EFFORT-VOICE RELATIONSHIPS IN INTERACTIONS WITH IMAGINARY OBJECTS IN HINDUSTANI VOCAL MUSIC

Stella Paschalidou
Dept. of Music Technology & Acoustics Eng.,
TEI of Crete, Greece
pashalidou@staff.teicrete.gr

Martin Clayton, Tuomas Eerola
Dept. of Music, Durham University, UK
martin.clayton@durham.ac.uk,
tuomas.eerola@durham.ac.uk

ABSTRACT

In Hindustani (dhrupad) vocal improvisation singers often engage with melodic ideas by manipulating intangible, imaginary objects with their hands while singing, such as through stretching, pulling and pushing. Such engagements ('MIIO' for Manual Interactions with Imaginary Objects) suggest that some patterns of change in the acoustic features relate to rudimentary interactions and the levels of effort that the respective objects may afford due to their physical properties. Through this work we seek to gain a deeper understanding of performance practice in the dhrupad music tradition in the specific cases where the singer seems to interact with imaginary objects, by examining whether effort and gesture types appear in an arbitrary fashion along with the voice or if they are related to the sound in a consistent way. The results suggest that a good part of the variance in both physical effort and gesture type can be explained through a small set of audio and motion features.

1. INTRODUCTION

In recent years we have seen a shift towards more embodied approaches in the study of music performance. However, while sound-producing gestures have drawn strong attention, studies on sound accompanying gestures (as in singing) have been extremely rare, even more so in terms of computational approaches (Luck & Toiviainen, 2008) and in the case of non-western 'oral' music traditions (Clayton & Leante, 2013) like the one portrayed here. Additionally, although physical effort has been stressed as one of the important aspects in music performance, systematic approaches in its role still remain limited.

Here effort is understood as a concept which reflects the active or passive attitude of the person in fighting against or giving in to the physical conditions that influence the movement while trying to achieve an intentional task (Hackney, 1998). During MIIIOs Hindustani singers seem to act as if they encounter an increased resistance upon their hands, presumably imitating the effort that would have been induced in handling real objects in our natural environment. Although MIIOs can be considered as founded on rudimentary knowledge of interacting with the environment, whether and how these are related to their melodic counterpart is a non-trivial question. Examining whether, how and to what extent effort and gesture types appear associated with the voice in a consistent way is where this work aims to offer a contribution.

2. METHOD

The work uses a mixed methodological approach, combining qualitative (ethnographic) and quantitative methods based on original material that was recorded in domestic spaces in India. This consists of interviews (8 vocalists), audio-visual material (4 vocalists) and motion (10-camera passive-marker Optitrack system) capture data (2 vocalists) of vocal improvisations by different dhrupad vocalists of the same music lineage. We argue that the sequential approach of combining the rich outcomes of qualitative methods with the compact results of quantitative methods can offer a more rigorous and comprehensive picture of the phenomena under study.

Real performances rather than designed experiments were used for the purposes of ecological validity. Singers were asked to improvise without any further instructions and were recorded only during the alap improvisation (the initial slow non-metered section, sung to a repertoire of non-lexical syllables), in order to concentrate on melodic factors rather than the metrical structure or lyrical content in the later stages of the rāga performance.

In the qualitative part of the analysis, we first applied a thematic analysis to the interview material in order to identify action-based metaphors, which informed the annotation process of the video material that followed. For the video analysis we relied on third-person observations that aimed at identifying, labeling and later classifying the audio-visual material in terms of recurrent types of MIIOs (categorical), as well as perceived effort levels (numerical) that appear to be exerted by the performer in a range between 0 (lowest) and 10 (highest). Such annotations were cross-validated by two choreographers.

In the quantitative part of the study, we used the cross-validated annotations as response values of linear models that were fit to measured movement and sound features in order to (a) estimate effort levels and (b) classify gestures as interactions with either elastic (through elasticity) or rigid (through weight/friction) objects. Two vocalists were used here, namely Afzal Hussain (rāga Jaunpurī) and Lakhan Lal Sahu (rāga Mālkauns). The features that were used for estimating the responses were computed by first extracting time-varying movement and audio features from the raw data, and then computing representative statistical global measures (such as mean, SD, min, max). We started by using the features reported in (Nymoen et al, 2013), but then a number of alternative features were also explored that were meant to raise the explained variance of the estimated responses.

3. RESULTS

Two variations of linear models were developed for each task (effort estimation and gesture classification):
(1) a model that best fits each individual performer, thus better reflecting the idiosyncratic aspect of each singer;
(2) a model that can better describe shared, more generic cross-performer behaviours.

3.1 Idiosyncratic schemes

3.1.1 Effort levels

Different idiosyncratic schemes of associating the perceived physical effort with acoustic and movement features were identified, that are based on the pitch space organisation of the rāga and the mechanical strain of voice production.

Hussain: Higher effort levels are required when the hands move slower and wider apart and with a larger speed variation. They are accompanied by melodic glides that start from lower degrees and ascend to higher degrees of the rāga scale within the boundaries of each individual octave, thus they are associated with characteristic qualities of the specific rāga. The use of 5 non-collinear audio and movement features in the linear models that were developed yielded a good fit (R^{2adj}) of about 60%.

Sahu: Higher bodily effort is required for hand movements that exhibit a larger variation of hand divergence (speed in moving the hands further apart), with a strong onset acceleration. They are accompanied by larger melodic glides that reach up to higher maximum pitches, reflecting the increased mechanical strain of voice production. As the alap is organised based on a gradual ascent towards the pitch climax, pitch is here also representative of the alap macro-structure. The use of 4 non-collinear audio and movement features in the linear model yielded an adequately good fit (R^{2adj}) of about 44%.

3.1.2 Gesture classification

Different modes of gesture class association with acoustic and movement features were identified, that are based on regions of particular interest in the rāga pitch space organization and analogous cross-domain morphologies.

Hussain: It is more likely that interactions with elastic objects (rather than rigid) are performed by hand gestures that exhibit a low absolute mean acceleration and a large variation in hands' divergence. They are associated with slower and larger melodic movements that ascend to a higher degree of the scale. Interestingly, the highest degree happens to be the most unstable degree of the scale (in rāga Jaunpurī), which imposes a subsequent pitch descent (i.e. a double pitch glide), similar to the change of direction observed by the hands when interacting with an elastic object. Thus, it could be suggested that MIIO types are associated with the grammatical rules of the rāga. The use of 5 non-collinear audio and movement features in the logistic models that were developed yielded a high classification rate (AUC) of about 95%.

Sahu: Interactions with elastic objects are more likely performed with pitch movements of a larger interval and larger duration and with the hands moving faster and remaining bound to each other. The use of 4 non-collinear audio and movement features in the logistic models yielded a high classification rate (AUC) of about 80%.

3.2 Generic scheme

3.2.1 Effort levels

Higher bodily effort levels are required by both singers for melodic movements that start from a lower and reach up to a higher pitch, reflecting the mechanical requirements of voice production. They are accompanied by movements which are slow on average but exhibit a large variation of speed, and in the specific case of Hussain when the hands move further apart. Two almost identical linear models were developed, yielding a good fit (R^2adj) of about 53% (with 5 features) for Hussain and 42% (with 4 features) for Sahu respectively.

3.2.2 Gesture classification

Interactions with elastic objects are more likely to be performed at lower pitches for larger melodic movements, and with the hands moving further apart for Hussain and less apart but faster in the case of Sahu. Two almost identical general logistic models were developed, yielding a good fit (AUC) of about 86% (with 3 features) for Hussain and 78% (with 4 features) for Sahu respectively.

4. CONCLUSIONS

MIIOs offer a special case where motor imagery is "materialised" through physical actions directed towards an imagined object. Despite the flexible character of music-movement correspondences, there is ample evidence of more generic associations that are not necessarily performer-specific or stylistic. I suggest that the vocalists' capacity of imagining musical sound is facilitated through the retrieval of motor programs and image schemata from well-known real interactions with real objects and that this may be exactly the reason for which imaginary objects are employed.

As much as bringing the advantages of ecological validity in combining ethnographic data with exact measurements of real performances, the approach that was followed has also posed important challenges and limitations, such as the limited dataset. Larger datasets of multiple performers, performances and rāgas for each performer would be beneficial for enabling a more systematic comparison between performers, performances and rāgas.

5. REFERENCES

Hackney, P. (2003). *Making connections: Total body integration through Bartenieff fundamentals*. Routledge.

Luck, G., & Toiviainen, P. (2008). Exploring Relationships between the Kinematics of a Singer's Body Movement and the Quality of Their Voice. *Journal of interdisciplinary music studies*, 2.

Nymoen, K., Godøy, R. I., Jensenius, A. R., & Torresen, J. (2013). Analyzing correspondence between sound objects and body motion. *ACM Transactions on Applied Perception (TAP), 10(2), 9*.

Clayton M., & Leante L. (2013). *Experience and Meaning in Music Performance* (Martin Clayton, Byron Dueck & Laura Leante (eds.)). Oxford University Press, USA, 188-207.