Neuropsychologia 51 (2013) 500–505

Contents lists available at SciVerse ScienceDirect

Neuropsychologia

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

Interaction between musical emotion and facial expression as measured by event-related potentials

Keiko S. Kamiyama [a,b,c], Dilshat Abla [d], Koichi Iwanaga [e], Kazuo Okanoya [a,c,f,*]

[a] Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1Komaba, Meguro-Ku, Tokyo 153-8902, Japan
[b] Research Fellow of the Japan Society for the Promotion of Science, Tokyo, Japan
[c] Emotional Information Joint Research Laboratory, 2-1 Hirosawa, Wako-Shi, Saitama 351-0198, Japan
[d] Noninvasive BMI Unit, BSI-Toyota Collaboration Center, RIKEN Brain Science Institute, 2-1 Hirosawa, Wako-Shi, Saitama 351-0198, Japan
[e] Department of Design, Graduate School of Engineering, Chiba University, Chiba, Japan
[f] Japan Science Technology Agency, ERATO, Okanoya Emotional Information Project, 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan

ARTICLE INFO

Article history:
Received 27 April 2012
Received in revised form
30 October 2012
Accepted 27 November 2012
Available online 5 December 2012

Keywords:
Music
Facial expression
Emotion
Electroencephalogram (EEG)
N400

ABSTRACT

We examined the integrative process between emotional facial expressions and musical excerpts by using an affective priming paradigm. Happy or sad musical stimuli were presented after happy or sad facial images during electroencephalography (EEG) recordings. We asked participants to judge the affective congruency of the presented face–music pairs. The congruency of emotionally congruent pairs was judged more rapidly than that of incongruent pairs. In addition, the EEG data showed that incongruent musical targets elicited a larger N400 component than congruent pairs. Furthermore, these effects occurred in nonmusicians as well as musicians. In sum, emotional integrative processing of face–music pairs was facilitated in congruent music targets and inhibited in incongruent music targets; this process was not significantly modulated by individual musical experience. This is the first study on musical stimuli primed by facial expressions to demonstrate that the N400 component reflects the affective priming effect.

1. Introduction

Musical performers might be able to make good use of their facial expressions to communicate their expressive intentions. Some performers might consciously put this possibility into practice. On the other hand, some performers occasionally play joyful music with knitted brows. We cannot easily conclude that this emotional conflict has a meaningful effect on listeners, since music might be processed independently of nonmusical factors such as facial expressions. When one is having a conversation, it is natural to watch the face of the person to whom one is speaking; also, it is rational to utilize visual information supplied by the face as well as auditory information supplied by the voice in order to understand the emotional states of others. In fact, de Gelder and Vroomen (2000) demonstrated that identification of the emotion expressed by tone of voice was subconsciously biased in the direction of a simultaneously presented face, and vice versa. Does musical processing automatically refer to the facial information, as does speech processing?

Musical meaning can be summarized in terms of three fundamental classes: extramusical, intramusical, and musicogenic meanings (Koelsch, 2011). The emotional states expressed by musical pieces are classified as indexical extramusical meaning. In this study, we addressed the meaning that emerges from the interpretation of indexical musical sign quality, in specific perception of happiness and sadness expressed by music. Several properties of music affect its emotional expression: tempo, mode, loudness, pitch, interval, melody, rhythm, timbre, articulation, etc. (Gabrielsson & Lindström, 2001). For instance, major keys and fast tempi have been associated with happiness, and minor keys and slow tempi with sadness (Gagnon & Peretz, 2003). In almost all musical pieces, the way in which these elements are combined influences the emotional meaning conveyed to the listeners. Furthermore, Gagnon and Peretz (2003) showed that these properties do not contribute equally to musical emotion by demonstrating that tempo was a more important factor than mode when nonmusicians judged happy–sad emotion in music. Thus, people show similarities in how they recognize musical emotion, at least with regard to the basic emotions expressed by music (for review, see Trainor & Corrigall, 2010). According to Leaver and Halpern (2004), nonmusicians could not accurately

* Corresponding author at: The University of Tokyo, Department of Life Sciences, Graduate School of Arts and Sciences, 3-8-1 Komaba, Meguro Ku, Tokyo 153-8902, Japan. Tel.: +81 3 5454 6266; fax: +81 3 5454 6725.
E-mail addresses: kkamiyama@brain.riken.jp (K.S. Kamiyama), dilshata@gmail.com (D. Abla), iwanaga@faculty.chiba-u.jp (K. Iwanaga), kazuookanoya@gmail.com (K. Okanoya).

0028-3932/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.neuropsychologia.2012.11.031

classify melodies according to major/minor labels, but they could discriminate the same melodies in terms of affective labels. However, they reported that minor chords were evaluated as more sad in musicians than in nonmusicians, although this difference between groups was not reflected in blood oxygenation level-dependent (BOLD) contrasts in the brain (Pallesen et al., 2005). These findings indicate that musical training may influence the acoustic or emotional processing of music, although even nonmusicians are able to recognize basic emotions in music.

The cross-modal integration process between music and visual stimuli, especially the influence of music on visual recognition, has been documented in recent years. Baumgartner, Esslen, and Jäncke (2006) showed that music enhanced the emotional experience evoked by affective pictures through the use of EEG α power density and other physiological indicators. Baumgartner, Lutz, Schmidt, and Jäncke (2006) also presented fearful and sad pictures either alone or combined with congruent emotional musical excerpts, and observed clear functional and structural dissociation between the combined condition and the picture-only condition. Furthermore, Logeswaran and Bhattacharya (2009) proved that images of facial expressions induced varying ERP amplitudes depending on the relationship between facial stimuli (happy, sad, or emotionally neutral facial images) and the preceding musical stimuli (happy or sad music). Tan, Spackman, and Bezdek (2007) showed that the viewers tended to interpret film characters' emotions in ways that were consistent with the emotional contents of music presented before or after the characters in the film. Jolj and Meurs (2011) also reported that happy music made participants believe that they saw a happy face on a screen on which random noise was displayed even when there was actually no facial image present. These findings indicate that musical perception is integrated into the emotional processing and judgment of visual stimuli such as pictures of objects, scenery, and human faces. In the literature on audiovisual integration using music as the auditory stimuli, researchers have often been interested in how the visual stimuli are processed. In contrast, the bidirectional effect between voice and face recognition has been confirmed through behavioral and neurophysiological studies (de Gelder, Böcker, Tuomainen, Hensen, & Vroomen, 1999; de Gelder & Vroomen, 2000; Grossmann, Striano, & Friederici, 2006; Pourtois, de Gelder, Vroomen, Rossion, & Crommelinck, 2000; Pourtois, Debatisse, Despland, & de Gelder, 2002; Spreckelmeyer, Kutas, Urbach, Altenmüller, & Münte, 2006). The emotional compatibility between face and music, however, was investigated practically in a previous study, which proved universal recognition of the basic emotions expressed by music. In the study by Fritz et al. (2009), participants evaluated the emotion conveyed by music by matching musical excerpts with the emotions in facial images expressing happiness, sadness, and fear. Nevertheless, the constantly changing nature of music makes it difficult to measure the brain activities, which are responsive to the emotional processing induced by musical events. To solve this issue, some research projects have utilized very short musical excerpts and measured the electrical brain responses thereto using EEG (Daltrozzo & Schön, 2008; Goerlich, Witteman, Aleman, & Martens, 2011). In behavioral studies, the ability for very short musical excerpts to express emotion has been discussed from the viewpoints of musicogenic and extra-musical meanings. Bigand, Vieillard, Madurell, Marozeau, and Dacquet (2005) demonstrated that musical excerpts lasting only 1 s induced consistent emotional responses, regardless of participants' musical experience levels. Krumhansl (2010) found that judgments of the emotional contents of sing clips lasting 300 ms, 400 ms, and 15 s were highly consistent. Also, Peretz, Gagnon, and Bouchard (1998) showed that nonmusicians could distinguish the emotional character (happiness vs. sadness) of presented music even when its duration was 250 ms. Thus, the use of short musical excerpts as emotional auditory stimuli has been validated.

In this study, we used a priming paradigm to reveal the effects of facial expressions on emotional processing of music. The priming paradigm was initially applied in a linguistic domain, and then it became commonly used in non-linguistic contexts. Kutas and Hillyard (1980) investigated the ERPs elicited by words at the end of sentences, and they found that contextually inappropriate words induced a large N400, which occurs 400 ms after the onset of the target words. N400 was distributed over the scalp, and was strongest in the central parietal region. It has also been reported that N400 was sensitive to the Cloze probability and that its amplitude was an inverse function of how strongly participants expected the terminal word (Kutas & Hillyard, 1984). Furthermore, Bentin, McCarthy, and Wood (1985) demonstrated that during a lexical decision task, target words primed by related words were identified more quickly and accurately than were unprimed words. In addition, the primed words elicited a more negative EEG deflection around 400 ms after the onset of word stimuli compared with that elicited by unprimed words. Apart from these semantic priming effects, affective priming effects also occur depending on the affective relatedness of word pairs. The behavioral and electrophysiological data obtained by semantic and affective priming paradigms are quite compatible, but these effects can be clearly distinguished. Blair et al. (2006) compared individuals with psychopathy and comparison individuals in terms of their behavioral performance in semantic and affective priming paradigms; less affective priming was observed in the psychopathy group relative to the comparison group. In contrast, the two groups showed comparable levels of semantic priming.

In addition to linguistic stimuli, it has been reported that nonlinguistic stimuli such as pictures, odors, facial expressions, and musical elements or excerpts elicit priming effects. In studies on musical meaning, which applied priming paradigms, a negative ERP component (N400) reflected the processing of extramusical meaning (Koelsch, 2011). As in the linguistic domain, music can contribute to both the semantic and affective priming effects. The following semantic priming effects between musical and linguistic stimuli have been reported: in a previous study in which music excerpts and sentences were used as primers and words as targets, it was found that music and sentences could possess semantic relationships with particular words such as "wideness" and "blossom" (Koelsch et al., 2004). Likewise, Daltrozzo and Schön (2008) found that semantically unrelated word–music pairs enhanced the amplitude of N400 when using 1-s music excerpts and words such as "magic" and "courage." The affective priming effects of words, musical elements, and musical excerpts have been indexed in terms of the N400 effect in EEG recordings and in terms of response latency and accuracy in behavioral tests. Sollberger, Reber, and Eckstein (2003) showed that the consonance vs. dissonance of musical chords influenced the affective evaluation of target words. Their participants evaluated target words more quickly and accurately when words were preceded by similarly valenced chords. In addition, Steinbeis and Koelsch (2008) demonstrated that targets consisting of both music and language, which were incongruous to the preceding affective context, enhanced the amplitude of the N400 component. The finding that low-level musical features convey meaningful representations is also supported by the results of Painter and Koelsch (2011). They manipulated the timber of tones and demonstrated that the N400 effects were found for sound and word targets following sound and word primes. Musical segments created from real musical pieces have also been used as auditory stimuli in affective priming paradigms. Goerlich et al. (2011) created musical excerpts and speech stimuli of length 600 ms and