Acta Psychologica xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/ locate/actpsy

Parafoveal semantic information extraction in traditional Chinese reading☆

Jie-Li Tsai [a,*], Reinhold Kliegl [b], Ming Yan [b,**]

[a] Department of Psychology and Research Center for Mind, Brain, and Learning, National Chengchi University, Taiwan
[b] Department of Psychology, University of Potsdam, Germany

ARTICLE INFO

Article history:
Received 3 February 2012
Received in revised form 5 June 2012
Accepted 6 June 2012
Available online xxxx

Keywords:
Semantic preview benefit
Chinese reading
Eye movements

ABSTRACT

Semantic information extraction from the parafovea has been reported only in simplified Chinese for a special subset of characters and its generalizability has been questioned. This study uses traditional Chinese, which differs from simplified Chinese in visual complexity and in mapping semantic forms, to demonstrate access to parafoveal semantic information during reading of this script. Preview duration modulates various types (identical, phonological, and unrelated) of parafoveal information extraction. Parafoveal semantic extraction is more elusive in English; therefore, we conclude that such effects in Chinese are presumably caused by substantial cross-language differences from alphabetic scripts. The property of Chinese characters carrying rich lexical information in a small region provides the possibility of semantic extraction in the parafovea.

1. Introduction

Reading is a highly complex and automatized task involving dynamic adjustments of eye movements to visual and language-related properties of the reading material in foveal and parafoveal vision. A current issue is whether semantic information can be extracted from parafoveal words in alphabetic scripts (see Altarriba, Kambe, Pollatsek, & Rayner, 2001, and Rayner, Balota, & Pollatsek, 1986, for failure to find such evidence in English; see Hohenstein, Laubrock, & Kliegl, 2010, and Hohenstein & Kliegl, submitted for publication, for positive evidence in German). Parafoveal semantic extraction has been reported for simplified Chinese (Yan, Richter, Shu, & Kliegl, 2009; Yan, Zhou, Shu, & Kliegl, 2012; Yang, Wang, Tong, & Rayner, 2012). Simplified Chinese is used in mainland China and was designed to simplify character complexity and reduce the number of traditional Chinese characters. However, traditional Chinese is still widely used globally, primarily in Taiwan and Hong Kong. Because of the difference between the two writing systems, parafoveal semantic extraction should not be assumed to be significant in traditional Chinese. This study addresses this question.

Chinese is an important script, not only because it is used by the largest number of people and is one of the oldest languages in human history, but also because of fundamental differences from alphabetic scripts. Differing from Western languages such as English and German, Chinese is a logographic writing system. Its basic writing units, characters, are written in a series of square-shaped forms of the same horizontal extent in a passage of text, but of different levels of visual complexity (roughly the number of strokes). Chinese orthography lacks explicit word boundaries (such as spaces); thus, the Chinese written text is a serial string of characters with punctuation, but without space to tell readers where a word begins or ends.

A theoretical understanding of parafoveal information processing during reading Chinese has witnessed recent progress. In the boundary paradigm (Rayner, 1975), with the type of information of the to-be-fixated word $N+1$ available during a fixation on word N under experimental control, different preview words are replaced by the target word once the eyes cross an invisible boundary located between words N and $N+1$. Preview benefit (PB) is calculated as a positive difference between fixation durations on the target word $N+1$ when preview is denied, compared with when different types of preview are available during previous fixations. Parafoveal preview benefits have been consistently demonstrated for orthographic (Inhoff, 1990) and phonological (Pollatsek, Lesch, Morris, & Rayner, 1992) word features in alphabetic script and in simplified and traditional Chinese sentence reading studies (Liu, Inhoff, Ye, & Wu, 2002; Tsai, Lee, Tzeng, Hung, & Yen, 2004; see Chen & Tsang, submitted for publication, for a review).

2. Parafoveal semantic preview benefit in simplified Chinese

Chinese is particularly well suited for demonstrating semantic parafoveal processing. First, compared to alphabetic languages, Chinese

☆ This research was funded by Deutsche Forschungsgemeinschaft, grant KL 955/18 to Reinhold Kliegl and Jie-Li Tsai, by a grant from the Taiwan National Science Council (NSC 99-2420-H-004 -002-, 99-2410-H-004-091-MY2) to Jie-Li Tsai, and by the Open Research Fund of the State Key Laboratory of Cognitive Neuroscience and Learning of P.R. China to Ming Yan. We thank Hua Shu, Keith Rayner and Albrecht Inhoff for their helpful comments and Pei-ju Kao for her effort during data collection. Data and R scripts used in this work are available at the Potsdam Mind Research Repository (http://read.psych.uni-potsdam.de/pmr2).

* Correspondence to: J.-L. Tsai, Department of Psychology, National Chengchi University, No. 64, Sec. 2, ZhiNan Rd., Wenshan, Taipei 11605, Taiwan ROC. Tel.: +886 2 29393091x62978.
** Correspondence to: M. Yan, Department of Psychology, University of Potsdam, Karl-Liebknecht-Str. 24/25, 14476 Potsdam-Golm, Germany.
E-mail addresses: jltsai@nccu.edu.tw (J.-L. Tsai), mingyan@uni-potsdam.de (M. Yan).

0001-6918/$ – see front matter © 2012 Published by Elsevier B.V.
doi:10.1016/j.actpsy.2012.06.004

characters are mapped more closely to meaning than to phonology (see Hoosain, 1991, for a summary). Second, most Chinese words are only one or two characters long. Because a Chinese character typically occupies the space of three letters in alphabetic languages (Tsai & McConkie, 2003), on average, word $N+1$ is closer to the fixation point on word N in Chinese than in alphabetic languages. Third, the absence of inter-word spaces brings the upcoming word into a much less eccentric position compared to alphabetic scripts. These features of Chinese scripts facilitate parafoveal processing.

The most striking finding of parafoveal processing in Chinese is extracting semantic information from the first parafoveal word to the right of the current fixation (Yan et al., 2009). Such an effect has been consistently elusive in English reading experiments (Altarriba et al., 2001; Rayner et al., 1986). Yan et al. (2009) used visually simple and highly frequent non-compound characters as previews, which are a set of characters composed of single components or radicals. In two subsequent studies, the semantic preview effect was also established with compound characters. Yang et al. (2012) found a cross-experiment interaction between semantic preview benefit and contextual plausibility: the semantic preview benefit measured in single fixation duration (SFD) was stronger when the preview words were plausible from prior sentence context than when they were not. Using a neutral context, Yan, Zhou, et al. (2012) reported a semantic preview benefit "purely" from bottom-up processing without top-down contextual support. They also demonstrated parafoveal semantic extraction from the radical/sub-lexical level.

3. Traditional and simplified Chinese comparison

Although simplified Chinese reading consistently reports the semantic preview benefit, parafoveal semantic extraction should not be assumed to also exist in traditional Chinese. Modern traditional Chinese characters are adapted and evolved from characters used more than 2000 years ago, retaining structural visual complexity. Mainland China designed a set of simplified Chinese characters in the 1950s to reduce the structural complexity of many traditional Chinese characters and to reduce the number of characters. The simplification covered approximately 40% of the characters (Table 1). Over 30% of the simplified characters were a reduction in the number of strokes or structural complexity from traditional Chinese characters.

Compared to simplified Chinese, structural complexity can lead to two opposite predictions for extracting parafoveal semantic information in reading traditional Chinese. First, higher visual complexity may imply a smaller perceptual span (the area within which useful visual information can be extracted during a fixation on word N), and thus, diminish effective parafoveal semantic extraction of words $N+1$ in traditional Chinese. Alternatively, higher visual complexity also implies that the viable visual structure can be informative for mapping form to semantic, facilitating parafoveal semantic extraction. The benefit of simplification, that is, simplified Chinese, could be traded by the cost of morphological or semantic ambiguity. For example, 10% of simplified characters involve mapping to multiple (two to at most four) traditional characters, increasing semantic/morphemic ambiguity in simplified Chinese. Certain simplified characters are visually similar to other characters that are easily distinguished in traditional Chinese. For instance, and are similar to each other in simplified Chinese, but the first character is written as in traditional Chinese, which is extremely different from . Additional lexical representations activated by visually similar entries lead to more competitive morphemic information during lexical activation of simplified Chinese characters. Oppositely, the distinctive structure of traditional Chinese can map to morpheme or semantic information with less ambiguity. Therefore, semantic access to word $N+1$ might be direct and efficient to show a reliable semantic preview effect in reading traditional Chinese.

Table 1
Statistics of mapping between traditional Chinese (TC) and simplified Chinese (SC) characters based on 5619 commonly used traditional Chinese characters.

TC to SC mapping	SC counts	SC proportion		TC counts	TC proportion	Example
1 to 1	5,039	94.7%	identical	3353	59.67%	(凰:凰);(媳:媳)
			simplified	1686	30.01%	(鳳:凤);(媽:妈)
2 to 1	264	5.0%	simplified	528	9.40%	(仆,僕:仆);(佛,彿:佛)
3 to 1	12	0.2%	simplified	36	0.64%	(係,系,繫:系)
4 to 1	4	0.1%	simplified	16	0.28%	(台,檯,臺,颱:台)
total	5,319		total	5619		

4. Preview duration modulation and preview cost

In the boundary paradigm, the previews are always either available or denied for the entire duration of the fixation prior to the boundary. Because fixation durations on the preboundary words N are under the readers' control, the traditional measure of preview benefit averages trials with different preview durations and may mask a duration-dependent time course of parafoveal information activation. Including preview duration as a covariate in a linear mixed model [LMM]) allows us to test this potential dependency.

The traditional term "preview benefit" presupposes that fixation durations on target words following random-letter strings or unrelated word previews serve as an adequate baseline for identical or related preview conditions. Thus, there is an implicit assumption that fixation durations in such reference conditions are constant across preview duration. However, recent studies have shown that this may not be the case: processing random-letter strings or unrelated words may interfere with efficient processing of target words. Therefore, Yan, Risse, Zhou, and Kliegl (2012) and Kliegl, Hohenstein, Yan, and McDonald (2012) introduced the term "preview cost" when fixation durations or refixation rates on a target word increase with fixation duration on the pre-target word (before the boundary) in unrelated or non-word conditions. "Savings" in fixation durations on target words for identical or related over unrelated or non-word previews may also be because of combined genuine preview benefit and preview cost.

5. Experiment

Other than identical and unrelated previews, the present experiment included three critical conditions using characters that were semantically, orthographically, or phonologically related to the target characters. We expected that semantic preview benefit would generalize from simplified to traditional Chinese script, despite substantial differences in visual complexity and morphological/semantic ambiguity. We also tested modulation of preview effects by fixation durations on the pre-boundary word.

6. Method

6.1. Participants

Fifty undergraduate and graduate students ($M=22.8$ years, ranging from 18 to 27 years) at National Chengchi University were paid to participate in the experiment. Data from an additional 11 participants were excluded because of either (a) low correct rates of comprehension questions (<80%; 2 participants) or (b) over half of the experimental trials contaminated by blinks or late display change (9 participants) in which the display change from preview to target character occurred later than 10 ms after fixation onset. The participants were all native Chinese speakers with normal or corrected-to-normal vision.

Please cite this article as: Tsai, J.-L., et al., Parafoveal semantic information extraction in traditional Chinese reading, *Acta Psychologica* (2012), doi:10.1016/j.actpsy.2012.06.004

Table 2
Means (M) and standard deviation (SD) of frequency and number of strokes for preview and target characters in each condition.

	Preview type									
	Identical/target		Orthographic		Phonological		Semantic		Unrelated	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Example	乘		乖		成		坐		或	
Pronunciation	cheng2		guei1		cheng2		tzuo4		huo4	
Frequency	1107	1801	871	1665	1414	2146	1268	1803	1136	1690
Strokes	7.2	4.1	6.9	4.3	6.5	3.1	7.5	4.8	7.6	4.7

Note. Each target character was embedded in a two-character word, for instance, 乘涼 (translated as *shade*).

6.2. Design and materials

The 48 sets of target and preview characters were adapted from Yan et al. (2009). All simplified Chinese characters were transformed into traditional Chinese characters, including five types of preview characters for each target character: identical, orthographically related, phonologically related, semantically related, and unrelated. The proportions of characters transformed from simplified Chinese into traditional Chinese for the preview conditions were 27%, 25%, 21%, 31%, and 38%, respectively. In most cases, transformed traditional characters orthographically differed from the matched simplified characters, for example, the traditional character   versus the simplified character . Sixteen of 48 characters in the orthographically related preview condition from Yan et al. (2009) were only valid in simplified Chinese, but not in traditional Chinese. These characters were replaced by new characters that were orthographically similar to the targets; other targets and preview characters were valid after transformation. Character frequency was obtained from Academic Sinica Balanced Corpus (Academia Sinica Balanced Corpus, 2004) in Taiwan (Table 2). No difference existed among preview types for number of strokes and character frequency [All $Fs<1$].

Each target character was embedded as the first character of a two-character word. Two-character target words were identical to Yan et al. (2009) and their mean frequency was 27 per million words according to Academia Sinica Balanced Corpus (2004). All the non-identical characters yielded pseudo-word previews in the parafovea. The sentences were also identical to Yan et al. (2009), except that seven words were changed to words with similar meanings because of word usage difference between mainland China and Taiwan. Sentence length ranged from 20 to 25 characters ($M=22.5$) and the target character was located in the range of the 8th to 15th character in a sentence. Participants read each sentence only once, and the five preview types of the 48 targets were approximately counter balanced across participants (Table 3).

6.3. Apparatus

An Eyelink (SR Research, Osgoode, ON, Canada) 1000 eye-tracking system with a sampling rate of 1000 Hz was used to record gaze positions during the experiment. Sentences were displayed on a single line at the middle location at the height of a 19-in. ViewSonic PT795 monitor (1024 × 768 pixels in resolution and 120 Hz in vertical refresh rate). The experimental program was implemented in Matlab (MathWorks, Natick, MA) using the Psychtoolbox 2.54 (Brainard, 1997; Pelli, 1997) and the Eyelink toolbox (Cornelissen, Peters, & Palmer, 2002) to display the stimuli and to communicate with eyetracker core libraries. Detecting the eye position across the boundary took 3 or 4 ms and an additional 8 ms (maximum) to complete a display change.

The sentences were displayed in black on a light gray background. The size of each character presented on the screen was 32 × 32 pixels, with a four-pixel space between characters. The viewing distance was 70 cm, and the character width and the space before it subtended 0.9° of the visual angle.

6.4. Procedure

Participants were instructed to read sentences normally for comprehension and to answer a comprehension question after some of the sentences. After establishing the eye-tracker followed by a 5-point calibration procedure, each trial started with a cross at the location of the first character in the sentence as the initial fixation point. Once the participants fixated on the cross, the experimenter pressed a button to present the sentence on the screen. Prior to the participants' eyes crossing the invisible boundary, the preview character was presented in the target character location. Fig. 1 shows that when the eyes moved over the invisible boundary, the preview character was immediately replaced by the target character. The participants read each sentence at their own pace and were instructed to fixate on a square box at the lower right corner of the screen to indicate that they had finished reading and understood the sentence. The participants then pressed a button to make the sentence disappear, and the fixation cross for the next trial appeared. In approximately 40% of the trials, a comprehension question appeared after the sentence disappeared. The participants were required to provide a yes or no answer as to whether the comprehension question paraphrased the sentence they had just read. Participants correctly answered 89% of all questions ($SD=6\%$). Every participant was required to read 144 sentences, including 48 trials and 96 fillers. The overall procedure of the experiment was approximately 40–55 min.

6.5. Data analysis

Fixations were determined with an algorithm for binocular saccade detection (Engbert & Kliegl, 2003). Sentences containing a blink or loss

Table 3
Means (M) and standard errors (SE) of first-fixation duration (FFD), single-fixation duration (SFD), and gaze duration (GD) on target words.

	Preview type									
	Identical		Orthographic		Phonological		Semantic		Unrelated	
	M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
FFD	261	7	294	7	303	7	291	7	301	7
SFD	265	8	299	8	315	8	293	8	315	8
GD	305	10	361	10	374	10	350	10	370	11

Identical Preview

醫生提醒市民們傍晚乘涼時儘量少去草叢茂密的地方。

*

Orthographically Related Preview

醫生提醒市民們傍晚乖涼時儘量少去草叢茂密的地方。

*

Phonologically Related Preview

醫生提醒市民們傍晚成涼時儘量少去草叢茂密的地方。

*

Semantically Related Preview

醫生提醒市民們傍晚坐涼時儘量少去草叢茂密的地方。

*

Unrelated Preview

醫生提醒市民們傍晚或涼時儘量少去草叢茂密的地方。

*

Target

醫生提醒市民們傍晚乘涼時儘量少去草叢茂密的地方。

N N+1

*

Fig. 1. An example of Chinese sentences using the boundary paradigm. The target sentence is translated as follows: *The doctor reminded the citizens that they should avoid lush grass when they enjoy evening shade*. The preview characters (乘, 乖, 成, 坐 or 或) occupy the first character position of the target word $N+1$ and are replaced by the target character (乘) as soon as the reader's eyes cross the invisible boundary located between words N (傍晚) and $N+1$ (乘涼).

of measurement were deleted (8%). Analyses were based on right-eye fixations. The fixation duration measures in the analysis included first-fixation durations (FFDs; the first fixation on a word, irrespective of the number of fixations), single-fixation durations (SFDs; cases in which a word was inspected with exactly one fixation), and gaze durations (GDs; the sum of fixations during the first reading of the word).

Inferential statistics were based on planned comparisons for the three related and identical previews with the unrelated preview as reference. Estimates were based on a linear mixed model (LMM) for durations and a generalized linear mixed model (GLMM) for skipping with crossed random effects for subjects and items using the *lmer* program of the *lme4* package (Bates & Maechler, 2010) in the R environment for statistical computing and graphics (R Development Core Team, 2010). Effects larger than twice their standard errors were interpreted as significant at the 5% level (with the number of subjects and the large number of observations for each subject, the *t*-statistic [M/SE] effectively corresponds to the *z*-statistic). Analyses of residuals and inspection of duration distributions strongly suggested that log-transformation was required to meet LMM assumptions. Therefore, we used log-transformed durations for LMMs. Analyses for untransformed and log-transformed durations yielded the same significance pattern.

7. Results

7.1. Skipping and regression rate

The identical condition had a significantly higher skipping rate ($M=14\%$, $b=.548$, $SE=.225$, $z=2.43$, $p=.015$) and lower regression rate ($M=5\%$, $b=-1.539$, $SE=.283$, $z=5.44$, $p<.001$) than the unrelated condition (skipping rate: 9% and regression rate: 18%). Furthermore, the three related conditions all had (marginally) lower regression rates than the unrelated condition ($M=13\%$, $b=.387$, $SE=.209$, $z=1.85$, $p=.064$; $M=12\%$, $b=.537$, $SE=.216$, $z=2.49$, $p=.013$ and $M=10\%$, $b=.680$, $SE=.223$, $z=3.05$, $p=.002$; for the orthographic, phonological, and semantic previews, respectively).

7.2. Preview benefits

FFDs and SFDs shorter than 60 ms or longer than 600 ms and GDs longer than 800 ms were excluded (7% of all valid trials). Trials with regressions from words *N* were excluded (12%) because they might reflect incomplete parafoveal processing of previews during fixations on pre-target words *N*. Finally, we excluded the late display-change trials in which display change occurred later than 10 ms after fixation onset because readers should be more likely to perceive a display change or a flash at this time (4%). After the three levels of data screening, 1688, 1284, and 1688, observations on the target word contributed to the following FFD, SFD, and GD analyses, respectively.

7.2.1. Identical preview benefits

Relative to unrelated previews, identical preview benefits were highly significant in all three duration measures (PB = 40 ms, $b=-.127$, $SE=.022$, $t=-5.65$; PB = 51 ms, $b=-.161$, $SE=.026$, $t=-6.31$, and PB = 65 ms, $b=-.190$, $SE=.028$, $t=-6.80$, for FFD, SFD, and GD analyses, respectively).

7.2.2. Orthographic preview benefits

Fixation durations on target words also decreased significantly from previewing characters that were orthographically related to the targets: There was a significant preview benefit of 16 ms in SFD analysis ($b=-.056$, $SE=.026$, $t=-2.19$). The orthographic preview benefits in FFD and GD measures were in the expected numerical trends, but not statistically reliable [abs(t)-values < 1.2].

7.2.3. Phonological preview benefits

Fixation durations on target words for phonological previews were not statistically different from those for unrelated previews (PB = −2 ms, $b=0.01$, SE = 0.02, $t=0.49$; PB = 0 ms, $b=-0.002$, SE = 0.03, $t=-0.08$; PB = −4 ms, $b=-0.01$, SE = 0.03, $t=0.31$, for FFD, SFD, and GD analyses, respectively). None of the analyses on phonological preview effects were significant.

7.2.4. Semantic preview benefits

Fixation durations on target words following semantically related previews were estimated to be 23 ms and 19 ms shorter in SFD and GD analyses, compared with the unrelated conditions ($b=-.074$, $SE=.026$, $t=-2.83$ and $b=-.063$, $SE=.028$, $t=-2.26$). This semantic preview benefit was not significant for FFDs ($t=-1.21$). The reliable differences between semantically related and unrelated

Please cite this article as: Tsai, J.-L., et al., Parafoveal semantic information extraction in traditional Chinese reading, *Acta Psychologica* (2012), doi:10.1016/j.actpsy.2012.06.004

preview conditions indicate that parafoveal semantic information extraction also exists in traditional Chinese reading.

7.3. Preview duration and pre-target word frequency modulation effects

The goal of these analyses was to test whether the fixation duration on the pre-target word *N* modulates preview benefit in reading traditional Chinese. Similar to Kliegl et al. (2012) and Yan, Risse, et al. (2012), these effects translate into interactions between a continuous predictor of fixation duration (FFD, SFD, or GD) on pre-target word *N* and treatment contrasts for semantic, orthographic, phonological, and identical previews using the unrelated preview as a reference category in LMMs. Fixation durations on pre-target words were filtered according to the same criteria.

Preview benefit may not only depend on preview duration, but also on frequency of the pre-target word. For example, compared to a less frequent pre-target word, a more highly frequent word implies an increased perceptual span because of a lower processing load and is likely to result in a larger preview benefit (Henderson & Ferreira, 1990; Kliegl, Nuthmann, & Engbert, 2006; Yan, Kliegl, Shu, Pan, & Zhou, 2010).

We report an extended model using GDs on pre-target word *N* and its frequency as an independent variable (both were log-transformed and centered) and SFDs on target word $N+1$ as a dependent variable. The trial selection led to 1001 observations. The main effects were evaluated at the mean of the log preview GD and at the mean of the log frequency of the pre-target word. Similar trends (not always significant) were also present for models using other combinations of dependent and independent variables.

7.3.1. Identical preview benefit

The main effect of identical preview was highly significant in the extended model ($b=-.175$, $SE=.029$, $t=-6.03$). The identical preview benefit (the difference between the lines for unrelated and identical conditions in Fig. 2) was also significantly modulated by preview GD: The identical preview benefit increased significantly with preview duration. The null hypothesis that the two lines are parallel was rejected ($b=-.161$, $SE=.074$, $t=-2.17$).

7.3.2. Preview cost

The increase in identical preview benefit mainly resulted from an increase in SFD on target word $N+1$ with increasing preview duration for the unrelated preview ($b=.140$, $SE=.054$, $t=2.58$). The numeric decrease in SFD on target word $N+1$ with increasing preview duration for the identical preview was not significant by itself ($b=-.021$, $SE=.053$, $t=-.39$). Thus, in line with previous studies (Kliegl et al., 2012; Yan, Risse, et al., 2012), the increase in SFD with preview duration for unrelated preview condition in this study may be interpreted as evidence for a preview cost.

7.3.3. Phonological preview benefit

Although the phonological preview benefit was not reliable by itself nor significantly interacted with preview duration, a significant three-way interaction existed for phonological preview benefit ($b=-.083$, $SE=.032$, $t=-2.64$). As shown in the left panel of Fig. 3, the difference between unrelated and phonologically related previews (the phonological preview benefit) enlarged with increasing gaze duration for high-frequency pre-target words ($b=-.224$, $SE=.116$, $t=-1.93$), whereas, as shown in the right panel of Fig. 3, the preview duration modulation effect on phonological preview benefit was unreliable when pre-target words were of low frequency ($b=.101$, $SE=.138$, $t=0.73$).

8. Discussion

Researchers have reported parafoveal semantic extraction for alphabetic scripts in German, but not in English. The inconsistency might be due to language-specific visual features (capitalized initial letters of German nouns that might attract attention; Hohenstein et al., 2010) or a more regular mapping between orthography and phonology in German, which presumably facilitates lexical access in an orthography-to-phonology-to-semantics route (Hohenstein & Kliegl, submitted for publication). In contrast, presumably because of a substantial cross-language difference between Chinese and alphabetic scripts, such semantic preview benefit has been consistently reported in simplified Chinese reading (Yan et al., 2009; Yan, Zhou, et al., 2012; Yang et al., 2012). The present study demonstrates (a) access to parafoveal semantic information extending to traditional Chinese, and

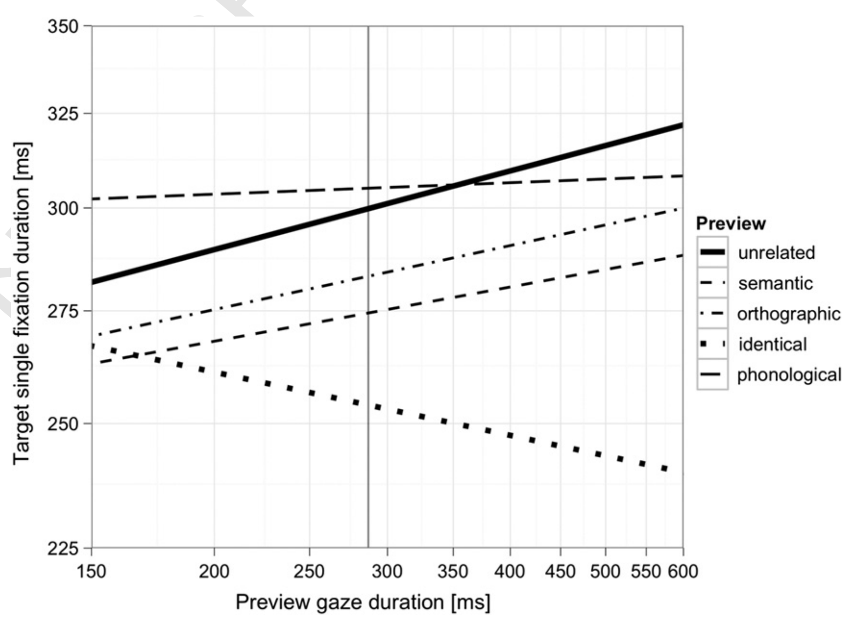

Fig. 2. Linear regression of SFD on word $N+1$ as a function of GD on word *N* for unrelated (solid), orthographic (short-dashed), phonological (long-dashed), semantic (dot-dashed), and identical (dotted) character previews using logarithmic scales for both axes. The vertical line indicates the mean log SFD on word *N* (the value at which main effects are evaluated). Between-subject and between-item differences for the dependent variable and covariance in the LMM were removed prior to regressions. The figure was generated with ggplot2 (Wickham, 2009).

Q2

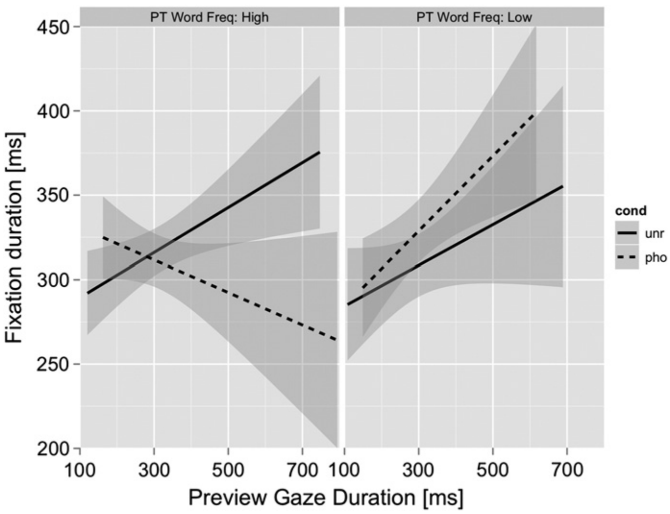

Fig. 3. Single fixation duration on target words as a function of preview duration for more frequent (left panel) and less frequent (right panel) pre-target words. Errorbands show 90% confidence intervals.

(b) preview duration modulation effects on various types (identical, phonological, and unrelated) of parafoveal information extraction.

8.1. Semantic preview benefit in traditional Chinese

Using the boundary paradigm, recent studies have demonstrated reliable semantic preview benefits from word $N+1$ (Yan et al., 2009; Yan, Zhou, et al., 2012; Yang et al., 2012) and a preview benefit at the morphological level (Yen, Tsai, Tzeng, & Hung, 2008). However, this is the first study to report reliable facilitation of fixation duration because of parafoveal semantic overlap during prior fixations in traditional Chinese reading. Compared with simplified Chinese, the advantage for semantic preview benefits in traditional Chinese (less ambiguity in mapping orthography to semantic or phonology) might overcome the disadvantage of heavier decoding caused by the higher visual complexity of characters during parafoveal processing. The semantic preview benefit obtained with visually complex characters in traditional Chinese is an important result. Because of the frequency and visual complexity constraints of the material used in Yan et al. (2009), these authors indicated that "generalizability (of semantic preview benefit) to Chinese characters at large still remains to be established" (p. 564). The present study delivered this result (see also Yan, Zhou, et al., 2012).

The present study results also serve as an example to support a recent view by Frost (accepted for publication), who argued that a particular language has specific optimization information that the orthographic code conveys. The failure of finding semantic benefit in English might indicate that its orthography has been optimized for extracting phonological—but not semantic—information during language evolution. Both simplified and traditional Chinese are well prepared for early and fast semantic access. Visual and lexical factors may be responsible for the advantage of Chinese semantic processing. At the visual level, a small region occupied by the Chinese character conveys morphological or semantic information. At the lexical level, in many cases, the components or radicals of characters provide categorical meaning for the specific meaning of characters. For example, characters with the radical "扌" mean actions with hands (for hitting; for stirring). These unique features may facilitate semantic processing when reading Chinese script (Yan, Zhou, et al., 2012).

Acquiring useful semantic information from parafoveal visible words has assumed a major role in the current debate on oculomotor models of eye movement control during reading. Extracting semantic information from the parafovea appears to be more compatible with guidance from attentional gradient (GAG) models such as SWIFT (Engbert, Nuthmann, Richter, & Kliegl, 2005; Schad & Engbert, 2012), which assumes distributed lexical processing in the perceptual span and generally allows semantic extraction to go beyond the currently fixated word if the parafoveal words fall into the perceptual span. Because of the language-specific characteristics of Chinese scripts, our results may not be taken as strong evidence against the sequential attention shift (SAS) models such as E-Z Reader (Reichle, Pollatsek, & Rayner, 2012), which was developed mainly for English reading.

8.2. Time course of parafoveal information extraction

8.3. Identical preview benefit

Identical preview benefit in SFD increased over preview duration. This increase in the difference between identical and unrelated previews was caused by a numeric decrease of duration from identical preview and a significant increase of duration from unrelated preview. With identical previews, readers acquired beneficial and facilitative information during the entire preview duration; the accumulation of such information contributed to an increasing identical preview benefit over preview duration. The result is in agreement with those of Yan, Risse, et al. (2012) and Schroyens, Vitu, Brysbaert, and d'Ydewalle (1999), who reported a larger preview benefit on target word $N+1$ with increasing pre-target durations when they presented a triad of words with an invisible boundary for display change located between the first and second words and a manipulated preview of word $N+1$ during pre-boundary fixations on word N.

8.3.1. Preview cost

A preview of unrelated characters revealed a significant increase in SFD on target words. Such an increase with preview duration likely reflects interference with target word processing because of earlier or ongoing processing of the unrelated preview. The unrelated preview condition has traditionally been considered an adequate baseline for computing preview benefit. However, together with recent studies (Kliegl et al., 2012; Yan, Risse, et al., 2012), the increasing processing difficulty of target words induced by unrelated or non-word previews has suggested that classical preview benefits may partially arise as a consequence of preview cost associated with long previews.

8.3.2. Phonological preview benefit

Phonological preview benefits have been consistently shown in GD but not in FFD during reading of Chinese sentences (Liu et al., 2002; Tsai et al., 2004; Yan et al., 2009). The present study further demonstrated that the phonological preview effect of non-compound traditional characters requires both long preview duration and high parafoveal processing efficiency afforded by high-frequency pre-target words. Conditional phonological preview effects could be due to non-compound characters not having phonetic radicals to provide phonological information. The present study and the study by Yan et al. (2009) may not serve as fair tests on the phonological effects in Chinese reading because of the small percentage of non-compound characters and the absence of phonological radicals in these studies. Effective information extraction from parafoveal radicals has been shown by both Tsai et al. (2004) and Yan, Zhou, et al. (2012). The results obtained by Tsai et al. (2004; Experiment 2) indicated that the consistency of phonetic radicals to the character's sound could affect phonological preview benefit. Similarly, Yan, Zhou, et al. (2012) also demonstrated that Chinese readers could extract useful semantic information from semantic radicals of the preview characters. Both studies indicated activation of sub-lexical information during parafoveal processing. Further research must investigate whether Chinese phonetic and semantic radicals of compound characters can function to show different sub-lexical preview effects interacting with preview time or pre-target frequency.

Please cite this article as: Tsai, J.-L., et al., Parafoveal semantic information extraction in traditional Chinese reading, *Acta Psychologica* (2012), doi:10.1016/j.actpsy.2012.06.004

9. Conclusion

This study reports a reliable semantic preview benefit and preview duration modulation effect on identical and phonological preview benefits. The results generalize previous findings of access to parafoveal semantic information from simplified to traditional Chinese and extend our understanding of the effects of preview duration on parafoveal information extraction.

10. Uncited reference

Rayner, 2009

References

Academia Sinica Balanced Corpus (2004). *[CDROM]: Academia Sinica, Taipei, Taiwan.* .

Altarriba, J., Kambe, G., Pollatsek, A., & Rayner, K. (2001). Semantic codes are not used in integrating information across eye fixations in reading: Evidence from fluent Spanish–English bilinguals. *Perception & Psychophysics*, *63*, 875–890.

Bates, D., & Maechler, M. (2010). *lme4: Linear mixed-effects models using S4 classes (R package version 0.999375-32) [Computer software].* Vienna, Austria: R Foundation for Statistical Computing.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436, http://dx.doi.org/10.1163/156856897X00357.

Chen, H. -C., Tsang, Y. -K. (submitted for publication). Eye movement control in reading: Logographic Chinese versus alphabetic scripts.

Cornelissen, F. W., Peters, E., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers*, *34*, 613–617.

Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research*, *43*(9), 1035–1045.

Engbert, R., Nuthmann, A., Richter, E., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, *112*, 777–813.

Frost, R. (accepted for publication). Towards a universal model of reading. Behavioral and Brain Sciences.

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *16*, 417–429.

Hohenstein, S., Kliegl, R. (submitted for publication). Semantic preview benefit during reading.

Hohenstein, S., Laubrock, J., & Kliegl, R. (2010). Semantic preview benefit in eye movements during reading: A parafoveal fast-priming study. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*, 1150–1170.

Hoosain, R. (1991). *Psycholinguistic implications for linguistic relativity: A case study of Chinese.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Inhoff, A. W. (1990). Integrating information across eye fixations in reading: The role of letter and word units. *Acta Psychologica*, *73*, 281–297.

Kliegl, R., Hohenstein, S., Yan, M., & McDonald, S. A. (2012). How preview space/time translates into preview cost/benefit for fixation durations during reading. *The Quarterly Journal of Experimental Psychology*, http://dx.doi.org/10.1080/17470218.2012.658073.

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology. General*, *135*, 13–35.

Liu, W., Inhoff, A. W., Ye, Y., & Wu, C. (2002). Use of parafoveally visible characters during the reading of Chinese sentences. *Journal of Experimental Psychology. Human Perception and Performance*, *28*, 1213–1227.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442, http://dx.doi.org/10.1163/156856897X00366.

Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology. Human Perception and Performance*, *18*(1), 148–162.

R Development Core Team (2010). *R: A language and environment for statistical computing.* Wien: R Foundation for Statistical Computing (Computer software).

Rayner, K. (1975). The perceptual span and peripheral cues during reading. *Cognitive Psychology*, *7*, 65–81.

Rayner, K. (2009). The Thirty Fifth Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*, 1457–1506.

Rayner, K., Balota, D. A., & Pollatsek, A. (1986). Against parafoveal semantic preprocessing during eye fixations in reading. *Canadian Journal of Psychology*, *40*, 473–483.

Reichle, E. D., Pollatsek, A., & Rayner, K. (2012). Using E-Z Reader to simulate eye movements in non-reading tasks: A unified framework for understanding the eye-mind link. *Psychological Review*, *119*, 155–185.

Schad, D., & Engbert, R. (2012). The zoom lens of attention: Simulating shuffled versus normal text reading using the SWIFT model. *Visual Cognition*, *20*, 391–421.

Schroyens, W., Vitu, F., Brysbaert, M., & d'Ydewalle, G. (1999). Eye movement control during reading: Foveal load and parafoveal processing. *The Quarterly Journal of Experimental Psychology*, *52A*, 1021–1046.

Tsai, J. L., Lee, C. Y., Tzeng, O. J. L., Hung, D. L., & Yen, N. S. (2004). Use of phonological codes from Chinese characters: Evidence from processing of parafoveal preview when reading sentences. *Brain and Language*, *91*, 235–244.

Tsai, J. L., & McConkie, G. W. (2003). Where do Chinese readers send their eyes? In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 159–176). Oxford: Elsevier.

Yan, M., Kliegl, R., Shu, H., Pan, J., & Zhou, X. (2010). Parafoveal load of word $n+1$ modulates preprocessing of word $n+2$. *Journal of Experimental Psychology. Human Perception and Performance*, *36*, 1669–1676, http://dx.doi.org/10.1037/a0019329.

Yan, M., Richter, E. M., Shu, H., & Kliegl, R. (2009). Chinese readers extract semantic information from parafoveal words during reading. *Psychonomic Bulletin & Review*, *16*, 561–566, http://dx.doi.org/10.3758/PBR.16.3.561.

Yan, M., Risse, S., Zhou, X., & Kliegl, R. (2012). Preview fixation duration modulates identical and semantic preview benefit in Chinese reading. *Reading and Writing*, *25*, 1093–1111, http://dx.doi.org/10.1007/s11145-010-9274-7.

Yan, M., Zhou, W., Shu, H., & Kliegl, R. (2012). Lexical and sub-lexical semantic preview benefits in Chinese reading. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, http://dx.doi.org/10.1037/a0026935 (Advance online publication).

Yang, J., Wang, S., Tong, X., & Rayner, K. (2012). Semantic and plausibility effects on preview benefit during eye fixations in Chinese reading. *Reading and Writing*, *25*, 1031–1052, http://dx.doi.org/10.1007/s11145-010-9281-8.

Yen, M. -H., Tsai, J. -L., Tzeng, O. J. -L., & Hung, D. L. (2008). Eye movements and parafoveal word processing in reading Chinese sentences. *Memory & Cognition*, *36*, 1033–1045.

Please cite this article as: Tsai, J.-L., et al., Parafoveal semantic information extraction in traditional Chinese reading, *Acta Psychologica* (2012), doi:10.1016/j.actpsy.2012.06.004