962 (including Introduction, Results, and Discussion)

Abstract

We reported a large-scale Internet-based experiment to investigate the impact of emotion information on Chinese word segmentation, in which participants completed an overlapping ambiguous string (OAS) segmentation task and the Chinese version of Beck Depression Inventory-II in a counterbalanced order. OAS is a three-character string (ABC) in which the middle character can form a distinct word with both the character on its left side (word AB) and the character on its right side (word BC). Participants were presented with isolated OASs and were asked to report the word they identified first. Emotional OAS was constructed by a combination of a neutral word and an emotional word, with the neutral and emotional words sharing character B. We orthogonally manipulated the valence of the emotional words (positive vs. negative) and their position in the OAS (left-side vs. right-side). The results showed that compared to neutral words, both positive and negative words were more likely to be segmented, and this segmentation outcome was not affected by readers with different depression tendencies. These findings suggest that emotion information can influence word segmentation, and that both positive and negative emotions take precedence over neutral words in the word segmentation process. This study provides a new perspective and evidence to understand the impact of emotion information on word processing.

Keywords: emotion information, word segmentation, lexical competition, negative, positive

Introduction

Unlike alphabetic languages such as English, Chinese texts do not have inter-word spaces to demarcate word boundaries. Consequently, Chinese readers need to group continuous characters into words (i.e., word segmentation) for successful reading comprehension (Li et al., 2009). A key issue in this regard is to understand what kind of information Chinese readers employ for word segmentation. While recent studies have shown that Chinese readers use linguistic factors such as word frequency to segment words (Ma et al., 2014), the impact of high-level information such as the emotion information carried by words on word segmentation remains unclear. In the present study, we aimed to investigate this question.

The absence of inter-word spaces in Chinese texts can lead to ambiguity in word boundaries in certain cases. One such form of ambiguity is known as *overlapping ambiguous string* (OAS). Typically, an OAS consists of three characters (ABC, denoting the characters from left to right), where the middle character can form distinct words with the characters on both its left (word AB) and its right (word BC) (Luo et al., 2002). Previous studies have examined how Chinese readers segment words using the OAS segmentation task and eye-tracking technique. Ma et al. (2014) found that word frequency plays a crucial role in word segmentation, and higher frequency words are more likely to be segmented. Additionally, due to the left-to-right reading direction in Chinese, words AB were found to be more likely to be segmented into words than words BC, all other factors being equal (Huang & Li, 2020; Huang et al., 2021; Ma et al., 2014). Huang et al. (2021) also obtained similar results as mentioned above in the OAS segmentation task. In their study, participants were presented with isolated OASs and were asked to report the word they identified first. These findings highlight the impact of word frequency and left-side word advantage on the word segmentation process.

How do Chinese readers use the information presented above to segment words? According to a computational model, the *Chinese reading model* (*CRM*), word identification and word segmentation are a unified process (Li & Pollatsek, 2020). Specifically, all characters within the perceptual span are activated and processed in parallel, and all words composed of these characters are also activated. Spatially overlapping words, such as words AB and BC where character B overlaps, compete with each other, and only one word can win the competition. Once a word wins the competition, it is identified and segmented from the text simultaneously.

An unsolved question in the literature is whether Chinese readers utilize high-level information such as the emotion information carried by words to segment words. Emotions play an important role in daily lives, human survival and biological regulation (Hinojosa et al., 2010; Kousta et al., 2009; Kuperman et al., 2014). Extensive research has been conducted to understand how emotion information affects word processing, mainly focusing on two questions: (1) the time course of emotional word processing (e.g., Herbert et al., 2008; Schacht & Sommer, 2009; Scott et al., 2009; Zhang et al., 2014), and (2) whether positive and negative emotions have similar effects on word processing (e.g., Estes & Adelman, 2008; Kuperman et al., 2014; Scott et al., 2012). However, these studies have not provided sufficient evidence to understand whether Chinese readers use emotion information during word segmentation. Moreover, previous studies have found mixed results regarding the two aforementioned questions.

Regarding the time course of emotional word processing, several ERP studies have indicated the early impact of emotion information on word processing. They found that the amplitude of early ERP components such as P1 and P2 components elicited by emotional words was greater compared to neutral words (Kanske & Kotz, 2007; Scott et al., 2009; Zhang et al., 2014). In contrast, other studies have not found any ERP differences between emotional and neutral words in early time window, suggesting that the emotion information of words may not be activated in the early stages of processing (Herbert et al., 2008; Schacht & Sommer, 2009).

Previous studies have yielded inconsistent results regarding whether positive and negative emotions have similar effects on word processing. Some studies have reported a processing advantage for both positive and negative words compared to neutral words. The results of these studies support the *emotion advantage account*, which suggests that emotional stimuli, due to their motivational significance, capture attention earlier and maintain a processing advantage over neutral stimuli, regardless of their valence (Kousta et al., 2009; Lang et al., 1990, 1997; Scott et al., 2012). On the other hand, other studies have only observed a processing advantage for negative words. The results of these studies support the *negative advantage account*, which posits that survival is heavily related to withdrawing from potential negative situations, so that negative stimuli are allocated more attentional resources and therefore processed faster than neutral stimuli (Cacioppo & Gardner, 1999; Kuperman et al., 2014; Pratto & John, 1991). A third perspective emerges from studies indicating processing disadvantage associated with negative words and is known as the *negative disadvantage account*. This view suggests that due to defense mechanisms that freeze activity in the presence of threat, negative stimuli result in a general slowdown in processing (Algom et al., 2004; Estes & Adelman, 2008; Estes & Verges, 2008).

Despite the impact of emotion information on word processing has been extensively examined, its impact on Chinese word segmentation remains unclear. To understand the effects of emotion information on word segmentation, we investigated the following three questions in this study. The first question was whether emotion information could have an impact on words before they are identified. One possibility is that emotion information exerts its effect only after words are identified. Alternatively, emotion information can exert influence before words are identified. If emotion information can indeed play a role before word identification, we can then delve into whether negative and positive emotions have similar or distinct effects on the word segmentation process. Furthermore, previous studies have shown that depressive states typically increase sensitivity to negative stimuli (Chen et al., 2014; Dai & Feng, 2012; Nandrino et al., 2004). Thus, we also examined whether readers with different depression tendencies have similar or distinct effects in the segmentation process involving emotional words.

To address these questions, we conducted a large-scale Internet-based study in which participants completed an OAS segmentation task and the *Chinese version of Beck Depression Inventory-II* (BDI-II-C; Beck et al., 1996; Wang et al., 2011) in a counterbalanced order. In the OAS segmentation task, as a means of detecting the lexical competition outcome, participants were presented with isolated OASs and asked to report the word they first identified. Emotional OAS was formed by combining a neutral word with an emotional word, which overlapped in character B and were comparable in word frequency. We orthogonally manipulated both the valence of emotional words (positive vs. negative), and their position in the OAS (left-side vs. right-side) (see Table 1 for an example). Participants completed the BDI-II-C either before or after the OAS segmentation task. The BDI-II-C is a widely used self-evaluation scale designed to assess the severity of depressive symptoms. Based on the total scores of the BDI-II-C, participants were divided into either high or low depression tendency group.

Different hypotheses make different predictions about the outcomes of word segmentation. If the emotion information is activated only after words are identified, then the emotion information of words will not influence the outcome of word segmentation. According to CRM, once a word wins the lexical competition, it is identified and segmented simultaneously. Thus, if the emotion information is activated only after a word is identified, it has no chance to influence the lexical competition and hence the outcome of word segmentation. Therefore, regardless of the location of the emotional words, participants should segment the left-side word more frequently, due to the left-side word advantage when all other factors are equal. In contrast, if the emotion information can be activated before a word is fully identified, then the emotion information of words will influence the outcome of word segmentation. That is, if the emotion information is partially activated as part of the word representation during word identification process, then it may affect lexical competition and hence the outcome of word segmentation. As a result, the outcomes of word segmentation should vary depending on the location of the emotional word.

Now let us turn to the issue about the impacts of different emotion valences. According to the emotion advantage account, both positive and negative words should be more likely to be segmented than neutral words. However, according to the negative advantage account, negative words should be more likely to be segmented compared to neutral words, which is not expected for positive words. Furthermore, according to the negative disadvantage account, negative words should be less likely to be segmented compared to neutral words, which is not expected for positive words.

As for individual differences, if readers with high depression tendencies are more sensitive to negative stimuli (Chen et al., 2014; Dai & Feng, 2012; Nandrino et al., 2004), then the probabilities of negative words being segmented should be higher for readers with high depression tendencies compared to those with low depression tendencies. On the contrary, if readers with different depression tendencies exhibit similar performance in word segmentation, no significant difference should be observed in the probabilities of segmenting negative words between two depression tendency groups.

Methods

Participants

A total of 413 college students (198 females; age range 18–30 years, M = 21.84 years, SE = 0.12) participated in the experiment. Among them, 18 participants were excluded from data analysis because their accuracy rate on fillers was less than 85% (see the Procedure section for more information about fillers). All participants were native Chinese speakers with normal or corrected-to-normal vision and without a neurological history or language disorder diagnosis.

To explore the individual differences in the OAS segmentation involving emotional words, participants were divided into either high or low depression tendency group based on the total scores of the BDI-II-C. Because a total score of 15 is the optimal cutoffs for detecting depressive disorder (Yang & Xiong, 2016), participants who obtained scores greater than or equal to 15 were classified into the high depression tendency group ($N = 81$), while those with scores less than 15 were classified into the low depression tendency group ($N = 314$). This categorization allowed for the examination of potential differences in the OAS segmentation task performance between these two groups.

To determine the ideal number of participants, we estimated the priori power of the study by using the *powerSim* and *powerCurve* functions of the *simr* package (Green & MacLeod, 2018) within the *R* Environment for Statistical Computing (R Core Team, 2022). First, we conducted a pilot study with 36 participants and analyzed the data using a generalized linear mixed-effects model, in which the probability of AB-C segmentation was the dependent variable. As an effect of interest to us, we used the contrast that tests whether high and low depression tendency groups differ significantly in the left-side negative condition as the fixed effect (see more details in the Data Analysis section). Then, based on the pilot data, we explored how the power varies as a function of the number of participants. The results indicated that the power estimate of 400 participants was 99% (95% CI: 94.55%–99.97%). Therefore, 413 participants were suitable for a well-powered within-subjects design (Brysbaert & Stevens, 2018; Cohen, 2013).

Materials

We selected a total of 88 emotional words and paired them with 88 neutral words, thus constituting 88 emotional OASs, which did not form any valid phrases in Chinese. Emotional words were positioned either on the left-side or the right-side of each OAS, with half of them being positive and the other half being negative. Therefore, this experiment was a 2 (emotion: negative vs. positive) X 2 (location: left-side vs. right-side) within-participant design, resulting in 22 OASs in each condition (see Table 1). For example, in the left-side negative condition, a negative word such as "忽视" (meaning *neglect*) was paired with a neutral word such as "视觉" (meaning vision) to form an OAS "忽视觉"; whereas in the right-side negative condition, a negative word such as "分歧" (meaning *disagreement*) was paired with a neutral word such as "学分" (meaning *credit*) to form an OAS "学分歧". In the left-side positive condition, a positive word such as "礼物" (meaning *gift*) was paired with a neutral word such as "物体" (meaning *object*) to form an OAS "礼物体"; whereas in the right-side positive condition, a positive word such "天使" (meaning *angel*) was paired with a neutral word such as "先天" (meaning *innate*) to form an OAS "先天使".

Table 1

Examples of Stimuli in the Emotional Group

Condition	Word AB	Word BC	OAS
Left-side negative	忽视 *Neglect*	视觉 *Vision*	忽视觉
Right-side negative	学分 *Credit*	分歧 *Disagreement*	学分歧
Left-side positive	礼物 *Gift*	物体 *Object*	礼物体
Right-side positive	先天 *Innate*	天使 *Angel*	先天使

Note. The translation corresponding to each word is presented in italics.

Evaluation of the Words in the Emotional Group. We selected emotional words from the Chinese Affective Words System (Wang et al., 2008) and a database of the affective norms (Yao et al., 2017). Both databases use a Likert scale ranging from 1 to 9, with 1 indicating *very negative/unpleasant*, and 9 indicating *very positive/pleasant*. For positive word group, we selected words with values higher than 6.5, while for negative word group, we selected words with values lower than 3.5. We selected a total of 268 emotional words from these databases.

To assess the emotional valence of the words, we further conducted a norming study on the selected emotional words. A total of 29 participants who did not participate in the main experiment were presented with these words and were asked to rate the pleasure level of each word on a 7-point Likert scale (1= *very unpleasant*, 7 = *very pleasant*). We chose negative words rated below 2.5, and chose positive words rated above 5.5. In addition, we selected emotional words with a medium frequency ranging from 1 to 50 occurrences per million (Lexicon of common words in contemporary Chinese 2008). We also controlled for the arousal level of emotional words in different conditions (Left-side negative: $M = 5.61$, $SE = 0.17$; Right-side negative: $M = 5.56$, $SE = 0.14$; Left-side positive: $M = 5.54$, $SE = 0.14$; Right-side positive: $M = 5.66$, $SE = 0.19$; Fs <1). As a result, a total of 88 emotional words were used in the main experiment, with 50% of the positive and 50% of the negative words.

Each emotional word was paired with a neutral word that was comparable to the emotional word in terms of word frequency and the accumulated number of strokes of the words (referred to as "word strokes" hereafter) (ps > .100; see Table 2). According to the norming study mentioned above, neutral words were also rated on a 7-point Likert scale for pleasure level, with all neutral words receiving a rating in the range of 3 to 5. Emotional valence ratings for neutral words were comparable for each OAS in all four conditions (ps

> .150; see Table 2). However, for the emotional words, the emotional valence ratings were significantly higher for positive words (M = 6.03, SE = 0.04) than for negative words (M = 1.99, SE = 0.05; $F(1, 21) = 4931.99$, $p < .001$), while other effects were not significant (Fs < 1; see Table 2). For each OAS, there was a significant difference in the emotional valence ratings between the emotional words and their paired neutral words (Left-side negative: $t(21) = -16.97$, $p < .001$; Right-side negative: $t(21) = -21.21$, $p < .001$; Left-side positive: $t(21) = 15.31$, $p < .001$; Right-side positive: $t(21) = 18.36$, $p < .001$).

Properties of Words in the Control Group. We included the other 88 OASs as a control group, where words AB and BC were neutral words and were referred to as the *control condition*. And these OASs also did not form any valid phrases in Chinese. For example, a control OAS “分类似” had a left-side neutral word “分类” (meaning *classification*) and a right-side neutral word “类似” (meaning *similar*). The 29 participants who participated in the norming study described above also rated these neutral words on a 7-point Likert scale (1= *very unpleasant*, 7 = *very pleasant*). The results showed no significant differences between words AB (M = 4.16, SE = 0.03) and BC (M = 4.13, SE = 0.03), as all ratings ranged from 3 to 5 ($t(174) = 0.60$, $p = .552$). Similar to words AB and BC in the emotional group, words AB and BC in the control group were also medium-frequency words ranging from 1 to 50 occurrences per million (Lexicon of Common Words in Contemporary Chinese Research Team, 2008). Word frequency ($t(174) = -0.76$, $p = .449$) and word strokes ($t(174) = -0.37$, $p = .714$) were comparable between words AB (word frequency: M = 7.99, SE = 0.63; word strokes: M = 14.69, SE = 0.34) and BC (word frequency: M = 8.73, SE = 0.77; word strokes: M = 14.88, SE = 0.36).

The control group provided a baseline for the OAS segmentation, in which no emotional words were embedded. To compare the segmentation outcomes between emotional words and neutral words, we randomly divided the control group into two subgroups: left-side control group and right-side control group[1]. These subgroups corresponded to the conditions in which the emotional words were positioned on the left or right side for further data analysis.

[1] The properties of the two subgroups of stimuli did not differ significantly. Moreover, the probabilities of AB-C segmentation were comparable between the two subgroups (M = .54, SE = .004 for the left-side control group; M = .54, SE = .004 for the right-side control group).

Table 2

Properties of Words AB and BC in the Emotional Group

	Negative – Neutral		Neutral – Negative		Positive – Neutral		Neutral – Positive	
	Word AB	Word BC	Word AB	Word BC	Word AB	Word BC	Word AB	Word BC
Word frequency	8.32 (1.82)	8.30 (2.07)	8.05 (1.37)	7.94 (1.26)	7.56 (1.28)	8.20 (1.31)	7.74 (1.20)	7.92 (0.99)
Word strokes	14.95 (0.89)	14.86 (0.82)	14.82 (0.48)	14.91 (0.66)	15.95 (0.97)	15.55 (1.07)	16.23 (0.80)	15.86 (0.77)
Emotional valence rating	2.02 (0.08)	4.01 (0.08)	4.14 (0.10)	1.97 (0.08)	6.03 (0.06)	4.20 (0.09)	4.04 (0.10)	6.02 (0.07)
	[1.3–2.5]	[3.4–4.9]	[3.2–4.9]	[1.3–2.5]	[5.5–6.6]	[3.3–4.9]	[3.2–4.7]	[5.6–6.7]

Note. Negative – Neutral represents the left-side negative condition; Neutral – Negative represents the right-side negative condition; Positive – Neutral represents the left-side positive condition; Neutral – Positive represents the right-side positive condition. Standard errors are given in parentheses. Ranges are given in brackets. The unit of word frequency is the number of occurrences per million.

Procedures

Participants were instructed to complete the task online through the *Wenjuanxing* platform (https://www.wjx.cn/). Wenjuanxing is an online survey platform designed to create, distribute, and analyze surveys. It provides a user-friendly interface for researchers to design customized questionnaires and invite respondents to participate through various channels. Wenjuanxing has been widely applied in the field of scientific research and is a valuable tool for researchers seeking to collect and interpret data through structured surveys.

After agreeing to participate, participants were required to sign a consent form. Then, they were instructed to complete an OAS segmentation task and the BDI-II-C in a counterbalanced order. For the OAS segmentation task, participants were given a questionnaire and instructed to report the word they first identified. Each participant was presented with 176 OASs (comprising the emotional group and the control group) in a randomized order. They also read 30 fillers, which were unambiguous regarding word boundaries. These fillers were included to detect whether participants were serious about performing the task. The fillers were three-character strings and consisted of two types, with each type accounting for half: one in which only the first two characters formed a word (e.g., "启动各", in which the first two characters form the word "启动" meaning *start*) and one in which only the last two characters formed a word (e.g., "里信纸", in which the last two characters form the word "信纸", meaning *stationery*). Furthermore, participants completed the BDI-II-C either before or after the OAS segmentation task. The 21 items in the BDI-II-C aim to measure depressive symptoms such as despair, irritability, guilt, energy loss, and changes in appetite or sleep habits, each with four options indicating varying levels of symptom frequency or severity. For each item, participants were instructed to choose the statement that best represented their current state based on their experiences over the past two weeks among four options.

Transparency and Openness

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. All data, code, and materials are publicly available at Science Data Bank and can be accessed at https://doi.org/10.57760/sciencedb.psych.00130. This study's design and its analysis were not pre-registered.

Data Analysis

The mean accuracy of the fillers was 98%, indicating that participants completed the questionnaire with care. Data was analyzed using *generalized mixed-effects models* (Baayen et al., 2008). The segmentation outcomes (coding response AB-C as 1 and A-BC as 0) were used as the dependent variable in the models. Following Schad et al. (2020), we used a generalized linear mixed-effects model to directly test the theoretically motivated hypotheses using the following five customized contrasts: 1) a contrast that tests the main effect of location (left-side vs. right-side), 2) a contrast that tests whether positive and negative conditions differ significantly in the left-side condition, 3) a contrast that tests whether negative and control conditions differ significantly in the left-side condition, 4) a contrast that tests whether positive and negative conditions differ significantly in the right-side condition, 5) a contrast that tests whether positive and control conditions differ significantly in the right-side condition, 6) a contrast that tests whether high and low depression tendency groups differ significantly in the left-side control condition, 7) a contrast that tests whether high and low depression tendency groups differ significantly in the left-side negative condition, 8) a contrast that tests whether high and low depression tendency groups differ significantly in the left-side positive condition, 9) a contrast that tests whether high and low depression tendency groups differ significantly in the right-side control condition, 10) a contrast that tests whether high and low depression tendency groups differ significantly in the right-side negative condition, and 11) a contrast that tests whether high and low depression tendency groups differ significantly in the right-side positive condition.

Following Barr et al. (2013), we used a maximal model that converged. We first constructed a model with a maximum random factor structure that specified participants and items as crossed random effects, including intercepts and slopes. When the maximal model failed to converge, we used a zero-correlation parameter model and dropped the random components that generated the minimum variances. The *glmer* function from the *lme4* package (Bates et al., 2015) was used in the *R* Environment for Statistical Computing (R Core Team, 2022). We reported regression coefficients (*b*s, which estimate the effect size), 95% confidence interval, standard errors (*SE*s), *z*-values, and corresponding *p*-values. We estimated and reported the p-values for the effects using the *summary* function from the *lmerTest* package (Kuznetsova et al., 2017).

Results

As shown in Figure 1 and Table 3, the probability of AB-C segmentation was significantly higher in the left-side condition ($M = .59$, $SE = 0.003$) compared to the right-side condition ($M = .46$, $SE = 0.003$). Specifically, for the emotional group, the probability of AB-C segmentation was significantly higher in the left-side condition ($M = .64$, *SE* = 0.004) compared to the right-side condition (*M* = .38, *SE* = 0.004).

We made further comparison between neutral and emotional words. For the left-side condition, the probability of AB-C segmentation was significantly higher in the positive condition (*M* = .70, *SE* = 0.005) than in the negative condition (*M* = .59, *SE* = 0.005); however, the probability of AB-C segmentation was comparable between the negative condition and the control condition (*M* = .54, *SE* = 0.004). For the right-side condition, the probability of AB-C segmentation was significantly lower in the negative condition (*M* = .36, *SE* = 0.005) than in the positive condition (*M* = .40, *SE* = 0.005); and the probability of AB-C segmentation was significantly lower in the positive condition than in the control condition (*M* = .54, *SE* = 0.005).

A significant difference between the high and low depression tendency groups was observed in three conditions: the left-side positive condition, the right-side control condition, and the right-side negative condition. Notably, in these three conditions, participants with high depression tendencies consistently exhibited lower probabilities of AB-C segmentation compared to those with low depression tendencies. This individual difference was not affected by the location of emotional words, suggesting that the presence of emotional words did not cause this observed variation.

Table 3

Results of Generalized Linear Mixed-Effect Models

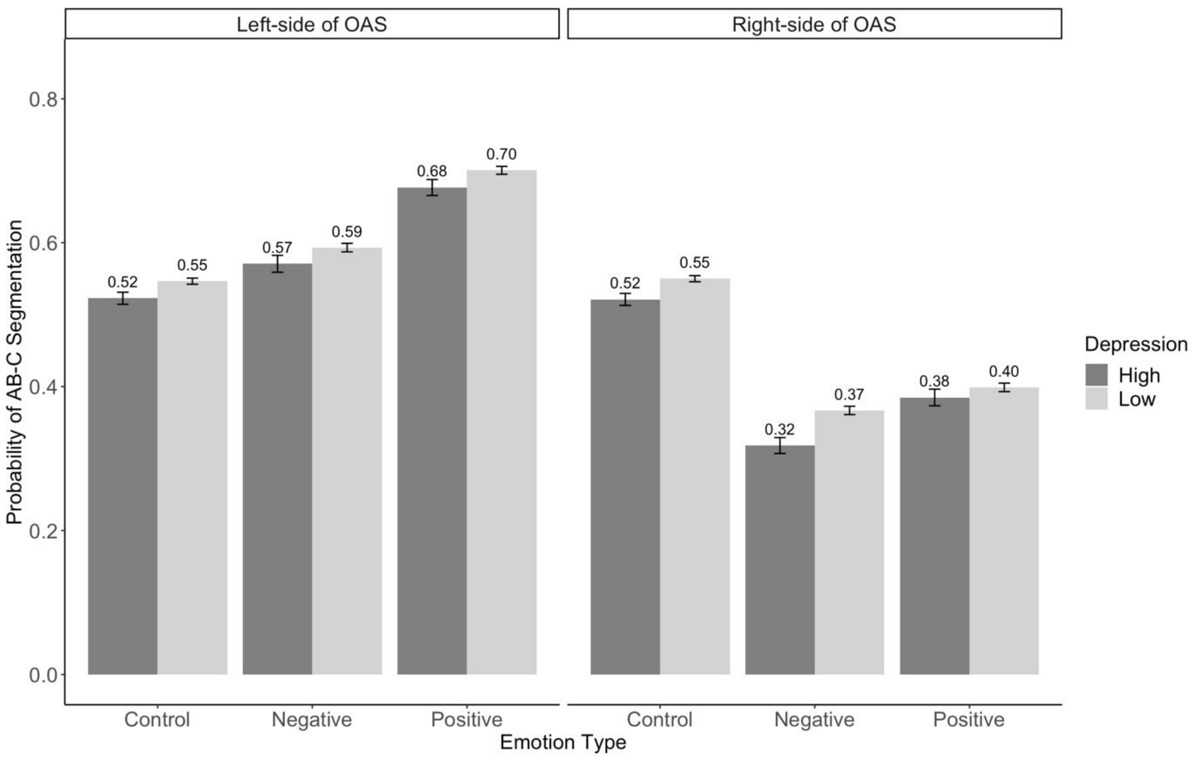
Fixed effects	*b*	*CI (0.95)*	*SE*	*z*	*p*
Left-side vs. Right-side	**0.62**	**[0.57, 0.68]**	**0.03**	**21.04**	**< .001**
Left-side: Positive vs. Negative	**0.71**	**[0.59, 0.82]**	**0.06**	**12.22**	**< .001**
Left-side: Negative vs. Control	−0.06	[−0.14, 0.02]	0.04	−1.37	.171
Right-side: Positive vs. Negative	**−0.33**	**[−0.44, −0.22]**	**0.06**	**−5.89**	**< .001**
Right-side: Positive vs. Control	**−0.79**	**[−0.87, −0.71]**	**0.04**	**−18.55**	**< .001**
Left-side control: high vs. low	0.14	[−0.06, 0.35]	0.10	1.40	.162

Left-side negative: high vs. low	0.19	[−0.03, 0.41]	0.11	1.69	.091
Left-side positive: high vs. low	**0.25**	**[0.03, 0.48]**	**0.11**	**2.19**	**.029**
Right-side control: high vs. low	**0.21**	**[0.01, 0.41]**	**0.10**	**2.05**	**.041**
Right-side negative: high vs. low	**0.26**	**[0.03, 0.48]**	**0.11**	**2.27**	**.023**
Right-side positive: high vs. low	0.10	[−0.12, 0.32]	0.11	0.91	.365

Note. Significant effects are shown in bold.

Figure 1

Probability of AB-C Segmentation Across Conditions



Note. Each text label represents mean probability of AB-C segmentation in each condition. Error bars represent standard errors.

Discussion

We conducted a large-scale Internet-based experiment in which participants completed an OAS segmentation task and the BDI-II-C to examine the impact of emotion information on Chinese word segmentation. Both the valence of emotional words and their position in the OAS were manipulated. We found a significant effect of the location of emotional words. Specifically, compared to neutral words, positive words were more likely to be segmented regardless of their position, while negative words were more likely to be segmented when positioned on the right side of the OAS. Readers with high depression tendencies were less likely to make AB-C segmentation than those with low depression tendencies, which was independent of the presence of emotional words.

The significant effect of the location of emotional words on segmentation outcomes suggests that emotion information can be activated before words are fully identified. Based on the CRM, once a word wins the lexical competition, it is identified and segmented simultaneously. If the emotion information is activated only after the words are identified, it has no chance to affect the lexical competition and hence the outcome of word segmentation. Thus, we would expect a left-side word advantage for the emotional group, regardless of where the emotional words are located, just as the control group did. However, our findings suggest that the position of the emotional words does affect the probability of AB-C segmentation. When the emotional words were located on the left side, the probability of AB-C segmentation was higher for the positive condition compared to the control condition. Conversely, when the emotional words were located on the right side, the probability of AB-C segmentation was lower in the positive and negative conditions compared to the control condition. These results indicate that the emotion information of words AB and BC has an impact before the words are fully identified. They provide evidence to support the argument that the emotion information impacts Chinese word segmentation, with emotional words being more likely to be segmented than neutral words.

The results of this study contribute to the understanding of the time course of emotional word processing. Previous studies exploring this issue have relied on measures classified as early or late measures to detect the effects of emotion information on early or late measures

(Herbert et al., 2008; Scott et al., 2009, 2012). However, the classification of measures into early or late stages may not reflect the exact time course of processing. To overcome this limitation, we employed the word segmentation phenomenon as an entry point to provide a robust approach to addressing this question. Based on the CRM, which posits a unified process of word identification and word segmentation, we propose that if emotion information can influence lexical competition and hence the outcome of word segmentation, this would serve as strong evidence for the early impact of emotion information on word processing. This indicates that emotion information can be partially activated as part of word representation along with the word identification process, thus providing a more nuanced understanding of the time course of emotional word processing.

The present results further showed that compared to neutral words, both positive and negative words were more likely to be segmented (except for the left-side negative words). These findings support the emotion advantage account, but contradicted the negative advantage account and the negative disadvantage account. Based on the negative advantage account (Pratto & John, 1991), only negative words should be more likely to be segmented than neutral words due to the allocation of additional attention resources. Based on the negative disadvantage account (Kuperman et al., 2014), due to defense mechanisms against threat, negative words should be less likely to be segmented than neutral words. In contrast to these accounts, the emotion advantage account posits that both positive and negative words are considered as salient due to their motivational significance, resulting in processing advantages over neutral stimuli (Knickerbocker et al., 2015; Kousta et al., 2009). According to this account, both positive and negative words should be more likely to be segmented than neutral stimuli, which is confirmed by the results of the present study. This finding suggests that individuals can use both the positive and negative information of words when segmenting words.

We did not find that negative words were more likely to be segmented when positioned on the left side of the OAS compared to neutral words. A possible reason is that the segmentation priority of negative words is offset by the desire to avoid negative outcomes. Human behavior is influenced by two underlying motivational systems where emotion stem from: approach-appetitive and withdrawal-aversive (Bradley & Lang, 2000; Trew, 2011). Human have a desire to approach positive outcomes and a desire to avoid negative outcomes. Due to reading orientation, the words on the left side of OAS (i.e., AB) are given processing priority and more salient than the words on the right side of OAS (i.e., BC). Thus, when word AB is a negative word, readers might avoid to segment it. This avoidance tendency might offset the segmentation priority brought by negative words, resulting in non-significant differences between negative words and neutral words.

Although previous studies have shown that depressive states are typically associated with greater sensitivity to negative stimuli (Chen et al., 2014; Dai & Feng, 2012; Nandrino et al., 2004), the present study did not find the segmentation priority of positive and negative words was affected by readers with different depression tendencies. This finding is consistent with the study by Knickerbocker et al. (2015), which found that the level of depressive symptoms did not interact with the emotion effect. Although we did not find any individual differences in segmenting emotional words, we did find some individual differences in performing OAS segmentation task. Specifically, participants with high depression tendencies consistently exhibited lower probabilities of AB-C segmentation compared to those with low depression tendencies. This individual difference might reflect a higher avoidance tendency among individuals with high depression tendencies. Depression has been shown to be closely related to avoidance, as it increases avoidance and escape activity (Ferster, 1973; Trew, 2011). Thus, in the present study, individuals with high depression tendencies might be more likely to avoid making responses than those with low depression tendencies. Since words AB are given processing priority due to reading orientation, they might avoid making response to words AB, resulting in lower probabilities of AB-C segmentation than those with low depression tendencies. Additionally, the impact of emotion information on word segmentation might be so strong that it reaches the ceiling level, making it difficult to show individual differences. The exact reason needs further research.

The findings of the present study are suggestive for the Chinese word segmentation model. According to the CRM, both words AB and BC were activated and competed with each other for segmentation (Li & Pollatsek, 2020). In addition to word frequency and the left-side word advantage, the present findings suggest that the CRM should consider how emotion information affects this lexical competition process. One possible way is to assume that emotion information is partially activated along with the activation of word units, without waiting for the entire word to be fully activated. The emotion information carried by words can increase their activation, which makes them more likely to win lexical competition against neutral words.

Conclusion

In this study, we conducted a large-scale Internet-based experiment utilizing the OAS segmentation task and the BDI-II-C to investigate the impact of emotion information on Chinese word segmentation. The results showed that both positive and negative words were more likely to be segmented compared to neutral words, and this segmentation outcome is not affected by readers with different depression tendencies. These results highlight the significance of emotion information in word segmentation and suggest that Chinese reading models need to take this factor into account. Moreover, our findings indicate that emotion information can have an impact even before complete word identification, supporting the early effect of emotion information on word processing. Additionally, both positive and negative words show priority during word segmentation process, supporting the emotion advantage account.

Declarations

Funding

This research was supported by a grant from the National Natural Science Foundation of China (32371156, 31970992). This work was also jointly funded by the National Natural Science Foundation of China (NSFC) and the German Research Foundation (DFG) in Project Crossmodal Learning, NSFC 62061136001/DFG TRR-169. Linjieqiong Huang was supported by China Postdoctoral Science Foundation (2022M723362) and the Scientific Foundation of Institute of Psychology, Chinese Academy of Sciences (E2CX6625CX).

Conflict of interest statement

The authors declare no conflicts of interest.

Data availability statement

The data, code and materials of this study are publicly available at Science Data Bank and can be accessed at https://doi.org/10.57760/sciencedb.psych.00130.

Ethics approval

The study was approved by the ethics committee of the Institute of Psychology, Chinese Academy of Sciences.

Consent to participate

All participants gave their informed consent prior to their inclusion in the study.

Consent for publication

The participant has consented to the submission of their data to the journal.

Author contributions

Linjieqiong Huang: Conceptualization, Methodology, Software, Investigation, Formal analysis, Visualization, Writing-Original draft, Writing-Reviewing & Editing.

Xiangyang Zhang: Writing-Reviewing & Editing.

Xingshan Li: Conceptualization, Methodology, Formal analysis, Visualization, Supervision, Writing-Reviewing & Editing, Funding acquisition.

References

Algom, D., Chajut, E., & Lev, S. (2004). A rational look at the emotional stroop phenomenon: A generic slowdown, not a stroop effect. *Journal of Experimental Psychology: General, 133*(3), 323–338. https://doi.org/10.1037/0096-3445.133.3.323

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

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

Beck, A. T., Steer, R. A., & Brown, G. (1996). Beck depression inventory–II. *Psychological assessment*.

Bradley, M. M., & Lang, P. J. (2000). Emotion and motivation. In J. T. Cacioppo, L. G. Tassinary, & G. Berntson (Eds.), Handbook of Psychophysiology (pp. 602–642). New York: Cambridge University Press.

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, *1*(1), 1–20. https://doi.org/10.5334/joc.10

Cacioppo, J. T., & Gardner, W. L. (1999). Emotion. *Annual Review of Psychology, 50*, 191–214. https://doi.org/10.1146/annurev.psych.50.1.191

Chen, J., Ma, W., Zhang, Y., Wu, X., Wei, D., Liu, G., ... & Zhang, Z. (2014). Distinct facial processing related negative cognitive bias in first-episode and recurrent major depression: Evidence from the N170 ERP component. *PLoS One*, *9*(10), e109176. https://doi.org/10.1371/journal.pone.0109176

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.

Dai, Q., & Feng, Z. (2012). More excited for negative facial expressions in depression: Evidence from an event-related potential study. *Clinical Neurophysiology*, *123*(11), 2172–2179. https://doi.org/10.1016/j.clinph.2012.04.018

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*(2), 557–565. https://doi.org/10.1016/j.cognition.2008.03.003

Ferster, C. B. (1973). A functional analysis of depression. *American Psychologist*, *28*, 857–870. https://doi.org/10.1037/h0035605

Green, P., & MacLeod, C. J. (2016). simr: An R package for power analysis of generalised linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493–498. https://doi.org/10.1111/2041-210X.12504

Herbert, C., Junghofer, M., & Kissler, J. (2008). Event related potentials to emotional adjectives during reading. *Psychophysiology, 45*(3), 487–498. https://doi.org/10.1111/j.1469-8986.2007.00638.x

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

Huang, L., & Li, X. (2020). Early, but not overwhelming: The effect of prior context on segmenting overlapping ambiguous strings when reading Chinese. *Quarterly Journal of Experimental Psychology, 73*(9), 1382–1395. https://doi.org/10.1177/1747021820926012

Huang, L., Staub, A., & Li, X. (2021). Prior context influences lexical competition when segmenting Chinese overlapping ambiguous strings. *Journal of Memory and Language, 118*, 104218. https://doi.org/10.1016/j.jml.2021.104218

Kanske, P., & Kotz, S. A. (2007). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain Research, 1148*, 138–148. https://doi.org/10.1016/j.brainres.2007.02.044

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

Kousta, S.-T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition, 112*(3), 473–481. https://doi.org/10.1016/j.cognition.2009.06.007

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

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*, 1–26.

https://doi.org/10.18637/jss.v082.i13

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological Review, 97*(3), 377–395. https://doi.org/10.1037/0033-295X.97.3.377

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). Motivated attention: Affect, activation, and action. In P. J. Lang, R. F. Simons, & M. T. Balaban (Eds.), *Attention and orienting: Sensory and motivational processes* (Vol. 97, pp. 97–135). Lawrence Erlbaum Associates Publishers.

Lexicon of Common Words in Contemporary Chinese Research Team. (2008). *Lexicon of common words in contemporary Chinese*. Commercial Press.

Li, X., & Pollatsek, A. (2020). An integrated model of word processing and eye-movement control during Chinese reading. *Psychological Review, 127*(6), 1139–1162. https://doi.org/10.1037/rev0000248

Li, X., Rayner, K., & Cave, K. R. (2009). On the segmentation of Chinese words during reading. *Cognitive Psychology, 58*(4), 525–552. https://doi.org/10.1016/j.cogpsych.2009.02.003

Luo, X., Sun, M., & Tsou, B. K. (2002). Covering ambiguity resolution in Chinese word segmentation based on contextual information. In P*roceedings of the 19th international conference on computational linguistic* (Vol. 1., pp. 1–7). Association for Computational Linguistics.

Ma, G., Li, X., & Rayner, K. (2014). Word segmentation of overlapping ambiguous strings during Chinese reading. *Journal of Experimental Psychology: Human Perception and*

Performance, 40(3), 1046–1059. https://doi.org/10.1037/a0035389

Nandrino, J. L., Dodin, V., Martin, P., & Henniaux, M. (2004). Emotional information processing in first and recurrent major depressive episodes. *Journal of Psychiatric Research*, *38*(5), 475–484. https://doi.org/10.1016/j.jpsychires.2004.03.002

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

R Core Team. (2022). *R: A language and environment for statistical computing* (Version 4.2.2) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org.

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, *110*, Article 104038. https://doi.org/10.1016/j.jml.2019.104038

Schacht, A., & Sommer, W. (2009). Time course and task dependence of emotion effects in word processing. *Cognitive, Affective, & Behavioral Neuroscience, 9*(1), 28-43. https://doi.org/10.3758/CABN.9.1.28

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

Scott, G. G., O'Donnell, P. J., Leuthold, H., & Sereno, S. C. (2009). Early emotion word processing: Evidence from event-related potentials. *Biological Psychology, 80*(1), 95–104. https://doi.org/10.1016/j.biopsycho.2008.03.010

Trew, J. L. (2011). Exploring the roles of approach and avoidance in depression: An integrative model. *Clinical Psychology Review*, *31*(7), 1156–1168. https://doi.org/10.1016/j.cpr.2011.07.007

Wang, Y., Zhou, L., & Luo, Y. (2008). The pilot establishment and evaluation of Chinese affective words system. *Chinese Mental Health Journal, 22*(8), 608–612.

Wang, Z., Yuan, C., Huang, J., Li, Z., Chen, J., Zhang, H, ... & Xiao, Z. (2011). Reliability and validity of the Chinese version of Beck Depression Inventory-II among depression patients. *Chinese Mental Health Journal*, 25(6), 476–480.

Yao, Z., Wu, J., Zhang, Y., & Wang, Z. (2017). Norms of valence, arousal, concreteness, familiarity, imageability, and context availability for 1,100 Chinese words. *Behavior Research Methods, 49*, 1374–1385. https://doi.org/10.3758/s13428-016-0793-2

Yang, W., & Xiong, G. (2016). Screening for adolescent depression: Validity and cut-off scores for depression scales. *Chinese Journal of Clinical Psychology*, *24*(6), 1010–1015.

Zhang, D., He, W., Wang, T., Luo, W., Zhu, X., Gu, R., ... & Luo, Y. J. (2014). Three stages of emotional word processing: An ERP study with rapid serial visual presentation. *Social Cognitive and Affective Neuroscience*, *9*(12), 1897–1903. https://doi.org/10.1093/scan/nst188