

Thinking Sound and Body-Motion Shapes in Music: Public Peer Review of “Gesture and the Sonic Event in Karnatak Music” by Lara Pearson

ROLF INGE GODØY
University of Oslo, Department of Musicology

ABSTRACT: It seems that the majority of research on music-related body motion has so far been focused on Western music, so this paper by Lara Pearson on music-related body motion in Indian vocal music is a most welcome contribution to this field. But research on music-related body motion does present us with a number of challenges, ranging from issues of method to fundamental issues of perception and multimodal integration in music. In such research, thinking of perceptually salient features in different modalities (sound, motion, touch, etc.) as shapes seems to go well with our cognitive apparatus, and also be quite practical in representing the features in question. The research reported in this paper gives us an insight into how tracing shapes by hand motion is an integral part of teaching Indian vocal music, and the approach of this paper also holds promise for fruitful future research.

Submitted 2013 May 14; accepted 2013 May 29.

KEYWORDS: *Indian vocal music, body motion, shape, multimodality*

CHALLENGES OF RESEARCH ON MUSIC-RELATED BODY MOTION

ALTHOUGH we have seen a growing number of publications in recent decades documenting the many and seemingly robust links between sound and body motion in music, publications on such relationships in non-Western musical cultures are relatively few. Questions as to the universality vs. culture-specificity of various links between sound and body motion clearly call for more research into non-Western musical cultures, and Lara Pearson’s paper is a very welcome contribution in this direction. Besides giving us a view of music-related body motion in a (for many of us) not so familiar musical culture, this paper also raises a number of more general issues concerning music-related body motion, sound perception, motion perception, cross-modality, music theory, as well as issues of research methods, some of which I would like to comment on in this public peer review.

As for research methods, studying music-related body motion is intrinsically challenging because sound and motion (the main elements here) are both ephemeral, i.e. they are both time-dependent phenomena that we in research somehow have to convert to more ‘solid’ representations that enable closer scrutiny. The approach in Lara Pearson’s paper is that of an in-depth analysis of a vocal lesson documented in a video recording. For studying music-related body motion video is the most readily available and also most unobtrusive research technology (at least for the moment): that one (or several) video camera(s) can just quietly record an entire session of music-related body motion is great for the ecological validity of the performance situation, however it does present challenges in terms of extracting more precise spatiotemporal body motion data afterwards. There are some processing methods for extracting such data from video (e.g., Jensenius, 2013), but Lara Pearson has in the present study chosen to do a frame-by-frame tracking of the hand motion, resulting in clear and informative motion trajectory images of the teacher’s hand. With the addition of still images displaying postures at different stages of the motion trajectories, as well as combined signal-based pitch extraction graphs and Western (approximate) notation of the pitches, we get a good basis for following the presentation in the paper. However, it would have been useful to have a clearer indication of the temporal direction in the trajectory tracings, and also some information (graphs) about the velocity of the hand motions so that we could have some information on the rhythmic articulation as well. Also, if there were prominent changes in the velocities, acceleration graphs could have been useful. As for segmentation, it seems that the phrases, or chunks, were presented by the teacher, i.e. that there was no need for chunking the sound and body motion in the study (something that very often is an additional challenge in research on music-related body motion).

That said, we all still have substantial challenges in representing music-related motion and sound, both with regard to conserving the dynamic nature of motion and sound, and with regard to

capturing the multiple concurrent features and their nuances, and this clearly calls for much effort and research collaboration in the future. We could claim that the core of this representational challenge has to do with the relationships between the discrete and the continuous in our thinking, so now some thoughts on this.

DISCRETE VERSUS CONTINUOUS

The paper reports a teaching situation, but it would be interesting to know whether the hand motions presented in this paper are only used for teaching, or whether they also occur in actual performance situations. In both teaching and performance, the intriguing question is then: what is the main utility of these hand motions? Are they first of all a mnemonic tool for performance, or rather, a pedagogical visualization of the music? In the latter case, it seems to be related to our present day Western *sol-fa* methods, something that in turn seems to point back to ancient practices of *chironomy* (in Jewish and Christian sacred music). With my limited knowledge of Indian music theory, it seems that there is a categorical pitch system, however much more elaborated in view of nuances (inflexions and transitions) and context constellations than in Western music theory. This should mean that in the teaching session reported here, there is a combination of the more or less categorical, or even ‘discrete’, notions of pitch, with some more continuous transitions between, and within, the tones, combined with the smooth accompanying hand motions made by the teacher. This is a general issue also in various genres of Western music: pitches and onsets may be more or less discrete as categories and events in time (as represented in Western notation), whereas the sound-producing body motions are continuous.

It would not be farfetched to assume that this duality of the discrete (or discontinuous and categorical) and the continuous is present in most (if not all) instances of music: in spite of Western notation’s representations of tones symbolically as discrete points in time and as (ideally perceived) discrete pitches, any musical sound will needless to say have extensions (long or short), have a set of concurrent envelopes (of dynamics, as well as of pitch and of timbre fluctuations), and occur in some context, typically be ‘smeared’ by neighboring tones, i.e. be subject to what we would include in the phenomenon of *coarticulation*. Briefly stated, coarticulation is the contextual smearing of not just sonic events, but also of sound-producing and other music-related body motions, meaning that otherwise singular events are fused into more superordinate trajectories of sound and motion (see Godøy, Jensenius, & Nymo, 2010, for details).

Coarticulation concerns both the sonic output of sound-producing actions in music (and eminently so in spoken language) and the sound-producing body motions. Any sonic event is always included in some kind of action-trajectory, a trajectory that usually starts before the onset of the sound and continues through, and often also after the sound: the singer inhales and shapes the vocal apparatus before the onset of the sound, letting the vocal apparatus return to equilibrium after the sound has ended; the percussion player lifts the hand/mallet and makes a trajectory towards the instrument before the impact and sound onset, and often continues with a rebound after the impact, etc.

Interestingly, this duality of the discrete and the continuous was similarly at the root of Pierre Schaeffer’s music theory (Schaeffer, 1966), where the discrete (and symbolic) was contrasted with the continuous (and concrete), a distinction that also is relevant for music-related body motion (see Godøy, 2006, for embodied extensions of Schaeffer’s ideas). Furthermore, the basic element in Schaeffer’s music theory was precisely that of thinking shapes, meaning that any perceived feature of music could be conceptualized as a shape, e.g., the overall dynamic, timbral, and pitch features, as well as a number of sub-features, e.g., various fluctuations within the sonic object, as demonstrated with sound examples in Schaeffer (1998).

SOUND-BODY MOTION RELATIONSHIPS

In my own research presentations, I have sometimes suggested that music could be defined as “sound + body motion”. The point with this rather simplistic definition is to highlight the close links between sound and body motion, as has been suggested by the so-called *motor theory of perception* (see Galantucci, Fowler, & Turvey, 2006, for an overview). Various variants of this theory have been presented since the 1960s, and with the advent of brain observation methods in the past decades it has become quite clear that perception is closely linked with motor control in the human brain. From both neurological and behavioral studies there is now converging evidence suggesting that auditory perception also entails a mental simulation of the sound-producing motions and/or other motions related to the sound (see Godøy, 2010, for an overview).

This means that perception may be regarded as an active process where we extract useful information from sensory input by way of simulating and/or tracing whatever it is that we focus our attention on. This means also that most listeners, regardless of training and/or previous experience, will tend to apply some body motion imagery to whatever they perceive. In some cases, the capacity for such simulation and/or outright imitation can be quite impressive, as the phenomena of so-called *scat*

singing and *beat-boxing* attest to. In particular with beat-boxing, we can often observe an astonishing capacity to imitate non-vocal sounds (e.g., percussion sounds, DJ scratching, as well as various environmental sounds) suggesting a very extensive capacity for active perception.

Furthermore, we can often observe cases of so-called *air-instrument* performance, as for instance in *air-guitar* or *air-drums*. We previously tried to make some more systematic studies of air-instrument performance because we believed this testified to the active nature of perception (as well as to the extensive knowledge of sound-producing body motion by very many, if not most, listeners). As an extension of this, we have also done a number of so-called sound-tracing studies with the aim of finding out more about how listeners, with different levels of musical training, spontaneously trace the shapes of what they perceive as salient features of the music (Nymoën, Tørresen, Godøy, & Jensenius, 2012).

With these experiences of music-related body motion in mind, I think it is reasonable to see a triangular relationship between sound, vision, and sense of motion (motion, which in turn may be seen as composite, including both proprioception and haptics). The hand motions of the vocal teacher reported in this paper could then be seen as situated within this triangle where sound, vision and motion are perceptually interconnected.

The role of hand motion here is significant for several reasons. For one thing, it is a visualization of the melodic contour in a teaching situation, and we could also suspect that it is a mnemonic tool for musicians, thus also a tool for structuring the music as such. In more general terms, it could be tempting to speak of “manual cognition” in the sense that hand motion seems to have a fundamental cognitive function, as has been suggested by research on language and evolution (Rizzolatti & Arbib, 1998) and on cognitive processes in general: hand motions help us think (Goldin-Meadow, 2003).

Monitoring the motions in the vocal apparatus, as mentioned towards the end of the paper, is a very interesting idea, but it seems to me that more research is needed both on the activity of the vocal apparatus and its neurocognitive basis. However, for the moment we could understand the manual rendering of vocal sound as a case of so-called *motor equivalence* (Kelso, Fuchs, Lancaster, Holroyd, Cheyne, & Weinberg, 1998). Motor equivalence means that one set of effectors can do the job of another set of effectors, e.g., I can open a door with my shoulder or my foot if my hands are occupied with carrying suitcases. In our context of music-related body motion, this would mean that the hands could try to render some task or feature of the vocal apparatus, in a way the reverse of beat-boxing where the vocal apparatus is trying to render sound which is usually produced by the hands on instruments. In both cases, the crucial point is that of rendering some perceptually salient features “originally” belonging to one set of effectors with another set of effectors.

IT'S ALL ABOUT SHAPES

Taking into account outcomes of research on music-related body motion of the past couple of decades, it would in my opinion not be unreasonable to suggest that most, maybe all, musical features could be conceived of (and possibly also represented) as shapes:

- Pitch contours, both at the timescale of melodic contours and the smaller timescales of ornaments, fluctuations (vibrato) and various other inflections and contextual transitions (e.g., portamento).
- Dynamic contours, also at different timescales, ranging from relative large (e.g., a long-drawn crescendo) to smaller (e.g., the dynamic profile of single tones as well as various fluctuations within the tones).
- Timbral contours, ranging from slow or even quasi-stationary as in vowels and instrumental formants, to faster, as in vowel or formantic changes (e.g., the *wah-wah* mute opening-closing on brass instruments, readily rendered vocally by opening-closing the mouth).
- Tempo curves of all kinds at different timescales, including musical expressivity (e.g., rubato and groove).

But shapes could also be useful in representing more composite and ‘higher level’ features such as:

- Overall density of events, pitch space, dynamic range, etc.
- Overall sensations of affect, e.g., calm (with long, protracted shapes) vs. agitated (with shorter, more ‘choppy’ shapes), etc.

In all such thinking of shapes, the intriguing and also very productive element is the *shifts* in our minds between more or less “static” images of sound and motion, and the real-time temporal unfolding of sound and motion. This allows us mentally to zoom in and out of the music we are exploring, generating different temporal perspectives and thereby enhancing our knowledge of musical features. Besides giving us an insight into Indian vocal music, I see this paper as also an effort in the direction of such an understanding of shape and motion in music, and I look forward to Lara Pearson’s future research here.

REFERENCES

- Galantucci, B., Fowler, C.A., & Turvey, M.T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin and Review*, Vol. 13, No. 3, pp. 361-377.
- Godøy, R.I. (2006). Gestural-sonorous objects: embodied extensions of Schaeffer's conceptual apparatus. *Organised Sound*, Vol. 11, No. 2, pp. 149-157.
- Godøy, R.I. (2009). Geometry and effort in gestural renderings of musical sound. In: M. Sales Dias, S. Gibet, M. M. Wanderley, & R. Bastos (Eds.), *Gesture-Based Human-Computer Interaction and Simulation*. 7th International Gesture Workshop 2007, Lecture Notes in Artificial Intelligence, Vol. 5085. Berlin, Heidelberg: Springer-Verlag, pp. 205-215.
- Godøy, R.I. (2010). Images of sonic objects. *Organised Sound*, Vol. 15, No. 1, pp. 54-62.
- Godøy, R.I., Jensenius, A.R., & Nymoen, K. (2010). Chunking in music by coarticulation. *Acta Acustica united with Acustica*, Vol. 96, No. 4, pp. 690-700.
- Goldin-Meadow, S. (2003). *Hearing Gesture: How Our Hands Help Us Think*. Cambridge, MA: Harvard University Press.
- Jensenius, A.R. (2013). Some video abstraction techniques for displaying body movement in analysis and performance. *Leonardo: Journal of the International Society for the Arts, Sciences and Technology*, Vol. 46, No. 1, pp. 53-60.
- Kelso, J.A.S., Fuchs, A., Lancaster, R., Holroyd, T., Cheyne, D., & Weinberg, H. (1998). Dynamic cortical activity in the human brain reveals motor equivalence. *Nature*, Vol. 392, No. 6678, pp. 814-818.
- Nymoen, K., Tørresen, J., Godøy, R.I., & Jensenius, A.R. (2012). A statistical approach to analyzing sound tracings. In: S. Ystad, M. Aramaki, R. Kronland-Martinet, K. Jensen, & S. Mohanty (Eds.), *Speech, Sound and Music Processing: Embracing Research in India*. Lecture Notes in Computer Science, Vol. 7172. Berlin, Heidelberg: Springer-Verlag, pp. 120-145.
- Rizzolatti, G., & Arbib, M.A. (1998). Language within our grasp. *Trends in Neuroscience*, Vol. 21, No. 5, pp. 188-194.
- Schaeffer, P. (1966). *Traité des Objets Musicaux*. Paris: Éditions du Seuil.
- Schaeffer, P. (1998). *Solfège de l'Objet Sonore*. (first published in 1967, with sound examples by G. Reibel, and B. Ferreyra). Paris: INA/GRM.