Neuroscience 168 (2010) 757–768

PERCEPTION OF CHINESE POEM AND ITS ELECTROPHYSIOLOGICAL EFFECTS

W. LI AND Y. YANG*

State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, 4A Datun Road, Chaoyang District, Beijing, 100101, PR China

Abstract—A neural correlate for phrase boundary perception in language has recently been identified as a reliable and replicable brain effect. It is called the closure positive shift (CPS) and has an equivalent in the perception of music (music CPS). Nevertheless, either in language or in music, this component is elicited by phrase boundary embedded in sentence or melody. Poetry, as the interlude of language and music, is a special kind of discourse and promising material to explore prosodic boundary processing beyond sentence level. The aim of the present study was to investigate the cognitive processing of hierarchical prosodic boundaries in Chinese Tang poem using rhythm matching task. There are generally four hierarchical levels in each poem, including foot boundary, phonological phrase boundary, intonational phrase boundary, and couplet boundary. The electrophysiological results indicated that all the prosodic boundaries of different levels in poems could give rise to the CPS reflecting prosodic phrasing. Furthermore, as the prosodic hierarchical level became higher, the onset latency of the CPS got longer, suggesting the influence of retrospective processing of former information. With regard to the amplitude, we analyzed the CPS amplitude in every 100 ms time window. It was showed that phonological phrase boundary elicited higher CPS amplitude as compared to that evoked by couplet boundary in an earlier time window, whereas in a later time window both of them were lower than the CPS correlated to intonational phrase boundary. The present results further shape our understanding of the CPS component and its relation to the processes involved in prosodic phrasing.

Key words: closure positive shift, foot boundary, phonological phrase boundary, intonational phrase boundary, couplet boundary.

During comprehension of a spoken sentence, listeners rely not only on semantic and syntactic information, but also on prosodic cues. Prosody indicates which element is stressed, and marks the grouping of the constituents (i.e. intonational phrases). The processing of intonational phrase boundaries (IPBs) has been found to correlate with a specific event-related potential (ERP) component, a bilateral, centroparietal positive deflection following the end of an intonational phrase called the closure positive shift (CPS) (Steinhauer et al., 1999). This component is also evoked by materials that were stripped of semantic and syntactic information, such as pseudoword sentences (Pannekamp et al., 2005), filtered speech materials (Steinhauer and Friederici, 2001) and hummed sentences (Pannekamp et al., 2005). These studies provide evidence that CPS is a reflection of purely prosodic processing, and dissociated from other segmental information (Friedrich et al., 2004; Pannekamp et al., 2005; Steinhauer, 2003). Recently, a study was conducted to explore the processing of hierarchical prosodic boundaries in Mandarin Chinese sentence. It was found that both IPB and phonological phrase boundary (PPB) in sentence could evoke CPS, whereas prosodic word boundary couldn't. Furthermore, PPB elicited CPS with earlier onset latency and somewhat lower amplitude as compared to the one evoked by IPB (Li and Yang, 2009). Besides, it has also been shown that this waveform is a reliable marker for prosodic phrasing during silent reading based on comma rules (Steinhauer, 2003; Steinhauer and Friederici, 2001), and the perception of musical phrase structure (Knösche et al., 2005; Nan et al., 2006; Neuhaus et al., 2006). Some differences, however, exist among variants of CPS. To be specific, the speech CPS lasts between approximately 500 and 1000 ms, with the mean amplitude varying between 3.4 and 4.6 μV, while the CPS elicited by commas has reduced amplitude (approximately 2 μV) and shorter duration (200–400 ms) (Steinhauer, 2003). For the CPS of music phrasing, it bears the same amplitude and topographic distribution as the language CPS but different latency (500 and 600 ms after the offset of a phrase boundary) and duration (100–200 ms) (Knösche et al., 2005).

One thing should be noted is that the studies introduced above and other relevant ones (Boecker et al., 1999; Friedrich et al., 2001; Schirmer et al., 2002) have mainly focused on the neural basis of prosodic processing within sentences. Nevertheless, either reading or verbal communication is mostly realized in the form of discourse instead of words or loose strings of single sentences. Thus, it is also necessary to explore the neurophysiological basis of prosodic processing beyond sentence level. It has been indicated that discourse can be segmented reliably (Nakatani et al., 1995; Swerts, 1997). In addition to other linguistic cues, such as lexical and syntactic ones, prosody is assumed to have a distinctive function in discourse segmentation, although the relationship between intonational features and discourse elements is very complex. Several decades of research have resulted in numerous findings on how discourse level meaning can be conveyed by acoustic-prosodic properties of speech, such as pitch range, pause duration, amplitude, speaking rate and intonational prominence (Nakatani et al., 1995). The most important and fully investigated

*Corresponding author. Tel: +86-10-64888629; fax: +86-10-64872070.
E-mail address: yangyf@psych.ac.cn (Y. Yang).

Abbreviations: ANOVA, analysis of variance; CB, couplet boundary; CPS, closure positive shift; EEG, electroencephalography; EOG, electrooculogram; ERP, event related potential; FB, foot boundary; IPB, intonational phrase boundary; PPB, phonological phrase boundary; TW, time window.

0306-4522/10 $ - see front matter
doi:10.1016/j.neuroscience.2010.03.069

prosodic means for structuring longer utterances are pitch range and pause duration. Larger-scale information units tend to be phonetically encoded by means of higher pitch reset values and longer pauses, while lower pitch reset values and shorter pauses appear between intonational phrases dealing with the same topic (Grosz and Hirschberg, 1992; Krivokapi, 2007; Hirschberg and Nakatani, 1996; Swerts, 1997; Swerts and Geluykens, 1993).

One question unsettled is that the neural basis of prosodic processing in discourse has been rarely studied in any detail. One reason may be due to the fact that the study of discourse prosody using neurophysiological methods is relatively complicated and hard to manipulate as compared to that of words or sentences. To the best of our knowledge, there are only two ERP studies concerning prosodic processing in discourse. In order to investigate the immediate interplay of syntactic and prosodic information, Kerkhofs and colleagues (2007) embedded target sentence in a discourse context, which induced the expectation of either the presence or absence of a syntactic break right at a prosodic break (Kerkhofs et al., 2007). They found that CPS was reduced in size when a prosodic break was aligned with a syntactic break than against each other. Although this study was conducted in discourse, the ERP signature (CPS) of prosodic processing was still elicited by the prosodic break embedded in the target sentence rather than a boundary at higher levels, such as the prosodic boundary at the end of a sentence or paragraph. This was also the case for another study using German dialogues (Toepel et al., 2007), in which CPS was also induced by prosodic boundaries inserted within the domain of sentence. Therefore, it is essential to study the neural basis of prosodic boundaries processing beyond sentence.

As one special kind of discourse, Chinese Tang poem[1] refers to poetry generated and flourished in Tang Dynasty (618–907 A.D.), the Golden Age of Chinese poetry. They are strict in the number of sentences in a poem and the number of characters in each sentence (Wang, 2005). The four tones of Chinese language are divided into "ping2" (level tone and rising tone) and "ze4" (entering tone and falling tone), which appear alternatively in each sentence of a poem. Due to these formal requirements, the deep structure of poetry is rich in prosodic hierarchies while its surface structure remains simple and terse. It seems promising to explore the neural mechanism of prosodic processing at discourse level using this kind of material.

For the present study, seven-character-quatrain was used, which is composed of four sentences and seven characters in each sentence. Each quatrain is composed of two couplets (paragraphs), and divided by a big break, namely the couplet boundary (CB). In each paragraph, an IPB[2] is inserted between the two sentences. It is obvious that CB ranks higher than IPB in prosodic hierarchy. Seven characters in a sentence are divided into two parts such that the last three characters are separated from the preceding four ones and the boundary between them is represented as PPB. The last three characters constitute an independent prosodic unit, termed "three-character-foot," which is a significant characteristic of Chinese poetry. The preceding four characters are composed of two foots, and the boundary between them is the foot boundary (FB), which was taken as the baseline condition in the present study. To sum up, there are four layers in the prosodic hierarchy of a quatrain: FB and PPB embedded in sentence, IPB between sentences in a couplet, and CB between the two couplets in a quatrain. To be clear, they are exemplified in the following poem, and numbers were used to mark the sentences/lines within the poem. Specifically, FBs in the four sentences of each poem were labeled as FB1, FB2, FB3 and FB4 respectively. Similarly, PPBs in four sentences were termed PPB1, PPB2, PPB3 and PPB4, while the prosodic boundaries at the end of the sentences within the two couplets were marked as IPB1 and IPB2.

Song lyric, generated in Song Dynasty (960–1127 A.D.), is another kind of poetry in China. In contrast with Tong poem, it has a very flexible structure (Wang, 2005). Specifically speaking, it is composed of lines of irregular lengths, and variable number of characters in each sentence (from three to seven characters). Therefore, it is hardly possible for listeners to anticipate how many sentences a lyric contains and how many characters a sentence has. Song lyrics were used as fillers in the present study.

Analyses of reading and spontaneous monologues have shown that discourse units are often demarcated prosodically (Swerts, 1997), and listeners are able to identify discourse boundaries of different strength on the basis of prosodic cues (Swerts, 1997; Wang et al., 2006). It is reasonable to expect that CPS which functions as segmenting incoming sentence or melody may also be evoked by prosodic boundaries in discourse. Hence, the goal of the present study is to investigate the cognitive processing of hierarchical prosodic boundaries in discourse, using seven-character-quatrain, a special kind of discourse as stimulus material. More specifically, we aimed to explore: (1) whether different prosodic boundaries in quatrains could evoke CPS; if they can, (2) whether prosodic boundaries at the same hierarchical level elicited the same brain effect irrelevant to their positions in a quatrain; and (3) whether CPS induced by different hierarchical prosodic boundaries will present systematic differences in their amplitude and latency. It is hypothesized that all prosodic boundaries in quatrains, irrespective of their hierarchical level, could induce CPS, and prosodic boundaries at the

[1] Chinese Tang poem is named according to its generation era. It can be sorted by its length and the number of characters included in each sentence. According to the length, they can be separated into jue2 ju4 (quatrain) including four sentences, lv4 shi1 including 8 sentences and pai2 lv4 including more than 10 sentences. According to the number of characters involved in each sentence, they are divided into five characters (wu3 yan2), six characters (liu4 yan2, this case is not common), and seven characters (qi1 yan2) separately.

[2] Intonational phrase boundary can be signaled as embedded in a sentence or in the end of a sentence. In the current study, the prosodic boundary in the end of a sentence was termed as IPB.

same level should elicit the same ERP effect. Moreover, we expect that as the prosodic hierarchical level becomes higher, the onset latency of the positive shifts may be delayed, while the amplitude will be stronger.

EXPERIMENTAL PROCEDURES

Participants

All subjects were native speakers of Mandarin Chinese without any visual or hearing disorders according to self-report. They all participated for reimbursement. Nineteen right-handed students (mean age 24.5 years, 10 males) took part in the first experiment, while 20 right-handed students (mean age 22.4 years, 10 males) were tested in the second experiment.

Stimuli

Sixty quatrains and 60 lyrics were read by a male speaker of standard Chinese according to his own understanding and recorded in a soundproof chamber, at a sampling rate of 22 kHz.

These quatrains were selected from a set of 100 quatrains through a pretest to ensure unfamiliarity. In the pretest the quatrains were presented visually, while participants (16 students) were required to select the ones they knew well. A quatrain would be excluded if more than four subjects identified it as familiar. It was suspected that the familiarity with the stimuli might influence the processing of the quatrains and its prosodic phrasing. If listeners were familiar with a quatrain, they would know what was coming without listening. Furthermore, unfamiliar quatrains were generally difficult to comprehend so that listeners might only perceive their rhythm. Therefore, the brain responses would largely reflect the prosodic phrasing rather than the syntactic and semantic processing. Since lyrics were unknown to most people, we didn't conduct a separate test to examine their familiarity.

As examples of the stimuli used in the present study, a quatrain (left) and a lyric (right) were shown below. To be clear, we used "/" to segment each word, while the Pinyin and English translations for isolated words and whole sentences were presented below each Chinese sentence. In Experiment 1, 60 quatrains were presented to listeners. To ensure that listeners were not accustomed to the rhythm of these quatrains, 15 lyrics randomly selected from 60 lyrics were intermixed. In Experiment 2, 53 quatrains were included, wherein 50 were used as experimental material and three as practice material. The material distribution was the same for lyrics.

日光/斜照/集灵台，
ri4guang1/xie2zhao4/ji2ling2tai2
the sun/gone slanting over/a lordly roof
The sun has gone slanting over a lordly roof.

红/树/花/迎/晓露开。
hong2/shu4/hua2ying2/xiao3lu4/kai1
red/branches/blossoming/leaned toward/the dew
And red-blossoming branches have leaned toward the dew.

昨夜/上皇/新/授/箓。
zuo3ye4/shang4huang2/xin1/shou4/lu4
last night/the Emperor/new/summoned/a favorite
Since the Emperor last night summoned a new favorite.

太真/含笑/入/帘来，
tai4zhen1/han2xiao4/ru4/lian2lai2
Lady Yang's/smile/came through/the curtains
And Lady Yang's bright smile came through the curtains.

少年/痛饮，
Shao4nian2/tong4yin3,
While young/I drank till drunk
While young,/I drank till drunk,

忆/向/吴江/醒。
yi4/xiang4/wu2jiang1/xing3.
I remember/I found in face/the Southern Stream/awake from dream
Awake from dream, I remember I found in face the Southern Stream.

明月/团团/高树/影，
ming2yue4/tuan2tuan2/gao1shu4/ying3
The moon/full/the tall tree/cast shadow.
The full moon cast shadow of the tall tree.

十里/水/沉/烟冷。
shi2li3/shui3/chen2/yan1leng3
For miles and miles/water/deep/on mist-veiled
On mist-veiled water deep flowing for miles and miles.

Chinese is a tone language, whose intonational contour has been well-acknowledged to be composed of two lines. The top line, consisting of the maximal values of pitch register for each syllable, reflects the pitch accent placement, whereas the bottom line, comprising the minimal pitch values of pitch register, indicates the prosodic structure (Shen, 1985). For analyzing the tonal properties of prosodic boundaries, the minimal pitch values for the pre-boundary and post-boundary syllables (F1 and F2 respectively) were extracted, and the size of pitch reset (F2 minus F1) was also measured. The duration of pre-boundary syllable (D1) and duration of pre-boundary syllable plus following silence (D2) were acquired respectively. Fig. 1 shows the global acoustic features of seven-character-quatrain in pitch (Fig. 1A) and duration (Fig. 1B).

One-way analysis of variance (ANOVA) was used to look inside the acoustic properties of these quatrains statistically. Firstly, one-way ANOVA with FB in each sentence of these poems run on D1 (also D2, there was no pause for this condition) and pitch reset respectively. The results revealed no significant main effect for both D1, $F(3,208)=0.374$, $P>0.05$, and pitch-reset, $F(3,208)=10.69$, $P>0.05$. Secondly, we compared the acoustic characteristics of PPBs embedded in the four sentences. For D1, the main effect of condition was not significant, $F(3,196)=2.154$, $P>0.05$, although PPB3 was somewhat shorter than other PPBs. Nevertheless for D2, ANOVA revealed a significant main effect of condition, $F(3,196)=11.210$, $P<0.0001$. For multiple comparisons, Bonferroni corrections were used, which also applies to the statistical analyses of ERP data in the current study. Multiple comparisons showed that PPB3 was significantly shorter than PPB2 and PPB4, while PPB1 was significantly shorter than PPB4, all $Ps<0.005$. The results of pitch reset also showed that the main effect of condition was significant, $F(3,196)=19.951$, $P<0.0001$, and multiple comparisons revealed that the reset value for PPB3 was significantly higher than other PPBs, $Ps<0.0001$ (for the

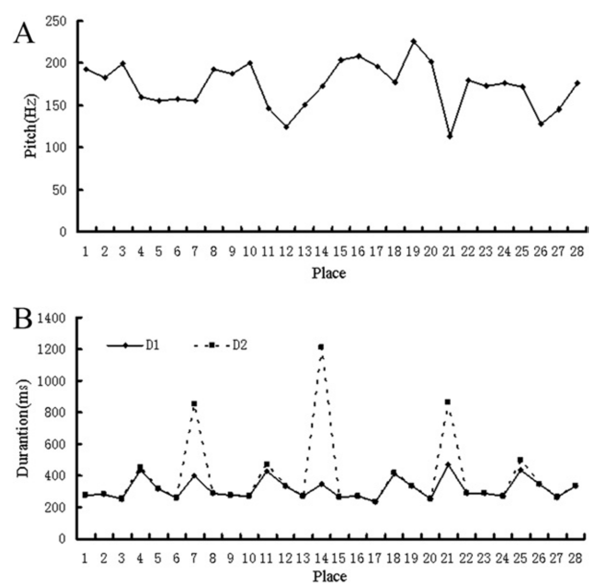

Fig. 1. Prosodic parameters. (A) Upper: minimal pitch for 28 characters in all the poems; (B) Lower: the duration (ms) of preboundary syllable (D1) and the duration of pre-boundary syllable + silence (D2).

Table 1. The acoustic parameters for different prosodic boundaries, with standard error in parentheses

Acoustic parameters	PPB1	PPB2	PPB3	PPB4	IPB1	IPB2	CB
D1 (ms)	435 (9)	431 (8)	410 (7)	435 (9)	398 (7)	469 (9)	346 (8)
D2 (ms)	451 (10)	469 (11)	418 (9)	501 (11)	851 (15)	863 (17)	1211 (19)
Pitch reset (Hz)	−4 (7)	−22 (6)	48 (11)	−39 (9)	41 (10)	68 (11)	31 (11)

acoustic properties, see Fig. 1, Table 1). Finally, we analyzed the acoustic parameters of IPBs and CB. For D1, the main effect of condition was significant, $F(2,147)=59.339$, $P<0.0001$. Multiple comparisons indicated that both IPB1 and IPB2 were significantly longer than CB, and IPB2 was longer than IPB1, all $Ps<0.001$. For D2, the main effect of condition was also significant, $F(2,147)=139.975$, $P<0.0001$. Multiple comparisons indicated that CB was significantly longer than IPB1 and IPB2, $P<0.001$. As for pitch reset, statistical analysis revealed that the main effect of condition was significant, $F(2,147)=3.326$, $P<0.05$, and multiple comparisons indicated that IPB2 was significantly higher than CB, $P<0.005$ (for the acoustic properties, see Fig. 1, Table 1).

Procedures

In Experiment 1, participants were instructed to listen carefully (and repeatedly, if needed) to the poems, while the written versions with commas or periods at the end of each sentence were also presented. Participants were told to label the prosodic hierarchy of the poems according to their acoustic characteristics before the experiment. They were asked to mark the size of the prosodic boundaries on a 4-point scale, with 3 being the highest and 0 the lowest, using the dichotomous approach. In practical terms, they were advised to find the most distinguishable boundaries in a poem first and marked it with 3, thus dividing a poem into two parts. And then they were instructed to determine the next biggest boundaries in these two parts and mark them with 2. In this way, they would finally determine four layers for each poem.

In Experiment 2, subjects were tested in one session lasting about 45–60 min, while reclining in a comfortable chair. All the materials were presented aurally to the listeners in a pseudo-randomized order in three blocks of 33 or 34 poems, while electroencephalography (EEG) was recorded. The blocks were counterbalanced across runs. In each block, the same type of material was presented in no more than three consecutive trials. Each poem began with a fixation cross at the center of the monitor to minimize eye movements. After 300 ms, a poem was presented aurally, while the cross remained on the screen. Subjects were instructed to listen carefully and perceive the rhythm of the poems attentively. As soon as a poem ended, a number (0, 1, 2, 3 or 4) was presented at the center of the screen. "0" indicated that there was no task to complete, while the number 1, 2, 3, or 4 indicated that a "sentence melody" would be presented as soon as they pressed the spacebar, each number indicating the corresponding line/sentence within the poem. Their task was to answer whether the melody was consistent with the poem they had just heard by pressing the keyboard as accurately as possible in 20% of the poems. For example, if number "2" appeared on the screen, it meant that a pure prosodic sentence melody would be presented, and they should determine whether it was consistent with the rhythm of the second sentence of the poem they had just heard. This task was set to have the listeners focused on the rhythm of the poems. The number 1/2/3/4 and the use of right/left hand to press the button were balanced across the entire experiment. At the beginning of the session, subjects were exposed to six practice poems (three quatrains and three lyrics) with EEG recorded, and these two types of poems were presented in a pseudorandom order. Participants were instructed before the beginning of the experiment to avoid eye blinking and other body movements during acoustical presentation of the poems.

ERP recordings

EEG was continuously recorded from 64 cap-mounted Ag/AgCl electrodes. The left mastoid served as reference during the recording. Vertical electrooculogram (EOG) was recorded bipolary by placing electrodes above and below the left eye, and horizontal EOG was recorded bipolary by placing electrodes at the outer canthus of each eye. The electrode impedance was less than 5 kΩ. Recordings were amplified with a high cutoff of 100 Hz and sampling rate was 500 Hz. Offline the EEG signals was re-referenced to average mastoids.

Data analysis

EEG data were processed with the software NeuroScan 4.3. The EEG and EOG signals were screened off-line for eye movements and electrode drifts. The data were filtered with a low-pass filter of 40 Hz. Epochs comprised the 200 ms preceding and 1500 ms following the pre-boundary syllable onsets.

Grand average ERP wave shapes were first analyzed by visual inspection in order to identify time window (TW) of interest. Except for FBs, the listeners provided positive shifts in response to all conditions (PPBs, IPBs, and CB) in the TW of 600 to 1100 ms after the acoustic onset of the pre-boundary syllables. Thus this TW was selected for analysis. For statistical analysis, three-way ANOVA was performed with mean amplitude as dependent variable, with condition, region (anterior–posterior topography: anterior—F3, FZ, F4; central—C3, CZ, C4; posterior—P3, PZ, P4), and hemisphere (left–right topography: left—F3, C3, P3, midline—FZ, CZ, PZ, right—F4, C4, P4) as independent factors. In cases in which interactions of the condition with topological factors could be observed, separate analyses were computed for hemispheres or regions. In order to verify that the four FBs with no acoustic differences indeed have no difference in brain response evoked, we conducted a consecutive ANOVA with mean amplitude in latency bins of 100 ms width (e.g. 100–200 ms, 200–300 ms, etc.) as dependent factor, with condition (FB1, FB2, FB3, and FB4), region and hemisphere as independent factors. Besides, two ANOVAs were conducted to statistically identify that different prosodic boundaries elicited positive shifts. The first ANOVA was performed with mean amplitude as dependent factor, with condition (FB, PPB1, PPB2, PPB3, and PPB4), hemisphere and region as independent factors. Based on visual inspection of the average waveforms, it seemed that PPB1 elicited higher amplitude as compared to other PPBs. According to our hypothesis, there would be an amplitude increase for the positive shift as the prosodic hierarchical level became higher. Then if the amplitude elicited by IPBs or CB were higher than the deflection evoked by PPB1, it would surely be higher than other PPBs. Thus, the second ANOVA was performed with mean amplitude as dependent factor, with condition (FB, PPB1, IPB1, IPB2 and CB), hemisphere and region as independent factors. In order to explore the time course of differences between the four PPBs, we averaged amplitude modulations for successive 100-ms TWs starting from stimulus onset. Thus, fifteen 100-ms TW analyses were performed with condition, hemisphere, and region as repeated-measures factors. For significant effects, a second ANOVA was performed summarizing significant consecutive TWs into one larger time. The same analyses were also conducted to explore the time course of differences between different hierarchical prosodic

Table 2. Marking results for different hierarchical prosodic boundaries in the poem by 19 subjects

Labelling level	FB	PPB	IPB	CB
0	96%	4.14%	1.56%	0.52%
1	4%	94.14%	1.61%	0.21%
2	0%	1.61%	96.04%	1.46%
3	0%	0.10%	0.78%	97.82%

boundaries (PPB1, IPB1, IPB2, and CB). Because of the increased likelihood of Type I errors associated with the large number of analyses, only effects that reached significance in two consecutive time windows were considered significant. To establish the onset of the positive shifts, a series of onset analyses were conducted in consecutive mean amplitude latency bins of 10 ms width (e.g. 100–110 ms, 110–120 ms, etc.) for PPBs, IPBs and CB as compared to FB. *P*-values were reported after Greenhouse–Geisser correction for non-sphericity.

RESULTS

Experiment 1

The results were presented in Table 2. As can be seen, participants never labeled FBs as level 2 or 3, and IPBs and CB were rarely marked as level 0 or 1, while PPBs may be signaled as 0 (probability of 4.14%). Generally speaking, listeners perceived CB in the middle of the poems to be the strongest prosodic boundary, followed by IPBs and PPBs, while FBs were the weakest ones.

Experiment 2

For behavioral results, the accuracy rate of participants' melody matching was 67.34%, well above chance level, although the task was difficult for them (according to post-experimental debriefings). For the electrophysiological effect, as shown in Fig. 2, FBs displayed negativities starting around 200 ms and lasting at least until the end of the 500 ms epoch. The statistical analysis revealed that there was no significant main effect of condition ($P>0.1$), and no interactions were provable between the factors condition×hemisphere and condition×region (see Fig. 2). Therefore, FB4 was selected randomly as reference for other statistical explorations.

As shown in Fig. 3, FB displayed a negativity, while the PPB conditions all showed positivities, starting around 500 ms and terminating after 1 s or more. The statistical analysis indicated that the main effect of condition was significant, $F(4,76)=30.501$, $P<0.0001$. Multiple comparisons indicated that FB ($M=-2.212$ μV, SE=0.140) differed significantly from PPB1 ($M=1.115$ μV, SE=0.131), PPB2 ($M=0.434$ μV, SE=0.091), PPB3 ($M=0.765$ μV, SE=0.123) and PPB4 ($M=0.868$ μV, SE=0.107), all $Ps<0.0001$, whereas none of the comparisons among PPBs were significant. Although the condition×region interaction suggested topographic differences, the condition effect was broadly distributed over the scalp. Furthermore, a condition×hemisphere interaction was also observed, $F(8,152)=3.813$, $P<0.0001$. Simple effect analysis revealed that the main effect of condition was significant at the left, $F(4,76)=16.14$, $P<0.0001$, midline, $F(4,76)=31.65$, $P<0.0001$, and the right hemisphere, $F(4,76)=24.46$, $P<0.0001$. Results of multiple comparisons showed that condition effect was broadly distributed over the scalp, while there was also significant difference between PPB1 to PPB2 and PPB3, $P<0.01$, at the right hemisphere.

As shown in Fig. 4, FB displayed a negativity, while PPB, IPBs and CB all showed positivities, starting around 500 ms and terminating after 1 s or more. For the shifts elicited by these different hierarchical prosodic boundaries, a main effect of condition was observed, $F(4,76)=22.715$, $P<0.0001$ (Fig. 4). Multiple comparisons indicated that FB was significantly different from IPB1 ($M=0.696$ μV, SE=

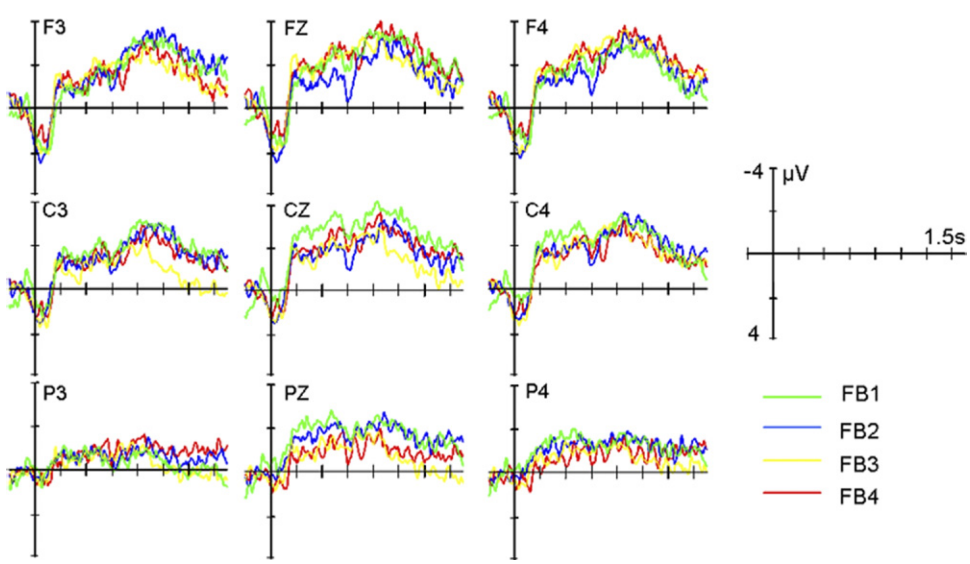

Fig. 2. Grand average waveforms over participants ($n=20$), time locked to the onset of pre-boundary syllable at selected channels for FB1, FB2, FB3 and FB4, four conditions are superimposed. For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.

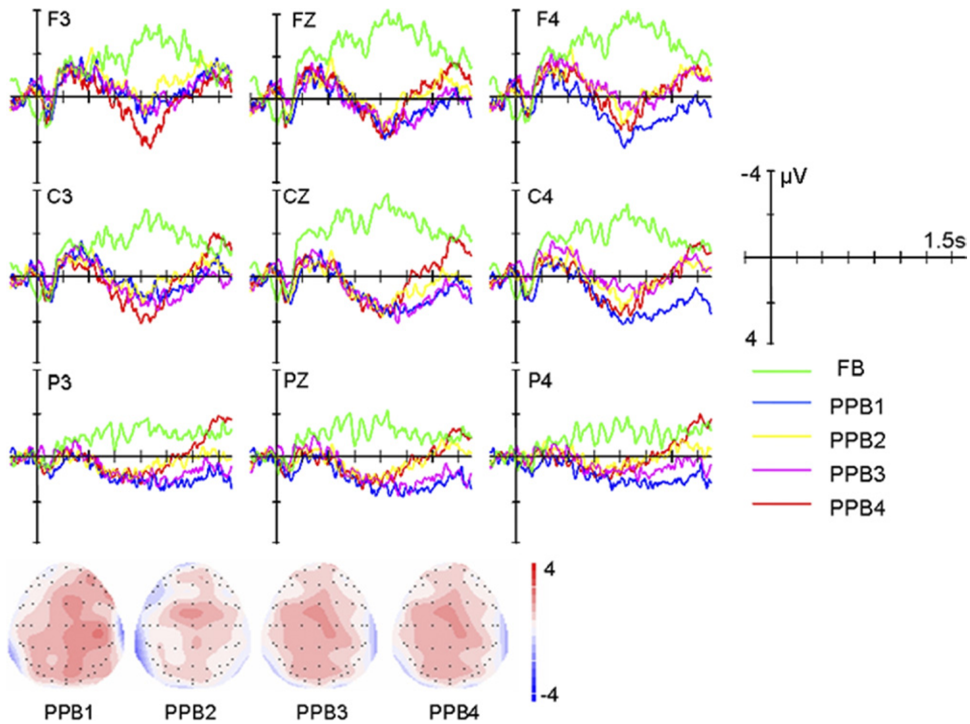

Fig. 3. Grand average waveforms over participants (n=20), time locked to the onset of pre-boundary syllable at selected channels for FB, PPB1, PPB2, PPB3 and PPB4, five conditions are superimposed. Topography of the ERP effects for PPBs during 600–1100 ms is showed below. For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.

0.118), IPB2 (M=0.992 μV, SE=0.126), and CB (M=0.433 μV, SE=0.134), all Ps<0.005, whereas other comparisons were not significant. Although the condition× region and condition×hemisphere interaction effects

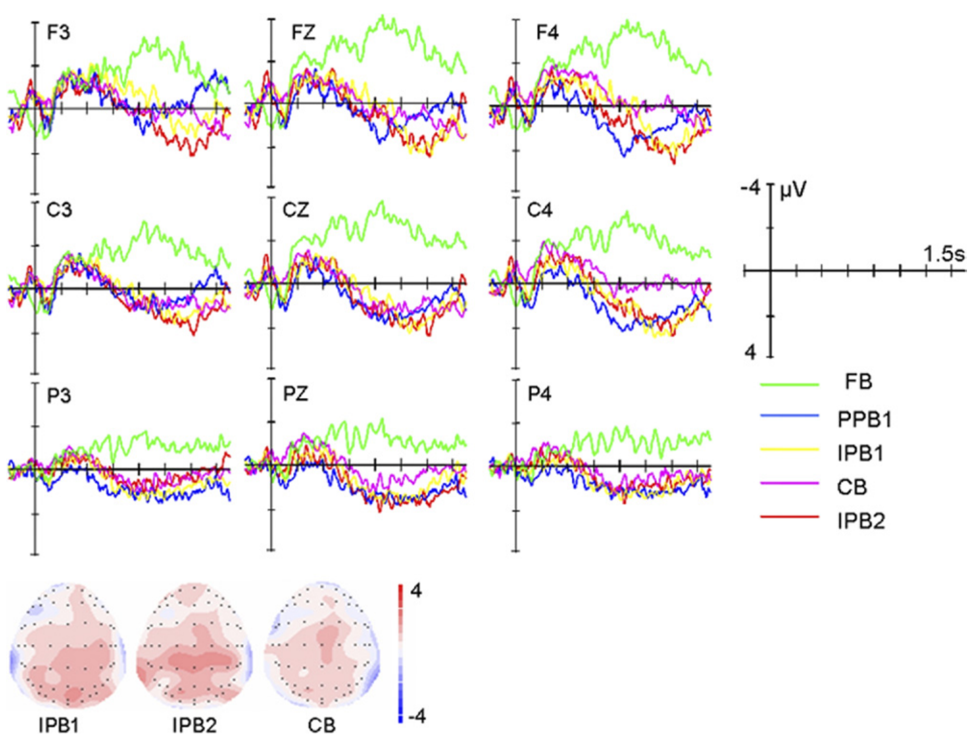

Fig. 4. Grand average waveforms over participants (n=20), time locked to the onset of pre-boundary syllable at selected channels for FB, PPB1, IPB1, IPB2 and CB, five conditions are superimposed. Topography of the ERP effects for IPB1, IPB2 and CB during 600–1100 ms is showed below. For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.

Table 3. Results of repeated-measures ANOVAs: Effect for PPBs

TW (ms)	Cond (3,57)	Cond×Hem (6,114)	Cond×Reg (6,114)	Cond×Hem×Reg (12,228)
0–100	$F=1.926$	$F=1.193$	$F<1$	$F=1.726$
100–200	$F=5.004$ $P=0.004$	$F=1.327$	$F<1$	$F=1.261$
200–300	$F<1$	$F<1$	$F<1$	$F=1.399$
300–400	$F<1$	$F=1.793$	$F=1.243$	$F=1.262$
400–500	$F=1.045$	$F=2.144$	$F<1$	$F=2.046$
500–600	$F<1$	$F=1.563$	$F<1$	$F=1.497$
600–700	$F<1$	$F=1.57$	$F<1$	$F=1.359$
700–800	$F=1.179$	$F=2.954$ $P=0.01$	$F<1$	$F=1.476$
800–900	$F=1.675$	$F=3.563$ $P=0.003$	$F=1.67$	$F=1.382$
900–1000	$F=1.938$	$F=3.081$ $P=0.008$	$F<1$	$F=1.226$
1000–1100	$F=4.071$ $P=0.011$	$F=2.936$ $P=0.011$	$F<1$	$F=1.039$
1100–1200	$F=4.506$ $P=0.006$	$F=2.201$ $P=0.048$	$F<1$	$F=1.104$
1200–1300	$F=3.208$ $P=0.03$	$F=2.405$ $P=0.032$	$F=1.086$	$F=1.403$
1300–1400	$F=3.673$ $P=0.017$	$F=2.889$ $P=0.012$	$F=1.877$	$F=1.556$
1400–1500	$F=3.966$	$F=1.976$	$F=2.735$	$F=1.236$

Cond, condition; Hem, hemisphere; Reg, region.

suggested topographic differences, the condition effect was broadly distributed over the scalp.

The first consecutive 100-ms TW analyses were conducted to explore the time course of differences between the four PPB conditions (see Table 3). Mean amplitudes of significant consecutive TWs were then averaged across the significant time interval and entered into a second ANOVA. Exploratory analyses revealed a significant main effect of condition between 1000 and 1400 ms. Multiple comparisons indicated that the amplitude of PPB1 was higher than PPB4 in this TW. Besides, condition×hemisphere interaction was detected in both TW 700 to 1000 ms, $F(6,114)=2.694$, $P<0.05$, and TW 1000 to 1400 ms, $F(6,114)=2.614$, $P<0.05$. For the first TW, the analysis of simple effect revealed that the main effect of condition was not significant at the left, $F(3,57)=1.86$, $P>0.05$, and central, $F(3,57)<1$, but significant at the right hemisphere, $F(3,57)=2.88$, $P<0.05$. Post hoc analysis revealed that positive shift elicited by PPB1 was higher than that evoked by PPB2 and PPB3 in amplitude. For the second TW, simple effect analysis revealed that the main effect of condition was not significant at the left, $F(3,57)=1.86$, $P>0.05$, but significant at the central, $F(3,57)=4.50$, $P<0.01$, and the right hemisphere, $F(3,57)=4.00$, $P<0.05$. Post hoc analysis revealed that PPB1 elicited higher amplitude than PPB2 and PPB4 at both the midline and right hemisphere, while PPB3 induced higher amplitude than PPB4 at the midline.

The second consecutive 100-ms TW analyses were conducted to explore the time course of differences between different hierarchical prosodic boundaries (see Table 4). Exploratory analyses revealed a condition×hemisphere×region interaction in TW 600 to 1000 ms, $F(12,228)=2.066$, $P<0.05$, and a condition×region interaction in TW 1100 to 1400 ms, $F(3,57)=3.142$, $P<0.01$. For the condition×hemisphere×region interaction, further analyses revealed that the main effect of condition was significant in centro–frontal area [F4: $F(3,57)=3.00$, $P<0.05$; C4: $F(3,57)=3.22$, $P<0.05$]. Multiple comparisons indicated that PPB1 elicited higher amplitude than IPB1 and CB in this TW. To clear up the condition×region interaction, separate analyses for each region were computed. The results showed that the main effect of condition was significant at the anterior, $F(3,57)=3.25$, $P<0.05$, but not significant at the central, $F(3,57)=1.75$, $P>0.05$, and the posterior area, $F(3,57)<1$. Multiple comparisons indicated that PPB1 induced positive shift with lower amplitude as compared to both IPB1 and IPB2, and the deflection elicited by CB was also lower than that correlated to IPB2.

The analysis of onset latency revealed that the CPS elicited by PPB1 started in 450–460 ms latency range and 460–470 ms for the other PPBs followed by a long and uninterrupted series of bins. The CPS elicited by IPB1 and IPB2 started in 510–520 ms and 500–510 ms respectively. CB evoked a CPS with a latency onset of approximate 710 ms, which is best seen at midline electrodes Cz and Pz.

DISCUSSION

The present study aimed to investigate the cognitive processing of hierarchical prosodic boundaries in discourse. Specifically, we focused on an electrophysiological component which is known as signaling the online processing of prosodic boundary in adults and infants (Pannekamp et al., 2005; Steinhauer et al., 1999). As a special kind of discourse, Chinese Tang poem is marked by strict tonal patterns and rhyme schemes, constituting a well-structured prosodic hierarchy. Participants in this study listened carefully to these poems and completed the rhythm match-

Table 4. Results of repeated-measures ANOVAs: Effect for different hierarchical prosodic boundaries

TW (ms)	Cond (3,57)	Cond×Hem (6,114)	Cond×Reg (6,114)	Cond×Hem×Reg (12,228)
0–100	F=2.155	F<1	F<1	F=1.218
100–200	F=5.803 P=0.002	F<1	F<1	F<1
200–300	F=1.625	F=1.072	F<1	F=1.028
300–400	F<1	F<1	F<1	F=1.256
400–500	F=1.346	F=1.912	F<1	F=1.63
500–600	F=1.486	F=2.012	F<1	F=1.716
600–700	F=1.469	F=2.657 P=0.019	F=1.009	F=1.99 P=0.026
700–800	F<1	F=2.146	F=1.914	F=2.139 P=0.016
800–900	F=1.549	F=2.235 P=0.045	F=3.49 P=0.003	F=2.056 P=0.021
900–1000	F<1	F=1.571	F=1.358	F=1.978 P=0.027
1000–1100	F=1.721	F=1.955	F=1.676	F=1.774
1100–1200	F=2.436	F=2.559 P=0.047	F=2.888 P=0.024	F=2.012
1200–1300	F=2.127	F=1.731	F=3.689 P=0.002	F=2.154
1300–1400	F<1	F=1.637	F=3.329 P=0.005	F=2.105
1400–1500	F<1	F=1.180	F=2.731	F=2.242

ing task. Although the task was difficult for them, their accuracy rate was well above chance level. For the electrophysiological response, it was shown that PPBs in the sentences elicited reliable positive shifts, which closely resembled the CPS component established by previous studies (Isel et al., 2005; Li and Yang, 2009; Pannekamp et al., 2005; Steinhauer, 2003; Steinhauer et al., 1999). More interestingly, positive shifts were also evoked whenever listeners detected IPBs and CB between sentences and couplets. All of these shifts were associated with prosodic boundary processing, had a positive polarity, and displayed a bilateral fronto–central scalp distribution. Given all these characteristics, we defined the shifts elicited by these breaks as CPS. In contrast to previous studies (Isel et al., 2005; Li and Yang, 2009; Pannekamp et al., 2005; Steinhauer, 2003; Steinhauer et al., 1999; Steinhauer and Friederici, 2001), we observed CPS effects for the processing of prosodic boundaries higher than within-sentence breaks. As listeners couldn't detect the syntax and semantics of the poems presented, the current results further confirmed the purely prosodic nature of this component. The most profound effect was on the onset latency, which was longest for CB (710–720 ms time range), next longest for IPBs (500–520 ms time range), and shortest for PPBs (450–470 ms time range). It seemed that the higher the prosodic hierarchical level, the later the onset latency of CPS. As for the CPS elicited by the same hierarchical level, it was showed that PPB1 induced higher amplitude as compared to other PPBs mainly in the right hemisphere, while there was no significant difference between the positive shifts evoked by IPB1 and IPB2. Moreover, the CPS amplitude didn't increase as the boundary hierarchy became higher, but demonstrated a dynamic change of patterns over time. That is, at right fronto–central electrodes the CPS amplitude for PPB was higher than that for CB and IPB1in an earlier TW (600–1000 ms), while in a later TW (1100–1400 ms) the positive deflections correlated to IPBs were higher than that elicited by PPB and CB in the anterior area. The present results will be discussed and explained in the following section.

As outlined in the introduction, previous research about prosodic processing in discourse has mainly used behavioral experiments, concerning questions including acoustic-phonetic correlates of prosody and its relation to discourse structures. Few studies have been conducted to explore the neural basis of prosodic processing using electrophysiology techniques. Discourse was often viewed as a structured composition of hierarchically different entities, consisting of paragraphs, which can be further split into sentences, clauses, and so on. For the prosodic phrasing results below the level of sentence, de Pijper and Sanderman (1994) reported that the strength of prosodic boundaries in spoken utterances perceived by listeners is not simply a binary feature, as their continuous pause, pitch and declination reset vary in a seemingly graded manner (de Pijper and Sanderman, 1994). A similar picture for phrasing emerged from the present materials. Specifically, intonational phrases and couplets (paragraphs), as larger-scale information units in the quatrains, tend to be phonetically encoded by means of higher pitch reset values and longer pauses. On the other hand, the phonological phrases in the sentences were always signaled by lower (or no) pitch reset values and somewhat shorter pauses. Although acoustic variance existed for different prosodic boundaries, they all functioned as segmenting the incoming speech and elicited positive shifts stably. Most importantly, the finding that prosodic boundaries at the discourse level, namely IPBs and CB, could evoke CPS has extended conditions for its elicitation. Whereas in previous studies the CPS had been considered as reflecting the closure of intonational phrases in the sentence level (Isel et al., 2005; Li and Yang, 2009; Pannekamp et al., 2006; Steinhauer, 2003; Steinhauer et al., 1999; Steinhauer and Friederici, 2001), the present results indicate that whenever listeners decompose continuous speech into relevant structural units such as phrases, sentences or paragraphs, this positive effect will be induced. In other words, the condition for the CPS elicitation is not restricted by the size of prosodic units.

The CPS onset latency

As pointed out in the introduction, higher prosodic hierarchy is generally accompanied with longer pause duration in discourse. This is also the case for the present poems. To be specific, the pause duration for PPBs was never longer than 100 ms, while it was about 400 ms for IPBs and 850 ms for CB. Coincidentally, the onset latency of CPS increased as the prosodic hierarchical level became higher. It appeared that the CPS onset latency was significantly influenced by the pause duration in the vicinity of prosodic boundary. Nevertheless, when we further compared the latency data to D1 and D2, it was found that for IPBs and CB, the CPS onset was already present during the pause. Besides, the trigger point was set to the onset of the last syllable before the respective boundary. Thus, it seemed implausible to attribute the latency differences merely to pause duration. During sentence processing, it has been indicated that pause length influences the CPS onset latency (Li and Yang, 2009; Steinhauer et al., 1998). It is important to note, however, the result was obtained by comparing the positive shifts elicited by the same prosodic boundary with and without pause. This was somewhat different from the comparisons among brain effect evoked by different prosodic boundaries, which vary in preboudary lengthening and pitch reset, except for pause length. Future studies are needed to explore what role a pause may play in boundary perception, and its relationship to other acoustic parameters.

From another point of view, the CPS latency shift, which was observed among prosodic boundaries at different hierarchies, seems to provide evidence for the chunking activity and the retrospective processing of information before boundary as proposed in previous studies (Knösche et al., 2005; Neuhaus et al., 2006). This assumption is based on the fact that prosodic boundary at higher level, which is generally accompanied with longer pause, always includes more preceding information, that is, rhythm, meter, metrical patterns, etc. It seemed that the earlier CPS reflects a faster processing of the shorter foregoing phrase as a lower amount of information was processed. Likewise, when a prosodic boundary like CB was detected, the brain presumably recognizes that more information has to be closed, and accordingly CPS is induced in a relative later time window. This is also quite in line with a previous study conducted to investigate phrase perception in music (Neuhaus et al., 2006).

The CPS amplitude for prosodic boundaries at the same hierarchical level

It seemed that at least for poems, prosodic boundary processing was influenced by its specific place in a discourse. This is reflected by the CPS amplitude correlated to the prosodic breaks at the same hierarchical level. Theoretically, the prosodic boundaries at the same hierarchical level should elicit identical ERP effect, since the cognitive processing for these boundaries was basically alike. However, visual inspection together with statistical results revealed that PPB1 elicited higher CPS amplitude as compared to other PPBs (see the results section). If we inspect the quatrains, and take into consideration the characteristics of neural activity, it will be found that the reason might lie in the properties and role of the stimuli. Although the four sentences in a poem function differently in the elaboration of a theme, that is, introduction (qi3), elucidation of the theme (cheng2), transition to another viewpoint (zhuan3) and summing up (he2) respectively, their difference in rhythmic pattern was hardly distinguishable. It is agreed that neural activity is usually reduced when stimuli are repeated (Grill-Spector et al., 2006). In the present study, when the first sentence of a poem was presented aurally, the stimulus was relatively novel, and participants were also highly aroused. As they listened to the subsequent sentences, the electrical activity in brain cells reduced given that the rhythm of each sentence varied not too much. Accordingly, the brain effect as reflected by the CPS amplitude decreased. It seems that something like "prosodic priming" may function in poem processing similar to syntactic or semantic priming, which refers to the phenomenon that the amplitude of the electrophysiological effect was reduced in target word/sentence when it followed a prime word/sentence (Kutas and Federmeier, 2000; Tooley et al., 2009). This point can be further evidenced from the amplitude comparison between PPB3 and PPB4. As the first couplet was closed, it was followed by a long pause, and the speech signal stopped stimulating the neuronal populations temporarily. When the participants in our study perceived the first sentence after CB, it would be a relatively fresh stimulus, and accordingly the PPB embedded in (namely PPB3) the sentence elicited greater brain effect as compared to the later one (PPB4). Nevertheless, this is only our speculation, and further studies are needed to investigate whether prosodic priming indeed exists and under what conditions it will appear in detail.

As to the difference between PPB1 and other PPBs in the right hemisphere, it is quite compatible with the existing accounts which propose that temporal resolution for the identification of phonemes is mainly processed in the left hemisphere, whereas spectral resolution for the recognition of tonal patterns is dominated by the right hemisphere (Schirmer and Kotz, 2006; Zatorre et al., 2002). In Chinese poetry, tones as well as stresses create rhythm, constituting the most important element of poetry. The CPS lateralization in the current study seems to change with the variation of stimulus properties, which is also in line with previous findings. It has been indicated that CPS is observed in all kinds of sentence materials including pseudoword sentences (Pannekamp et al., 2005), filtered speech materials (Steinhauer and Friederici, 2001), and hummed speech (Pannekamp et al., 2005). It is typically distributed bilaterally with a shift to the right hemisphere for hummed speech, which was preserved the prosody of natural speech while major aspects of semantic and syntactic information was removed (Pannekamp et al., 2005). It should be noted that the rhythm matching task used in the present study might have had some influence on the observed brain activation patterns. Researchers have shown that the processing of prosodic information was lateralized as a function of particular task demands (Plante et al., 2002; Strelnikov et al., 2006). In the present study, although segmental information is included in the poems, the task induced an attentional focus on pure prosody processing, and led to a right lateralization for CPS component.

The CPS amplitude for different hierarchical prosodic boundaries

The result of behavioral experiment showed that listeners inclined to label CB in the middle of the quatrains as the strongest prosodic boundary, and next strongest for IPBs and PPBs in-order. According to previous studies, the impression of "phrasedness" seemes to be the most important cue for determining the CPS amplitude (Li and Yang, 2009; Steinhauer and Friederici, 2001), and therefore, the CPS amplitude is supposed to increase as the boundary levels become higher. In the ERP experiment, however, this hypothesis was not proved. It was even tentatively smaller for the CPS elicited by CB. For the seeming contradiction between the behavioral and ERP results, we assume that participants adopted different perceptual strategies when doing behavioral and ERP experiments. In Experiment 1, listeners were advised to use the "bisection method" to label the prosodic hierarchy of quatrains. Participants could listen to the poems repeatedly, while looking at the visually presented materials in the mean time. It is probable that in such a setting they were more likely to be influenced by the top–down process related to prosodic phrasing, that is, they may retrieve the general schemata of the poems stored in long-term memory. Furthermore, since quatrains have well-balanced structure, listeners may prefer to label the prosodic boundaries in the middle of the poems as the biggest ones, which seemed easier for them to complete the marking task.

In the ERP experiment, however, participants had only one chance to listen to each poem, and more importantly, they were asked to complete the rhythm-matching task. In this situation, they would be more likely to make use of acoustic boundary markers for the bottom-up processing of the speech input. During musical processing, it has been indicated that for strong phrase boundaries (longer boundary tones and longer pauses), musicians revealed a stronger CPS, while non-musicians showed a greater early negativity. However, if the pause is prolonged at the expense of the boundary tone, these effects are less pronounced (Neuhaus et al., 2006). In the current study, pause lengths were on average longest for CB, whereas preboundary syllable duration was shorter as compared to PPBs and IPBs. Preboundary lengthening is an essential and classical cue for listeners to perceive phrase boundaries in both language and music (Knösche et al., 2005; Wang et al., 2004; Wightman et al., 1992), and the boundary tone length has a strong influence on the CPS (Neuhaus et al., 2006). Hence, the present study indicates that a long pause alone does not necessarily produce a large CPS. Prolonged boundary syllables modify CPS and should be considered as a relevant boundary marker of prosodic phrasing in discourse. Besides, in music area, it has been stated that CPS not only reflects the detection of the phrase boundary, but also mirrors the memory- and attention-related processes, which are necessary for the transition from one phrase to the next (Knösche et al., 2005; Nan et al., 2006, 2008). Similar to music processing, boundary processing in poem may also be seen as a complex process comprising several cognitive subcomponents, like the detection of prosodic boundary cues, the reorientation of attention, the integration of information between prosodic units, etc. For CB in the poem, its acoustic parameter includes pre-final lengthening, pitch reset, and more importantly, pause with long duration (about 850 ms). We assume that such a time distance between couplets maybe so long that both verses were perceived as two relatively separate units. Moreover, the subjects in the ERP experiment were asked to complete the rhythm matching task, which implicitly guided the listeners to pay close attention to the inner rhythm of each sentence, while ignoring the relationship between sentences or couplets. Thus, when participants perceived CB, they were prone to expend relatively less cognitive resource to reorientate their attention and integrate the two couplets, and thus the brain effect indicated by CPS amplitude was somewhat reduced.

It can be seen that the CPS amplitude for CB and IPB1 was lower than PPB1 in an earlier TW (600–1000 ms), while in a later TW (1100–1400 ms) both PPB1 and CB elicited lower amplitude than the IPBs. It appeared that the CPS elicited by PPB1 reached its crest in an earlier TW as compared to the two IPBs. This is proved by an additional peak latency analysis, revealing that the peak latency of positive shift related to PPB1 (890 ms) was significantly earlier than that elicited by IPB1 (917 ms) and IPB2 (925 ms) ($P<0.005$). As has been noted, CPS is not a simple component only reflecting the detection of phrase boundary, but also an indicator of higher cognitive processes, involving memory (e.g., the closed phrase has to be stored in memory as a unified entity, and temporal integration between prosodic units) and attention (e.g., the attention focus has to be shifted and directed toward the next phrase). In the present study, PPB is lower in hierarchical level, which demands less time to complete all the relevant sub-cognitive processes. As a result, the CPS elicited by PPB1 reached its peak amplitude in an earlier TW as compared to IPBs, which are higher in prosodic hierarchical level. At least in the present study, it was indicated that although all prosodic boundaries of different hierarchy could elicited CPS, they are different in their amplitude and latency, reflecting different demands on attention and memory systems.

The present ERP results indicate that CPS occurred immediately when the prosodic boundaries in quatrains were perceived. If we compare the present results with previous ones in the domain of language and music, it can be seen that correlates for phrase boundary perception (i.e., the music CPS, the language CPS, and the poem CPS) resemble each other in amplitude and topographic distribution. The CPS in different domains could therefore represent a general form of information processing independent of modality. This also gives support to a recent conclusion drawn by Besson and Schon (2001), who suggested that special ERP components might reflect more general cognitive principles rather than being domain-specific. However, latency and duration in detail cannot be easily matched among these CPS in different domains. The language CPS and poem CPS seem to start immediately at the phrase boundary and last between approximately 500 and 1000 ms, while the music CPS was detected after the offset of a phrase boundary in a TW between 500 and 600 ms. The difference in latency between the CPS in language (and poem) and music may be due to the fact that, in language and poem, phrase boundary is signaled by parameter changes, which are already present before the onset of the pause. The influence of these early cues, that is, prefinal lengthening and, in particular, pitch movement, might be greater for phrase boundaries in language and poem as compared to music.

The current experiment used seven character quatrains to investigate the perception of prosodic boundaries at the discourse level. Although they give us a good opportunity to explore this question, there also exist some limitations. The simple structure of the quatrain was too straight or easy for the native listeners to predict, although lyrics, as the second type of Chinese poetry, were intermixed as the filler materials to interfere the expectation. This kind of prediction may influence the normal perception of the prosodic boundaries. We suppose that if other discourses were used, the result might be different to some degree, and would be better to answer the question how listeners perceive various prosodic boundaries in discourse. Besides, if these quatrains were presented to foreigners, what the brain response would be like is a problem deserving further exploration. It is suggested that CPS is elicited by different music styles irrespective of the relationship between the cultural background of the subjects and the cultural style of the music (Nan et al., 2006). However, it has also been reported that L1 and L2 speakers of German showed totally different ERP effects when listening to hummed speech (Isel F and Friederici AD, 2005). How L1/L2 speakers of tonal language (e.g. Chinese) will respond to boundary processing, especially in poems, is currently unknown. Further research may address these issues to better understand the mechanism of prosodic phrasing and the nature of CPS.

Acknowledgments—This research was supported by the National Natural Science Foundation of China (60775026), Scientific Research Foundation for Young Scientists, Institute of Psychology, Chinese Academy of Sciences (Y0CX052S01), and the Foundation of State Key Laboratory of Brain and Cognitive Science (The cognitive significance of the closure positive shift and its related prosodic cues). We are grateful to the two anonymous reviewers for their valuable comments and advice to earlier versions of this paper. We also thank Xiaohong Yang, Werner Sommer, and Yiya Chen for the draft revision.

REFERENCES

Besson M, Schon D (2001) Comparison between language and music. Ann N Y Acad Sci 930:232–258.

Boecker KB, Bastiaansen MC, Vroomen J, Brunia CH, de Gelder B (1999) An ERP correlate of metrical stress in spoken word recognition. Psychophysiology 36(6):706–720.

de Pijper J, Sanderman AA (1994) On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues. J Acoust Soc Am 96:2037–2047.

Friedrich CK, Alter K, Kotz SA (2001) An electrophysiological response to different pitch contours in words. Neuroreport 12(15): 3189–3191.

Friedrich CK, Kotz SA, Friederici AD, Gunter TC (2004) ERP reflect lexical identification in word fragment priming. J Cogn Neurosci 16(4):541–552.

Grill-Spector K, Henson R, Martin A (2006) Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn Sci 10: 14–23.

Grosz B, Hirschberg J (1992) Some intonational characteristics of discourse structure. In: Proceedings of the International conference on spoken language processing, Vol. 1 (Ohala JJ, Nearey TM, Derwing BL, Hodge MM, Wiebe GE, eds), pp 429–432. Banff, Canada.

Hirschberg J, Nakatani CH (1996) A prosodic analysis of discourse segments in direction-giving monologues. In: Proceedings of the 34th annual meeting on Association for Computational Linguistics, pp. 286–293. Association for Computational Linguistics Morristown, NJ, USA.

Isel F, Alter K, Friederici AD (2005) Influence of prosodic information on the processing of split particles: ERP evidence from spoken German. J Cogn Neurosci 17(1):154–167.

Isel F, Friederici AD (2005) The processing of prosodic boundaries in a second language: ERP evidence from French late bilinguals in spoken German. In: Proceedings of the Annual Meeting of Cognitive Neuroscience Society, New York.

Kerkhofs R, Vonk W, Schriefers H, Chwilla DJ (2007) Discourse, syntax, and prosody: the brain reveals an immediate interaction. J Cogn Neurosci 19(9):1421–1434.

Knösche TR, Neuhaus C, Haueisen J, Alter K, Maess B, Witte OW, Friederici AD (2005) Perception of phrase structure in music. Hum Brain Mapp 24(4):259–273.

Krivokapi J (2007) Prosodic planning: effects of phrasal length and complexity on pause duration. J Phon 35(2):162–179.

Kutas M, Federmeier KD (2000) Electrophysiology reveals semantic memory use in language comprehension. Trends Cogn Sci 4(12): 463–470.

Li W, Yang Y (2009) Perception of prosodic hierarchical boundaries in mandarin Chinese sentences. Neuroscience 158(4):1416–1425.

Nakatani CH, Hirschberg J, Grosz BJ (1995) Discourse structure in spoken language: studies on speech corpora. In: Proceedings of the AAAI spring symposium on empirical methods in discourse interpretation and generation, pp 106–112.

Nan Y, Knösche TR, Friederici AD (2006) The perception of musical phrase structure: a cross-cultural ERP study. Brain Res 1094(1):179–191.

Nan Y, Knösche TR, Zysset S, Friederici AD (2008) Cross-cultural music phrase processing: an fMRI study. Hum Brain Mapp 29(3):312–328.

Neuhaus C, Knösche TR, Friederici AD (2006) Effects of musical expertise and boundary markers on phrase perception in music. J Cogn Neurosci 18(3):472–493.

Pannekamp A, Toepel U, Alter K, Hahne A, Friederici AD (2005) Prosody-driven sentence processing: an event-related brain potential study. J Cogn Neurosci 17(3):407–421.

Pannekamp A, Weber C, Friederici AD (2006) Prosodic processing at the sentence level in infants. Neuroreport 17(6):675–678.

Plante E, Creusere M, Sabin C (2002) Dissociating sentential prosody from sentence processing: activation interacts with task demands. Neuroimage 17(1):401–410.

Schirmer A, Kotz SA, Friederici AD (2002) Sex differentiates the role of emotional prosody during word processing. Cogn Brain Res 14(2):228–233.

Schirmer A, Kotz SA (2006) Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. Trends Cogn Sci 10(1):24–30.

Shen J (1985) Pitch range and intonation of Chinese tone. In: Experimental phonetics of Chinese (Ling D, Wang LJ, eds) pp 75–107. Beijing: Pecking University Press [in Chinese].

Steinhauer K (2003) Electrophysiological correlates of prosody and punctuation. Brain Lang 86(1):142–164.

Steinhauer K, Alter K, Friederici AD (1998) Don't blame it (all) on the pause: Further ERP evidence for a prosody-induced garden path in running speech. Proceedings of the 5th International Conference on Spoken Language Processing, Vol. 5, (pp. 2187–2190). Canberra: Australian Speech Science and Technology Association (ASSTA).

Steinhauer K, Alter K, Friederici AD (1999) Brain potentials indicate immediate use of prosodic cues in natural speech processing. Nat Neurosci 2(2):191–196.

Steinhauer K, Friederici AD (2001) Prosodic boundaries, comma rules, and brain responses: the closure positive shift in ERP as a universal marker for prosodic phrasing in listeners and readers. J Psycholinguist Res 30(3):267–295.

Strelnikov KN, Vorobyev VA, Chernigovskaya TV, Medvedev SV (2006) Prosodic clues to syntactic processing—a PET and ERP study. Neuroimage 29(4):1127–1134.

Swerts M (1997) Prosodic features at discourse boundaries of different strength. J Acoust Soc Am 101:514–521.

Swerts M, Geluykens R (1993) The prosody of information units in spontaneous monologue. Phonetica 50(3):189–196.

Toepel U, Pannekamp A, Alter K (2007) Catching the news: processing strategies in listening to dialogs as measured by ERP. Behav Brain Funct 3(53):1–13.

Tooley KM, Traxler MJ, Swaab TY (2009) Electrophysiological and behavioral evidence of syntactic priming in sentence comprehension. J Exp Psychol Learn Mem Cogn 35(1):19–45.

Wang B, Lv S, Yang Y (2004) Acoustic analysis on prosodic hierarchical boundaries of Chinese. Acta Acust 29:29–36.

Wang B, Yang Y, Lv S (2006) The acoustic characteristics of clause, sentence and paragraph boundaries in chinese putonghua. Acta Acust 25(4):348–359.

Wang L (2005) Chinese prosody, 2nd ed. Shanghai: Shanghai Education Publishing Group.

Wightman CW, Shattuck HS, Ostendorf M, Price PJ (1992) Segmental durations in the vicinity of prosodic phrase boundaries. J Acoust Soc Am 91:1707–1717.

Zatorre RJ, Belin P, Penhune VB (2002) Structure and function of auditory cortex: music and speech. Trends Cogn Sci 6(1):37–46.

(Accepted 31 March 2010)
(Available online 9 April 2010)