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)

The Coupling of Gesture and Sound: The Kinematics of Cross-Modal Matching for Hand Conducting Gestures and Accompanying Vocal Sounds

Aysu Erdemir,[1] Erdem Erdemir,[2] Emelyne Bingham,[3] Sara Beck,[1] John Rieser[1]

[1]*Psychology and Human Development in Peabody College, Vanderbilt University, USA*
[2]*Electrical Engineering and Computer Science, Vanderbilt University, USA*
[3]*Blair School of Music, Vanderbilt University, USA*

aysu.erdemir@vanderbilt.edu, erdem.erdemir@vanderbilt.edu, emelyne.m.bingham@vanderbilt.edu, saralynbeck@gmail.com, j.rieser@vanderbilt.edu

ABSTRACT

Background

Physical movement of musicians and conductors alike play important role in music perception. Recently computational techniques have been developed to look at the relationship between movement kinematics and musical features. Areas that are investigated included musicians' bodily movements (Thompson & Luck, 2008), dance movements (Burger et al., 2010), and conductors' gestures (Luck & Toiviainen, 2006). This study adds to previous literature by exploring cross-modal similarities between hand movements and accompanying sound profiles across persons.

Aims

The study was designed to identify whether there was a predictable mathematical relationship between hand gestures performed by an expert conductor and vocal responses of a general adult sample. Our empirical work has found that adults automatically vary their utterance of the syllable /dah/ in a way that matches the motion characteristics of the hand gestures being observed, but the physical nature of this relationship remained unclear. Thus, this study specifically aimed at bringing out correlations among the kinematic features of the gestures and their accompanying sounds using a computational data quantification method.

Method

We asked adults (n=36), some with and others without musical training, to produce the syllable /dah/ repeatedly while watching the video of an expert conductor as she performed four different right hand gestures, namely flicks, punches, floats and glides, at constant tempo. The audio recordings were imported into Matlab from which we extracted an RMS (root mean square) amplitude profile for each participant, and then averaged across all participants to get an overall average RMS amplitude profile for each of the four gestures.

In order to examine the motion data, we captured the movements of the conductor using a high-resolution motion capture system (Vicon) as she performed the four different gestures. A total of 5 markers were attached to the right hand and arm of the conductor. The position data for each marker were represented as trajectories of dynamic stick figures from which the basic kinematic components of velocity were extracted using a numerical differentiation algorithm. The kinematic features were, then, compared to their amplitude counterparts in the audio tracks. Specifically, the structural similarities between the velocities and the RMS amplitude were examined using a series of correlation analyses.

Results

Correlation analyses showed very strong relations among the velocity profiles of the movements and their accompanying sound-energy profiles. Specifically, the velocity profile of the gestures recorded from the hand markers significantly correlated with the RMS amplitude profile of their audio counterparts, r = .96 for flicks, r = .81 for punches, r = .90 for floats, and r = .97 for glides. Additional analyses revealed that the velocity profiles were very similar for flicks & punches, and floats & glides. This can be used to explain why observers found it more difficult to distinguish flicks from punches, and floats from glides in the audio recordings. Deeper analysis showed that average velocity until the peak (first ~15 ms) reliably predicted the RMS amplitude in their auditory counterparts, i.e. faster initial speed caused louder responses. In other words, flicks & punches, and floats & glides could be differentiated based on their initial velocity & amplitude, which correlated significantly (r = .98).

Conclusions

Relationships between a conductor's hand gestures and their accompanying vocal sounds have been quantified computationally. The observed structural similarity between the movement and sound data might be due to a direct mapping of the visual representation of observed action onto one's own motor representation which is reflected in its resultant auditory effects. The visual effects from the kinematics of movement patterns are supposed to be automatically translated into auditory responses that are similar in shape. We hypothesize that this association is mediated through the motor system and/or so-called mirror neurons in the brain.

Keywords

Music and Movement, Conducting, Gestures, Movement Kinematics, Motion-Capture.

REFERENCES

Burger, B., Thompson, M.R., Saarikallio, S., Luck, G., & Toiviainen, P. (2010). Influence of musical features on characteristics of music-induced movements. In Proceedings of the 11th International Conference on Music Perception and Cognition. Seattle, WA.

Luck, G. P. B. & Toiviainen, P. (2006). Ensemble musicians' synchronization with conductors' gestures: an automated feature-extraction analysis. Music Perception, 24(2), 195-206.

Thompson, M. R., & Luck., G. (2008) Exploring relationships between expressive and structural elements of music and pianists' gestures. Proceedings of the fourth Conference on Interdisciplinary Musicology, Thessaloniki, Greece.