

# Context Effects in Processing of Chinese Academic Words: An Eye-Tracking Investigation

Yu-Cin Jian

*National Central University, Zhongli,  
Taiwan*

Ming-Lei Chen

*National Hsinchu University of  
Education, Taiwan*

Hwa-wei Ko

*National Central University, Zhongli,  
Taiwan*

## ABSTRACT

This study investigated context effects of online processing of Chinese academic words during text reading. Undergraduate participants were asked to read Chinese texts that were familiar or unfamiliar (containing physics terminology) to them. Physics texts were selected first, and then we replaced the physics terminology with familiar words; other common words remained the same in both text versions. Our results indicate that readers experienced longer rereading times and total fixation durations for the same common words in the physics texts than for the corresponding texts. Shorter gaze durations were observed for the replaced words than the physics terminology; however, the duration of participants' first fixations on these two word types did not differ from each other. Furthermore, although the participants performed similar reading paths after encountering the target words of the physics terminology and replaced words, their processing duration of the current sentences was very different. They reread the physics terminology more times and spent more reading time on the current sentences containing the physics terminology, searching for more information to aid comprehension. This study showed that adult readers seemed to successfully access each Chinese character's meaning but initially failed to access the meaning of the physics terminology. This could be attributable to the nature of the formation of Chinese words; however, the use of contextual information to comprehend unfamiliar words is a universal phenomenon.

It is not unusual for adult readers to learn new common words almost daily. Previous studies have confirmed that the primary way people learn new words is by using contextual information to infer the meaning of the unfamiliar words (Carlisle, Fleming, & Gudbrandsen, 2000; Landauer & Dumais, 1997; Nagy, Anderson, & Herman, 1987). The interactive-compensatory model (Stanovich, 1980) was proposed to explain context effects; this was an extension of Rumelhart's (1977) interactive processes model for reading. In addition to synthesizing bottom-up (e.g., text information) and top-down (e.g., readers' background, lexical, and syntactic knowledge) processes, Stanovich's interactive-compensatory model supplemented the concept of compensation, asserting that it was not necessarily the case that higher level processes awaited the completion of the lower level processes. Instead, readers encountering obstacles to any particular process during reading would rely on information from the other levels.

Does this apply to new academic words? People frequently encounter unknown academic words presented in newspapers and popular science books without any formal definition or explanation. How do readers infer the meaning of these new words? Although some studies have investigated the processes involved in learning new words (Kuhn

& Stahl, 1998; Nagy et al., 1987; Williams & Morris, 2004), less research has been conducted on how people process or learn new academic words.

Empirical studies have been conducted using words or sentences as stimuli to investigate context effects in word recognition (Carlisle et al., 2000; Nagy et al., 1987; Stanovich, 1980, 1984; Stanovich, West, & Feeman, 1981). These experiments often used semantic priming paradigms to measure participants' reaction times in word recognition. Participants read aloud sentences in which contextual information was relevant, irrelevant, or neutral to the target words; the reaction time for naming the target word was recorded. Readers consistently showed processing time advantages for the target words in the relevant contextual condition. This result indicated that contextual information facilitated semantic prediction of the target word (Kim & Goetz, 1994; Schwantes, 1982; Stanovich et al., 1981; West, Stanovich, Feeman, & Cunningham, 1983).

In general, the results of empirical studies concerning context effects showed that readers with poor decoding skills could be led to rely more on contextual information than readers with good decoding skills could (Kim & Goetz, 1994; Schwantes, 1982; Stanovich, 1980, 1984; Stanovich et al., 1981; West et al., 1983) and that younger readers displayed larger context effects than older readers did (Kim & Goetz, 1994; Perfetti, Goldman, & Hogaboam, 1979; Stanovich et al., 1981). The interactiv-compensatory model could explain all of these results.

Although considerable research has been conducted on context effects, most studies have addressed the context effect in word recognition, and few have investigated the process of context effects in word processing (Brusnighan & Folk, 2012).

Eye tracking has long been used to study online reading processes (Rayner, 1998). In recent years, studies have used eye tracking to examine context effects in word recognition (Brusnighan & Folk, 2012; Chaffin, Morris, & Seely, 2001). One of the first of these studies, conducted by Chaffin et al., used eye movement data to examine how readers build the meanings of novel words from the context in which they are presented. Undergraduate students' eye movements were recorded while they read sentences containing high-familiarity, low-familiarity, and novel words in the same sentence contexts. The results showed that participants took longer to read novel and low-familiarity words than high-familiarity words within the same informative context, suggesting that adult readers inferred new words' meanings from contextual information. Chaffin et al. also found that readers spent a longer total reading time on informative context than uninformative context, implying that readers distinguished between information that was informative or uninformative to the target words.

Unlike Chaffin et al. (2001), who manipulated only word familiarity in the same context to investigate context effects, Brusnighan and Folk (2012) used different contextual information in their reading materials and added morphemic information as another independent variable to investigate how adult readers used contextual and morphemic information when they encountered novel compound words. Readers in this study read sentences containing novel and familiar English compound words that were either semantically transparent or opaque in informative and neutral sentence contexts. Each sentence frame contained two sentences: Sentence 1 provided information for discerning the meaning of the target word, and sentence 2 contained a synonym of the target word. Brusnighan and Folk hypothesized that if readers established the target word's meaning through the contextual information in sentence 1, then they would not rely on the contextual information presented next, in sentence 2.

Readers' eye movements for target words and sentences were analyzed. The results for target words showed that in the informative context, reading time was significantly longer during participants' first pass through familiar opaque compound words than for familiar transparent compound words; reading time was also longer for familiar transparent words presented within the neutral context than within the informative context. These findings indicated that initial processing time for familiar compound words was influenced by semantic transparency in informative sentence contexts and that contextual information facilitated familiar word recognition. The results for sentences showed that participants' reading time was unaffected by the synonymous anaphors presented in sentence 2, under either the informative or the neutral contexts. This suggested that adult readers had inferred the meaning of the target word while reading sentence 1; therefore, when they subsequently read sentence 2, no additional cognitive resources or reading time were required to infer the meaning of the target word.

These eye movement findings confirmed the action of the context effect during the process of word recognition. Although these data documented an interpretational process of the context effect acting on English words, it is not yet clear whether this would generalize to a different writing system, such as written Chinese.

There are great differences between alphabetic (e.g., English) and logographic (e.g., Chinese) writing systems in terms of written units, structure, space utilization, and sentence organization (Hoosain, 1992). Chinese is a character-based language in which each printed character occupies an equal square space, and one or more characters in combination can form words. Most Chinese characters carry their own independent meaning; this type of character also constitutes a morpheme.

Chinese words can be divided into single-morpheme and compound words. A single-morpheme word is a character with meaning (e.g., 水 [water], or 人 [people]) or more than one character combined to communicate one meaning (e.g., 蚱蜢 [grasshopper]). A compound word is composed of two or more morphemes (e.g., 學生 [learn/person = student], 老師 [old/master = teacher]).

Due to its nature, Chinese is a writing system with highly productive ways of creating new words. Many new words are introduced each century, and a large number of modern Chinese words are compound words (Packard, 2000; Ramsey, 1987). The majority of Chinese words (approximately 70%) are composed of two characters, and a much smaller proportion (approximately 20%) comprise a single character or more than two characters (approximately 10%; Academia Sinica Taiwan, 1997).

Adult readers can easily deduce the meaning of a common compound word from its component characters (or morphemes). However, in domains such as science, many words carry domain-specific meanings. Sometimes academic words have specific denoted meanings that are difficult for readers to infer from the individual characters' meanings. For example, the word 電感 (inductance) is composed of two characters: 電 (electricity) and 感 (to feel).

A Chinese text is composed of serial words (one-character words, two-character words, three-character words, and so on) with no boundaries between them. Therefore, comprehending Chinese text is a complicated process that involves segmenting words to delineate sentences; this information is already provided in the text in alphabetic reading. Moreover, if the topic is difficult and unfamiliar to a reader, it is more challenging to segment words in the very beginning. Compared with alphabetic writing systems, in which spaces in the text indicate where eye fixations occur as word recognition proceeds (Rayner, 1998; Rayner & Pollatsek, 1981), Chinese word segmentation must first be completed before lexical identification can occur (Shen et al., 2012).

Consider a popular science text that usually consists of familiar common words (e.g., 圍繞 [around], 我們 [our], 的 [a bound subordinate], 宇宙 [universe], 吸引力 [attraction]), and unfamiliar academic terminology (e.g., 暗物質 [dark matter], 超伴子 [superpartner], 光子 [photon], 中性伴子 [neutralino]). See the Appendix for samples of the reading material used in this study, in Chinese and English.

Although written Chinese provides no word spacing information to assist eye movement and lexical identification, adult readers use their lexical knowledge to locate words and complete word segmentation before word processing. This theory is supported by research suggesting that Chinese readers preferred viewing

locations on Chinese words (Chen & Ko, 2011; Li, Rayner, & Cave, 2009; Yan, Kliegl, Richter, Nuthmann, & Shu, 2010), even with unfamiliar Chinese academic words (Jian & Ko, 2012).

In a Chinese reading study conducted by Jian and Ko (2012), increased processing time and longer rereading times were observed for common words placed with academic physics terminology in physics texts. In theory, readers should not reread familiar common words so frequently; this rereading behavior implies that readers experienced difficulty with reading comprehension (Inhoff & Wu, 2005; Rayner, 1998; Rayner & Juhasz, 2004). This unusual phenomenon might be caused by the readers searching for contextual information to clarify and interpret unfamiliar physics terminology, resulting in increased rereading time for these familiar common words.

The purpose of this study was thus to investigate the processes of the context effect in the identification of Chinese academic words during text reading. We approached the problem differently from previous studies that used naming tasks and manipulated different contextual information relevant or irrelevant (or neutral) to the target words (Kim & Goetz, 1994; Schwantes, 1982; Stanovich et al., 1981; West et al., 1983). We adopted Chaffin et al.'s (2001) paradigm (keeping the same contextual information but different target words) for the following reasons. If participants showed different eye movement patterns between the different contextual information presented in the various experimental conditions, then it would not be clear whether to attribute this difference to the effect of experimental manipulation or the different contextual information itself. It is reasonable to expect that reading a different context would result in different eye movement patterns. Therefore, in this study, we analyzed the same interest areas of contextual information across different experimental conditions to eliminate the abovementioned possibility. Furthermore, this study used texts as reading material rather than the two-sentence paradigm used by Chaffin et al. Text reading gives us the opportunity to determine which section in the text was searched by readers to gather the contextual information necessary to comprehend the target words.

In particular, we wanted to examine how readers processed contextual information when they encountered different target words. The target words were physics terminology and replaced words embedded in physics texts and corresponding texts, respectively. Based on findings from previous studies, both adults (Stanovich & West, 1981) and children (Perfetti et al., 1979; Stanovich et al., 1981) showed a greater context effect with difficult words than with simple words. We expected that undergraduate participants in this study would show a greater context effect with physics

terminology (difficult words) than with replaced words (easy words). This study also describes the reading processes by which readers used contextual information in texts to learn academic words.

## Methods

### Participants

Fifty native Chinese speakers from the National Central University in Taiwan volunteered to participate in the experiment for a monetary reward. All participants were enrolled in the College of Liberal Arts or College of Management and reported having no regular habit of reading science material; therefore, they were expected to have less background knowledge of physics. All participants had normal or corrected-to-normal vision.

### Materials and Design

We provided same-context texts for readers by manipulating target words that were either physics terminology or replaced words. Because academic text is not easily written, we first selected physics texts from *Scientific American* magazine (Chinese edition), and each text ranged from 190 to 220 characters in length. A physics professor was invited to identify the academic physics terminology within these texts. We then replaced the physics terminology with familiar words matching the format of the physics texts (subsequently referred to as corresponding texts). The topics of the corresponding texts included astronomy, nutrition, and information technology.

Six physics texts containing 40 instances of academic physics terminology and six corresponding texts containing 40 instances of replaced words were thus created. Both text types shared the same 456 common words. The six physics texts were labeled A1–A6, and the corresponding texts were labeled B1–B6. Half of the participants read even-numbered physics texts and odd-numbered corresponding texts, while the other half read odd-numbered physics texts and even-numbered corresponding texts. Thus, each participant read three physics and three corresponding texts.

To ensure the readability of both types of texts, we examined them using Chinese latent semantic analysis (Chen, Wang, & Ko, 2009). Latent semantic analysis is an objective method used in previous research to confirm readability in various texts. Two evaluation indicators (vocabulary richness and coherence of sentences) were applied to all texts. The statistical results showed no significant difference between the paired versions (same context text with physics terminology or replaced words):  $p_s > .10$ . We confirmed that the physics texts and the corresponding texts shared similar readability.

### Apparatus

Participants' eye movements were recorded using EyeLink II. The headband was adjusted for each participant, and eye movements were calibrated and validated until the average error in gaze position was less than  $0.5^\circ$  of visual angle. A chin bar was used to minimize head movement. Viewing was binocular, but eye movements were recorded from the right eye only. The sampling rate was 250 Hz (250 per second refresh rate). Texts were presented one at a time on a 19" LCD monitor, and the entire text was visible on the screen; no page scrolling was required. The size of each Chinese character in the text displayed was  $24 \times 24$  pixels; each character subtended  $1^\circ$  of visual angle at a distance of approximately 65 cm during the reading tasks. The eye-tracking apparatus was sensitive enough to detect eye movement from character to character.

### Procedure

Participants were instructed to read the texts for comprehension; the reading time was self-paced. Each participant read the six texts one at a time on the screen in random order. When the participants finished reading one text, they pressed a button to terminate the display. To ensure their attention to the texts, we instructed them to answer a yes–no comprehension question that appeared on the screen after terminating the text display. Participants first performed two practice trials to learn the experimental procedure. After participants indicated that they understood the procedure, 9-point calibration and validation procedures were initiated for each participant. Participants were also instructed to keep their heads still throughout the experimental procedure. Each participant completed the experiment in approximately 20–30 minutes.

## Results

We conducted two sets of analyses: measurements of eye movement at the global and local levels. Global analyses provide a measure of the overall reading difficulty associated with the texts (Li, Liu, & Rayner, 2011; Shen et al., 2012), and local analyses reflect the processes involved for the specific target words (Andrews, Miller, & Rayner, 2004; Chaffin et al., 2001; Jian & Ko, 2012; Shen et al., 2012; Williams & Morris, 2004). Individual fixations shorter than 100 ms were excluded from the analyses as in previous studies (Andrews et al., 2004; Chen & Ko, 2011; Jian & Ko, 2012), representing approximately 2% of the data.

### Global Analyses

We computed four global indicators of eye movement averaged across each text presentation to reflect the difficulty of processing:

1. *Total reading time*: The sum of all fixations during one text reading
2. *Mean fixation duration*: The average duration of all fixations occurring for a text
3. *Mean saccade length*: The distance between two successive fixations is called saccade length. Mean saccade length is the average length of all saccades occurring while reading a text.
4. *Number of regressive saccades*: The sum of all regressive saccades for a text

A *t*-test was carried out on each of the eye movement measurements to compute error variance over participants ( $t_1$ ) and items ( $t_2$ ) for each of these measures.

Table 1 lists participants' global eye movement measurements for the physics and corresponding texts. These results show longer mean total reading times for the physics texts than for the corresponding texts:  $t_1(49) = 12.29$ ,  $p < .001$ , Cohen's  $d = 1.34$ ;  $t_2(10) = 4.22$ ,  $p = .002$ , Cohen's  $d = 2.44$ . Mean fixation duration was longer for the physics texts than for the corresponding texts:  $t_1(49) = 10.34$ ,  $p < .001$ , Cohen's  $d = 0.51$ ;  $t_2(10) = 3.49$ ,  $p = .006$ , Cohen's  $d = 2.02$ . Mean saccade length was shorter for the physics texts compared with the corresponding texts:  $t_1(49) = -6.00$ ,  $p < .001$ , Cohen's  $d = -0.40$ ;  $t_2(10) = -2.18$ ,  $p = .054$ , Cohen's  $d = -1.25$ . Finally, the number of regressive saccades was much greater for the physics texts than for the corresponding texts:  $t_1(49) = 9.66$ ,  $p < .001$ , Cohen's  $d = 0.95$ ;  $t_2(10) = 3.78$ ,  $p = .004$ , Cohen's  $d = 2.18$ .

These data clearly indicate that the physics texts were more difficult for the readers to process. Participants spent 22 seconds more total reading time on each physics text than each corresponding text. This longer reading time indicates that much more time was spent on the scientific words as well as the common words. It is reasonable to speculate that this was due to readers' reliance on common words to further their understanding of the text. To test this hypothesis, we

conducted local analyses of the Chinese word units in the physics and corresponding texts as follows, including academic physics terminology versus the replaced words, and the words common to both texts.

## Local Analyses

For the local analyses, the measurements of initial and late processing time were assessed for all analyzed areas, including the physics terminology, the replaced words, and the words common to both types of text. Initial processing time includes (1) the first fixation duration (the duration of the first fixation on a word), which might often reflect the first constituent processing of a compound word (Rayner, 1998), and (2) gaze durations (the sum of all fixations on a word prior to leaving the word), which represents the initial processing in decoding a whole word (Brusnighan & Folk, 2012; Jian & Ko, 2012). Late processing time includes (1) total fixation durations (the sum of all fixations on a word), and (2) rereading time (the sum of all fixations excluding the first gaze durations on a word). Both eye movement indicators reflect reanalysis of the target words because of lack of comprehension or the integration of information relevant to the target words (Chaffin et al., 2001; Rayner, 1998; Williams & Morris, 2004).

## Comparison of Readers' Eye Movements for Physics Terminology and Replaced Words Between the Physics and Corresponding Texts

### Initial Processing Time

Table 2 shows that participants' first fixation durations did not differ significantly between the physics terminology and the replaced words,  $ps > .05$ , but gaze duration was longer for the physics terminology than for the replaced words:  $t_1(49) = 4.34$ ,  $p < .001$ , Cohen's  $d = 0.62$ ;  $t_2(78) = 2.55$ ,  $p = .013$ , Cohen's  $d = 0.57$ .

**TABLE 1**  
Global Eye Movement Measurements for Readers When Reading the Physics and Corresponding Texts

Eye movement indicators	Physics texts: <i>M</i> ( <i>SD</i> )	Corresponding texts: <i>M</i> ( <i>SD</i> )
Mean total reading time (s)	77.16 (22.69)	50.10 (17.42)
Mean fixation duration (ms)	228.50 (21.01)	218.10 (19.70)
Mean saccade length (character space)	4.54 (0.88)	4.93 (1.04)
Number of regressive saccades	88.17 (36.78)	57.65 (26.70)

**TABLE 2**  
Local Eye Movements (ms) While Reading Physics Terminology Versus Replaced Words in the Physics and Corresponding Texts

Eye movement indicators	Physics terminology: <i>M</i> ( <i>SD</i> )	Replaced words: <i>M</i> ( <i>SD</i> )
First fixation durations	234.88 (32.57)	224.56 (32.24)
Gaze durations	307.77 (68.06)	268.96 (56.75)
Rereading time	626.13 (285.40)	347.95 (218.13)
Total fixation durations	933.91 (312.88)	616.90 (236.65)

## Late Processing Time

Table 2 shows that participants had longer rereading time for the physics terminology than for the replaced words:  $t_1(49) = 7.80, p < .001$ , Cohen's  $d = 1.10$ ;  $t_2(78) = 4.42, p < .001$ , Cohen's  $d = 0.99$ . Participants also had longer total fixation durations:  $t_1(49) = 8.05, p < .001$ , Cohen's  $d = 1.14$ ;  $t_2(78) = 4.34, p < .001$ , Cohen's  $d = 0.97$ .

We performed a supplemental analysis to verify that first fixation durations for the two word types did not differ significantly because adult readers were able to decode the individual characters that comprised the unfamiliar physics words. Each Chinese character in the physics words and replaced words was considered an analysis unit. The results showed no significant differences in first fixation durations or gaze durations for the characters of the two word types:  $ps > .05$ . In addition, based on the *Academia Sinica Balanced Corpus of Modern Chinese* (Academia Sinica Taiwan, 1997), we calculated the frequency of characters appearing in both word types, which did not differ significantly:  $p > .05$ . The above findings suggest that the participants were as capable of recognizing individual characters within the unfamiliar physics words as they were of recognizing common characters in the replaced words.

## Comparison of Readers' Eye Movements for the Same Common Words in the Physics and Corresponding Texts

### Initial Processing Time

Table 3 shows longer first fixation durations for the same common words in the physics texts than for the corresponding texts:  $t_1(49) = 4.70, p < .001$ , Cohen's  $d = 0.36$ ;  $t_2(910) = 4.21, p < .001$ , Cohen's  $d = 0.28$ . Participants also had longer longer gaze durations:  $t_1(49) = 4.15, p < .001$ , Cohen's  $d = 0.39$ ;  $t_2(910) = 4.61, p < .001$ , Cohen's  $d = 0.31$ . However, effect sizes were very small, ranging from .28 to .39.

**TABLE 3**  
Local Eye Movements (ms) While Reading the Same Common Words in the Physics and Corresponding Texts

Eye movement indicators	Physics texts: M (SD)	Corresponding texts: M (SD)
First fixation durations	226.64 (23.63)	218.11 (23.63)
Gaze durations	250.00 (37.94)	236.38 (31.03)
Rereading time	282.59 (117.39)	169.09 (83.06)
Total fixation durations	532.58 (135.89)	405.47 (94.55)

## Late Processing Time

Table 3 shows longer rereading times for the same common words in the physics texts than in the corresponding texts:  $t_1(49) = 9.08, p < .001$ , Cohen's  $d = 1.12$ ;  $t_2(910) = 10.48, p < .001$ , Cohen's  $d = 0.69$ . Participants also had longer total fixation durations:  $t_1(49) = 9.13, p < .001$ , Cohen's  $d = 1.09$ ;  $t_2(910) = 10.05, p < .001$ , Cohen's  $d = 0.67$ .

## Eye Movements for the Next Fixation on Target Physics Words and Replaced Words

We further sought to identify where readers would search for contextual information when they encountered a physics word. We calculated the percentage of first and second fixations that followed each initial physics word fixation to examine the process of the context effect during text reading. To establish a baseline comparison, we also conducted the same analysis of fixations across locations of the corresponding texts. Five sentence locations were identified: (1) the present sentence, containing the target words (physics or replaced words); (2) the preceding sentence, before that which contained the target words; (3) the following sentence, following that in which the target words were located; (4) other sentences, those far from the target words; and (5) missing data, a location outside the texts.

### Initial Processing Stage

We calculated the percentage of fixations on sentence locations that followed the initial fixation on each target word when it was first encountered to examine the context effect during text reading (see Table 4). The results showed a main effect of sentence location:  $F(4, 176) = 697.11, p < .001, \eta^2 = .93$ . The participants tended to locate their fixations on the present sentence rather than the other sentences after leaving the target words, whether physics words or replaced words, but there was neither a main effect of word types nor an interaction effect of sentence locations and word types:  $ps > .10$ .

### Late Processing Stage

We calculated the percentage of fixations that counted as the second fixation on the reencountered physics terminology or the replaced words to examine the process of using contextual information to comprehend target words during text reading (see Table 4). These results also showed a main effect of sentence locations:  $F(4, 176) = 900.86, p < .001, \eta^2 = 0.95$ . The participants tended to locate their fixations on the present sentence rather than on other sentences after leaving the target words, whether physics or replaced words. There was

**TABLE 4**  
**The Percentage (%) of Fixations That Were the**  
**Fixation Following Each Physics Term and Replaced**  
**Words Encountered by Participants in the Initial and**  
**Late Processing Stages**

Sentence locations	Physics texts: M (SD)	Corresponding texts: M (SD)
<i>Initial processing stage</i>		
Missing data	1 (2)	1 (2)
Current sentences	74 (17)	74 (16)
Preceding sentences	4 (6)	5 (6)
Following sentences	14 (15)	11 (7)
Other sentences	7 (7)	9 (9)
<i>Late processing stage</i>		
Missing data	1 (2)	2 (4)
Current sentences	70 (11)	73 (13)
Preceding sentences	5 (4)	4 (5)
Following sentences	14 (6)	13 (10)
Other sentences	10 (8)	8 (8)

neither a main effect of word type nor an interaction effect of sentence location and word type:  $ps > .10$ .

Although the reading pathway for next fixations after leaving the target words was similar between the physics texts and the corresponding texts, it was obvious that the processing demands were not equal for both word types because the participants spent more rereading time on the physics terminology than on the replaced words, as mentioned previously. Therefore, we further investigated how many times the participants needed to reread the different target words to process them. The results showed that 79% of the physics terminology needed to be reread, significantly outnumbering the replaced words, of which approximately 63% were reread:  $\chi^2(1, N = 50) = 51.35, p < .001$ . Moreover, each physics word was reread on average 2.18 times, whereas the replaced words were reread an average of approximately 1.29 times. They were processed significantly differently:  $t(49) = 7.18, p < .001$ , Cohen's  $d = 1.08$ .

## Discussion

We investigated the context effect by examining the online processing of Chinese academic words during text reading. Previous context effect-related studies focused on the outcomes of the context effect (Perfetti & Lesgold, 1977; Rahman & Bisanz, 1986) or measured

reaction times to reflect the effect (Carlisle et al., 2000; Nagy et al., 1987; Stanovich, 1980, 1984; Stanovich et al., 1981) rather than the process. Several findings in this study will be discussed with regard to reading theory and empirical research in the literature.

First, reading time was longer and saccade length shorter when participants read the academic texts. This increased processing time was observed not only among the unfamiliar physics terminology but also among the common words in the physics texts.

By analyzing eye movement data from a temporal perspective, we can distinguish the reading process from the initial stage to the late stage. Under the same contextual conditions, readers' first fixation durations on target words, whether physics terminology or replaced words, were not significantly different. Gaze durations, however, were longer for physics terminology than for replaced words, implying that readers needed more time to decode the whole physics term during the initial reading stage.

We were surprised to find that participants' first fixations and gaze durations were longer on the same common words in the physics texts than in the corresponding texts. As previously mentioned, the common words used in this study were conventional and easy for adult readers to recognize; therefore, this would not be a result of any difficulty in recognizing these common words but rather that the physics context led readers to read these words carefully during the initial processing stage to help them interpret the physics terminology. It was evident that readers were aware of the difficulties they encountered; their actions in seeking help from the context began in the early reading stage and continued to the late stage.

In the late processing stage, rereading time for the physics terminology was longer than for the replaced words, as expected. Further, we found that rereading time for the same common words in the physics texts was also longer than in the corresponding texts. These findings were similar to the results of Jian and Ko's (2012) study, in which graduate students who were not physics majors demonstrated longer rereading times and total fixation durations on common words in physics texts than did graduate students who majored in physics. This indicated that readers with little background knowledge relied on the common words surrounding the unfamiliar physics terminology to help them infer their meaning; this process occurred at the late integrating stage. Readers tried to establish a definition of the academic physics terminology by checking common words. That is, adult readers used contextual information to assist their inference of the meaning of unfamiliar physics terminology during academic text reading. This phenomenon was demonstrated by other researchers who argued that readers

generate elaborative inferences online from known words (Chaffin et al., 2001; Garrod, O'Brien, Morris, & Rayner, 1990).

In general, adult readers decoded the meaning of the Chinese academic physics terminology at the beginning stage while reading physics texts. Then, they reread these words and the surrounding common words during the late processing stage. The process is slightly different from alphabetic word processing. The initial processing time of a word is influenced primarily by the time required to access the word's meaning. As readers complete lexical access, their eyes move on to the next location (Morrison, 1984; Reichle, Pollatsek, Fisher, & Rayner, 1998). Previous researchers using eye tracking to investigate alphabetic word processing found that readers demonstrated longer first fixation durations, gaze durations, and rereading time on unfamiliar (or novel) words than on familiar words in the sentences (Chaffin et al., 2001; Juhasz & Rayner, 2003; Williams & Morris, 2004).

In this study, we used Chinese texts as reading materials and found that readers did not show different first fixation durations for physics terminology and replaced words. We speculate that the properties of the Chinese writing system might account for the difference found between these two types of words. As mentioned earlier, a Chinese word is composed of one or more characters; it is very common to see an unfamiliar academic word composed of familiar characters. First fixation duration is an indicator of the first constituent processing of a compound word (Rayner, 1998). Participants showed no different first fixation durations between the physics and replaced words in this study. This might be because it was not difficult for readers to identify each component character of either the physics or replaced words. One finding to support this claim is that the frequency of the characters in the physics terminology and replaced words did not differ significantly:  $p > .10$ .

Another piece of evidence was that first fixation durations were approximately 220–250 ms regardless of whether participants were reading physics terminology or common words in the physics texts, consistent with the results of previous Chinese reading research (Chen & Ko, 2011; Jian & Ko, 2012; Li et al., 2011; Yan et al., 2010). In addition, when analyzing Chinese characters as a unit, the first fixation durations showed no difference between physics and replaced words, further supporting the previous statement. These data concur with Jian and Ko's finding that readers with either high or low physics knowledge did not differ on first fixation duration while reading unfamiliar physics terminology; all results fell within the range of 250 ms.

Nevertheless, decoding the academic physics terminology itself was not sufficient for comprehension.

Readers took contextual information and tried to comprehend the physics terminology by rereading both the physics and the common words. These findings were further supplemented by the behavior indicating that participants were seeking the contextual information. First, once readers' fixation left the physics terminology, most shifted their eyes within the present sentence (containing the physics terminology) and to the next sentence (the sentence following the present sentence) to search for more information. They moved their eyes comparatively less to the previous sentence from the current sentence. Although the participants performed similar reading paths for the corresponding texts, their processing duration for the present sentences was different. The readers needed to reread the physics terminology more times than the replaced words. The percentage of rereading for each physics word was also higher than for the replaced words.

In a way, these data corresponded with the conclusion of Chaffin et al. (2001). Their research showed that readers spent less total reading time on uninformative contexts and more total reading time on informative contexts because readers were able to distinguish which contextual information was informative or uninformative to the target novel words. It is reasonable for readers to spend more time (including rereading time) on the present sentence due to its immediate connection to the target word. Then, the readers turned to the next sentence for more information. Finally, they sought help from the previous sentence or other places. The reason for not returning to the previous sentence first after reading the present sentence might be because the readers had learned the contents of the previous sentence and found that they could not learn more. They thus moved their eyes to the following and other sentences. This is supported by the increased percentage of fixations on sentences other than the previous sentence (see Table 4). If this is the case, the metacognitive ability to monitor one's comprehension in reading was indirectly observed in this study.

In sum, this study described the processes by which adult readers read academic words during text reading. First, while reading physics texts, participants seemed to succeed at decoding the characters' meanings but failed to access the meaning of the physics terminology. Readers then continued to reread sentences and seek help from the common words surrounding the physics terminology to comprehend the unfamiliar terms. Although orthographic processing may differ because of the logographic nature of written Chinese, the syntactic and semantic processing of sentences and texts appears to be similar.

In theory, this study used academic words to demonstrate the nature of the formation of Chinese words. Reading Chinese characters fluently does not imply

comprehension of a Chinese word. Moreover, this study responds to the claim of the interactive-compensatory model (Stanovich, 1980) that readers utilize top-down information to compensate for deficiencies at any level of the reading process, especially for unfamiliar or difficult words. That is, readers relied on contextual information to compensate for a lack of knowledge required to comprehend the unfamiliar academic words.

One limitation of the eye movement analysis in this study was that the positioning of the physics target words was not controlled. In the experimental texts, some of the physics terminology was within a single sentence, and others were dispersed across several sentences. In the event that two physics terms were closely located in a sentence, fixation on the second physics word could have been affected by the location of the first physics word, causing a spillover effect. This would particularly be the case for low-frequency words because they prolong the gaze on the present word and the next word (Rayner & Duffy, 1986). In a way, there is a limitation to explain physics terminology processing of eye movements, especially for those physics terms located near each other.

We did not control the target words' positions in this study for two reasons. First, controlling the position of words in text of this nature is difficult: Scientific text usually has a standardized form, and changing word positions may distort the syntactic structure, reducing the readability and fluency of the text. Second, the text structure was the same between the physics and the nonphysics texts, so if any position effect occurred, it would affect both the physics words and the replaced words. Therefore, there was no systematic bias in analyzing the eye movement data in this study.

## NOTE

This study was supported by grants from the National Science Council in Taiwan (NSC 100-2511-S-008-009, NSC 100-2410-H-134-027).

## REFERENCES

- Academia Sinica Taiwan. (1997). *Academia Sinica balanced corpus of modern Chinese* (Version 3.0). Taipei, Taiwan: Language and Knowledge Processing Group, Institute of Information Science, Academia Sinica.
- 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. doi:10.1080/09541440340000123
- Brusnighan, S.M., & Folk, J.R. (2012). Combining contextual and morphemic cues is beneficial during incidental vocabulary acquisition: Semantic transparency in novel compound word processing. *Reading Research Quarterly*, *47*(2), 172–190. doi:10.1002/RRQ.015
- Carlisle, J.F., Fleming, J.E., & Gudbrandsen, B. (2000). Incidental word learning in science classes. *Contemporary Educational Psychology*, *25*(2), 184–211. doi:10.1006/ceps.1998.1001
- Chaffin, R., Morris, R.K., & Seely, R.E. (2001). Learning new word meanings from context: A study of eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(1), 225–235. doi:10.1037/0278-7393.27.1.225
- Chen, M., & Ko, H. (2011). Exploring the eye-movement patterns as Chinese children read texts: A developmental perspective. *Journal of Research in Reading*, *34*(2), 232–246. doi:10.1111/j.1467-9817.2010.01441.x
- Chen, M., Wang, H., & Ko, H. (2009). Using latent semantic analysis to create a Chinese semantic space and the validation of psychological reality [in Chinese]. *Chinese Journal of Psychology*, *51*(4), 415–435.
- Garrod, S., O'Brien, E.J., Morris, R.K., & Rayner, K. (1990). Elaborative inferencing as an active or passive process. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(2), 250–257. doi:10.1037/0278-7393.16.2.250
- Hoosain, R. (1992). Psychological reality of the word in Chinese. In H.-C. Chen & O. J.L. Tzeng (Eds.), *Language processing in Chinese* (pp. 111–130). Amsterdam: North-Holland.
- Inhoff, A.W., & Wu, C. (2005). Eye movements and the identification of spatially ambiguous words during Chinese sentence reading. *Memory & Cognition*, *33*(8), 1345–1356. doi:10.3758/BF03193367
- Jian, Y.-C., & Ko, H.-W. (2012). Investigating the effects of background knowledge on Chinese word processing during text reading: Evidence from eye movements. *Journal of Research in Reading*. Advance online publication. doi:10.1111/j.1467-9817.2012.01534.x
- Juhász, B.J., & Rayner, K. (2003). Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(6), 1312–1318. doi:10.1037/0278-7393.29.6.1312
- Kim, Y.H., & Goetz, E.T. (1994). Context effects on word recognition and reading comprehension of poor and good readers: A test of the interactive-compensatory hypothesis. *Reading Research Quarterly*, *29*(2), 178–188. doi:10.2307/747810
- Kuhn, M.R., & Stahl, S.A. (1998). Teaching children to learn word meanings from context: A synthesis and some questions. *Journal of Literacy Research*, *30*(1), 119–138. doi:10.1080/10862969809547983
- Landauer, T.K., & Dumais, S.T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*(2), 211–240. doi:10.1037/0033-295X.104.2.211
- Li, X., Liu, P., & Rayner, K. (2011). Eye movement guidance in Chinese reading: Is there a preferred viewing location? *Vision Research*, *51*(10), 1146–1156. doi:10.1016/j.visres.2011.03.004
- Li, X., Rayner, K., & Cave, K.R. (2009). On the segmentation of Chinese words during reading. *Cognitive Psychology*, *58*(4), 525–552. doi:10.1016/j.cogpsych.2009.02.003
- Morrison, R.E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(5), 667–682. doi:10.1037/0096-1523.10.5.667
- Nagy, W.E., Anderson, R.C., & Herman, P.A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, *24*(2), 237–270. doi:10.3102/00028312024002237
- Packard, J.L. (2000). *The morphology of Chinese: A linguistic and cognitive approach*. New York: Cambridge University Press. doi:10.1017/CBO9780511486821
- Perfetti, C.A., Goldman, S.R., & Hogaboam, T.W. (1979). Reading skill and the identification of words in discourse context. *Memory & Cognition*, *7*(4), 273–282. doi:10.3758/BF03197600
- Perfetti, C.A., & Lesgold, A.M. (1977). Discourse comprehension and sources of individual differences. In M.A. Just & P.A. Carpenter (Eds.), *Cognitive processes in comprehension* (pp. 141–183). Hillsdale, NJ: Erlbaum.

- Rahman, T., & Bisanz, G.L. (1986). Reading ability and use of a story schema in recalling and reconstructing information. *Journal of Educational Psychology, 78*(5), 323–333. doi:10.1037/0022-0663.78.5.323
- Ramsey, S.R. (1987). *The languages of China*. Princeton, NJ: Princeton University Press.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422. doi:10.1037/0033-2909.124.3.372
- Rayner, K., & Duffy, S.A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14*(3), 191–201. doi:10.3758/BF03197692
- Rayner, K., & Juhasz, B. (2004). Eye movements in reading: Old questions and new directions. *European Journal of Cognitive Psychology, 16*(1/2), 340–352. doi:10.1080/09541440340000385
- Rayner, K., & Pollatsek, A. (1981). Eye movement control during reading: Evidence for direct control. *The Quarterly Journal of Experimental Psychology, 33*(4), 351–373.
- Reichle, E.D., Pollatsek, A., Fisher, D.L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*(1), 125–157. doi:10.1037/0033-295X.105.1.125
- Rumelhart, D.E. (1977). Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance* (Vol. 6, pp. 573–603). Hillsdale, NJ: Erlbaum.
- Schwantes, F. (1982). Text readability level and developmental differences in context effects. *Journal of Literacy Research, 14*(1), 5–12.
- Shen, D., Liversedge, S.P., Tian, J., Zang, C., Cui, L., Bai, X., et al. (2012). Eye movements of second language learners when reading spaced and unspaced Chinese text. *Journal of Experimental Psychology: Applied, 18*(2), 192–202. doi:10.1037/a0027485
- Stanovich, K.E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16*(1), 32–71. doi:10.2307/747348
- Stanovich, K.E. (1984). The interactive-compensatory model of reading: A confluence of developmental, experimental, and educational psychology. *Remedial and Special Education, 5*(3), 11–19. doi:10.1177/074193258400500306
- Stanovich, K.E., & West, R.F. (1981). The effect of sentence context on ongoing word recognition: Tests of a two-process theory. *Journal of Experimental Psychology: Human Perception and Performance, 7*(3), 658–672. doi:10.1037/0096-1523.7.3.658
- Stanovich, K.E., West, R.F., & Feeman, D.J. (1981). A longitudinal study of sentence context effects in second-grade children: Tests of an interactive-compensatory model. *Journal of Experimental Child Psychology, 32*(2), 185–199. doi:10.1016/0022-0965(81)90076-X
- West, R.F., Stanovich, K.E., Feeman, D.J., & Cunningham, A.E. (1983). The effect of sentence context on word recognition in second- and sixth-grade children. *Reading Research Quarterly, 19*(1), 6–15. doi:10.2307/747333
- Williams, R., & Morris, R. (2004). Eye movements, word familiarity, and vocabulary acquisition. *European Journal of Cognitive Psychology, 16*(1/2), 312–339. doi:10.1080/09541440340000196
- Yan, M., Kliegl, R., Richter, E.M., Nuthmann, A., & Shu, H. (2010). Flexible saccade-target selection in Chinese reading. *The Quarterly Journal of Experimental Psychology, 63*(4), 705–725. doi:10.1080/17470210903114858

Submitted April 1, 2013

Final revision received June 13, 2013

Accepted June 15, 2013

**YU-CIN JIAN** is a postdoctoral fellow at the Graduate Institute of Learning and Instruction at the National Central University, Zhongli, Taiwan; e-mail jianyucin@gmail.com.

**MING-LEI CHEN** is an assistant professor at the Graduate Institute of Taiwan Languages and Language Education at the National Hsinchu University of Education, Taiwan; e-mail mingleichen@gmail.com.

**HWA-WEI KO** (corresponding author) is a professor at the Graduate Institute of Learning and Instruction at the National Central University, Zhongli, Taiwan; e-mail hwawei@cc.ncu.edu.tw.

## APPENDIX

### Sample Physics Text

圍繞我們的宇宙，與表面上所看到的並不相同。天文學家與物理學家持續在收集暗物質存在的證據。目前，超對稱在解釋暗物質方面頗具吸引力，因為它假設了一系列全新的粒子：每個已知的基本粒子都有一個超伴子。這些新粒子的質量都大於已知的粒子，其中有幾個是冷暗物質的恰當候選者。最引人注意的粒子是中性子，它是下列粒子的超伴子的綜合體：光子、Z玻色子，也許還有其他類型的粒子。中性子這個名詞的發音很像微中子，它們雖有一些共通特質，但在其他方面是相當不同的。

The universe around us is not what it appears to be. Astronomers and physicists have steadily gathered circumstantial evidence for the existence of dark matter. By far, supersymmetry is an attractive explanation for dark matter because it postulates a whole new family of particles—one superpartner for every known elementary particle. These new particles are all heavier than known particles. Several

are natural candidates for cold dark matter. The one that gets the most attention is the neutralino, which is an amalgam of the superpartners of the photon, the Z-boson, and perhaps other particle types. Neutralino sounds much like *neutrino*, and the two particles indeed share various properties, but they are otherwise quite distinct.

### Sample Corresponding Text

圍繞我們的宇宙，與表面上所看到的並不相同。天文學家與氣象學家持續在收集大犬座存在的證據。目前，星座學在解釋大犬座方面頗具吸引力，因為它解釋了一系列神話的故事：希臘式神話中的每一隻狗都是一個星座。這些星座群的熱度都大於一般的恆星，其中有幾個是最亮恆星的恰當候選者。最引人注意的恆星是天狼星，它是與下面這顆伴星所形成的一個雙星系統：白矮星，也許還有其他類型的星星。天狼星這個恆星的亮度很像老人星，它們雖有一些共通特質，但在其他方面是相當不同的。

The universe around us is not what it appears to be. Astronomers and meteorologists have steadily gathered circumstantial evidence for the existence of Canis Major. By far, *constellation* is an attractive explanation for Canis Major because it explains a series stories of mythology—one dog for one constellation in Greek mythology. These constellations are a higher temperature than general fixed stars.

Several are natural candidates for bright, fixed stars. The one that gets the most attention is Sirius, which is a binary system of the stars of the companion star, the white dwarf, and perhaps other star types. The brightness of Sirius is like the Canopus, and the two stars indeed share various properties, but they are otherwise quite distinct.