Head Movements Encode Emotions During Speech and Song

identify the emotions of vocalists from their head movements alone.

Experiment 1

In Experiment 1, vocalists' head motion was recorded with optical motion capture while they spoke and sang four statements, each with five emotional intentions (very happy, happy, neutral, sad, and very sad). In the speaking condition, vocalists memorized the statements and produced them when cued. In the singing condition, vocalists sang the same statements with a seven-note melody that maintained the same frequencies (note pitches) across emotion conditions. On the basis of previous findings that visual cues for the head pose affect perceived emotion (in the absence of vocalization; Atkinson et al., 2004; Boone & Cunningham, 1998), we hypothesized that happy displays would exhibit upward turning of the head and that sad displays would show downward turning of the head. Previous research has also found that the magnitude of facial feature movements is correlated with the emotional intensity of the experienced event (

To confirm that trials selected for the perceptual task were representative of the full sample, we examined acoustic measures including mean F_0, vocal intensity, and duration (see Analyses: Vocal acoustics for a description of acoustic analyses). These acoustic measures were selected because they are primary cues to vocal emotion (Cowie et al., 2001; Scherer, 2003). Mean values of the seven unselected vocalist–channel–emotion trials were compared with those of the selected trials using separate paired-samples t tests. The perceptual task stimuli were not significantly different from the full sample on any of the acoustic measures: F_0, $t(119)$ 1.23, p .22; vocal intensity, $t(119)$ 0.2,

$$R_z(\ ) = \begin{bmatrix} \cos & \sin & 0 \\ \sin & \cos & 0 \\ 0 & 0 & 1 \end{bmatrix},\ R_y(\ ) = \begin{bmatrix} \cos & 0 & \sin \\ 0 & 1 & 0 \\ \sin & 0 & \cos \end{bmatrix},\ R_x(\ ) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos & \sin \\ 0 & \sin & \cos \end{bmatrix}. \quad (3)$$

When factorized, the final rotation operation R^{World}_{Head} was defined by Equation 4:

$$R^{World}_{Head} = \begin{bmatrix} \cos\ \cos & \sin\ \cos\ \ \cos\ \sin\ \sin & \sin\ \sin\ \ \cos\ \sin\ \cos \\ \sin\ \cos & \cos\ \cos\ \ \sin\ \sin\ \sin & \cos\ \sin\ \ \sin\ \sin\ \cos \\ \sin & \cos\ \sin & \cos\ \cos \end{bmatrix}. \quad (4)$$

Head rotation angles (in radians) were extracted from the rotated data: yaw ( ) rotation around the *z* axis, which corresponded to turning the head left or right; pitch ( ) rotation around the *y* axis, which corresponded to tilting the head up to the ceiling or down to the floor; and roll ( ) rotation around the *x* axis, which corresponded to leaning the head to either shoulder. These measures are defined, respectively, in Equation 5:

$$\text{yaw}(\ ) = \tan^{-1}\frac{r_{yx}}{r_{xx}},\ \text{pitch}(\ ) = \sin^{-1}(r_{zx}),\ \text{roll}(\ ) = \tan^{-1}\frac{r_{zy}}{r_{zz}}. \quad (5)$$

Translational motion of the participant's head (in mm), representing the three axes of motion were defined as *horizontal* (x) movement along the *x* axis, which corresponds to moving forward or backward; *lateral* (y) movement along the *y* axis, which corresponds to moving left or right; and *vertical* (z) movement along the z axis, which corresponds to moving up or down.

Analyses.

Head movements. Motion data were set to zero for a "neutral resting" position of the participant's head during a 2,000-ms baseline window prior to each trial onset, using a baseline subtraction procedure. The modal value within this baseline window was subtracted from displacement values during the trial time line. The baseline-adjusted marker trajectories represented how the head position deviated throughout the trial from its resting position.

Functional data analyses were performed using the Functional Data Analysis toolbox (Ramsay & Silverman, 2010). Occasional missing data were interpolated (less than 0.0001% of data). Order 6 B-splines were fit to the second derivative of marker trajectories with a ratio of 1:4 knots to data samples. The data were smoothed using a roughness penalty on the fourth derivative (10^{-8}). Feature registration (time warping) was used to align trajectories to the syllable boundaries (six events) as determined from the acoustic analyses, described in Analyses: Vocal acoustics. Functional data were resampled to produce 75 equally spaced data points per syllable (300 ms at 250 Hz) for the first six syllables, with the final syllable resampled to 150 data points (600 ms at 250 Hz). This derivation enabled a syllable-matched comparison of head movements across speech and song. Functional analyses of variance (fANOVA) were used to examine whether differences were present in the motion trajectories at each time point (see also Livingstone et al., 2012). Significance levels were corrected for multiple comparisons with False discovery rate using the Benjamini–Hochberg–Yekutieli procedure for dependent statistical tests, with a q value of .05 (Benjamini & Hochberg, 1995; Benjamini & Yekutieli, 2001). We report mean *F* statistic, *p* value, and eta-squared values across time regions that reached statistical significance (p .05). Motion effect sizes are reported as eta-squared values.

Vocal acoustics. Acoustic recordings were analyzed with Praat (Boersma & Weenink, 2010). Speech utterances were segmented at syllable boundaries and were coded by a rater; 8% of the samples were checked by a second rater (mean interrater boundary time difference 0.0026 s, *SD* 0.0024 s). Boundaries were determined by changes in the spectrogram, and in the F_0 and acoustic intensity contours. Singing trials were inspected by the raters to confirm that vocalists had sung the prescribed melody. F_0 contours were extracted using an autocorrelation algorithm (ac) in Praat. Mean F_0 values across the entire utterance were analyzed with a mixed-design analysis of variance (ANOVA). When Mauchly's sphericity test was significant, Greenhouse-Geisser's correction was applied. Acoustic effect sizes report partial eta-squared values. All reported means are accompanied by 95% confidence intervals in square brackets. Pairwise comparisons were adjusted using Bonferonni correction. All statistical tests were conducted in Matlab 2013b and SPSS 20.0.0 (IBM Corp, 2011).

Perceptual task. Data from the perceptual task were coded as 1 (*correct*) when the emotional category selected by the listener matched the emotional category that the vocalist had been in-pitch, and roll rotation angles (see Atkinson et al., 2004; Keltner, 1995, for an examination of nonvocal head movement during emotional expression).

A three-way repeated-measures fANOVA was conducted on vocalists' horizontal head measures (deviation from resting position of head) by channel (2 levels: speech, song), emotion (5 levels: very happy, happy, neutral, sad, very sad), and statement (4 statements mentioned in the Stimuli and Apparatus section). No effect of channel or its interactions was found in the analysis, confirming that speech and song exhibited similar horizontal head movements. Figure 3a shows the mean horizontal head displacement (forward-backward movement, x) values across all trials by emotion conditions. Regions of statistical significance (p .05) are indicated by the black horizontal bar in the time line. The main effect of emotion was found during the majority of vocal sound productions, F(4, 44) 7.17, p .01, ^{2} .05. Happy and sad utterances were characterized by larger forward movement of the head than were emotionally neutral utterances, with very happy utterances appearing to move forward quickly at the onset of vocalization. Very happy and very sad productions also appeared to exhibit larger forward displacement relative to their happy and sad counterparts, respectively. No effect of statement or its interactions was found in the analysis, confirming that lexical variability did not affect horizontal head movements.

The previous analysis suggested that the velocity of horizontal head movements may vary with the expressed emotion. To examine this relationship, we conducted a follow-up analysis that involved a three-way repeated-measures fANOVA on the velocity of vocalists' horizontal head movements (forward–backward, x). No effect of channel or its interactions was found in the analysis, confirming that speech and song exhibited similar velocities of horizontal movements. Figure 3b shows the mean velocity of vocalists' horizontal head movements across all trials by emotion conditions. A main effect of emotion was found, F(4, 44) 9.26, p .01, ^{2} .13, which began 100 ms prior to vocal onset and continued for 220 ms during the start of vocal sound production. These results suggest that the forward-moving velocity of head movements distinguished the vocalized emotion and that very happy utterances exhibited greater forward-moving velocity than did other emotions. No effect of statement or its interactions was found in the analysis, confirming that lexical variability did not affect the velocity of horizontal head movements.

A three-way repeated-measures fANOVA was conducted on vocalists' vertical head measures (moving up–down, z). No effect of channel or its interactions was found in the analysis, confirming that speech and song exhibited similar vertical movements. Figure 3c shows vocalists' mean vertical head displacement (deviation from resting position) across all trials by emotion conditions. A

Figure 3. Main effects of emotion conditions on four aspects of head movements in Experiment 1. Each trajectory line is the functionally time-warped mean across all actors, vocal channels, statements, and repetitions (192 trials per trajectory line). Zero represents the neutral "at rest" position of the head. Dashed vertical lines between vocal onset and offset indicate syllable boundaries. Black horizontal lines below trajectories indicate regions of significance at p .05. Error bars are indicated by shaded regions around trajectory lines. (a) Mean horizontal displacement of head movement (forward–backward, x). (b) Mean horizontal velocity of head movement (forward–backward, X). (c) Mean vertical displacement of head (up–down, z). (d) Mean rotational pitch of head movement (looking up–down).

significant main effect of emotion was found throughout vocalization, which continued long after vocal sound ended, $F(4, 44)$ 8.47, p .01, ^{2} .11. Importantly, these effects began 3,196 ms prior to the onset of vocal sound, confirming that vocalists' head movements are not only a result of vocal sound production. These results indicated that vocalists vertically raised their head during happy and neutral vocalizations and lowered their head during sad because movements succeeding vocalization are directly influenced by the preceding movements used in the control of F_0. We regressed the mean F_0 of an utterance on the means of the six degrees of head motion during the time period from stimulus presentation to vocal onset for that same utterance. A stepwise regression with backward elimination was performed. For the stepping method criteria, the p value for including a variable was set at .05 and the p value for excluding a variable at .10. The model with the final adjusted R^2 was chosen. Head movements in the absence of vocal sound were found to explain a significant proportion of variance in F_0, R^2 0.02, $F(3, 959)$ 7.67, p .001. Horizontal, .099, $t(959)$ 3.09, p .002; vertical, .088, $t(959)$ 2.74, p .006; and yaw, .078, $t(959)$ 2.44, p .015, movements were found to significantly predict F_0 values. However, although head movements in the 7 s preceding vocal production were related to their subsequent F_0, the movements explained a much smaller percentage of the variance than did the movements that occurred during vocal sound production (2% vs. 71.7%, respectively). This relationship is to be expected, because movements preceding vocal sound and during F_0 production both reflect the expressed emotion. These results further confirm that vocalists' F_0 are related to their head movements, but only to an extent.

Perceptual task. Mean response rates for each of the three emotion categories in the perceptual task are shown in Table 1. Participants' mean accuracy across all utterances was M .71, SE .02. To determine whether observers' accuracy scores difvocalists’ head movements function as a visual form of emotional communication. We examined this hypothesis further in the next experiment.

Experiment 2

In Experiment 2, viewers identified vocalists’ emotional intent from silent visual displays of their head movement during speech and song. Silent video recordings of vocalists’ heads with facial features occluded were presented, such that only whole-head motion cues were present. Because similar head movements for speech and song were found in Experiment 1, we hypothesized that observers would reliably distinguish the intended emotion of the vocalists from their head motion with similar levels of accuracy for both speech and song.

Method

Participants. Twenty-four native English-speaking adults (17 female and seven male, mean age 21.6 years, *SD* 3.9) were recruited from the McGill community. Participants had received varied amounts of private music instruction (*M* sponse measure (neutral, happy, sad). Participants also completed a background questionnaire about their singing, acting, and musical experience. Participation in the experiment took approximately 30 min.

Data were coded as 1 (correct) when the emotional category motion. Motion capture analysis demonstrated that vocalists' head movements showed similar patterns of translational displacement and velocity and rotational movement for a given emotional intent (very happy, happy, neutral, sad, very sad) across speech and song, transcending changes in fundamental frequency that were systematic (fixed) in song but free in speech. The head movement trajectories were also comparable across four different sentences, confirming that head movements were not affected by lexical variability. Emotion-related head movements also occurred immediately prior to and after sound production, with rotational pitch (floor-or-ceiling looking) and vertical head motion beginning 3–4 s before the onset of vocalization and continuing for 2.4 s after vocalization had ended. The influence of head movements—presented without facial features or sound—on viewers' subsequent identification of emotional intent provides further evidence that head motion during speech and song functions as a visual form of emotional communication.

Vocalists' fundamental frequency (F_0) varied predictably with vocalists' emotional intent in speech, increasing for happy and very happy utterances relative to neutral, sad, and very sad vocalizations (Cowie et al., 2001; Scherer, 2003). In contrast, F_0 remained constant across emotions in song, varying only with the prescribed melody. Changes in vocalists' head movements correlated with changes in their F_0; the variance in F_0 accounted for by the six degrees of freedom of head motion was lower for song, in which F_0 is fixed by the melody, than for speech. The fact that other factors influence head movements beyond F_0 is underscored by the presence of emotionally distinct head movements in the absence of F_0 (right before or after vocalizations). The variation seen in head movements during vocalized sound may reflect orofacial movements of the jaw (Tasko & McClean, 2004; Vatikiotis-Bateson & Ostry, 1995), which are tied to the production of vocal intensity, and suprasegmental features of speech such as stress and prominence (Hadar, Steiner, Grant, & Rose, 1983; Munhall et al., 2004). Thus, vocal head movements observed here seemed to arise from at least two distinct sources—changes in acoustic parameters of the voice and changes in the expression of emotion. These results provide a novel extension to models of head motion in which linguistic and acoustic motor factors are the primary determinants of head movements (Munhall et al., 2004; Thompson & Russo, 2007; Thompson et al., 2010; Yehia et al., 2002, 1998).

Vocalists' emotions were differentiated not by a single movement trajectory but by a collection of head movement cues. Happy vocalizations exhibited larger and faster forward movement, raising of the head, and ceiling-looking rotational pitch. Conversely, sad expressions had slower forward movement, lowering of the head, and floor-looking rotational pitch. Neutral expressions exhibited attenuated movement along these dimensions. The emotional intensity of vocalizations was further differentiated by the magnitude of head movements: More-intense vocalizations generally showed larger movement patterns than did their less-intense counterparts, with very happy utterances exhibiting larger and faster forward movements and more ceiling-looking rotation relative to happy vocalizations and very sad utterances exhibiting more-forward movement and lowering of the head and greater floor-looking rotation relative to sad vocalizations. Encoding of emotion through a collection of cues, in this case the dimensions, magnitude, and direction of head movements, is also thought to be the basis of emotional communication in facial expressions, where expressions can be described through the combination of muscle action units (Ekman & Friesen, 1978). Similarly, emotional communication in music and speech acoustics is thought to occur through variations in acoustic cues (Juslin & Laukka, 2001, 2003; Livingstone et al., 2010; Scherer, 1995, 2003). Importantly, Experiment 2 confirmed that observers could identify vocalists' intended emotions from the presence of head movements alone. These results support the hypothesis that head movements function as part of an emotional code that is shared by encoders (vocalists) and decoders (observers; Ekman & Friesen, 1969; Elfenbein & Ambady, 2002). The use of head gestures by vocal performers may therefore function to support the emotional message conveyed by the voice, complementing the auditory signal with motion.

The present study contributes several novel findings to the understanding of emotion. First, specific patterns of head movements were found to reflect a vocalist's emotional intent for expressions of happiness and sadness. The relationship between head pose and an individual's emotional state has been posited (Atkinson et al., 2004; Darwin, 1872/1965; Matsumoto & Willingham, 2009; Wallbott, 1998), with prior work finding a link between head tilt and the nonverbal expression of self-conscious emotions such as pride and shame (Keltner, 1995; Nelson & Russell, 2014; Tracy & Robins, 2004). This is the first study to our knowledge to quantify the extent of these movement parameters, which included direction, rotation, and velocity of head motion, and that these movements distinguished happy and sad expressions of emotion, along with the intensity of these expressions. Second, we found that emotional head movements occurred before and after vocal sound production. Ceiling- and floor-looking rotational pitches and vertical movement of the head began prior to the onset of sound. When vocalization was initiated, forward movement of the head co-occurred with emotionally distinct forward-moving velocity with more-pronounced rotational pitch and vertical head movements. Although previous research has noted a relationship between rotational pitch and head motion (Atkinson et al., 2004; Boone & Cunningham, 1998), we are not aware of any reference to forward movement of the head in connection with emotion; expressive head movements accompanying vocalization may be qualitatively different from those occurring in a nonvocal context. Third, observers accurately identified a vocalist's emotional intent solely on the basis of head movements. To our knowledge, this is the first demonstration of emotion identification based solely on dynamic displays of head motion. Observers in a previous study (Ekman, 1965) identified emotions emotional recognition is unknown. In contrast, the present study used video recordings of participants to capture the full range of expressive head movements, without the influence of facial features.

An ability to recognize emotional intent from head motion alone may also have applications in the area of automatic emotion recognition. With the rapid pace of development in artificial intelligence and robotics, computers are increasingly used in security situations, where video cues in the absence of audio cues is common, and as virtual or robotic agents that interact with humans, an area referred to as socially interactive robotics (Fong, Nourbakhsh, & Dautenhahn, 2003; Ishiguro, 2007). The addition of expressive vocal head movements may improve the realism of vocalizing humanoid robots and on-screen virtual avatars (Busso, Deng, Grimm, Neumann, & Narayanan, 2007), helping to close the realism gap referred to as the "uncanny valley" (Mori, MacDorman, & Kageki, 2012). That head movements encode a vocalist's emotional intentions may have applications in the area of automated affect recognition. Decoding an individual's emotional state in situations of visual noise and facial occlusion may be improved by using the information encoded in vocal head movements (Fasel & Luettin, 2003).

The present study has two notable limitations. Head movements were examined for only two emotional categories: happiness and sadness, with a neutral control condition. These emotions were selected because they exhibit low rates of confusion in perceptual studies. Happiness and sadness reflect differences in emotional valence and are thought to embody positive and negative valence, respectively (Frijda, 1986). Differences between the positive emotional conditions (very happy, happy) and negative emotional conditions (very sad, sad) may reflect aspects of emotional valence (upward for positive, downward for negative), rather than distinct categories of basic emotions. Differences in head tilt may also reflect the role of self-conscious emotions (Keltner, 1995; Nelson & Russell, 2014; Tracy & Robins, 2004). An important avenue for future work would be to investigate a broader set of discrete emotions, including anger, fear, disgust, and surprise. Because the present work reported on several novel aspects of vocal emotional head movements, including the effect of intensity on size of head movements, we theorize that a broader set of distinct emotions can be conveyed to observers by a vocalist's head movements.

A second limitation was the use of induced emotional expressions (sometimes called "simulated" or "posed" in the facial expression literature, see Bänziger et al., 2012; Gosselin, Kirouac, & Dore, 1995), instead of naturally occurring "spontaneous" displays of emotion. A criticism of instructed expressions is that they can be exaggerated, leading to inflated rates of observer agreement relative to spontaneous displays (Cohn & Schmidt, 2004; Elfenbein & Ambady, 2002; Motley & Camden, 1988; Wagner, MacDonald, & Manstead, 1986). The use of spontaneous expression paradigms may yield attenuated patterns of head movement relative to those reported in Experiment 1 or result in lower rates of observer agreement than those found in Experiment 2, or both. Induction procedures like the one used in the current study are increasingly used to obtain more-genuine expressions of emotion, while retaining the control required for a within-subject design (see also Bänziger et al., 2012; Douglas-Cowie et al., 2007; Ebner, Riediger, & Lindenberger, 2010; Kaulard et al., 2012; Livingstone et al., 2014). The comparison of induced and spontaneous displays of vocal emotional communication is an important avenue for further research. Video-based motion tracking, which does not require the use of motion capture markers, will likely be required to enable natural recordings of live music performances and real-world speech, although there remain many unsolved technical and theoretical issues when using naturalistic recordings (Douglas-Cowie, Campbell, Cowie, & Roach, 2003; Scherer, 2003).

Conclusion

In two experiments, we demonstrated that vocalists' head movements encode emotional information in speech and song and that observers could identify emotions on the basis of head movements alone. Seasoned oratory and singing performers are renowned for their dramatic and expressive use of head and facial movements. These head movements may therefore function in part to support the acoustic signal, communicating emotion through motion.

References

Ambadar, Z., Cohn, J. F., & Reed, L. I. (2009). All smiles are not created equal: Morphology and timing of smiles perceived as amused, polite, and embarrassed/nervous. *Journal of Nonverbal Behavior, 33,* 17–34. http://dx.doi.org/10.1007/s10919-008-0059-5

Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception, 33,* 717–746. http://dx.doi.org/10.1068/p5096

Bänziger, T., Mortillaro, M., & Scherer, K. R. (2012). Introducing the Geneva Multimodal expression corpus for experimental research on emotion perception. *Emotion, 12,* 1161–1179. http://dx.doi.org/10.1037/a0025827

Bassili, J. N. (1978). Facial motion in the perception of faces and of emotional expression. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 373–379. http://dx.doi.org/10.1037/0096-1523.4.3.373

Bassili, J. N. (1979). Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face. *Journal of Personality and Social Psychology, 37,* 2049–2058. http://dx.doi.org/10.1037/0022-3514.37.11.2049

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B: Methodological, 57,* 289–300.

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics, 29,* 1165–1188.

Boersma, P., & Weenink, D. (2010). Praat: Doing phonetics by computer (Version 5.2.15) [Computer program]. Retrieved from http://www.praat.org/

Boone, R. T., & Cunningham, J. G. (1998). Children's decoding of emotion in expressive body movement: The development of cue attunement. *Developmental Psychology, 34,* 1007–1016. http://dx.doi.org/10.1037/0012-1649.34.5.1007

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436. http://dx.doi.org/10.1163/156856897X00357

Brunswik, E. (1956). *Perception and the representative design of psychological experiments.* Berkeley: University of California Press.

Bugental, D. B. (1986). Unmasking the "polite smile": Situational and personal determinants of managed affect in adult-child interaction. *Personality and Social Psychology Bulletin, 12,* 7–16. http://dx.doi.org/10.1177/0146167286121001

Busso, C., Deng, Z., Grimm, M., Neumann, U., & Narayanan, S. (2007). Rigid head motion in expressive speech animation: Analysis and syn-

thesis. IEEE Transactions on Audio, Speech, and Language Processing, 15, 1075–1086.
Cacioppo, J. T., Petty, R. E., Losch, M. E., & Kim, H. S. (1986). Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. Journal of Personality and Social Psychology, 50, 260–268. http://dx.doi.org/10.1037/0022-3514.50.2.260
Carlo, N. S., & Guaitella, I. (2004). Facial expressions of emotion in speech and singing. Semiotica, 2004, 37–55. http://dx.doi.org/10.1515/semi.2004.036
Cohn, J. F., & Schmidt, K. L. (2004). The timing of facial motion in posed and spontaneous smiles. International Journal of Wavelets, Multiresolution, and Information Processing, 2, 121–132. http://dx.doi.org/10.1142/S021969130400041X
Coltheart, M. (1981). The MRC Psycholinguistic Database. Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 33, 497–505. http://dx.doi.org/10.1080/14640748108400805
Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., & Taylor, J. G. (2001). Emotion recognition in human-computer interaction. IEEE Signal Processing Magazine, 18, 32–80. http://dx.doi.org/10.1109/79.911197
Cunningham, D. W., & Wallraven, C. (2009). Dynamic information for the recognition of conversational expressions. Journal of Vision, 9, 7. http://dx.doi.org/10.1167/9.13.7
Dalla Bella, S., Peretz, I., Rousseau, L., & Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. Cognition, 80(3), B1–B10. http://dx.doi.org/10.1016/S0010-0277(00)00136-0
Darwin, C. (1965). The expression of emotions in man and animals. Chicago, IL: University Chicago Press. (Original work published 1872)
Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. Psychology of Music, 21, 103–113. http://dx.doi.org/10.1177/030573569302100201
Dittrich, W. H., Troscianko, T., Lea, S. E., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. Perception, 25, 727–738. http://dx.doi.org/10.1068/p250727
Douglas-Cowie, E., Campbell, N., Cowie, R., & Roach, P. (2003). Emotional speech: Towards a new generation of databases. Speech Communication, 40, 33–60. http://dx.doi.org/10.1016/S0167-6393(02)00070-5
Douglas-Cowie, E., Cowie, R., Sneddon, I., Cox, C., Lowry, O., McRorie, M., . . .

Wallbott, H. G., & Scherer, K. R. (1986). Cues and channels in emotion recognition. Journal of Personality and Social Psychology, 51, 690–699. http://dx.doi.org/10.1037/0022-3514.51.4.690

Wanderley, M. M., Vines, B. W., Middleton, N., McKay, C., & Hatch, W. (2005). The musical significance of clarinetists' ancillary gestures: An exploration of the field. Journal of New Music Research, 34, 97–113. http://dx.doi.org/10.1080/09298210500124208

Wiener, M., Devoe, S., Rubinow, S., & Geller, J. (1972). Nonverbal behavior and nonverbal communication. Psychological Review, 79, 185–214. http://dx.doi.org/10.1037/h0032710

Yehia, H. C., Kuratate, T., & Vatikiotis-Bateson, E. (2002). Linking facial animation, head motion and speech acoustics. Journal of Phonetics, 30, 555–568. http://dx.doi.org/10.1006/jpho.2002.0165

Yehia, H. C., Rubin, P., & Vatikiotis-Bateson, E. (1998). Quantitative association of vocal-tract and facial behavior. Speech Communication, 26, 23–43. http://dx.doi.org/10.1016/S0167-6393(98)00048-X

Received March 25, 2015
Revision received June 16, 2015
Accepted June 24, 2015