Parafoveal Processing of Phonology and Semantics during the Reading of Korean Sentences

Ming Yan [(1)], Aiping Wang [(2)], Hosu Song [(2)], Reinhold Kliegl [(3)]

(1) Department of Psychology, University of Macau, Taipa, Macau

(2) School of Psychology, Beijing Normal University, Beijing, People's Republic of China

(3) Department of Psychology, University of Potsdam, Potsdam, Germany

Author Note

This research was supported by the Ministry of Education in China Project of Humanities and Social Sciences (15YJA190007). The authors thank Hua Shu and Wei Zhou for their efforts on the project, and Denis Drieghe and an anonymous reviewer for their helpful comments. Ming Yan and Aiping Wang equally contributed to this work. Correspondence concerning this article should be sent to Ming Yan (e-mail: mingyan@um.edu.mo). Data and R scripts used in the current study are available at the Potsdam Mind Research Repository http://read.psych.uni-potsdam.de/pmr2/index.php?option=com_content&view=article&id=184) . Declarations of interest: none.

ABSTRACT

The present study sets out to address two fundamental questions in the reading of continuous texts: Whether semantic and phonological information from upcoming words can be accessed during natural reading. In the present study we investigated parafoveal processing during the reading of Korean sentences, manipulating semantic and phonological information from parafoveal preview words. In addition to the first evidence for a semantic preview effect in Korean, we found that Korean readers have stronger and more long-lasting phonological than semantic activation from parafoveal words in second-pass reading. The present study provides an example that human mind can flexibly adjust processing priority to different types of information based on the linguistic environment.

Keywords: semantic, phonological, preview benefit, Korean

Processing Parafoveal Words in Korean

During reading of continuous texts, various types of information can be obtained not only from the currently fixated foveal words but also from upcoming parafoveal words. One topic that has been explored in the past decades is the type and priority of parafoveal linguistic information available for processing even before a word is fixated. The topic is typically tested with the gaze-contingent boundary-paradigm (Rayner, 1975): during fixations on foveal pre-target word N-1, the parafoveal preview of an upcoming target word N is presented using either its original form, a word related to the target, or a completely unrelated (non-)word. During a saccade from word N-1 to N crossing an invisible boundary between these words, the correct word N is revealed. Preview benefit is indicated by shorter fixation-durations on word N for identical/related over unrelated previews. In classic experiments in English, preview benefit has been consistently demonstrated for orthographic (Inhoff, 1990; Inhoff & Tousman, 1990; Rayner, 1975) and phonological overlaps (e.g., Pollatsek, Lesch, Morris, & Rayner, 1992). The first sentence reading experiments using the boundary-paradigm on Asian scripts extended such orthographic and phonological preview benefits to Chinese (Liu, Inhoff, Ye, & Wu, 2002; Tsai, Lee, Tzeng, Hung, & Yen, 2004).

Among different types of linguisitc information (including orthographic, phonological, semantic and syntactic information), whether semantics can be accessed from parafoveal words has been under considerable debate. There was no evidence for parafoveal processing of semantically-related preview in English for many years (e.g., Inhoff, 1982; Inhoff & Rayner, 1980). Using the boundary-paradigm, Rayner, Balota and Pollatsek (1986; see Rayner, Schotter, & Drieghe, 2014, for a replication) found that, fixation-duration did not statistically differ between semantically-related previews (*tune-song*) and unrelated previews (*door-song*). In contrast, Pollatsek et al. (1992) found early phonological preview benefit from homophone previews (*reins-rains*) over non-homophone visually similar baseline (*ruins-rains*). On the basis of this contrast of a non-significant effect in one study and a significant result in another, it was argued (1) that parafoveal semantic information is not extracted and (2) that parafoveal processing of phonology precedes that of semantics in English. These results have also been interpreted as support of the phonological mediation account which assumes that access to semantics depends on phonological activation (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). However, the conclusion is based on different target words and different readers. Additionally, from a statistical perspective, "*the difference between 'significant' and 'not significant' is not itself significant*" (Gelman & Stern, 2006, p.328).

To our knowledge for alphabetic scripts, there is so far no direct comparison between phonological and semantic preview benefits using within-item and within-subject design. There is, however, such evidence from a character-based script. Yan, Richter, Shu and Kliegl (2009) showed that processing priority among different types of parafoveal information depends on the nature of the writing system. They reported the first evidence for early parafoveal activation of semantics - and earlier than that of phonology during the reading of Chinese sentences (see also Tsai, Kliegl, & Yan, 2012): semantic preview benefit appeared early on in FFD (first-fixation duration, duration of the first-fixations on a word irrespective of number of fixations) whereas phonological preview benefit appeared only in GD (gaze duration, accumulative duration of fixations during first-pass reading). Apparently, effects that emerge within the first fixation on a word are assumed to occur earlier than those that require refixations (Inhoff, 1984). Additionally, a direct comparison between these two conditions showed a larger effect of semantics in FFD (Yan et al., 2009). Apparently, as a logographic script, the direct mapping from orthography to semantics in Chinese promotes early semantic activation. Thus, the results derived from classic experiments in English, that parafoveal processing is sensitive to phonological but not sensitive to semantic information, do not generalize to Chinese.

The work on semantic preview benefit in Chinese has inspired researchers to re-explore the topic in alphabetic scripts. Hohenstein, Laubrock and Kliegl (2010) argued that, as compared to English, shallower orthographic-depth in German leads to faster phonological decoding, which in turn facilitates access to semantics. This makes it possible for German readers to efficiently extract semantic knowledge from parafoveal words. Indeed, across their four experiments using a combination of fast-priming and boundary-paradigm in which the parafoveal preview was presented only for a limited amount of time but not during the whole fixation on pre-target word, semantic preview benefit was first demonstrated in alphabetic scripts. Their follow-up work established semantic preview benefit using the standard boundary-paradigm (Hohenstein & Kliegl, 2014). In English, evidence for semantic preview benefit is also reported but typically depends on high semantic feature overlap with the target (Schotter, 2013; Veldre & Andrews, 2016a), or strong context with highly predictable targets (Schotter & Jia, 2016; Veldre & Andrews, 2016b). Although semantic preview benefit is now widely acknowledged, existing experimental evidence is limited to some European languages with Roman alphabet (English and German).

The present study reports an attempt to extend parafoveal semantic processing to Korean alphabet, whose cognitive processes have been largely understudied. At a first glance, the spatial layout of Korean sentences largely resembles Chinese except that the former comes with spaces between words/phrases. In both scripts each character represents a single syllable and occupies the same horizontal extent in the text. Historically, the Chinese writing system played an important role in the evolution of the Korean language. A significant amount of the Korean vocabulary is borrowed from Chinese. The exact proportion of Sino-Korean vocabulary is estimated between 50% and 65% (Sohn, 2001; 2006). Using pure Korean characters (Hangul) without Chinese characters (Hanja) to write became common since 1970 when the Korean government has been encouraging citizens to stop using Hanja and replace them with Hangul (Sohn, 2001). Nowadays, Hanja is seldom used in daily life but only for the purposes of historical and linguistic studies. As another script with shallow orthographic-depth, the letter–phoneme correspondence in Korean (i.e., Hangul letters) is highly consistent. There are only two studies on the parafoveal processing of Korean words. Kim, Radach, and Vorstius (2012) found that parafoveally presented inappropriate syntactic case markers led to prolongation of fixations on target-words. Their results concluded effective parafoveal processing of syntactic information in Korean. Wang, Yeon, Zhou, Shu and Yan (2016) demonstrated that during the reading of Chinese sentences, semantically-related Korean parafoveal previews facilitated foveal processing of Chinese target-words, reporting the first evidence for cross-language semantic preview effect. One minor potential limitation of these two studies is that neither of them tested parafoveal semantic preview benefit in a *single-language* reading experiment.

The classic boundary paradigm was only designed to compare different types of previews and did not take into consideration preview-time (Kliegl, Hohenstein, Yan, & McDonald, 2013; Li, Wang, Mo, & Kliegl, 2018; Marx, Hawelka, Schuster, & Hutzler, 2015; Yan, Risse, Zhou, & Kliegl, 2012a). Apparently, the longer time spent to process parafoveal preview, the more parafoveal information is acquired.

Preview-time is under the readers' own control and thus cannot be manipulated experimentally. The effect of preview-time, however, can be tested by including fixation-duration on the pretarget-word as a covariate. A 'preview cost' effect from processing unrelated previews, as reflected by increasing fixation duration/times on the target-word with increasing preview-time, has been consistently demonstrated. The preview-time analyses provide theoretical and methodological advance over the traditional analyses, because it examines time course of parafoveal information activation, unfolding effects over multiple processing stages.

The present experiment combined the design ideas reviewed above. First, if semantic preview benefit is a language-universal phenomenon and can be generalized from Roman alphabet to Korean alphabet, we should observe such an effect during the monolingual reading of Korean sentences. Second and more importantly, we test parafoveal priority of phonological versus semantic information using a within-item and within-subject design rather than cross-experiment comparisons. Third, with preview-time analyses we aim at testing time course of parafoveal processing in Korean.

Method

Participants

Forty-four readers (mean age = 22.2 year-old, *SD* = 1.86, 22 females) with normal or corrected-to-normal vision participated in the eye-tracking experiment. Three separate groups of 12, 12 and 20 readers participated in pretests for phonological and semantic relatedness and for predictability norming data. All participants were native Korean students in Beijing Normal University originally from South Korea. Experimental procedures were approved by the Ethics Committee of the School of Psychology, Beijing Normal University. Participants gave their written informed consent prior to the experiment, which conformed to the tenets of the Declaration of Helsinki.

Material

Eighty quadruplets of two-character words were selected for identical, semantically-related, phonologically-related and unrelated previews, which were paired to the same target-word. The non-identical conditions were matched for word frequency ($F < 1$; National Institute of Korean Language, 2005; Table 1) and for visual complexity as indexed by stroke frequency (Zhang, Zhang, Xue, Liu, & Yu, 2007; $F < 1$), which calculates the maximum number of crossed strokes among 6 different slices cutting through the word. Semantic and phonological relatedness between each of the three non-identical previews and the target-word was evaluated on two 5-point scales. The previews were related to the targets only on the desired dimensions: Semantic relatedness was high only for the semantic previews [$F(2,158) = 1696.697$, $p < .001$]. Likewise, was phonological relatedness [$F(2,158) = 1279.897$, $p < .001$]. The whole target-word region including suffix characters was 3.1 characters ($SD = 0.6$) long. The manipulation of phonological similarity was achieved via visual similarity. Apparently, in all alphabetic writing systems especially those with shallow orthographic depths such as Korean, phonological similarity must be achieved via visual similarity. It is not possible to orthogonally manipulate phonological and orthographical similarity.

For each target-word, two different sentence frames were constructed to increase observations and thus statistical power. The target-preceding sentence contexts (including pretarget-words, which were always two-character words) were constructed to be non-predictive for different types of previews in order to minimize top-down processing. As confirmed in a cloze test, preview words were equally unpredictable [$F(1,159) = 2.597$, $p > 0.1$]. Words N-1 and N were never among the first/last three words in the sentences.

Table 1. Preview word properties

	Type of Parafoveal Preview			
	Identical	Semantic	Phonological	Unrelated
Example	바다	대양	빠따	교류
Pronunciation	bada	daeyang	ppatta	gyolyu
Translation	sea	ocean	stick	alternating current
Word frequency	94.1 (138.4)	96.4 (151.9)	90.5 (244.3)	96.3 (228.7)
Visual Complexity	3.07 (0.57)	2.88 (0.43)	2.90 (0.42)	2.85 (0.45)
Predictability	4.0% (6.2%)	0.5% (2.5%)	0.1% (1.6%)	0.0% (0.0%)
Sem. relatedness	N.A.	3.5 (0.5)	1.0 (0.0)	1.0 (1.0)
Pho. relatedness	N.A.	1.0 (0.1)	3.2 (0.5)	1.0 (1.0)

Means (and standard deviations in parentheses) of word frequencies, visual complexity, predictability, semantic and phonological relatedness to the target word.

Apparatus

Participants' eye-movements were recorded with an Eyelink Desktop system running at 1000Hz. Each sentence was presented in a single line on a 21-inch ViewSonic G220f monitor (resolution: 1280-by-1024 pixels; frame rate: 100Hz) using font Batang. Participants were seated 65cm from the monitor, with their heads positioned by a forehead and chin rest. Each character subtended 1.0 degree of visual angle. All recordings and calibrations were done monocularly based on the right eye; viewing was binocular.

Procedure

Gaze-positions were calibrated with a 9-point grid (error < 0.5°). Fixation on the initial fixation-point initiated presentation of the next sentence, with its first character occupying the fixation-point. Participants were instructed to read the sentences silently for comprehension, then fixate a dot in the lower-right corner of the monitor, and finally press a joystick button to signal completion. The gaze-contingent display-change technique was adopted to manipulate the parafoveal visibility of word N (Figure 1). Forty sentences were followed by easy yes-no comprehension questions and the participants correctly answered 89% of them (*SD* = 7%). Thirty participants reported "flashes" on the screen, but they could not report the exact number of trials or what they had seen exactly.

사고로 함몰된 어선은 모두 바다 속에 깊이 잠겼다. (identical preview)
*
사고로 함몰된 어선은 모두 대양 속에 깊이 잠겼다. (semantic preview)
*
사고로 함몰된 어선은 모두 빠따 속에 깊이 잠겼다. (phonological preview)
*
사고로 함몰된 어선은 모두 교류 속에 깊이 잠겼다. (unrelated preview)
*
사고로 함몰된 어선은 모두 바다 속에 깊이 잠겼다. (target sentence)
*
Word N-1 Word N

Figure 1. A set of example sentences. The parafoveal preview of target word N is identical, semantically related, phonologically related or unrelated to the target word, prior to fixations on it, as indexed by asterisks. The previews are immediately replaced by the correct target word once the reader's eyes cross an invisible vertical boundary that follows word N-1. The target sentence translates as: *All of the fishing boats that were caught in the accident were deeply submerged in the sea.*

Data Analysis

Fixations were determined with an algorithm for saccade-detection (Engbert & Kliegl, 2003). Trials were removed due to participants' blinks, coughing or tracker errors (n = 478, 7%). Target words with FFDs shorter than 60 ms or longer than 600 ms and GDs longer than 1000 ms were removed (n = 220, 3%). Additionally, 193 trials (3%) with regressions from the pre-target words were discarded because they may reflect incomplete parafoveal processing of the previews. Finally, 595 trials (9%) in which display-changes were triggered during fixations were excluded because readers were more likely to perceive display-changes or flashes. Together we kept 4840 observations for analyses reported in the Appendix.

Estimates are based on linear mixed models (LMMs) using the lme4 package (Version 1.1-19; Bates, Maechler, Bolker, & Walker, 2015) in the R environment (Version 3.5.1; R Development Core Team, 2018). We included the following fixed-effects for our theoretical interests, (a) a treatment-contrast with the unrelated condition as the baseline (therefore the three levels of the contrast represent identical, semantic and phonological preview benefits, respectively), (b) log-transformed and centered single-fixation duration (where a word was inspected with exactly one fixation) on the pretarget-word as an index of preview-time and (c) their interactions to test the time course of parafoveal processing. The main reason for choosing single-fixations, as in previous studies (Li et al., 2018; Yan et al., 2012a; Yan & Sommer, 2019), is that they carry few mislocated fixations (Nuthmann, Engbert, & Kliegl, 2005). For this purpose, 614 trials, in which the pre-target words were skipped or fixated more than once, were additionally removed from the main analyses below. Fixed-effects of quadratic fixation-location (i.e., inverted optimal viewing position effect; Vitu, McConkie, Kerr, & O'Regan, 2001) and launch site, which have been well documented to influence fixation-duration during reading (e.g., Kliegl, Nuthmann, & Engbert, 2006; Nuthmann et al., 2005), were also included for the sake of statistical control. Their significance are not in the focus of the present study and thus are not reported in the main text. Results from a simpler model with the fixed-effects including only the experimental manipulation but no covariates are reported in the Appendix, in which data analysis was not restricted to those trials when the pre-target word was fixated exactly once.

Analyses of LMM residuals consistently suggested a need for log-transformation of durations (Kliegl, Masson, & Richter, 2010). We started with LMMs with full random-effects, including subject- and item-related variance components for intercepts and random-slopes for fixed-effects. These models were always too complex and not supported by the data. Therefore, parsimonious LMMs were selected following Bates, Kliegl, Vasishth, and Baayen (2015; see also Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2017). In general, there was never sufficient evidence for variance components related to fixation location and interaction terms; there were also no reliable correlation parameters. In addition, small variance parameters were removed until the LMM was supported by the data using the lme4::rePCA function. Table 2 (fixed effects) and Table 3 (variance components) report the estimates of model parameters for the final LMMs. Significance of fixed effects (p-values) reported in Table 2 is based on Satterthwaite's degrees of freedom method (using the *lmerTest* package; Kuznetsova, Brockhoff, & Chritensen, 2017). Significance of effects was also consistent with 95% confidence intervals based on profiled likelihoods (using the lme4:profile function).

Results

Based on 4226 observations in which the pre-target words were fixated only once, Korean readers spent less time fixating target-words following identical than unrelated previews (FFD: 42 ms, *b* = -0.131, *SE* = 0.014, *t* = -9.24; GD: 74 ms, *b* = -0.201, *SE* = 0.020, *t* = -10.27; total reading time, TRT, sum of all fixations on a word including regressive fixations: 87 ms, *b* = -0.236, *SE* = 0.020, *t* = -11.90). Such a canonical effect of identical preview benefit has been consistently documented across different scripts and therefore suggests that the present reading data are comparable to earlier studies. In agreement with the main hypothesis of the present study, the readers also spent less time on target-words after semantically-related (GD: 13 ms, *b* = -0.044, *SE* = 0.016, *t* = -2.78; TRT: 10 ms, *b* = -0.035, *SE* = 0.016, *t* = -2.23) and phonologically-related previews (GD: 21 ms, *b* = -0.048, *SE* = 0.015, *t* = -3.17; TRT: 30 ms, *b* = -0.067, *SE* = 0.015, *t* = -4.40). Such semantic (6 ms, *b* = -0.023, *SE* = 0.015, *t* = -1.55) and phonological preview benefits (8 ms, *b* = -0.021, *SE* = 0.013, *t* = -1.70) appeared in the same numerical trends for FFD but were not statistically significant. In a post-hoc comparison with semantic condition as the reference, TRT were estimated to be 20 ms shorter in the phonological than semantic condition (*b* = -0.032, *SE* = 0.015, *t* = -2.10).

The estimates of interaction terms in Table 2 and Figure 2 show that, for log gaze durations, the identical and phonological preview benefits (i.e., the *differences* between the lines for unrelated and identical/phonological conditions) significantly increased with preview-time (*b* = -0.157, *SE* = 0.050, *t* = -3.15 and *b* = -0.100, *SE* = 0.049, *t* = -2.06). However, unlike for Chinese reading (Li et al., 2018; Yan, Risse, et al., 2012a), we did not observe a significant change in semantic preview benefit across preview time. In addition, we replicated ‘preview cost’ effect: fixation-duration on the target-word following unrelated previews increased with increasing preview-time (FFD: $b = 0.081$, $SE = 0.033$, $t = 2.48$; GD: $b = 0.209$, $SE = 0.042$, $t = 4.93$; TRT: $b = 0.131$, $SE = 0.038$, $t = 3.45$), suggesting that more parafoveally acquired irrelevant information interferes with foveal processing of the target word, replicating results from previous studies (e.g., Kliegl et al., 2013; Yan, et al., 2012a). Figure 2 displays partial effects of the interactions between preview benefit and log preview time for log target-word gaze duration (see Table 2 for tests of parallelism of gaze-duration lines). Partial effects are based on the LMM of log gaze durations, that is after statistically controlling for fixed effects of launch site and fixation location as well as for individual and item differences as listed in Table 3.

Table 2. Linear mixed model estimates for durations: Fixed effects

	First fixation duration			Gaze duration			Total reading time		
Fixed effect	Est	SE	p-value	Est	SE	p-value	Est	SE	p-value
Unrelated	5.637	0.025	<.001 *	5.859	0.032	<.001 *	5.949	0.037	<.001 *
PPB	-0.021	0.013	.089	-0.048	0.015	.002 *	-0.067	0.015	<.001 *
SPB	-0.023	0.015	.128	-0.044	0.016	.006 *	-0.035	0.016	.027 *
IPB	-0.131	0.014	<.001 *	-0.201	0.020	<.001 *	-0.236	0.020	<.001 *
PT	0.081	0.033	.014 *	0.209	0.042	<.001 *	0.131	0.038	<.001 *
FL ^ 1	0.780	0.372	.036 *	-3.760	0.459	<.001 *	-4.723	0.458	<.001 *
FL ^ 2	-4.475	0.310	<.001 *	1.021	0.378	.007 *	2.036	0.374	<.001 *
LS	-0.130	0.036	<.001 *	-0.124	0.045	.009 *	-0.036	0.042	.406
PPB x PT	-0.016	0.040	.695	-0.100	0.049	.040 *	-0.014	0.048	.769
SPB x PT	0.046	0.041	.256	0.003	0.048	.943	0.060	0.047	.207
IPB x PT	-0.051	0.040	.208	-0.157	0.050	.002 *	-0.089	0.050	.075

Note. PPB = Phonological Preview Benefit, SPB = Semantic Preview Benefit and IPB = Identical Preview Benefit, PT = Preview Time, FL^1= Fixation Location (linear), FL^2 = Fixation Location (quadratic), LS = Launch Site. Analyses were based on 44 subjects and 160 items.

Table 3. Linear mixed model estimates for durations: Variance components

Variance component	FFD	GD	TRT
Subject – unrelated	0.153	0.194	0.229
Subject – PPB	0.045	*NA*	*NA*
Subject – SPB	*NA*	*NA*	*NA*
Subject – IPB	0.046	0.081	0.079
Subject – lauch site	0.159	0.214	0.188
Subject – prev dur	0.093	0.146	0.085
Item – unrelated	0.041	0.074	0.086
Item – PPB	0.017	*NA*	0.029
Item – SPB	0.048	0.063	0.063
Item – IPB	*NA*	*NA*	0.063
Item – launch site	0.063	*NA*	*NA*
Item – prev dur	*NA*	*NA*	*NA*
Residual	0.282	0.342	0.338

Note. Values are standard deviations (i.e., square root of variance). PPB = Phonological Preview Benefit, SPB = Semantic Preview Benefit and IPB = Identical Preview Benefit. FFD = First-Fixation Duration, GD = Gaze Duration and TRT = Total Reading Time. The analyses were based on 44 subjects and 160 items. *NA* = variance component not supported by data.

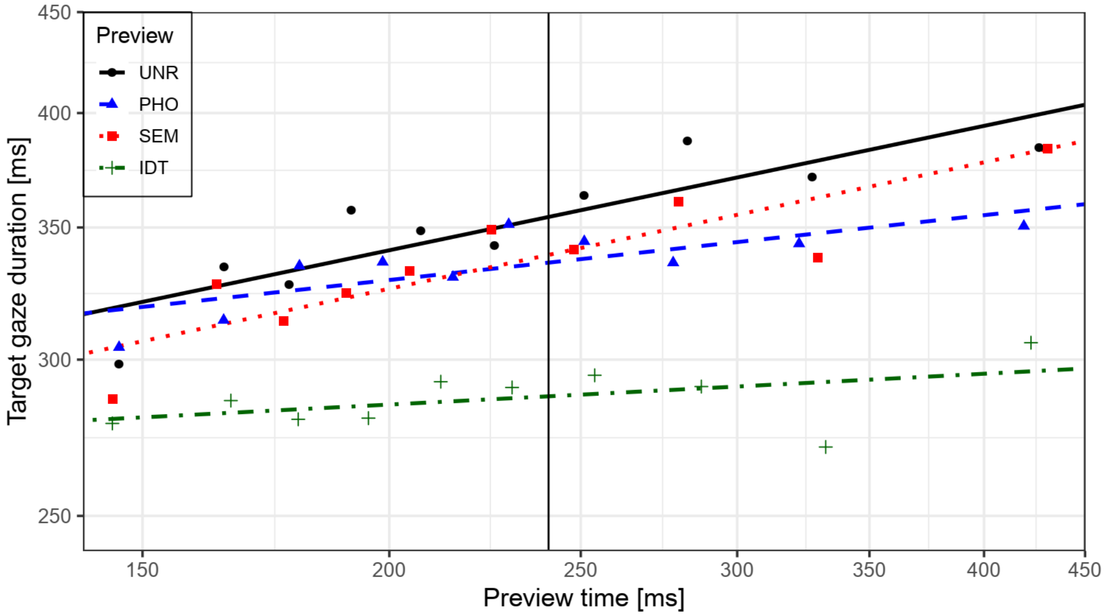

Figure 2. Visualization of experimental effects (UNR - unrelated: solid line and circle points; PHO - phonologically-related: dashed line and triangle points; SEM - semantically-related: dotted line and square points; IDT - and identical: dot-dashed line and cross points). Regression lines are based on partial-effect estimates of LMM for gaze durations; points are means of these estimates computed for decile bins on preview time (mean number of fixations / point = 106, range: 99-114). The vertical line indicates the mean log single-fixation duration on word N-1.

Discussion

In the present study we investigated parafoveal processing during the reading of Korean sentences, aiming at (1) extending semantic preview benefit from logograph (Chinese) and Latin alphabets (German and English) to Korean alphabet and (2) exploring processing priority between different types of parafoveal information. So far, no attempt has been made to differentiate priority between semantic and phonological information during parafoveal processing of alphabetic scripts. Since both types of information has been shown to influence parafoveal processing, it remains an open question whether their relative importance is equivalent. The present study sets out to address these theoretical issues by providing experimental evidence on parafoveal semantic and phonological activation during the reading of Korean sentences.

Yan et al. (2009) argued that the nature of a writing system decides the type and priority among different parafoveal information that can be processed parafoveally. Indeed, semantic preview benefit has been consistently demonstrated in Chinese (Li et al., 2018; Tsai et al., 2012; Yan, Pan, Bélanger, & Shu, 2015; Yan, Zhou, Shu, & Kliegl, 2012b; Yang, Wang, Tong, & Rayner, 2012; Yen, Tsai, Tzeng, & Hung, 2008), arguably due to its optimization for fast and direct semantic but not phonological activation (Hoosain, 1991; Reilly & Radach, 2012). Although semantic preview benefit was traditionally elusive in English, as in the case of German, the shallow orthographic depth (i.e., highly transparent letter-to-phoneme correspondence) leads to faster phonological decoding and facilitates access to semantics (Hohenstein et al., 2010; Hohenstein & Kliegl, 2014). In contrast, semantic preview benefit in English is limited to synonym previews (e.g., Schotter, 2013; Veldre & Andrews, 2016a, 2016b). The present findings apparently not only contribute to a growing body of evidence for semantic preview benefit, but also generalize this effect to a very different writing system.

The Korean Hangul is very unique. First, it is similar to Chinese that each character represents a single syllable and occupies the same horizontal and vertical extent. Korean words are typically much shorter than English words. In this case, upcoming words are on average less eccentric, enabling Korean readers to benefit from the higher visual acuity. Second, Korean is similar to German with respect to orthographic-depth: Both scripts have highly regular letter-to-phoneme correspondences. Therefore, it may be easier to observe semantic preview in Korean and German than in English (Hohenstein et al., 2010). Obviously, although semantic preview benefits are seemingly of similar magnitude in Chinese (in a range of 12 ms to 31 ms in GD; Pan et al., 2016; Tsai et al., 2012; Yan et al., 2009; 2012b) and German (in a range of 13 ms to 31 ms in GD; Hohenstein & Kliegl, 2014) /Korean (19 ms in GD), the effects arise from very different mechanisms: the Chinese orthography directly represents meaning while access to semantics in German/Korean is indirect via phonology.

Finally, with respect to the spatial layout, Korean is similar to both German and English but different from Chinese, Japanese and Thai, in that it is written with spaces between words/phrases (Lee & Ramsey, 2000). Research on Chinese reading indicates that additional processing cost is associated with word boundary ambiguity due to the absence of word spacing (e.g., Inhoff & Wu, 2005; Yan & Kliegl, 2016). Apparently, the explicit word knowledge in Korean may save more attentional resources for parafoveal processing. Taken together, linguistic properties such as high information density, shallow orthographic-depth and spaces between words/phrases may have jointly provided favorable conditions for parafoveal, especially semantic processing.

As an alphabetic writing system with shallow orthographic depth, Korean letters map to syllables in a highly transparent way. Therefore, Korean words that are phonologically similar are also visually similar. Previous studies showed that, shallow orthographic-depth leads to fast phonological decoding (Hohenstein, et al., 2010; Hohenstein & Kliegl, 2014; Wang et al., 2016), activating words' phonological representations. Pollatsek et al.'s (1992) finding of phonological preview benefit from homophone previews over non-homophone visually similar baseline strongly indicates that it is unlikely for alphabetic readers to maintain the orthographic forms of words without phonological activation, otherwise there would not have been any difference between these two conditions. Based on these findings, it is reasonable to interpret the effects to phonological rather than orthographic processing especially for those appeared in TRT. In the present study, we did not aim at completely distinguishing orthographic and phonological effects, due to the nature of the language itself. Nevertheless, future studies on writing systems with deep orthographic depths, such in English, are in better positions to tease them apart.

The alphabetic nature of Hangul implies that the well-known sequential orthography-phonology-semantics activation with access to word meaning in a relatively late stage (Van Orden, 1987) holds true for Korean. Developmental studies showed that phonological awareness is important for Korean and English children but less so for Chinese (Cho, McBride-Chang, & Park, 2008; McBride-Chang et al., 2005). As long as parafoveal processing is concerned, existing evidence in Chinese consistently shows that semantics precedes phonology (Tsai et al., 2012; Yan et al., 2009). In contrast to Chinese and as the first direct comparison between parafoveal phonology and semantics in an alphabetic script, the present study showed that parafoveal phonological information remains activated for a long period of preview-time, indicating long lasting activation of parafoveal phonology, or less disruption from phonologically than semantically related previews in a late processing stage. Indeed, we found numerically stronger phonological than semantic preview effect across different fixation measures. We consider the difference of 20 ms between the phonological and semantic conditions in TRT as evidence supporting advantage in parafoveal phonological over semantic extraction. It is of great theoretical interest and importance for future research to determine the relative priority of parafoveal processing in other scripts.

References

Bates, D., Kliegl, R., Vasishth, S., & Baayen, R.H. (2015). Parsimonious mixed models. arXiv: 1506.04967 [stat.ME]

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67,1-48.

Cho, J.R., McBride-Chang, C., & Park, S. G. (2008). Phonological Awareness and Morphological Awareness: differential Associations to Regular and Irregular Word Recognition in Early Korean Hangul Readers. *Reading and Writing: An Interdisciplinary Journal, 21*, 255-274.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*, 204-256.

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

Gelman, A., & Stern, H. (2006). The difference between „significant“ and „not significant” is not itself statistically significant. *The American Statistician, 60,* 328-331.

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.

Hohenstein, S., & Kliegl, R. (2014). Semantic preview benefit during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 166-190. doi:10.1037/a0033670

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

Inhoff, A.W. (1982). Parafoveal word perception: A further case against semantic preprocessing. *Journal of Experimental Psychology: Human Perception and Performance, 8*, 137-145.

Inhoff, A.W. (1984). Two stages of word processing during eye fixations in the reading of prose. *Journal of Verbal Learning and Verbal Behavior, 23*, 612–624.

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

Inhoff, A.W. & Rayner, K. (1980). Parafoveal word perception: A case against semantic preprocessing. *Perception & Psychophysics, 27*, 457-464.

Inhoff, A.W. & Tousman, S. (1990). Lexical integration across saccades in reading. *Psychological Research, 52*, 330–337.

Inhoff, A.W., & Wu, C. (2005). Eye movements and the identification of spatially ambiguous words during Chinese sentence reading. *Memory & Cognition, 33*, 1345-1356.

Kim, Y.S., Radach, R. & Vorstius, C. (2012). Eye movements and parafoveal processing during reading in Korean. *Reading and Writing, 25*, 1053-1078.

Kliegl, R., Hohenstein, S., Yan, M., & McDonald, S.A. (2013). How preview space/time translates into preview cost/benefit for fixation durations during reading. *Quarterly Journal of Experimental Psychology, 66,* 581-600.

Kliegl, R., Masson, M.E.J., & Richter, E.M. (2010). A linear mixed model analysis of masked repetition priming. *Visual Cognition, 18*, 655-681.

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.

Kuznetsova, A., Brockhoff, P.B., & Christensen, R.H.B. (2017). *lmerTest* package:

Test in linear mixed effects models. *Journal of Statistical Software, 82,* 1-26.

Lee, I., & Ramsey, R. (2000). *The Korean language*. Albany, NY: State University of New York Press.

Li, N., Wang, S., Mo, L., & Kliegl, R. (2018). Contextual constraint and preview time modulate the semantic preview effect: Evidence from Chinese sentence reading. *Quaterly Journal of Experimental Psychology, 71,* 241-249.

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.

Marx, C., Hawelka, S., Schuster, S., & Hutzler, F. (2015). An incremental boundary study on parafoveal preprocessing in children reading aloud: Parafoveal masks overestimate the preview benefit. *Journal of Cognitive Psychology*, 27, 549–561.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language, 94*, 305-315.

National Institute of Korean Language (2005). Modern Korean Usage Frequency Survey. http://www.korean.go.kr

Nuthmann, A., Engbert, R., & Kliegl, R. (2005). Mislocated fixations during reading and the inverted optimal viewing position effect. *Vision Research, 45,* 2201-2217.

Pan, J., Laubrock, J., & Yan, M. (2016). Parafoveal processing in silent and oral reading: Reading mode influences the relative weighting of phonological and semantic information in Chinese. *Journal of Experimental Psychology: Learning, Memory and Cognition, 42*(8), 1257-1273. doi: 10.1037/xlm0000242

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.

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

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

Rayner, K., Schotter, E.R., & Drieghe, D. (2014). Lack of semantic parafoveal preview benefit in reading revisited. *Psychonomic Bulletin & Review, 21*, 1067–1072.

Reilly, R., & Radach, R. (2012). The dynamics of reading in non-Roman writing systems. *Reading and Writing, 25,* 935-950.

Schotter, E.R. (2013). Synonyms provide semantic preview benefit in English. *Journal of Memory and Language, 69*, 619-633.

Schotter, E.R., & Jia, A. (2016). Semantic and plausibility preview benefit effects in English: Evidence from eye movements. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 42*, 1839–1866.

Sohn, Ho-Min (2001). *The Korean Language.* Cambridge, MA: Cambridge University Press.

Sohn, Ho-Min (2006). *Korean language in culture and society.* Boston: Twayne Publishers.

Tsai, J.-L., Kliegl, R., & Yan, M. (2012). Parafoveal semantic information extraction in Traditional Chinese reading. *Acta Psychologica, 141,* 17-23. doi:10.1016/j.actpsy.2012.06.004

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

Van Orden, G.C. (1987). A rows is a rose: Spelling, sound, and reading. *Memory & Cognition, 15, 181–198*. doi:10.3758/BF03197716

Veldre, A., Andrews, S. (2016a). Semantic preview benefit in English: Individual differences in the extraction and use of parafoveal semantic information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42(6)*, 837-854.

Veldre, A., & Andrews, S. (2016b). Is semantic preview benefit due to relatedness or plausibility? *Journal of Experimental Psychology: Human Perception and Performance, 42*, 939–952.

Vitu, F., McConkie, G.W., Kerr, P., O'Regan, J.K. (2001). Fixation location effects on fixation durations during reading: an inverted optimal viewing position effect. *Vision Research, 41,* 3513-3533

Wang, A., Yeon, J., Zhou, W., Shu, H., & Yan, M. (2016). Cross-language parafoveal semantic processing: Evidence from Korean-Chinese bilinguals. *Psychonomic Bulletin & Review, 23(1)*, 285-290. doi:10.3758/s13423-015-0876-6

Yan, M., & Kliegl, R. (2016). CarPrice versus CarpRice: Word boundary ambiguity influences saccade target selection during the reading of Chinese sentences. *Journal of Experimental Psychology: Learning, Memory and Cognition, 42(11)*, 1832-1838. doi:10.1037/xlm0000276

Yan, M., Pan, J., Bélanger, N. & Shu, H. (2015). Chinese deaf readers have early access to parafoveal semantics. *Journal of Experimental Psychology: Learning, Memory and Cognition, 41,* 254-261. doi:10.1037/xlm0000035

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. doi:10.3758/PBR.16.3.561

Yan, M., Risse, S., Zhou, X., & Kliegl, R. (2012). Preboundary duration modulates semantic preview benefit for word n+1 and n+2 in Chinese reading. *Reading and Writing, 25,* 1093-1111.

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, 38,* 1069-1075. doi:10.1037/a0026935

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. doi: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 and Cognition*, *36*, 1033-1045. doi:10.3758/MC.36.5.1033

Zhang, J.Y., Zhang, T., Xue, F., Liu, L., & Yu, C. (2007). Legibility variations of Chinese characters and implications for visual acuity measurement in Chinese reading population. *Investigative Ophthalmology & Visual Science, 48*, 2383-2390.

Appendix

In this section we document results from a basic LMM without any covariates. There were no significant experimental effects on skipping probabilities of the pre-target and target words (|*z*-values| < 1; Table A1). Based on 4840 observations, Korean readers showed identical preview benefits in FFD (41ms, *b* = -0.126, *SE* = 0.017, *t* = -7.40), GD (79ms, *b* = -0.212, *SE* = 0.022, *t* = -9.81) and TRT (96ms, *b* = -0.256, *SE* = 0.021, *t* = -12.22). The readers also showed semantic preview benefit in GD (18ms, *b* = -0.056, *SE* = 0.014, *t* = -3.71) and TRT (14ms, *b* = -0.046, *SE* = 0.015, *t* = -3.03) and phonological preview benefit in GD (23ms, *b* = -0.056, *SE* = 0.015, *t* = -3.67) and TRT (35ms, *b* = -0.082, *SE* = 0.015, *t* = -5.39). Semantic and phonological preview benefits were not statistically significant in FFD (6ms, |*t*-values| < 1.3). The advantage of phonological over semantic preview benefit in TRT was estimated to 21ms (*b* = -0.036, *SE* = 0.016, *t* = -2.26). Finally, consistent with preview benefit effects, our readers made less refixations of the target words in the identical (*b* = -0.641, *SE* = 0.115, *z* = -5.57, *p* < .001), phonologically related (*b* = -0.234, *SE* = 0.096, *z* = -2.44, *p* = .015) and semantically related conditions (*b* = -0.188, *SE* = 0.095, *z* = -1.97, *p* = .048) than in the unrelated condition. Their refixation probabilities of the pre-target words did not differ across conditions (|*z*-values| < 1).

Table A1. Fixation measures for pre-target and target words.

	Type of Parafoveal Preview			
	Identical	Semantic	Phonological	Unrelated
SP-PT	11.6 (14.1)	11.1 (13.2)	11.9 (15.3)	11.6 (14.3)
RP-PT	11.6 (11.0)	12.3 (10.8)	11.4 (10.0)	11.8 (10.5)
SP-TW	3.5 (5.5)	2.8 (5.5)	2.5 (6.1)	2.7 (6.2)
RP-TW	23.5 (16.4)	30.8 (18.6)	29.6 (18.1)	34.0 (20.5)
FFD-TW	258 (39)	294 (61)	293 (53)	300 (58)
GD-TW	308 (64)	370 (96)	363 (85)	389 (95)
TRT-TW	331 (87)	414 (129)	394 (105)	430 (118)

Means (and standard deviations in parentheses) for skipping probability (SP) and refixation probability (RP) in percent for the pre-target (PT) and target words (TW), as well as first-fixation durations (FFD), gaze durations (GD) and total reading time (TRT) in ms for the target words. Values are computed across participant means.