Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music, July 23-28, 2012, Thessaloniki, Greece
Cambouropoulos E., Tsougras C., Mavromatis P., Pastiadis K. (Editors)

Multisensory Perception of Six Basic Emotions in Music

Ken-ichi Tabei,[*1] Akihiro Tanaka[#2]

[*]*Department of Dementia Prevention and Therapeutics, Graduate School of Medicine, Mie University, Japan*
[#]*Department of Psychology, Tokyo Woman's Christian University, Japan*
[1]tabei@clin.medic.mie-u.ac.jp, [2]akihiro@lab.twcu.ac.jp

ABSTRACT

The interaction between auditory and visual information is known to influence emotion judgments by using audiovisual speech stimuli (i.e., face–voice combination). In contrast, little is known about how emotion perception changes when the musician's facial and bodily movements can be seen as well as heard. In the present study, we applied a paradigm often used in face–voice emotion perception to music performance to examine the interaction between musical sound and facial and bodily movements in perceiving emotion from music performance. Results showed that the performances in the Audio (A), Visual (V), and Audio-Visual (AV) conditions were dependent on the combination of instruments and emotions: angry expression by cellists and sad expression by violinist were perceived better in the V condition, while disgust expression by pianist were perceived better in the AV condition. While previous studies have shown that visual information from facial expression facilitates the emotion perception from emotional prosody in speech, that of musician's facial and bodily movements did not necessarily enhance the emotion perception from musical sound. This pattern suggests that multisensory perception of emotion from music performance may be different from that from audiovisual speech.

Background

As is the case with audiovisual speech (i.e., facial expression combined with an affective voice fragment), we integrate auditory and visual information of music performance in many contemporary musical idioms such as opera, musical theatre, and rock concerts.

However, fewer studies have investigated the audiovisual emotion perception of body–sound combination by means of music performance than that of face-voice combination. Thus, little is known about how emotion perception changes when the musician's facial and bodily movements can be seen as well as heard.

Visual information is important not only in the emotion perception of the audiovisual speech but also in music performance. Davidson (1993) investigated the emotional information conveyed by the movements of the piano, the violin and the cello performers when s/he was asked to play the same piece in three different expressive manners (deadpan, projected and exaggerated). The results revealed that participants could distinguish performer's expression only by the movement of the marker given to the head, elbows, wrists, knees, and ankles and on each hip and shoulder. Thompson et al. (2008) examined whether facial expressions of performers influence the emotional information of sung materials. Participants judged the emotional valence of audio-visual presentations of songs that consisted of major and minor third intervals. The results showed that participants' judgments were influenced by facial expressions.

Auditory information plays a dominant role in the emotion perception of music performance. Petrini et al. (2010) examined the nature of the audio-visual integration of the emotion by using the drum and the saxophonist's ad-lib performance. From the original bimodal congruent displays, the audio-only, visual-only, and audiovisual incongruent conditions (obtained by combining the two signals both within and between instruments) were derived. Participants judged the perceived emotion (happiness, sadness, anger, fear, disgust, surprise and neutral) and rated the strength of each emotion. The results indicated that auditory information dominated the visual information in the perception of emotion expression. Further, the participants were asked to either pay attention to the musicians' movements or to the sound when judging the perceived emotions. The result showed no effect of visual information when judging the sound. This auditory dominance is different from the previous research that used audiovisual speech and indicated that visual information dominated the auditory information in the perception of emotion expression (e.g., Collignon et al., 2008).

Aims

While the research using face and voice suggested that visual information dominates the auditory information in the perception of emotional expression (e.g., Collignon et al., 2008), the research that used musical sound and bodily movements of music performance suggested the priority of the auditory information (e.g., Vines et al. 2006; Petrini et al., 2010). Most of previous studies hardly examined the relative contribution of visual and auditory information of music performance. The present study thus attempts to assess the multisensory perception of emotional expressions of music using violin, cello, piano and vocal performances. Participants were presented with audiovisual congruent, audio-only and visual-only stimuli, and judged whether the excerpts expressed happiness, sadness, anger, fear, disgust, or surprise.

Method

Participants

Seventeen participants (three males, mean age = 20.4, *SD* = 1.06) were recruited. The participants had played an instrument for 3 to 19 years. All participants signed written informed consent.

Stimuli

The audiovisual stimuli for this experiment were created from video recordings of music performance. Eight university students majoring in piano, vocal, cello or violin played the brief excerpts (Gabrielsson, 1996) by expressing six basic emotions (happy, sad, angry, fear, disgust and surprise: Ekman and Friesen, 1975). The brief excerpts were in a major or minor key. They were free to vary whatever variables they

wanted - such as tempo, timing, dynamics, articulation, phrasing, vibrato, attack, and timbre.

Procedure

Participants were presented with a video recording of a performance and were asked to judge the perceived emotion from a list of six categories (happy, sad, angry, fear, disgust and surprise). In A condition, participants were instructed to judge the emotion of the musical sounds. In V condition, participants were instructed to judge the emotion of the facial and bodily movements. In AV condition, participants were instructed to judge the emotion of the musical sound and the facial and bodily movements. The order of the conditions was counterbalanced across participants. Participants were engaged in 288 trials in total (2 excerpts × 8 performers × 6 emotions × 3 conditions).

Results

Interaction of modality, instrument and emotion

A three-way within-subjects analysis of variance with the factors of presentation modality (A, V, AV), instrument (piano, vocal, cello and violin) and emotion (happy, sad, angry, fear, disgust and surprise) revealed a significant three-way interaction, ($F(30,480) = 2.525$, $p < .001$). The result showed that performances in A, V and AV conditions were dependent on the combination of instruments and emotions. Specifically, accuracy in angry expression by cellists was higher in V condition than in other conditions ($p < .05$); Accuracy in disgust expression by pianist was higher in AV condition than in A condition ($p < .05$); Accuracy in happy expression by violist was higher in the V condition than in A condition; sad expression by violist was higher in V condition than in other conditions ($p < .05$); Accuracy in happy expression by vocalist was higher in AV and V conditions than in A condition ($p < .05$).

Relative priority of visual or auditory information

In order to compare an influence from visual information toward auditory information and vice versa, we examined the relative priority of the visual and auditory information of the music performance. One-Sample t-tests for comparing the influence from visual information toward auditory information (AV - A) and vice versa (AV - V) were carried out by using accuracy rate summarizing the emotion. The result of vocal stimuli showed that the influence from the visual information to the auditory information (AV - A = 9%) was more than vice versa (AV - V = -1%, $t(16) = 3.04$, $p < .01$). A significant difference was not found in the result summarizing instruments ($t(16) = 1.46$, $p > .05$).

Facilitates the emotion perception of AV condition

While previous studies have shown that visual information from facial expression facilitates the emotion perception from emotional prosody in speech, that from musician's facial and bodily movements did not necessarily enhance the emotion perception from musical sound in this experiment. There are several reasons why music performance hardly enhanced emotion in AV condition. One of them, we consider that facial and bodily movements and musical sound may not express the same emotion. The stimuli that participants judged different emotions between V and A condition were 32 percent of the total. This percentage was the higher than audiovisual speech (9.7%; Tanaka, 2011). The result may indicate that visual and auditory information of musical performance did not express the same emotion as audiovisual speech did.

Conclusions

The present study attempts to assess the multisensory nature of the emotion perception by using violin, cello, piano and vocal performances. Participants were presented with audiovisual congruent, audio-only and visual-only conditions, and asked to judge whether the excerpts expressed happy, sad, angry, fear, disgust or surprise. The result suggested that the performances in the Audio (A), Visual (V), and Audio-Visual (AV) conditions were dependent on the combination of instruments and emotions. While previous studies have shown that visual information from facial expression facilitates the emotion perception from emotional prosody in speech, that of musician's facial and bodily movements did not necessarily enhance the emotion perception from musical sound. This pattern suggests that multisensory perception of emotion from music performance may be different from that from audiovisual speech.

Keywords

Emotion perception, Basic emotions, Audiovisual integration, and Music performance

Acknowledgments

This work was supported by Strategic Information and Communications R&D Promotion Programme (10210311).

REFERENCES

Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., & Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain research*, 1242, 126–135.

Davidson, J. W. (1993). Visual Perception of Performance Manner in the Movements of Solo Musicians. *Psychology of Music*, 21(2), 103–113.

Ekman, P., & Friesen, W. V. (2003). Unmasking the face. a guide to recognizing emotions from facial clues (p. 212).

Gabrielsson, A. (1996). Emotional Expression in Music Performance: Between the Performer's Intention and the Listener's Experience. *Psychology of Music*, 24, 68-91.

Petrini, K., & McAleer, P. (2010). Audiovisual integration of emotional signals from music improvisation does not depend on temporal correspondence. *Brain research*. 1323:139-148.

Tanaka, A. (2011). Multisensory communication of emotion by faces and voices. *Cognitive Studies*, 18(3), 1-12.

Thompson, W. F., Russo, F. A., & Quinto, L. (2008). Audio-visual integration of emotional cues in song. *COGNITION & EMOTION*, 22(8), 1457–1470.

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, 101(1), 80–113.