

N400 lexicality effect in highly blurred Chinese words: evidence for automatic processing

Quanhong Wang^{a,b} and Jiajin Yuan^{a,b}

^aKey Laboratory of Cognition and Personality (SVU), MoE and ^bSchool of Psychology, Southwest University, Chongqing, China

Correspondence to Quanhong Wang, School of Psychology, Southwest University, Beibei, Chongqing 400715, China
Tel: +86 23 6825 2309; e-mail: quanhong177@yahoo.com

Received 21 October 2007; accepted 29 October 2007

This study investigated whether N400 was sensitive to primed, highly degraded Chinese compound words (two-character pairs) that were not consciously identifiable, by recording event-related potentials for lightly or highly blurred words and pseudowords while volunteers performed a lexical-decision task. The results showed that pseudowords elicited significantly larger N400 responses than words did, over the central–parietal recording sites, regardless of the stimulus-degradation levels. Moreover, the N400

latency was proportional to the stimulus-degradation level. The N400 effects thus remained observable even when the character pairs were highly degraded and consciously unidentifiable, in addition to the N400 effects that were observed for lightly blurred and intact character pairs. These findings supported the view that the N400 component indexes an automatic processing, providing evidence against the postlexical account. *NeuroReport* 19:173–178 © 2008 Wolters Kluwer Health | Lippincott Williams & Wilkins.

Keywords: automatic processing, Chinese words, event-related potentials, N400, stimulus degradation

Introduction

N400, a negative deflection peaking at about 400 ms after target onset, is a well known component of event-related potential (ERP); it was first observed when a semantically inappropriate word appeared unexpectedly at the end of a sentence [1]. N400 effects are found not only in sentence contexts but also in word-pair contexts. When asked to perform a lexical-decision task, volunteers typically respond more quickly and accurately to a target with a related prime than to one with an unrelated prime: the N400 amplitude is smaller for the related target than for the unrelated target [2,3]. N400 effects were observed not only during semantic-judgment tasks but also during nonsemantic judgments about target stimuli [4–7]. Enhanced N400 is elicited even by nonrhyming words or pseudowords [2,3,8], low-frequency words [9], new words [10], concrete words [7], or words with more orthographic neighbors [11] relative to their counterparts (e.g. rhyming words, high-frequency words, etc.). Words have been found to produce smaller N400 responses even than pseudowords, when no explicit priming context was present [11,12]. A view of the integrative and conscious processes, as well as of the semantic or postlexical processes, is helpful for many researchers in interpreting the N400s found in various studies [3,6].

The postlexical account is, nevertheless, at odds with the findings of many studies demonstrating N400-priming effects when postlexical processes were considered inoperative (e.g. short stimulus-onset asynchrony) [4,5,13]. These findings are better explained with a prelexical, automatic process underlying the N400 effects [14], as they showed that a smaller N400 effect was elicited by primed target

words than by unprimed target words when the stimulus-onset asynchrony between primes and targets was quite short, suggesting that the lexical detector for a target could benefit from the ‘automatic spreading activation’ elicited by a just-presented prime word. The prelexical account has, nonetheless, been challenged in many studies. Holcomb [3] found that the stimulus-degradation level interacted with the behavioral-priming effect [reaction time (RT)], but did not interact with the N400-priming effect, suggesting that the N400 effect did not involve a prelexical process presumably sensitive to stimulus degradation. In addition, despite a significant behavioral-priming effect, the N400-priming effect disappeared when the prime item of a pair was masked [15,16]. Deacon *et al.* [17] and Kiefer and Spitzer [18], however, consistently found a N400 priming effect, regardless of whether the prime was masked or not [6,19,20].

The finding that Chinese compound pseudowords produced greater N400 effects than compound words can be explained by virtue of the postlexical integration of morphemes (meanings of characters) within both compound words [12]. Using a lexical-decision task, this study investigated whether the N400 lexicality effect could occur with primed, highly blurred, two-character Chinese compound words and pseudowords. If an N400 lexicality effect occurs for consciously unidentifiable, highly degraded, two-character pairs as it does for consciously identifiable pairs, it would be evident that the N400 directly reflects an automatic process. In contrast, failure to find any N400 lexicality effect on highly degraded character pairs would strengthen the case for the N400 being a direct reflection of the postlexical process.

Methods

Participants

As paid volunteers, 14 adults (seven women and seven men) aged 18–25 years ($M=21.7$ $SD=2.4$) participated in the study. All volunteers reported normal or corrected-to-normal vision, no neurological or psychological disorders, and all were naive to electrophysiological studies and linguistics. All the volunteers, who gave a written informed consent form, were right-handed as reported.

Stimuli

The stimuli in the lexical-decision task were two-character Chinese words (e.g. 芦笋, asparagus) and pronounceable two-character Chinese pseudowords (e.g. 猫烟). A pseudoword consisted of two Chinese characters (e.g. 猫, cat; 烟, smoke) that neither made up a word nor were found occurring together in a two-character vocabulary [21]. Each stimulus was drawn in black in the center of a white square window (245×245 pixels, 5.2° horizontally and vertically) and saved as a BMP image. Each image was fragmented at blurredness levels 3 and 6, using the fragmentation procedure of Snodgrass and Hirshman [22]; the stimuli at level 3 were consciously made unidentifiable (only 9.9% of the words and 8.1% of the pseudowords at level 3 were correctly identified; see the section on Behavioral procedure); in contrast, the stimuli at level 6 were degraded just a little, in such a way that their performance levels approached those of intact stimuli. Each image was presented in the center of the screen of a 17-inch monitor, 85 cm away from the volunteer's eyes. The size of each word was 2.2° vertically and 4.8° horizontally; the font and size of the characters corresponded to Song Ti No. 119 (2.2° horizontally and vertically).

Behavioral procedure

The stimuli in the lexical-decision task were delivered via a PC while the volunteers were seated in a quiet room, at a distance of 85 cm from the computer screen. At each trial, the volunteer first viewed a fixation mark for 200 ms. A blank screen was then presented, designed to avoid stimulus-onset expectation, for randomly varying durations of between 500 and 2000 ms; this, in turn, was followed by the onset of a stimulus. The volunteer was then asked to judge as quickly and accurately as possible whether the stimulus was a word or a pseudoword by pressing one of two keys. The stimulus onset was terminated either by pressing a key or it was terminated when 2000 ms had elapsed, which allowed for a long intertrial interval. Assignment of keys and response hands was counterbalanced across volunteers.

Two tasks had to be performed during each of the six blocks: the two tasks were the same, with the exception that the first task required responses to intact stimuli and the second to blurred stimuli. Thirty intact words and 30 intact pseudowords were used, with each being presented only once during each block. The onset sequences of the words, as well as pseudowords, were at random for each volunteer. In total, 180 intact words and 180 intact pseudowords, as well as 180 blurred words and 180 blurred pseudowords, were thus used in this experiment. Once the volunteers completed the lexical decisions for the 60 intact stimuli, the second task was initiated. Volunteers then made distinctions between words and pseudowords for the corresponding

blurred stimuli, the number of which was matched over all degradation levels for each stimulus type. Similarly, each of the blurred stimuli appeared only once, and the onset sequence of these blurred stimuli was at random for each volunteer. The purpose of the first task was to prime the processing of the blurred stimuli subsequently presented in the second task; blurredness levels at which the word stimuli appeared were counterbalanced across volunteers. A 1-min break was introduced between two successive blocks, to avoid overtiring the volunteers, as that might obscure the lexicality effect on ERPs. Furthermore, several practice trials with extra stimuli were used to familiarize the volunteers with the procedure, and only volunteers whose responses had been accurate during practice could participate in the formal experiment.

To determine the degree to which the word stimuli at level 3 were consciously identified, identification rates were measured with extra volunteers. The procedure for this behavioral test was similar to that of the ERP experiment, except that it was run without any electroencephalographic (EEG) recordings: volunteers were simply asked to type either the name of each stimulus or a 'no' (if they could not venture a guess).

Recording and data analysis

The EEG was recorded from 64 scalp sites using tin electrodes mounted in an elastic cap (Brain Product GmbH, Germany), with the references on the left and right mastoids and a ground electrode on the medial frontal aspect. Vertical eye movements were recorded with electrodes placed on the supraorbital and infraorbital ridges of the right eye, and horizontal eye movements were recorded with electrodes placed lateral to the outer canthi of both eyes. The impedance of all electrodes was less than $5\text{ k}\Omega$. The EEG and electrooculogram signals were amplified using a bandpass of 0.01–100 Hz, and were continuously digitized at 500 Hz/channel.

The average epoch was 1200 ms, including a 200-ms prestimulus baseline. The EEGs were averaged separately for words and pseudowords, each stimulus type at integrity level, and at blurredness levels 6 and 3. Only segments with correct responses were averaged, and the averaged numbers of averaging segments were 66 and 65 for level-6 stimuli, and 32 and 51 for level-3 stimuli. Trials with contaminations from electrooculogram artifacts, amplifier clipping, bursts of electromyogram activity, or peak-to-peak deflections exceeding $\pm 80\ \mu\text{V}$ were excluded from averaging.

As shown by Figs 1, 3, and 4, the amplitude differences between word and pseudoword conditions were largest over the central and centroparietal scalp areas during the level-3 condition. The following eight electrode points, based on the grand-average and topographical maps of the ERPs during the level-3 condition, were thus chosen for statistical analyses: FCz, C1, Cz, C2, CP1, CPz, CP2, and Pz. For the data from the second task, averaged amplitudes in each time window were analyzed using three-way repeated-measures analysis of variance (ANOVA). The ANOVA factors were stimulus type (words and pseudowords), blurredness (levels 3 and 6), and electrode site (the chosen electrodes) for the second task, whereas they were stimulus type and electrode site for the first task. For all analyses, the degrees of freedom of the F ratio were corrected according to the Greenhouse–Geisser method [23].

Results

Reaction time and accuracy

A two-way repeated-measures ANOVA on RT and accuracy data showed significant interaction effects between stimulus type and blurredness level [$F(1,13)=9.13$, $P<0.01$; $F(1,13)=13.22$, $P<0.01$]. Simple-effect ANOVA showed that the mean RTs to words (1017.0 ms) and pseudowords (1038.2 ms) were not significantly different from each other at level 3. At level 6, however, the RT to words (792.8 ms) was significantly faster than to pseudowords (919.9 ms), $F(1,13)=48.82$, $P<0.001$.

The accuracy rate of key pressing was significantly higher for pseudowords (74.4%) than for words (38.2%) at level 3, $F(1,13)=16.03$, $P<0.01$. For level 6, there was no significant difference between words (94.7%) and pseudowords (89.1%), $F(1,13)=1.06$, $P=0.32$. An additional experiment was conducted on a small group of extra volunteers, so as to evaluate individual confidence in word identification, for level 3, on a scale of 5 points. The d' was 0.87 and the criterion for judging the stimuli as words was 1.65, demonstrating a strong tendency for volunteers to judge stimuli at level 3 as pseudowords rather than as words, which is consistent with the fact that the accuracy was lower for words than for pseudowords.

Event-related brain potentials

Figure 1 illustrates average ERPs for intact words and pseudowords at Cz. As seen from Fig. 1, pseudowords elicited more negativity than words after approximately 300 ms, and this difference was most prominent at the 320–580-ms interval. The subtraction of the waveforms of ERPs evoked by words from those of the ERPs evoked by pseudowords generated a conspicuous N400 component, which peaked at around 462 ms after stimulus onset; the differences in N400 amplitudes were largest over the midline centroparietal recording sites. Indeed, the repeated-measures ANOVA on the averaged amplitudes in the 320–580-ms time window showed main effects of stimulus type [$F(1,13)=15.49$, $P<0.01$] and electrode site [$F(1,13)=11.29$, $P<0.01$]. Despite this finding, no significant interaction effect between stimulus type and electrode site was observed [$F(7,91)=1.03$, $P=0.37$]. The averaged

amplitudes were greater at C2, CP2, and Pz than at C1 [$F(1,14)=52.13$, $P<0.01$], CP1 [$F(1,13)=5.68$, $P<0.05$], and FCz [$F(1,13)=5.68$, $P<0.05$], respectively.

More importantly, Fig. 2 showed a greater negative deflection for blurred pseudowords than for blurred words at both levels 6 (about 200–600-ms interval) and 3 (790–890-ms interval). The analysis on the averaged amplitudes in the 200–400-ms interval showed a significant interaction of stimulus type and blurredness level [$F(1,13)=5.58$, $P<0.05$]. Simple-effects analyses showed that only at level 6, pseudowords produced more negativity than words did [$F(1,13)=27.67$, $P<0.001$], suggesting that the difference between pseudowords and words, a lexicality effect, started earlier for level 6 (at around 200 ms) than for level 3 and intactness level (Fig. 1). The analyses in the 400–500-ms interval showed significant main effects of stimulus type [$F(1,13)=10.98$, $P<0.01$], blurredness level [$F(1,13)=17.01$, $P<0.01$], and electrode site [$F(7,91)=10.88$, $P<0.01$], and also significant interaction effects between blurredness level and stimulus type [$F(1,13)=6.23$, $P<0.03$], and between blurredness level and electrode site [$F(7,91)=6.25$, $P<0.01$]. The level-6 condition elicited larger amplitudes than the level-3 condition; the amplitudes were larger at the posterior relative to the anterior sites; and they were larger at the right-hand-side versus the left-hand-side sites. Further analyses of the interaction effect between blurredness level and stimulus type demonstrated a significant lexicality effect for level 6 [$F(1,13)=24.00$, $P<0.001$], whereas the lexicality effect for level 3 was not significant ($F<1$). Analyses on averaged amplitudes during the 500–600-ms interval revealed a similar pattern of results (Table 1). In addition, as shown in Table 1, the reversed pattern of the interaction effect of blurredness level by stimulus type was observed during the 790–890-ms interval, at which point a significant lexicality effect was observed for level 3 but not for level 6.

It is thus clear that amplitude differences between word and pseudoword conditions were significant during the level-6 condition at the 200–600-ms time interval, whereas only level 3 witnessed significant amplitude differences between word and pseudoword conditions at 790–890 ms.

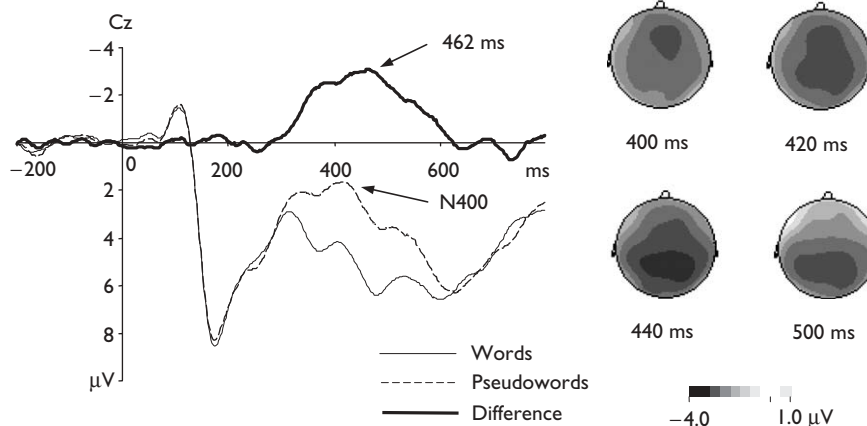


Fig. 1 Left: grand average event-related potentials to intact words (solid line) and pseudowords (dashed line), and the pseudoword-minus-word difference wave (bold line) at Cz. Right: topographical maps of the voltage amplitudes for the pseudoword-minus-word difference wave (peaked at 462 ms) for intact verbal stimuli at the 400, 420, 440, and 500 ms.

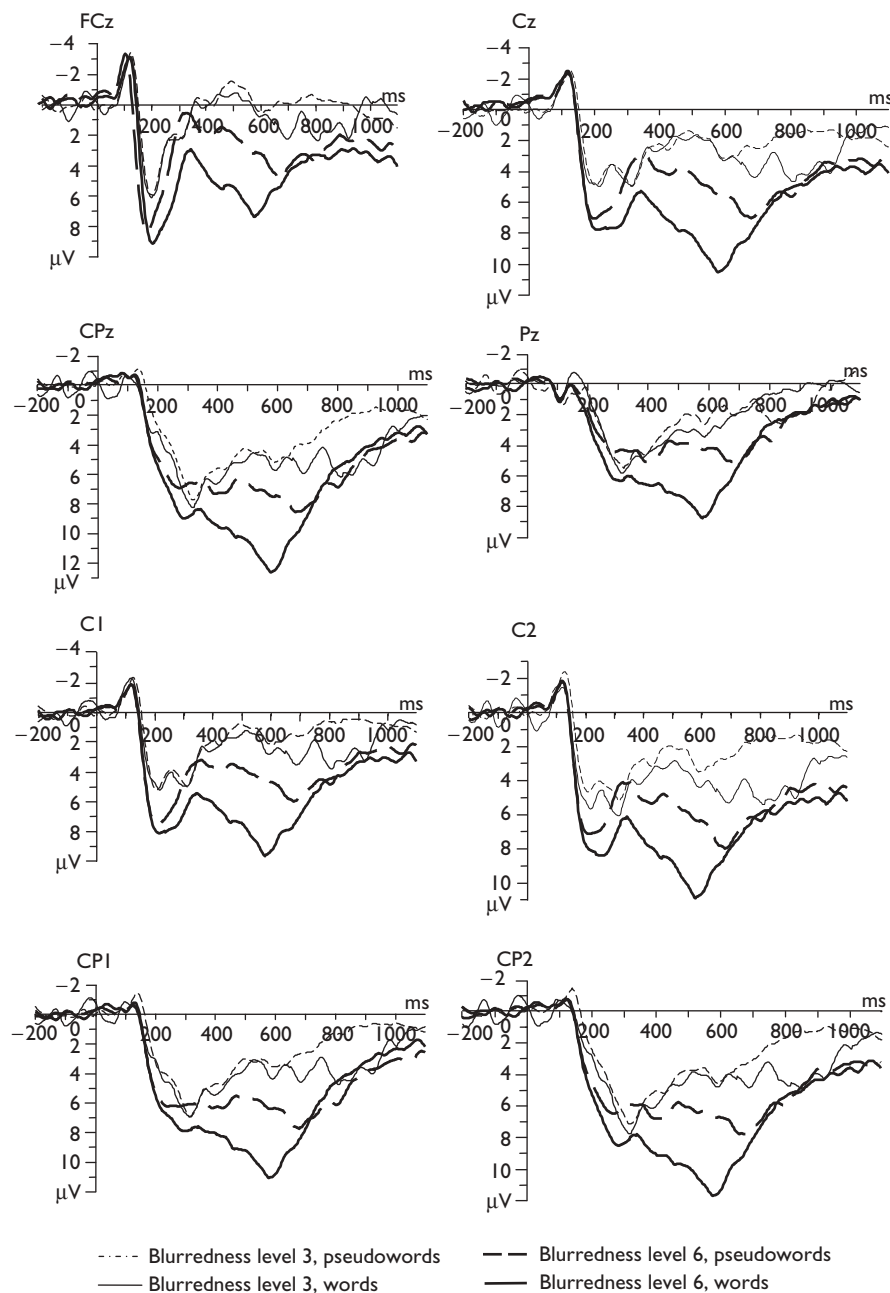


Fig. 2 Grand average ERPs elicited at FCz, Cz, CI, C2, CPI, CP2, CPz, and Pz for all conditions of the second task.

Discussion

Chinese compound pseudowords produced a more negative ERP deflection than Chinese compound words when they were intact or lightly blurred (level 6). A similar effect was found when all the words were so highly blurred (level 3) that they were consciously unidentifiable.

The behavioral data showed very low accuracy rates in identifying words (9.9%) and pseudowords (8.1%) at level 3: these accuracy rates were far below the criterion for conscious identification, as adopted by Deacon *et al.* [17]. As indicated by the statistical results and by Fig. 4, pseudowords elicited significantly more negativity than words did, in the 790–890-ms interval during level 3. This enhanced negativity was manifested as a delayed N400

component, the topographical map of which resembles the classical centroparietal distribution [1], in the pseudoword-minus-word difference wave. The N400 effects elicited during the level-3 conditions were similar to those observed during the level-6 and intact conditions in both waveform morphology and topographical distribution (compare Fig. 4 with Figs 1 and 3), except for a major difference in their peak latencies. In fact, the peak latency of the N400 difference elicited by the level-3 condition was much longer than in the other two conditions (the averaged peak latency of N400 was 812 ms for level 3, 580 ms for level 6, and 462 ms for intact words). As the response hands for the words and pseudowords were counterbalanced across volunteers, and as all volunteers were self-reportedly right-handed, the

N400 effects observed in this study index a lexical processing. Moreover, as the responses that each volunteer made to words and pseudowords were the same across tasks and across blurredness conditions, the N400 latency variations across blurredness conditions ought not to be explained by differences in decisional and response readiness for the two types of stimuli [24]. Rather, this effect ought to be explained by stimulus degradation: the more degraded the stimuli, the longer the latency. Holcomb [3] observed that degraded stimuli produced a longer N400 latency than intact stimuli did, which also agrees with these findings. Holcomb attributed this effect to a relatively early word-level processing that is sensitive to stimulus degradation.

The finding that the N400 lexicality effect occurred for highly degraded words and pseudowords that were consciously unidentifiable provides evidence that N400 directly reflects an automatic process. From masked priming studies, it is now generally accepted that N400 is exclusively sensitive neither to any postlexical semantic process nor to the influence of automatic priming. Holcomb *et al.* [19] proposed a stern criterion as proof of a purely automatic process: he claimed that only by finding evidence of N400 priming effects when the evoking (target) stimulus itself, rather than the priming stimulus, is processed automatically can the N400 exclusively or directly reflect

an automatic process. This study followed a similar approach, to show that the N400 directly reflects an automatic process. Specifically, there were two characters (morphemes) in every compound word or pseudoword, and the first was treated as a prime and the second as a target. The processing of both characters was made automatic by severe degradation. In this way, any conscious, semantic, or postlexical process, such as morphemic integration of the prime and the target, was prevented. The finding that the N400 lexicality effect occurred for highly blurred words and pseudowords, in the same manner as it did for consciously recognizable words and pseudowords, was evidence that N400 directly reflected an automatic, rather than a postlexical, process.

This study replicated the findings of previous studies that pseudowords produced greater N400 effects than words without explicit contexts [11,12]. As this study showed that N400 lexicality effects occurred even for highly degraded character pairs that could not consciously be recognized, N400 probably reflected a lexical process. In fact, Deacon *et al.* [25] explicitly dropped the term 'prelexical process', which was considered misleading, and adopted the term 'lexical process', to mean that it is a lexical, rather than a prelexical, process that accounts for the N400 priming effect. According to Deacon *et al.* [25], N400 is proportional to the amount of orthographic or phonological processing neces-

Table I Summary of the ANOVA results on main effects of word status (lexicality) and blurredness level, their interaction (WS × BL), simple word status effects (the parentheses indicate significant blurredness levels), electrode site main effect, and interactions of electrode site and word status (RS × WS) and of electrode site and blurredness level (RS × BL) for the 400–500, 500–600, and 790–890-ms intervals, respectively

Time (ms)	Word status		Blurredness level		WS × BL		Simple word status		Electrode site		RS × WS		RS × BL	
	F	P	F	P	F	P	F	P	F	P	F	P	F	P
400–500	10.98	0.01	17.01	0.01	6.23	0.03	24.0	0.001 (L6)	10.88	0.01	1.33	0.28	6.25	0.01
500–600	20.12	0.01	23.63	0.001	13.06	0.01	73.66	0.001 (L6)	9.50	0.01	1.76	0.16	3.92	0.04
790–890	7.86	0.02	4.99	0.05	5.55	0.04	11.64	0.01 (L3)	5.70	0.01	1.01	0.38	1.54	0.24

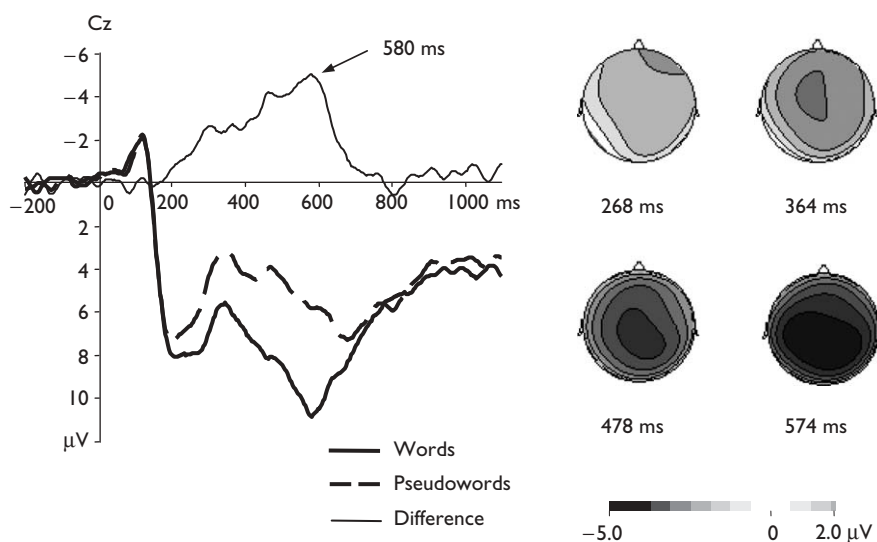


Fig. 3 Left: grand average ERPs to words (solid line) and pseudowords (dashed line) at blurredness level 6, and the pseudoword-minus-word difference wave (thin line) at Cz. Right: Topographical maps of the voltage amplitudes for the pseudoword-minus-word difference wave (peaked at 580 ms) for the blurredness level at 268, 364, 478, and 574 ms.

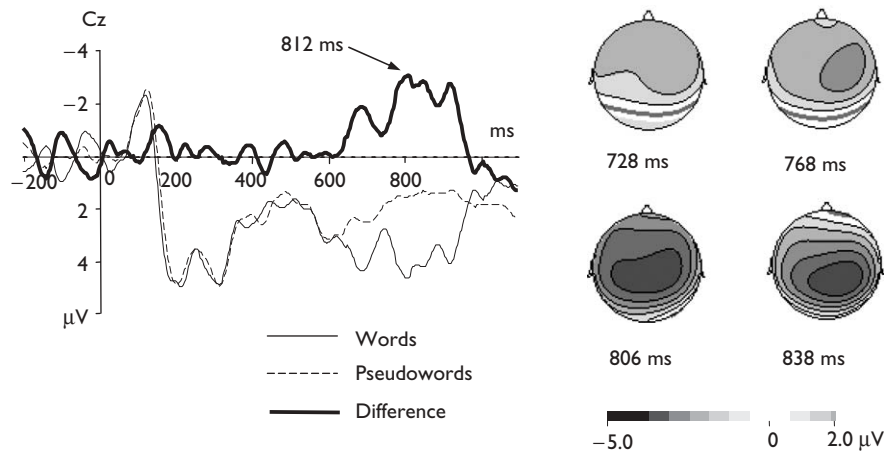


Fig. 4 Left: grand average ERPs to words (solid line) and pseudowords (dashed line) at blurredness level 3 and the pseudoword-minus-word difference wave (bold line) at Cz. Right: Topographical maps of the voltage amplitudes for pseudoword-minus-word difference wave (peaked at 812 ms) for the blurredness level at 728, 768, 806, and 838 ms.

sary to extract meaning from a word, which is reduced in the presence of related contextual words.

Conclusion

This study demonstrated that two-character Chinese pseudowords elicited greater N400 than did words, even when they were too highly degraded to be consciously identifiable. The N400 lexicality effect thus occurs independently of stimulus degradation, which provides evidence for a direct automatic process underlying the N400 lexicality effect and against the postlexical account.

Acknowledgements

The research for this study was sponsored by the Doctoral Foundation of Southwest University (SWNUB 2004019) and the National Key Discipline of Basic Psychology in Southwest University (NSKD06004).

References

- Kutas M, Hillyard SA. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 1980; **207**:203–205.
- Bentin S, McCarthy G, Wood CC. Event-related potentials associated with semantic priming. *Electroencephalogr Clin Neurophysiol* 1985; **60**:343–355.
- Holcomb PJ. Semantic priming and stimulus degradation: implications for the role of the N400 in language processing. *Psychophysiology* 1993; **30**:47–61.
- Besson M, Fischler I, Boaz T, Raney G. Effects of automatic associative activation on explicit and implicit memory tests. *J Exp Psychol Learn Mem Cogn* 1992; **18**:89–105.
- Kutas M, Hillyard SA. An electrophysiological probe of incidental semantic association. *J Cogn Neurosci* 1989; **1**:38–49.
- Misra M, Holcomb PJ. Event-related potential indices of masked repetition priming. *Psychophysiology* 2003; **40**:115–130.
- West WC, Holcomb PJ. Imaginal, semantic, and surface-level processing of concrete and abstract words: an electrophysiological investigation. *J Cogn Neurosci* 2000; **12**:1024–1037.
- Rugg MD. Event-related brain potentials in phonological matching tasks. *Brain Lang* 1984; **23**:225–240.
- Barber H, Vergara M, Carreiras M. Syllable-frequency effects in visual word recognition: evidence from ERPs. *NeuroReport* 2004; **15**:545–548.
- Rugg MD, Nagy ME. Event-related potentials and recognition memory for words. *Electroencephalogr Clin Neurophysiol* 1989; **72**:395–406.
- Holcomb PJ, Grainger J, O'Rourke T. An electrophysiological study of the effects of orthographic neighborhood size on printed word perception. *J Cogn Neurosci* 2002; **14**:938–950.
- Wang Q, Tao W, Wu D. Effects of relationships between Chinese morphemes on N400 component of event-related potentials (2007, submitted).
- Anderson J, Holcomb P. Auditory and visual semantic priming using different stimulus onset asynchronies: an event-related brain potential study. *Psychophysiology* 1995; **32**:177–190.
- Van Petten C, Kutas M. Electrophysiological evidence for the flexibility of lexical processing. In: Simpson G, editor. *Word and sentence*. Amsterdam: North Holland Press; 1991. pp. 129–174.
- Brown CM, Hagoort P. The processing nature of the N400: evidence from masked priming. *J Cogn Neurosci* 1993; **5**:34–44.
- Ruz M, Madrid E, Lupianez J, Tudela P. High density ERP indices of conscious and unconscious semantic priming. *Cogn Brain Res* 2003; **17**:719–731.
- Deacon D, Hewitt S, Yang C-M, Nagata M. Event-related potential indices of semantic priming using masked and unmasked words: evidence that the N400 does not reflect a post-lexical process. *Cogn Brain Res* 2000; **9**:137–146.
- Kiefer M, Spitzer M. Time course of conscious and unconscious semantic brain activation. *NeuroReport* 2000; **11**:2401–2407.
- Holcomb PJ, Reder TL, Misra M, Grainger J. The effects of prime visibility on ERP measures of masked priming. *Cogn Brain Res* 2005; **24**:155–172.
- Kiefer M, Brendel D. Attentional modulation of unconscious 'automatic' processes: evidence from event-related potentials in a masked priming paradigm. *J Cogn Neurosci* 2006; **18**:184–198.
- Institute of Language Teaching and Research. *Statistics and analyses of Mandarin Chinese vocabulary [in Chinese]*. Beijing: Beijing Language Institute Press; 1984.
- Snodgrass JG, Hirshman E. Theoretical explorations of the Bruner-Potter (1964) interference effect. *J Mem Lang* 1991; **30**:273–293.
- Greenhouse WW, Geisser S. On methods in the analysis of profile data. *Psychometrika* 1959; **24**:95–112.
- Yuan J, Zhang QL, Chen AT, Li H, Wang Q, Zhuang ZCX, Jia SW. Are we sensitive to valence differences in emotionally negative stimuli? Electrophysiological evidence from an ERP study. *Neuropsychologia* 2007; **45**:2764–2771.
- Deacon D, Dynowska A, Ritter W, Grose-Fifer J. Repetition and semantic priming of nonwords: implications for theories of N400 and word recognition. *Psychophysiology* 2004; **41**:60–74.