International Symposium on Performance Science
ISBN 978-94-90306-01-4

Emotional lingering: Facial expressions of musical closure

William Forde Thompson[1], Rachel Bennetts[1], Bojan Neskovic[1], and Caroline Palmer[2]

[1] Department of Psychology, Macquarie University, Australia
[2] Department of Psychology, McGill University, Canada

We report evidence that singers maintain emotional facial expressions after vocalization has terminated, introducing a form of emotional lingering. Emotional lingering extends and complements acoustic signals of emotion, providing a visual signal of musical closure. We first describe evidence from production studies that emotional facial expressions continue beyond the acoustic dimension of music. We next describe evidence that perceivers are sensitive to facial expressions that occur beyond the production of sound, and that such signals carry reliable emotional information. We note that audiovisual experiences of music are compatible with current understandings of multisensory integration in the central nervous system. As such, investigations of performance should include consideration of facial expressions and other performance gestures.

Keywords: music cognition; emotion; singing; closure; facial expression

Music is a powerful medium for conveying emotion (Juslin and Sloboda 2001). Emotional information is carried not only in musical sound; it is also signaled in the facial expressions of performers (Livingstone *et al.* 2009, Thompson *et al.* 2005, Thompson *et al.* 2008). This paper concerns the nature and significance of facial expressions during sung music performance, focusing on expressions that occur immediately following the cessation of vocal production.

Effective singing requires a stage in which plans for vocal and emotional expression are created, coordinated, and implemented. Vocal planning facilitates accurate production of pitch, timing, intensity, and vocal quality (timbre). It recruits a range of muscle groups including the vocal cords, throat, face, and other parts of the body that support vocal production. The

nature and time course of vocal planning can be inferred by examining subtle movements in these muscle groups that occur prior to sound production. Such movements reflect a dynamic process in which action and decision-making are tightly coupled (McKinstry *et al.* 2008).

Muscular activity associated with emotional communication overlaps with muscular activity used for effective singing, but it is not identical. While a performer is singing it may be difficult to disentangle movements that support accurate vocal production from those that support emotional communication. Similarly, movements that occur immediately prior to vocal production combine plans for accurate vocal production with plans for emotional communication.

Once vocal production has terminated, however, the bodily movements and facial expressions that persist cannot be interpreted as a reflection of vocal production mechanisms. Emotional facial expressions linger beyond the time it takes to deactivate muscular activity, often lasting several seconds or more. Such activity can only be interpreted as a constituent of emotional communication and is referred to as *emotional lingering*.

Emotional lingering can be investigated in both production and perception studies. First, facial expressions and other movements that occur following the cessation of vocal production can be captured and analyzed using technologies such as motion capture and electromyography. These strategies can be used to confirm the existence of emotional lingering and determine its properties and time course. Second, perceptual studies can be used to determine how perceivers are affected by emotional lingering phenomena. Do facial expressions that persist beyond the termination of sound provide reliable emotional information, and do they influence emotional interpretations of the music?

Livingstone *et al.* (2009) confirmed that singers maintain emotional facial expressions beyond the cessation of vocal production. Participants were recorded with motion capture or electromyography (EMG) as they watched and imitated phrases of emotional singing. All were shown audiovisual recordings of sung phrases performed with happy, sad, or neutral emotional expressions. Their task was to imitate the target stimulus, emphasizing the emotion expressed. Facial expressions for happy and sad sung phrases were significantly different from facial expressions for neutral sung phrases. However, these differences were observed not only during the imitation, but also before and after vocal production. The authors concluded that facial expressions not only support music performance during vocal production, but also reflect processes of emotional *planning* and *lingering*.

In the current investigation, we examined whether the facial expressions that occur after vocal production provide reliable emotional information to listeners. If so, it would support the notion that such expressions can be used to clarify, underscore, or modify emotional intentions. A full description of the investigation will be described in a forthcoming publication.

METHOD

Participants

Seventeen first year psychology students took part in the experiment in return for course credit. All had normal or corrected to normal vision.

Materials

Stimuli were excerpts of video recordings of facial expressions used by four singers (Livingstone *et al.* 2009). Stimuli were presented as full video or point-light display (13 markers). Full videos were recorded using a Sony DV camera, positioned roughly 45 degrees to the right of participants at a distance of 2.5 m. Motion capture data were collected from 13 markers on the face using an 8-camera Vicon optical motion capture system (Oxford Metrics), calibrated to less than 1 mm error. Six stimuli were chosen from each singer, depicting two sung phrases for each of three emotions: happy, neutral, and sad. Clips from three time periods or *epochs* were excerpted from each trial: pre-production (three beats prior to vocalization, or 2.4 s), production (the middle three beats during vocalization), and post-production (the first three beats after the singer stopped singing). Each epoch was depicted in two formats: video and point-light display. Videos were edited in *Final Cut Pro* (Apple) to reveal only the face and shoulders of the singer on a neutral (beige) background. Point-light displays were produced by processing motion capture data for each trial in *Matlab* (Mathworks) and consisted of black dots on a white background.

The experiment was run on a *MacBook Pro* laptop, using *Experiment Creator 1.4* (available from WFT's website). Stimuli were presented on a 24" Samsung monitor. Participants were seated roughly 50-100 cm from the screen.

Procedure

Each participant completed 144 trials (4 singers, 2 phrases, 3 epochs, 3 emotions, 2 display types). Display type (video or point-light display) was blocked and counterbalanced and the order of presentation within blocks was ran-

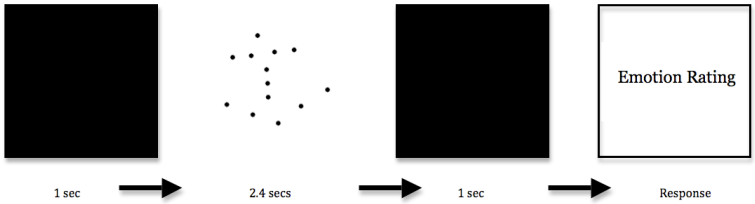

Figure 1. Timeline of a trial involving point-light display. Each black dot represents a marker used to track facial motion. Markers were placed on the forehead, temples, nose-bridge, nose-tip, eyebrows (2 each), and lips (upper, lower, and corners).

domized. Participants completed six practice trials for each display type before completing each block. Illustrated in Figure 1, each trial began with 1 s of black screen, followed by the 2.4 s stimulus and then another 1 s of black screen. Participants then rated the emotion perceived on a scale from -3 (very sad), to 0 (neutral), to +3 (very happy). A computer mouse was used to select ratings displayed on the monitor. Once participants rated the stimulus they could progress to the next trial when ready.

RESULTS

Preliminary analyses revealed significant interactions involving display type, motivating separate analyses for video and point-light display conditions. For each, an Analysis of Variance (ANOVA) was conducted with repeated measures on "emotion" (happy, sad, neutral) and "epoch" (pre-production, production, post-production).

For full video trials, there was a main effect of emotion ($F_{2,32}$=235.82, p<0.001) with high ratings assigned to happy trials (M=1.59, SE=0.10), intermediate ratings assigned to neutral trials (M=-0.47, SE=0.06), and low ratings assigned to sad trials (M=-1.29, SE=0.11). Although this general pattern was observed in all three epochs, there was a significant interaction between emotion and epoch ($F_{4,64}$=20.48, p<0.001). This interaction is illustrated in Figure 2 (left panel), which displays mean ratings of emotion for each of the three epochs. Analysis of data for the lingering epoch separately revealed a main effect of emotion ($F_{2,32}$=177.31, p<0.001), and planned comparisons verified that ratings for the three emotions differed significantly from each other. Thus, facial expressions that lingered beyond song production provided reliable emotional cues.

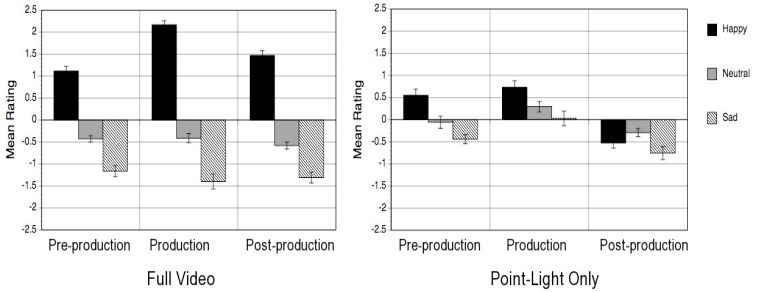

Figure 2. Mean emotion ratings for full video (left panel) and point-light display (right panel), and standard error bars.

For point-light display, there was a main effect of emotion ($F_{2,32}$=25.05, p<0.001), with highest ratings assigned to happy trials (M=0.25, SE=0.09), intermediate ratings assigned to neutral trials (M=-0.02, SE=0.08), and lowest ratings assigned to sad trials (M=-0.39, SE=0.11). There was also a significant interaction between emotion and epoch ($F_{4,64}$=4.60, p<0.01), illustrated in Figure 2 (right panel). Analysis of the post-production epoch separately revealed a main effect of emotion ($F_{2,32}$=5.38, p<0.05). However, planned comparisons indicated that only the neutral and sad conditions differed significantly in emotion ratings. During the post-production epoch, point-light display did not provide participants with sufficient information to differentiate happy facial expressions from other facial expressions.

DISCUSSION

Results confirm that facial expressions that linger beyond the cessation of vocal production contain reliable emotional cues. Indeed, when full video was available, discrimination of emotional facial expressions was observed prior, during, and after vocalization. Such facial expressions of emotion may function to shape a listener's interpretation of the music (Thompson *et al.* 2008).

When only point-light display was available, discrimination was observed prior and during vocalization, but participants could not discriminate happy expressions from other emotional intentions in the post-production epoch, indicating that marker movement alone carried insufficient information for decoding happy expressions in that epoch. It is possible that isolated happy emotional expressions are ambiguous in the lingering epoch because they begin with raised lip corners and end with a neutral expression, implying

movement away from a positive emotional expression. Alternatively, in the absence of recognizable facial expressions, movement that comes to a stop may generally have negative connotations.

Facial expressions of emotion are likely to be integrated with auditory signals of emotion, yielding audiovisual experiences of music. Such experiences are consistent with current understandings of multisensory integration in the central nervous system, and underscore the importance of considering facial expressions and other performance gestures in studies of music and emotion.

Acknowledgments

This research was supported by an ARC Discovery Grant DP0987182 awarded to the first and last authors. Alex Chilvers, Catherine Greentree, and George Narroway provided research assistance.

Address for correspondence

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

References

Juslin P. N. and Sloboda J. A. (2001). *Music and Emotion*. Oxford: Oxford University 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.

McKinstry C., Dale R., and Spivey M. (2008). Action dynamics reveal parallel competition in decision making. *Psychological Science, 19,* pp. 22-24.

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

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