

The Effect of Lyrics on Musical Emotion Processing: A Behavioral and ERP Study (Postprint)

Authors: Zhang Weixia, Wang Wanqi, Zhou Linshu, Jiang Cunmei

Date: 2018-10-26T00:00:00+00:00

Abstract

This study investigates the influence of lyrics on musical emotion processing. Experiment 1 employed an affective priming paradigm, with music excerpts with and without lyrics as prime stimuli, and face pictures that were either congruent or incongruent with the musical emotion as target stimuli; the listeners' task was to judge the emotion of the target faces both quickly and accurately. The results showed that, regardless of whether the music contained lyrics, listeners' responses were faster and more accurate under congruent conditions than under incongruent conditions, indicating that listeners can process the emotional information conveyed by music. Experiment 2 further explored the neural mechanisms underlying the influence of lyrics on musical emotion processing through electrophysiological methods. The findings revealed that although listeners exhibited priming effects for both lyrical and non-lyrical music emotion processing, the non-lyrical music condition produced an N400 effect in the 250–450 ms time window, whereas the lyrical music condition elicited an LPC effect in the 500–700 ms time window, suggesting that lyrics influence the temporal course of the brain's processing of musical emotion. The results of this study will provide, to a certain extent, a basis for exploring the relationship between music and language.

Full Text

Preamble

Effects of Lyrics on the Processing of Musical Emotion: A Behavioral and ERP Study

ZHANG Weixia¹, WANG Wanqi², ZHOU Linshu³, JIANG Cunmei³

(¹ Department of Psychology, College of Education, Shanghai Normal University;

² School of Finance and Business, Shanghai Normal University; ³ Music College, Shanghai Normal University, Shanghai 200234, China)

Abstract

This study investigated the influence of lyrics on the processing of musical emotion. In Experiment 1, using an affective priming paradigm, musical excerpts with and without lyrics served as prime stimuli, and facial pictures with emotions either congruent or incongruent with the musical emotion served as target stimuli. Participants were asked to judge the emotion of the target faces as quickly and accurately as possible. The results showed that, regardless of whether the music contained lyrics, participants responded faster and more accurately under congruent conditions than incongruent conditions, indicating that listeners can process emotional information conveyed by music. Experiment 2 further explored the neural mechanisms underlying the influence of lyrics on musical emotion processing using electrophysiological measures. The findings revealed that although both lyrical and non-lyrical music conditions produced priming effects, the non-lyrical music condition elicited an N400 effect in the 250–450 ms time window, whereas the lyrical music condition elicited an LPC effect in the 500–700 ms time window. These results suggest that lyrics affect the temporal course of neural processing of musical emotion. The present findings provide empirical evidence for understanding the relationship between music and language.

Keywords: musical emotion; language; lyrics; time course; ERPs

Classification Code: B842

Music and language are unique to human society and represent two important means of human communication. As early as 1871, Darwin proposed the protolanguage hypothesis, suggesting that language and music may share a common origin. For early humans, the primary function of both language and music might have been emotional expression (Thompson, Marin, & Stewart, 2012). With increasing socialization, language and music gradually diverged and evolved along different paths (Mithen, 2006; Perlovsky, 2011). Language developed into a symbolic communication system with explicit semantics, whereas music became an important medium for emotional expression (Jackendoff, 2009).

Music may have emerged before language, and humans might have communicated through music-like means before language developed (Darwin, 1871; Levman, 1992). Does linguistic information with explicit semantics facilitate the emotional expression of music? Or, as an independent communication system, does music's emotional expression not rely on semantic information, which might instead interfere with its emotional conveyance? Songs contain both lyrics and melody, possessing characteristics of both language and music. In musical works, performers can present only the melody (e.g., singing with meaningless syllables like “la”) or present both melody and lyrics simultaneously. Therefore, many studies have examined songs to reveal how explicit semantic information influences the processing of musical emotion, which may help clarify the relationship between language and music in emotional processing.

Previous research has primarily addressed this question at the behavioral level. Studies have shown that lyrics indeed affect listeners' processing of musical emotion (Ali & Peynircioglu, 2006; Mori & Iwanaga, 2013; Stratton & Zalanowski, 1994). For example, Stratton and Zalanowski (1994) used the song "Why I Was Born" to create three versions: lyrics only, music only, and lyrics with piano accompaniment. Participants rated the pleasantness of the stimuli and completed mood measures before and after listening. The results showed that participants rated the non-lyrical version as significantly more pleasant than the two lyrical versions. Non-lyrical music reduced negative emotions and enhanced positive emotions. However, when lyrics were present, the opposite pattern emerged: participants' negative emotions increased while positive emotions decreased. These findings suggest that lyrics, whether presented alone or with melody, can influence listeners' emotional processing of music. The impact of lyrics on musical emotion processing has also been confirmed by computational studies, which found that including lyrical information in algorithms affects the accuracy (Laurier, Lartillot, Eerola, & Toiviainen, 2008) and hit rates (Hu, Downie, & Ehmann, 2009) of automatic music mood classification.

However, other studies have found that listeners' emotional experiences before and after listening to music with lyrics did not change significantly, regardless of whether the lyrics were spoken or sung (Galizio & Hendrick, 1972). Similarly, by measuring listeners' mood changes before and after music listening, Sousou (1997) found that lyrics did not affect listeners' emotional experience of music.

The discrepancies in previous findings may stem from two sources. First, the experimental conditions varied across studies. Specifically, Galizio and Hendrick (1972) and Sousou (1997) did not include a non-lyrical control condition, whereas other studies (Ali & Peynircioglu, 2006; Brattico et al., 2011; Stratton & Zalanowski, 1994) included such a condition. Second, researchers may not have used objective measures to quantify listeners' processing of musical emotion. Most previous studies employed subjective rating methods, such as the Multiple Affect Adjective Check List-Revised (Stratton & Zalanowski, 1994) or self-designed rating scales (Sousou, 1997; Mori & Iwanaga, 2013), with rating content varying from pleasantness (Stratton & Zalanowski, 1994; Mori & Iwanaga, 2013) to intensity (Ali & Peynircioglu, 2006). These methodological differences have made it difficult to compare results across studies.

Based on these considerations, the present study used electrophysiological methods to investigate the influence of lyrics on musical emotion processing, building upon behavioral research. Experiment 1 examined whether listeners can process emotions conveyed by music with and without lyrics using cognitive-behavioral methods. If Experiment 1 demonstrates that listeners can process emotions regardless of lyrical content, Experiment 2 further investigates whether there are differences in how the brain processes emotions conveyed by these two types of music using electrophysiological methods. Both experiments employed an affective priming paradigm. The musical stimuli consisted of 120 excerpts performed by a vocalist, creating two conditions: a lyrical condition (sung with Chinese

lyrics) and a non-lyrical condition (sung with meaningless syllables “la”). Both versions served as prime stimuli, with emotional face pictures that were either affectively congruent or incongruent with the musical emotion serving as target stimuli. If music can prime the processing of emotional faces, it implies that listeners can process the emotional information conveyed by music. If the pattern of emotional processing differs between lyrical and non-lyrical music conditions, it suggests that lyrics influence musical emotion processing.

Experiment 1

Experiment 1 investigated at the behavioral level whether listeners can process emotional information from music with and without lyrics. To exclude the influence of familiarity, all musical stimuli were selected from European opera excerpts, as the participants were non-musically trained university students who primarily listen to popular music and rarely encounter European opera. Additionally, to control for timbre effects, all musical stimuli were performed by human voice. In the non-lyrical version, the vocalist sang using meaningless syllables “la,” while in the lyrical version, the vocalist sang Chinese lyrics (published Chinese translations). To ensure the validity of the two versions, three pre-tests were conducted (see Stimuli and Procedure section for details). If both lyrical and non-lyrical music can prime listeners’ processing of emotional faces, it would demonstrate that listeners can process emotional information conveyed by both types of music.

Participants

Forty non-musically trained university students voluntarily participated in the experiment. All participants were right-handed, had normal hearing, normal or corrected-to-normal vision, no history of neurological or psychiatric disorders, and signed informed consent forms prior to the experiment. Data from eight participants were excluded due to misunderstanding of instructions, leaving 32 valid participants (mean age = 24.47 ± 1.65 years, 15 male).

Stimuli and Procedure

The 120 original musical stimuli were selected from European opera excerpts, with half expressing happiness and half expressing sadness. All stimuli were recorded by a vocalist with 18 years of professional bel canto training. Each original musical excerpt was performed with both Chinese lyrics and meaningless syllables “la,” creating lyrical and non-lyrical versions. Before recording, the vocalist was instructed to carefully consider the performance style for each excerpt, striving for consistency between the two versions in dynamics (the performer’s handling of loudness variations), rubato (the performer’s flexible timing within certain limits), phrasing (the breathing points or phrase divisions), and overall performance quality (the overall sound quality).

The recorded music was edited using Adobe Audition CS6 (Adobe Systems Inc)

and Goldwave. The musical stimuli had an average duration of 17 seconds (range: 10-25 seconds), were monophonic, with a sampling rate of 22.050 kHz and 16-bit resolution. The average loudness was normalized to -7 dB, with a 1-second fade-out.

To ensure stimulus validity, three pre-tests were conducted. The first pre-test ensured that the two versions did not differ significantly in performance quality and style. Sixteen music majors (with 18 years of professional music training) rated the consistency between lyrical and non-lyrical versions in rubato, dynamics, phrasing, and overall performance quality (1 = very different, 4 = uncertain, 7 = very similar). Only stimuli with average scores above 4 were selected. The second pre-test ensured that listeners could clearly hear the lyrics in the lyrical version. Sixteen non-musician university students who did not participate in the formal experiment rated the clarity of the lyrics (1 = unclear, 3 = uncertain, 5 = clear). Only stimuli with average scores above 4 were selected. After these two pre-tests, 80 stimuli remained for each version. The third pre-test ensured that the relationship between prime music and target faces was affectively congruent or incongruent. From the Chinese Facial Affective Picture System (Gong, Huang, Wang, & Luo, 2011), 80 sad and 80 happy faces were selected as potential target stimuli, with equal numbers of male and female faces. Using Adobe Photoshop CS, face images were adjusted to 102×768 pixels with 16-bit resolution. Each of the 160 musical stimuli was presented twice, paired with both emotionally congruent and incongruent faces, creating 320 music-face pairs. Sixteen non-musician university students who did not participate in the formal experiment rated the congruence of each pair (1 = very incongruent, 5 = uncertain, 9 = very congruent). Pairs scoring above 7 were considered affectively congruent, while those scoring below 3 were considered incongruent. Ultimately, 60 stimuli for each music version met the criteria, forming four conditions: lyrical music-congruent, lyrical music-incongruent, non-lyrical music-congruent, and non-lyrical music-incongruent (see Figure 1 [Figure 1: see original paper]), with 60 trials per condition.

A 2 (congruence: congruent, incongruent) \times 2 (lyrics: lyrical music, non-lyrical music) repeated-measures ANOVA on the emotion congruence ratings for the selected music-face pairs revealed a significant main effect of congruence, $F(1, 59) = 2318.45$, $p < 0.001$, $\eta^2 = 0.98$. Whether for lyrical music (congruent: $M = 7.37$, $SD = 0.53$; incongruent: $M = 2.70$, $SD = 0.57$) or non-lyrical music (congruent: $M = 7.33$, $SD = 0.55$; incongruent: $M = 2.66$, $SD = 0.56$), congruent pairs were rated significantly higher than incongruent pairs. No other effects were significant (all $ps > 0.09$). These results indicate that, regardless of lyrical content, affectively congruent and incongruent conditions differed significantly. Additionally, for the final 60 lyrical music stimuli selected through pre-test 3, the average lyric clarity rating was 4.42 ($SD = 0.23$). The consistency scores between lyrical and non-lyrical versions were as follows: rubato (5.44 ± 0.48), dynamics (5.56 ± 0.47), phrasing (5.56 ± 0.47), and overall performance quality (5.52 ± 0.44).

The 240 pairs were distributed into two stimulus lists using a Latin square design, with 120 trials per list, containing four experimental conditions. Each participant completed only one list. Stimuli were presented in a pseudo-random order, with the same prime or target stimulus separated by at least eight trials. During the experiment, a 1000 ms black fixation point was followed by the prime music played through Philips SHM1900 headphones. Immediately after the music ended, an emotional face picture appeared. Participants judged the emotion of the face as quickly and accurately as possible, pressing “F” for happy and “J” for sad. The mapping of emotion type (happy/sad) to response keys (F/J) was counterbalanced across participants. After responding, participants pressed the spacebar to begin the next trial. Four practice trials were provided before the formal experiment to ensure familiarity with the procedure. To further exclude familiarity effects, after the formal experiment, participants were asked to report the titles of the musical excerpts. If a participant could report more than one keyword from a title, it would indicate familiarity with that excerpt. No participant could report any keywords from the titles.

Results and Discussion

Figure 2 [Figure 2: see original paper] shows the mean accuracy (a) and reaction time (b) across the four experimental conditions. For accuracy, a 2 (lyrics: lyrical music, non-lyrical music) \times 2 (congruence: congruent, incongruent) repeated-measures ANOVA revealed a significant main effect of congruence, $F(1, 31) = 9.80$, $p = 0.004$, $\eta^2 = 0.24$, indicating higher accuracy in the congruent condition ($M = 97.86\%$, $SD = 2.74$) than in the incongruent condition ($M = 94.71\%$, $SD = 6.79$). The main effect of lyrics was also significant, $F(1, 31) = 7.64$, $p = 0.01$, $\eta^2 = 0.20$, showing higher accuracy in the non-lyrical condition ($M = 97.17\%$, $SD = 4.03$) than in the lyrical condition ($M = 95.40\%$, $SD = 6.39$). The interaction was not significant ($p > 0.05$).

For reaction time, a 2 (lyrics: lyrical music, non-lyrical music) \times 2 (congruence: congruent, incongruent) repeated-measures ANOVA showed a significant main effect of congruence, $F(1, 31) = 14.38$, $p = 0.001$, $\eta^2 = 0.32$, with faster responses in the congruent condition ($M = 1051.83$ ms, $SD = 389.35$) than in the incongruent condition ($M = 1166.45$ ms, $SD = 503.22$). The main effect of lyrics was significant, $F(1, 31) = 4.42$, $p = 0.04$, $\eta^2 = 0.13$, with longer reaction times in the lyrical condition ($M = 1136.59$ ms, $SD = 486.83$) than in the non-lyrical condition ($M = 1081.69$ ms, $SD = 415.88$). The interaction between lyrics and congruence was not significant ($p > 0.05$).

Consistent with previous behavioral findings (Goerlich, Wittman, Aleman, & Martens, 2011; Wang & Qin, 2016; Zhang, Li, Gold, & Jiang, 2010), these results demonstrate that participants responded faster and more accurately under affectively congruent conditions compared to incongruent conditions. This indicates that both lyrical and non-lyrical music can prime listeners’ processing of emotional faces, meaning that listeners can process emotions from both types of music, consistent with previous research (Morton & Trehub, 2007).

Experiment 2

The results of Experiment 1 showed that listeners can process emotional information from music regardless of lyrical content. However, behavioral experiments cannot reveal the dynamic temporal changes in brain processing of musical emotion. Therefore, Experiment 2 used electrophysiological methods to further explore the neural mechanisms through which lyrics influence musical emotion processing. Only one previous study has used functional magnetic resonance imaging (fMRI) to investigate the neural mechanisms of lyrics' influence on musical emotion processing (Brattico et al., 2011). Unfortunately, that study used vocal music with lyrics and instrumental music without lyrics, making it impossible to exclude the role of timbre in musical emotion processing. Indeed, numerous studies have shown that timbre significantly influences musical emotion processing (Behrens & Green, 1993; Hailstone et al., 2009; Franco, Chew, & Swaine, 2017). The present experiment investigated whether lyrics affect the neural processing of musical emotion when timbre is controlled.

Previous studies using affective priming paradigms have focused on the N400 and late positive component (LPC). Research has shown that affectively incongruent conditions elicit larger N400 responses than congruent conditions, reflecting the brain's detection of emotional conflict (Schirmer, Kotz, & Friederici, 2002; Zhang, Lawson, Guo, & Jiang, 2006) and the integration of emotional information (Kamiyama, Abla, Iwanaga, & Okanoya, 2013; Zhang et al., 2010). Similarly, incongruent conditions elicit larger LPC responses than congruent conditions (Herring, Taylor, White, & Crites, 2011; Werheid, Alpay, Jentzsch, & Sommer, 2005; Zhang et al., 2010), suggesting that processing affectively incongruent conditions requires greater attentional engagement (Zhang, Kong, & Jiang, 2012; Zhang et al., 2010). Given that Experiment 1 and previous behavioral studies (Morton & Trehub, 2007) have shown that listeners can process emotional information from both lyrical and non-lyrical music, and that lyrics influence musical emotion processing (Ali & Peynircioglu, 2006; Stratton & Zalanowski, 1994), we hypothesized that both lyrical and non-lyrical music would produce priming effects (i.e., N400 or LPC effects). However, the priming effects might differ between the two conditions.

Method

Participants Twenty non-musically trained university students participated. Four participants were excluded due to excessive EEG artifacts, leaving 16 valid participants (mean age = 23.88 ± 1.36 years, 7 male). All participants were right-handed, had normal hearing, normal or corrected-to-normal vision, no history of psychiatric disorders or brain damage, and signed informed consent forms. Participants received compensation after the experiment.

Stimuli and Procedure The stimuli were identical to those in Experiment 1. To avoid behavioral responses interfering with EEG data collection, the procedure differed slightly from Experiment 1. In Experiment 2, the emotional face

picture presented after the music did not disappear upon keypress but remained for a fixed duration of 1000 ms. Event-related EEG markers were time-locked to the onset of the face presentation. After the face disappeared, a response screen appeared, and participants judged whether the emotions expressed by the music and face were congruent or incongruent, pressing “F” for congruent and “J” for incongruent. The mapping of congruence (congruent/incongruent) to response keys (F/J) was counterbalanced across participants. Response time was unlimited, and participants pressed the spacebar to begin the next trial after responding. Six practice trials were provided before the formal experiment. As in Experiment 1, to exclude potential familiarity effects, participants were asked to report the titles of the musical excerpts after the formal experiment. No participant could report any keywords from the titles.

EEG Recording and Data Analysis EEG signals were recorded using a Biosemi 64-channel Active Two electrode cap at a sampling rate of 2048 Hz. External electrodes were placed at the left and right outer canthi and above and below the left eye to record horizontal and vertical electrooculograms. Impedance was kept below 20 k Ω during recording. Offline analysis involved re-referencing to the average of bilateral mastoids, applying a 0.1-30 Hz bandpass filter (24 dB/oct), and automatic ocular artifact correction using BESA software. Epochs were segmented from 200 ms before to 1000 ms after target onset, with a 200 ms pre-stimulus baseline. Trials with amplitude changes exceeding \pm \$120 V or incorrect responses were excluded.

Based on hemisphere and region, nine regions of interest (ROIs) were selected for analysis (left anterior: FP1, AF3, F3, F5, F7; midline anterior: FPz, AFz, Fz; right anterior: FP2, AF4, F4, F6, F8; left central: C1, CP1, FC3, C3, CP3; midline central: FCz, Cz, CPz; right central: C2, CP2, FC4, C4, CP4; left posterior: P3, P5, PO3, PO7, O1; midline posterior: Pz, POz, Oz; right posterior: P4, P6, PO4, PO8, O2). Repeated-measures ANOVAs were conducted separately for midline and lateral sites.

For midline electrodes, congruence (congruent, incongruent), lyrics (lyrical music, non-lyrical music), and region (anterior, central, posterior) were within-subject factors. Lateral analyses included hemisphere (left, right) as an additional within-subject factor. The mean amplitude across all electrodes within each ROI was calculated for further analysis. Only statistically significant or marginally significant results for the main experimental variables (lyrics, congruence) are reported. Simple effects analyses were conducted following significant interactions, with pairwise comparisons corrected using Bonferroni correction. The Greenhouse-Geisser correction was applied when sphericity assumptions were violated.

Results

Behavioral Results A 2 (lyrics: lyrical music, non-lyrical music) \times 2 (congruence: congruent, incongruent) repeated-measures ANOVA on accuracy re-

vealed no significant effects (all $ps > 0.05$). Accuracy was 85.83% in the lyrical condition and 83.02% in the non-lyrical condition, indicating that participants performed the task attentively in both conditions.

ERP Results Figure 3 [Figure 3: see original paper] presents ERP waveforms for non-lyrical (a) and lyrical (b) music priming conditions. Figure 4 [Figure 4: see original paper] shows difference wave topographies (incongruent minus congruent) for non-lyrical and lyrical conditions in the 250–450 ms (a) and 500–700 ms (b) time windows. Based on visual inspection and previous research on N400 (Daltrozzo & Schön, 2009; Kamiyama et al., 2013) and LPC (Herring et al., 2011; Werheid et al., 2005; Zhang et al., 2010) in affective priming paradigms, we defined 250–450 ms and 500–700 ms post-target onset as the N400 and LPC time windows, respectively.

A repeated-measures ANOVA on mean N400 amplitude with congruence (congruent, incongruent), lyrics (lyrical music, non-lyrical music), and region (anterior, central, posterior) as factors revealed a significant lyrics \times congruence interaction, $F(1, 15) = 8.48$, $p = 0.01$, $p^2 = 0.36$. Simple effects analysis showed that for non-lyrical music, incongruent trials elicited larger N400 amplitudes than congruent trials, $F(1, 15) = 5.17$, $p = 0.04$, $p^2 = 0.26$. However, for lyrical music, no significant difference in N400 amplitude was found between congruent and incongruent conditions, $F(1, 15) = 2.90$, $p = 0.11$. The lateral analysis also revealed a significant lyrics \times congruence interaction, $F(1, 15) = 7.80$, $p = 0.02$, $p^2 = 0.34$. Simple effects analysis indicated that non-lyrical music priming produced larger N400 amplitudes for incongruent than congruent trials, $F(1, 15) = 6.81$, $p = 0.02$, $p^2 = 0.31$, whereas lyrical music showed no such difference, $F(1, 15) = 2.18$, $p = 0.16$. No other effects related to lyrics or congruence were significant (all $ps > 0.35$).

Similarly, a repeated-measures ANOVA on mean LPC amplitude revealed a significant lyrics \times congruence interaction in the midline analysis, $F(1, 15) = 7.47$, $p = 0.02$, $p^2 = 0.33$. Simple effects analysis showed that for lyrical music, incongruent trials elicited larger LPC amplitudes than congruent trials, $F(1, 15) = 6.90$, $p = 0.02$, $p^2 = 0.32$. In contrast, for non-lyrical music, no significant difference in LPC amplitude was found between incongruent and congruent trials, $F(1, 15) = 0.78$, $p = 0.39$. The lateral analysis also revealed a significant lyrics \times congruence interaction, $F(1, 15) = 6.20$, $p = 0.03$, $p^2 = 0.29$. Simple effects analysis indicated that lyrical music priming produced larger LPC amplitudes for incongruent than congruent trials, $F(1, 15) = 5.17$, $p = 0.04$, $p^2 = 0.26$, whereas non-lyrical music showed no such difference, $F(1, 15) = 0.82$, $p = 0.38$. No other effects related to lyrics or congruence were significant (all $ps > 0.21$).

The ERP results of Experiment 2 showed that lyrical music priming elicited an N400 effect in the 250–450 ms time window, while non-lyrical music priming elicited an LPC effect in the 500–700 ms time window. These findings indicate that regardless of lyrical content, listeners can judge the emotional relationship between prime and target stimuli. However, the temporal course of processing

emotions from lyrical and non-lyrical music differs, suggesting that emotional processing of lyrical music lags behind that of non-lyrical music.

General Discussion

The present study investigated the neural mechanisms underlying the influence of lyrics on musical emotion processing through two experiments. The behavioral results showed that, regardless of lyrical content, participants responded faster and more accurately under congruent than incongruent conditions, indicating that listeners can process emotional information conveyed by both lyrical and non-lyrical music. The ERP results further revealed that although both conditions produced priming effects, non-lyrical music elicited an N400 effect in the 250–450 ms time window, whereas lyrical music elicited an LPC effect in the 500–700 ms time window. These findings demonstrate that lyrics affect the temporal course of neural processing of musical emotion, suggesting that emotional processing of lyrical music lags behind that of non-lyrical music.

The classic N400 is typically associated with semantic processing, with semantically incongruent words eliciting larger N400 responses than congruent words (Kutas & Hillyard, 1980), reflecting conceptual integration processes (Brown & Hagoort, 1993; Kutas & Federmeier, 2000, 2011). Recent studies have also shown that in affective priming paradigms, affectively incongruent conditions elicit larger N400 responses than congruent conditions (Schirmer et al., 2002; Zhang et al., 2006, 2010), indicating that integrating emotional information requires more cognitive resources (Kamiyama et al., 2013; Zhang et al., 2010). Research has also demonstrated that when prime stimuli are short musical excerpts (Daltrozzo & Schön, 2009; Goerlich et al., 2011; Koelsch et al., 2004) or chords (Steinbeis & Koelsch, 2011), the brain can elicit larger N400 responses to incongruent emotional information. The present study found that processing emotions from non-lyrical music elicited an N400 effect, which likely resulted from the activation of emotional meaning (Daltrozzo & Schön, 2009; Eder, Leuthold, Rothermund, & Schweinberger, 2011). Indeed, the presentation of prime stimuli pre-activated emotional representations related to the target at the conceptual level, reducing N400 amplitudes in congruent conditions (Goerlich et al., 2012). In incongruent conditions, the emotional representations related to the target were not pre-activated, requiring more cognitive resources to integrate emotional information from non-lyrical music and facial expressions, which manifested as increased N400 amplitudes.

Consistent with previous affective priming studies (Herring et al., 2011; Wang & Qin, 2016; Werheid et al., 2005; Zhang et al., 2010, 2012), the present results show that processing emotions from lyrical music elicited an LPC effect. Researchers have suggested that larger LPC responses in incongruent conditions during affective priming reflect greater attentional engagement (Herring et al., 2011; Werheid et al., 2005; Zhang et al., 2010). Therefore, similar to previous research, the LPC in the present study may reflect attentional resource allocation. Additionally, previous studies have found that the LPC also reflects integra-

tion processes (Juottonen, Revonsuo, & Lang, 1996). In memory research on language and music, when linguistic and melodic information are presented simultaneously, listeners integrate them as a whole rather than processing them separately (Serafine, Davidson, Crowder, & Repp, 1986). In the present study, when music contained lyrics, listeners needed to integrate melody and lyrics and possibly match this integrated representation with the emotional face, thereby eliciting an LPC effect. However, it should be noted that although emotional processing of lyrical music lagged behind non-lyrical music (as evidenced by the latency differences between LPC and N400 effects), the presence of priming effects indicates that listeners could correctly understand and process emotional information from both types of music.

As two distinct ERP components, N400 and LPC are generally considered to represent different cognitive processes (Ibáñez et al., 2010; Juottonen et al., 1996; Rohaut & Naccache, 2017). Furthermore, previous research suggests that processing emotional information from lyrical and non-lyrical music may involve different neural mechanisms (Brattico et al., 2011). Therefore, although we propose that both N400 and LPC reflect integration processes to some extent, the integration processes they represent should differ. Based on the origin of emotions, the emotion duality model posits that two evaluation systems operate during emotional responses to stimuli: the automatic evaluating system and the reflective evaluating system (Jarymowicz & Imbir, 2015). Reflective responses require language-based processing and cannot occur without linguistic participation (Imbir, Spustek, & Żygierewicz, 2016). The presence or absence of linguistic information is precisely the difference between lyrical and non-lyrical music in the present study. Therefore, the LPC elicited by lyrical music priming may represent a reflective response, whereas the N400 elicited by non-lyrical music priming may primarily reflect an automatic response. In other words, the LPC in the present study may reflect reflective integration processing, while the N400 may reflect relatively automatic integration processing.

Through two experiments, the present study found that the involvement of lyrics caused a lag in emotional processing of music, supporting perspectives from music philosophy. Music philosophers argue that pure music (without lyrics) conveys emotion more rapidly and directly than lyrical music (Yu, 2000; Zhang & Wang, 1992) because language explicitly conveys emotional information through a propositional system (Erickson, 2005; Jankélévitch & Abbate, 2004), whereas pure music lacks such linguistic semantics, allowing it to bypass the “translation” process in the propositional system and thus transmit emotional information more rapidly and directly.

Regarding the relationship between language and musical emotion processing, consistent with previous behavioral findings (Ali & Peynircioglu, 2006; Mori & Iwanaga, 2013; Stratton & Zalanowski, 1994), the present results demonstrate that lyrics influence musical emotion processing. It should be noted that the present study investigated the effect of lyrics by comparing emotional processing of music with and without lyrics. Although the experimental design cannot

directly address whether language and music interact in emotional processing, the results suggest that they may share certain mechanisms. Specifically, unlike the non-lyrical condition, the lyrical condition elicited an LPC effect, which primarily resulted from the involvement of lyrics. In the lyrical condition, lyrics and melody were presented simultaneously. Since lyrics influenced emotional processing of the music, this implies that language and music may interact in emotional processing. Future research should further investigate this possibility.

In conclusion, the present study demonstrates that listeners can process emotional information conveyed by music regardless of lyrical content. However, non-lyrical music elicited an N400 effect in the 250–450 ms time window, whereas lyrical music elicited an LPC effect in the 500–700 ms time window. These findings indicate that lyrics affect the temporal course of neural processing of musical emotion, with emotional processing of lyrical music lagging behind that of non-lyrical music. The present findings provide empirical evidence for understanding the relationship between music and language.

References

- Ali, S. O., & Peynircioglu, Z. F. (2006). Songs and emotions: Are lyrics and melodies equal partners? *Psychology of Music*, 34(4), 511–534.
- Behrens, G. A., & Green, S. B. (1993). The ability to identify emotional content of solo improvisations performed vocally and on three different instruments. *Psychology of Music*, 21(1), 20–33.
- Brattico, E., Alluri, V., Bogert, B., Jacobsen, T., Vartiainen, N., Nieminen, S., & Tervaniemi, M. (2011). A functional MRI study of happy and sad emotions in music with and without lyrics. *Frontiers in Psychology*, 2, 308.
- Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience*, 5(1), 34–44.
- Daltrozzo, J., & Schön, D. (2009). Conceptual processing in music as revealed by N400 effects on words and musical targets. *Journal of Cognitive Neuroscience*, 21(10), 1882–1892.
- Darwin, C. (1871). *The descent of man and selection in relation to sex*. London, UK: John Murray.
- Erickson, D. A. (2005). *Language, ineffability and paradox in music philosophy* (Unpublished doctoral dissertation). Simon Fraser University.
- Franco, F., Chew, M., & Swaine, J. S. (2017). Preschoolers' attribution of affect to music: A comparison between vocal and instrumental performance. *Psychology of Music*, 45(1), 131–149.
- Galizio, M., & Hendrick, C. (1972). Effect of musical accompaniment on attitude: The guitar as a prop for persuasion. *Journal of Applied Social Psychology*, 2(4), 350–359.

- Goerlich, K. S., Witteman, J., Aleman, A., & Martens, S. (2011). Hearing feelings: Affective categorization of music and speech in alexithymia, an ERP study. *PloS One*, 6(5), e19501.
- Goerlich, K. S., Witteman, J., Schiller, N. O., Heuven, V. J. V., Aleman, A., & Martens, S. (2012). The nature of affective priming in music and speech. *Journal of Cognitive Neuroscience*, 24(8), 1725-1741.
- Gong, X., Huang, Y. X., Wang, Y., & Luo, Y. J. (2011). Revision of the Chinese facial affective picture system. *Chinese Mental Health Journal*, 25(1), 40-46. [龚栩, 黄宇霞, 王妍, 罗跃嘉. (2011). 中国面孔表情图片系统的修订. 中国心理卫生杂志, 25(1), 40-46.]
- Hailstone, J. C., Omar, R., Henley, S. M., Frost, C., Kenward, M. G., & Warren, J. D. (2009). It's not what you play, it's how you play it: Timbre affects perception of emotion in music. *Quarterly Journal of Experimental Psychology*, 62(11), 2141-2155.
- Herring, D. R., Taylor, J. H., White, K. R., & Crites, S. L. (2011). Electrophysiological responses to evaluative priming: The LPP is sensitive to incongruity. *Emotion*, 11(4), 794-806.
- Hu, X., Downie, J. S., & Ehmann, A. F. (2009, October). Lyric text mining in music mood classification. Paper presented at the International Society for Music Information Retrieval Conference, Kobe, Japan.
- Ibáñez, A., Manes, F., Escobar, J., Trujillo, N., Andreucci, P., & Hurtado, E. (2010). Gesture influences the processing of figurative language in non-native speakers: ERP evidence. *Neuroscience Letters*, 471(1), 48-52.
- Imbir, K. K., Spustek, T., & Żygierewicz, J. (2016). Effects of valence and origin of emotions in word processing evidenced by event related potential correlates in a lexical decision task. *Frontiers in Psychology*, 7, 271.
- Jackendoff, R. (2009). Parallels and nonparallels between language and music. *Music Perception*, 26(3), 195-204.
- Jankélévitch, V., & Abbate, C. (2004). *Music and the ineffable*. Princeton, NJ: Princeton University Press.
- Jarymowicz, M. T., & Imbir, K. K. (2015). Toward a human emotions taxonomy (based on their automatic vs. reflective origin). *Emotion Review*, 7(2), 183-188.
- Juottonen, K., Revonsuo, A., & Lang, H. (1996). Dissimilar age influences on two ERP waveforms (LPC and N400) reflecting semantic context effect. *Cognitive Brain Research*, 4(2), 99-107.
- Kamiyama, K. S., Abla, D., Iwanaga, K., & Okanoya, K. (2013). Interaction between musical emotion and facial expression as measured by event-related potentials. *Neuropsychologia*, 51(3), 500-505.

- Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: Brain signatures of semantic processing. *Nature Neuroscience*, 7(3), 302-307.
- Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, 4(12), 463-470.
- 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.
- Kutas, M., & Hillyard, S. A. (1980). Reading between the lines: Event-related brain potentials during natural sentence processing. *Brain and Language*, 11(2), 354-373.
- Laurier, C., Lartillot, O., Eerola, T., & Toiviainen, P. (2008, May). Multimodal music mood classification using audio and lyrics. Paper presented at the Seventh International Conference on Machine Learning and Applications, Copenhagen, Denmark.
- Levman, B. G. (1992). The genesis of music and language. *Ethnomusicology*, 36(2), 147-170.
- Mithen, S. (2006). *The singing Neanderthals: The origins of music, language, mind and body*. London: Weidenfeld & Nicolson.
- Mori, K., & Iwanaga, M. (2013). Pleasure generated by sadness: Effect of sad lyrics on the emotions induced by happy music. *Psychology of Music*, 42(5), 643-652.
- Morton, B. J., & Trehub, S. E. (2007). Children's judgements of emotion in song. *Psychology of Music*, 35(4), 629-639.
- Perlovsky, L. (2011). Musical emotions: Functions, origins, evolution. *Physics of Life Reviews*, 7(1), 2-27.
- Rohaut, B., & Naccache, L. (2017). Disentangling conscious from unconscious cognitive processing with event-related EEG potentials. *Revue Neurologique*, 173(7-8), 521-528.
- Serafine, M. L., Davidson, J., Crowder, R. G., & Repp, B. H. (1986). On the nature of melody-text integration in memory for songs. *Journal of Memory and Language*, 25(2), 123-135.
- Schirmer, A., Kotz, S. A., & Friederici, A. D. (2002). Sex differentiates the role of emotional prosody during word processing. *Cognitive Brain Research*, 14(2), 228-233.
- Sousou, S. D. (1997). Effects of melody and lyrics on mood and memory. *Perceptual & Motor Skills*, 85(1), 47-56.

Steinbeis, N., & Koelsch, S. (2011). Affective priming effects of musical sounds on the processing of word meaning. *Journal of Cognitive Neuroscience*, 23(3), 604-621.

Stratton, V. N., & Zalanowski, A. H. (1994). Affective impact of music vs. lyrics. *Empirical Studies of the Arts*, 12(2), 173-184.

Thompson, W. F., Marin, M. M., & Stewart, L. (2012). Reduced sensitivity to emotional prosody in congenital amusia rekindles the musical protolanguage hypothesis. *Proceedings of the National Academy of Sciences*, 109(46), 19027-19032.

Wang, Y., & Qin, Z. (2016). Affective priming by simple geometric shapes: Evidence from event-related brain potentials. *Frontiers in Psychology*, 7, 917.

Werheid, K., Alpay, G., Jentzsch, I., & Sommer, W. (2005). Priming emotional facial expressions as evidenced by event-related brain potentials. *International Journal of Psychophysiology*, 55(2), 209-219.

Yu, R. Y. (2000). *An introduction to modern western music philosophy*. Changsha, China: Hunan Education Press. [于润洋. (2000). 现代西方音乐哲学导论. 长沙: 湖南教育出版社.]

Zhang, Q., & Wang, C. Z. (1992). *Foundation in musical aesthetics*. Beijing, China: People's Music Publishing House. [张前, 王次炤. (1992). 音乐美学基础. 北京: 人民音乐出版社.]

Zhang, Q., Kong, L., & Jiang, Y. (2012). The interaction of arousal and valence in affective priming: Behavioral and electrophysiological evidence. *Brain Research*, 1474, 60-72.

Zhang, Q., Lawson, A., Guo, C., & Jiang, Y. (2006). Electrophysiological correlates of visual affective priming. *Brain Research Bulletin*, 71(3), 316-323.

Zhang, Q., Li, X., Gold, B. T., & Jiang, Y. (2010). Neural correlates of cross-domain affective priming. *Brain Research*, 1329, 142-151.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv – Machine translation. Verify with original.