Movement Sonification of Musical Gestures: Investigating Perceptual Processes Underlying Musical Performance Movements

Jesper Hohagen
jesper.hohagen@uni-hamburg.de

Clemens Wöllner
clemens.woellner@uni-hamburg.de

Institute of Systematic Musicology, University of Hamburg

ABSTRACT

Truslit (1938) developed a theory on the gestural quality of musical interpretations. Self-other judgment paradigms of visual point-light movements allow elucidating action-perception coupling processes underlying musical performance movements as described by Truslit. Employing movement sonification with a continuous parameter mapping approach may further show parallels between the audio information of music, physical movements, and audio information based on sonified movement parameters. The present study investigates Truslit's hypothesis of prototypical musical gestures by comparing free movements and movements following detailed instructions recorded by a 12-camera optical motion capture system. The effects of watching these movements and listening to the sonification were tested within a multimodal self-other recognition task. A total of 26 right-handed participants were tracked with a motion capture system while executing arm movements along with Truslit's (1938) original musical examples. The second experimental part consisted of a multimodal self-other perception judgment paradigm, presenting sequences to the same participants (matched with those of two other participants, unbeknown to them) under four different conditions. Signal detection analyses of the self-other recognition task addressed judgment sensitivity by calculating for individual participants. While self-recognition was successful for visual, audiovisual and still image examples, movement sonification did not provide sufficient detail on performer's agency. Nevertheless, a number of relevant sonification parameters is discussed.

1. BACKGROUND

Alexander Truslit proposed a theory of gestalt and movement that highlights the gestural quality of musical interpretations [1]. In his work he attempted to investigate the connection between the inner shape and motion of music and the perceptual processes of listeners while executing movements along with the music. Thereby, Truslit posed questions of the prototypicality of musical gestures. Since Bruno Repp's [2] synopsis of Truslit's work, researchers have increasingly developed methods and paradigms to study listeners' responses to musical gestures. Truslit himself presented movement graphs to selected musical pieces, based on dynamic and agogic information, assuming that these motion trajectories are valid intersubjectively. In contrast to Becking [3], who assumed that there are distinct and stable motor pulses in the works of different composers, Truslit hypothesized that listeners can learn to *feel* and then reproduce the shape and motion lying inside the music. However, there are still open questions concerning the common factors of musical and movement parameters, for example on how individuals perceive musical motion while they perform movements related to the music. Further empirical examinations are required to test this theory of a prototypicality of musical movements on a descriptive and comparative level.

In order to reach insights into the perceptual processes of music listeners and performers, asking for self-other judgments of visual point-light movements appears to be a promising method [4, 5]. Beside such a perception paradigm allowing the study of action-perception coupling [6], movement sonification may provide listeners with intuitive feedback on musical movement features. So far, movement sonification has mainly been applied in artistic performances [7] or sport and rehabilitation science [8, 9]. To our knowledge, our study is the first to systematically investigate the motion of Truslit's gestures with sonification of gestures and a self-other perception paradigm.

2. AIMS AND HYPOTHESES

In a first study (Table 1), we investigated Truslit's hypothesis of prototypical musical gestures by comparing free movements to Truslit's original sound examples with movements following a visual presentation and detailed verbal instructions. The second study tested the effects of watching point-light displays and listening to the sonification of movements with a multimodal self-other judgment paradigm.

Along with various analogous experimental tasks, we expect differences in expression of the movements before and after instruction. Moreover, we assume that this variation is higher in musically experienced participants in comparison with non-musicians, according to the formers' capacity to receive and process musical material and transfer it to appropriate movements.

Referring to results of established self-recognition tasks, we assumed that self-identification of visual displayed movements would be above chance. We expected higher scores of self-recognition for the unconstrained movements compared to post-instruction movements and a better performance in judging self and others movements by musicians in comparison with non-musicians.

Study	Conditions	Stimuli
Study 1 *(Performance task)*	Block 1: free condition	1. Wagner – *Gebet der Elisabeth* 2. Verdi – *Celeste Aida*
	Block 2: instruction condition	*Broken Chord C major, staccato* – bassoon
	Block 3: after instruction condition	Same as in 1st block
Study 2 *(Perception task)*	1. Visual (v)	Animation of visual point-light displays
	2. Auditory (a)	Movement data sonification
	3. Still image (si)	Still Image of movement trajectory graph
	4. Audio-visual (av)	Animation of visual point-light displays & sonification

Table 1. Overall research design

3. STUDY 1

3.1 Participants

In Study 1, a total of 26 right-handed participants (age: M=27.35, SD=4.06; 30,8% female, 13 musicians) were recruited to take part in the performing sessions.

3.2 Materials

23s cuttings of three selected original Truslit recordings (1. *Gebet der Elisabeth* – Richard Wagner, 2. *Celeste Aida* – Giuseppe Verdi, 3. *Mondnacht* – Robert Schumann) were presented during the first and the third blocks of the recording sessions. In addition, another 7s-short original Truslit piece (*Broken Chord C major, staccato* – bassoon) was played within the instruction part during the second block. While listening to the *broken chord* piece, an original Truslit drawing of a set definition movement trajectory was presented on a screen (see Fig. 1).

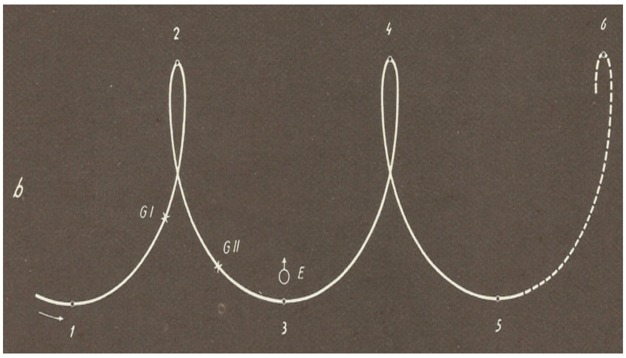

Figure 1. Original Truslit movement graph of *Broken Chord C major, staccato* – bassoon, used for visual instruction

3.3 Design and procedure

Performing sessions in Study 1 were divided in two main blocks, each with three different movement trials linked to three different musical stimuli. In between these two recording blocks, we integrated a shorter instruction block. We had a 2x2 repeated-measured design with two within-participant factors of instruction (Block 1: free movements along with three different musical stimuli; Block 3: after instruction, movements following the same musical stimuli as in the first block) and musical excerpt (Wagner; Verdi).

Participants' movements were tracked by a 12-camera 3D OptiTrack® motion capture system. For the recording sessions of the conducting-like gestures such as in Truslit (1938), a single marker was placed around the index finger of participants' right hand.

In each trial of the first block (free condition) participants received the instruction to "follow with your right arm freely the melody of the song you listen to". After a practice trial, participants performed three trials along with three original Truslit songs. Each time, participants first listened to the song before moving their arm to the melody of the song in order to become familiar with the style of the musical pieces. The three songs were presented randomly across participants to control order effects.

As part of the Block 2, participants first listened to 7s pieces of a broken chord sequence (2x) before they following the melody (2x) as explained. Second, participants listened to the same piece three times consecutively while looking at the appropriate Truslit trajectory graph (see Fig. 1) on a screen. The third part of this second session was a replication of the first one, but with the verbal instruction beforehand to "follow the melody with the index finger right in the way you saw it on the screen".

The third block (after instruction condition) was a replication of the first block without any additional instructions. Again, the musical pieces were played in randomized order.

3.4 Results

In order to assess differences between Block 1 (free) and Block 3 (after instruction) as well as between the two musical excerpts, we analyzed the averaged global measures of their index finger movement lines in terms of movement velocity, acceleration, jerk and cumulative distance (for descriptive analysis of movement parameters see Table 2).

Movement parameters	Conditions	Averaged global measures
Velocity	Free - Wagner	M=0.023m/s
	After inst. - Wagner	M=0.017m/s
	Free - Verdi	M=0.021m/s
	After instr. - Verdi	M=0.020m/s
Acceleration	Free - Wagner	M=-0.005m/s²
	After inst. - Wagner	M=-0.005m/s²
	Free - Verdi	M=-0.006m/s²
	After instr. - Verdi	M=-0.010m/s²
Jerk	Free - Wagner	M=0.033m/s³
	After inst. - Wagner	M=0.003m/s³
	Free - Verdi	M=-0.048m/s³
	After instr. - Verdi	M=-0.076m/s³
Cumulative distance	Free - Wagner	M=1.995m
	After inst. - Wagner	M=1.959m
	Free - Verdi	M=2.815m
	After instr. - Verdi	M=2.820m

Table 2. Averaged global measures of movement parameters

The Schumann song was not taken into consideration for statistical data analysis due to large gaps in some motion capture data.

A 2x2 repeated-measures ANOVA indicated significant differences between the velocity values of the Blocks 1 and 3 regarding the factor Instruction (F{1, 25}=5.40, p=.029, η^2=.177), indicating that participants moved more quickly in the free condition compared to the post-instruction block.

When comparing jerk values, we found a significant difference between the Wagner and Verdi piece, (F{1, 25}=8.56, p=.007, η^2=.255). The fact that participants jittered more during the Wagner song can be explained my musical parameters, i.e. the melody line of the Wagner piece is less active, so participants may have struggled to hold the line with their finger and started to shake.

Furthermore, there are highly significant differences in cumulative distance values between Verdi and Wagner movements (F{1, 25}=20.17, p<.001, η^2=.447). As observed for the results above, this difference shows again the effect of musical features based on the more active melody line of the Verdi piece compared to Wagner, leading to a higher distance travelled for Verdi movements.

No significant interaction effects between the factors musical excerpt and instruction were found. Furthermore, participants were relatively consistent in their movement styles irrespective of the Truslit-based instructions. In a similarly vein, analysis of covariance indicated no effects of musical experience or preferences ratings for the music on movement characteristics.

4. STUDY 2

4.1 Participants

In Study 2, 23 (age: M=27.43; SD=4.29; 30.4% female, 11 musicians) out of the 26 participants from the first study took part in the perception task experiment.

4.2 Materials

The multimodal self-other recognition task consisted of four different display conditions, based on the movement trajectory data recorded within the performing trials of Study 1.

4.2.1 Animated visual point-light displays (v)

For preparing the first condition of the self-other perception paradigm, we used the data of movement trajectories of the index finger marker from the performing sessions of the first study. For each participant, we created four 2D 10s normalized video clips with point-light displays of the finger marker (end points of data on both axes were standardized, in order to avoid recognition skills based on maximum movement amplitude) from the original 23s excerpts of Study 1 (1st clip: free – Wagner; 2nd clip: free – Verdi; 3nd clip: post-instr. – Wagner; 4th clip: post-instr. – Verdi). Video animations were created by using the MoCap-Toolbox [10] for Matlab®.

4.2.2 Sonifications of movement data (a)

The second condition for the perception task contained again four 10s clips per participant, but this time data of movement trajectories was matched with a continuous auditory feedback. For preparing these sonification sequences the continuous parameter mapping method [11] was applied, i.e. vertical position data (y-axis) of finger movements were matched with the pitch of a continuous synthesizer and horizontal position data (x-axis) were matched with stereo panning of the same sound. With this sonification mapping participants just “heared” their own movements captured in the first study, and during the presentation of the sonification clips the screen was black. Sonification sequences were programed by using Python® and the audio synthesis programming language SuperCollider.

4.2.3 Still Images of movement trajectories (si)

In the third perception condition, the same animation process as in the first condition was applied, but we created four 10s still images of the movement trajectories for each participant, comparable with the movement graphs of Truslit. For an exemplary trajectory graph of one participant see Fig. 2 (free – Wagner) and Fig. 3 (after inst. – Wagner).

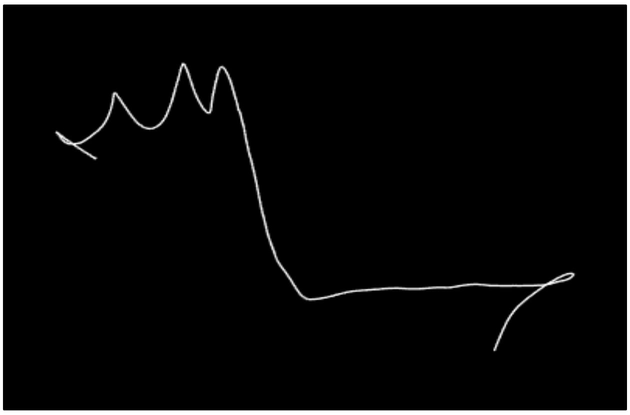

Figure 2. Exemplary trajectory graph for free Wagner movements (10s) of index finger (Block 1)

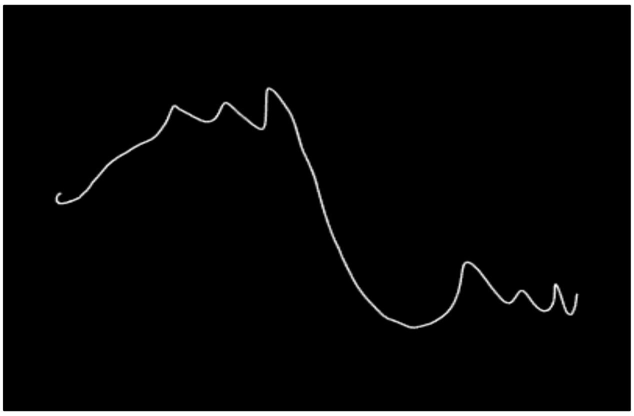

Figure 3. Exemplary trajectory graph for post-instruction Wagner movements (10s) of index finger (Block 3)

4.2.4 Visual point-light and auditory displays (av)

The fourth condition was a combination of the first (v) and second (a) perception condition, i.e. participants watched the visual point-light displays on a screen while listening to the sonification of the same movement data at the same time – again 4x 10s clips.

4.3 Design and procedure

9 months after the recording sessions of Study 1, 23 out of the 26 participants took part on a perception task experiment. We choose this long interlude time to avoid movement memory effects, which means that participants would have recognized their own movements based on their memory skills.

Within the self-other recognition paradigm, the multimodal movement displays of one participant were matched with sequences of two other participants by height and sex [12]. This method is useful to avoid recognition effects based on body information that can be interpreted from the video clips. Overall, each participant watched and listened to four 10s sequences of oneself and four sequences of the two other participants unbeknown to them. Every clip was presented twice across the four conditions (12 clips x 4 conditions x 2 presentations of each clip = 96 clips, divided in four condition blocks), thus, the total test time was around 60 minutes.

After watching or listening the 10s sequences, participants judged whether they had perceived their own movements or those of someone else. In addition, they answered how sure they were in their judgments and how expressive as well as how fluent the movements appeared to them (on 7-point Likert Scales from 1 - "very secure/expressive/fluent" to 7 - "not at all").

Participants were not informed whether the sequences displayed free or post-instructed respectively Wagner or Verdi movements. The two original Truslit excerpts were not played while the participants watched the point-light displays of their movement performances.

4.4 Results

Analyses of the self-other recognition task addressed judgment sensitivity by calculating d-prime (d') scores for individual participants, i.e. we substract z-transformated false alarm rates (participants incorrectly assume that they perceive the displayed movement as their own) from hit rates (correct self-recognition). We assumed that self-identification of visual and auditory displayed movements is above chance in all conditions. One-sample t-tests revealed that self-recognition was successful in three conditions: v ($t_{22}=2.21$, $p<.05$); si ($t_{22}=2.45$, $p<.05$) and av ($t_{22}=2.46$, $p<.05$). These results show the ability to recognize one's own movements even in a perception task in which body information was strongly reduced, indicating that kinematic information sufficed for participants' recognition accuracy.

No significant results were found for the auditory display condition, indicating that participants could not map the movement sonification intuitively with the shape of their movements.

Within the self-other recognition task, musicians scored significantly higher in the visual ($t_{21}=2.29$, $p<.05$) and audiovisual conditions ($t_{21}=2.31$, $p<.05$) compared with non-musicians. Thus, musicians in this study possessed advanced skills in recognizing their own music-related movements potentially based on enhanced action-perception-coupling for these musical tasks.

5. CONCLUSIONS

While there were large inter-individual differences in the movement trajectories of participants in Study 1, analyses revealed a high consistency in the repeated-measures condition, so that individuals performed comparable movements across trials. These results were unexpected, considering the clear movement instructions in Block 2 based on Truslit's motion shapes. However, results of significant differences between performance conditions (free – after instruction) indicates small effects of moving intuitively on movement velocity. Furthermore, musical characteristics seem to influence movement execution, so that participants travelled longer resp. moving more while listening to the song with a complex melody line, that is Verdi's *Celeste Aida*. On the other hand, we see typical movement characteristics in terms of jitter, while moving to a melody with longer tone sequences. Thus, Wagner's *Gebet der Elisabeth* leads to a significantly higher jerk compared to Verdi. In further studies, we will focus on such correlations between musical features and movement parameters, so we will include more spatial and temporal parameters of the recorded movement trajectories as well as musical and acoustical analysis of the original Truslit samples.

Results of the self-other recognition task indicate a common perceptual basis that is grounded in human movements and lies beyond individual percepts of music, but just in terms of visual perception processes. Musicians tend to recognize their movements more often correctly compared to non-musicians, possibly showing advanced musical perception processes due to their expertise in moving while making music. The sonification used in this study did not lead to a self-recognition above chance. Further methods will be employed an auditory display method that tries to get on a deeper layer of perceptional processes while listening to sonification of movement data. Therefore, an evaluation study of different sonification mappings, sounds and styles appears to be a promising approach.

6. REFERENCES

[1] A. Truslit, Gestaltung und Bewegung in der Musik. Ein tönendes Buch vom musikalischen Vortrag und seinem bewegungserlebten Gestalten und Hören. Vierweg, 1938.

[2] B. Repp, "Music as Motion: A Synopsis of Alexander Truslit´s (1938) Gestaltung und Bewegung in der Musik. Psychology of Music, 1993, pp. 48-72.

[3] G. Becking, Der musikalische Rhythmus als Erkenntnisquelle, Benno Filser, 1928.

[4] V. Sevdalis, and P.E. Keller, „Cues for self-recognition in point-light displays of actions performed in synchrony with music", Consciousness and Cognition, 2010, pp. 617-626.

[5] C. Wöllner, „Self-recognition of highly skilled actions: A study of orchestral conductors", Consciousness and Cognition, 2012, pp. 1311-1321.

[6] W. Prinz, „Perception and action planning", European Journal of Cognitive Psychology, 1997, pp. 1-20.

[7] A. Renault, C. Charballier, and S. Chagué, „3dinmotion – a Mocap Based Interface for Real Time Visualisation and Sonification of Multi-User Interactions", in NIME International Conferences on New Interfaces for Musical Expression, 2014, pp. 495-496.

[8] N. Schaffert, Sonifikation des Bootsbeschleunigungs-Zeit-Verlaufs als akustisches Feedback im Rennrudern, Logos, 2011.

[9] A.O. Effenberg, U. Fehse, and A. Weber, „Movement Sonification: Audiovisual benefits on motor learning", in B.G. Bardy, J. Lagarde, D. Mottet (eds.), BIO Web of Conferences. The International Conference SKILLS, 2011, pp. 1-5.

[10] B. Burger, and P. Toiviainen, „MoCap Toolbox – A Matlab toolbox for computational analysis of movement data", in R. Bresin (Ed.), Proceedings of the 10th Sound and Music Computing Conference, (SMC), 2013, pp. 172-178.

[11] F. Grond, and J. Berger, „Parameter mapping sonification", in T. Hermann, A. Hunt, J.G. Neuhoff (eds.), The Sonification Handbook, 2011, pp. 363-398

[12] F. Loula, S. Prasad, K. Harber, and M. Shiffrar, „Recognizing people from their movement, in J. of Exp. Psych., 2005, pp. 210-220.