Exploring Music-Related Gestures by Sound-Tracing. A Preliminary Study.

Rolf Inge Godøy, Egil Haga, and Alexander Refsum Jensenius
Department of Musicology, University of Oslo, P.B. 1017 Blindern, N-0315 Oslo, Norway
{r.i.godoy, egil.haga, a.r.jensenius}@imv.uio.no

Abstract

This is an exploration of listeners association of gestures with musical sounds. The subjects listen to sounds that have been chosen for various salient features, and the tracing movements made by the subjects are recorded and subsequently compared in view of common features in the tracings.

1. Introduction

What kinds of gestures do listeners and performers associate with various musical sounds? This is a very extensive question and concerns our basic understanding of music as a phenomenon, and furthermore, concerns various practical applications such as gestural control of new musical instruments. For this reason, we are in our current musical gestures research[1] carrying out a number of observation studies where listeners are asked to make gestures to musical excerpts according to various given instructions. In this paper, we shall report some preliminary findings of what we have called 'sound-tracing', meaning tracing gestures that listeners make with a pen on a digital tablet in response to various sound fragments.

2. Sound-related gestures

Obviously, there are many kinds of sound-related gestures, and we have seen various schemes for gesture classification in music [1]. In our current research, we have initially distinguished between sound-producing and sound-accompanying gestures, where sound-producing gestures denote gestures necessary for the generation of musical sound, and sound-accompanying gestures denote various body movements as in dancing or more freely gesticulating to the music [2]. The borders between these categories may be blurred, something that also became apparent in the present preliminary study of sound-tracing.

A sound-producing gesture could be defined as an excitatory or modulatory gesture, i.e. a gesture transferring energy from the human body to some resonating medium (e.g. hitting, stroking, bowing) or modifying the resonance (e.g. moving a mute, changing the bow position). However, sound-producing gestures may also involve movement of effectors to and from positions of energy transfer, as in piano playing when the hand and/or elbow moves along the horizontal plane of the keyboard in order to

[1] http://musicalgestures.uio.no

position the fingers for hitting specific keys. In this case, the hand and/or elbow movement could also be understood as following the overall pitch contour of the music in the piano playing, hence be equivalent to what we would call a sound-tracing gesture. In fact, the hand movement could be seen as reflecting the coarse pitch contour of the music (in moving along the keyboard), hence could be understood as a 'low-pass' version of the more detailed, note-by-note, version of the music. In view of listeners sound-tracing gestures, it is important to see that there are different levels of acuity, i.e. detailed, note-level as well as more coarse phrase-level gestures, concurrently embedded in sound-producing gestures.

Furthermore, we may observe so-called 'ancillary' gestures in performance [1], gestures that are not indispensable, yet apparently performance-facilitating both with regards to expressivity and motor control. For this reason, we believe that although sound-producing gestures are based on energy transfer, and hence on bodily effort, they extend into 'higher level' schematic images of musical sound where it is not quite clear what is 'strictly necessary' for sound-production and what are more expressive, emotive, or even theatrical kinds of gestures. Also, the multitude of sound-accompanying gestures such as in dance or various types of free movement to the music, will of course often relate to the music by synchrony (pulses, cyclic movement, etc.), pitch contours, timbral evolutions, etc., and may be reflected in the spontaneous sound-tracing images that listeners make.

Although we recognize the great diversity of gestures that may be evoked in any listener by musical sound, we still believe there is a foundation for sound-gesture relationships in the energy features of the sound, i.e. in what we could call the overall envelope of musical sounds. For this reason, we have initially chosen to present listeners with rather short fragments of musical sound where the overall envelopes of the sounds are quite salient. From an effort point of view, we are suggesting three categories of excitatory gestures, based on what we consider distinct biomechanical phenomena in sound-production:

- *Impulsive* (could also be called 'ballistic'), meaning a spike of effort followed by relaxation and/or rebound, such as in hitting a drum, a piano key, a rapid *sforzato* bow movement, etc.
- *Sustained*, meaning a continuous effort, such as in a continuous bow movement, continuous blowing, etc.
- *Iterative*, meaning a repetitive movement that is so rapid as to fuse into one gesture, such as in a piano or harp glissando, hence actually an intermediate category between impulsive and sustained.

These sound-producing categories may be correlated to the three main categories of sonorous objects in Pierre Schaeffer's typological system, what he called *impulsive*, *sustained* and *iterative* sonorous objects [3, 4]. And furthermore, following Schaeffer into his morphological ideas, i.e. that which concerns the internal features, and the evolution over time of these features, of the sonorous object, we have a framework for associating various modulatory gestures with sound-features such as pitch, texture, and timbre [5]. The line of reasoning here is that we try to focus initially on what we consider closely production related gestural features of sound fragments, and then later on explore what other, and perhaps not so immediately sound-production related gestures, various musical fragments may evoke in listeners.

3. Embodied sound perception

The many and close links between sound and gesture could be understood within the framework of what is now often called *embodied cognition*, meaning that perception and cognition of sound is based on both neurocognitive constraints and massive ecological experience of sound-gesture relationship, including explicit, as well as more implicit, knowledge of causality, resonant features of entities, and sound-producing gestures [6]. One essential idea here is that we as listeners develop various generalized schemata for sound-perception based on innumerable instances of sound-generation, generalized schemata which we use both in the perception of familiar and unfamiliar sounds.

More specifically for sound-related gestures, the so-called 'motor theory' of perception suggested several decades ago that language perception is closely linked with mental images of the sound-producing gestures of the vocal apparatus [7]. The idea of motor theory has been contested, however recent neurocognitive research seems to increasingly lend support to the idea that motor images are indeed integral to sound-perception [8, 9], and similar close links between sound and motor images have also been reported in music [10, 11]. Based on the research briefly mentioned here, as well as on our own previous studies [2], we believe motor images may play an important role in the perception and cognition of musical sound, and notably so, for listeners with different levels of expertise in music. Furthermore, this means that motor imagery, including images of effort and kinematics, could be integral to our mental images of musical sound, and that knowledge of sound also has components of so-called *procedural knowledge* ('knowing how').

In particular, we have been interested in exploring the flexible nature of sound-related gestural images, meaning their applicability to rather different types of musical sound as well as different effectors, e.g. how the inflections in a vocal sound may be reproduced by a listener as a tracing gesture with one hand. We are here exploring the idea of *motor equivalence* [12], meaning that an action may be carried out by different effectors: We could play a flute tune on the piano, retaining some of the features of the flute tune, e.g. the pitches, durations, contour, etc., but of course also loosing other features, e.g. timbre and various expressive features.

This phenomenon of loosing some, retaining some, in the 'translation' of a fragment of musical sound from one setting to another, raises the question of *variable acuity* in the perception and cognition of musical sound. By the expression 'variable acuity' we refer to the variable amount of detail resolution in a gestural rendering of musical sound, meaning that there is a continuum from an 'exact' rendering of all detail (e.g. pitches, durations, dynamics, expressive features, etc.) to rather coarse or approximate renderings (e.g. retaining only the overall contour and/or sensation of the musical fragment).

4. Sound-tracing studies

Equipped with the twin ideas of *motor equivalence* and *variable acuity*, we have in the present study of sound-gesture relationships set out to explore how listeners with different levels of musical expertise will make gestural renderings of musical sounds. In contrast to our previous studies on 'air instrument' playing where subjects were asked to make gestures as if they were playing [2], we have in the present study asked the subjects to just simply make gestures, that they believed corresponded well with the sounds they heard, with a pen on a digital tablet. We did in other words not ask the subjects to make what they believed were sound-producing gestures, but we did not exclude that possibility either. The reason for being deliberately unspecific here, was that we wanted to see what kinds of gestures different subjects would spontaneously associate with various musical sounds.

However, we very carefully designed two sets of sounds, the first set containing 30 sounds varying in duration between 2 and 6 seconds, and the second set containing 20 sounds, also varying in duration between 2 and 6 seconds. The first set contained sounds that were rather limited with regards to features and sub-chunks, i.e. should not contain several successive envelopes, whereas the second set contained sounds that were progressively more complex with regards to features as well as sometimes containing sub-chunks.

Sounds from traditional musical instruments, as well as electronic and environmental ('concrete') sounds were used, inspired by Schaeffer's abovementioned typological principles, and chosen or edited according to Schaeffer's scheme for overall envelopes, i.e. as *impulsive sounds*, *continuous sounds*, or *iterative sounds*. With regards to pitch content, we also chose to follow Schaeffer's typology, i.e. *stable*, *changing/unstable*, or *undefined*, and similarly for timbral-textural content, i.e. also here *stable*, *changing/unstable*, or *undefined*.

In all 9 subjects participated in this initial sound-tracing study. 4 of the subjects were undergraduate or graduate music students, hence could be assumed to have rather high levels of musical expertise, 2 subjects were non-musicians, but had some musical training, and 3 subjects had little or no musical training. After being explained the purpose of the sound-tracing studies, the subjects were instructed to make what they believed were the most appropriate gesture for each sound on a Wacom A-4 digital tablet, immediately after the sound had finished playing. The subjects had no visual feedback of what they were drawing on the tablet, something we had decided on in order to have the subjects focus as much as possible on their tracing gestures. Each sound was followed by a pause of the same length as the just heard sound for doing the tracing, and the subjects were asked to preferably hear the entire sound fragment before they started making the tracing gesture. Between the first and the second sets of sounds, there was a short break, and after the second set

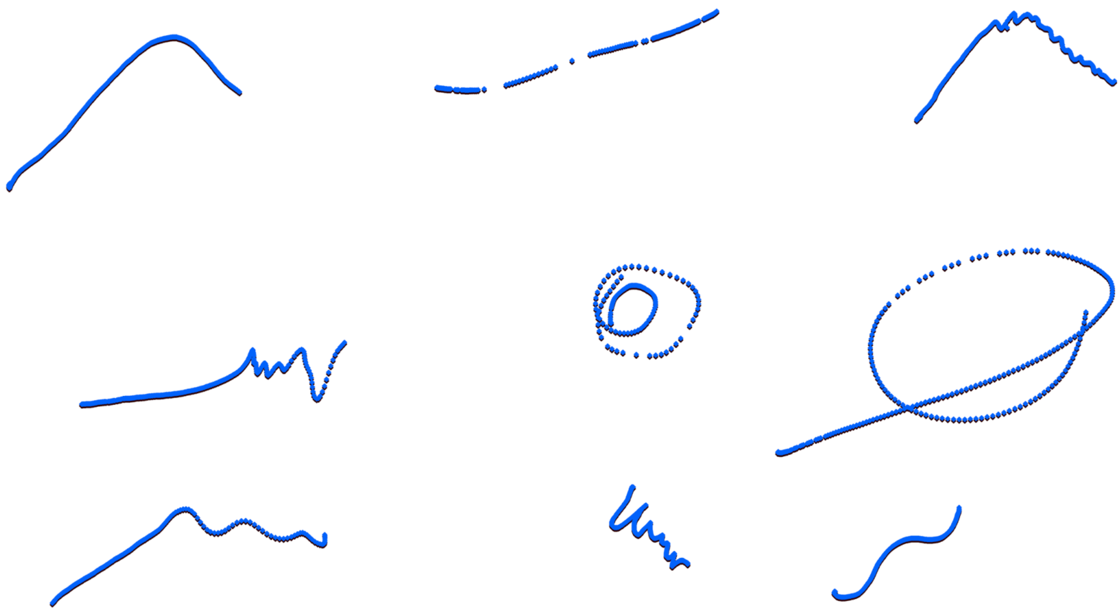

Figure 1. Synoptic presentation of sound-tracings of a fragment from Schaeffer's *Solfège de l'Objet Sonore* [13]. The tracings are arranged here according to the decreasing levels of musical training of the subjects starting from top left corner going to the right and downwards in the rows.

of sounds, the subjects were asked about their opinions of doing such sound-tracing. The MAX/MSP/Jitter software was used for the playback of the sounds as well as for the collection of the tablet data, and the tablet data was further processed in Matlab. The sound-tracing sessions as well as the subsequent interviews with the subjects about their reactions to the sound-tracing tasks, were all recorded on video as well.

These interviews were important because we were interested in how the subjects reacted to the task of 'translating' the experience of the sound-fragments into sound-tracing images. The subjects' reactions were quite varied and not easy to summarize, however several of the subjects mentioned that they felt the two-dimensional surface of the tablet to be a limitation, and that they would have preferred to do sound-tracing as free three-dimensional movement. Also, some of the subjects felt that with the increasing number of concurrent features in the second set of sounds, they were forced to leave out features in their tracings because several things were happening in parallel. – These reactions were not surprising, and we are already planning further sound-tracing studies with free 3-dimensional bi-manual movements using the Polhemus electromagnetic tracking system.

As for the data collected from the tablet, we shall do some more processing[2], but our initial overall impression is that there seems to be a fair amount of agreement between the tracings of the short fragments with rather

[2] We are currently exploring various schemes for comparison, and the presentation of this paper will include a QuickTime movie with all the subjects tracings with all the sounds presented synoptically as in figures 1 and 2.

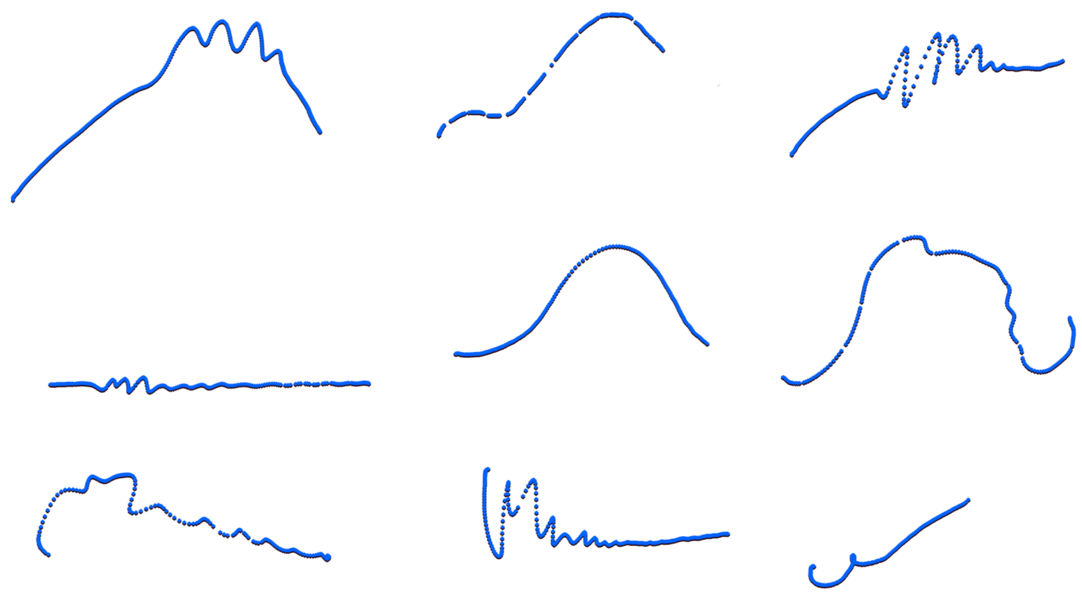

Figure 2. Synoptic presentation of sound-tracings of the entire composite object from Schaeffer's *Solfège de l'Objet Sonore* [13]. Also here, the tracings are arranged according to the decreasing levels of musical training of the subjects starting from top left corner going to the right and downwards in the rows.

unambiguous envelopes in overall energy or pitch, however there seemed to be more divergence amongst the shorter fragments that had more complex timbral-textural content as can be seen in figure 1. This is in fact an example from Schaeffer's *Solfège de l'Objet Sonore*, more precisely the second sub-object of a longer object type called 'composite object' ('objet composite') [13]. On the other hand, there seemed to be more agreement in the tracings of the longer fragment, i.e. the whole composite object where the two sub-objects are fused into what we could call one coarticulated gesture as may be seen in figure 2.

5. Conclusions

Given the imposed restrictions on the sound material in this study, i.e. both the shortness and the feature salience of the sounds, the fair amount of consistency in some of the responses is perhaps not so surprising. For instance, it is no great surprise that an ascending pitch will be traced as an ascending curve by most people in our culture, or that a percussive onset followed by a long decay will be traced as a steep slope followed by a long descent. Yet this kind of study of 'gestural primitives' [14] is in our opinion far from trivial in that it can be applied to systematic and large-scale studies of sound-gesture relationships as well as to more complex objects, as the results of the second set of sounds here seems to indicate. The strategy of using 'custom made' sounds for exploring sound-gesture relationships seems promising, and we shall in future studies extend this to a number of textural-timbral features as well.

In particular, we believe it is useful to explore further how subjects with different levels of expertise make gestures corresponding to various non-conventional musical sounds, i.e. various electroacoustic, synthetic, 'concrete', and/or ecological sounds, sounds which our western musical conceptual apparatus is not well equipped to handle. But also, we believe sound-tracing explorations of traditional, 'notatable' western music could be useful in revealing more of how people perceive higher level, i.e. chunk-level and phrase-level features of musical sound.

6. References

[1] C. Cadoz and M. Wanderley, "Gesture-Music" In: M. Wanderley and M. Battier (eds.) *Trends in Gestural Control of Music*. Ircam, Paris, 2000.

[2] R. I. Godøy, E. Haga, and A. Jensenius, "Playing 'Air Instruments': Mimicry of Sound-producing Gestures by Novices and Experts." In S. Gibet, N. Courty, and J.-F. Kamp (eds.): Gesture in Human-Computer Interaction and Simulation. 6th International Gesture Workshop, GW 2005, Berder Island, France, May 18-20, 2005, Revised Selected Papers. LNAI 3881. Springer-Verlag, Berlin Heidelberg, 256–267, 2006.

[3] P. Schaeffer, *Traité des Objets Musicaux*. Éditions du Seuil, Paris, 1966.

[4] P. Schaeffer, P. (with sound examples by Reibel, G. and Ferreyra, B.), *Solfège de l'Objet Sonore*. INA/GRM, Paris, 1998 (first published in 1967).

[5] R. I. Godøy, "Gestural-Sonorous Objects: Embodied Extensions of Schaeffer's Conceptual Apparatus", Organised Sound, 2006 (in press)

[6] D. Rocchesso and F. Fontana, (eds.), *The Sounding Object*. Edizioni di Mondo Estremo, Firenze, 2003.

[7] A. M. Liberman and I. G. Mattingly, "The Motor Theory of Speech Perception Revised", *Cognition* 21, 1985, pp. 1-36.

[8] L. Fadiga, L. Craighero, G. Buccino, and G. Rizzolatti, "Speech Listening Specifically Modulates the Excitability of Tongue Muscles: a TMS Study", *European Journal of Neuroscience* 15, 2002, pp. 399-402.

[9] E. Kohler, C. Keysers, M. A. Umiltà, L. Fogassi, V. Gallese, G. Rizzolatti, "Hearing Sounds, Understanding Actions: Action Representation in Mirror Neurons", *Science*, vol. 297, 2002, pp. 846-848.

[10] J. Haueisen and T. R. Knösche, "Involuntary Motor Activity in Pianists Evoked by Music Perception", *Journal of Cognitive Neuroscience 13* (6), 2001, pp. 786–792.

[11] R. J. Zatorre and A. R. Halpern, "Mental Concerts: Musical Imagery and Auditory Cortex", *Neuron* 47, 2005, pp. 9–12.

[12] D. Rosenbaum, *Human Motor Control*. Academic Press, San Diego, 1991.

[13] CD3, track 3 in [4].

[14] Choi, I., "Gestural Primitives and the Context for Computational Processing in an Interactive Performance System", In: M. Wanderley and M. Battier (eds.) *Trends in Gestural Control of Music*. Ircam, Paris, 2000.