International Symposium on Performance Science
ISBN 978-2-9601378-0-4

Rhythmic entrainment in communicative, dyadic improvisation

Tommi Himberg[1,2], Marc R. Thompson[2], and Satinder P. Gill[3]

[1] Brain Research Unit, O.V. Lounasmaa Laboratory, Aalto University, Finland
[2] Finnish Centre of Excellence in Interdisciplinary Music Research, Department of Music, University of Jyväskylä, Finland
[3] Centre for Music and Science, Faculty of Music, University of Cambridge, UK

In our everyday interactions with others, we rhythmically entrain with the movements of each other's bodies and voices, and this entrainment seems to share a quality with that of musical interaction. In order to understand this quality, we have taken the case of improvisation where both musical and linguistic interaction are considered as performance, and compare how we entrain when we jointly create music and stories together. Using a combination of qualitative and quantitative analysis of video and motion capture data, we have identified salient rhythmic moments (SRMs) that are heightened moments of rhythmic and empathic connection. A perceptual experiment has tested the effects of these SRMs on the perceived qualities of the performance.

Keywords: improvisation; movement; entrainment; communication; pragmatics

We automatically entrain gestures, expressions, and body movements with each other in social situations (Kendon 1970, Chartrand and Bargh 1999). Entrainment of body movements and body sway suggests that entrainment is a basic building block of interpersonal understanding and communication (Shockley *et al.* 2003, Shockley *et al.* 2009).

Music and dance as joint action foreground this rhythmic entrainment (Himberg and Thompson 2011). As with conversation, musical improvisation is based upon conventions and rules, and the dynamics of the interaction unfold in unpredictable ways as performers negotiate the shared musical time. As part of this process, there are moments of heightened rhythmic and empathic connection that we call *salient rhythmic moments* (SRM; Gill

2007). These mark pivotal points in the interaction. Here, we present preliminary results from a production task whereby participant dyads interacted with each other through verbal exchanges and improvisatory musical performances. A related perception task, in which observers reported the temporal locations of SRMs within the production task, is also detailed.

The data were collected using video recordings and motion capture. Motion capture provides an accurate, three-dimensional representation of movements, allowing patterns of interaction to be studied in detail (Thompson 2012). This complements the analysis of film footage, which has a rich complexity of information, thereby affording greater clarity for identifying SRMs. Our aims are to: (1) combine qualitative observational analysis and quantitative movement analysis methods to identify SRMs in dyadic musical improvisation, and explore their kinematic properties; (2) explore the roles of auditory and visual modalities in generating SRMs by comparing interaction in face-to-face and non-face-to-face conditions; and (3) investigate the effects of SRMs on perceived qualities of the performance.

METHOD

Participants

For the Interaction study, 9 pairs of participants were recruited from the Department of Music, University of Jyväskylä. With the exception of one pair, they consisted of friends. For the Perception study, 26 students were recruited from the University; 21 had prior dance training, 10 were semi-professional or professional musicians, and 7 had no musical training.

Materials

In the Interaction study, participants were asked to freely co-create stories or improvise using shakers in two-minute segments. For "inspiration," they were provided with a soundscape and a matching, slowly changing video of still pictures of natural places (e.g. forests, beaches) or urban places (e.g. street corners, restaurants). Optical motion capture, audio, and video recordings of the two minute performances were captured.

In the Perception study, participants viewed 12 video excerpts from the Interaction study, lasting 30 minutes. The excerpts were selected for the quantity of movement and observable interaction, contrasting the most and least movement. 30 second segments (within the 60- to 90-second period of each trial) were presented in random order within the Max software (www.cycling74.com).

Procedure

For the Interaction study participants changed into motion capture (mocap) suits with reflective markers on suits, hands, and feet. The experiment started with a practice session. Each pair performed eight trials, half facing and half non-facing. Participants were instructed to take turns in half of the trials, and in the others were asked to perform freely. The starts and ends of trials were signaled by a beep from the loudspeakers.

For the Perception study participants received a verbal definition of a Salient Rhythmic Moment and practiced using the Max interface. Excerpts were presented in a random order. Participants were permitted to re-watch an excerpt up to six additional times. They were instructed to click on a virtual button within the Max patch when they perceived a SRM. This recorded the temporal location of SRMs. After clicking, participants were asked to rate Cooperation, Synchronicity, and Ease of Interaction for each excerpt using a 1-7 Likert scale, and rate the level of difficulty of scoring.

RESULTS

Interaction study

The kinematic analysis of the SRMs is currently on-going. Windowed cross-correlation (WCC) was used to look at body entrainment at a general level. This analysis is good in unveiling the temporal evolution of entrainment, and possible lags between participants (Himberg and Thompson 2011).

For example, in Figure 1 the WCC of wrist movements in a free, facing, musical task shows the sensitivity of the WCC to changes in the pulse and lag between the participants. At the beginning of this interaction there was a sequence of tentative attempts (74 to 100 seconds) to engage, mostly initiated by the female participant, rather like a “game.” The WCC shows no clear pattern of entrainment. At around 100 seconds the female participant initiated a clear pulse (in the video she could be seen tapping her foot to this beat), and they locked in synchrony (100 to 112 seconds).

After this, the male participant continued the “game” from earlier on, deliberately missing his turns and shifting to a slower tempo. The WCC shows this as a changing lag (downwards pattern of high correlations) between 112 and 120 seconds. The trial ended with an accented hit, taken in turns.

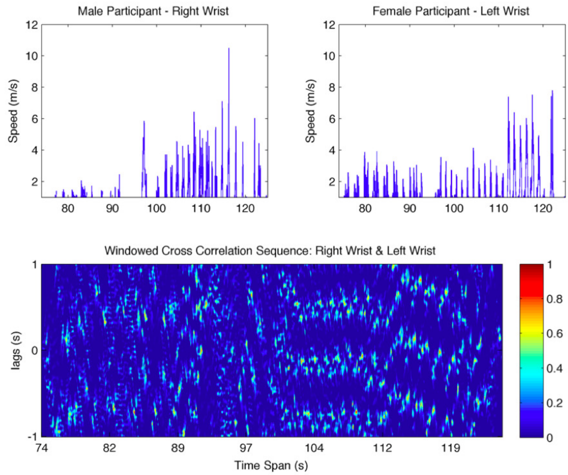

Figure 1. Windowed cross-correlation of hand movements in a music task. The top panels show velocities of wrist markers, and the bottom shows the WCC of these time-series. Parallel bands of high correlation reflect periodicity of rhythmic patterns. (See full color version at www.performancescience.org.)

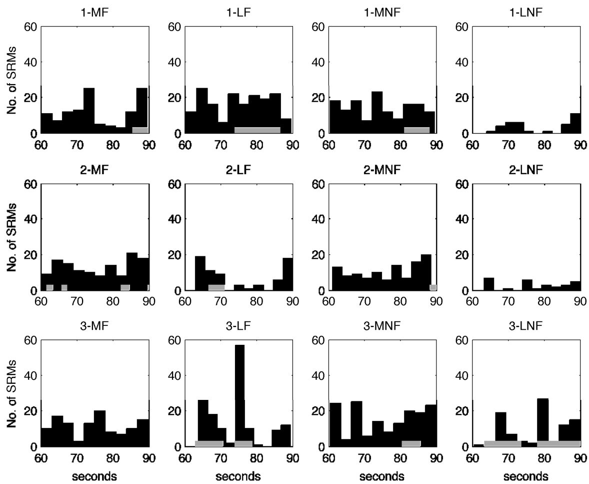

Figure 2. Analysis of SRMs reported. SRMs were perceived quite uniformly throughout each excerpt, with the exception of 3-LF, in which there was large agreement about one SRM. The LNF condition yielded the smallest number of perceived SRMs.

Perception study

The recorded temporal locations of perceived SRMs made using the Max patch have been plotted in histograms in Figure 2. The rows represent the three performer dyads observed by participants and the columns are the four performance conditions: music facing (MF), language facing (LF), music non-facing (MNF), and language non-facing (LNF). Each bar represents a time span of two seconds. In the bottom row, the plot labeled 3-LF shows a general consensus among the 25 participants that there was an SRM at around the 70s mark. The other excerpts produced a good number of perceived SRMs, which are more scattered. The story-telling, non-facing condition warranted the smallest number of SRMs (right-most column).

Based on post-experiment questionnaires, 48% of participants reported that they found it easier to perceive SRMs in the music trials than in the language trials; however, there were clear findings for the language non-facing (LNF) results, and differences between music facing (MF) and language facing (LF).

The grey bars at the bottom of each plot indicate the periods of heightened corporeal interaction within participant dyads, as coded by one of the authors of this study—an expert in interaction analysis. The figure thus provides a comparison between expert and novice SRM rating. For instance, in plot 1-LF (pair 1, language facing), the location between 73 and 86 seconds shows that the participants overall observed an extended period of interaction, which was supported by the expert rating. Likewise, in plots 1-LNF and 2-LNF the expert observer coded no SRMs, which was reflected in the participants' comparatively few SRMs for these trials.

DISCUSSION

Complementary qualitative and quantitative analyses were used to discover the dynamics and kinematics of interaction and mutual sense-making, evident in the occurrence of SRMs. WCC analysis revealed dynamics of interaction, sections of entrainment, and a shifting lag between performers. The expert analysis of SRMs and the novice participants' annotations showed good agreement. Originally, SRMs had been described after meticulous expert analysis, but our results suggest that they are perceivable for laypersons as well, supporting the theory that they serve an important communicative role in structuring and making sense of interactions, e.g. in gauging the level of mutual understanding.

The analysis is on-going. Taking advantage of the results of the quantitative analysis and the perceptual experiment, computational analysis of the

kinematics and movement coordination is on the way, with the aim of constructing a computational model of SRMs.

Acknowledgments

We thank our students Sarah Faber and Anna Fiveash for their help in collecting data for the perceptual study.

Address for correspondence

Tommi Himberg, Brain Research Unit, O.V. Lounasmaa Laboratory, Aalto University Puumiehenkuja 2 B, Otaniemi, Espoo 00076 Aalto, Finland; *Email:* tommi.himberg@aalto.fi

References

Chartrand T. and Bargh J. (1999). The chameleon effect: The perception-behavior link and social interactions. *Journal of personality and social psychology, 76,* pp. 893–910.

Gill S. P. (2007). Entrainment and musicality in the human system interface. *AI & Society, 21(4)*, pp. 567–605

Himberg T. and Thompson M. R. (2011). Learning and synchronising dance movements in south african songs – cross-cultural motion-capture study. *Dance Research, 29(2)*, pp. 305–328.

Kendon A. (1970). Movement coordination in social interaction: Some examples described. *Acta psychologica, 32,* pp. 101–125.

Shockley K., Santana M. -V., and Fowler C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology-Human Perception and Performance, 29(2),* pp. 326–332.

Shockley K., Richardson D. C., and Dale R. (2009). Conversation and coordinative structures. *Topics in Cognitive Science, 1,* pp. 305–319.

Thompson M. R. (2012). *The application of motion capture to embodied music cognition research.* Unpublished doctoral thesis, University of Jyväskylä, Finland.