Cognitive Neural Mechanisms of Poetic Tonal Patterns and Semantic Processing

Authors: Zhang Jingjing, Shi Ying, Deng Shanwen, Li Jiabin, Chen Qingrong, Chen Qingrong

Date: 2025-12-01T00:00:00+00:00

Abstract

Music is the very life of poetry. For ancient Chinese poetry, tonal patterns represent a typical expression of poetic musicality. This study selected pentasyllabic regulated verse as experimental materials, simultaneously manipulating the tonal patterns and semantic features of poetry, and employed EEG technology to investigate the cognitive neural mechanisms of tonal perception and poetic meaning comprehension during the reading of ancient Chinese poetry. Behavioral results demonstrated a main effect of tonal patterns as well as an interaction between tonal patterns and semantics. ERP results further revealed the presence of tonal effects during poetry reading, manifesting as an early P200 effect, followed by an N400 effect and a late LPC effect; moreover, tonal perception and poetic meaning comprehension modulated and integrated with each other in the middle and late stages, as evidenced by an interaction on the N400 and LPC. Additionally, to uncover the relationship between neural activity and tonal/semantic features, EEG data were subjected to decoding analysis, which revealed that neural networks trained on EEG data could effectively classify different types of poetic lines. The experimental results consistently indicate that during poetry reading, individuals anticipate harmony in the sounds of characters themselves as well as coordination between sound and meaning, and such anticipation influences early phonological representation and subsequent sound-meaning integration in poetry.

Full Text

Cognitive Neural Mechanisms of Tonal Patterns and Semantic Processing in Poetry

ZHANG Jingjing[1], SHI Ying[1], DENG Shanwen[1], LI Jiabin[2], CHEN Qingrong[1,3]

([1] School of Psychology, Nanjing Normal University, Nanjing 210097, China)
([2] School of Chinese Language and Literature, Nanjing Normal University, Nanjing 210097, China)
([3] Adolescent Education and Intelligence Support Lab of Nanjing Normal University, Laboratory of Philosophy and Social Sciences at Universities in Jiangsu Province, Nanjing, 210097)

Abstract

Music has long been considered the lifeblood of poetry. For classical Chinese poetry, tonal patterns represent a quintessential expression of poetic musicality. The present study employed five-character regulated verse as experimental materials, simultaneously manipulating both tonal patterns and semantic features to investigate the cognitive neural mechanisms underlying tonal perception and poetic comprehension during the reading of classical Chinese poetry using EEG technology. Behavioral results demonstrated a main effect of tonal patterns and an interaction between tonal patterns and semantics. ERP findings further revealed the presence of tonal effects during poetry reading, manifesting as an early P200 effect, followed by an N400 effect and a late LPC effect. Moreover, tonal perception and poetic comprehension mutually modulated and integrated during middle-to-late processing stages, as evidenced by interactions on the N400 and LPC components. Additionally, to elucidate the relationship between neural activity and tonal/semantic features, we conducted decoding analyses on EEG data, which revealed that neural networks trained on EEG data could effectively classify different types of poetic lines. The convergent results indicate that readers anticipate both phonological harmony and the coordination between sound and meaning during poetry reading, with such expectations influencing early phonological representation and subsequent sound-meaning integration.

Keywords: poetry, music, tonal patterns, semantics, EEG

Received date: January 25, 2025
This research was supported by the National Social Science Fund Major Project (21&ZD288).

Corresponding author: CHEN Qingrong, E-mail: jscqr80@sina.com

1 Introduction

Music is the lifeblood of poetry (Zhu Guangqian, 2012). Across cultures, poetry and music originated as unified art forms, with poems frequently sung and songs accompanied by music (Aristotle, 1968; Nietzsche, 2003). As the arts diverged, poetry evolved toward literal meaning, entering an era of "words without tunes." When words and melodies existed side by side, poetry' s musicality resided in its tune; after words separated from tunes, poetry' s music had to be found within the words themselves. Using classical Chinese regulated verse as an example, to embody linguistic rhythm, regulated poetry adheres to strict conventions regarding character count, tonal patterns, and rhyme schemes (Wang Li, 2012).

As a tonal language, Chinese possesses a unique prosodic feature in its classical poetry: the tonal pattern system. Classical Chinese had four tones—level (平), rising (上), departing (去), and entering (入)—where "level" refers to the level tone (including modern yinping and yangping), while "oblique" (仄) encompasses the remaining three tones. The tonal pattern rules for five- and seven-character regulated verse involve alternating level and oblique tones within single lines (e.g., "仄仄平平仄") and contrasting patterns between couplet lines (e.g., "平平仄仄平" versus "仄仄平平仄"). Additionally, the second character of a subsequent couplet' s first line must match the tone of the second character of the preceding couplet' s second line to avoid monotonous repetition (Wang Li, 2012). This systematic alternation of tones creates a "chanting" quality in poetry without musical accompaniment, thereby endowing it with intrinsic musicality. For instance, in the line "人间四月芳菲尽，山寺桃花始盛开" (The human world' s April flowers have faded, the mountain temple' s peach blossoms just begin to bloom), changing the second line to "山寺桃花始盛放" (mountain temple' s peach blossoms just begin to bloom) disrupts the flow despite minimal semantic change. Moreover, the four tones differ in phonetic quality, enabling tonal patterns to create harmony both within individual characters and between sound and meaning. For example, describing galloping horses often employs sonorous, abrupt sounds, while flowing water suits light, smooth tones (Zhu Guangqian, 2012). The first four lines of Han Yu' s "Listening to Master Ying Play the Zither" — "昵昵儿女语，恩怨相尔汝；划然变轩昂，猛士赴敌场" (Intimate as children' s whispers, grudges and kindnesses exchanged; Suddenly transforming into grandeur, warriors rushing to the battlefield)—demonstrate how tonal qualities generate harmonious impressions and synchronize sound with meaning.

The emergence of cognitive poetics in the 1980s began to examine how readers perceive prosodic features in poetry (Zhang Yehong, 2015; Tsur, 1983). Building on his analysis of Western poetry, Jakobson (1960) proposed the poetic function theory, which posits that poetic texts prompt readers to consciously perceive the linguistic structures and patterns inherent in the text, particularly its prosodic features. Previous experimental research on poetry reading has primarily focused on typical prosodic features of Indo-European poetry, such as stress and rhythm. These studies have found that reading poetry, compared to prose, activates not only reading-related brain regions but also additional areas associated with rhythm and music processing, such as the left and right superior temporal gyri (Zeman et al., 2013). Furthermore, as poetic difficulty increases, readers allocate more attentional resources to formal features, including certain prosodic elements (Yaron, 2002). Given readers' sensitivity to poetic prosody, prosodic anomalies (irregular stress patterns and rhyme violations) cause significant reading disruption and trigger re-reading of local text segments (Beck & Konieczny, 2020).

Some researchers have examined prosodic feature processing in Chinese poetry, such as pauses and rhymes, which together with tonal patterns constitute the typical prosodic features of classical Chinese poetry (Wang Li, 2012; Zhu Guangqian, 2012). By comparison, tonal patterns represent a global tonal organization rule across the poem, rhyme serves as an echo mechanism at line endings, and pauses create segmentation and beat grouping at the reading level. Li and Yang (2010) investigated the processing of prosodic boundary hierarchies in classical Chinese poetry using seven- or five-character quatrains as materials and found that different hierarchical boundaries could elicit closure positive shift (CPS) components, indicating that participants could perceive pauses at different positions in classical Chinese poetry. Eye-tracking studies on rhyme (Chen Qingrong, Yang Yiming, 2017) have shown that rhyme violations affect both early eye-movement measures (e.g., first fixation duration, gaze duration) and late measures (e.g., total fixation duration, total fixation count), while electrophysiological research (Chen et al., 2016) has also found that lines with rhyme violations elicit larger P200 and N400 components than rhyming-appropriate lines.

Zhang et al. (2020) used seven-character quatrains as experimental materials to simultaneously examine the processing of rhyme and tone in poetry, finding that tone and rhyme violations were reflected in P200, N400, and LPC components, and that these effects were modulated by participants' musical experience (Zhang et al., 2020) and experimental tasks (Zhang et al., 2021). Although these studies addressed tone processing in Chinese poetry, they did not employ lines strictly adhering to tonal pattern rules, nor did they consider changes between ancient and modern pronunciations, and thus were not specifically investigations of tonal pattern rules. Second, while manipulating rhyme and tone, these studies did not control for the degree of semantic reasonableness across lines, meaning that different experimental conditions varied not only in prosodic patterns but also potentially in meaning, making it difficult to disentangle which factor drove the observed effects.

In addition to prosodic features, cognitive poetics also concerns the comprehension of poetic meaning. According to genre-specific hypotheses (Jakobson, 1960) and traditional literary theory (Culler, 1975), reading and appreciation modes are clearly influenced by genre characteristics, leading to distinct processing differences across text genres with different external features. For poetry, readers direct attention to specific linguistic patterns, particularly prosodic features, resulting in slower reading speeds and greater difficulty in semantic comprehension (Hanauer, 1998). Moreover, Hanauer (1998) proposed that "similarity in sound is often treated as similarity in meaning," and the prosodic features of classical Chinese poetry create obvious phonological similarities. Based on these perspectives, semantic processing in poetry is modulated by prosodic factors.

Theoretical debates regarding the relationship between prosody and semantics have centered on the independence versus interactivity of processing stages and modes. On one hand, traditional views hold that semantic processing dominates comprehension, with prosody serving merely as an auxiliary feature. For instance, modular theory (Fodor, 1983) posits that prosodic processing constitutes a relatively independent, encapsulated module whose output is transmitted to higher-order central systems to assist semantic construction. Consequently, prosodic processing occurs only at early perceptual stages, and its interaction with semantics is unidirectional and limited, with semantic information exerting no influence on prosodic processing. On the other hand, the three-stage neurocognitive model (Friederici, 2002, 2012) and the interactive activation model (McClelland & Rumelhart, 1981) argue that prosodic processing is not modular but rather engages in real-time bidirectional interaction with semantic processing. These models emphasize that prosody and semantics are processed in parallel, not limited to early stages, and that prosodic cues can dynamically modulate semantic processing and vice versa.

Existing empirical studies have tested these theoretical perspectives. In Indo-European languages (e.g., English or German), research has primarily manipulated stress position (Rothermich et al., 2012; Rothermich & Kotz, 2013) and syllable duration (Magne et al., 2007), whereas in Chinese, studies have targeted phrase rhythm (e.g., "种蒜" [1+1] vs. "种植蒜" [2+1] rhythm; Luo & Zhou, 2010) or sentence rhythm (e.g., regular two-character word contexts vs. irregular patterns; Li et al., 2019). These studies have found that prosodic violations often elicit N100 (Li et al., 2019), N400 (Luo & Zhou, 2010; Magne et al., 2007), or late positive components such as LPC (Luo & Zhou, 2010; Magne et al., 2007; Rothermich et al., 2012), and that prosody and semantics mutually influence each other on N400 and LPC components (Li et al., 2019; Luo & Zhou, 2010; Rothermich et al., 2012). Among these, early ERP effects primarily reflect the detection and identification of prosodic violations, whereas late positive components involve more integrative and reanalytic processes of prosodic structure.

Despite these findings generally supporting the three-stage neurocognitive model and interactive activation model, several limitations remain. First, features such as stress, syllable duration, and phrase/sentence rhythm primarily involve sound intensity and duration— "time-oriented" prosodic features (e.g., forming beats and groups)—whereas "pitch-oriented" prosodic features (e.g., tone and intonation) require further scientific investigation (Hirst & Di Cristo, 1998; Ladd, 2008). Second, manipulations of lexical-level stress or syllable duration are often highly coupled with lexical semantics, such that prosodic changes may render words unnatural or even unrecognizable, thereby directly or indirectly interfering with lexical identification and semantic integration processes (Rothermich et al., 2012). Additionally, manipulations of sentence rhythm regularity, such as stress patterns in Indo-European languages (Rothermich et al., 2012; Rothermich & Kotz, 2013) and two-character word patterns in Chinese (Li et al., 2019), do not represent inherent, natural rules of language, and people may have difficulty perceiving or effectively integrating these patterns, thereby obscuring the true relationship between prosodic and semantic processing.

Based on these considerations, the present study employed classical regulated verse as experimental materials, simultaneously manipulating both tonal patterns and semantics, and used EEG technology to investigate the cognitive neural mechanisms of prosody and poetic meaning during poetry reading. The study' s innovations are threefold. First, tonal patterns primarily rely on fundamental frequency (F0) patterns and represent "pitch-oriented" prosodic features; this study is the first to examine the mutual influence between tonal combinations and semantic processing, providing an important supplement to theoretical debates on the prosody-semantics relationship. Second, manipulations of poetic tonal pattern rules, such as "山寺桃花始盛开/山寺桃花始盛放," alter tonal patterns without changing word meanings, thereby providing a cleaner prosodic manipulation. Moreover, classical poetry represents authentic cultural materials that embody natural patterns of tone usage in actual language, offering greater ecological validity than laboratory-constructed sentences and reducing artificial manipulation biases.

We hypothesized that tonal pattern effects would emerge at early stages, with irregular tonal patterns eliciting larger N100 or P200 amplitudes than regular patterns. On subsequent N400 and LPC components, we expected interactions between tonal pattern perception and semantic comprehension. N100 and P200 are early ERP components evoked by visual or auditory stimuli, sensitive to acoustic features (Näätänen & Picton, 1987), reflecting the matching of stimulus input with perceptual representations stored in memory (i.e., "perceptual matching" ; Luck & Hillyard, 1994). N400 is a classic ERP index of semantic processing (Brown & Hagoort, 1993; Hagoort et al., 2004; Kutas & Federmeier, 2011; Lau et al., 2008; Nieuwland et al., 2020), though some reading or speech comprehension studies have also linked N400 to phonological processing, possibly reflecting expectations about phonological or suprasegmental information (Chen et al., 2016; Coch et al., 2005; Rothermich et al., 2012; Rugg, 1984; Zhang et al., 2023). LPC is a late positive component appearing 500–1000 ms post-stimulus with widespread or central-parietal distribution, initially thought to reflect syntactic structure reanalysis (Friederici, 1995; Hagoort et al., 1999; Osterhout & Holcomb, 1992) or syntactic integration difficulty (Kaan et al., 2000; Kaan & Swaab, 2003), though subsequent research has linked LPC to semantic or prosodic integration and reanalysis (Brouwer et al., 2012, 2017). Simultaneously, we conducted decoding analyses on EEG data; if neural networks trained on EEG data could effectively classify different types of poetic lines, this would further confirm that tonal patterns and poetic meaning constitute important features in classical Chinese poetry processing.

2 Method

2.1 Participants

We selected participants majoring in classical Chinese literature. To ensure participants possessed substantial poetry experience, we administered the "Ancient Chinese Poetry Knowledge Test Questionnaire" (Chen Qingrong, Dou Fen, 2021) and selected those scoring above 60. Using G*Power 3.1 to calculate required sample size, we determined that at least 36 participants were needed to detect a medium effect size ($f = 0.25$) at a significance level of 0.05 with 95% statistical power (Cohen, 1988). To account for potential participant exclusion due to low behavioral accuracy or excessive EEG artifacts, and to capture possible smaller effects, we ultimately recruited 48 participants (24 female, 24 male) aged 18–28 years (M = 21.25, SD = 2.29). Poetry experience scores ranged from 60–98 (M = 80.83, SD = 11.89). All participants were right-handed native Chinese speakers with normal or corrected-to-normal vision and no history of neurological or psychiatric disorders. Participants provided informed consent and received compensation after the experiment. The study was approved by the Nanjing Normal University Biomedical Ethics Committee (approval number: NNU202410030).

2.2 Experimental Design and Materials

We employed a 2 (semantically congruous vs. incongruous) × 2 (tonally regular vs. irregular) within-subjects design, with dependent variables including participants' response accuracy and EEG signals elicited by critical words in the poems.

We selected 100 couplets from the Complete Tang Poems as experimental materials. These couplets strictly adhered to tonal pattern rules and showed no tonal changes between ancient and modern pronunciations (Luo Zhufeng, 1993). To eliminate potential rhyme effects, no rhyming occurred between the two lines of each couplet. Since these lines were semantically congruous and tonally regular, they constituted the first experimental condition (semantic + tone +, abbreviated S+T+). Because the experiment required participants to judge poem semantics, we selected relatively easy lines to ensure comprehension. To prevent familiarity-based expectations from affecting normal processing, we chose unfamiliar lines. Ten classical literature majors who did not participate in the formal experiment rated the difficulty and familiarity of the lines using five-point scales (1 = very easy/very unfamiliar, 5 = very difficult/very familiar). Results showed an average difficulty of 3 (SD = 0.29), indicating moderate difficulty, and low familiarity.

We modified the final character of the second line in each couplet to create three additional conditions: semantically congruous but tonally irregular (S+T−), semantically incongruous but tonally regular (S−T+), and both semantically incongruous and tonally irregular (S−T−). Character frequency of critical words did not differ significantly across conditions ($p = 0.10$). Example materials are shown in Table 1 .

After material creation, we conducted a semantic reasonableness rating. Twenty-six classical literature majors who did not participate in the formal experiment rated the semantic reasonableness of lines across the four conditions, with each participant rating only one version of each couplet. Using a five-point scale (1 = very unreasonable, 5 = very reasonable), mean ratings were: S+T+ (M = 4.52, SD = 0.73), S+T− (M = 4.40, SD = 0.72), S−T+ (M = 2.08, SD = 1.28), and S−T− (M = 2.09, SD = 1.22). Linear mixed model (LMM) analysis revealed a significant main effect of semantic reasonableness ($F = 674.00$, $p <$

0.001) and no significant interaction between semantic and tonal reasonableness (F = 1.36, p = 0.25), indicating that semantically incongruous lines were rated significantly lower than congruous lines, and that semantic reasonableness was unaffected by tonal patterns. The main effect of tonal reasonableness was not significant (F = 0.90, p = 0.34), confirming that tonally regular and irregular lines did not differ in semantic reasonableness, thus ensuring that tonal effects in the formal experiment would not be confounded by semantic factors.

2.3 Procedure

The experiment was conducted in a soundproof, electrically shielded laboratory. Each trial began with a fixation cross "+" presented at the center of the screen for 1000 ms, followed by segmented presentation of the poetic lines. Each segment was displayed for 600 ms with a 400 ms interval between segments, except for the final character of each line and the punctuation mark (comma or period), which had a 600 ms interval. Punctuation was presented separately to minimize sentence-final effects on EEG signals at critical words (Just & Carpenter, 1980; Luo & Zhou, 2010). After sentence presentation, a 1000 ms blank screen appeared, followed by four response options: "tonally regular and semantically congruous," "tonally irregular and semantically congruous," "tonally regular and semantically incongruous," and "tonally irregular and semantically incongruous." Participants were instructed to click on the option that best matched their understanding of the poem (Chen Qingrong, Yang Yiming, 2017). Option positions were counterbalanced so that each option appeared in each quadrant with 25% probability. The order of options was also counterbalanced across participants: half saw semantics listed before tonal patterns (e.g., "semantically congruous, tonally regular"), while the other half saw tonal patterns first (e.g., "tonally regular, semantically congruous"). After responding, a 1500 ms blank screen preceded the next trial. The experimental procedure is illustrated in Figure 1

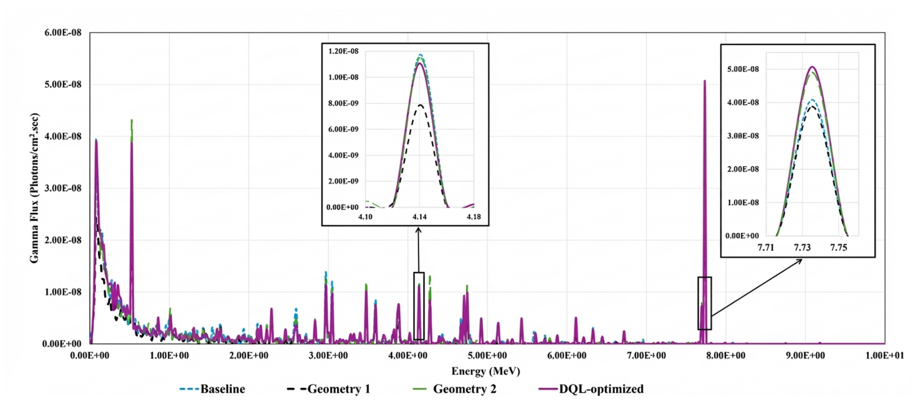

Figure 1: Figure 1

.

The formal experiment comprised four sets of materials, with each participant completing one set to ensure each couplet was presented only once per participant. Each set contained 100 experimental couplets (25 per condition) and 40 additional seven-character regulated verses as filler materials to prevent response bias. Thus, each participant viewed 140 couplets in pseudorandom order, with no more than two consecutive trials from the same condition. The experiment was divided into two blocks of 70 couplets each, with block order counterbalanced across participants and brief rest periods between blocks. Participants completed 10 practice trials before the formal experiment. The entire session lasted approximately 2 hours including preparation.

2.4 Data Analysis

2.4.1 Behavioral Data Analysis Accuracy was analyzed using logistic linear mixed-effects models. The dependent variable was participants' keypress response (0/1), with semantic reasonableness and tonal reasonableness as fixed effects and maximal random slopes and intercepts for both subjects and items (Barr et al., 2013). Since we did not match stroke counts of critical characters across the four conditions during material creation, stroke count was included as a covariate in the model: model <- glmer(response ~ semantic * tone + stroke + (1 + semantic * tone | sub) + (1 | item), family = binomial, control = glmerControl(optimizer = 'bobyqa'), data). In the model, response represents participants' keypress response, semantic represents semantic reasonableness, tone represents tonal reasonableness, stroke represents stroke count, subject represents participant, and item represents poetic line. If interactions were present, simple effects were analyzed using the emmeans package (Lenth & Lenth, 2018).

2.4.2 EEG Recording and Preprocessing EEG data were recorded using a 64-channel NeuroScan system with an Ag/AgCl electrode cap conforming to the international 10–20 system. Online bandpass filter parameters were set to 0.05–400 Hz, with a sampling rate of 1000 Hz. Horizontal electrooculogram (HEOG) was recorded from electrodes placed at the outer canthi of both eyes, and vertical electrooculogram (VEOG) from electrodes above and below the left eye. All electrodes were online-referenced to the left mastoid, with impedance maintained below 5 kΩ.

EEG data were offline-analyzed using Python' s MNE software package (Gramfort et al., 2013). Data were offline-referenced to the average of the left and right mastoids, then filtered with a 0.1–30 Hz digital bandpass filter. Independent component analysis was used for artifact correction, combined with manual inspection to remove ocular and muscular artifacts. Epochs were segmented from 200 ms before to 1000 ms after critical word onset, with baseline correction using the 200 ms pre-stimulus interval. Trials with amplitudes exceeding $±$75 V at any electrode were rejected. On average, 24.40 (0.96) trials were retained for S+T+, 24.67 (0.78) for S+T−, 24.33 (1.19) for S−T+, and 24.48 (0.71) for S−T−.

2.4.3 ERP Analysis Based on inspection of ERP waveforms, analysis windows were defined as N100 (80–120 ms), P200 (150–250 ms), N400 (250–450 ms), and LPC (500–700 ms). ERP analyses were conducted using linear mixed-effects models (LMM), which provide more robust parameter estimates than traditional ANOVA (Brauer & Curtin, 2018; Winter, 2013). The LMM included four factors: semantic reasonableness (congruous vs. incongruous), tonal reasonableness (regular vs. irregular), hemisphere (left vs. right), and region (anterior, central, posterior). Hemisphere and region were combined to form six regions of interest (ROIs), each containing 5–6 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). To control for individual differences and material variance, random intercepts for subjects and items, as well as random slopes for semantic and tonal reasonableness for subjects, were included in the model, along with stroke count as a covariate.

The final model was: model <- lmer(amplitude ~ semantic * tone * region * hemisphere + stroke + (1 + semantic * tone | subject) + (1 | item), control = lmerControl(optimizer = 'bobyqa'), data), where amplitude represents mean amplitude in the target time window, semantic represents semantic reasonableness, tone represents tonal reasonableness, region represents ROI, hemisphere represents hemisphere, stroke represents stroke count, subject represents participant, and item represents poetic line. If interactions with experimental conditions reached significance, simple effects were analyzed using the emmeans package. For multiple comparisons, p-values were adjusted using Tukey' s method.

2.4.4 CNN-LSTM-Based EEG Classification Analysis ERP analysis represents an encoding approach that reveals how the brain processes poetic tonal and semantic dimensions. Simultaneously, we conducted decoding analyses on EEG data to identify different poem types from EEG signals. Decoding analysis employs multivariate modeling and leverages fine-grained information in EEG signals to capture local patterns and complex features, better reflecting overall brain activity patterns (Cox & Savoy, 2003; Kriegeskorte & Douglas, 2019; Tong & Pratte, 2012). This bidirectional analysis, approaching from different angles, provides powerful evidence for understanding how the brain processes tonal rules and poetic semantics (Kriegeskorte & Douglas, 2019).

For decoding analysis, we employed a Convolutional-Long Short-Term Memory (CNN-LSTM) network to classify different poem types from EEG data (Bashivan et al., 2015; Tabar et al., 2017), with the procedure illustrated in Figure 2

. If the model could reliably classify the four poem types on test datasets, EEG signals must contain decodable information about semantics and tonal rules.

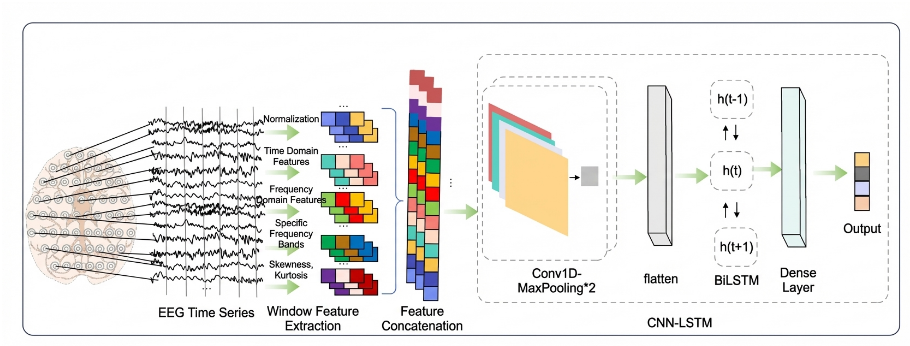

Figure 2: Figure 2

Following ERP analysis, we extracted 150–700 ms from raw signals as the analysis window (e.g., Borra & Magosso, 2021; Li et al., 2024; Tian et al., 2024). We then used 5-fold cross-validation, partitioning the dataset into five mutually exclusive subsets, training on four folds and validating on the remaining fold each round, cycling five times and averaging all test results as the model' s comprehensive evaluation. During preprocessing, sliding windows (0.1 s duration, 85% overlap) segmented signals, with z-score normalization per channel. For feature extraction, temporal features included standard deviation, mean, extrema, peak-to-peak values, skewness, and kurtosis to reflect amplitude fluctuations and distribution patterns. Frequency-domain features used Welch' s method to compute spectral energy distribution. Additionally, mean power across different frequency bands (delta, theta, alpha, beta, gamma) characterized spectral information, and spectral entropy quantified signal complexity. These five feature categories were concatenated as model input.

Figure 2 illustrates the CNN-LSTM-based EEG classification analysis pipeline. We employed a CNN-LSTM hybrid network to integrate local feature extraction with temporal dynamic modeling. The spatial feature extraction component used a CNN architecture with 60-channel raw EEG signals as input. The first convolutional layer used 64 one-dimensional kernels of width 5, followed by BatchNorm1d and ReLU activation; the second layer expanded channels to 128 with the same kernel width. Both layers employed 2$×$2 max pooling, reducing temporal dimension from 240 to 120 and then 60. Progressively increasing channel numbers enhanced extraction of complex spatial features, while gradual downsampling compressed temporal information while preserving sensitivity to subtle temporal variations.

The temporal modeling component consisted of two bidirectional LSTM layers, each with 128 hidden units. The bidirectional structure captured both forward and backward temporal information, outputting a 256-dimensional feature vector, with dropout between layers to mitigate overfitting. Subsequently, two fully connected layers reduced feature dimension from 256 to 128, then mapped to 4 output categories. Intermediate layers included BatchNorm1d, ReLU activation, and 0.2 dropout rate to improve training stability and generalization. Finally, a Softmax layer output class probability distributions.

3 Results

3.1 Behavioral Results

As shown in Figure 3

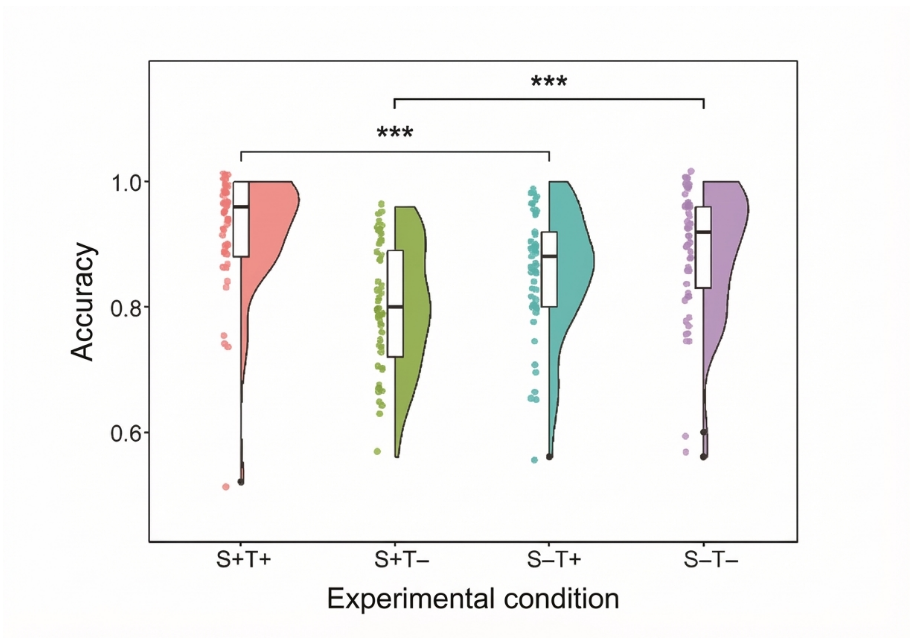

Figure 3: Figure 3

, mean accuracy rates and standard deviations across the four conditions were: S+T+ (M = 92.25%, SD = 0.09), S+T− (M = 80.42%, SD = 0.10), S−T+ (M = 85.42%, SD = 0.10), and S−T− (M = 88.75%, SD = 0.10). Logistic linear mixed model analysis revealed a significant main effect of tonal reasonableness (^{2}(1) = 7.01, p = 0.008), with higher accuracy for tonally regular than irregular lines. The interaction between tonal and semantic reasonableness was significant, ^{2}(1) = 52.45, $p < 0.001$. Simple effects analysis showed that when tonal patterns were regular, accuracy was significantly lower for semantically incongruous than congruous lines (OR = 0.37, SE = 0.09, z = -4.04, $p < 0.001$, 95% CI = [0.31, 0.45]); when tonal patterns were irregular, accuracy was significantly higher for semantically incongruous than congruous lines (OR = 2.48, SE = 0.53, z = 4.28, $p < 0.001$, 95% CI = [0.89, 6.94]).

Collapsing all error trials (Figure 4 [FIGURE:4]), the numbers of specific error types across conditions were: S+T+ = 93 trials, S+T− = 236 trials, S−T+ = 176 trials, S−T− = 135 trials. The two most common error types were misclassifying S+T− lines as S−T− (179 trials) and misclassifying S−T+ lines as S+T+ (143 trials).

3.2 ERP Results

Grand-averaged ERPs at critical words across the four conditions are shown in Figure 5 [FIGURE:5]. In the N100 window (80-120 ms), no significant main effects or interactions were observed ($Fs < 1.07$, $ps > 0.34$).

In the P200 window (150-250 ms), the interaction between tonal reasonableness and region was significant, $F(2, 27851.3) = 3.71$, $p = 0.024$. Simple effects tests (Figure 6a [FIGURE:6]) showed that at anterior scalp regions ($\beta = 0.51$, SE = 0.22, z = 2.26, p = 0.02, 95% CI = [0.07, 0.95]), tonally irregular lines elicited larger P200 than regular lines, with a marginal effect at central regions ($\beta = 0.39$, SE = 0.22, z = 1.75, p = 0.080, 95% CI = [-0.05, 0.83]), and no significant difference at posterior regions (p = 0.88).

In the N400 window (250-450 ms), the main effect of semantic reasonableness was significant ($F(1, 48.9) = 108.73$, $p < 0.001$, $p^2 = 0.69$), with semantically incongruous lines eliciting larger negativities than congruous lines. The interaction between semantic and tonal reasonableness was significant, $F(1, 47.9) = 7.18$, $p = 0.01$, $p^2 = 0.13$. Simple effects tests (Figure 6b) showed that when tonal patterns were regular, semantically incongruous lines elicited larger negativities than congruous lines ($\beta = -2.39$, SE = 0.27, z = -8.86, p < 0.001, 95% CI = [-2.92, -1.86]); when tonal patterns were irregular, the semantic effect persisted but was significantly attenuated ($\beta = -1.32$, SE = 0.27, z = -4.99, p < 0.001, 95% CI = [-1.84, -0.80]). Thus, regular tonal patterns facilitated poetic meaning processing.

In the LPC window (500-700 ms), the main effect of tonal reasonableness was significant ($F(1, 48.9) = 4.04$, $p = 0.05$), with tonally irregular lines eliciting larger LPC than regular lines. The interaction between semantic reasonableness and region was significant ($F(2, 27852.4) = 4.47$, $p = 0.01$); simple effects tests showed that semantically incongruous lines elicited larger negativities than congruous lines at anterior regions ($\beta = -0.64$, SE = 0.33, z = -1.97, p = 0.05, 95% CI = [-1.28, -0.002]), with no significant semantic effects at central or posterior regions ($ps > 0.64$). The interaction between semantic and tonal reasonableness was significant, $F(1, 48.3) = 9.61$, $p = 0.003$, $p^2 = 0.17$. Simple effects tests revealed that the tonal pattern effect was modulated by semantic reasonableness (Figure 6c): when semantics were incongruous, tonally irregular lines elicited larger LPC than regular lines ($\beta = 1.24$, SE = 0.37, z = 3.35, p < 0.001, 95% CI = [0.51, 1.96]); when semantics were congruous, the tonal pattern effect disappeared (p = 0.21).

3.3 CNN-LSTM Classification Results

Using convolutional neural networks to classify the four poem types from EEG signals, classification performance is shown in Figure 7 [FIGURE:7]. In Figure 7(1), precision refers to the proportion of samples predicted as a certain class that actually belong to that class, reflecting model accuracy in positive predictions; recall refers to the proportion of actual class members correctly predicted as such, reflecting coverage of positive samples; F1-Score is the harmonic mean of precision and recall, comprehensively evaluating both metrics. These are commonly used metrics in multi-class classification tasks. In Figure 7(2), true positive rate refers to the proportion of actual positives correctly identified, false positive rate reflects the proportion of actual negatives misclassified as positives, and AUC is the area under the ROC curve, reflecting overall classification ability, with the red dashed line indicating the random classifier reference.

As shown, AUC values for each category ranged from 0.95–0.97, with macro-averaged AUC (arithmetic mean of all category AUCs) and micro-averaged AUC (AUC calculated after merging all category samples) both reaching 0.96, fully validating the model' s discriminative ability across poem types. Permutation tests on precision, recall, and F1-Score showed that all real classification results were significantly higher than the random probability of 25% ($ps < 0.001$), indicating that the model could reliably identify lines with different tonal and semantic features.

4 Discussion

Poetry is an art form synthesizing language and music; "those who cannot appreciate the music of poetry may never grasp its subtleties" (Zhu Guangqian, 2012). Using EEG technology, the present study presented experienced poetry readers with strictly tonal-pattern-conforming five-character quatrains while manipulating both tonal and semantic reasonableness to investigate the cognitive neural mechanisms of tonal patterns and poetic meaning during classical Chinese poetry reading. Behavioral results showed a significant main effect of tonal reasonableness and an interaction between tonal and semantic reasonableness. ERP results further revealed that tonal effects in poetry reading occurred approximately 200 ms after critical word presentation, manifesting as an early P200 effect and interactions between tonal patterns and semantics on N400 and LPC components. Moreover, neural networks trained on EEG data effectively classified the four experimental conditions, further demonstrating that tonal rules and poetic meaning constitute important features in poetry processing. These findings indicate that readers develop specialized preferences for tonal regularity when reading classical Chinese poetry, and that this musical function of poetry interacts with meaning comprehension during middle-to-late processing stages.

The influence of tonal patterns on poetic meaning processing was evident in behavioral results. When tonal patterns were regular, participants showed greater difficulty judging semantically incongruous lines, consistent with research on rhyme processing in poetry reading (Chen et al., 2016; Chen Qingrong, Yang Yiming, 2017) and prosodic processing in spoken discourse comprehension (Rothermich et al., 2012). Error pattern analysis revealed that when tonal patterns were regular, participants tended to misjudge semantically incongruous lines as congruous. Conversely, when tonal patterns were irregular, accuracy for semantically congruous lines decreased, again demonstrating that poetic meaning comprehension is influenced by prosodic features, with irregular tonal patterns leading readers to perceive semantics as also unreasonable.

ERP results further revealed that tonal effects in poetry reading occurred approximately 200 ms after critical word presentation. Tonally irregular lines elicited larger P200 amplitudes than regular lines because participants generated top-down prosodic expectations based on poetry reading experience. When presented lines violated tonal patterns, the mismatch between bottom-up input and stored tonal representations produced expectation errors that incurred processing costs, resulting in larger P200 amplitudes (Chen et al., 2016). Although previous studies have found that early ERP components including P200 reflect formal expectation processing (e.g., phonological, orthographic), debate continues regarding the stage of language processing at which P200 reflects prosodic expectations. According to the sensory hypothesis, prosodic expectation effects on P200 reflect the influence of prosodic expectations on the lowest-level, prelexical processing based on basic perceptual features (Dambacher et al., 2009; Dikker et al., 2010; Dikker et al., 2009), whereas the word recognition hypothesis posits that prosodic expectations affect the lexical processing stage where phonological information is accessed (Connolly & Phillips, 1994; Nieuwland, 2019). Whether the P200 effect reflects prelexical or lexical processing requires further investigation, but it clearly demonstrates that tonal-pattern-based prosodic expectations influence early processing stages of line reading.

On the subsequent N400 component, poetic semantic processing was modulated by tonal reasonableness. Studies manipulating German sentence stress patterns (Rothermich et al., 2012) or Chinese phrase rhythm (Luo & Zhou, 2010) have similarly found interactions between prosodic features and semantics on N400, with regular prosodic patterns facilitating sentence processing. Previous research on poetry rhyme and semantic processing has also reported comparable results, where regular rhyme patterns modulated N400 effects for poetic semantics (Chen et al., 2016). Poetry reading involves a clear poetic function, wherein poetic texts direct readers to attend more to prosodic features (Hanauer, 1998). The poetic function focuses on internal connections between poetic words, manifested as structural features of similarity, regularity, and repetition. Tonal patterns, as a typical prosodic feature of classical Chinese poetry, exhibit clear structures of similarity, contrast, and repetition. Based on implicit or explicit learning experiences, readers consciously attend to tonal structures when reading classical regulated verse, expecting oblique tones after level tones and vice versa. When encountered characters violate tonal rules, like discovering an unexpectedly high or low step while climbing stairs, attentional resources are rapidly allocated to prosodic features, thereby attenuating semantic effects.

The interactive influence of tonal patterns and poetic meaning persisted into the LPC time window. Tonally irregular conditions elicited larger LPC than regular conditions, consistent with previous phonological and prosodic studies showing that phonological or tonal violations evoke larger LPC (Zhang Zhenghua et al., 2021; Hu et al., 2012; Schirmer et al., 2005; Zhang et al., 2023). Similarly, manipulating pitch of final words in sentences or final notes in music also elicits larger LPC (Magne et al., 2006; Schön et al., 2004). These findings collectively suggest that LPC may reflect conflict detection and resolution (Bornkessel-Schlesewsky & Schlesewsky, 2008; Kim & Osterhout, 2005) and more general integration and reanalysis processes (Brouwer et al., 2012; Liu et al., 2010).

Astésano et al. (2004) manipulated sentence intonation using splicing techniques, such as combining statement beginnings with question endings to create intonationally incongruous conditions. Although this prosodic manipulation was somewhat coupled with pragmatic factors (Rothermich et al., 2012), its control over global pitch patterns is highly comparable to the present tonal pattern manipulation. That study' s results closely mirror ours: intonation violations elicited larger late positivities and interacted with semantics, producing larger prosodic effects when semantics were incongruous. During late stages, readers mobilized cognitive resources to attempt reanalysis of incongruous lines, consciously analyzing prosodic structural rules of similarity and repetition to reconstruct the nonlinear hierarchical structure of poetic texts (Hanauer, 1998). Specifically, when semantics were incongruous, irregular tonal patterns exacerbated overall disharmony, yielding larger LPC effects; when semantics were congruous, poetic meaning buffered prosodic imperfections, eliminating the effect.

Notably, Chen et al. (2016) found that rhyme violations did not elicit LPC effects. This processing difference may stem from two factors: First, rhymes have stronger binding force and influence in lexical identification and semantic constraints, making rhyme processing relatively faster than tonal processing (Hu et al., 2012; Li et al., 2014; Tong et al., 2008). Second, rhyme and tonal pattern are distinct prosodic features in poetry. Rhyme primarily involves phonological matching at word endings, representing relatively local phonological template detection that typically does not require large-scale structural reintegration. In contrast, tonal patterns constitute a set of tonal rules within and across lines, representing more global structural rules that require reanalysis or repair of poetic prosodic structure, thereby eliciting LPC in late stages.

We also employed CNN-LSTM networks to classify different poem types from EEG data. This decoding analysis represents the inverse process of encoding, aiming to reveal encoded content, sometimes described as "reading the brain" or "cracking the neural code" (Haynes & Rees, 2006; Tong & Pratte, 2012). EEG classification analysis revealed that the four experimental conditions could be reliably distinguished from EEG data, indicating that semantic and tonal features of poetry exist in EEG signals in a format utilizable by classifiers (Kriegeskorte & Douglas, 2019). Thus, decoding analysis further revealed the information encoded by the brain about poetry, confirming that tonal rules and poetic meaning constitute important features in classical Chinese poetry processing.

The present study simultaneously manipulated tonal patterns and semantic features of classical Chinese poetry, using EEG to observe the cognitive neural mechanisms of prosody and poetic meaning during poetry reading. Results demonstrated tonal pattern effects in classical Chinese poetry reading, wherein tonal patterns influenced lexical phonological representation at early stages and constrained and integrated with poetic meaning at middle-to-late stages, ultimately achieving poetry comprehension and appreciation. Moreover, neural network models trained on EEG data effectively distinguished the four experimental conditions. These findings indicate that readers prefer rhythmically harmonious and well-patterned lines in poetry reading, consequently expecting harmony between sound and meaning. The results validate the neurocognitive poetics model (Jacobs, 2015), which posits interactions between foreground features (e.g., tonal patterns, rhyme) and background features (e.g., lexical and sentence comprehension) in poetry reading, and further reveal the temporal dynamics and patterns of interaction between prosodic features and poetic meaning. The findings also illuminate how pitch-oriented prosodic features like tonal combinations co-process with semantics, supporting the three-stage neurocognitive model and interactive activation model: prosodic effects operate across different stages of language processing rather than being limited to early stages, and prosody and semantics exhibit bidirectional dynamic interactions rather than unidirectional flow.

The present study focused on experienced poetry readers; future research should examine how general readers process poetic tonal patterns and semantics and identify differences and connections between these groups. Additionally, according to the dual-pathway neural model (Schwartze & Kotz, 2013), the human brain possesses two pathways for processing rhythmic and content dimensions of stimuli, representing a fundamental brain function. Future research should investigate the characteristics of separate processing pathways for poetic prosody and meaning and their potential interaction patterns. Finally, given poetry and music' s close associations in prosodic structure and emotional expression, subsequent studies should systematically analyze similarities and differences in their cognitive processing and examine how musical experience influences poetry prosody and emotional processing mechanisms.

References

Aristotle, A. (1968). *Poetics* (Vol. 9). Clarendon Press Oxford.

Astésano, C., Besson, M., & Alter, K. (2004). Brain potentials during semantic and prosodic processing in French. *Cognitive Brain Research, 18*(2), 172-184.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255-278.

Bashivan, P., Rish, I., Yeasin, M., & Codella, N. (2015). Learning representations from EEG with deep recurrent-convolutional neural networks. *arXiv preprint arXiv:1511.06448*.

Beck, J., & Konieczny, L. (2020). Rhythmic subvocalization: An eye-tracking study on silent poetry reading. *Journal of Eye Movement Research, 13*(3), 1-40.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on "semantic P600" effects in language comprehension. *Brain Research Reviews, 59*(1), 55-73.

Borra, D., & Magosso, E. (2021). Deep learning-based EEG analysis: investigating P3 ERP components. *Journal of Integrative Neuroscience, 20*(4), 791-811.

Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods, 23*(3), 389-411.

Brouwer, H., Crocker, M. W., Venhuizen, N. J., & Hoeks, J. C. (2017). A neurocomputational model of the N400 and the P600 in language processing. *Cognitive Science, 41*, 1318-1352.

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about semantic illusions: rethinking the functional role of the P600 in language comprehension. *Brain Research, 1446*, 127-143.

Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience, 5*(1), 34-44.

Chen, Q., Zhang, J., Xu, X., Scheepers, C., Yang, Y., & Tanenhaus, M. K. (2016). Prosodic expectations in silent reading: ERP evidence from rhyme scheme and semantic congruence in classic Chinese poems. *Cognition, 154*, 11-21.

Coch, D., Grossi, G., Skendzel, W., & Neville, H. (2005). ERP nonword rhyming effects in children and adults. *Journal of Cognitive Neuroscience, 17*(1), 168-182.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge.

Connolly, J. F., & Phillips, N. A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience, 6*(3), 256-266.

Cox, D. D., & Savoy, R. L. (2003). Functional magnetic resonance imaging (fMRI) "brain reading": Detecting and classifying distributed patterns of fMRI activity in human visual cortex. *NeuroImage, 19*(2), 261-270.

Culler, J. (1975). *Structuralist poetics: Structuralism, linguistics and the study of literature*. Routledge.

Dambacher, M., Rolfs, M., Göllner, K., Kliegl, R., & Jacobs, A. M. (2009). Event-related potentials reveal rapid verification of predicted visual input. *PLoS One, 4*(3), e5047.

Dikker, S., Rabagliati, H., Farmer, T. A., & Pylkkänen, L. (2010). Early occipital sensitivity to syntactic category is based on form typicality. *Psychological Science, 21*(5), 629-634.

Dikker, S., Rabagliati, H., & Pylkkänen, L. (2009). Sensitivity to syntax in visual cortex. *Cognition, 110*(3), 293-321.

Fodor, J. A. (1983). *The modularity of mind.* MIT press.

Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language, 50*(3), 259-281.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in cognitive sciences, 6*(2), 78-84.

Friederici, A. D. (2012). The cortical language circuit: From auditory perception to sentence comprehension. *Trends in Cognitive Sciences, 16*(5), 262-268.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T. L., Parkkonen, L., & Hämäläinen, M. S. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience, 7*, 267.

Hagoort, P., Brown, C. M., & Osterhout, L. (1999). The neurocognition of syntactic processing. *The Neurocognition of Language*, 273-316.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science, 304*(5669), 438-441.

Hanauer, D. (1998). The genre-specific hypothesis of reading: Reading poetry and encyclopedic items. *Poetics, 26*(2), 63-80.

Haynes, J., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience, 7*(7), 523-534.

Hirst, D., & Di Cristo, A. (Eds.). (1998). *Intonation systems: A survey of twenty languages*. Cambridge University Press.

Hu, J., Gao, S., Ma, W., & Yao, D. (2012). Dissociation of tone and vowel processing in Mandarin idioms. *Psychophysiology, 49*(9), 1179-1190.

Jacobs, A. M. (2015). Neurocognitive poetics: methods and models for investigating the neuronal and cognitive-affective bases of literature reception. *Frontiers in Human Neuroscience, 9*, 186.

Jakobson, R. (1960). Linguistics and poetics. In T. Sebeok (Ed.), *Style in language* (pp. 350-377). Cambridge, MA: MIT Press.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review, 87*(4), 329-354.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes, 15*(2), 159-201.

Kaan, E., & Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience, 15*(1), 98-110.

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language, 52*(2), 205-225.

Kriegeskorte, N., & Douglas, P. K. (2019). Interpreting encoding and decoding models. *Current Opinion in Neurobiology, 55*, 167-179.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*(1), 621-647.

Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge University Press.

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience, 9*(12), 920-933.

Lenth, R., & Lenth, M. R. (2018). Package 'lsmeans'. *The American Statistician, 34*(4), 216-221.

Li, W., Wang, L., & Yang, Y. (2014). Chinese tone and vowel processing exhibits distinctive temporal characteristics: An electrophysiological perspective from classical Chinese poem processing. *PLOS ONE, 9*(1), e85683.

Li, W., & Yang, Y. (2010). Perception of Chinese poem and its electrophysiological effects. *Neuroscience, 168*(3), 757-768.

Li, X., Shao, X., Xia, J., & Xu, X. (2019). The cognitive and neural oscillatory mechanisms underlying the facilitating effect of rhythm regularity on speech comprehension. *Journal of Neurolinguistics, 49*, 155-167.

Li, S., Zhang, T., Yang, F., Li, X., Wang, Z., & Zhao, D. (2024). A Dynamic Multi-Scale Convolution Model for Face Recognition Using Event-Related Potentials. *Sensors, 24*(13), 4368.

Liu, Y., Li, P., Shu, H., Zhang, Q., & Chen, L. (2010). Structure and meaning in Chinese: An ERP study of idioms. *Journal of Neurolinguistics, 23*(6), 615-630.

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Physiology, 31*(3), 291-308.

Luo, Y., & Zhou, X. (2010). ERP evidence for the online processing of rhythmic pattern during Chinese sentence reading. *NeuroImage, 49*(3), 2836–2849.

Magne, C., Astésano, C., Aramaki, M., Ystad, S., Kronland-Martinet, R., & Besson, M. (2007). Influence of syllabic lengthening on semantic processing in spoken French: behavioral and electrophysiological evidence. *Cerebral cortex, 17*(11), 2659–2668.

Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience, 18*(2), 199–211.

McClelland, J. L., & Rumelhart, D. E. (1981). An Interactive Activation Model of Context Effects in Letter Perception: I. An Account of Basic Findings. *Psychological Review, 88*(5), 375-407.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Physiology, 24*(4), 375–425.

Nietzsche, F. (2003). *The birth of tragedy.* Oxford University Press.

Nieuwland, M. S. (2019). Do 'early' brain responses reveal word form prediction during language comprehension? A critical review. *Neuroscience & Biobehavioral Reviews, 96*, 367-400.

Nieuwland, M. S., Barr, D. J., Bartolozzi, F., Busch-Moreno, S., Darley, E., Donaldson, D. I., Ferguson, H. J., Fu, X., Heyselaar, E., & Huettig, F. (2020). Dissociable effects of prediction and integration during language comprehension: Evidence from a large-scale study using brain potentials. *Philosophical Transactions of the Royal Society B, 375*(1791), 20180522.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*(6), 785-806.

Rothermich, K., & Kotz, S. A. (2013). Predictions in speech comprehension: fMRI evidence on the meter–semantic interface. *NeuroImage, 70*, 89–100.

Rothermich, K., Schmidt-Kassow, M., & Kotz, S. A. (2012). Rhythm' s gonna get you: Regular meter facilitates semantic sentence processing. *Neuropsychologia, 50*(2), 232–244.

Rugg, M. D. (1984). Event-related potentials in phonological matching tasks. *Brain and language, 23*(2), 225–240.

Schirmer, A., Tang, S. L., Penney, T. B., Gunter, T. C., & Chen, H. C. (2005). Brain responses to segmentally and tonally induced semantic violations in Cantonese. *Journal of Cognitive Neuroscience, 17*(1), 1–12.

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Physiology, 41*(3), 341-349.

Schwartze, M., & Kotz, S. A. (2013). A dual-pathway neural architecture for specific temporal prediction. *Neuroscience & Biobehavioral Reviews, 37*(10), 2587-2596.

Tabar, Y.R., & Halici, U. (2017). A novel deep learning approach for classification of EEG motor imagery signals. *Journal of Neural Engineering, 14*(1), 016003.

Tian, X., Zhu, L., Zhang, M., Wang, S., Lu, Y., Xu, X., Jia W., Zheng, Y., & Song, S. (2024). Social anxiety prediction based on ERP features: A deep learning approach. *Journal of Affective Disorders, 367*, 545-553.

Tong, F., & Pratte, M. S. (2012). Decoding patterns of human brain activity. *Annual Review of Psychology, 63*(1), 483-509.

Tong, Y., Francis, A. L., & Gandour, J. T. (2008). Processing dependencies between segmental and suprasegmental features in mandarin Chinese. *Language and Cognitive Processes, 23*(5), 689-708.

Tsur, R. (1983). What is cognitive poetics? Katz Research Institute for Hebrew Literature, Tel Aviv University.

Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications. *arXiv preprint arXiv:1308.5499.*

Yaron, I. (2002). Processing of obscure poetic texts: mechanisms of selection. *Journal of Literary Semantics, 31*(2), 147-166.

Zeman, A., Milton, F., Smith, A., & Rylance, R. (2013). By heart: An fMRI study of brain activation by poetry and prose. *Journal of Consciousness Studies, 20*(9-10), 132-158.

Zhang, Z., Zhang, H., Sommer, W., Yang, X., Wei, Z., & Li, W. (2023). Musical training alters neural processing of tones and vowels in classic Chinese poems. *Brain and Cognition, 166*, 105952.

Chen, Q., & Dou, F. (2021). Constructing a Questionnaire to Test the Knowledge of Ancient Chinese Poetry: A Preliminary Result. *Journal of Nanjing Normal University (Social Science Edition), 3*, 73-82. [陈庆荣, 窦芬. (2021). 古诗知识经验测评问卷的初步编制. 南京师大学报 (社会科学版), 3, 73-82.]

Chen, Q., & Yang, Y. (2017). Cognitive Mechanisms in Reading Ancient Chinese Poetry: Evidence from Eye Movements. *Social Sciences in China, 3*, 48-76+205. [陈庆荣, 杨亦鸣. (2017). 古诗阅读的认知机制: 来自眼动的证据. 中国社会科学 (3), 48-76+205.]

Luo, Z. (1993). *Chinese Dictionary Compendium.* Shanghai: Publishing House of Chinese Dictionary Compendium. [罗竹风. (1993). 汉语大词典. 上海: 汉语大词典

出版社.]

Wang, L. (2012). *The metrics of Chinese poetry.* Beijing: Zhonghua Book Company. [王力. (2012). 诗词格律. 北京：中华书局.]

Zhang Y. H. (2015). Cognitive Poetics and Interdisciplinary Research on Literary Comprehension. *Journal of Tsinghua University (Philosophy and Social Sciences), 30*(2), 139-147+190. [张叶鸿. (2015). 认知诗学与跨学科文学理解研究. 清华大学学报 (哲学社会科学版), 30(2), 139-147+190.]

Zhang, Z., Han, M., Zhang, F., & Li, W. (2020). Musical training improves rhythm integrative processing of classical Chinese poem. *Acta Psychologica Sinica, 52*(7), 847-860. [张政华, 韩梅, 张放, 李卫君. (2020). 音乐训练促进诗句韵律整合加工的神经过程. 心理学报, 52(7), 847-860.]

Zhang, Z., Zhang, H., Wang, B., Zheng, Z., Zhao, L., & Li, W. (2021). The processing of the tone and the vowel in poems under different tasks: Evidence from ERPs. *Studies of Psychology and Behavior, 19*(6), 728-735. [张政华, 张航, 王兵, 郑子龙, 赵黎明, 李卫君. (2021). 不同任务下声调和韵母在诗句中的加工过程——来自 ERP 的证据. 心理与行为研究, 19(6), 728-735.]

Zhu, G. Q. (2012). *On Poetry.* Beijing: Zhonghua Book Company. [朱光潜. (2012). 诗论. 北京：中华书局.]