Cognition 154 (2016) 11–21

Contents lists available at ScienceDirect

Cognition

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

Prosodic expectations in silent reading: ERP evidence from rhyme scheme and semantic congruence in classic Chinese poems

Qingrong Chen [a,b,*], Jingjing Zhang [c,*], Xiaodong Xu [d], Christoph Scheepers [e], Yiming Yang [f,*], Michael K. Tanenhaus [b,*]

[a] School of Psychology, Nanjing Normal University, China
[b] Department of Brain and Cognitive Sciences, University of Rochester, NY, USA
[c] State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, China
[d] School of Foreign Languages and Cultures, Nanjing Normal University, China
[e] Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, UK
[f] School of Linguistic Sciences, Jiangsu Normal University, China

ARTICLE INFO

Article history:
Received 8 July 2014
Revised 4 November 2015
Accepted 13 May 2016
Available online 23 May 2016

Keywords:
Prosody
Poem
Rhyme scheme
Semantics
P200
N400
N400-like

ABSTRACT

In an ERP study, classic Chinese poems with a well-known rhyme scheme were used to generate an expectation of a rhyme in the absence of an expectation for a specific character. Critical characters were either consistent or inconsistent with the expected rhyme scheme and semantically congruent or incongruent with the content of the poem. These stimuli allowed us to examine whether a top-down rhyme scheme expectation would affect relatively early components of the ERP associated with character-to-sound mapping (P200) and lexically-mediated semantic processing (N400). The ERP data revealed that rhyme scheme congruence, but not semantic congruence modulated the P200: rhyme-incongruent characters elicited a P200 effect across the head demonstrating that top-down expectations influence early phonological coding of the character before lexical-semantic processing. Rhyme scheme incongruence also produced a right-lateralized N400-like effect. Moreover, compared to semantically congruous poems, semantically incongruous poems produced a larger N400 response only when the character was consistent with the expected rhyme scheme. The results suggest that top-down prosodic expectations can modulate early phonological processing in visual word recognition, indicating that prosodic expectations might play an important role in silent reading. They also suggest that semantic processing is influenced by general knowledge of text genre.

1. Introduction

A growing body of work suggests that expectations are important in language processing with readers and listeners generating predictions that affect processing at multiple levels. When the incoming input has features consistent with expectancies, processing is facilitated (e.g. Altmann & Kamide, 1999; Arai & Keller, 2013; Bicknell, Elman, Hare, McRae, & Kutas, 2010; Brown, Salverda, Dilley, & Tanenhaus, 2011; DeLong, Urbach, & Kutas, 2005; Farmer, Christiansen, & Monaghan, 2006; Federmeier, 2007; Kamide, Altmann, & Haywood, 2003; Kimball, 1975; Levy, 2008; Staub & Clifton, 2006; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005; see Kamide, 2008 for an overview of anticipatory effects in the sentence processing). For example, in current generative approaches to language processing, these expectations play an important role in instantiating hypotheses that try to explain the input as it unfolds, and updating hypotheses based on the degree of mismatch between model-based expectations and the input (Farmer, Brown, & Tanenhaus, 2013; Farmer, Yan, Bicknell, & Tanenhaus, 2015; Jaeger & Snider, 2013; Kuperberg & Jaeger, 2016).

Recently, there has been emerging evidence that form-based expectation can affect early processing in spoken language comprehension. For example, expectations about prosodic structure, including expectations conditioned on rate, sequences of pitch accents and stress patterns influence the earliest moments of

* Corresponding authors at: School of Psychology, Nanjing Normal University, 210097, China; Institute of Psychology, Chinese Academy of Sciences, 100101, China; School of Linguistic Sciences, Jiangsu Normal University, 221009, China; Department of Brain and Cognitive Sciences, University of Rochester, NY 14627, USA.

E-mail addresses: jscqr80@sina.com (Q. Chen), Zhangjingjing@psych.ac.cn (J. Zhang), yangym@jsnu.edu.cn (Y. Yang), mtan@mail.bcs.rochester.edu (M.K. Tanenhaus).

http://dx.doi.org/10.1016/j.cognition.2016.05.007
0010-0277/© 2016 Elsevier B.V. All rights reserved.

spoken word recognition, affecting the parsing of the speech stream and recognition of spoken words (Breen, Dilley, McAuley, & Sanders, 2014; Brown, Salverda, Dilley, & Tanenhaus, 2015; Brown et al., 2011; Reinisch, Jesse, & McQueen, 2011). In spoken language there is a clear rationale for why prosodic expectations might help structure the input. Word boundaries are not clearly marked in the input and the duration of syllables and other fine-grained variations in acoustic/phonetic cues are conditioned on speech rate and prosodic structure (Steinhauer, Alter, & Friederici, 1999; see Dahan, 2015 for a recent review).

The possible role of prosodic expectations in silent reading is less well understood. There are also compelling reasons why prosodic expectations might play a much less important role in reading than they do in spoken language processing. Unlike spoken language, the rate of the input in reading is not affected by prosodic factors. Rather, the reader controls the timing of the input. In addition, for alphabetic orthographies, word boundaries are clearly marked. Therefore, prosodic expectations would not be play the same role in segmentation as they do in spoken language. Moreover, many aspects of prosody are explicitly signaled by orthographic information such as commas and capitalization.

Nonetheless, there is some evidence that reader's expectations about the phonological *form* of a word might affect early components of visual word recognition. For example, Farmer et al. (2006) created a phonological typicality metric for nouns and verbs. In contexts that strongly bias a noun, more typical nouns are processed more quickly than less typical nouns. These form-based typicality effects are found in early components of the MEG response (Dikker & Pylkkanen, 2011) and they affect first-pass eye-movement measures (Farmer et al., 2015). It is important to note, however, that in alphabetic languages orthographic and phonological representations are strongly correlated. Therefore, these typicality effects might be due to expectations about the orthographic form of word categories rather than phonological typicality effects *per se* (Farmer et al., 2015; Tanenhaus & Hare, 2007).

A growing body of literature suggests that prosodic representations (though not necessarily expectations) are computed in silent reading (for a recent review see Clifton, 2015). The phonological structure of a word (e.g., its syllable structure) is accessed in visual word recognition. In addition, more abstract aspects of prosody such as lexical stress, information structure marked by punctuation and capitalization, and even implicit prosody tied to phrasal grouping affect early eye-movement measures. It is unclear, however, whether these reflect prosodic expectations, and if so, what their relative timing is with respect to (post-lexical) phonological representations that would be made available as a by-product of lexical access.

As noted earlier, in alphabetic languages it is difficult to rule out the possibility that prosodic expectations might be orthographically mediated. Consider, for example, syllabic representations, which play a central role in prosody. One crucial aspect of a syllable is the rime (e.g., the vowel and final consonant(s) in a CVC). An expectation for the form of the rime could be mediated by an expectation for an orthographic form because syllables with similar rhymes would share orthographic features. Similar arguments can be made for other prosodic expectations.

In contrast, in some languages like Chinese, which have logographic orthographies, an expectation for a rime need not map onto an expectation for an orthographic form (We're setting aside characters that have phonetic radicals). Therefore, an expectation for a particular prosodic feature, such as a rime, need not map onto an expectation for a character or features that would be shared by the characters that have syllables with that rime. However, in order to create prosodic expectations that cannot be mapped onto orthographic/visual expectations it is also necessary to use contexts in which an expectation for a form occurs independently of an expectation for a particular word. Otherwise, semantic expectations would re-introduce the possibility that the effects were orthographically/visually mediated. We created such contexts by using classic Chinese poems.

Rhyme scheme is one of the most typical prosodic features of classic Chinese poetry (e.g., Obermeier et al., 2013). As shown in example 1a, the poem lines have an identical rhyme scheme (i.e., '伤 sh*ang*' and '荒 hu*ang*') according to classical poem rules.

1a. 十/年/归/客/但/心/伤/,
Shi/nian/gui/ke/dan/xin/***shang***/,
Ten/year/return/people/heart/sad/,
The returnee feels so sad after ten years,
三/径/无/人/已/自/荒/。
San/jing/wu/ren/yi/zi/***huang***/.
Three/path/no/people/already/natural/overgrown with weeds/.
Many paths that were not treaded by people have been overgrown with weeds.

Most Chinese readers are familiar with the rhyme scheme of the classical poetry like 1a but they are unfamiliar with the poems. Therefore, when people read the first line, they can anticipate a rhyming character in the second line. Because, as mentioned earlier, Chinese orthography is character-based, and critical characters do not share sub-lexical information about the pronunciation of that character, rhyme expectations would not be reflected in expectations about the orthography of the target word. As a consequence, Chinese readers will have a strong expectation for a character that maps onto a particular rime in a specific position in the text. However, they will not be able to predict either the particular word or the orthographic properties of the character.

The primary goal of the current paper was to examine if, and if so, when prosodic expectations are used in reading. In particular we were interested in how they might interface with the processes that make available phonological representations in silent reading. There are two possibilities. The first hypothesis is that syllable expectations that are induced by the rhyme scheme are evaluated using *post-lexical* phonological representations that are made available as a consequence of lexical access. A second, and more intriguing hypothesis is that prosodic expectations are used in conjunction with phonological information that is made available early in the word recognition process (i.e., pre-lexically). This would allow prosodic representations to be constructed and used before a word is fully recognized. Moreover, it would allow top-down prosodic expectations to influence visual word recognition. And, because word and character boundaries are not marked by spaces, prosodic expectations could, in principle, help readers segment the input.

The post-lexical and pre-lexical hypotheses can be clarified by focusing on the predictions they make about the relative time course of prosodic and semantic expectation effects. The post-lexical hypothesis predicts that semantic and syllable expectation effects should have a similar time-course, whereas the pre-lexical hypothesis would predict that syllable effects would precede semantic effects.

We used stimuli drawn from classic Chinese poems where, because of the well-known rhyme scheme, the last word of the first line of the poem creates an expectation that the word in the same position in the second line will rhyme with that word. We manipulated the congruity of the rhyme scheme by replacing the poem-final character in the second line (e.g., '荒 hu*ang*') with rhymed (e.g., '香 xi*ang*') or unrhymed characters (e.g., '悠 y*ou*'). In addition, we manipulated whether the word was semantically congruous or incongruous with the meaning of the poem. We monitored ERPs because they allow us to compare the time course of rhyme effects with the time course of semantic congruency effects.

The rationale for using ERPs is the following. There is considerable evidence that phonological representations are accessed during visual word recognition in Chinese. If prosodic expectations are evaluated as part of post-lexical processing, then rhyme incongruence effects should be reflected in an "N400-like" response, with timing similar to that of classic semantic and lexical effects on N400, but with a more right-hemisphere distribution (Coch, Grossi, Skendzel, & Neville, 2005; Coch, Hart, & Mitra, 2008; Rugg, 1984a, 1984b) than is found with a classic semantically or lexically-mediated N400. If, however, rhyme expectations interface with pre-lexical phonological representations, rhyme expectation effects would be reflected in an earlier component and would manifest themselves before semantic congruency effects. For the reasons detailed below, we hypothesized that any such effects would be reflected in the P200 component.

Behavioral studies have demonstrated clear effects of character consistency, where character consistency is operationalized as the consistency with which a character maps onto a particular syllable. The highest consistency characters map onto only one syllable, whereas lower consistency characters map onto more than one syllable. Naming and lexical decision times are longer to low consistency characters compared to high consistency characters, for both high and low frequency words and characters (e.g., Fang, Horng, & Tzeng, 1986; Hue, 1992; Lee, Tsai, Su, Tzeng, & Hung, 2005). These effects of consistency are robust, even when all other variables are tightly controlled (Lee, Hsu, Chang, Chen, & Cho, 2015). Consistency effects are generally interpreted as reflecting phonological mapping, a process typically considered to be an early, pre-lexical component of visual word recognition. ERP evidence is consistent with this pre-lexical interpretation (Lee et al., 2007), showing that low-frequency characters that are less phonetically consistent (e.g., '缀 zhui', '掇 duo', and '辍 chuo') result in larger P200 responses than characters that are more phonetically consistent (e.g., '琬, 惋, 腕, 蜿, 豌, 碗, 婉 wan'). These findings have been interpreted as support for a two-stage model of lexical access (Perfetti, Liu, & Tan, 2005) in which the P200 effect reflects phonetic radical interference at an early phonological level.

If the P200 effects reflect pre-lexical processes involved in assigning a phonological representation to a character, then congruency with rhyme expectations would be likely to modulate the P200. If rhyme scheme expectancy facilitates assigning a syllabic structure to the character, then it should be reflected in a smaller P200 for the character that matches the rhyme scheme. Moreover, if these effects are due to phonological form-based expectations that occur prior to semantic processing, we would not expect to see either effects of semantic congruity or interactions with rhyme scheme congruity (Huang, Yang, Zhang, & Guo, 2014). Therefore, evidence that rhyme scheme expectation affects P200 would provide clear evidence for top-down phonological effects in early visual word recognition. These might or might not, be followed by post-lexical N400-like effects.

Our design also allows us to examine when and to what extent prosodic expectancy affects semantic processing in silent reading. Semantic processes often considered to be largely determined by the content of an utterance, i.e., the congruity of the meaning of a word with the prior context. If content-based semantic processing is a separate processing stream in reading poems, we would expect to see semantic congruity effects regardless of rhyme scheme. These congruity effects would be consistent with a general understanding of factors that influence the N400 (see Kutas & Federmeier, 2011, for a review of N400 component).

However, the form of the representations that readers construct and what dimensions are considered to affect meaning, might depend on the genre of a text. In narrative text, meaning is likely to be strongly tied to the (denotative) content of the text. However, according to the genre-specific hypothesis of Jakobson (1960) and conventionalist theories of literature (Culler, 1975, pp. 1–368), meaning might vary with the genre of a text. In particular, one view of semantic processing in poetry suggests that "similarity in sound may be taken as a similarity in meaning or the connotative associations of a particular linguistic sign" (Hanauer, 1998, p. 65). If this is the case, the perception of semantic congruity might be strongly modulated by rhyme scheme expectation. If the rhyme incongruent target is viewed as semantically anomalous in the context of poem context, this would increase the difficulty of accessing information and semantic integration processing and therefore would yield a significant N400 component (e.g., Baggio, Van Lambalgen, & Hagoort, 2008; Federmeier & Kutas, 1999; Hagoort, Hald, Bastiaansen, & Petersson, 2004), for even the semantically congruous characters. Thus, the difference of N400 amplitude for the character with consistent semantics compared to the character with inconsistent semantics would depend upon rhyme congruence. This would also be consistent with the neurocognitive poetics model of literary reading proposed in Jacobs (2014).

In sum, we used classical Chinese poems to induce a rhyme scheme expectation for a particular word and used ERPs as a dependent measure because they allow us to study the time course of prosodic effects. We manipulated whether or not the target word was congruent with the rhyme scheme expectation and whether or not it was semantically congruent with the content of the poem. We asked whether rhyme scheme expectation influences P200, which reflects phonological processes, such as assigning a syllabic structure to a character, that are typically considered to reflect early "pre-lexical" processing, predicting that rhyme congruence would be reflected in a smaller P200. We also examined whether rhyme scheme expectation would modulate the N400 and used the properties of our stimuli to examine the interaction between semantic congruity and rhyme scheme congruity (e.g., Luo & Zhou, 2010).

2. Methods

2.1. Participants

Nineteen students from Nanjing Normal University (10 females, mean age 24.1, age range 23–26) participated in the experiment. All participants were right-handed as determined by the Chinese Handedness Questionnaire (Li, 1983). They were all native Mandarin Chinese speakers who had normal or corrected-to-normal vision, no history of neurological or language impairments and no previous exposure to the experimental items. This study was approved by the Academic Committee of the School of Psychology, Nanjing Normal University. All participants provided informed consent prior to participation and were paid for their participation.

2.2. Materials and design

As shown in Table 1, the stimuli consisted of four versions of classical poems, which differed in whether or not the target word matched the rhyme scheme (Rhyme+ or Rhyme−) and was either semantically congruous (Semantic+ or Semantic−). We will refer to the conditions as: SEM+ RHY+, SEM+ RHY−, SEM− RHY+, and SEM− RHY−. One hundred and forty-eight Chinese seven-character poems of the Tang Dynasty were used as sources for the SEM+ RHY+ condition, where every two lines rhymed in the final characters (e.g., '山sh*an*' and 攀p*an*' in Table 1). To generate the remaining three versions per poem (i.e., SEM+ RHY−, SEM−

Table 1
Experimental conditions and exemplar poems.

Condition	Example	Rhyme scheme	Semantic
SEM+ RHY+	君王/台榭/枕/巴/山/,/万丈/丹梯/尚/可/攀/。 Junwang/taixie/zhen/ba/shan/,/ wanzhang/danti/shang/ke/***pan***/。 King/pavilion/pillow/ba/mountain/ ,/one hundred thousand feet/ purple ladder/still/able/***climb***/	Acceptable	Congruous
SEM+ RHY−	君王/台榭/枕/巴/山/,/万丈/丹梯/尚/可/登/。 Junwang/taixie/zhen/ba/shan/,/ wanzhang/danti/shang/ke/***deng***/。 King/pavilion/pillow/ba/mountain/ ,/one hundred thousand feet/ purple ladder/still/able/***clamber***/	Abnormal	Congruous
SEM− RHY+	君王/台榭/枕/巴/山/,/万丈/丹梯/尚/可/参/。 Junwang/taixie/zhen/ba/shan/,/ wanzhang/danti/shang/ke/***can***/。 King/pavilion/pillow/ba/mountain/ ,/one hundred thousand feet/ purple ladder/still/able/***participate***/	Acceptable	Incongruous
SEM− RHY−	君王/台榭/枕/巴/山/,/万丈/丹梯/尚/可/窥/。 Junwang/taixie/zhen/ba/shan/,/ wanzhang/danti/shang/ke/***kui***/。 King/pavilion/pillow/ba/mountain/ ,/one hundred thousand feet/ purple ladder/still/able/***peep***/	Abnormal	Incongruous

Note: Exemplar poems are given in Chinese Pinyin and English literal glosses. The critical characters are in bold and italic. Poems are segmented by "/" between two segments. The approximate English translation of this poem is *Of the royal pavilion mountain ba is the seat, onto which the one hundred thousand feet purple ladder can lead*.

RHY+, and SEM− RHY−), the final characters in the second lines were modified (see also pre-testing of materials in the next section). In this way, a total of 592 items were generated (see Table 1 for examples). Using a Latin square counterbalancing scheme, four presentation lists were created so that each list contained 37 items per condition, with no item appearing more than once per list. One hundred and forty-eight filler poems of various types were also added to each list. None of the filler items were seven-character Tang-dynasty poem. Fillers were included for two reasons. First, it allowed us to provide a range of types of Chinese classical poems with different characteristics. This reduced the predictability of the experimental items and increased the naturalness of the stimulus set. Second, the filler poems were used to reduce response biases that might have occurred if there were an unequal number of trials requiring positive and negative responses (e.g., Hsu, Tsai, Lee, & Tzeng, 2009; Zhang, Zhang, & Kong, 2009).

The four sets of target characters (e.g., '攀 *pan*', '登 *deng*', '参 *can*', and '窥 *kui*' in Table 1) were matched in stroke number and in lexical frequency based on the Modern Chinese Frequency Dictionary (Institute of language teaching & research, 1986). The mean frequencies were: SEM+ RHY+: 735 ($SD = 136$), SEM+ RHY−: 512 ($SD = 108$), SEM− RHY+: 503 ($SD = 189$), and SEM− RHY−: 404 ($SD = 141$) occurrences per million, respectively. Repeated-measure ANOVAs revealed that there was no significant main effect of frequency for rhyme scheme ($F(1, 147) = 2.76$, $p = 0.10$), semantic congruency ($F(1, 147) = 2.01$, $p = 0.16$), or their interaction ($F(1, 147) = 0.34$, $p = 0.56$). The mean number of strokes per character for the four different conditions was 8.97 ($SD = 3.4$) for SEM+ RHY+, 8.64 ($SD = 3.2$) for SEM+ RHY−, 8.82 ($SD = 3.65$) for SEM− RHY+, and 8.6 ($SD = 3.34$) for SEM− RHY−, respectively. Again, the analysis showed that neither the main effects of rhyme scheme and semantics nor the interaction of these variables approached significance, $Fs < 1.30$, $ps > 0.25$. In each poem, the final characters of the second lines in the four conditions were balanced to eliminate any possible effects of ambiguity of the word category (e.g., Rodd, Davis, & Johnsrude, 2005) and the effect of tone (e.g., Krishnan, Gandour, & Bidelman, 2010). Final characters all had the same spatial structure (e.g., left-right, top-down, etc.).

2.3. Pre-testing of materials

Twenty-five participants who did not participate in main experiment were asked to rate the familiarity and difficulty of the experimental poems, each on a scale ranging from 1 to 5. Participants were instructed to assign a rating of 1 to poems that were "completely unfamiliar" or "very easy", respectively, and to assign a rating of 5 to poems that were "completely familiar" or "very difficult", respectively. The original poems had a mean familiarity rating of 1.58 ($SD = 0.52$) and mean difficulty rating of 2.23 ($SD = 0.47$), respectively.

Another thirty participants rated the semantic plausibility of the final characters in the second line of each poem for each condition. Participants were instructed to assign a rating of 1 to the critical character in the second line if it was "completely unacceptable" and to assign a rating of 5 to the critical character in the second line if it was "completely acceptable". The average rating results were as follows: SEM+ RHY+: $M = 4.27$, $SD = 0.22$, SEM+ RHY−: $M = 4.20$, $SD = 0.29$, SEM− RHY+: $M = 1.61$, $SD = 0.22$, and SEM− RHY−: $M = 1.54$, $SD = 0.21$. The analysis revealed a main effect of semantic congruency, with higher semantic plausibility ratings for the semantically congruent character, $F(1, 29) = 3295.62$, $p < 0.001$. Although the semantic plausibility of the character with acceptable rhyme scheme was numerically higher than the character with the abnormal rhyme scheme, the difference was not significant, $F(1, 29) = 2.41$, $p = 0.13$. The interaction of these variables was also not significant, $F(1, 29) = 0.02$, $p = 0.89$.

To assess the relative strength of semantic and rhyme scheme expectations, a written poem line completion study was performed, with thirty participants being instructed to write a completion for each second line with the first character coming to mind (e.g., Table 1 '君王台榭枕巴山, 万丈丹梯尚可_。'). The mean percentage of completions rates for the four conditions ordered from highest to lowest were: SEM+ RHY+ ($M = 37.16\%$, $SD = 25.54\%$) > SEM− RHY+ ($M = 28.48\%$, $SD = 17.67\%$) > SEM− RHY− ($M = 18.31\%$, $SD = 14.40\%$) > SEM+ RHY− ($M = 16.05\%$, $SD = 14.18\%$). There was a significant main effect of rhyme scheme, $F(1, 147) = 107.94$, $p < 0.001$. As illustrated in Fig. 1, the interaction between the rhyme scheme and semantics was also significant, $F(1, 147) = 11.41$, $p = 0.001$. Post-hoc analysis revealed that the percentage of SEM+ RHY+ characters was higher than the percentage of SEM+ RHY− characters, $F(1, 147) = 62.36$, $p < 0.001$. The percentage of SEM− RHY− character was lower than the percentage of SEM− RHY+ characters, $F(1, 147) = 39.25$, $p < 0.001$. However, the main effect of semantics and the difference between SEM+ RHY− and SEM− RHY− characters was not significant, $Fs < 2.36$, $ps > 0.12$. In sum, readers strongly expect that the second line will end with a word that fits with the acceptable rhyme scheme, with semantic congruency being less important.

2.4. Procedure

The ERP experiment was conducted in a sound proof, electrically shielded room. Participants were seated in a comfortable chair with a viewing distance of approximately 100 cm and instructed to read the poems silently. At this distance, each Chinese character subtended 1° of the visual angle. All these poems were

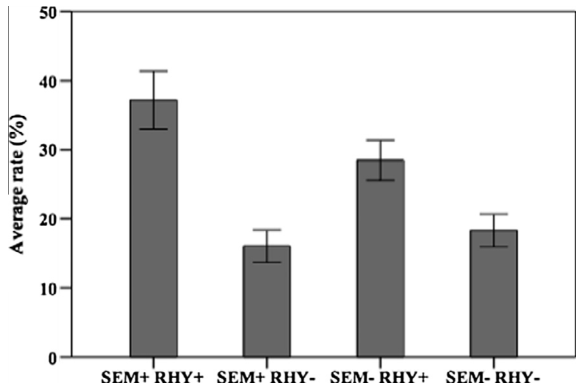

Fig. 1. Percentage of occurrence per condition. Error bars show standard errors.

presented in a pseudorandom order and poems from the same condition were presented in no more than two consecutive trials.

The poems were presented segment by segment at the center of a 17-in. computer screen with a 1024 × 768 pixel resolution. Segments are marked for the sample materials in Table 1. On each trial, a fixation cross appeared at the center of the computer screen for 1000 ms and was immediately replaced by the first segment of the sentence. Each segment was presented for 600 ms with an inter-stimulus interval (ISI) of 400 ms between every two segments. However, the rhyming character in each poem line (e.g., '山 *shan*' and '攀 *pan*' in Table 1) was presented for 600 ms and the interval between this character and the punctuation (e.g., commas and period) was 600 ms. The separate presentation of the punctuation was to reduce the effect of sentence-final wrap-up processes on the ERP responses to the character (Just & Carpenter, 1980; Luo & Zhou, 2010). Participants were instructed to press the "Z" key or "M" key when a question appeared on the computer screen to rate whether the poem was expressed in a natural way or whether it was semantically acceptable. The 1000 ms delayed response procedure was used in order to avoid contamination of the ERP components corresponding to the critical character from being affected by motor potentials coming from response preparation and execution. The assignment of the "yes" or "no" response to the "Z" key or "M" key was counterbalanced across participants. Participants completed ten practice trials before the main experimental block. The entire experiment lasted approximately 2 h.

2.5. EEG recordings

EEG was continuously recorded from 64 cap-mounted Ag/AgCl electrodes (Brain Products GmbH, Munich, Germany) mounted on an elastic cap that were located in the Standard International 10–20 System (Jasper, 1958). The vertical electroculogram (VEOG) was recorded from an electrode placed to the supraorbital region of the right eye. The horizontal electroculogram (HEOG) was recorded from an electrode attached to the outer canthi of the left eye. The EEGs were referenced online to the tip of the nose and were re-referenced offline to the average of the left and right mastoids. Electrode impedances were kept below 5 kΩ. Electrophysiological signals were amplified with a band pass 0.016–70 Hz via BrainAmps (Brain Products GmbH, Munich, Germany) and were digitized at 500 Hz.

2.6. EEG data analysis

EEG data were analyzed using the Brain Vision Analyzer 2.0 (Brain Products GmBH, Gilching, Munich). ERPs were computed for each participant over an epoch from 200 ms before to 1200 ms after the onset of the critical Chinese characters, and the average of the data in the 200-ms duration before the onset served as the baseline. Only correctly answered trials were analyzed. Furthermore, trials contaminated by EOG, muscle artifacts, amplifier saturation, or other artifacts were excluded by using a semiautomatic artifact rejection (automatic criterion: signal amplitude exceeding ±70 μV, afterwards followed by a manual check).

On the basis of visual inspection of the ERP waveforms and consecutive analyses of 50 ms time windows, the following four time windows were chosen for the statistical analyses of the ERP: 0–100, 100–300, 300–500, and 500–1000 ms. Repeated-measure analysis of variances (ANOVAs) were conducted on the mean amplitude of the ERP components elicited by the critical character in the second poem line (e.g., '攀 *pan*' in Table 1). The electrodes in the middle and lateral lines were computed, separately. The repeated measure ANOVA has three factors in the middle line analysis, which included Rhyme scheme acceptability (RHY+ and RHY−), Semantic Congruity (SEM+ and SEM−), and Electrodes (Fz, FCz, Cz, CPz, Pz, POz), and four factors in the lateral line analysis, including Rhyme scheme acceptability, Semantic Congruity, Hemisphere (left and right), and Region (anterior, central and posterior). Lateral electrodes were organized into six regions of interest (ROI), each having five or six representative electrodes: left anterior (F1, F3, F5, FC1, FC3, FC5), left central (C1, C3, C5, CP1, CP3, CP5), left posterior (P1, P3, P5, PO3, PO7), right anterior (F2, F4, F6, FC2, FC4, FC6), right central (C2, C4, C6, CP2, CP4, CP6) and right posterior (P2, P4, P6, PO4, PO8). All significant interaction effects were followed by post hoc simple effect comparisons. The Greenhouse–Geisser correction was performed when required (Greenhouse & Geisser, 1959). To control for multiple comparisons, *p*-values were adjusted using the Bonferroni correction (Keppel, 1982).

3. Results

3.1. Behavioral results

Average accuracy rates per condition were as follows: 87.34% for the SEM+ RHY+ condition (SD = 8.06%), 79.09% for the SEM+ RHY− condition (SD = 11.14%), 69.7% for the SEM− RHY+ condition (SD = 16.33%) and 80.23% for the SEM− RHY− condition (SD = 13.36%). An omnibus ANOVA, including Rhyme scheme and Semantics as factors, was computed on the percentage of correct answers. The results showed that there was a significant main effect of Semantics, $F\,(1, 18) = 3.98$, $p < 0.05$, $\eta_p^2 = 0.18$, indicating participants made more errors in responding to semantically incongruent lines than to congruent lines. More importantly, a significant interaction was found between Rhyme scheme and Semantics, $F\,(1, 18) = 23.69$, $p < 0.001$, $\eta_p^2 = 0.57$. Results of post hoc tests showed that the mean number of correct response for the SEM− RHY+ condition was significantly lower than the SEM+ RHY+ condition, $t\,(18) = -4.2$, $p < 0.01$, and SEM− RHY− condition, $t\,(18) = -3.63$, $p < 0.01$. However, there was no significant difference between the SEM+ RHY− and SEM− RHY− conditions, $t\,(18) = -0.23$, $p = 0.82$. In addition, the accuracy rate in the SEM+ RHY− condition was numerically higher than the SEM− RHY+ condition, but this difference did not reach significance, $t\,(18) = 1.68$, $p > 0.2$.

3.2. ERP results

ERPs time-locked to the critical character are displayed in Fig. 2 for the four experimental conditions. ANOVAs were conducted on the mean amplitude of the ERP in the selected time windows.

The analysis showed that there were neither significant main effects nor interactions in the mean amplitudes for either the 0- to 100-ms (*F*s < 1.7, *p*s > 0.15) or the 100- to 300-ms time window (*F*s < 2.9, *p*s > 0.11).

3.2.1. N1 and P200 analyses

We measured the N1 and P200 peaks and used a ±20 ms time window around those peaks.[1] The analysis indicated that neither the main effects (*F*s < 2.43, *p*s > 0.10) nor interactions (*F*s < 2.58, *p*s > 0.10) were significant in the 85 to 125-ms time window. Then, based on the previous relevant literature (e.g., Hagoort & Kutas, 1995, for an overview of language related ERP components, p. 110) and visual inspection of average ERP waveforms (see Figs. 2 and 3), the 150- to 250-ms time window was chosen for statistical analyses of the P200 component. There was a significant main effect of Rhyme congruity [$F(1, 18) = 22.32$, $p < 0.001$, $\eta_p^2 = 0.55$] for midline analysis and $F(1, 18) = 23.02$, $p < 0.001$, $\eta_p^2 = 0.56$ for lateral analysis. Neither the main effect of Semantic congruity nor the interaction of Rhyme congruity and Semantic congruity approached significance, *F*s < 0.36, *p*s > 0.14. Although the interaction between Hemisphere and Region was marginally significant [$F(2, 36) = 3.66$, $p = 0.06$, $\eta_p^2 = 0.17$], follow-up comparisons did not reveal any significant differences in Hemisphere or Region, *F*s < 2.95, *p*s > 0.10. Additionally, no other main effects or interactions were significant, *F*s < 2.27, *p*s > 0.14.

In sum, then, there was a clear effect of Rhyme congruity on the P200 component, with the P200 component being larger for characters that violated the Rhyme scheme expectancy. In contrast, there were no reliable effects of Semantic congruity on the P200. This is the predicted pattern if rhyme expectations influence early (pre-lexical) phonological processes in visual word recognition.

3.2.2. N400-like and N400 analyses

There was a significant main effect of Rhyme congruity in the 300- to 500-ms time window (see Fig. 2). The characters that violated expectations based on the rhyme scheme elicited larger negative components than the characters that matched the rhyme scheme, $F(1, 18) = 9.2$, $p = 0.007$, $\eta_p^2 = 0.34$ for midline electrodes ($d = -1.09$ μV), and $F(1, 18) = 6.92$, $p = 0.02$, $\eta_p^2 = 0.28$ for lateral electrodes ($d = -0.73$ μV). Moreover, the interaction between the Rhyme scheme and Hemisphere was significant, $F(1,18) = 5.05$, $p = 0.04$, $\eta_p^2 = 0.22$. Further analyses showed that the main effect of Rhyme scheme was only significant over the right hemisphere ($d = -0.93$ μV), $F(1,18) = 8.84$, $p = 0.008$, $\eta_p^2 = 0.33$, with its maximum in the right posterior regions ($d = -1.15$ μV). This pattern corresponds to the expected N400-like effects of Rhyme scheme expectation.

There was also a significant main effect of semantic congruency, $F(1, 18) = 12.7$, $p = 0.002$, $\eta_p^2 = 0.41$ for midline sites ($d = -1.04$ μV) and $F(1, 18) = 14.09$, $p = 0.001$, $\eta_p^2 = 0.44$ for lateral sites ($d = -0.96$ μV). The characters in the semantically incongruous poem lines elicited more negative responses than the semantically congruous characters. This interaction of this effect was marginally significant with Electrode in the midline analysis, $F(5, 90) = 3.17$, $p = 0.06$, $\eta_p^2 = 0.15$, and with Region in the lateral analysis, $F(2, 36) = 3.18$, $p = 0.08$, $\eta_p^2 = 0.15$. Further comparisons revealed that the effect of semantic congruency was significant on Cz, CPz, Pz, and POz in the midline sites (*F*s > 13, *p*s < 0.003) and in the central and posterior regions (*F*s > 17, *p*s < 0.001).

[1] Following one reviewer's suggestion, we also calculated the N1 in a ±15 ms time window around its peaks. There were no significant main effects and interactions, *F*s < 2.25, *p*s > 0.13.

3.2.3. Interactions between rhyme scheme and semantic congruency

In the 300- to 500-ms time range, the interaction between Rhyme congruity and Semantic congruity was significant both at the midline electrodes (see Figs. 2 and 4), $F(1, 18) = 4.67$, $p < 0.05$, $\eta_p^2 = 0.21$, and the lateral electrodes, $F(1, 18) = 4.98$, $p < 0.05$, $\eta_p^2 = 0.22$. There was a larger semantic congruency effect when the final character was consistent with the Rhyme scheme ($d = -1.7$ μV for the midline electrodes and $d = -1.46$ μV for the lateral electrodes) compared to when the final character was inconsistent with the Rhyme scheme ($d = 0.44$ μV for the midline electrodes and $d = -0.46$ μV for the lateral electrodes). Semantic incongruence resulted in a larger negative component compared to semantic congruence when participants processed the Rhyme congruent character, $F(1, 18) = 12.8$, $p < 0.01$, $\eta_p^2 = 0.42$ in the midline analysis, and $F(1, 18) = 13.11$, $p < 0.01$, $\eta_p^2 = 0.42$ in the lateral analysis. However, there was no significant difference between the semantically congruent and incongruent characters when the character violated the expected Rhyme scheme in the midline and lateral analyses, *F*s < 3.2, *p*s > 0.1. In the poems with semantically congruent final words, thyme scheme violations elicited a larger negative component than poems that were consistent with the rhyme scheme both at midline and lateral electrodes, $F(1, 18) = 16.31$, $p = 0.001$, $\eta_p^2 = 0.48$, and $F(1, 18) = 14.3$, $p = 0.001$, $\eta_p^2 = 0.44$, respectively.

In sum, then, semantic congruency effects on the N400 were observed only for the characters that were consistent with the expected Rhyme scheme. This is consistent with predictions we derived from the genre-specific hypothesis.

3.2.4. Later positivity analyses

A main effect of Semantic congruity was detected in the 500- to 1000-ms time window (see Fig. 2) in the midline analysis ($d = 1.15$ μV), $F(1, 18) = 20.95$, $p < 0.001$, $\eta_p^2 = 0.54$, and in the lateral analysis ($d = 0.54$ μV), $F(1, 18) = 6.35$, $p = 0.021$, $\eta_p^2 = 0.26$. As illustrated in Fig. 2, semantically incongruous characters elicited larger positive components than semantically congruous characters at both midline and lateral electrodes, showing a late positivity effect. There was neither a significant main effect (*F*s < 0.47, *p*s > 0.5) nor an interaction with Rhyme scheme in this time window (*F*s < 2.7, *p*s > 0.1).

4. Discussion

Using classic Chinese poems to create expectations based on rhyme scheme, we used ERPs to examine the time-course of rhyme congruency and semantic congruency effects. A violation of the rhyme scheme expectation produced more positive amplitude of the P200 component across the head, with its maximum in the left frontal sites, and with no effects of semantic congruency. Rhyme scheme incongruent characters also induced increased negativity in the 300–500 ms time window that was mainly distributed in the right hemisphere. Semantically incongruent characters elicited an increased negativity in this time window, with its maximum in the central and posterior regions—a distribution that is consistent with that of the classic N400. However, the semantic effect was modulated by rhyme scheme expectation. When the poem had an acceptable rhyme scheme, semantically incongruent characters elicited a larger negative response relative to semantically congruent characters. However, there was no difference between these two conditions when the poem had an abnormal rhyme scheme, suggesting that rhyme scheme modulated semantic processing during online poem comprehension. Additionally, there was a larger late positive component in the 500- to 1000-ms time window in semantically incongruent poems than in semantically congruent

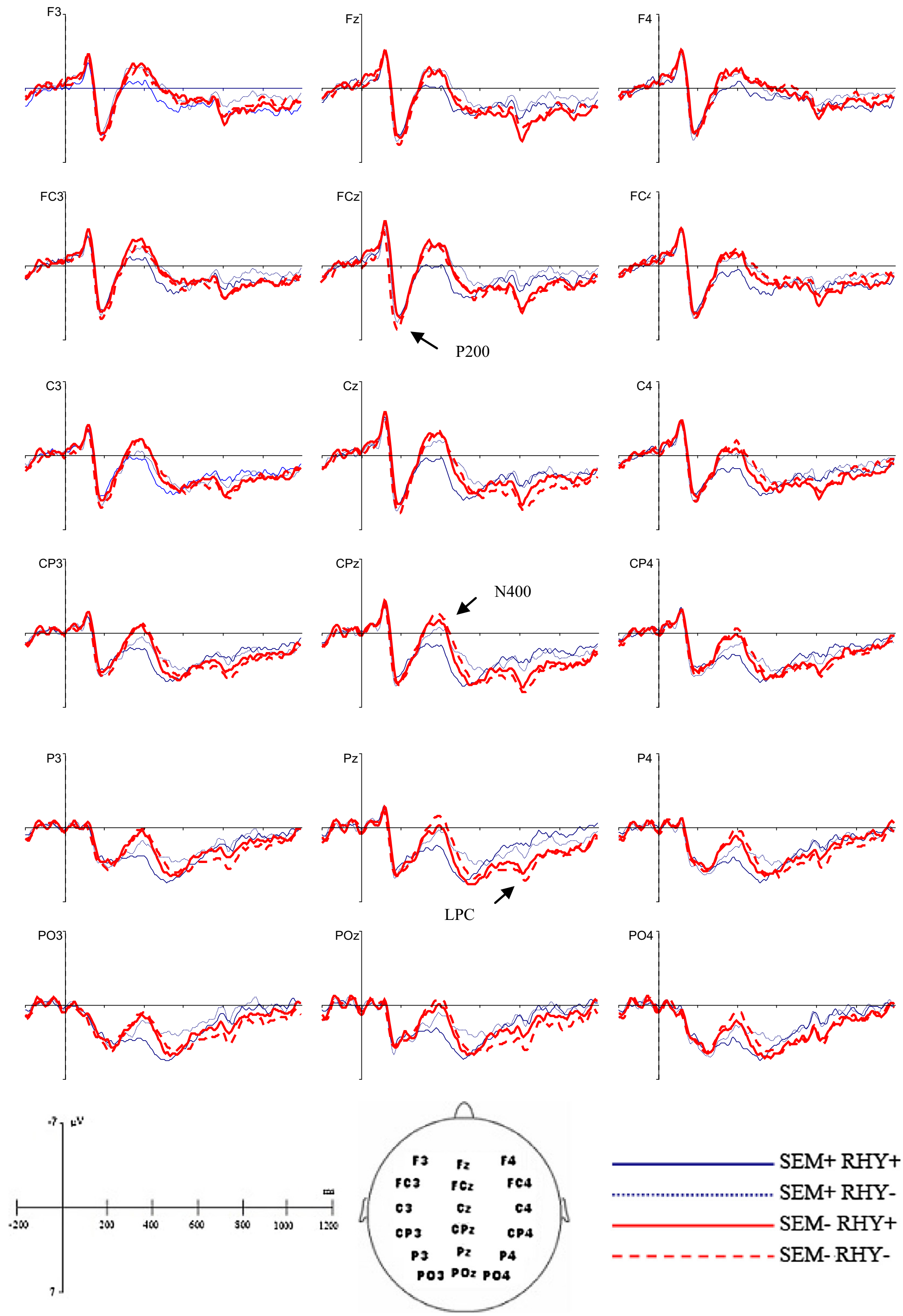

Fig. 2. Grand average ERP in response to the final character in the second poem line from 18 representative electrodes over the 0–1200 ms time window. The onset of the critical character is aligned to the zero in the time line. SEM+ RHY+ = semantically congruent with acceptable rhyme scheme; SEM+ RHY− = semantically congruent with unacceptable rhyme scheme; SEM− RHY+ = semantically incongruent with acceptable rhyme scheme; SEM− RHY− = semantically incongruent with unacceptable rhyme scheme.

poems, suggesting that the second-pass processing (i.e., re-computing) was more demanding for the incongruent poem line than for the congruent poem line. We discuss each of the results in turn.

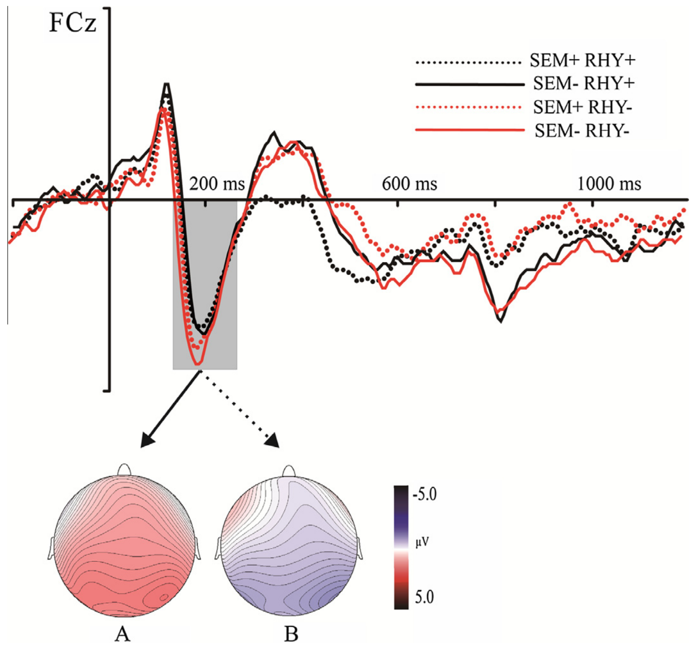

Fig. 3. P200 component and topographical distributions of ERP differences in the 150- to 250-ms time window. Top: The P200 component at FCz electrode. Bottom: (A) The effect of Rhyme scheme: unacceptable rhyme scheme minus acceptable rhyme scheme. (B) The effect of Semantic congruity: incongruous semantics minus congruous semantics.

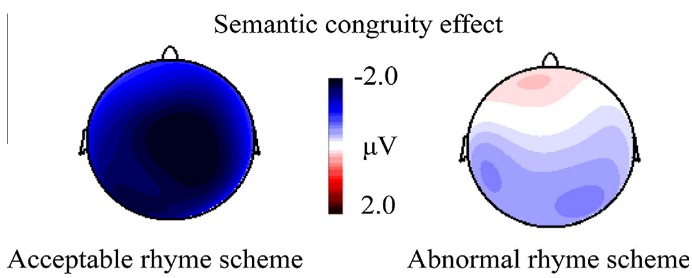

Fig. 4. Impact of rhyme scheme on semantics. Topographic distributions of semantic congruous effect in the acceptable and abnormal rhyme schemes. Mean amplitude differences between semantically incongruous and congruous characters in the 300-to- 500-ms time window.

4.1. Rhyme scheme expectation and P200

The most important result to emerge from this study is that rhyme expectations affected P200, an ERP component that has been linked to the mapping of a character onto a phonological representation. The fact that P200 was modulated by rhyme scheme expectations but not semantic congruity suggests that top-down prosodic expectations can modulate early perceptual processes before lexical-semantic processing during silent reading. One plausible mechanism is that expectations facilitate assigning a syllabic structure to the character and therefore reduces the uncertainty of the phonological features.

To the best of our knowledge, our study is the first to demonstrate early effects of top-down sound-based expectations in the processing stream for silent reading. It is important to note that we were able to find clear evidence for top-down prosodic expectations because we used classic Chinese poems to create expectations based on rhyme scheme, in the absence of expectations for a particular word or character, thus ruling out an explanation that the rhyme expectation effects were orthographically mediated. This finding is surprising given many theoretical perspectives about how phonology is computed in visual word recognition, and more generally how prosody might be used in silent reading.

Furthermore, this result raises a number of important questions for future research. One question is whether early use of prosodic expectations in silent reading is a special property of character-based orthographies because prosody might play a similar role in segmentation when the orthography does not clearly mark word boundaries. Evaluating this possibility will require further research comparing rhyme expectation effects in languages with alphabetic and character-based orthographies. One possible avenue of research would be to use ERPs to examine how rhyme expectation might modulate feedback consistency effects (Lee et al., 2015; Ziegler, Petrova, & Ferrand, 2008).

One intriguing possibility is that assigning prosodic representations in silent reading operates in parallel with more content-based processing. As sound-based representations become available prosodic representations are constructed and are part of the merging representations constructed in silent reading.

4.2. Right-lateralized rhyme scheme congruity effects

One possible source of evidence for parallel processing streams is that also observed Rhyme congruence effects in the 300–500 ms time window. This was reflected in right-lateralized negative component with timing similar to that of classic semantic and lexical effects on N400 (Kutas & Federmeier, 2011). Similar effects have been reported in previous studies using rhyme-judgment and rhyme-priming tasks (e.g., Coch et al., 2005; Rothermich, Schmidt-Kassow, & Kotz, 2012; Rugg, 1984a, 1984b). When rhyme expectations are violated, they may generate "error signals" that contain information about the difference between what was predicted and the syllabic structure that becomes available as a consequence of lexical access.(e.g., Chang, Dell, & Bock, 2006; Farmer et al., 2013, 2015; Jaeger & Snider, 2013). Crucially, these rhyme effects were also distinct from any effects of semantic congruence. Whether or not the early (P200) and later rhyme effects reflect different processes remains a question for future research.

4.3. Prosodic-semantic interaction in the N400

In addition to these rhyme-based effects we found that the more classic N400 effect (e.g., Holcomb & Neville, 1991; Kutas & Hillyard, 1980), which is generally considered to reflect semantic processing, was modulated by rhyme scheme congruency. One way of understanding these effects is that semantic incongruity modulated the N400 effect during the reading of poem lines with plausible rhyme scheme. However, content-based semantic incongruity did not evoke a typical N400 effect in poem lines which had an abnormal rhyme scheme.

The present interaction suggests that what is considered to be meaning-based processing may depend in part upon the genre of a text. This is important because a preponderance of the data we have in language processing comes from restricted linguistic genres. Varying the type of language studied in psycholinguistic experiments can sometimes highlight important aspects of language processing that are otherwise ignored. For example, a large body of work assumed taking into account an interlocutor's perspective, involves processing with respect to common ground. However, this research used declarative sentences, often framed as commands (e.g., Keysar, Barr, Balin, & Brauner, 2000). Taking perspective in processing information questions involves attending to information that is privileged and not information that is shared. Listeners are adept at doing so during real-time processing in cooperative tasks (e.g., Brown-Schmidt, Gunlogson, & Tanenhaus, 2008).

In declarative sentences, particularly those used in most psycholinguistic experiments examining reading, congruity of content (i.e., plausibility) is a reasonable stand-in for semantic processing. However, in other literary genre's sound and meaning might be much more tightly bound together, a proposal argued for in genre-specific approaches to comprehension (Culler, 1975; Jakobson, 1960). In poem reading, it has been argued that typical prosodic features (e.g., rhyme scheme) likely attract more attention and are more strongly activated than other information (e.g., Carminati, Stabler, Roberts, & Fischer, 2006; Hanauer, 1996; Obermeier et al., 2013; Tillmann & Dowling, 2007; Yaron, 2002). Rhyme scheme likely induces an expectation of coherence (e.g., similarities, regularities and repetitions) that modulates semantic processing. Moreover, similar sounds may be taken as having similar meanings or connotative associations. According to this perspective, any character that violates an obligatory rhyme scheme would be semantically incongruous. Content-based incongruity effects might, then, only occur when the rhyme scheme suggests that the character can be integrated into the poem, leading to classic N400 effects (see Jacobs, 2014 for a neurocognitive model of poetry reading).

4.4. Semantic congruity effects in the 500–1000 ms time interval

These late effects of semantic congruity are similar to that found in some previous studies in which semantic manipulations led to larger N400 accompanied by larger late positive components (i.e., LPC) in several experimental tasks, such as congruent versus incongruent sentence completion and translation recognition tasks (e.g., Thornhill & Van Petten, 2012; Van Petten & Luka, 2012). Based on the current behavioral and electrophysiological findings, we think that the LPC likely reflects cognitive effort spent on attempting to check and re-compute some aspects of the unusual poems (cf. Kolk & Chwilla, 2007; Salmon & Pratt, 2012).[2] For example, when both the rhyme and meaning were incongruent, participants might have increased uncertainty about whether they had correctly read the critical character in the previous line (Levy, Bicknell, Slattery, & Rayner, 2009).

5. Conclusions

By using classic poems to create a prosodic expectation for a particular syllabic structure, we demonstrated that rhyme scheme expectations affect processes manifested by the P200, which has previously been argued to reflect orthographic to phonological mapping, during an early "pre-lexical" process in silent reading. Top-down expectations affected the P200 even though the rhyme scheme expectations did not map onto a particular orthography (character). These findings shed light on the interplay between top-down expectation and bottom-up processes in pre-lexical time windows. We believe that they open the door to a new class of questions about the role of prosodic expectations in silent reading, including some potentially important differences in how prosody might be used differently in silent reading in languages with alphabetic and (deep) character-based orthographies.

We also observed clear effects of rhyme scheme congruency on a right-lateralized N400-like component. Moreover, semantically anomalous poems only induced a significant N400 effect in the acceptable rhyme scheme condition, suggesting that content-based semantic processing might not be dissociable from sound-based coherence in poetry reading. This result suggests that modulating expectations using different language genres may be a fruitful approach for exploring both novel and more classic issues in language processing.

Acknowledgements

This work was supported by grants from the Major State Basic Research Development Program of China (973 Program, 2014CB340502), the Major Program of the Social Science Foundation of China (10&ZD126), the Natural Science Foundation of China (31100814, 31271196, and 31300929), the Postdoctoral Science Foundation of China (2014M560433), the Postdoctoral Science Foundation of Jiangsu Province of China (1401061B). Michael K. Tanenhaus was supported by NIH grants HD 27206 and HD 073890.

References

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*, 247–264.
Arai, M., & Keller, F. (2013). The use of verb-specific information for prediction in sentence processing. *Language and Cognitive Processes, 28*, 525–560.

[2] A related explanation would emphasize that this LPC may be similar to the P600 effect which has been interpreted to be associated with task difficulty (e.g., Kim & Osterhout, 2005; see Kuperberg, 2007 for a review of neural mechanisms of language comprehension). This explanation is inconsistent with the findings from the behavioral and ERPs results. For example, SEM− RHY− characters induced larger LPC than SEM+ RHY− characters (see Fig. 2). However, the average accuracy rates showed that SEM+ RHY− characters were more difficult than SEM− RHY− characters. Additionally, if LPC is induced by task related difficulty, larger positivity should be evoked by the SEM− RHY+ characters compared to the SEM− RHY− characters, since the former task is more difficult than the later, as confirmed by the higher accuracy performance in the SEM− RHY− sentences than the SEM− RHY+ sentence. However, we failed to find differences between these two conditions on ERP amplitude.

Baggio, G., Van Lambalgen, M., & Hagoort, P. (2008). Computing and recomputing discourse models: An ERP study. *Journal of Memory and Language, 59*, 36–53.
Bicknell, K., Elman, J. L., Hare, M., McRae, K., & Kutas, M. (2010). Effects of event knowledge in processing verbal arguments. *Journal of Memory and Language, 63*, 489–505.
Breen, M., Dilley, L. C., McAuley, J. D., & Sanders, L. D. (2014). Auditory evoked potentials reveal early perceptual effects of distal prosody on speech. *Language, Cognition and Neuroscience, 29*, 1132–1146.
Brown, M., Salverda, A. P., Dilley, L. C., & Tanenhaus, M. K. (2011). Expectations from preceding prosody influence segmentation in online sentence processing. *Psychonomic Bulletin & Review, 18*, 1189–1196.
Brown, M., Salverda, A. P., Dilley, L. C., & Tanenhaus, M. K. (2015). Metrical expectations from preceding prosody influence perception of lexical stress. *Journal of Experimental Psychology: Human Perception and Performance, 41*, 306–323.
Brown-Schmidt, S., Gunlogson, C., & Tanenhaus, M. K. (2008). Addressees distinguish shared from private information when interpreting questions during interactive conversation. *Cognition, 107*, 1122–1134.
Carminati, M. N., Stabler, J., Roberts, A. M., & Fischer, M. H. (2006). Readers' responses to sub-genre and rhyme scheme in poetry. *Poetics, 34*, 204–218.
Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113*, 234–272.
Clifton, C. Jr., (2015). The roles of phonology in silent reading: A selective review. In L. Frazier & E. Gibson (Eds.), *Explicit and implicit prosody in sentence processing: Studies in honor of Janet dean Fodor* (pp. 161–176). New York: Springer.
Coch, D., Grossi, G., Skendzel, W., & Neville, H. (2005). ERP nonword rhyming effects in children and adults. *Journal of Cognitive Neuroscience, 17*, 168–182.
Coch, D., Hart, T., & Mitra, P. (2008). Three kinds of rhymes: An ERP study. *Brain and Language, 104*, 230–243.
Culler, J. (1975). *Structuralist poetics: Structuralism, linguistics and the study of literature*.London: Routledge.
Dahan, D. (2015). Prosody and language comprehension. *Wiley Interdisciplinary Reviews: Cognitive Science, 6*, 441–452.
DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*, 1117–1121.
Dikker, S., & Pylkkanen, L. (2011). Before the N400: Effects of lexical-semantic violations in visual cortex. *Brain and Language, 118*, 23–28.
Fang, S. P., Horng, R. Y., & Tzeng, O. J. L. (1986). Consistency effects in the Chinese character and pseudo-character naming tasks. In H. S. R. Kao & R. Hoosian (Eds.), *Linguistics, psychology, and the Chinese language* (pp. 11–21). Hong Kong: Center of Asia Studies.
Farmer, T. A., Brown, M., & Tanenhaus, M. K. (2013). Prediction, explanation, and the role of generative models in language processing. *Behavioral and Brain Sciences, 36*, 211–212.
Farmer, T. A., Christiansen, M. H., & Monaghan, P. (2006). Phonological typicality influences on-line sentence comprehension. *Proceedings of the national academy of sciences of the United States of America, 103*, 12203–12208.
Farmer, T. A., Yan, S., Bicknell, K., & Tanenhaus, M. K. (2015). Form-to-expectation matching effects on first-pass eye movement measures during reading. *Journal of Experimental Psychology: Human Perception and Performance, 41*, 958–976.
Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44*, 491–505.
Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language, 41*, 469–495.
Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*, 95–112.
Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science, 304*, 438–441.
Hagoort, P., & Kutas, M. (1995). Electrophysiological insights into language deficits. In F. Paller & J. Grafman (Eds.). *Handbook of neuropsychology* (Vol. 10, pp. 105–134). Amsterdam: Elsevier.
Hanauer, D. (1996). Integration of phonetic and graphic features in poetic text categorization judgments. *Poetics, 23*, 363–380.
Hanauer, D. (1998). The genre-specific hypothesis of reading: Reading poetry and encyclopedic items. *Poetics, 26*, 63–80.
Holcomb, P. J., & Neville, H. J. (1991). Nature speech processing: an analysis using event-related brain potentials. *Psychobiology, 19*, 286–300.
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 potentials study. *Brain and Language, 108*, 56–66.
Huang, X. J., Yang, J. C., Zhang, Q., & Guo, C. Y. (2014). The time of course of spoken word recognition in Mandarin Chinese: A unimodal ERP study. *Neuropsychologia, 63*, 165–174.
Hue, C. W. (1992). Recognition processes in character naming. In H. C. Chen & O. J. L. Tzeng (Eds.), *Language processing in Chinese* (pp. 93–107). Amsterdam: North-Holland.
Institute of language teaching and research (1986). *Frequency dictionary of modern Chinese*.Beijing: Beijing Language Institute.
Jacobs, A. M. (2014). Towards a neurocognitive poetics model of literary reading. In R. M. Willems (Ed.), *Towards a cognitive neuroscience of natural language use* (pp. 135–159). Cambridge: Cambridge University Press.
Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition, 127*, 57–83.
Jakobson, R. (1960). Linguistics and poetics. In T. Sebeok (Ed.), *Style in language* (pp. 350–377). Cambridge, MA: MIT Press.
Jasper, H. (1958). The ten-twenty electrode system of the international Federation. *Electroencephalography and Clinical Neurophysiology, 10*, 371–375.
Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329–354.
Kamide, Y. (2008). Anticipatory processes in sentence processing. *Language and Linguistics Compass, 2*, 647–670.
Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49*, 133–159.
Keppel, G. (1982). *Design and analysis*.Englewood Cliffs, NJ: Prentice-Hall.
Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Sciences, 11*, 32–38.
Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language, 52*, 205–225.
Kimball, J. (1975). Predictive analysis and over-the-top parsing. In J. Kimball (Ed.), *Syntax and semantics* (pp. 155–179). New York: Academic Press.
Kolk, H., & Chwilla, D. J. (2007). Late positivities in unusual situations. *Brain and Language, 100*, 257–261.
Krishnan, A., Gandour, J. T., & Bidelman, G. M. (2010). The effects of tone language experience on pitch processing in the brainstem. *Journal of Neurolinguistics, 23*, 81–95.
Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenge to syntax. *Brain Research, 1146*, 23–49.
Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition & Neuroscience, 31*, 32–59.
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*, 621–647.
Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*, 203–205.
Lee, C. Y., Hsu, C. H., Chang, Y. N., Chen, W. F., & Cho, P. C. (2015). The feedback consistency effect in Chinese character recognition: Evidence from a psycholinguistic norm. *Language and Linguistics, 16*, 535–554.
Lee, C. Y., Tsai, J. L., Chan, W. H., Hsu, C. H., Hung, D. L., & Tzeng, O. J. L. (2007). Temporal dynamics of the consistency effect in reading Chinese: An event-related potentials study. *NeuroReport, 18*, 147–151.
Lee, C. Y., Tsai, J. L., Su, E. C. I., Tzeng, O. J. L., & Hung, D. L. (2005). Consistency, regularity, and frequency effects in Naming Chinese characters. *Language and Linguistics, 6*, 75–107.
Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106*, 1126–1177.
Levy, R., Bicknell, K., Slattery, T., & Rayner, K. (2009). Eye movement evidences that readers maintain and act on uncertainty about past linguistic input. *Proceedings of the national academy of sciences of the United States of America, 106*, 21086–21090.
Li, T. (1983). Handedness questionnaire (Chinese version). *Acta Psychology, 15*, 268–276.
Luo, Y. Y., & Zhou, X. L. (2010). ERP evidence for the online processing of rhythmic pattern during Chinese sentence reading. *NeuroImage, 49*, 2836–2849.
Obermeier, C., Menninghaus, W., Von Koppenfels, M., Raettig, T., Schmidt-Kassow, M., Otterbein, S., & Kotz, S. (2013). Aesthetic and emotional effects of meter and rhyme in poetry. *Frontier in Psychology, 4*, 1–9.
Perfetti, C. A., Liu, Y., & Tan, L. H. (2005). The lexical consistency model: Some implications of research on Chinese for general theories of reading. *Psychological Review, 112*, 43–59.
Reinisch, E., Jesse, A., & McQueen, J. M. (2011). Speaking rate from proximal and distal contexts is used during word segmentation. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 978–996.
Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex, 15*, 1261–1269.
Rothermich, K., Schmidt-Kassow, M., & Kotz, S. A. (2012). Rhythm's gonna get you: Regular meter facilitates semantic sentence processing. *Neuropsychologia, 50*, 232–244.
Rugg, M. D. (1984a). Event-related potentials and the phonological processing of words and non-words. *Neuropsychologia, 22*, 435–443.
Rugg, M. D. (1984b). Event-related potentials in phonological matching tasks. *Brain and Language, 23*, 225–240.
Salmon, N., & Pratt, H. (2012). A comparison of sentence- and discourse-level semantic processing: An ERP study. *Brain and Language, 83*, 367–383.
Staub, A., & Clifton, C. Jr., (2006). Syntactic prediction in language comprehension: Evidence from either...or. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 425–436.
Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience, 2*, 191–196.
Tanenhaus, M. K., & Hare, M. (2007). Phonological typicality and sentence processing. *Trends in Cognitive Sciences, 11*, 93–95.

Thornhill, D. E., & Van Petten, C. (2012). Lexical versus conceptual anticipation during sentence processing: Frontal positivity and N400 ERP components. *International Journal of Psychophysiology, 83*, 382–392.

Tillmann, B., & Dowling, W. J. (2007). Memory decreases for prose, but not for poetry. *Memory & Cognition, 35*, 628–639.

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 443–467.

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology, 83*, 176–190.

Yaron, I. (2002). Processing of obscure poetic texts: Mechanisms of selection. *Journal of Literary Semantics, 31*, 133–170.

Zhang, Q., Zhang, J. X., & Kong, L. (2009). An ERP study on the time course of phonological and semantic activation in Chinese word recognition. *International Journal of Psychophysiology, 73*, 235–245.

Ziegler, J. C., Petrova, A., & Ferrand, L. (2008). Feedback consistency effects in visual and auditory word recognition: Where do we stand after more than a decade? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 643–661.