Read Writ (2018) 31:607–625
https://doi.org/10.1007/s11145-017-9800-y

The time course of incremental word processing during Chinese reading

Junyi Zhou[1,2] · Guojie Ma[3] · Xingshan Li[1] · Marcus Taft[4]

Published online: 20 November 2017

Abstract In the current study, we report two eye movement experiments investigating how Chinese readers process incremental words during reading. These are words where some of the component characters constitute another word (an embedded word). In two experiments, eye movements were monitored while the participants read sentences with incremental words whose first two characters (Experiment 1) or last two characters (Experiment 2) constituted a word (referred to respectively as "head-embedded" and "tail embedded"). Reading times on these words were longer when the frequencies of the embedded words were lower. However, this was only seen on first fixation duration for head-embedded words. These results suggest that embedded words are activated when Chinese readers process incremental words, and that this activation is earlier for a head-embedded word than for a tail-embedded word. These results support a hierarchical model which assumes that the representation for whole word is activated via the representation of its constituent morphemes.

Keywords Word processing · Eye movements · Chinese reading

✉ Xingshan Li
lixs@psych.ac.cn

1 CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, University of Chinese Academy of Sciences, 16 Lincui Road, Beijing 100101, China

2 School of Physical Education and Sport Science, Fujian Normal University, 1 Keji Rd., Minhou District, Fuzhou 350117, Fujian, China

3 Shaanxi Key Laboratory of Behavior and Cognitive Neuroscience, School of Psychology, Shaanxi Normal University, Xi'an, China

4 School of Psychology, University of NSW, Sydney, Australia

Introduction

Unlike most alphabetic writing systems, there are no explicit word boundaries in the Chinese writing system, such as spaces between words. Thus, Chinese readers have to group characters into words using high-level knowledge of the language. Even so, Chinese readers have no difficulties in reading text (Bai, Yan, Liversedge, Zang, & Rayner, 2008; Liversedge et al., 2014; Zang, Liang, Bai, Yan, & Liversedge, 2013), and the question of how this is achieved has attracted much interest in recent times (e.g., Li, Rayner, & Cave, 2009; Ma, Li, & Rayner, 2014; Zang et al., 2015). In the present work, we focus on how Chinese readers process *incremental words*, which are multi-character words containing a subset of characters that constitute another, related word (referred to as the *embedded word*). Such words require more processing resources for readers in order to determine whether to combine the embedded word with the upcoming character or to start the next word with that upcoming character than those three-character words where none of the two-character combinations can constitute words (e.g. 维生素, *vitamin*). For example, in a word containing three characters ABC (e.g. 老板娘, *landlady*), two of its characters AB (e.g. 老板, *boss*) constitute an embedded word that is transparently related in meaning to the whole word. In such cases, it may well be challenging for readers to decide whether AB constitutes an independent word in the sentence context, or whether it combines with the character C to constitute another word. Ultimately, readers will need to adopt the latter analysis since it is the one that will be appropriate for the sentence context. This raises two interesting questions that need to be explored: (1) whether the embedded word (a morphological constituent) is activated during the processing of a Chinese incremental word, and (2) if the embedded word is activated, whether it is activated earlier, later, or simultaneously in relation to the activation of the full form of the word.

Similar questions have been asked with regard to compound words during the reading of alphabetic text. Many psycholinguistic studies have been conducted to examine whether morphological constituents of compound words are activated during the reading of alphabetic scripts (e.g., Andrews, Miller, & Rayner, 2004; Beauvillain, 1996; Gagné & Spalding, 2009; Hyönä & Pollatsek, 1998; Juhasz, Starr, Inhoff, & Placke, 2003). Several of these studies have found effects of constituent frequency (e.g., Hyönä et al., 1998; Kuperman, Bertram, & Baayen, 2008; Kuperman, Schreuder, Bertram, & Baayen, 2009), showing that reading time is longer for words with low-frequency constituents than for words with high-frequency constituents. The results of these studies have usually been taken as evidence to support the activation of constituents during compound word processing.

Several previous studies in Chinese reading have also found that the lexical characteristics of the morphological constituents influence processing of Chinese two-character compound words (e.g., Cui et al., 2012; Janssen, Bi, & Caramazza, 2008; Tsang & Chen, 2010, 2013, 2014; Wang & Peng, 1999; Yan, Tian, Bai, & Rayner, 2006; Zhou, Marslen-Wilson, Taft, & Shu, 1999). However, it is noteworthy that the single characters contained within a two-character word do not always carry unambiguous meaning. For example, in the word老师 (*teacher*), the first character 老 could have several meanings (e.g., *old*, *always*, or *experienced*). In contrast, the two-character embedded word in a three-character incremental word always possesses an unambiguous meaning which, therefore, provides us with an opportunity to study how the full form of a word and its morphological constituents are activated during natural reading.

How do Chinese readers segment the incremental words in Chinese scripts? One possibility is the *unidirectional parsing hypothesis* considered by Inhoff and Wu (2005), which assumes that grouping characters into words is strictly sequential from left to right in Chinese reading. According to this hypothesis, characters on the left have priority in processing, but if such left-to-right parsing turns out to be inappropriate, it will impede the segmentation of the character string. Perfetti and Tan (1999) embedded overlapping ambiguous strings into sentences where the correct segmentation style should always be A-BC. The control sentence was identical to the experimental sentence except that the first character of the ambiguous strings was substituted by another synonym so that there was no overlapping ambiguity in the control sentence and the meaning of sentence was unchanged. They found that reading times after the overlapping ambiguous strings were longer than that in the control condition. This result showed that subjects might combine the first two characters of the ambiguous strings into a word in first-pass reading so that they have to correct the segmentation when they read the later part of the sentence which then leads to the processing cost.

However, other studies in Chinese have generated findings that the unidirectional parsing hypothesis cannot explain. Inhoff and Wu (2005) found that even if a character can legally combine with the character on its left to constitute a word, it will affect reading performance when it can also combine with the character on its right. In their study, different four-character strings were embedded in the same sentence frames where both the first two and last two characters constituted a word. In the ambiguous condition the central two characters also constituted a two-character word while in the control conditions they did not. For example, in the ambiguous condition, the four characters ABCD (e.g. 专科学生, *college students*) could constitute three two-character words AB (专科, *college*), BC (科学, *science*), and CD (学生, *students*). However, in the unambiguous control conditions, the central two characters in both ABEF (专科毕业, *college graduation*) and GHCD (外地学生, *nonlocal students*) did not constitute a word. They found that first-pass reading times and total reading times were longer in the ambiguous condition than in either of the control conditions. The unidirectional parsing hypothesis predicts that once a character is assigned to the character on its left to constitute a word, it will not be assigned to the character on its right even if it can combine with it to form a word. So, if parsing strictly follows character order from left to right, the lexical status of BC should have no influence on the processing of ABCD. Therefore, the results of this study oppose the unidirectional parsing hypothesis.

Ma et al. (2014) also found that the unidirectional parsing hypothesis cannot explain the processing of Chinese overlapping ambiguous strings. They inserted Chinese overlapping ambiguous strings (ABC, where both AB and BC are two-character words) into one of two sentence frames so that it could be either segmented as AB-C or A-BC depending on the sentence context. They found that Chinese readers were more likely to segment the string as A-BC (rather than AB-C) when the word on the right (BC) had a higher frequency than the left-hand one (AB). More recently, Ma, Pollatsek, Li, and Li (2017) investigated whether readers can perceive a word even when it is composed of noncontiguous characters. Chinese four-character strings ABCD in which both AB and CD were two-character words were used as materials, and it was found that the status of AC as a word slowed the reading of CD even though AC were not adjacent to each other. These studies appear to show, then, that word boundaries are not solely determined by characters on the left side of the character strings. Thus, none of these results can be explained by a unidirectional model.

Li et al. (2009) proposed an alternative account of Chinese word segmentation for which the processing of incremental words provides potential evidence. According to this computational model, Chinese word recognition is an interactive process involving multiple levels. The first is a visual perception level that abstracts visual features from the stimulus. The second is a level that recognizes characters using perceptual information from the first level, with multiple character recognizers that work in parallel. Finally, there is the word recognition level that allows segmentation of the sentence into words. This level receives information from both the character recognizers and the lexicon, and sends feedback to the character recognition level. After the characters in the perceptual span are processed in parallel, the activated character information feeds forward to the word-processing level, and all of the words associated with these characters are activated. These activated words then compete for a single winner and as a word unit wins the competition, it is ultimately recognized and the word is therefore treated as a single entity. As this model borrows some assumptions of the interactive-activation framework of McClelland and Rumelhart (1981), it also assumes that representations within the same level will compete with each other while related representations between different levels will facilitate each other.

The model proposed by Li et al. (2009) can successfully explain the results of previous studies. When incremental words are processed, the two-character and three-character words are in competition with each other. However, feedforward information from three characters will provide greater activation than feedforward information from only two characters, which means that three-character words are more likely to win the competition. The frequency of the embedded words will also influence the time needed to settle the competition. Therefore, a two-character word will provide stronger competition to a three-character word in which it is embedded if it is relatively high in frequency. As a result, the model predicts that the identification of an incremental word will be delayed if it contains a high-frequency embedded word more than if it contains a low-frequency one. Previous studies have obtained results consistent with this. Weingartner, Juhasz, and Rayner (2012) found that higher-frequency embedded words interfered with the processing of a target word during reading, while Davis, Perea, and Acha (2009) found longer gaze durations and total fixations on target words that had a higher frequency embedded word in outer position than those that had no embedded word.

Another possible alternative is the hierarchical model proposed by Taft (2003, 2004) where lexical representations exist at a lemma level that mediates between form representations and meaning. The model assumes that the lemma for a polymorphemic word is activated via the lemmas of its constituent morphemes, which means that the whole word and its constituent morphemes are not activated in parallel. When incremental words are processed, the lemma for the whole word will be activated through the lemma for the embedded word if they are related in meaning. Therefore, the frequency of the embedded words will also influence the time for processing the whole word, but in the reverse direction to the model proposed by Li et al. (2009). Specifically, as the lemma for a high-frequency embedded word is more readily activated than the lemma for a low-frequency one, the higher the frequency of the embedded word, the faster the activation of the lemma for a whole word that contains it. As a result, the model predicts that the identification of an incremental word will be facilitated when it contains a high-frequency embedded word compared to when it contains a low-frequency one.

Research on Japanese trimorphemic compounds by Miwa, Libben and Ikemoto (2017) is relevant here. Participants were asked to perform a lexical decision task while their eye movements were recorded. On first fixation durations, the results showed an effect of whole-word frequency, but not of embedded-word frequency which was only observed on second subgaze durations. In addition, no interaction effect between whole-word frequency and embedded-word frequency was found on any measures. These results were taken to mean that the whole word is activated earlier than the embedded word rather than later or simultaneous with it, which is not consistent with the hierarchical account. The present study will further test the prediction of the hierarchical model.

As can be seen, theoretical consequences follow from an examination of the relative time course of activation of the whole word and its constituents. A study by Kuperman et al. (2009) has explored the temporal order of activation of full forms of Dutch compounds and their morphological constituents within a corpus study of uncontrolled eye movements. One of their major findings was that first fixation and gaze durations were shorter for high-frequency than low-frequency compounds, and were shorter for words with high-frequency than with low-frequency left-hand constituents. Moreover, the effect of compound frequency was larger for words with low-frequency than with high-frequency left-hand constituents. It was concluded from these results that the activation of the compound's full form coincides with the activation of the compound's left-hand constituent, and the interaction of the impact of their frequency shows that these two representations are processed simultaneously and are not independent. As such, Kuperman et al. (2009) favored a model which allows access to multiple routes simultaneously, namely, one that is more in line with the account of Li et al. (2009) than the hierarchical model where the whole word is accessed via its constituents.

Note, however, that the study of Kuperman et al. (2009) was a corpus analysis drawing conclusions from correlational statistics rather than controlled experiments, and Rayner et al. (2007) have argued that there are problems in interpreting such analyses. Even though a correlation indicates that there is an effect of a variable, it is difficult for one to know how much of the effect is due to this variable or other confounds. It is also difficult to assess how much of the joint variance is due to one or more variables. For these reasons, Rayner et al. (2007) believed that controlled experiments are a better source of data because they allow for stronger inferences about causality. Thus, it can be argued that the results from a corpus analysis are not sufficient and that well-controlled experiments are needed to further understand the time course of the activation of constituents.

Another important issue to be considered is the influence of embedding position on activation of morphological constituent. Research on Finnish compound words has shown that there is an early effect of first-constituent frequency and an absence of an early effect of second-constituent frequency (Bertram & Hyönä, 2003; Pollatsek, Hyönä, & Bertram, 2000), revealing that the activation of constituents varies with position. Constituents on the left are activated at an earlier stage than constituents on the right. Bertram and Hyönä (2003) put forward a visual acuity principle to account for the results of previous experiments whereby whatever information is available in foveal vision is processed. Therefore, when the whole word is within foveal vision, word recognition takes place through the whole-word representation. However, when only the first constituent is available in foveal vision, word recognition takes place through the first-constituent representation, followed by the second-constituent representation and whole-word representation.

In the present study, two eye movement experiments were conducted. In Experiment 1, we used incremental words whose first two characters constitute a related word while their last two characters do not (i.e., “head-embedded” incremental words). Two sets of word pairs were included. In one set, whole-word frequency was manipulated while embedded-word frequency was held constant, while in the other set, embedded-word frequency was manipulated while whole-word frequency was held constant. These manipulations enabled us to examine whether the frequency of the whole word and embedded word affected fixation durations on the target word. The model proposed by Li et al. (2009) predicts a whole-word frequency effect, as well as greater inhibition of whole-word recognition with higher embedded-word frequency. As discussed earlier, this model assumes that characters in the perceptual span activate all of the words they can constitute, and these compete for a single winner. Therefore, higher embedded-word frequency implies a higher level of activation of the embedded word, which will hinder the whole word from winning the competition. In contrast, the hierarchical model (Taft, 2003, 2004) predicts a positive effect of both whole-word and embedded-word frequency. Because the whole word is activated via its constituents, faster activation of the embedded word will facilitate recognition of the whole word.

Experiment 2 examined whether there is any difference between the activation of embedded words located at different positions. In contrast to Experiment 1, the embedded words of Experiment 2 constituted the last two characters of the incremental words rather than their first two characters (i.e., “tail-embedded” incremental words). Since Chinese readers read from left to right, one might expect only the left-hand embedded words to be processed in combination, with no impact of tail-embedding. To examine this, pairs of words were selected in such a way that embedded-word frequency was manipulated while whole-word frequency was held constant. If an embedded word is activated despite being on the right-hand side of the incremental word, the frequency of that embedded word would be expected to affect fixation durations on the target word. On the other hand, if a tail-embedded word is not activated at all, its frequency should have no effect on eye movement measures.

Experiment 1 (head-embedded words)

Method

Participants

Twenty-eight undergraduate students (all native speakers of Chinese, 13 female and 15 male) from colleges around the Institute of Psychology, Chinese Academy of Sciences were paid to participate in the present study. All had either normal or corrected-to-normal vision. Their ages ranged from 20 to 28 years, with an average of 23 years.

Materials

Forty-four pairs of three-character incremental words were selected as target items. Their first two characters could form an independent word (which was a morphological constituent), while the last two characters could not (e.g., 广告画, *poster* where 广告 means *advertisement*). The words were divided into two sets: an embedded-word frequency set (22 pairs) and a whole-word frequency set (22 pairs). We obtained the frequency data of all the items from the Chinese Linguistic Data Consortium (2003). In the embedded-word frequency set, we manipulated the frequency of the embedded word (which was defined as the frequency of the embedded word when used as a word alone) while matching for whole-word frequency such that the words in each pair differed substantially in embedded-word frequency while their whole-word frequency was equated. In the whole-word frequency set, we manipulated the frequency of whole word while matching for embedded-word frequency such that the words in each pair differed substantially in whole-word frequency while their embedded-word frequency was equated. The characteristics of the stimuli are presented in Table 1. Character frequency and stroke number in each character position of the 3-character word were matched to the extent that they were not significantly different between two conditions.

For both sets, we embedded each pair of words into the same sentence frame without causing any ambiguity or implausibility (see Table 2), whereby we could examine whether the frequency of the embedded word or the whole word affected fixation durations. Ten native Chinese speakers from universities in Beijing near the Institute of Psychology, Chinese Academy of Sciences who did not participate in the main experiment were recruited to assess the plausibility of the sentences. They were assigned to one of two counterbalanced lists and were asked to rate the plausibility of each sentence on a 7-point scale (1 = very implausible; 7 = very plausible). Importantly, there was no difference in plausibility between the two

Table 1 Stimulus Characteristics of Experiment 1 (word frequency refers to occurrences per million)

Stimulus characteristic	Whole-word set				Embedded-word set			
	High	Low	t	p	High	Low	t	p
Mean whole-word frequency	37.36 (19.03)	1.34 (0.85)	8.96	< 0.001	0.62 (0.46)	0.58 (0.44)	1.35	0.19
Mean embedded-word frequency	194 (182.9)	195 (183.2)	0.91	0.37	88.09 (49.61)	1.59 (0.73)	8.18	< 0.001
Mean first character frequency	2178 (2674.6)	2145 (2682.4)	0.39	0.69	760 (627.1)	846 (618.3)	0.96	0.35
Mean second character frequency	1934 (1233.0)	2076 (1114.9)	0.93	0.36	753 (587.5)	754 (784.1)	0.01	0.99
Mean third character frequency	2888 (2777.2)	2217 (2145.8)	1.53	0.14	1508 (2280.0)	1414 (2189.9)	0.59	0.56
Mean first character stroke	8.0 (3.0)	7.2 (2.8)	1.44	0.17	7.7 (3.0)	8.5 (3.3)	1.12	0.28
Mean second character stroke	7.2 (1.9)	6.9 (1.7)	0.68	0.5	8.8 (2.5)	8.5 (2.9)	0.55	0.59
Mean third character stroke	7.5 (2.1)	7.8 (2.1)	0.59	0.56	8.6 (3.4)	8.4 (3.2)	0.37	0.71

Table 2 Sample stimuli of Experiment 1

Stimuli set	Condition	Target	Sentence
Embedded-word set	High	广告画	马杰坐在书桌前构思广告画所要采用的构图样式
(22 pairs)		poster	(Ma Jie is sitting in front of his desk and thinking about the style of composition for the poster)
	Low	水粉画	马杰坐在书桌前构思水粉画所要采用的构图样式
		gouache	(Ma Jie is sitting in front of his desk and thinking about the style of composition for the gouache)
Whole-word set	High	艺术家	这位固执的艺术家在会上努力说服大家接受他的观点
(22 pairs)		artist	(The obstinate artist tried to persuade everyone to accept his views at the meeting)
	Low	改革家	这位固执的改革家在会上努力说服大家接受他的观点
		reformer	(The obstinate reformer tried to persuade everyone to accept his views at the meeting)

whole word conditions (6.1 for the HF condition, 6.3 for the LF condition; $t < 1$) and the two embedded word conditions (6.2 for the HF condition, 6.3 for the LF condition; $t < 1$).

Forty-four sentence frames were constructed, and these sentences were presented in two blocks, with the embedded-frequency word set in one block, and the whole-word frequency set in the other. Each block also contained six practice sentences, 22 experimental sentences, and 29 fillers. The filler sentences contained no incremental words, and were included simply to minimize the chances that the participants would notice the purpose of the experiment. The order of sentences was randomized within each block, and the order of blocks was counterbalanced across subjects using a Latin square design. Each sentence was presented in one of the two presentation conditions for half of the subjects. The length of all the sentences ranged from 21 to 30 characters. None of the characters of the critical words appeared in the first five or the last five character positions of the sentence.

Procedure

The materials were presented on a 21-inch CRT monitor (Sony G520; resolution: 1024 × 768 pixels; refresh rate: 150 Hz) connected to a DELL PC. Every sentence was displayed on a single line in Song 20-point font in white (RGB: 255, 255, 255) on a black background (RGB: 0, 0, 0). Participants were seated at a viewing distance of 58 cm from the computer monitor, with each character subtending a visual angle of approximately 0.7°. A chin rest and a forehead rest were used to stabilize participants' heads. Participants read sentences binocularly, but only the right eye was monitored. Eye movements were recorded by an SR Research Eyelink 1000 eye tracking system with a sampling rate of 1000 Hz.

After the participants had read the experimental instructions and a brief description of the apparatus, the chair was adjusted to make them feel comfortable and the eye tracker was calibrated using a three-point calibration and validation procedure. The maximal error of validation was below 0.5° in visual angle. At the beginning of each trial, a white character-sized square appeared on the left side of the computer screen, which indicated the position of the first character in the sentence. Once the participant successfully fixated on the white square, a sentence was displayed. The participants were asked to read the sentences silently and to answer some comprehension questions following approximately one-third of the sentences. After reading each sentence, the participant pressed a response button to start the next trial.

Results and discussion

The accuracy of answers to the comprehension questions was high (93%) which indicates that the participants understood the sentences well. Trials in which participants made more than three blinks in the whole sentence or made one or more blinks on the target word were excluded from the analysis. This resulted in a loss of 2.36% of the trials. Fixations with durations longer than 1000 ms or shorter than 80 ms (approximately 2% of all fixations) were also excluded from the analysis.

We report eye movement measures for the embedded-word region (i.e., the first two characters of the target word) and the whole-word region (i.e., the three characters of the target word). For both regions, the following eye movement measures were analyzed: first fixation duration, gaze duration (the sum of all first-pass fixations on the region before moving to another region), and total reading time (the sum of all fixations on the target region including any regressions to it).

Data were analyzed using linear mixed-effects models (LMM). In the LMMs, embedded-word frequency (high vs. low) or whole-word frequency (high vs. low) was entered as a fixed effect, and subjects and items were entered as crossed random effects, including intercepts and slopes (see Baayen, Davidson, & Bates, 2008). The high-frequency condition was set arbitrarily as the intercept (baseline condition), and the low-frequency condition was compared with it. The lmer function from the lme4 package (Bates, Maechler, Bolker & Walker, 2014) was used within the R Environment for Statistical Computing (R Development Core Team, 2016). For fixation duration measures, we used linear mixed-effect regression and report regression coefficients (b) which estimate the effect size, and the t-value of the effect coefficient. We also estimated and report the p-values for the effects by using the summary function from lmerTest package (Kuznetsova, Brockhoff & Christensen, 2014).

Whole-word frequency set

Eye movement measures and results of linear mixed effects analyses on these measures with the whole-word frequency manipulation are reported in Table 3. There were significant effects of whole-word frequency for all eye movement measures in both the whole-word and embedded-word regions. First fixation, gaze duration, and total reading time were all shorter in the high-frequency condition than in the low-frequency condition suggesting that activation of the representation of the whole word continues throughout all stages of processing.

Table 3 Eye movement measures (means and standard errors) and results of the linear mixed effects models for these measures on two interested region with a manipulation of whole-word frequency in Experiment 1

Region	Measure	High	Low	*b*	*SE*	*t*	*p*
Whole word	First fixation duration	259 (9.0)	286 (8.4)	28.23	7.35	3.84	< 0.001
	Gaze duration	351 (16.6)	400 (16.5)	54.28	13.74	3.95	< 0.001
	Total time	490 (33.3)	562 (29.4)	80.29	21.22	3.78	< 0.001
Embedded word	First fixation duration	256 (9.8)	283 (8.6)	26.96	7.62	3.54	< 0.001
	Gaze duration	279 (12.5)	315 (12.0)	36.22	10.81	3.35	< 0.001
	Total time	373 (21.5)	413 (23.1)	44.5	18.42	2.42	0.02

Embedded-word frequency set

Eye movement measures and results of linear mixed effects analyses on these measures with the embedded-word frequency manipulation are reported in Table 4. As with whole-word frequency, significant effects of embedded-word frequency were found in both the whole-word and embedded-word regions for all eye movement measures. Again, first fixation and gaze durations were shorter in the high-frequency condition than in the low-frequency condition, along with shorter total reading times. These results indicate that the embedded word, like the whole word, is activated at an early stage of processing and this activation is maintained throughout the processing of the whole word.

Experiment 2 (tail-embedded words)

Method

Participants

Twenty-eight undergraduate students (all native speakers of Chinese, 13 female and 15 male) from the same participant pool used in Experiment 1 were paid to

Table 4 Eye movement measures (means and standard errors) and results of the linear mixed effects models for these measures on two interested region with a manipulation of embedded-word frequency in Experiment 2

Interest region	Measure	High	Low	*b*	*SE*	*t*	*p*
Whole word	First fixation duration	263 (6.9)	284 (7.9)	23.83	8.36	2.85	< 0.01
	Gaze duration	386 (17.3)	431 (18.4)	48.01	15.3	3.14	< 0.01
	Total time	530 (35.9)	618 (32.8)	87.22	22.73	3.84	< 0.001
Embedded word	First fixation duration	262 (7.1)	285 (8.2)	24.33	8.89	2.74	0.01
	Gaze duration	291 (9.9)	346 (13.5)	54.9	12.17	4.51	< 0.001
	Total time	387 (26.0)	491 (27.0)	102.06	20.61	4.95	< 0.001

participate in the present study. None of them participated in Experiment 1. Their ages ranged from 19 to 32 years, with an average of 23 years.

Materials

Thirty-four pairs of three-character incremental words were selected as target items. Unlike Experiment 1, the last two characters formed an independent word that was a morphological constituent of the incremental words, while the first two characters did not (e.g. 微电脑, *microcomputer*, where 电脑 means *computer*). Using frequency data from the Chinese Linguistic Data Consortium (2003), we manipulated the frequency of the embedded word when used as a word by itself, while matching for whole-word frequency. In this experiment, only embedded-word frequency was examined because there were insufficient items to manipulate whole-word frequency while holding embedded-word frequency constant. The characteristics of the stimuli are presented in Table 5. Character frequency and stroke number in each character position of the 3-character word were matched to the extent that they were not significantly different between the high and low frequency conditions.

As in Experiment 1, each pair of words was embedded into the same sentence frame (see Table 6), Sentence plausibility was again assessed on the basis of ratings on a 7-point scale by ten native Chinese speakers from universities in Beijing near Institute of Psychology, Chinese Academy of Sciences who did not participate in the main experiment. There was no difference in plausibility between the two conditions (5.8 for the HF condition, 5.9 for the LF condition; $t < 1$).

Two lists of items were constructed, with each participant reading one list that contained 6 sentences for practice, followed by 34 experimental sentences and 32 filler sentences in a random order. Each sentence was presented in one of the two conditions for half of the subjects. The length of the sentences ranged from 20 to 26 characters. None of the characters of the critical words appeared in the first five or the last five character positions of the sentence.

Table 5 Stimulus Characteristics of Experiment 2 (word frequency refers to occurrences per million)

Stimulus characteristic	Whole-word set			
	High	Low	*t*	*p*
Mean whole-word frequency	0.49 (0.4)	0.54 (0.6)	1.03	0.31
Mean embedded-word frequency	108 (114)	1.95 (0.7)	5.42	< 0.001
Mean first character frequency	753 (796)	1035 (1322)	1.01	0.32
Mean second character frequency	1299 (982)	1068 (1872)	0.71	0.48
Mean third character frequency	1144 (1184)	1179 (1252)	0.14	0.89
Mean first character stroke	7.7 (2.9)	7.1 (2.5)	0.89	0.38
Mean second character stroke	7.4 (2.5)	8 (2.9)	0.92	0.37
Mean third character stroke	7.8 (2.9)	8.3 (3.2)	0.77	0.45

Table 6 Sample stimuli of Experiment 2

Condition	Target	Sentence
High embedded-word frequency	微电脑	这位政治家认为微电脑的出现极大地影响了世界经济模式
	Microcomputer	(This politician believes that the invention of the microcomputer dramatically influenced the nature of the world economy)
Low embedded-word frequency	金本位	这位政治家认为金本位的出现极大地影响了世界经济模式
	Gold standard system	(This politician believes that the invention of the gold standard system dramatically influenced the nature of the world economy)

Procedure

The procedure was the same as in Experiment 1.

Results and discussion

The accuracy of answers to the comprehension questions was again high (95%) which indicates that the participants were successfully processing the sentences for meaning. Trials in which participants made more than three blinks in the whole sentence or made one or more blink on the target word were excluded from the analysis. This resulted in a loss of 2.5% of the trials. Fixations with durations longer than 1000 ms or shorter than 80 ms (approximately 1% of all fixations) were also excluded from the analysis.

The data were analyzed in the same way as in Experiment 1, using LMM. In the LMMs, embedded-word frequency (high vs. low) was entered as a fixed effect, and subjects and items were entered as crossed random effects, including intercepts and slopes. The high-frequency condition was set arbitrarily as the intercept (baseline condition), and the low-frequency condition was compared with it.

Eye movement measures and results of linear mixed effects analyses on these measures with the embedded-word frequency manipulation are reported in Table 7. In neither the whole-word nor embedded word regions were first fixation durations significantly different between the high-frequency and low-frequency embedded word conditions. However, both gaze durations and total reading times were shorter in the high-frequency than in the low-frequency condition.

These results show that the frequency of an embedded word does have an impact on eye fixations when reading an incremental word, hence indicating that pairs of characters are processed together even when they are not the first two characters encountered in the reading of an incremental word. However, the fact that there was no effect of embedded-word frequency on first fixations suggests a later locus of processing of such tail-embedded words than was the case for head-embedded words.

Table 7 Eye movement measures (means and standard errors) and results of the linear mixed effects models for these measures on two interested region with a manipulation of embedded-word frequency in Experiment 2

Interest region	Measure	High	Low	*b*	*SE*	*t*	*p*
Whole word	First fixation duration	260 (6.9)	263 (8.5)	6.14	6.91	0.89	0.37
	Gaze duration	394 (19.9)	420 (23.9)	26.08	11.67	2.24	0.03
	Total time	569 (30.2)	623 (31.9)	85.53	17.65	4.85	< 0.001
Embedded word	First fixation duration	265 (8.9)	269 (10.4)	3.26	5.88	0.56	0.57
	Gaze duration	313 (11.7)	335 (20.0)	29.20	13.86	2.11	0.04
	Total time	418 (18.5)	500 (25.3)	59.52	21.42	2.78	0.01

General discussion

In the present study, we examined whether the reading of a Chinese incremental word entails the activation of its embedded word and. Such an examination sheds some light on understanding how Chinese readers process words when they have greater internal complexity during reading.

In Experiment 1, we manipulated the frequency of head-embedded incremental words in one item set, and the frequency of the embedded words themselves in another item set and found that first fixation durations, gaze durations and total reading times were shorter with higher frequency of either type, regardless of whether the region being measured was the target word or the embedded word. In Experiment 2, the embedded words were on the right of the whole words, and the results showed that the effect of embedded-word frequency was only observed on gaze durations and total reading times. In general, such results suggest that embedded words are activated when Chinese readers process incremental words. This outcome is consistent with previous studies (Bertram & Hyönä, 2003; Davis & Taft, 2005; Hyönä & Pollatsek, 1998; Kuperman, Bertram, & Baayen, 2008; Kuperman et al., 2009; Pollatsek, Bertram, & Hyönä, 2011; Taft, 1979, 2004; Taft & Forster, 1975, 1976) which showed that morphological constituents are activated during whole word processing.

The different results for head-embedded and tail-embedded words on the first fixation duration suggest that the representations of the former are activated at an earlier stage of processing than the latter. This difference can be accounted for by the visual acuity principle (Hyönä, 2012). Previous studies on eye fixations in Chinese compound word reading have shown that the preferred viewing location is located near the beginning of the word (Li, Liu, & Rayner, 2014; Zang et al., 2013). Therefore, it can be assumed that the embedded word is available in foveal vision at the initial processing stage of a head-embedded incremental word, but not for the initial processing stage of a tail-embedded incremental word. As such, the effect of embedded-word frequency will only be seen on first fixation durations for the head-embedded words.

The present results are also in support of the hierarchical model (Taft, 2003, 2004) which assumes that the representation for whole word is activated via the representation of its constituent morphemes. The embedded words were always activated whether or not they were in head-embedded or tail-embedded incremental words. Moreover, the early effect of embedded-word frequency for the former and the null effect of embedded-word frequency for the latter implies that the incremental words are more likely to be processed in a serial rather than parallel fashion. If incremental words were processed in a parallel fashion, embedded words would always be activated at an early stage whatever their location.

It should be noted that the results of Experiment 1 are inconsistent with the study of Miwa et al. (2017). Not only do we observe an effect of embedded-word frequency on first fixation, but we also find that the embedded-word frequency effect emerges earlier than the whole-word frequency effect. It is hard to explain why there is such a discrepancy between the two studies, but there were some important differences between them. First, the Miwa et al. (2017) results came from a corpus study in which the independent variables were not strictly manipulated. In contrast, the present study is a well-controlled experiment in which the independent variables were explicitly manipulated and other irrelevant variables were controlled. Second, Miwa et al. (2017) presented Japanese trimorphemic compounds in isolation while the words in the present study were embedded into sentence frames where the length of the character string in question is not physically demarcated, hence leading to different eye movement patterns. Finally, the task used by Miwa et al. (2017) was lexical decision where the existing words had to be discriminated from non-existing character strings. The potential strategies associated with performing such a task were eliminated in the present study by having participants simply read and comprehend sentences.

Although the present results demonstrate that both whole-word and embedded-word frequency have an effect on eye fixations, they do not support the prediction of competition between the embedded word and the whole word. The model of Li et al. (2009) predicts that higher embedded-word frequency should hinder the whole word from winning the competition, hence leading to longer fixation durations on the whole word. Instead, the results showed that fixation durations on the whole word were shorter when the embedded word was of higher frequency.

These results are very different from previous evidence generated from the study of overlapping ambiguous strings in Chinese reading. Ma et al. (2014) found that fixation durations on overlapping ambiguous character-strings ABC were longer when the word BC had a higher frequency than the word AB even though AB-C was the correct segmentation for the context. The present results reveal that even though the embedded word and the whole word are both activated, they do not compete with each other, but instead are in a facilitatory relationship. One possible reason for the discrepancy between the present study and that of Ma et al. (2014), is that the latter used overlapping ambiguous strings (ABC) as materials where there are two possible segmentations (i.e. AB-C or A-BC) depending on the context in which the string occurs. Therefore, the competition between AB and BC reflects the interference to ultimate segmentation from the alternative segmentation. In contrast, the incremental words used in the present study have only one correct segmentation that fits the context (i.e. segmenting incremental word from the rest of the context), hence eliminating competition.

The finding of a facilitatory effect of embedded-word frequency is consistent with the predictions of the hierarchical model (Taft, 1979, 2004; Taft & Forster, 1975, 1976). This model assumes that activation of the lemma for the embedded word (AB) combines with the activation of the lemma for the third character (C) to activate the lemma for the whole word. Therefore, the ease of activation of the lemma unit for the embedded word will influence the activation of the lemma unit for the whole word, hence explaining the observed effect of embedded-word frequency.

It should be noted that the concept of incremental word used in the present study is quite different from the concept of words that contain embedded words that are not semantically related, such as *hat* in *that* (Bowers, Davis, & Hanley, 2005). Given the nature of three-character incremental words in Chinese, the embedded word and the whole word in the current study were always semantically related. Because there are many of this kind of word in Chinese, the question of how they are processed is just as much of interest as when they are unrelated. Indeed, how to process this kind of word has proven a challenge for computer scientists when designing an artificial natural language processing system (Huang and Zhao 2007; Wu, 2007). Therefore, it is useful to address how Chinese readers process such words during reading even if it means that the present research may not be directly comparable to the previous studies that have examined semantically opaque embedded words (e.g., Inhoff, & Wu, 2005; Ma, et al., 2014). Therefore, further research is needed to examine the processing of incremental words whose embedded word and whole word have a different semantic relationship (e.g. 老虎钳, *pliers*, where 老虎 means *tiger*).

In conclusion, the results of the present research support the hypothesis of a hierarchical model where activation of the embedded word is a crucial step in the activation of an incremental word. The activation of the embedded word facilitates the activation of whole word rather than inhibiting it, as predicted by the model of Li et al. (2009). Moreover, the absence of an early frequency effect for embedded words in tail-embedded incremental words indicates that the position of embedding affects the time course of activation of embedded word when processing incremental words which implied that incremental words are likely to be processed in a serial fashion as claimed by hierarchical model. Overall, the present findings have implications for understanding the mechanisms of word recognition and word segmentation in Chinese reading.

Acknowledgements This research was supported by grants from the Natural Science Foundation of China (31571125), and the Beijing Advanced Innovation Center for Imaging Technology (BAICIT-2016018).

References

Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology, 16*(1–2), 285–311. https://doi.org/10.1080/09541440340000123.

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,* 390–412. https://doi.org/10.1016/j.jml.2007.12.005.

Bai, X., Yan, G., Liversedge, S. P., Zang, C., & Rayner, K. (2008). Reading spaced and unspaced Chinese text: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 34*(5), 1277–1287. https://doi.org/10.1037/0096-1523.34.5.1277.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed- effects models using Eigen and S4_. R package version 1.1-7.* http://CRAN.R-project.org/package=lme4.

Beauvillain, C. (1996). The integration of morphological and whole-word form information during eye fixations on prefixed and suffixed words. *Journal of Memory and Language, 35,* 801–820. https://doi.org/10.1006/jmla.1996.0041.

Bertram, R., & Hyönä, J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from eye movements when short and long finnish compounds. *Journal of Memory and Language, 48*(3), 615–634. https://doi.org/10.1016/S0749-596X(02)00539-9.

Bowers, J. S., Davis, C. J., & Hanley, D. A. (2005). Automatic semantic activation of embedded words: Is there a “hat” in “that”? *Journal of Memory and Language, 52,* 131–143.

Chinese Linguistic Data Consortium. (2003). 现代汉语通用词表 [Chinese lexicon] (CLDC-LAC-2003-001). Beijing, China: Tsinghua University, State Key Laboratory of Intelligent Technology and Systems, and Chinese Academy of Sciences, Institute of Automation.

Cui, L., Yan, G., Bai, X., Hyönä, J., Wang, S., & Liversedge, S. P. (2012). Processing of compound-word characters in reading Chinese: An eye-movement-contingent display change study. *Quarterly Journal of Experimental Psychology, 66*(3), 527–547. https://doi.org/10.1080/17470218.2012.667423.

Davis, C. J., & Taft, M. (2005). More words in the neighborhood: interference in lexical decision due to deletion neighbors. *Psychonomic Bulletin & Review*, *12*(5), 904–910.

Davis, C. J., Perea, M., & Acha, J. (2009). Re(de)fining the orthographic neighborhood: The role of addition and deletion neighbors in lexical decision and reading. *Journal of Experimental Psychology: Human Perception and Performance, 35*(5), 1550–1570. https://doi.org/10.1037/a0014253.

Gagné, C. L., & Spalding, T. L. (2009). Constituent integration during the processing of compound words: Does it involve the use of relational structures? *Journal of Memory and Language, 60*(1), 20–35. https://doi.org/10.1016/j.jml.2008.07.003.

Huang, C., & Zhao, H. (2007). Chinese word segmentation: A decade review. *Journal of Chinese Information Processing, 21*(2), 8–19.

Hyönä, J., & Pollatsek, A. (1998). Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception and Performance, 24*(6), 1612–1627.

Inhoff, A. W., & Wu, C. (2005). Eye movements and the identification of spatially ambiguous words during Chinese sentence reading. *Memory & Cognition, 33,* 1345–1356. https://doi.org/10.3758/BF03193367.

Janssen, N., Bi, Y., & Caramazza, A. (2008). A tale of two frequencies: Determining the speed of lexical access for Mandarin Chinese and English compounds. *Language and Cognitive Processes*. https://doi.org/10.1080/01690960802250900.

Juhasz, B. J., Starr, M. S., Inhoff, A. W., & Placke, L. (2003). The effects of morphology on the processing of compound words: Evidence from naming, lexical decisions and eye fixations. *British Journal of Psychology, 94*(Pt 2), 223–244. https://doi.org/10.1348/000712603321661903.

Kuperman, V., Bertram, R., & Baayen, R. H. (2008). Morphological dynamics in compound processing. *Language and Cognitive Processes, 23,* 1089–1132. https://doi.org/10.1080/01690960802193688.

Kuperman, V., Schreuder, R., Bertram, R., & Baayen, R. H. (2009). Reading polymorphemic Dutch compounds: Toward a multiple route model of lexical processing. *Journal of Experimental Psychology: Human Perception and Performance, 35*(3), 876–895. https://doi.org/10.1037/a0013484.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2014). *lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package)*: R package version 2.0-6. http://CRAN.R-project.org/package=lmerTest.

Li, X., Liu, P., & Rayner, K. (2014). Saccade target selection in Chinese reading. *Psychonomic Bulletin & Review*, *22*(2), 524–530.

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.

Liversedge, S. P., Zang, C., Zhang, M., Bai, X., Yan, G., & Drieghe, D. (2014). The effect of visual complexity and word frequency on eye movements during Chinese reading. *Visual Cognition, 22*(3–4), 441–457. https://doi.org/10.1080/13506285.2014.889260.

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.

Ma, G., Pollatsek, A., Li, Y., & Li, X. (2017). Chinese readers can perceive a word even when it's composed of noncontiguous characters. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 43*(1), 158–166. https://doi.org/10.1037/xlm0000298.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in basic findings. *Psychological Review, 88,* 375–407.

Miwa, K., Libben, G., & Ikemoto, Y. (2017). Visual trimorphemic compound recognition in a morphographic script. *Language, Cognition and Neuroscience, 32*(1), 1–20. https://doi.org/10.1080/23273798.2016.1205204.

Perfetti, C. A., & Tan, L. H. (1999). The constituency model of Chinese word identification. In J. Wang, A. W. Inhoff, & H.-C. Chen (Eds.), *Reading Chinese script: A cognitive analysis* (pp. 115–134). Mahwah, NJ: Erlbaum.

Pollatsek, A., Hyönä, J., & Bertram, R. (2000). The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology Human Perception & Performance*, *26*(2), 820–833.

Pollatsek, A., Bertram, R., & Hyönä, J. (2011). Processing novel and lexicalised Finnish compound words. *Journal of Cognitive Psychology, 23*(7), 795–810. https://doi.org/10.1080/20445911.2011.570257.

R Development Core Team. (2016). R: A language and environment for statistical computing [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from www.R-project.org.

Rayner, K., Pollatsek, A., Drieghe, D., Slattery, T. J., & Reichle, E. D. (2007). Tracking the mind during reading via eye movements: Comments on Kliegl, Nuthmann, and Engbert (2006). *Journal of Experimental Psychology: General, 136*(3), 520–537. https://doi.org/10.1037/0096-3445.136.3.520.

Taft, M. (1979). Recognition of affixed words and the word frequency effect. *The Memory & Cognition, 7*(4), 263–272. https://doi.org/10.3758/BF03197599.

Taft, M. (2003). Morphological representation as a correlation between form and meaning. In E. Assink & D. Sandra (Eds.), *Reading complex words* (pp. 113–137). Amsterdam: Kluwer.

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *The Quarterly Journal of Experimental Psychology Section A, 57*(4), 745–765. https://doi.org/10.1080/02724980343000477.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior, 14*(6), 638–647. https://doi.org/10.1016/S0022-5371(75)80051-X.

Taft, M., & Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior, 15,* 607–620. https://doi.org/10.1016/0022-5371(76)90054-2.

Tsang, Y.-K., & Chen, H.-C. (2010). Morphemic ambiguity resolution in Chinese: Activation of the subordinate meaning with a prior dominant-biased context. *Psychonomic Bulletin & Review, 17*(6), 875–881. https://doi.org/10.3758/PBR.17.6.875.

Tsang, Y.-K., & Chen, H.-C. (2013). Morpho-semantic processing in word recognition: Evidence from balanced and biased ambiguous morphemes. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 39*(6), 1990–2001. https://doi.org/10.1037/a0033701.

Tsang, Y.-K., & Chen, H.-C. (2014). Activation of morphemic meanings in processing opaque words. *Psychonomic Bulletin & Review, 6,* 1281–1286. https://doi.org/10.3758/s13423-014-0589-2.

Wang, C., & Peng, D. (1999). The role of surface frequencies, Cumulative morpheme frequencies, and semantic transparencies in the processing of compound words. *Acta Psychologica, 31*(3), 2–9.

Weingartner, K. M., Juhasz, B. J., & Rayner, K. (2012). Lexical embeddings produce interference when they are morphologically unrelated to the words in which they are contained: Evidence from eye movements. *Journal of Cognitive Psychology, 24*(2), 179–188. https://doi.org/10.1080/20445911.2011.604028.

Wu, A. (2007). Reflection on precedence relationship between word recognition and word segmentation. *Journal of Chinese Information Processing, 21*(3), 20.

Yan, G., Tian, H., Bai, X., & Rayner, K. (2006). The effect of word and character frequency on the eye movements of Chinese readers. *British Journal of Psychology, 97*(May), 259–268.

Zang, C., Liang, F., Bai, X., Yan, G., & Liversedge, S. P. (2013). Interword spacing and landing position effects during Chinese reading in children and adults. *Journal of Experimental Psychology: Human Perception and Performance, 39*(3), 720–734. https://doi.org/10.1037/a0030097.

Zang, C., Zhang, M., Bai, X., Yan, G., Paterson, K. B., & Liversedge, S. P. (2015). Effects of word frequency and visual complexity on eye movements of young and older Chinese readers. *Quarterly Journal of Experimental Psychology, 2006,* 1–17. https://doi.org/10.1080/17470218.2015.1083594.

Zhou, X., Marslen-Wilson, W., Taft, M., & Shu, H. (1999). Morphology, orthography, and phonology reading Chinese compound words. *Language and Cognitive Processes, 14*(5–6), 525–565.