International Symposium on Performance Science
ISBN 978-94-90306-02-1
© The Author 2011, Published by the AEC
All rights reserved

Expressive performance in music: Mapping acoustic cues onto facial expressions

William Forde Thompson

Department of Psychology, Macquarie University, Australia

The acoustic attributes conveyed in music are often ambiguous, and people vary in their sensitivity to such attributes. For this reason, expert musicians supplement performances with non-acoustic cues that support communication, including gestures and facial expressions. For musicians, facial expressions are often interpreted as emotional communication, but they reflect many other properties of music. Facial expressions provide information about phonetic information, pitch and interval size, tonality, closure, dissonance, and emotional states. How can continuous changes in facial expressions simultaneously reflect multiple dimensions of the auditory signal? In this article, I will introduce a model of music communication that explains why performers map acoustic information onto facial expressions and how these mappings influence the perceptions and experiences of music listeners.

Keywords: movement; music; emotion; perception; synchronization

Research on music performance typically focuses on the production of sound and resultant acoustic information. In a series of investigations, we have shown that music performers supplement acoustic signals of music with richly informative facial expressions and body movements. These movements not only provide phonetic information (Quinto *et al.* 2010), but they provide signals of emotion, dissonance, pitch structure, and phrasing. They also extend the time-period within which communication occurs: meaningful expressions are observed prior to and after the production of sound (Livingstone *et al.* 2009). Pre-production facial expressions may *prime* forthcoming acoustic signals for listeners, facilitating accurate perception and encoding. Post-production facial expressions may reinforce representations of structural and emotional signals. By extending the temporal window of communication, facial expressions and body movements provide an umbrella that surrounds the acoustic dimension of music, supporting and enriching auditory signals and creating a multimodal experience of music. Because rapidly changing acoustic signals are often ambiguous or difficult to decode, especially for musically untrained listeners, visual signals also function as a safety net for breakdowns in the transmission of acoustic information.

MAIN CONTRIBUTION

Recent research has revealed that the facial expressions and body movements of musicians are remarkably important for music experience. Thompson *et al.* (2008) investigated the significance of facial expressions for communicating emotion in music. Participants were presented with audio-visual presentations of two types of sung intervals: an ascending major third and an ascending minor third. Ascending major third intervals connote a positive emotion whereas minor thirds connote a negative emotion. In the congruent condition, audio-visual recordings of sung intervals were presented to participants in original form. In the incongruent condition, the video showing facial expressions accompanying the major third were re-synchronised with audio of the sung minor third, and vice versa. A group of participants judged the emotional valence of congruent and incongruent intervals. Sung major thirds were judged as more positive emotionally than sung minor thirds, as expected. However, judgments were also influenced by facial expressions. Participants judged both intervals as more positive when accompanied by facial expressions used to produce a major interval than a minor interval. The effect remained when participants were told to ignore visual information or when participants were given a challenging secondary task that involved attending to rapid sequences of numbers. These findings suggest that visual signals are integrated with auditory signals automatically in a way that does not vary with available attentional resources.

Facial expressions of emotion may be particularly informative because they extend beyond the temporal window within which acoustic signals of emotion are available. Livingstone *et al.* (2009) used motion capture or electromyography (EMG) to record the facial movements of singers. Singers were presented with audiovisual recordings of sung phrases performed with happy, sad, or neutral emotional expressions. They then imitated the recordings. Analysis of facial movements revealed reliable signals of emotion that occurred before, during, and after the production of sound. Perceivers of music, in turn, could reliably decode these visual signals of emotion (Thompson *et al.* 2009).

Facial expressions also allow listeners to assess consonance and dissonance in music. Thompson *et al.* (2005) selected twenty excerpts from audio-visually recorded performances of B. B. King playing blues guitar. In ten selections (visual condition 1), King's facial expressions and body movements conveyed musical dissonance: signals included wincing of the eyes, shaking of the upper body, and a rolling of the head. In the remaining ten selections, King's expressions were neutral (visual condition 2). Two groups of participants (n=26) were presented with the 20 excerpts and judged the level of dissonance in each. One group judged dissonance in audio-only presentations and the other judged dissonance in audio-visual presentations. Dissonance was defined as a discordant *sound* that suggested a need for resolution. For audio-only presentations, there was no significant difference between the two visual conditions in mean ratings of dissonance. That is, for the musically untrained participants in this study, the guitar sounds themselves did not predict the dissonant facial expressions made by B. B. King. For audio-visual presentations, there was a large significant difference in dissonance ratings between the two visual conditions: when guitar sounds were coupled with dissonant facial expressions and body movements, they were judged as acoustically dissonant. That is, visual signals arising from facial expressions and body movements guided acoustic judgments.

Facial expressions also carry information about musical structure. Thompson and Russo (2007) found that facial expressions reflect the size of sung melodic intervals. Participants observed silent videos of musicians singing 13 melodic intervals and judged the size of each interval the singer was imagined to be singing. Participants could discriminate intervals based on visual information alone. Facial and head movements were correlated with the size of sung intervals. More recently, Thompson *et al.* (2010) presented participants with silent video recordings of sung melodic intervals spanning 0, 6, 7, or 12 semitones. Again, interval sizes were discriminated based on visual information alone. Even when the auditory signal was made available, facial expressions still affected judgments of interval size, suggesting that visual signals are integrated with auditory information to form an overall sense of interval size. The effects of facial expressions remained when a challenging secondary task was introduced to consume attentional resources. The latter finding suggests that audio-visual *integration* of interval size information occurs independently of attention.

Facial expressions also reflect phrase structure. Ceaser *et al.* (2009) investigated whether musical performers use facial expressions to communicate a sense that a musical phrase has come to an end. Musicians hummed *Silent Night* with two endings. One version ended on the first note of the scale (doh) and conveyed a sense of closure. The other version ended on the fifth note of the scale and conveyed a lack of closure, as though the melodic phrase was unfinished. Fifteen participants were presented with video-only recordings of the hummed sequences and judged whether the (imagined) melody was closed (came to a satisfactory end) or unclosed (seemed unfinished). Accuracy was reliably above chance, indicating that participants were able to read expressions of musical closure from the facial expressions of the musicians.

IMPLICATIONS

What can explain this remarkable capacity of facial and body movement to convey multiple qualities of music, and what are the implications for understanding music cognition? Over the past decade, a body of theory and evidence has emerged concerning the cognitive-motor implications of music. This development suggests a *common-coding* framework for understanding the role of facial expressions and body movements in music perception (Prinz 1990). Specifically, it has been suggested that music has the capacity to engage cognitive-motor processes that function in human *synchronization* (Overy and Molnar-Szakacs 2009). Motor processes involved in synchronization, in turn, may be integrated with the perception of structural and emotional attributes of music. Music affords explicit synchronization in time (clapping, tapping) and pitch (singing along). However, implicit forms of synchronization may also occur in response to musical input (Overy and Molnar-Szakacs 2009). All synchronization involves motor processes, but such processes need not entail explicit or observable movements.

The facial expressions and body movements of performing musicians are explicit manifestations of the motor commands that are activated during the production of musical sounds. The qualities of those movements may reflect the degree of muscular change required in producing a musical event, and the degree of mental effort involved. Events that are unstable and poorly represented in memory require greater effort and motor commands may be less specified. Thus, singing a highly unstable pitch may lead to greater irrelevant muscular activity and apparent effort in the face than singing a highly stable pitch. The timing and duration of motor actions may also reflect the stability of mental representations of music. Action timings may be more precise for stable musical events than for unstable musical events.

Music perceivers readily decode facial movements, linking different movements to different musical events. According to Thompson and Quinto (in press), decoding is also facilitated by implicit synchronization during music listening. For example, an ascending interval may activate motor commands associated with the vocalization of that interval; these commands may then contribute to the recognition and classification of the interval.

The involvement of synchronization in music perception means that all musical events can have an emotional quality. The synchronization-feedback model proposed by Thompson and Quinto (in press) posits two processes that assist with goal-directed behavior. One is a behavior-guiding feedback process that registers errors and acts to correct the error. The second is a feedback loop that monitors discrepancy-reduction over time (i.e. monitoring the first process). The concurrent operation of both feedback systems, one controlling position and the other velocity, leads to rapid and effective synchronization to music.

Feedback from each system is experienced as emotion. Feedback from the behavior-guiding process leads to tension and prediction responses, discussed by Huron (2006). In the tension response, arousal is elicited as a target of synchronization is approached. In the prediction response, positive or negative feedback arises depending on whether synchronization with the target event is correctly aligned. Positive feedback rewards and reinforces alignment; negative feedback motivates increased effort in synchronization.

The second *monitoring* feedback process is maintained by emotional valence. When there is an increase in synchronization accuracy over time, positive feedback results. Otherwise, negative feedback results. Thus, moment-to-moment *arousal* and *reward* generated by the (first) behavior-guiding feedback process are combined with experiences of emotional *valence* generated by the (second) monitoring feedback process.

Together, the two synchronization-feedback processes continuously imbue music with emotional character, though other links between music and emotion have also been identified. Facial expressions and body movement are explicit instances of motor commands that occur not only in performers, but also in listeners. Such movements reflect musical structure, emotion, and a common bond between performers and listeners.

Acknowledgments

This work was supported by an ARC Discovery grant (DP0987182).

Address for correspondence

Bill Thompson, Department of Psychology, Macquarie University, Sydney, New South Wales 2109, Australia; *Email:* bill.thompson@mq.edu.au

References

Ceaser D. K., Thompson W. F., and Russo F. A. (2009). Expressing tonal closure in music performance: Auditory and visual cues. *Canadian Acoustics, 37,* pp. 29-34.

Huron D. (2006). *Sweet Anticipation*. Cambridge, Massachusetts, USA: MIT Press.

Livingstone S. R, Thompson W. F., and Russo F. A. (2009). Facial expressions and emotional singing: A study of perception and production with motion capture and electromyography. *Music Perception, 26*, pp. 475-488.

Overy K. and Molnar-Szakacs I. (2009). Being together in time: Musical experience and the mirror neuron system. *Music Perception, 26*, pp. 489-504.

Prinz W. (1990). A common coding approach to perception and action. In O. Neumann and W. Prinz (eds.), *Relationships between Perception and Action* (pp. 167-201). Berlin: Springer.

Quinto L., Thompson W. F., Russo F. A., and Trehub S. E. (2010). A comparison of the McGurk effect for spoken and sung syllables. *Attention, Perception and Psychophysics, 72*, pp. 1450-1454.

Thompson W. F., Bennetts R., Neskovic B., and Palmer C. (2009). Emotional lingering: Facial expressions of musical closure. In A. Williamon, S. Pretty, and R. Buck (eds.), *Proceedings of ISPS 2009* (pp. 359-364). Utrecht, The Netherlands: European Association of Conservatoires (AEC).

Thompson W. F., Graham P., and Russo F. A. (2005). Seeing music performance: Visual influences on perception and experience. *Semiotica, 156*, pp. 203-227.

Thompson W. F. and Russo F. A. (2007). Facing the music. *Psychological Science, 18*, pp. 756-757.

Thompson W. F., Russo F. A., and Livingstone S. L. (2010). Facial expressions of singers influence perceived pitch relations. *Psychonomic Bulletin and Review, 17,* pp. 317-322.

Thompson W. F., Russo F. A., and Quinto L. (2008). Audio-visual integration of emotional cues in song. *Cognition and Emotion, 22*, pp. 1457-1470.

Thompson W. F. and Quinto L. (in press). Music and emotion: Psychological considerations. In P. Goldie and E. Schellekens (eds.), *Philosophy and Aesthetic Psychology*. Oxford: Oxford University Press.