Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM 2009) Jyväskylä, Finland
Jukka Louhivuori, Tuomas Eerola, Suvi Saarikallio, Tommi Himberg, Päivi-Sisko Eerola (Editors)

Group synchronization of coordinated movements in a cross-cultural choir workshop

Tommi Himberg, Marc Thompson

Finnish Centre of Excellence in Interdisciplinary Music Research, University of Jyväskylä, Finland
tommi.himberg@jyu.fi, marc.thompson@jyu.fi

ABSTRACT

Group synchronisation in choir performance was studied using motion capture. The choir consisted of South-African experts and Finnish novices. 4+4 participants' movements were recorded during a performance of a traditional song and choreography. The aim of the study was to: 1) describe possible differences in movement implementation between the two groups; 2) investigate synchronicity between all participants and within and between each of the two groups; 3) visualise the dynamics of interpersonal influence. Preliminary results indicate that simple indicators such as cumulative distance and position as a function of time reveal differences in implementation of choreography between the groups. Movements were very synchronous, but cross-correlations of the vertical velocities of markers revealed a subtle phase difference between individuals and shows that the expert group was more coherent than the novices, as was expected. This could be due to having more attentional resources to spend on coordination, as the musical and movement material is more familiar to the experts than the novices. Finally, the interpersonal interaction was visualised using windowed cross-correlation. This suggests that the temporal relationship of movements reflects the participants' mutual adaptation to each other's subtle timing perturbations, and that this adaptation is a dynamic process.

I. INTRODUCTION

Choir singing, as one of the most common musical activities in the world is popular across cultures, social groups and ages (Nettl 2000). It's a social activity that promotes social cohesion, and even has been shown to have health benefits (Beck et al. 1999, Kreutz et al. 2004). It can take many forms – from an impromptu congratulatory rendering of "Happy Birthday To You" to the professional operatic chorus or a performance of traditional tribal song.

In this study, group synchronisation in choir performance by South-African "experts" and Finnish "novices" [1] was studied, using motion capture. In a choir workshop, the experts taught South-African traditional songs to the novices. This style of choir music fits especially well for movement research, as the songs have a movement choreography in addition to the "music" and lyrics. In this paper, a preliminary analysis of the movement data is presented. Three aspects are analysed: differences in the execution of the choreography between the experts and novices, group synchronisation and cohesion, and interpersonal influence.

A. Background – music performance in ensembles

Singing in a choir requires numerous cognitive abilities. In addition to the basic abilities underlying auditory perception and vocal production, a host of skills for coordinating and synchronising with other choir members are needed. A singer needs to constantly monitor the pitch, timbre and timing of his or her voice in relation to the others' voices, and make necessary adjustments. Moreover, those adjustments need to be made in a co-operative manner (Keller & Repp 2008), taking into account that the others are also aware of the asynchronies or mistuning and that they will also take corrective action.

Adjusting the pitch of one's voice to fit in the mix of voices is a separate issue, and outside the scope of this paper. Our focus is on the temporal synchrony and coordination of note onsets and syllables. Instead of analysing the auditory information, we focus on the synchronisation and coordination of choreographed movements to music. This temporal synchrony is more complex than synchronising (usually finger tapping) with an isochronous pulse, which is the usual paradigm for studying timing and synchronisation (see Repp (2005) for a review). The added complexity comes first of all from the lack of absolute referent - instead of following a metronome, the choir, even when conducted, abstracts it's own pulse and maintains it in interaction. Secondly, the choir "aims for" culturally and stylistically acceptable expressive timing rather than a rigid isochrony. Both these factors bring with them challenges for analysis (Rasch 1988).

Information input is turned into action in music performance via a motor program (Davidson & Correia 2002, 238). Motor programs are developed through training and rehearsals, and their hierarchical nature makes it possible to perform a large amount of fast movements in a sequence, and be able to control them (their timing, velocity etc.) to achieve desired effects in terms of musical expression, for instance. Or, importantly for this study, adjust to their partners' performance in order to keep the performance "unified" (Davidson & Correia 2002, 244).

Motor programs are used to implement the musical ideas and intentions that the performers have. A social situation such as choir singing requires coordination of these intentions, and results in alignment of mental states (Tomasello & Carpenter 2007). Shared intentionality (sharing goals, dividing roles, monitoring progress towards these goals) can call for or lead to shared representations, such as having the same, correctly aligned metrical and tonal schemata of the music being performed. In this study, the experts of a particular musical style would have developed better-fitting schemata, through exposure and training, having learned the cultural conventions of that style. This shared knowledge would benefit their cooperation. Conversely the novices would have schemata for their own music, but these might not fit South-African traditional music. This potential lack of fit could have an effect on how the movements were performed, as well as the accuracy of their timing.

These schemata modulate our expectations, which in turn direct our limited attentional resources to where they are likely to be needed. According to Jones (e.g. Large & Jones

[1] Expert and novice refer to the groups' familiarity with the musical material used.

1999), attention is dynamic and oscillates, entrained to the pulse of salient, periodic events that are being attended. Attentional resources also need to be spread between monitoring one's own actions and those of the others (Keller 1999, Sawyer 2005). The assumption is that novices, being less familiar with the material than the experts, will need more of those limited resources to monitor their own actions, therefore leading to a weaker inter-group coordination. The experts, armed with their superior knowledge of the musical material can dispense more attention for keeping together and expressive features.

B. Background – movement and motion capture in music research

In musical performance, body movements are usually needed for manipulation of the musical instrument and production of musical sounds, or in the case of singing, production of the singing voice. In addition to the movements necessary for producing the correct sequence of sounds, musicians use their bodies to communicate expressive ideas about the music (see e.g. Davidson & Correia 2002). The musical context in the current study is South African traditional choir performance, in which these expressive movements are emphasised and structured to form a movement choreography, a rehearsed dance that the whole choir performs while singing. This genre of choral music therefore is an ideal context to study movement synchrony in a group. The movements are pronounced, making them easy to track; in a choir everyone uses a similar "instrument", meaning that the movements needed to produce sounds are the same for all performers (cf. string quartet and the different movements needed to play the 'cello and the violin); as the genre of music is not familiar to people at the location of the research lab, it is easy to find a novice group to be compared with the experts, to get a glimpse of how learning those movements happens.

Prior research in ensemble playing has largely been phenomenological in nature, been based on interviews or for instance observations of practice sessions (Davidson & Good 2002; King 2006) or analysis of the verbal exchanges during practice (Ginsborg et al. 2006).

In recent years, researchers interested in music-related body movement have opted to analyze movement data derived from optical motion capture systems as opposed to traditional video. Although video can been analyzed computationally (Jensenius, Godøy & Wanderley, 2005; Castellano, Mortillaro, Cammurri, Volpe & Scherer, 2008), motion capture data can be analysed in three dimensions of movement and with a higher spatial and temporal precision (Wanderley, Vines, Middleton, McKay & Hatch, 2005; Eerola, Luck & Toiviainen, 2006; Luck &Toiviainen, 2006).

The use of such data allows one to quantify relationships between performers' movements. Motion capture data allows data to be analysed as a time series, from which kinematic features such as acceleration can be extracted. Through cross correlation analysis, one can investigate influence and synchronization between performances and body movement.

C. Aims

The aim is to study synchronisation of movements in groups of choir singers using motion capture. More specifically, we are interested in how well the movements of the whole group are synchronised, and if there is a difference between the novice and expert groups. Also, we are interested to see whether experts influence the novices, and how the movement synchrony evolves over time.

II. METHODS

A. Participants and materials

Emmanuel Lutheran Church Choir from Johannesburg, South Africa, while on a concert tour in Finland, agreed to facilitate an African choir workshop at the Department of Music in Jyväskylä. The Choir performs traditional African music, originating mainly from the various southern African cultures. Coordinated movement choreography is part and parcel of these songs, therefore making them especially suitable for a music and movement study.

Finnish participants were recruited from among the extended community of the Department of Music. They all had musical experience, some had experience of African music or general knowledge of the musical style, but none were familiar with the songs performed in the workshop.

The workshop was a two-day event, and consisted of half a dozen African songs and one Finnish song being learned and performed by the participants. The performance analysed for this paper was of a Sotho song called *Fiela Ngwanana*, or *Sweeping Girl*. The workshop was conducted in English.

B. Apparatus and procedure

Video and audio recordings were made. Movement data from selected sessions of the workshop was captured using an 8-camera Qualisys Motion Capture system, which recorded the trajectory of reflective markers attached to participants at a rate of 120 fps. For this particular task, 8 participants (4 expert, 4 novice) wore markers (head + feet) and were positioned at the centre of the room in a 2 by 4 –formation (Figure 1.), while the rest of the workshop participants also contributed to the music making but stayed outside the capture volume.

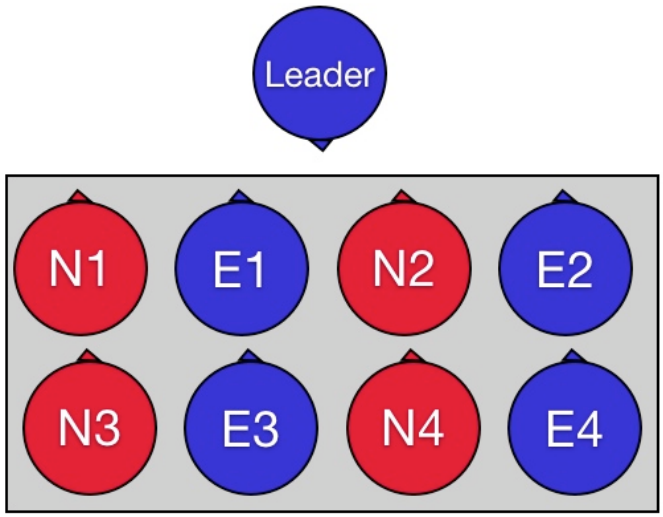

Figure 1. Participant positions. E1-4 refer to the experts, and N1-4 to the novices.

The leaders of the choir, with assistance of the other expert members, taught the novices how to perform a traditional Sotho song *Fiela Ngwanana*. Lyrics were written on a flip board, and the music was taught voice by voice and segment by segment aurally. The choreography of the movements was also taught so that the leader and members of ELC showed the movement first, and then everyone did them together. After a few practices, a performance of the song was recorded using the motion capture system. Due to technical constraints, the recording was made in three consequent measurements.

C. Analysis

The movement data acquired from the motion capture system was imported into MATLAB as time series data representing the location of each marker on the three dimensions of the Cartesian coordinate system. The computation of movement data was performed using the MoCap Toolbox (Toiviainen, 2008). First, the cumulative distance travelled by each marker was calculated to see the total amount of movement performed by each member of both groups. Our analysis then moved to the corporal relationships between and within the two groups (novice and expert) through windowed cross correlations. The plots created from the windowed cross correlation reveal the progression of influence between the expert and novice performers. From this perspective, one can gain an objective idea of the length to which the novices depends on the experts to stay synchronised.

III. RESULTS

A. Amount of movement

Figure 1 shows the cumulative distance travelled for each marker, averaged across the novice and expert groups. These results show that in terms of total amount of movement, there were no large differences between groups and that all participants implemented the choreography in a similar fashion. Between the head and feet, there is the least agreement in the total amount of movement for the head. This is a result of the head not having any set choreography in the performance.

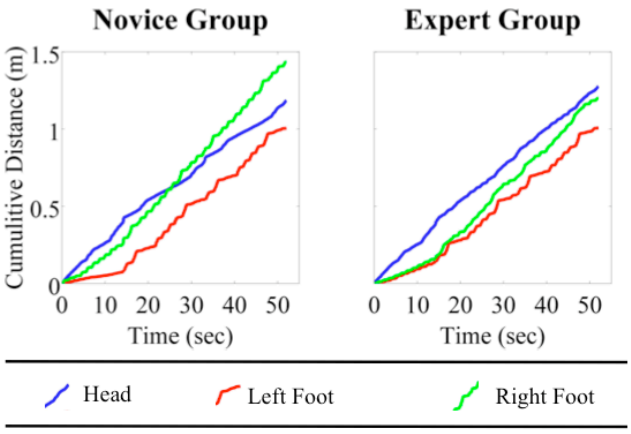

Figure 2. Cumulative distances travelled by markers, group averages of the whole song.

The right feet cover more distance than left feet. This is due to the movement choreography, which has two sections where left feet are relatively stationary while participants take steps forward with their right foot on every other beat. This movement is accompanied with a sweeping motion of the hand, as the song is about a girl sweeping a path. There is a slight difference between the groups in when that difference emerges. The novices implement the choreography by taking those steps forward with their right feet, while the expert group moves both feet, perhaps reflecting their focus on maintaining the feeling of the beat rather than over-emphasising the already salient feature of the choreography.

Figure 3 depicts a subtle difference between the expert participant number 1 and the novice participant number 1 (E1 and N1, see Figure 1). Looking at a snapshot of their performance, 5 seconds of march-like stepping from the middle of the B-section, lends further evidence to the hypothesis that novices focus on executing the "correct" choreography, taking the right steps at the right time, whereas experts are more able to express the "flow" of music.

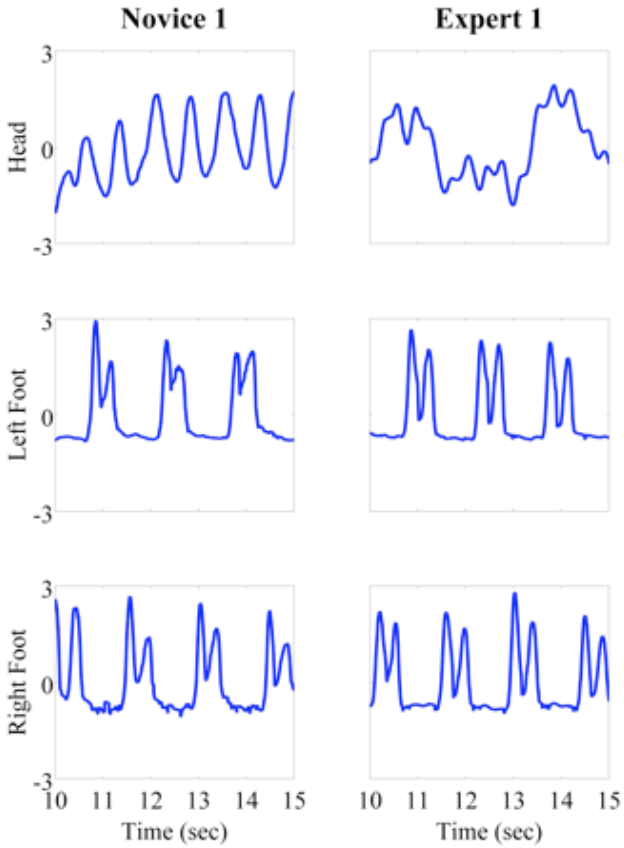

Figure 3. Head and foot acceleration – comparison of E1 and N1 in the "marching" section

For N1, the head movements are clearly periodic at the beat level. On every step made (a step is marked by a double-peaked rise in position of the foot markers) the head moves up and down. The expert seems to have a periodicity at a higher harmonic level (at bar level) in addition to this beat-level movement. Furthermore, this beat-level movement

is much smaller in amplitude than that of N1. This could reflect E1's better ability to focus on higher levels of the metrical hierarchy, possibly a reflection of knowing the structure of the song better, and having expectations of the meter that span further ahead than what N1 has. One way of characterising the N1 movement here compared to E1 is that N1 is more "stiff", as movements of her feet are directly reflected in head movements, whereas E1's body absorbs that energy, resulting in a smoother movement of her head. This issue should of course be studied in detail using a full-body marker set.

B. Synchronisation

Movement synchrony between participants was examined using cross-correlations of acceleration along the vertical dimension of the feet markers. The components of acceleration were estimated using he Savitzky-Golay smoothing FIR filter (Braci & Diop, 2003) with a window length of seven samples and a polynomial order of two. From the peaks in the resulting acceleration time-series provided the onset times for each step taken by the performers. Therefore, we expected correlations to indicate the amount of synchronization between performers. The whole length of the performance (three captures) was used in this analysis.

Normalized cross-correlations with a maximum lag value of +/-60 frames (equivalent to +/-0.5 seconds) were calculated between each of the four participants in the Expert Group to the four members of the Novice Group, resulting in a sequence of 121 low to medium coefficient values. Figure 4 depicts the maximum coefficient values (positive or negative) extracted from the sequences and plotted across the lag values. From Figure 4, it can be deduced that Novices 1, 2 and 4 have their maximum correlation at the zero point, meaning they were able to synchronize to the Expert Group.

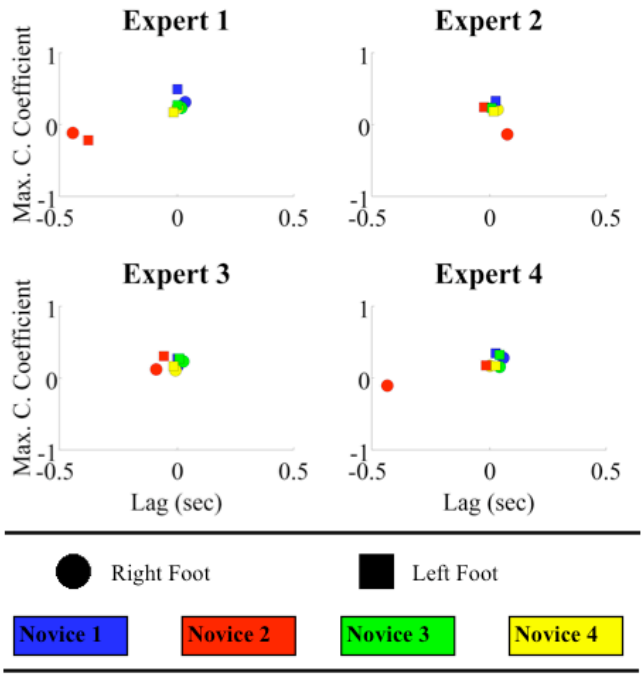

Figure 4. Maximum cross-correlations between experts and novices. Vertical acceleration of feet, the whole song.

Most maxima are either at or close to zero lag, indicating simultaneity – the participants are performing their movements at the same time. Some pairs reach relatively high correlations, which in turn indicates that they perform the movements in a very similar way. There are some outliers (red ball and square, representing the right and left feet of N2, respectively). The cross-correlation curve between N2 and E1 or E4 was fairly flat, and the maxima are not as pronounced as for the other pairs. In these outlier cases, there were local maxima at lag 0, and these were just slightly lower than the largest ones depicted in Figure 4. Therefore one should not over-stress the importance of these outliers.

Turning to between-group differences, Figure 5 shows cross-correlation maxima for the Expert and Novice groups separately. The experts are a tighter group, both in timing and in movement execution, whereas the novices have a wider range of movement relationships in their group.

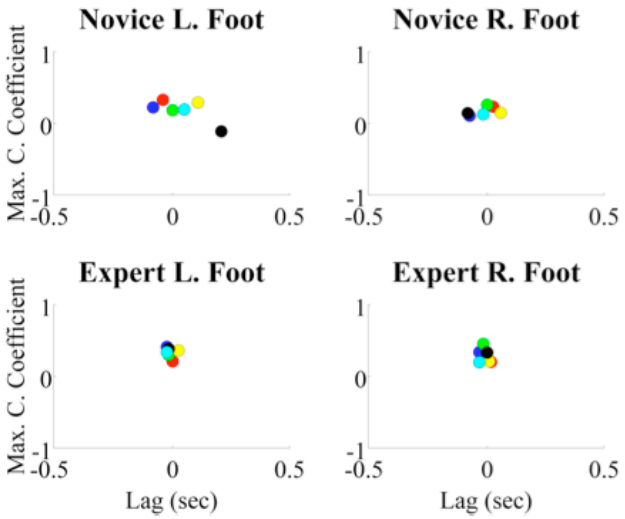

Figure 5. Maximum cross-correlations for Expert and Novice groups, the whole song. The colours represent the six possible correlations within each group.

C. Temporal dynamics of synchrony – inter-individual influence

To investigate how the dependencies between participants developed over time, the cross-correlation of vertical acceleration was calculated in a moving window. The cross-correlograms for E1 with N1 and N2 are presented in Figure 6. The graphs cover the same five-second "marching" section that is depicted in Figure 2.

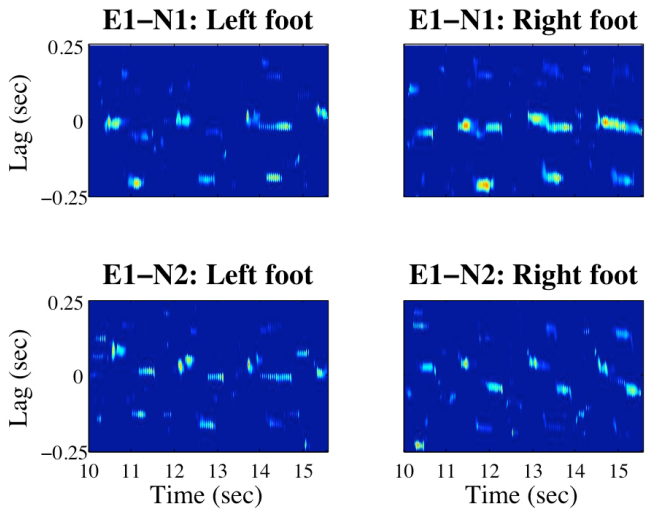

Figure 6. Cross-correlograms, E1 vs. N1 and N2 in the "marching" section. "Warmer" colours indicate higher correlations.

The cross-correlations at zero lag represent periods of rest between steps, when the foot is on the ground and the other foot is moving. It is interesting to notice that there is another row of high cross-correlations at negative lags, at times of the steps. This indicates that especially N1 has clearly been a bit late in their steps, compared to E1. This difference is about 20 frames, or 0.1 seconds. The pattern of E1-N2 cross-correlations is much less clear. There seems to be a repeating pattern in the temporal relationship of their movements, but correlations are lower than in the other pair. This was also shown in the analysis of overall cross-correlations in Figure 4.

Finally, a windowed cross-correlation analysis was performed for a longer section, the B-section of the song (the second capture). The beginning part of the second capture was left out of this analysis, as there was very little movement there. In addition to the cross-correlogram, Figure 7 has the sums of correlation values at each lag on the right panel. This shows a peak at zero, indicating synchrony, but also peaks at lags between 0.1 and 0.2 seconds, both negative and positive, indicating that both leading and lagging behind occur during the trial. This could be interpreted as an indicator of interpersonal influence, as participants were adjusting their movement timings based on previous movements.

IV. CONCLUSIONS

Group synchronisation in a choir performance of a southern African traditional song that had a movement choreography was studied using motion capture. The choir consisted of Finnish novices who had learned the song earlier that day by aural instruction, and of South-African experts who had this song as a number in their performing repertoire.

Data analysis indicates that the novices are more focused on taking the correct steps at the right times, whereas experts show a more fluent execution of the choreography. This was indicated for instance in the independence of their head movements, and a less pronounced difference in the distances travelled by each foot.

Synchronicity was analysed using cross-correlations. This analysis shows that the experts are closer together as a group than were the novices. This could be an indication of the expert group having more attentional resources to spare for interpersonal coordination, whereas the novices need to focus on remembering the correct sequences of movements. The experts also have stronger mental representations of the melody and lyrics and better rehearsed motor programs for the movements.

The temporal evolution of synchronisation can be investigated using windowed cross-correlation. In the example presented here, it shows that the novices lag a bit behind in their steps. Zooming out and looking at a longer section, it is possible to have a glance of the dynamics of interpersonal influence and mutual adaptation.

The analysis presented here is preliminary, and descriptive in nature. Only one performance of many that were recorded was analysed. Therefore these results are far from conclusive, but rather this study should be seen as a pilot that is flagging interesting aspects of movement synchrony in multi-cultural choir performance for future research.

In our view, these interesting aspects include:

- movement synchrony is continuously evolving during performance, as participants adapt to each other's movements. This is indicated by fluctuations of cross-correlation maxima in the windowed analysis.
- experts' movements are more synchronous and similar with each other than those of the Novice Group, indicated by smaller lags and higher cross-correlations.
- novices might be more focused on steps while experts who already know the choreography can interpret the movements more freely.

In summary, motion capture is a good tool for studying movement synchrony in music performance in an ecologically valid setting. The data discussed in this paper was collected from just three markers per participant. Therefore there is a lot of room for development and more complex data could be obtained simply by adding markers to hands, knees etc. This setup was chosen because the focus was in very basic synchronisation processes that we assumed were reflected in feet and head movements. Other kinds of choreographies and other research questions would of course require different setups, and cross-correlation analysis provides one way to study synchronicity and mutual influence within the group.

In the future, this methodology could be applied for other choirs and contexts. As mentioned above, a more complete marker set would allow a more detailed analysis of the quality of movements. Combined with quasi-experimental research designs, theories of embodied and social cognition could be put to a test. Analysis could be extended by for example subjecting the cross-correlation results to further numerical analysis. One possibility would be to look at the rate of change of the maximum lag as an indicator of stability of the relationship between participants.

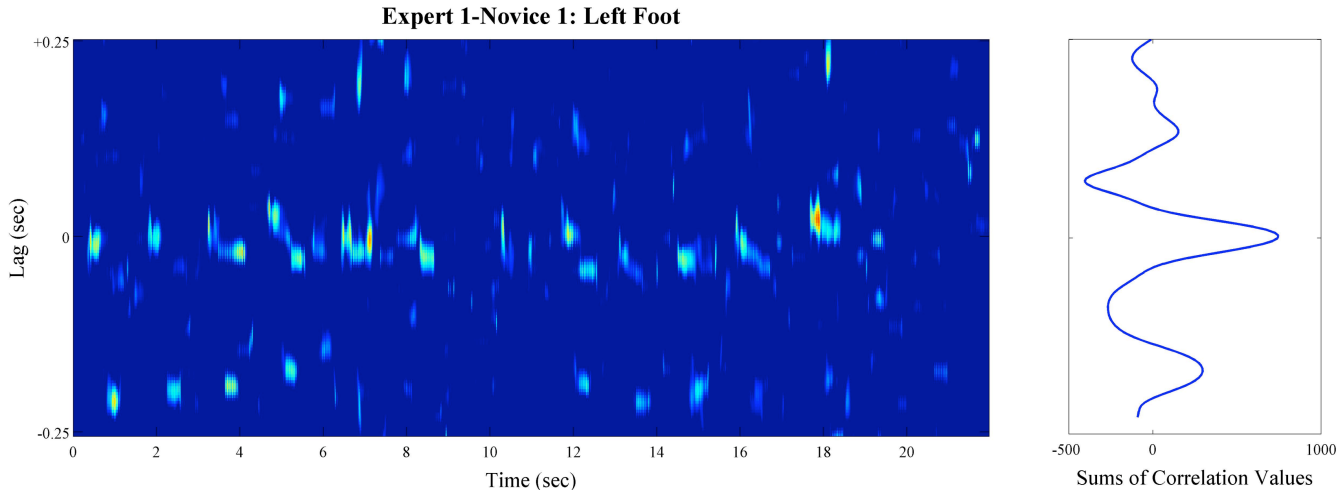

Figure 7. Cross-correlogram and sums of correlation values for each lag, Left feet of E1 & N1, mid-section of the song. "Warmer" colours indicate higher correlations.

ACKNOWLEDGMENTS

The authors would like to thank all the participants for their kind help and prof. Jukka Louhivuori for organising the workshop and making this study possible, and prof. Petri Toiviainen, Dr. Geoff Luck and Dr. Olivier Lartillot for their invaluable advise and help in analysing the data.

REFERENCES

Beck, R. J.; Cesario, T. C.; Yousefi, A. & Enamoto, H. (1999). Choral singing, performance perception, and immune system changes in salivary immunoglobulin A and cortisol. *Music Perception, 18*, 87–106.

Braci, M., & Diop, S. (2003). On numerical differentiation algorithms for nonlinear estimation. Paper presented at the IEEE International Conference on Decision and Control, Maui.

Ginsborg, J., Chaffin, R. & Nicholson, G. (2006a). Shared performance cues in singing and conducting: A content analysis of talk during practice. Psychology of Music, 34 (2), 167-192.

Castellano, G., Mortillaro, M., Camurri, A., Volpe, G. & Scherer, K. (2008). Automated analysis of body movement in emotionally expressive piano performances. *Music Perception 26*, 2, 103-120.

Davidson, J.W. & Good, J.M.M, (2002). Social and musical co-ordination between members of a strings quartet: an exploratory study. *Psychology of Music, 30*, 186-201

Davidson, J. W. & Correia, J. S. (2002). Body Movement. In Parncutt, R. & McPherson, G. (eds.). The Science and Psychology of Music Performance. New York, NY: Oxford University Press, 237-252.

Eerola, T., Luck, G., & Toiviainen, P. (2006). An investigation of pre-schoolers' corporeal synchronization with music. In M. Baroni, A. R. Addessi, R. Caterina & M. Costa (Eds.) Proceedings of the *9th International Conference on Music Perception & Cognition* (pp. 472-476). Bologna, Italy: ICMPC and ESCOM.

Goodman, E. (2003). Ensemble performance. In Rink, J. (ed.). Musical Performance - A Guide to Understanding. Cambridge: Cambridge University Press. 153-167.

Himberg, T. (2006). Co-operative Tapping and Collective Time-keeping - differences of timing accuracy in duet performance with human or computer partner. In Baroni, M.; Addessi, A. R.; Caterina, R. & Costa, M. (ed.) Proceedings of the 9th International Conference for Music Perception and Cognition ICMPC 9, 377.

Jensenius, A.R., Godøy, R.I. & Wanderley, M.M., (2005). Developing tools for studying musical gestures within the Max/MSP/Jitter environment. In *Proceedings of the International Music Computer Conference*. Barcelona: ICMC.

Keller, P. E. (1999). Attending in Complex Musical Interactions: The adaptive dual role of meter. Australian Journal of Psychology, 51, 66-175.

King, E. (2006). The Roles of Student Musicians in Quartet Rehearsals. Psychology of Music, 34(2), 263-283.

Kreutz, G.; Bongard, S.; Rohrmann, S.; Hodapp, V. & Grebe, D. (2004). Effects of Choir Singing or Listening on Secretory Immunoglobulin A, Cortisol, and Emotional State. Journal of Behavioral Medicine, 27(6), 623-635.

Large, E. W. & Jones, M. R. (1999). The Dynamics of Attending: How people track time-varying events. Psychological Review, 106, 119-159

Lartillot, O. & Toiviainen, P. (2007). MIR in Matlab (II): A Toolbox for Musical Feature Extraction From Audio. International Conference on Music Information Retrieval, Vienna, 2007.

Luck, G., Toiviainen, P., (2006). Ensemble Musicians' Synchronization with Conductor's Gestures: An Automated Feature-Extraction Analysis, *Music Perception, 24*, 189-199.

Nettl, B. (2001). An Ethnomusicologist Contemplates Universals in Musical Sound and Musical Culture. In Wallin, N. L.; Merker, B. & Brown, S. *(eds.) The Origins Of Music*, Cambridge, MA: The MIT Press, 463-472.

Rasch, R. (1988). Timing and Synchronization in Ensemble Performance. In Sloboda, J. (ed.). The Generative Processes in Music, University Press, 70-90.

Repp, B. H. (2005). Sensorimotor Synchronization: A Review of the tapping literature. Psychonomic Bulletin & Review, 12, 969-992.

Repp, B. H., & Keller, P. E. (2008). Sensorimotor synchronization with adaptively timed sequences. Human Movement Science, 27(3), 423-456.

Sawyer, R. K. (2005). Musical Conversation. In Miell, D. MacDonald, R. & Hargreaves, D. (eds.). Musical Communication. Oxford: OUP, 45-60.

Tomasello, M. & Carpenter, M. (2007). Shared intentionality. Developmental Science, 10, 121-125.

Toiviainen, P. (2008). *MoCap Toolbox Manual*. University of Jyväskylä. Jyväskylä, Finland. Available at http://www.jyu.fi/music/coe/materials/mocaptoolbox/MCTmanual

Toiviainen, P. & Eerola, T. (2003). Where is the Beat?: Comparison of Finnish and South African Listeners. In Kopiez, R.; Lehmann, A. C.; Wolther, I. & Wolf, C. (ed.) Proceedings of the 5th Triennial ESCOM Conference, Hanover University of Music and Drama, 501-504.

Wanderley, M. M., Vines, B., 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.