Applied Psycholinguistics (2022), **43**, 389–410
doi:10.1017/S0142716421000564

CAMBRIDGE UNIVERSITY PRESS

ORIGINAL ARTICLE

Interactions of stimulus quality and semantic context on N400 in visual word recognition

Yong Zhang[1]*, Min Xie[2,3], Youguo Chen[2,3], Rongmin Xiong[1], Change Yue[1], Shuqiong Wu[4], Feng Ji[5] and Quanhong Wang[2,3]*

[1]School of Foreign Languages, Southwest University of Political Science and Law, Chongqing 401120, China, [2]Faculty of Psychology, Southwest University, Chongqing 400715, China, [3]Key Laboratory of Cognition and Personality (SWU), MoE, Chongqing 400715, China, [4]Center for Linguistic, Literary and Cultural Studies, Sichuan International Studies University, Chongqing 400031, China and [5]Chongqing Institute of Foreign Studies, Chongqing 401120, China
*Corresponding authors. Emails: zhangyong@swupl.edu.cn; quanhong177@yahoo.com

(Received 14 December 2020; revised 28 October 2021; accepted 8 November 2021; first published online 04 January 2022)

Abstract

The joint effects of stimulus quality and semantic context in visual word recognition were examined with event-related potential (ERP) recordings. In one-character Chinese word recognition, we manipulated stimulus quality at two degradation levels (highly vs. slightly degraded) and semantic context at two priming levels (semantically related vs. unrelated). In a prime–target–probe trial flow, ERPs were recorded to the target character which was presented in either high or slight degradation and which was preceded by either a semantically related or unrelated prime character. The target character was then followed by a probe character which was either identical to or different from the target character. Subjects were instructed to make target–probe matching judgments. The ERP results demonstrated a degradation by priming interaction, with larger N400 semantic priming effects for slightly degraded targets. Moreover, the degradation effects were observed on the P200, N250, and N400. These findings provided evidence for the cascaded model of visual word recognition such that the visual processing cascaded into the semantic stage and thus interacted on the N400 amplitude. The results were compared to an earlier study with a null ERP degradation by priming interaction. The ramifications of these results for models of visual word recognition are discussed.

Keywords: visual word recognition; stimulus quality; semantic context; cascaded model; N400

Reading is a fundamental human capacity. The mechanisms by which the human brain decodes single words have been of special interest to a broad cross section of researchers. Despite much progress that has ever been made, there remains a matter of active debate concerning whether visual word recognition involves discrete and sequentially organized modules or a highly interactive system (e.g.,Carreiras et al., 2014; Heilbron et al., 2020; Price & Devlin, 2011; Whaley et al., 2016).

© The Author(s), 2022. Published by Cambridge University Press

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press

The temporally modular feedforward models propose that visual word recognition is processed in discrete, hierarchically organized modules in a feedforward manner (Coltheart & Curtis, 1993; Jobard, Crivello & Tzourio-Mazoyer, 2003; Levy et al., 2008). Feedforward theorists assert that only when orthographic access is completed does other higher-level linguistic representations (e.g., phonology, morphology, semantics) begin to proceed (Davis, 2012). However, according to the interactive activation (IA) models, mental processes are cascaded in a way that all processes are operating continuously, with the current (partial) output of one process immediately available as an input for the upcoming process (Mcclelland, 1979). Mcclelland and Rumelhart (1981) later extended this formulation in which ongoing activation of a higher-level process feeds back to a lower-level process. Connectionists posit a highly interactive system in which word recognition is achieved with bidirectional connections between lower- and higher-level representations throughout the entire orthographic–phonological–lexical–semantic network (Seidenberg, 2010).

A central focus of empirical studies in visual word recognition literature is the nature of the processes in which two or more independent variables produce additive or interactive effects on response times (RTs) or event-related potentials (ERPs) (Borowsky & Besner, 1993; Holcomb, 1993; O'Malley, Reynolds & Besner, 2007; Scaltritti, Balota & Peressotti, 2013; van Petten & Kutas, 1990; Yap & Balota, 2007). By the additive-factors logic (Sternberg, 1969), additive effects of two factors suggest that these two factors influence discrete stages of processing independently; in contrast, interactive effects are taken as evidence that these two factors influence at least one stage of processing in common. Stimulus quality, word frequency, and semantic context are the most studied three psycholinguistic variables in visual word recognition literature and reflect processes at the orthographic, lexical, and semantic levels, respectively (e.g., Aschenbrenner & Yap, 2019; Stolz & Besner, 1998; Yap, Balota & Tan, 2013). The joint effects among those three variables are often taken as evidence which can shed light on the feedforward versus interactive debate concerning the underlying cognitive mechanism of visual word recognition.

Behavioral research has presented a conundrum: stimulus quality and word frequency yield additive effects, but both variables interact with semantic context. One account to resolve the additive–interactive discrepancy may be that the nature of the process in visual word recognition is task dependent (e.g., O'Malley et al., 2007; Scaltritti et al., 2013; Yap & Balota, 2007). For example, Yap and Balota (2007) found that additive effects of stimulus quality and word frequency were only observed in the lexical decision task (LDT) while interactive effects were found in both pronunciation and semantic categorization tasks. As such, addictive effects may reflect task-specific lexical decision operations rather than the underlying mechanism of word recognition system (O'Malley et al., 2007; Scaltritti et al., 2013; Yap & Balota, 2007).

Different from behavioral measures which reflect an end state of processing in the system as a whole, electroencephalography (EEG) taps the synchronized neuronal activity over time evoked by some cognitive event. With this inherent advantage, EEG allows us to delve into the internal temporal dynamics underlying the neural processes involved in visual word recognition (Carreiras et al., 2014).

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press

Thus, the introduction of this time-sensitive technique with a temporal resolution in milliseconds can provide direct temporal evidence regarding how information flows in the brain circuits.

The N150/170 and N250 are among two earlier markers of visual word recognition. The left lateralized N150/170 is sensitive to orthographic stimuli but not to other stimuli such as symbols (Duñabeitia et al., 2012). This early ERP component is significantly larger to mismatches between prime and target letter case (Petit et al., 2005). The N250 is sensitive to orthographic similarity (i.e., the degree of prime–target orthographic overlap), being larger for targets that partly overlap with their primes than targets that completely overlap with their primes. The N250 has a widespread scalp distribution over frontal scalp areas (Grainger & Holcomb, 2009; Holcomb & Grainger, 2006). Moreover, N250 is also modulated by lexical factors such as letter identity or the phonological status of the letters (Duñabeitia, Perea & Carreiras, 2009). Traditionally, the N400, a negative-going potential that peaks centroparietally over the scalp approximately 400 ms after stimulus onset, is a reflection of post-lexical aspects of semantic analysis (Kutas & Hillyard, 1980). N400 amplitudes are found to be larger for words that are unrelated to their primes. The N400 is also sensitive to orthographic and/or phonological analysis (Deacon et al., 2004). Additionally, the P200, which is basically associated with visual processes, is reported to be sensitive to sublexical orthographic or phonological processing in visual Chinese character processing with a positive wave distributed at the frontal-central sites (Hsu et al., 2009; Kong et al., 2012; Kong et al., 2010).

However, there are only a couple of ERP studies with conflicting findings which have examined the joint effects among stimulus quality, word frequency, and semantic context in visual word recognition (Holcomb, 1993; van Petten & Kutas, 1990; Zhang, Zhou & Wang, 2020). Actually, the effects of stimulus degradation on behavioral and ERP measures of semantic priming were earlier examined by Holcomb (1993). In a block of trials, the target words or pseudowords, which were preceded by either semantically related or unrelated prime words, were presented in an intact manner (intact block) or a degraded manner (degraded block). Subjects were asked to make speeded lexical decisions. Behaviorally, subjects made faster and more accurate responses to semantically related target words than to unrelated ones (behavioral semantic priming effect), and faster and more accurate responses to intact target words than to degraded ones (behavioral stimulus degradation effect). The interaction between semantic context and stimulus quality was observed with the priming effects being greater when the target was degraded. At the electrophysiological level, although semantically unrelated target words were found to elicit larger N400 than related ones (ERP semantic priming effect), no significant ERP difference between the intact and degraded blocks was observed (ERP stimulus degradation effect) in the time domain of the N150/170, P200, and N400 components. Moreover, there was no significant interaction between these two factors on any ERP components. This earlier ERP result seemed to run counter to the IA models.

In contrast, the interactive effects of two variables reported by a couple of ERP studies (van Petten & Kutas, 1990; Zhang et al., 2020) provide evidence in favor of the IA models. An earlier ERP study examined the effects between word frequency and sentence context (i.e., linear word position in the sentence) in a reading task

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press

(van Petten & Kutas, 1990), wherein frequency effects reflect lexical-level processing and word position effects index the increasing influence of context at the post-lexical level. An interaction between these two factors was observed, with N400 frequency effects being prominent for the sentence-initial words, attenuated for the sentence-intermediate words, and absent for the sentence-final words. Thus, word frequency seemed not to influence N400 amplitude for words presenting late in a sentence. Moreover, in a recent ERP study we observed a significant interaction between stimulus quality and word frequency on N400-like amplitude in visual Chinese character recognition (Zhang et al., 2020). Specifically, in a trial, the target stimulus, which was either a one-character Chinese word of high frequency or low frequency, was presented in either high or slight degradation. Then a probe character followed which was either identical to or different from the target one. Subjects were required to make target–probe matching judgments. The behavioral results showed a significant degradation by frequency interaction with the RTs being slower for the target stimuli in high degradation than for those in slight degradation in the low-frequency condition. The ERP results also demonstrated a significant degradation by frequency interaction on N400-like amplitude. Therefore, both behavioral and ERP results converged on the interaction of these two factors.

Scarce evidence as there may be, the conflicting ERP findings from the absence of a stimulus degradation by semantic priming interaction in Holcomb's study (1993) and from the presence of interactions between stimulus quality and word frequency in our recent study (Zhang et al., 2020) and between sentence context and word frequency in an earlier study (van Petten & Kutas, 1990) prompt a revisit to the joint effects of stimulus quality and semantic context on ERP components. This helps to adjudicate the theoretical debate between temporally modular feedforward processing and highly interactive processing during visual word recognition on one hand and to delineate the above-mentioned behavioral conundrum on the other.

The present ERP study aimed to determine whether the joint effects of stimulus quality and semantic context were additive or interactive in visual word recognition. Stimulus quality refers to whether a visual stimulus is presented in an intact or degraded manner. Intact words are recognized faster and more accurately than degraded words (Becker, 1979). Meanwhile, words preceded by semantically related primes (e.g., DOG–CAT) are recognized faster and more accurately than words preceded by unrelated primes (e.g., MOON–CAT), yielding semantic priming effects (Meyer, Schvaneveldt & Ruddy, 1975). Stimulus quality is often considered to influence an early pre-lexical encoding stage (Besner & Roberts, 2003; Holcomb, 1993; Reynolds & Besner, 2004), while semantic priming effects are taken as a marker of a later semantic processing. It is well documented in behavioral studies that there is an interaction between stimulus quality and semantic context on the RTs, with larger semantic priming effects for degraded targets than for intact ones (e.g., Balota et al., 2008; Becker & Killion, 1977; Besner & Smith, 1992; Borowsky & Besner, 1993; Brown & Besner, 2002; Meyer et al., 1975; Thomas, Neely & O'Connor, 2012). However, some studies indicated additive effects (i.e., no interaction) between these two factors (e.g., Aschenbrenner & Yap, 2019; Brown & Besner, 2002; Ferguson, Robidoux & Besner, 2009; Robidoux, Stolz & Besner, 2010; Stolz & Neely, 1995).

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press

Despite addressing the identical variables, the present study differed from Holcomb's study (1993) in two aspects: first, different levels of degradation were used in the present study (i.e., slightly vs. highly degraded) and Holcomb's study (i.e., intact vs. slightly degraded). In Holcomb's study (1993), null ERP stimulus degradation effect was observed, suggesting that there might be a ceiling effect for the slightly degraded words which can be accessed as easily as intact words. However, in our previous study using slightly and highly degraded stimuli (Zhang et al., 2020), ERP stimulus degradation effect was found. Second, different tasks were used in the present study (i.e., delayed character-matching task, DCMT) and Holcomb's study (i.e., LDT). Using a LDT may elicit a decision-related P300 that overlaps with N400 and thus contaminates N400 (Deacon et al., 2004). Moreover, using a DCMT, our previous study found an interaction between stimulus quality and word frequency on N400 (Zhang et al., 2020).

Subjects were instructed to make target–probe matching (i.e., identical or different) judgments of one-character Chinese words in a prime–target–probe trial flow. We manipulated the target character at two levels of visual degradation (high, slight) and at two levels of semantic priming (related, unrelated). For the manipulation of stimulus degradation, the present study manipulated stimulus quality of the target character at two (high vs. slight) degradation levels, identical to our previous study (Zhang et al., 2020). For the manipulation of semantic context, the target character was either semantically related or unrelated to the prime character. Previous ERP research on visual word recognition has defined a series of components (for a review, see Grainger & Holcomb, 2009). Among these ERP components, the N250 and the N400 are of specific interest for the present study because the N250 is sensitive to orthographic similarity (Carreiras et al., 2009) and the N400 is modulated by semantic processing (Kutas & Hillyard, 1980). Based on the ERP findings in relevant previous studies (van Petten & Kutas, 1990; Zhang et al., 2020), we predict that there would be an interactive effect between stimulus quality and semantic context on the N400 component. If the joint effect of stimulus quality (i.e., the orthographic level) and semantic context (i.e., the semantic level) is interactive, then it suggests that the lower-level representation continuously cascades into the higher-level semantic network, thus yielding an interaction with bidirectional connections between orthographic and semantic nodes. Moreover, we predict that there would be N250 degradation effects, on the basis of our previous ERP study which found that the character degradation elicited N250-like effects (Zhang et al., 2020). Behaviorally, we predict that there would be an interaction of stimulus quality and semantic context on the RTs, based on previous studies (Scaltritti et al., 2013; Yap et al., 2013; Yap & Balota, 2007).

Materials and methods

Subjects

Twenty right-handed students (11 females, nine males; mean age = 21.6, $SD = 1.9$) from a local university were paid to participate in this study. All the subjects reported to have normal or corrected-to-normal vision and were native Chinese

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press

speakers. A written informed consent was collected from each subject. The study protocol was approved by the Ethics Committee of Southwest University, China.

Design and materials

The experiment followed a 2 (priming: related, unrelated) × 2 (degradation: high, slight) within-subject design. All stimuli were one-character Chinese words. We firstly chose 400 pairs of semantically related words out of 800 words as candidate stimuli (Beijing Language Institute, 1986). An independent cohort of 20 undergraduates assessed the semantic relatedness of these word pairs with a 7-point scale. Based on the screening of semantic relatedness ($M = 5.35$ points, $SD = 1.37$), we finally selected 560 words that yielded 280 prime–target pairs of the most semantically related words as the experimental stimuli (see Appendix). For priority consideration of semantic relatedness, the prime–target character frequency defied a match. As such, the prime character group (for frequency: $M = 210.85$ per million, $SD = 546.81$; for stroke count: $M = 9.94$, $SD = 3.22$) and the target character group (for frequency: $M = 40.72$ per million, $SD = 13.02$; for stroke count: $M = 10.10$, $SD = 3.34$) differed in frequency ($t = 5.229$, $p < .001$) but matched in stroke count ($t = 0.411$, $p > .05$).

From the master list, four sublists with each 70 prime–target related pairs were formed. Prime–target unrelated pairs were formed by rearranging related primes and targets in each sublist so that there was no prime–target semantic association. A Latin-square counterbalance design was used such that across subjects each sublist occurred in the related and unrelated condition and in the highly and slightly degraded condition an equal number of times, but within subjects each sublist was presented once.

One hundred and forty additional characters (for frequency: $M = 45.03$ per million, $SD = 45.97$; for stroke count: $M = 10.15$, $SD = 3.06$) and 140 characters out of the 280 targets were selected to serve as probes in a prime–target–probe trial flow. The target–probe pairs were pseudo-randomly arranged in each of four experimental conditions (i.e., related vs. unrelated, highly vs. slightly degraded), with one half of trials being the target–probe different and the other half identical.

For stimulus degradation manipulation, the target stimuli were configured as follows. Each target character in black font ST (Song Ti) 39 pt was at the center of a 245 × 245 pixel white background in a Macromedia Flash 8.0 environment (Macromedia Company, San Francisco, CA, USA) and saved as a smoothed bitmap image. The image was then impoverished with the PicFrag program (Snodgrass & Hirshman, 1991) into a highly (level 5) or slightly (level 6) degraded stimulus, with the proportions of unmasked pixels for two levels being 34% and 49%, respectively, using the formula to calculate the unmasked ratio Ps (exposure ratios) for the degraded images [i.e., P (block pixel, level n) = r (8 –- level n), where $r = 0.7$, the block pixel is 16 × 16 when the level is equal to 5 and 6]. Finally, the image size was readjusted to 45 × 45 pixels and further reversed into a white character in a black background with Adobe Photoshop (8.0 version). Likewise, the prime stimuli in font ST 39 pt and the probe stimuli in font XKT (Xing Kai Ti) 39 pt were similarly configured but without being impoverished. The prime and target stimuli were always the same in font, whereas the target and probe stimuli differed in font such

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press

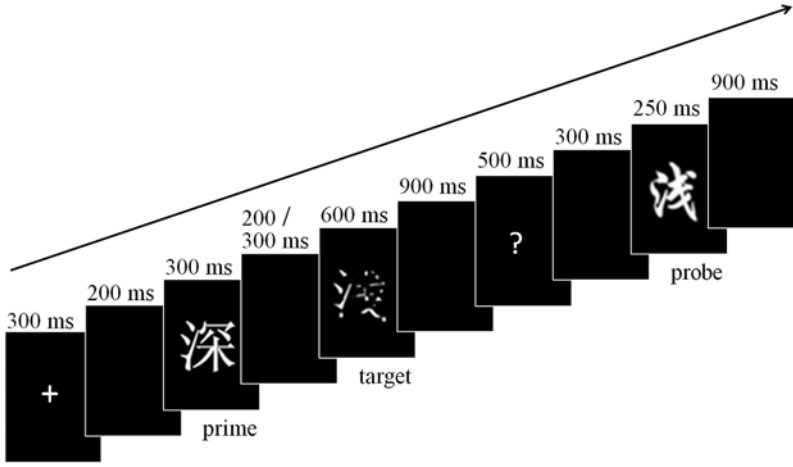


Figure 1. A prime–target–probe trial flow. The target character which is presented in either slight or high degradation is preceded by either a semantically related or unrelated prime character. The target character is then followed by a probe character which is either identical to or different from the target character. Subjects make target–probe matching judgments.

that subjects made target–probe matching on the basis of the characters' whole identities rather than their partial visual features. The size of each image was 1.43° both vertically and horizontally.

Procedure

Each subject completed four experimental blocks in which targets in related highly degraded (R-H), unrelated highly degraded (UR-H), related slightly degraded (R-S), or unrelated slightly degraded condition (UR-S) were randomly assigned to each block. A 16-trial practice block preceded the experiment using additional characters. There was a 3-min break between blocks.

Each trial followed a prime–target–probe flow (see Figure 1): a) a 300-ms cross fixation, followed by a 200-ms blank screen; b) a 300-ms prime character, followed by a blank screen for 200 ms or 300 ms at random; c) a target character in either high or slight degradation, which was either semantically related or unrelated to the preceding prime character, was presented for 600 ms, followed by a 900-ms blank screen; d) a 500-ms question mark, followed by a 300-ms blank screen; and e) a probe character, which was either identical to or different from the target, was presented for 250 ms, followed by a blank screen which remained on the screen for 900 ms or disappeared upon responses. Subjects were required to make target–probe matching (i.e., identical or different characters) judgments by button pressing as quickly and accurately as possible upon the probe presentation. The thumbs used for "yes (i.e., identical)" and "no (i.e., different)" responses were counterbalanced across subjects. Subjects were advised to blink only in the interval between the trials.

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press

ERP recording and data analysis

EEG was recorded from 64 scalp sites with tin electrodes mounted on an elastic cap (Brain Products GmbH, Germany). The EEG data were referenced to the left and right mastoids (i.e., averaged mastoid reference) and a ground electrode was placed on the medial frontal aspect (at AFz electrode). Moreover, vertical eye movements were recorded with electrodes placed on the supraorbital and infraorbital ridges of the right eye. Horizontal eye movements were measured with electrodes laterally placed on the outer canthi of both eyes. The EEG and electrooculogram (EOG) signals were amplified through a bandpass filter of 0.01–100 Hz and digitally sampled at 500 Hz per channel. The impedance of all electrodes was <5 kΩ.

ERPs were time-locked to the onset of the target, with the 200-ms pre-target period used as a baseline. The EEG pattern was segmented into epochs beginning at 200 ms before the target, which lasted for 800 ms after the target. For each condition with 70 trials, only epochs associated with correct responses were averaged. Excluded from the average were the trials with EOG artifacts (i.e., trials in which the EOG voltage exceeded ± 80 μV), amplifier clippings, bursts of electromyographic activity, or peak-to-peak deflections exceeding ± 80 μV, leaving the averaging numbers of the segments as 59.81 ($SD = 6.12$), 56.94 ($SD = 5.75$), 64.25 ($SD = 6.26$), and 63.38 ($SD = 6.07$) for R-H, R-S, UR-H, and UR-S conditions, respectively.

Results

Behavioral data

Linear mixed effects were performed for the analyses of RTs (including only correct responses) and accuracy, respectively. All analyses were carried out in R (v.3.1.2; R Core Team, 2018), using the lme4 (v.1.1-21; Bates et al., 2014) and lmerTest (v.3.1-0; Kuznetsova, Brockhoff & Christensen, 2017).

Response times

Response errors and RTs exceeding a threshold of three *SD*s by subject and by condition were excluded from analysis (11.5% of the data). RTs were entered as the dependent variable in a linear mixed model analysis (see Table 1). Degradation and priming were entered as fixed effects, while subject and item were entered as random effects. The analysis yielded a significant main effect of degradation (estimate = 29.51, $p < .001$), with subjects responding faster to slightly than to highly degraded characters (546 ms vs. 576 ms), and a significant main effect of priming (estimate = 15.68, $p < .001$), with subjects responding faster to related than to unrelated characters (553 ms vs. 569 ms). The interaction was not significant (estimate = 11.37, $p = .138$). Then, we conducted statistical power analyses for the interaction of stimulus degradation and semantic priming on the RTs using R software with pwr (https://CRAN.R-project.org/package=pwr). According to the equation, the effect size f was 0.316. We computed the ideal sample size (Cohen, 1988) as 40 given that the ideal statistical power for any study is considered to be 0.8 (Malone, Nicholl & Coyne, 2016) and the significance level is 0.05.

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press

Table 1. Results of the linear mixed models on mean RTs and response accuracy

	Fixed effects				Random effects	
					By subject	By item
	Estimate	*SE*	*t*-value	*P*	*SD*	*SD*
Response times (RT)						
Intercept	561.26	17.60	31.89	<.001	78.08	9.37
Degradation	29.51	3.83	7.70	<.001		
Priming	15.68	3.83	4.09	<.001		
Degradation × Priming	11.37	7.67	1.48	0.138		
Accuracy						
Intercept	0.91	0.012	75.68	<.001	0.05	0.02
Degradation	0.09	0.008	11.60	<.001		
Priming	0.03	0.008	3.24	0.001		
Degradation × Priming	0.02	0.015	1.55	0.121		

Note: All factors were coded using contrast coding as follows: degradation (-0.5 = slight, 0.5 = high), priming (−0.5 = related, 0.5 = unrelated). Model formula for response times: RT ~ 1 + priming + degradation + priming × degradation + (1 | subject) + (1 + | item). Model formula for accuracy: ACC ~ 1 + priming + degradation + priming × degradation + (1 | subject) + (1 + | item).

Response accuracy

To analyze the proportion of correct responses, we also conducted a linear mixed model analysis with degradation and priming entered as fixed effects and subject and item as random effects (see Table 1). The analysis yielded a significant main effect of degradation (estimate = 0.09, $p < .001$) with subjects responding more accurately to slightly than to highly degraded characters (94.9% vs. 93.4%) and a significant main effect of priming (estimate = 15.68, $p = .001$) with subjects responding more accurately to related than to unrelated characters (94.7% vs. 93.6%). The interaction was not significant (estimate = 0.02, $p = .121$). We also conducted statistical power analyses for the interaction of degradation and priming on the accuracy. The effect size *f* was 0.438 and the ideal sample size was 21 with 0.8 statistical power and at the 0.05 significance level.

ERP data

As shown in Figure 2, visual inspection of the grand mean waveforms and conventional windows for identifying ERP components in visual word recognition experiments (Holcomb & Grainger, 2006) revealed a positive peak in the 170–220 ms post-stimulus time window, a negative peak in the 220–280 ms time window, and a negative peak in the 300–500 ms time window. They were identified as P200, N250, and N400, respectively. Mean ERP amplitudes from each time window were computed for each participant for all four conditions. On the basis of previous ERP studies on visual word recognition and the scalp distributions

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press

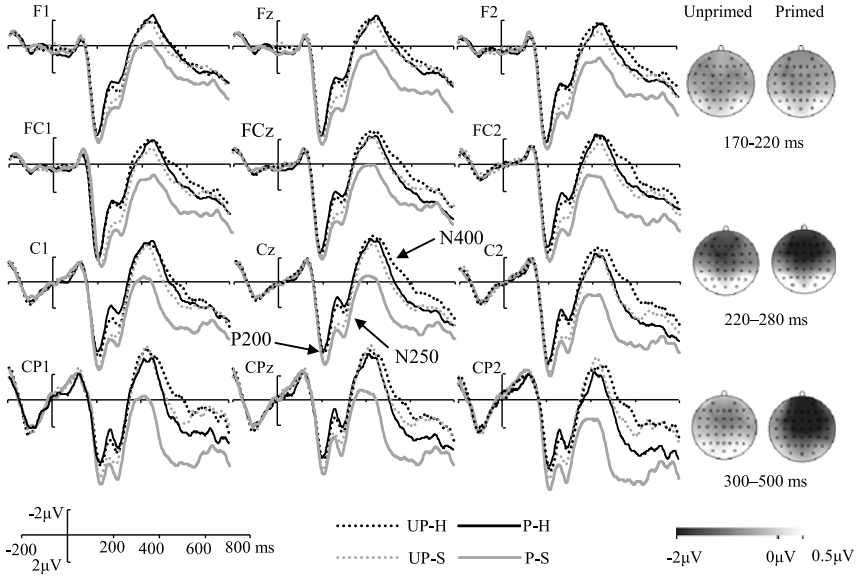


Figure 2. *Left*: Grand average ERP plots for unrelated highly degraded (UR-H), unrelated slightly degraded (UR-S), related highly degraded (R-H), and related slightly degraded (R-S) conditions at F1, Fz, F2, FC1, FCz, FC2, C1, Cz, C2, CP1, CPz, and CP2 electrodes. *Right*: Topographical maps of the voltage amplitudes for differences between slight and high degradation levels under unrelated and related conditions in the time windows of 170–220, 220–280, and 300–500 ms.

(P200: frontal-central; N250: frontal; N400: centroparietal) of the three components (Holcomb & Grainger, 2006; Wu et al., 2012; Zhang et al., 2020), we chose the following 12 electrode sites (i.e., F1, Fz, F2, FC1, FCz, FC2, C1, Cz, C2, CP1, CPz, CP2) to compute and analyzed the mean amplitudes of these components in the noted time windows. The measurement of the ERP amplitudes was conducted on data digitally filtered offline with a bandpass of 0.01–30 Hz. A three-way repeated-measures ANOVA with priming (unrelated vs. related), degradation (slight vs. high), and electrodes as within-subjects factors was conducted on mean amplitude. The degrees of freedom of the *f*-ratio were corrected according to the Greenhouse–Geisser method. Significant interactions were followed by pairwise comparisons, with Bonferroni correction.

170–220 ms (P200). The three-way repeated-measures ANOVA on mean amplitude with degradation, priming, and electrode site revealed only a significant main effect of degradation [F (1, 19) = 9.91, MSE = 11.01, $\eta 2 = 0.343$, $p = .005$], with higher P200 peaks for slightly than highly degraded targets regardless of priming conditions. Neither significant main effect of priming nor significant interaction between degradation and priming was observed [Fs < 1]. The statistical power analyses for the interaction showed the effect size f being 0.063 and the ideal sample size being 978 with 0.8 statistical power and at the 0.05 significance level. These findings suggested that there were only the degradation effects on the P200.

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press

220–280 ms (N250). The three-way repeated-measures ANOVA showed only a significant main effect of degradation [F (1, 19) = 38.27, MSE = 23.08, $\eta^2 = 0.668$, $p < .001$], with higher N250 amplitude for the highly than slightly degraded targets regardless of priming conditions. There was no significant main effect of priming [Fs < 1]. Moreover, there was no significant interaction between degradation and priming [F (1, 19) = 1.91, MSE = 0.99, $\eta^2 = 0.091$, $p = .183$]. As there was a trend-to-significance interaction, we conducted statistical power analyses for the interaction of degradation and priming on N250. The effect size f was 0.316 and the ideal sample size was 40 with 0.8 statistical power and at the 0.05 significance level. These findings suggested that there were only degradation effects.

300–500 ms (N400). The three-way repeated-measures ANOVA demonstrated significant main effects of degradation [F (1, 19) = 28.40, MSE = 389.34, $\eta^2 = 0.599$, $p < .001$] and priming [F (1, 19) = 55.39, MSE = 224.77, $\eta^2 = 0.745$, $p < .001$], with higher negativity for the highly than slightly degraded targets and for the unrelated than related targets. There was a significant interaction between degradation and priming [F (1, 19) = 12.26, MSE = 147.29, $\eta^2 = 0.392$, $p = .002$], suggesting that the lower-level visual representation cascades into the higher-level semantic network. The statistical power analyses for the interaction showed the effect size f being 0.803 and the ideal sample size being 7 with 0.8 statistical power and at the 0.05 significance level. Further analyses exhibited a significant priming effect on the slightly degraded targets [$p < .001$] but not on the highly degraded targets [Fs < 1], and a significant degradation effect on both related and unrelated targets [$p < .001$; $p = .040$]. These results suggest that there was a prominent priming effect only for the slightly degraded targets and that there were degradation effects regardless of the priming conditions. Moreover, a significant interaction between degradation and electrode site was observed [F (11, 209) = 4.29, MSE = 3.91, $\eta^2 = 0.184$, $p = .002$]. Further analyses were performed over each electrode and revealed larger N400 degradation effects at the centroparietal electrode sites (F1 [$p = .018$], Fz [$p = .026$], FCz [$p = .016$], FC1 [$p = .020$], FC1 [$p = .004$], Cz [$p = .001$], C1 [$p < .001$], C2 [$p = .004$], CPz [$p < .001$], CP1 [$p = .003$], CP2 [$p < .001$]. No other significant interactions were found.

Discussion

The present study investigated behavioral and neural correlates of the joint effects between stimulus quality and semantic context in visual word recognition. Subjects made target–probe matching decisions to one-character Chinese words which were presented in high or slight degradation and which were preceded by either semantically related or unrelated prime words. Analysis of ERPs revealed interactive effects of stimulus quality and semantic context on the N400 component. Moreover, we observed degradation effects on the P200, N250, and N400. However, at the behavioral level, there was no stimulus degradation by semantic priming interaction on RTs. We will now discuss the results in light of the theoretical and methodological considerations.

The ERP results demonstrated a stimulus degradation by semantic priming interaction, with more prominent N400 semantic priming effects for slightly than highly

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press

degraded targets. N400 is an index of semantic processing (Kutas & Hillyard, 1980). The N400 effects are interpreted in terms of post-lexical (semantic integration) versus pre-lexical (orthographic or phonological) processing (Deacon et al., 2004). Recently, Kutas and Federmeier (2011) assumed that lower-level and higher-level processes in word recognition intersect at the level of semantic access itself. In the present study, stimulus quality may be influencing an earlier pre-lexical encoding stage while semantic priming may be influencing a later semantic retrieval stage. Thus, the interaction between stimulus quality and semantic context on N400 component suggests that visual activations can be cascaded to semantic stages of processing, wherein the ongoing visual processing of degraded stimulus in the current stage begin at its early point in processing to affect the subsequent semantic stage of processing that modulates the N400 amplitude. Therefore, stimulus quality modulates the N400 amplitude and interacts with semantic priming on the N400 amplitude.

Our ERP result challenges strictly hierarchical feedforward views and aligns with the interactive cognitive architecture of visual word recognition. This finding is consistent with a couple of ERP evidence which favors the IA views of visual word recognition. Indeed, our previous ERP study in a target–probe matching task showed interactive effects between stimulus quality and word frequency, with larger N400 frequency effects for the target stimuli in high degradation than those in slight degradation (Zhang et al., 2020). Moreover, an earlier ERP study in a reading aloud task found an interaction between sentence context and word frequency on N400 (van Petten & Kutas, 1990).

The present electrophysiological evidence for the interaction between stimulus quality and semantic context on the N400 is theoretically relevant in the light of the broader debate on the extent to which visual word recognition is accessed in a hierarchical, feedforward manner or within an interactive framework. In the temporally modular feedforward models, word recognition processes occur in a hierarchical, feedforward fashion such that higher-level linguistic representations are not accessed until the lower-level orthographic access is completed (Coltheart & Curtis, 1993; Jobard et al., 2003; Levy et al., 2008). In contrast, the IA models propose that word recognition includes multiple levels of processing in a fully interactive network, with information flowing in both directions across the different levels (Mcclelland, 1979; Mcclelland & Rumelhart, 1981). For example, both the dual-route cascaded model (Coltheart et al., 2001) and the connectionist dual (Perry, Ziegler & Zorzi, 2007) process model implement cascaded activation within an IA framework and predict interactive effects between stimulus quality and semantic context. As predicted, in the present study we observed an interaction between a variable that affects the character/letter level (i.e., stimulus degradation) and a variable that affects higher-level semantic processing (i.e., semantic priming).

There is ample evidence from behavioral, electrophysiological (EEG), magneto-encephalographic (MEG), electrocorticographic (ECoG), and functional magnetic resonance imaging (fMRI) studies in favor of interactive framework as the more biologically plausible explanations for visual word processing. Behaviorally, higher-level linguistic representations, such as phonological (Rastle & Brysbaert, 2006), morphological (Deutsch et al., 2005), lexical (Duñabeitia et al., 2009), and semantic information (Frost, 2012), have been observed to influence the lower-level

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press

early pre-lexical encoding stage of visual word recognition. EEG studies also found that higher-level lexical–phonological–semantic information constrains orthographic form-level processing in the N250 (Carreiras et al., 2009; Carreiras, Vergara & Perea, 2007; Duñabeitia et al., 2009; Holcomb & Grainger, 2007). Notably, MEG (Woodhead et al., 2014), ECoG (Whaley et al., 2016), and fMRI (Heilbron et al., 2020) findings suggest the interaction between different levels of linguistic processing and thus challenge strictly hierarchical, feedforward views of visual word recognition. Taken together, the multi-disciplinary evidence is aligned with the interactive account of visual word recognition.

Though the degradation by priming interaction on the N400 amplitude in the present study seems to provide unambiguous evidence against the temporally modular feedforward models, we are here parsimonious to make an alleged assertion in favor of the IA model. Notably, in contrast to the ERP degradation by priming interaction found in the present study and the ERP degradation by frequency interaction observed in a previous study (Zhang et al., 2020), Holcomb (1993) did not observe the degradation by priming interaction on N400, thus presenting a particular challenge to the IA models.

One salient possibility for the ERP discrepancy may result from the difference in the manipulation of task demands (i.e., LDT vs. DCMT) between Holcomb's study (1993) and the present study. Actually, behavioral research has shown that the additive effects between two variables in visual word recognition are specific to LDT, merely reflecting task-specific operations (O'Malley et al., 2007; Scaltritti et al., 2013; Warrington et al., 2018; Yap & Balota, 2007; Zhang et al., 2020). Moreover, a few recent EEG and MEG studies examined the task effects using different task demands and found task-dependent effects (Chen et al., 2015; Faísca, Reis & Araújo, 2019; Wang & Maurer, 2017; Wang & Maurer, 2020; Yang et al., 2012). Therefore, the nature of the process in visual word recognition is likely to be modulated by the task context. Notably, both spatial and temporal evidence has pointed to that there is task-induced top-down modulation merely on early stage of visual-orthographic processing within 200 ms (Chen et al., 2015; Faísca et al., 2019; Wang & Maurer, 2020), which suggests that the N400 amplitude can reflect processes that are independent of task demands.

An alternative possibility may be attributed to language-specific attributes (Zhang et al., 2020). The semantic access of a visual word in alphabetic systems involves phonological mediation (Rayner et al., 2010), which possibly retards cascaded visual activation across subsequent stages to the stage of semantic processing. In contrast, Chinese visual processing can be sufficiently or rapidly cascaded to semantic stage of processing because there is direct semantic access to Chinese characters from orthography with little phonological mediation (Zhou & Marslen-Wilson, 1999).

Additionally, a third account may result from the manipulation of stimulus degradation levels. Specifically, in Holcomb's study (1993), the slightly degraded and intact targets showed no differences on any ERP components. No differences even at the early visual stage of processing suggest a ceiling effect: the visual processing of the slightly degraded targets ended rapidly in the LDT before it cascaded into semantic processing, as did the processing of intact targets. In contrast, we observed degradation effects on P200, N250, and N400, suggesting that subjects

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press

in the present study and a previous study (Zhang et al., 2020) would involve more effortful semantic access to the highly (as compared to slightly) degraded targets in an attempt to match the following probes. As a result, the visual processing could cascade into the semantic stage of processing and thus interacted on the N400 amplitude.

A reconciliation with the additive–interactive debate is to assume a remarkably adaptive lexical processor (Aschenbrenner & Yap, 2019). This lexical processor can be dynamically fine-tuned to psycholinguistic variables, yielding additive or interactive effects. Indeed, converging behavioral evidence suggests this to be the case. The extent to which processing levels in word recognition are additive or interactive may be affected by the experimental task demands (Yap & Balota, 2007), proportion of related primes (Aschenbrenner & Yap, 2019; Ferguson et al., 2009), presence or absence of nonwords (Masson & Kliegl, 2013; O'Malley & Besner, 2008), relative prime–target associative strength (Masson & Kliegl, 2013; Robidoux et al., 2010), and semantic association direction (Thomas et al., 2012). Additionally, variance across orthographies (Frost, 2012; Zhang et al., 2020) and stimulus degradation levels (Holcomb, 1993; Zhang et al., 2020) may also modulate visual word recognition.

Higher N250 amplitude for highly degraded targets than slightly degraded targets was observed in the present study. This negativity is thought to be involved in processes at the sublexical–lexical interface. The N250 is sensitive to the degree of prime–target orthographic overlap, being more negativity to targets that are unrelated to primes than those that are repeated primes (Grainger & Holcomb, 2009; Holcomb & Grainger, 2006). Wang et al. (2004) also found a N250-like component and explained that this component reflects the brain activity for detecting cognitive information mismatches or conflicts. In the same vein of this mismatch account, it is the degradation levels that modulate the N250 in the present study: the more the degraded stimuli, the higher the amplitude.

In addition, we observed higher P200 effects on slightly degraded targets than highly degraded targets. This positivity is thought to be an earlier marker in Chinese character recognition and reflects orthographic and phonological processing (Kong et al., 2012; Kong et al., 2010; Liu, Perfetti & Hart, 2003). For example, Kong et al. (2012) found that controls elicited higher P200 effects than visually similar characters. Interestingly, in contrast to logographic characters, the P200 has not been reported to involve lexical processing of visual word recognition in alphabetic scripts. However, the P200 modulation at the sublexical processing seems to be language general. For example, syllable frequency has been found to modulate the P200 component in studies on visual recognition in European languages (Chetail, Colin & Content, 2012; Kwon, Lee & Nam, 2011).

Behaviorally, we did not obtain a RT degradation by priming interaction, demonstrating a dissociation between RT and N400 amplitude. Interestingly, Holcomb (1993) observed a significant degradation by priming interaction with RTs but not with ERPs and explained the dissociation between behavioral and ERP measures in terms of different sets of cognitive operations involved. In our view, this apparent discrepancy between the RT and ERP results may stem from a couple of aspects. Firstly, the centroparietal N400 reflects processes that the orthographic information cascades into the semantic stage of processing, thus yielding an interaction between

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press

two levels of representation, whereas the RT data reflect the net result or end state of having recognized characters. Alternatively, we may attribute the absence of interaction with RTs more to the use of delayed target–probe matching task. Unlike the ERP measures which were time-locked to the target onset, the behavioral responses occurred after the probe presentation. Therefore, the behavioral results merely showed indirect RTs of processing the targets and thus should be interpreted with caution. Additionally, the absence of degradation by priming interaction on RT in the present study may be attributed to small sample size. The power analysis for the interaction of stimulus quality and semantic priming on the RT in the present study showed that the ideal sample size is 40. This is also true for the interaction of stimulus quality and semantic priming on N250. Therefore, the absence of interaction on the RTs or ERP components in the present study may result from a small sample size. However, with approximately 60 valid trials for each experimental condition, we obtained significant ERP results from 20 subjects in the interaction of stimulus quality and semantic priming on N400.

Some limitations need to be considered. First, as a few ERP studies found task-dependent effects using different task demands (Chen et al., 2015; Faísca et al., 2019; Wang & Maurer, 2017; Wang & Maurer, 2020; Yang et al., 2012), the present study used only a single task. As such, future ERP research is needed using both LDT and DCMT and making a direct comparison of effects from different task demands. Second, language-specific properties and degradation levels are needed to be controlled if we want to justify the interaction of stimulus degradation and semantic priming on the N400. Third, the difference in character frequency between prime and target stimuli may be a confounding factor that needed to be considered.

In summary, the interactive effects of stimulus quality and semantic context on N400 in the present study suggest that the impact of stimulus quality is not constrained to the pre-lexical processing stage. Rather, the direct impact of stimulus degradation can be further cascaded to the semantic processing stage, such that the stimulus quality and semantic context may influence at least a common processing stage at the semantic level, before all the processing is completed. However, the nature of the process in which stimulus quality and semantic priming produce joint effects in visual word recognition may be dynamically adaptive, depending on the task (Yap & Balota, 2007) or stimulus contexts (e.g., Aschenbrenner & Yap, 2019; Ferguson et al., 2009; Masson & Kliegl, 2013; O'Malley & Besner, 2008; Robidoux et al., 2010; Thomas et al., 2012). Despite much behavioral evidence, there has been little to no ERP evidence that can be used to justify this assumption about the adaptive lexical processor. As the adaptive modulation of the information flow for visual word recognition is mostly temporal in nature, the time-sensitive EEG technique can shed light on the temporal order of neural processes. More future ERP work is needed to depict the underlying mechanism of adaptive word recognition system involved in visual word recognition.

Declaration of competing interest. No potential conflict of interest was reported by the authors.

Acknowledgements. This study was funded by the National Social Science Foundation of China (21XYY001 and 18XYY003), the Humanities and Social Science Foundation of Chongqing Municipal Education Committee (YYK20316 and 19SKGH002).

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press

References

Aschenbrenner, A. J., & Yap, M. J. (2019). The influence of relatedness proportion on the joint relationship among word frequency, stimulus quality, and semantic priming in the lexical decision task. *Quarterly Journal of Experimental Psychology*, **72**(10), 2452–2461. doi: 10.1177/1747021819845317.

Balota, D. A., Yap, M. J., Cortese, M. J., & Watson, J. M. (2008). Beyond mean response latency: Response time distributional analyses of semantic priming. *Journal of Memory and Language*, **59**(4), 495–523. doi: 10.1016/j.jml.2007.10.004.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting Linear Mixed-Effects Models Using lme4. *Statistics & Computing,* arXiv:1406(1), 133–199. doi: 10.1007/0-387-22747-4_4.

Becker, C. A. (1979). Semantic context and word frequency effects in visual word recognition. *Journal of Experimental Psychology-human Perception and Performance*, **5**(5), 252–259. doi: 10.1037//0096-1523.5.2.252

Becker, C. A., & Killion, T. H. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology-human Perception and Performance*, **3**(3), 389–401. doi: 10.1037/0096-1523.3.3.389.

Beijing Language Institute (1986). Beijing Language Institute Press.

Besner, D., & Roberts, M. A. (2003). Reading nonwords aloud: results requiring change in the dual route cascaded model. *Psychonomic Bulletin & Review*, **10**(2), 398–404. doi: 10.3758/BF03196498.

Besner, D., & Smith, M. C. (1992). Models of visual word recognition: When obscuring the stimulus yields a clearer view. *Journal of Experimental Psychology-Learning Memory and Cognition*, **18**(3), 468–482. doi: 10.1037/0278-7393.18.3.468.

Borowsky, R., & Besner, D. (1993). Visual word recognition: a multistage activation model. *Journal of Experimental Psychology-Learning Memory and Cognition*, **19**(19), 813–840. doi: 10.1037/0278-7393.19.4.813.

Brown, M., & Besner, D. (2002). Semantic priming: On the role of awareness in visual word recognition in the absence of an expectancy. *Consciousness and Cognition*, **11**(3), 402–422. doi: 10.1016/S1053-8100(02)00008-9.

Carreiras, M., Armstrong, B. C., Perea, M., & Frost, R. (2014). The what, when, where, and how of visual word recognition. *Trends in Cognitive Sciences*, **18**(2), 90–98. doi: 10.1016/j.tics.2013.11.005.

Carreiras, M., Gillon-Dowens, M., Vergara, M., & Perea, M. (2009). Are vowels and consonants processed differently? event-related potential evidence with a delayed letter paradigm. *Journal of Cognitive Neuroscice*, **21**(2), 275–288. doi: 10.1162/jocn.2008.21023

Carreiras, M., Vergara, M., & Perea, M. (2007). ERP correlates of transposed-letter similarity effects: are consonants processed differently from vowels? *Neuroscience Letters*, **419**(3), 219–224. doi: 10.1016/j.neulet.2007.04.053.

Chen, Y., Davis, M. H., Pulvermüller, F., & Hauk, O. (2015). Early visual word processing is flexible: Evidence from spatiotemporal brain dynamics. *Journal of Cognitive Neuroscience*, **27**(9), 1738–1751. doi: 10.1162/jocn_a_00815.

Chetail, F., Colin, C., & Content, A. (2012). Electrophysiological markers of syllable frequency during written word recognition in French. *Neuropsychologia*, **50**(14), 3429–3439. doi: 10.1016/j.neuropsychologia.2012.09.044.

Cohen, J. (1988). *Statistical power ANALYSIS for the Behavioral sciences.* Lawrence Erlbaum Associates.

Coltheart, M., & Curtis, B. (1993). Models of reading aloud: dual-route and parallel distributed processing approaches. *Psychological Review*, **100**(4), 589–608. doi: 10.1037/0033-295X.100.4.589.

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**(1), 204–256. doi: 10.1037/0033-295X.108.1.204.

Davis, C. J. (2012). Developing a universal model of reading necessitates cracking the orthographic code. *Behavioral & Brain Sciences*, **35**(5), 283–284. doi: 10.1017/S0140525X12000039.

Deacon, D., Dynowska, A., Ritter, W., & Grose-Fifer, J. (2004). Repetition and semantic priming of nonwords: Implications for theories of N400 and word recognition. *Psychophysiology*, **41**(1), 60–74. doi: 10.1111/1469-8986.00120.

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press

Deutsch, A., Frost, R., Pollatsek, A., & Rayner, K. (2005). Morphological parafoveal preview benefit effects in reading: Evidence from Hebrew. *Language & Cognitive Processes*, **20**(1–2), 341–371. doi: 10.1080/01690960444000115.

Duñabeitia, J., Dimitropoulou, M., Grainger, J., & Hernández, J. (2012). Differential sensitivity of letters, numbers, and symbols to character transpositions. *Journal of Cognitive Neuroscience*, **24**(7), 1610–1624. doi: 10.1162/jocn_a_00180

Duñabeitia, J., Perea, M., & Carreiras, M. (2009). There is no clam with coats in the calm coast : Delimiting the transposed-letter priming effect *Quarterly Journal of Experimental Psychology*, **62**(10), 1930–1947. doi: 10.1080/17470210802696070.

Faísca, L., Reis, A., & Araújo, S. (2019). Early brain sensitivity to word frequency and lexicality during reading aloud and implicit reading. *Frontiers in Psychology*, **10**, 1–13. doi: 10.3389/fpsyg.2019.00830.

Ferguson, R., Robidoux, S., & Besner, D. (2009). Reading aloud: Evidence for contextual control over lexical activation. *Journal of Experimental Psychology-human Perception and Performance*, **35**(2), 499–507. doi: 10.1037/a0013162.

Frost, R. (2012). A universal approach to modeling visual word recognition and reading: Not only possible, but also inevitable. *Behavioral & Brain Sciences*, **35**(5), 263–329. doi: 10.1017/S0140525X12000635.

Grainger, J., & Holcomb, P. J. (2009). Watching the word go by: On the time-course of component processes in visual word recognition. *Language and Linguistics Compass*, **3**(1), 128–156.

Heilbron, M., Richter, D., Ekman, M., Hagoort, P., & Lange, F. P. D. (2020). Word contexts enhance the neural representation of individual letters in early visual cortex. *Nature Communications*, **11**(1), 321–331. doi: 10.1038/s41467-019-13996-4.

Holcomb, P. J. (1993). Semantic priming and stimulus degradation: implications for the role of the N400 in language processing. *Psychophysiology*, **30**(1), 47–61. doi: 10.1111/j.1469-8986.1993.tb03204.x

Holcomb, P. J., & Grainger, J. (2006). On the time course of visual word recognition: an event-related potential investigation using masked repetition priming. *Journal of Cognitive Neuroscience*, **18**(10), 1631–1643. doi: 10.1162/jocn.2006.18.10.1631

Holcomb, P. J., & Grainger, J. (2007). Exploring the temporal dynamics of visual word recognition in the masked repetition priming paradigm using event-related potentials. *Brain Research*, **1180**(8), 39–58. doi: 10.1016/j.brainres.2007.06.110.

Hsu, C.-H., Tsai, J.-L., Lee, C.-Y., & Tzeng, O. J.-L. (2009). Orthographic combinability and phonological consistency effects in reading Chinese phonograms: an event-related potential study. *Brain and Language*, **108**(1), 56–66. doi: 10.1016/j.bandl.2008.09.002.

Jobard, G., Crivello, F., & Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: a metaanalysis of 35 neuroimaging studies. *Neuroimage*, **20**(2), 693–712. doi: 10.1016/S1053-8119(03)00343-4

Kong, L., Zhang, B., Zhang, J. X., & Kang, C. (2012). P200 can be modulated by orthography alone in reading Chinese words. *Neuroscience Letters*, **529**(2), 161–165. doi: 10.1016/j.neulet.2012.09.028.

Kong, L., Zhang, J. X., Kang, C., Du, Y., Zhang, B., & Wang, S. (2010). P200 and phonological processing in Chinese word recognition. *Neuroscience Letters*, **473**(1), 37–41. doi: 10.1016/j.neulet.2010.02.014.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology*, **62**(1), 621–647. doi: 10.1146/annurev.psych.093008.131123.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, **207**(4427), 203–205. doi: 10.1126/science.7350657.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, **82**, 1–26. doi: 10.18637/jss.v082.i13.

Kwon, Y., Lee, Y., & Nam, K. (2011). The different P200 effects of phonological and orthographic syllable frequency in visual word recognition in Korean. *Neuroscience Letters*, **501**(2), 117–121. doi: 10.1016/j.neulet.2011.06.060.

Levy, J., Pernet, C., Treserras, S., Boulanouar, K., Berry, I., Aubry, F., Demonet, J. F., & Celsis, P. (2008). Piecemeal recruitment of left-lateralized brain areas during reading: A spatio-functional account. *Neuroimage*, **43**(3), 581–591. doi: 10.1016/j.neuroimage.2008.08.008.

Liu, Y., Perfetti, C. A., & Hart, L. (2003). ERP evidence for the time course of graphic, phonological, and semantic information in Chinese meaning and pronunciation decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **29**(6), 1231–1247. doi: 10.1037/0278-7393.29.6.1231.

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press

Malone, H. E., Nicholl, H., & Coyne, I. (2016). Fundamentals of estimating sample size. *Nurse Research*, **23**(5), 21–25. doi: 10.7748/nr.23.5.21.s5.

Masson, M. E. J., & Kliegl, R. (2013). Modulation of additive and interactive effects in lexical decision by trial history. *Journal of Experimental Psychology-Learning Memory and Cognition*, **39**(3), 898–914. doi: 10.1037/a0029180.

Mcclelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, **86**(4), 287–330. doi: 10.1.1.228.2731.

Mcclelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception (Part I). *Psychological Review*, **88**(5), 375–407. doi: 10.1.1.298.4480.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1975). Loci of Contextual Effects on Visual Word Recognition. In P. Rabbitt (Ed.), *Attention & Performance*. London, England: Academic Press.

O'Malley, S., & Besner, D. (2008). Reading aloud: qualitative differences in the relation between stimulus quality and word frequency as a function of context. *Journal of Experimental Psychology-Learning Memory and Cognition*, **34**(6), 1400–1411. doi: 10.1037/a0013084.

O'Malley, S., Reynolds, M. G., & Besner, D. (2007). Qualitative differences between the joint effects of stimulus quality and word frequency in reading aloud and lexical decision: extensions to Yap and Balota (2007). *Journal of Experimental Psychology-Learning Memory and Cognition*, **33**(2), 451–458. doi: 10.1037/0278-7393.33.2.451.

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: the CDP+ model of reading aloud. *Psychological Review*, **114**(2), 273–315. doi: 10.1037/0033-295X.114.2.273.

Petit, J., Grainger, J., Midgley, K., & Holcomb, P. (2005). On the time-course of processing in letter perception: A masked priming ERP investigation. *Psychonomic Bulletin and Review*, (13), 674–681. doi: 10.3758/BF03193980.

Price, C. J., & Devlin, J. T. (2011). The Interactive Account of ventral occipitotemporal contributions to reading. *Trends in Cognitive Sciences*, **15**(6), 246–253. doi: 10.1016/j.tics.2011.04.001.

R Core Team (2018). R: A Language and Environment for Statistical Computing. doi:10.1890/0012-9658(2002)083[3097:CFHIWS]2.0.CO;2.

Rastle, K., & Brysbaert, M. (2006). Masked phonological priming effects in English: Are they real? Do they matter? *Cognitive Psychology*, **53**(2), 97–145. doi: 10.1016/j.cogpsych.2006.01.002.

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2010). How Psychological Science Informs the Teaching of Reading. *Psychological Science in the Public Interest*, **2**(2), 31–74. doi: 10.1111/1529-1006.00004.

Reynolds, M., & Besner, D. (2004). Neighbourhood density, word frequency, and spelling-sound regularity effects in naming: similarities and differences between skilled readers and the Dual Route Cascaded Computational model. *Canadian Journal of Experimental Psychology*, **58**(1), 13–31. doi: 10.1037/h0087437.

Robidoux, S., Stolz, J., & Besner, D. (2010). Visual word recognition: Evidence for global and local control over semantic feedback. *Journal of Experimental Psychology-human Perception and Performance*, **36**(3), 689–703. doi: 10.1037/a0018741.

Scaltritti, M., Balota, D. A., & Peressotti, F. (2013). Exploring the additive effects of stimulus quality and word frequency: The influence of local and list-wide prime relatedness. *Quarterly Journal of Experimental Psychology*, **66**(1), 91–107. doi: 10.1080/17470218.2012.698628.

Seidenberg, M. S. (2010). Connectionist Models of Word Reading. *Current Directions in Psychological Science*, **14**(5), 238–242. doi: 10.1111/j.0963-7214.2005.00372.x

Snodgrass, J. G., & Hirshman, E. (1991). Theoretical explorations of the Bruner-Potter 1964 interference effect. *Journal of Memory and Language*, **30**(3), 273–293. doi: 10.1016/0749-596X(91)90037-K.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica*, **30**(69), 276–315. doi: 10.1016/0001-6918(69)90055-9.

Stolz, J. A., & Besner, D. (1998). Levels of representation in visual word recognition: A dissociation between morphological and semantic processing. *Journal of Experimental Psychology-human Perception and Performance*, **24**(6), 1642–1655. doi: 10.1037/0096-1523.24.6.1642.

Stolz, J. A., & Neely, J. H. (1995). When target degradation does and does not enhance semantic context effects in word recognition. *Journal of Experimental Psychology-Learning Memory and Cognition*, **21**(3), 596–611. doi: 10.1037/0278-7393.21.3.596.

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press

Thomas, M. A., Neely, J. H., & O'Connor, P. (2012). When word identification gets tough, retrospective semantic processing comes to the rescue. *Journal of Memory and Language*, **66**(4), 623–643. doi: 10.1016/j.jml.2012.02.002.

van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory and Cognition*, **18**(4), 380–393. doi: 10.3758/BF03197127.

Wang, F., & Maurer, U. (2017). Top-down modulation of early print-tuned neural activity in reading. *Neuropsychologia*, **102**, 29–38. doi: 10.1016/j.neuropsychologia.2017.05.028.

Wang, F., & Maurer, U. (2020). Interaction of top-down category-level expectation and bottom-up sensory input in early stages of visual-orthographic processing. *Neuropsychologia*, **137**, 107299. doi: 10.1016/j.neuropsychologia.2019.107299.

Wang, Y., Cui, L., Wang, H., Tian, S., & Zhang, X. (2004). The sequential processing of visual feature conjunction mismatches in the human brain. *Psychophysiology*, **41**(1), 21–29. doi: 10.1111/j.1469-8986.2003.00134.x.

Warrington, K. L., Mcgowan, V. A., Paterson, K. B., & White, S. J. (2018). Effects of aging, word frequency, and text stimulus quality on reading across the adult lifespan: Evidence from eye movements. *Journal of Experimental Psychology-Learning Memory and Cognition*, **44**(11), 1714–1729. doi: 10.1037/xlm0000543.

Whaley, M. L., Kadipasaoglu, C. M., Cox, S. J., & Tandon, N. (2016). Modulation of Orthographic Decoding by Frontal Cortex. *The Journal of Neuroscience*, **36**(4), 1173–1184. doi: 10.1523/JNEUROSCI.2985-15.2016.

Woodhead, Z. V. J., Barnes, G. R., Penny, W., Moran, R., Teki, S., Price, C. J., & Leff, A. P. (2014). Reading Front to Back: MEG Evidence for Early Feedback Effects During Word Recognition. *Cerebral Cortex*, **24**(3), 817–827. doi: 10.1093/cercor/bhs365.

Wu, Y., Mo, D., Tsang, Y.-K., & Chen, H.-C. (2012). ERPs reveal sub-lexical processing in Chinese character recognition. *Neuroscience Letters*, **514**(2), 164–168. doi: 10.1016/j.neulet.2012.02.080.

Yang, J., Wang, X., Shu, H., & Zevin, J. D. (2012). Task by stimulus interactions in brain responses during Chinese character processing. *Neuroimage*, **60**(2), 979–990. doi: 10.1016/j.neuroimage.2012.01.036.

Yap, M. J., & Balota, D. A. (2007). Additive and interactive effects on response time distributions in visual word recognition. *Journal of Experimental Psychology-Learning Memory and Cognition*, **33**(2), 274–296. doi: 10.1037/0278-7393.33.2.274

Yap, M. J., Balota, D., & Tan, S. E. (2013). Additive and interactive effects in semantic priming: Isolating lexical and decision processes in the lexical decision task. *Journal of Experimental Psychology-Learning Memory and Cognition*, **39**(1), 140–158. doi: 10.1037/a0028520.

Zhang, Y., Zhou, M., & Wang, Q. (2020). Interactions of stimulus quality and frequency on N400 in Chinese character recognition: Evidence for cascaded processing. *Neuroscience Letters*, **715**. doi: 10.1016/j.neulet.2019.134614.

Zhou, X., & Marslen-Wilson, W. (1999). Phonology, orthography, and semantic activation in reading Chinese. *Journal of Memory and Language*, **41**(4), 579–606. doi: 10.1006/jmla.1999.2663.

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press

Appendix

280 prime–target pairs of semantically related one-character Chinese words with the first character as the prime and the second as the target

Sublist 1	Sublist 2	Sublist 3	Sublist 4
深deep-浅low	始begin-末end	喉larynx-嗓throat	禽birds-畜cattle
聘employ-雇hire	涝flood-旱drought	奇odd-偶even	稠thick-稀thin
孝filial-忠loyal	棒club-棍stick	锅pot-灶oven	弩crossbow-弓bow
牧herd-渔fishing	沫foam-泡froth	孩child-娃baby	香fragrance-臭foul
清clear-浑turbid	顷100 mu-亩mu	膘fat-胖belly	素quiet-艳vivid
聚gather-凑gain	你you-俺i	冠first-亚second	抢rob-掠harry
落fall-涨rise	瓦tile-砖brick	苗sprout-芽bud	瞬moment-秒second
捡gather-拾pickup	囚prison-狱jail	升rise-跌fall	跳jump-蹦hop
建build-拆remove	广wide-狭narrow	饵bait-钓fish	娶wed-嫁marry
略omit-详detail	踹tread-踢kick	菊mum-梅plum	靶target-箭arrow
逗amuse-哄coax	好good-赖bad	娘mom-爹dad	柳osier-杨alamo
鞋shoe-袜sock	妖monster-仙god	凹concave-凸raised	抬raise-扛carry
瘪shrink-膨swell	咸corned-盐salty	聪smart-智wise	桔orange-苹apple
瞒hide-骗cheat	崖cliff-峡gorge	舌tongue-咽throat	名name-氏surname
字letter-词word	肝liver-肺lung	瞎blind-盲blind	甲first-丙third
痰phlegm-咳bex	渴dry-饮drink	匙key-锁lock	吵quarrel-嚷yell
爷grandpa-孙grandson	鲨shark-鲸whale	炸fry-烤bake	擒catch-纵release
碧jade-翠green	蕊stamen-瓣petal	麦wheat-稻rice	糖sugar-蜜honey
头head-脖neck	揭uncover-掀open	雪snow-霜frost	衫skirt-裤pants
丝silk-绒velvet	猩orang-猴monkey	俯overlook-仰look up	洒shed-泼spill
褂gown-衫shirt	妈mom-姨aunt	逸easy-悠leisure	褥bedding-炕kang
仆servant-奴slave	拦bar-挡block	友friend-戚kin	热hot-烫scalding
醒sober-醉drunk	腐rot-霉mildew	暮dusk-旦dawn	唇lip-齿tooth
孕pregnancy-胎fetus	晨morning-夕sunset	帕kerchief-巾towel	书book-页page
医doctor-诊diagnosis	掣tug-驰gallop	毫milli-厘centi	躯trunk-肢limb
瘸cripple-拐limp	墓grave-碑tomb	虫bug-蚁ant	枝twig-梢tip
闭close-敞open	鼠mouse-猫cat	倚rely-傍against	棺coffin-坟tomb
罚punish-奖award	鹊magpie-鸦crow	衣cloth-扣button	挚sincere-恳cordial
倦tired-乏weary	优good-劣bad	刚solid-柔soft	骨bone-筋vein
廉honest-贪venal	腿leg-臂arm	封seal-启open	鸭duck-鹅goose
偷steal-贼thief	蟹crab-虾shrimp	卖sell-购buy	削cut-刮scrape

(*Continued*)

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press

(*Continued*)

Sublist 1	Sublist 2	Sublist 3	Sublist 4
坐sit-蹲squat	脐navel-腹belly	累tire-疲fatigue	铁iron-铝aluminum
蚊midge-蝇fly	箱case-盒box	通open-堵block	扬raise-抑curb
粟millet-粱durra	虎tiger-狼wolf	福bliss-祸ruin	蚪tadpole-蛙frog
肚tripe-肠gut	赚earn-赔pay	柱pillar-桩stake	抓grab-捞gain
吉lucky-凶ominous	剁cut-砍chop	扯pull-撕tear	吐spit-吞gulp
讲talk-叙narrate	缩shrink-胀swell	潭pond-穴hole	炊cook-厨kitchen
饿hungry-饱full	帅handsome-酷cool	嘻giggle-嘿hey	狗dog-猪pig
窝nest-棚shack	溪rill-泉spring	盈gain-亏loss	云cloud-雾fog
扔throw-抛cast	阴shade-晴clear	尊adore-恭honor	坛jar-罐pot
绳rope-鞭lash	寡few-孤alone	锄hoeing-耕plough	牌poker-棋chess
叔uncle-舅uncle	稚baby-童child	近near-遥remote	掐pinch-拧nip
美beauty-丑clown	高tall-矮short	尺foot-丈ten feet	蝎scorpion-蛇snake
艇yacht-舱cabin	喂feed-饲feed	违fawn-奉flatter	贵noble-贱poor
钞bill-币coin	烟fume-酒wine	买buy-售sell	账bill-债debt
伸extend-屈bend	屑crumb-渣dreg	沟gutter-渠canal	臣minister-君king
苦bitter-甘sweet	府mansion-宅house	裙skirt-裳cloth	枕pillow-眠sleep
圆circle-扁flat	莲lotus-荷lotus	街street-巷alley	磷phosphor-硫sulfur
峰peak-岭ridge	棵a(an)-株a(an)	挖dig-填fill	迈stride-跨stride
杏apricot-桃peach	擦wipe-抹smear	宽wide-窄narrow	豹leopard-熊bear
呆dazed-愣stupefied	勇brave-畏fear	墙wall-檐eave	挑pick-拣select
室room-厅hall	谈discuss-聊talk	拴tie-捆bind	果fruit-核core
敲knock-砸smash	斧axe-劈rive	胃stomach-脾spleen	级class-届grade
雏young-幼young	蓝blue-紫purple	岁year-龄age	神deity-魔demon
捉catch-逮arrest	燕swallow-雀sparrow	怯quail-惧fear	酸sour-甜sweet
睁open-眯squint	梨pear-橘orange	储store-仓depot	羹soup-粥congee
椅chair-凳stool	蒸steam-煮boil	此this-彼that	吊hoist-悬hang
乳newborn-婴infant	牵lead-挽draw	贝conch-螺shell	刀knife-剑sword
碗bowl-筷chopsticks	讽sneer-嘲mock	螳mantis-蝗locust	袋bag-兜pocket
婿husband-妻wife	喊yell-唤call	老old-嫩young	柴wood-炭coal
冷cool-颤tremble	斜skew-歪slant	船ship-帆sail	茧cocoon-蝶papilio
饭rice-汤soul	杯cap-壶kettle	溢spill-漫overflow	蜕molt-蝉cicada
假fake-伪phony	盛thrive-衰decay	恩grace-仇enmity	欣happy-愉glad
纱yarn-绸silk	锤thump-钉beat	发hair-梳comb	聋deaf-哑dumb
砚inkslab-墨ink	蛹pupa-蚕silkworm	夫husband-媳wife	荣bloom-枯dry

(*Continued*)

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press

(*Continued*)

Sublist 1	Sublist 2	Sublist 3	Sublist 4
杭hangzhou-苏soochow	焰flame-烛candle	密thick-疏sparse	笛flute-琴piano
掌palm-拳fist	馍steamed bun-饼pie	酥soft-脆crisp	拽haul-揪pull
赵Zhao-李Li	僧monk-庙temple	雅elegant-俗low	惩punish-赏reward
横across-竖vertical	桶pail-缸urn	穿wear-披drape	鹰hawk-兔rabbit
药drug-疗heal	经longitude-纬latitude	匪bandit-盗robber	袋bag-兜pocket

Cite this article: Zhang, Y., Xie, M., Chen, Y., Xiong, R., Yue, C., Wu, S., Ji, F., and Wang, Q. (2022). Interactions of stimulus quality and semantic context on N400 in visual word recognition. *Applied Psycholinguistics* **43**, 389–410. https://doi.org/10.1017/S0142716421000564

https://doi.org/10.1017/S0142716421000564 Published online by Cambridge University Press