Data-Driven Design of Sound for Enhancing the Perception of Expressive Robotic Movement

Luke Dahl
Department of Music
University of Virginia
Charlottesville, Virginia 22904
lukedahl@virginia.edu

Jon Bellona
Department of Music
University of Virginia
Charlottesville, Virginia 22904
jpbellona@virginia.edu

Lin Bai
Electrical and Computer Engineering
University of Virginia
Charlottesville, Virginia 22904
lb7ss@virginia.edu

Amy LaViers
Mechanical Science and Engineering
University of Illinois
Champaign-Urbana, Illinois 61801
alaviers@illinois.edu

ABSTRACT

Since people communicate intentions and inner states through movement, robots can better interact with humans if they too can modify their movements to communicate changing state. These movements, which may be seen as supplementary to those required for workspace tasks, may be termed "expressive." However, robot hardware, which cannot recreate the same range of dynamics as human limbs, often limit expressive capacity. One solution is to augment expressive robotic movement with expressive sound. To that end, this paper presents a study to find a qualitative mapping between movement and sound. Musicians were asked to vocalize sounds in response to animations of a simple simulated upper body movement performed with different movement qualities, parametrized according to Laban's Effort System. Qualitative labelling and quantitative signal analysis of these sounds suggests a number of correspondences between movement qualities and sound qualities. These correspondences are presented and analyzed here to set up future work that will test user perceptions when expressive movements and sounds are used in conjunction.

CCS CONCEPTS

•Applied computing →Sound and music computing;

KEYWORDS

Human-robot interaction, laban effort system, movement quality, music, sound-motion, MIRtoolbox

ACM Reference format:
Luke Dahl, Jon Bellona, Lin Bai, and Amy LaViers. 2017. Data-Driven Design of Sound for Enhancing the Perception of Expressive Robotic Movement. In *Proceedings of ACM MOCO conference, London, United Kingdom, 28-30 June 2017 (MOCO'17),* 8 pages.
DOI: 10.475/123_4

DOI: 10.1145/3077981.3078047

1 INTRODUCTION

We typically think of robotic movement as efficient, inexpressive, and purely functional. Humans, however, often communicate intentions and affective states through qualitative aspects of their body movement. The same gesture can express conviviality, as when one waves to a friend on the street, or urgency, as when waving to get the attention of a cab. The difference is in *how* the gesture is executed.

Robots in human-facing roles could more effectively interact with humans if they too could express various intentions through their movement. Consider a robot that guides someone from one place to another. If this robot works in a kindergarten it would move gently, and if it worked in a police station it might move authoritatively. Ongoing research is being conducted with the goal of endowing robots with the ability to express different qualities in their movement [9, 13]. However, the physical limitations of a specific platform can reduce a robot's ability to sufficiently express differences in movement qualities [1].

Humans also communicate intentions and affective states sonically through non-verbal aspects of their vocalizations. These include non-verbal utterances as well as the varying rhythm and intonation of speech (i.e. prosody). For example, an enthusiastic or agitated person might sharply modulate the pitch and loudness of their voice in a way that mirrors sharp accelerations in the movements of their hands as they speak. Thus, by endowing robots with expressive sound we may improve people's perception of differing qualities in robotic movement. The question then is, what is the right sound to accompany a particular expressive movement?

This paper describes a study conducted in order to find correspondences between movement qualities and sound qualities. We start with animations of a simple movement performed with the qualities of each of the eight *Basic Effort Actions (BEAs)* proposed by movement theorist Rudolf Laban (described below). Musicians were presented with these animations, along with the Effort Action label, and were asked to vocalize a sound that matched the qualities of the movement and label. We then analyzed these sounds in order to find how sonic qualities vary with respect to changes in movement quality (specified using the Laban Effort Factors described below). This paper extends preliminary results presented at a robotics workshop [1].

2 BACKGROUND

A number of studies have been conducted on relationships between body movement and musical sounds. One common paradigm is to study movement responses to music or sound. For example, people's spontaneous dance movements when listening to samba and chacarera music was analyzed in [16], and soundtracing, a more constrained movement response to music, has been studied in [17]. Another paradigm is to ask people to mime the performance of musical instruments while listening to the sound of drums [2] or violin [23]. Similar work involved recording people making conducting gestures in time to prerecorded music [18]. The gestural and vocal imitation of preexisting sounds was analyzed to study the communication of sound information in [20]. Imagined movement responses to sound were studied in [4]. A discussion of various sound-related gestures and their correspondence to dynamic envelopes in sound, and an analysis of movement response to music is presented in [6]

Our study moves the other way by asking participants to generate sounds for preexisting movements and associated descriptive terms. By controlling the qualities in the movement stimuli, we hope to isolate the sound characteristics that correspond to a given movement quality. The description of movement we used in our study is the codified body of knowledge in dance and choreography known as Laban/Bartenieff Movement Studies (LBMS). Individuals with certifications in this work are called Certified Movement Analysts (CMAs). This approach is similar to work in [5] in that they use the Laban Effort Factors to organize movement, however their participants perform both movement and sound simultaneously. Their data is used to train a system that generates sound from dancers' movements in real time, whereas our work aims to generate sound from movement quality parameters.

Sonifying movement data according to user-chosen parameters is described in [3], and [10] created a game-like interaction for validating sonifications of movement.

Previous work has investigated the effect of varying robotic movement on humans' feelings. Flying robots can use their locomotion paths to communicate affective information [21]. And LBMS was used to study the impressions produced by robotic motion in [15]. Our work generating robotic movement by mapping from the components of Laban's Effort system to weighted parameters in an optimal control problem builds on [14].

2.1 Describing Movement Quality

The notion that the same 'movement' can be expressed in different ways is apparent in our language. We separate actions (verbs) and modifiers (adverbs) into distinct descriptors. We can, for example, 'walk' or 'paint' in many different ways, with different functional and expressive objectives achieved in each. A teacher will walk into the classroom in a very different way than a student who is late for the class. In this example the variation is *expressive*. On the other hand, the brush stroke required to paint an object is very different depending on the viscosity of the paint. In this example movement variation in movement is *functional* – yet relies on expressive quality to be achieved.

Table 1: Laban's Eight Basic Effort Actions

Movement	Time	Space	Weight
Gliding	Sustained	Direct	Light
Pressing	Sustained	Direct	Strong
Floating	Sustained	Indirect	Light
Wringing	Sustained	Indirect	Strong
Dabbing	Sudden	Direct	Light
Thrusting	Sudden	Direct	Strong
Flicking	Sudden	Indirect	Light
Slashing	Sudden	Indirect	Strong

We use the Effort system defined by Rudolf Laban [11, 22] to qualitatively specify the ways a movement's "tone" may vary. In this system, movements are characterized using the four Effort Factors: Space Effort, Time Effort, Weight Effort, and Flow Effort. We capitalize these terms to avoid possible confusion with other notions of these terms.

Space Effort describes the attention a movement pays to the environment. A movement can be Direct (as in dabbing paint onto a canvas), or Indirect (as in releasing a balloon). Time Effort describes the attitude towards initiation and finish of a movement; it can be Sudden (as in fencing) or Sustained (when moving something heavy). Weight Effort describes the attitude towards the mover's mass; it may be Strong (as in punching) or Light (when tapping an icon on a touch screen). Flow Effort describes the progression of a series of movements. We do not use Flow Effort in this study, since the movements we use can be considered a single movement.

Laban proposed the eight BEAs which are formed by the extreme values of three of the Effort Factors: Space Effort, Weight Effort, and Time Effort [11], as shown in Table 1. In our study we employ these eight BEAs – Dabbing, Flicking, Floating, Gliding, Pressing, Slashing, Thrusting, and Wringing – as a prototypical set of movements with different qualities with which people have preexisting experience to draw from.

3 METHODOLOGY

The goal of our study is to search for qualitative correspondences between movement and sound, so that we can endow expressive robotic movement with expressively coherent sound. To prepare for this study, the authors participated in movement-based training on the Effort Factors with a CMA.

We start with animations of a simple movement performed with the qualities of each of the eight BEAs. We presented musicians with these animations, along with the BEA label, and asked them to vocalize a sound that matched the qualities of the movement and label. We recorded these sounds and subjected them to analyses described in Section 4. The labels were necessary because the animations, though generated with different movement qualities, were not different enough for the participants to be able to easily distinguish the differences. Indeed, this is part of the motivation for conducting this study.

Next we describe the process for generating movement with varying qualities, followed by the details of gathering the data.

3.1 Generating Movement

We created animations of a stick figure moving from a pose where the hands are near the center line, to a pose where the arms are extended to the sides, and then returning to the first pose. Each movement lasted four seconds. The animation trajectories were calculated in MATLAB™as solutions to an optimal control problem leveraged from [14].

$$\min_u J \tag{1}$$

$$s.t.$$

$$\dot{x} = Ax + Bu \tag{2}$$

$$y = Cx \tag{3}$$

where $A \in \mathbf{R}^{n\times n}$, $B \in \mathbf{R}^{n\times m}$, and $C \in \mathbf{R}^{l\times n}$. The cost function is

$$J = \frac{1}{2}\int_0^{T_f} [(y-r)^T Q(y-r) + u^T Ru + \dot{x}^T P\dot{x}]dt + \frac{1}{2}(y-r)^T S(y-r)\Big|_{T_f} \tag{4}$$

We extended the design of the stick figure in [14] by making it have shoulders, elbows and wrists (Figure 1). Thus, we consider a 6-dimensional system described in Equations (2) and (3) with the state $x = [\theta_1, \dot{\theta}_1, \theta_2, \dot{\theta}_2, \theta_3, \dot{\theta}_3]$ and the input $u = [u_{\theta_1}, u_{\theta_2}, u_{\theta_3}]$. We can generate trajectories with different qualities by varying the Q, R, P, and S parameters in the cost function (4) according to the desired Laban Effort Factors.

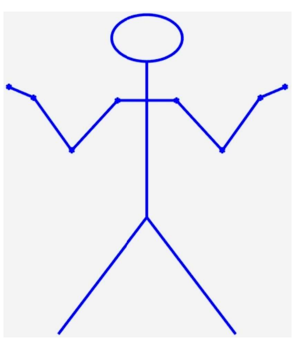

Figure 1: A still frame of the animation presented to participants.

3.2 Recording Sounds

Five graduate students and two music professionals in music composition from the University of Virginia participated in the study. They were recruited due to their significant experience performing and improvising music. The participants ranged in age from 24 to 46, with a median age of 32.

Participants were shown each of the eight animations, as well as the label of the Basic Effort Action for the animation (i.e. Gliding, Pressing, etc.) Participants were asked to vocalize a sound for each animation, such that their vocalization began at the start of the movement and lasted the duration of the movement. A three-second countdown was given before each animation to help with timing. Participants were allowed to practice before recording and to record up to three takes. After recording, we removed any takes with problems (e.g. clipping distortion). We then kept the last take, unless the participant indicated that a different take was best.

We built a custom software interface with Max/MSP™ to display the animation and record the participants vocalizations. All participants were recorded in an isolated studio environment using a Neumann TLM103 microphone with pop filter, and a Focusrite Scarlett 2i2 audio interface for microphone pre-amplification and analog-to-digital conversion.

4 ANALYSIS

Our goal in analyzing the recorded vocalizations is to discern whether there are general trends in how the sounds differ as movement quality varies. Our stimuli, the animations, were organized according to the Basic Effort Actions. However, what we would like to understand is how characteristics of sound change according to different values of the Laban Effort Factors: Time Effort, Space Effort, and Weight Effort. Thus, in the analyses that follow, we reorganize the data according to Effort Factors, which is a two step process. The first step is to apply qualitative labels to or calculate various quantities for the individual recordings. We then organize the data according to the Effort Factors. For example, when considering Time Effort we collect the data for the Gliding, Pressing, Floating, and Wringing movements into a group for Sustained Time Effort. The remaining movements (Dabbing, Thrusting, Flicking, and Slashing) are placed into a group for Sudden Time Effort. (It may help to review Table 1.) The groups allow us to compare how a given sonic quality differs between the Sustained and Sudden Time Efforts.

We performed two different analyses of the recordings. In the first analysis, we manually applied qualitative labels to each recording. In the second analysis, we performed signal analysis of the recordings in order to quantify various sonic qualities.

4.1 Qualitative Analysis

Our first analysis consisted of applying qualitative labels to the recordings. After listening to a sample of the recordings we chose to focus on the fundamental sonic attributes of pitch, amplitude (or loudness), and timbre, and after some discussion converged on the following labels.

The overall pitch of each sound is described using the labels *low, medium, high, none.* For the overall amplitude of each sound we use the labels *very soft, soft, medium, loud, very loud.* The labels for the overall timbre of each sound are *dark tone, dark noise, medium tone, medium noise, bright tone, bright noise.* 'Tone' refers to pitched sound, and 'noise' to non-pitched noisy sound, and 'dark', 'medium', and 'bright' refer in general to the relative distribution of energy across the spectrum (a preponderance of energy at low frequencies sounds dark whereas high frequencies sound bright). These categories are similar to those used in [17].

We also apply labels to the shape of how the pitch, amplitude, and timbre vary over the duration of each sound. These three attributes, which we call pitch shape, amplitude shape, and timbre shape,

Table 2: Sound categories and labels for qualitative analysis.

Pitch	Amplitude	Timbre	Shapes
None	Very Soft	Dark Tone	Sustained
Low	Soft	Dark Noise	Start Emphasis
Medium	Medium	Medium Tone	Middle Emphasis
High	Loud	Medium Noise	End Emphasis
	Very Loud	Bright Tone	Oscillating
		Bright Noise	Linear Increase
			Linear Decrease

are all described using the labels *start emphasis, middle emphasis, end emphasis, linear increase, linear decrease, sustained, oscillating*. 'Emphasis' refers to a sense of sudden intensification of the given quality, and 'start', 'middle', and 'end' refer to when an emphasis occurs.

Each of the four authors manually listened to and applied labels to each of the 56 recordings (7 musicians × 8 BEAs). For each attribute, a listener was allowed to apply only one label.

To facilitate our search for meaningful correspondences between movement and sound qualities we organized the label data according to the Effort Factors (as described above), and created histograms to compare the two values of each Factor for each quality label. By examining these histograms (see Figures 2,3,4) we notice the following.

4.1.1 Qualitative Encoding: Weight. In Figure 2 we see interesting relationships between the movement quality of Weight Effort (Light vs. Strong), and the sonic qualities of amplitude and timbre. For example, sounds that corresponded to BEAs with Light Weight Effort were more often labeled with soft amplitude, whereas sounds corresponding to BEAs with Strong Weight Effort were more often labeled medium or loud, suggesting a Weight Effort to amplitude correlation.

For timbre, Strong Weight sounds contained more dark tone and dark noise labels, whereas Light Weight Effort sounds had more mid and bright tone labels. In addition, from the three shape qualities we see that Strong Weight Effort sounds seems to have more end-emphasis, while Light Weight Effort sounds seemed to have more sustained and middle emphasis labels.

4.1.2 Qualitative Encoding: Time. For the Time Effort Factor (Sudden vs. Sustained), a few qualitative correlations appeared (see Figure 3). Vocalizations for Sustained movements were more often labeled with low and medium pitch, whereas for Sudden movements vocalizations pitch was more often labeled as 'none'. This suggests that Sustained movements may be more likely to have pitched sounds, and Sudden movements may be more noisy. The Timbre label counts support this finding, as sounds for Sudden Time Effort were more often labelled 'bright noise', and sounds for Sustained Time Effort were more often 'mid tone'.

When we look at amplitude we see that Sudden movement vocalizations were more often labelled 'loud', and that Sustained Time Effort had slightly more labels of 'medium' and 'soft'. In addition, from the three shape categories we see that in general Sudden movement vocalizations had more end- or middle-emphasis labels,

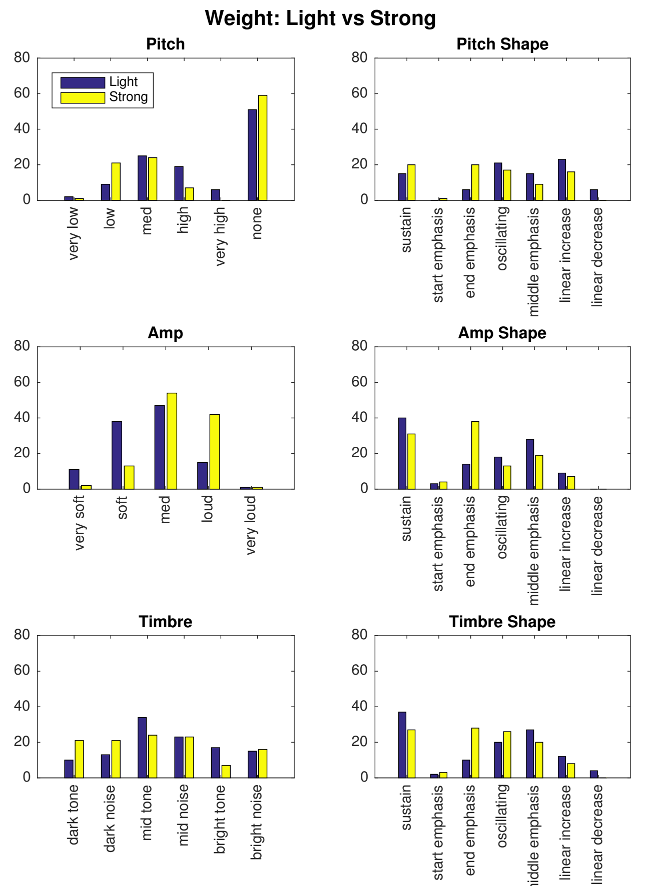

Figure 2: Qualitative label counts for Weight Effort comparing Light vs. Strong

whereas Sustained movement vocalizations were more often sustained.

4.1.3 Qualitative Encoding: Space. We did not notice any obvious differences between direct and indirect sound groups. The qualitative label histograms for space are shown in Figure 4.

4.2 Quantitative Analysis

The qualitative analysis described above suggests a number of relationships between movement qualities and sonic characteristics. However, it relies on a subjective labelling performed by the authors, and may be vulnerable to subtle biases. Thus, we sought to verify and possibly extend these findings by performing signal analyses of the recorded vocalizations.

Using the MIRtoolbox [12] library, we extracted the following audio features from each recording : amplitude envelope, spectral brightness, spectral centroid, spectral rolloff, spectral flatness, and zero-crossing rate. For each feature, the value was calculated for a series of time frames spanning the duration of the recording using MIRtoolbox's default frame length of 50 ms with 50% percent overlap between frames. Then for each recording, the mean value of the feature across all frames was calculated. It is these means

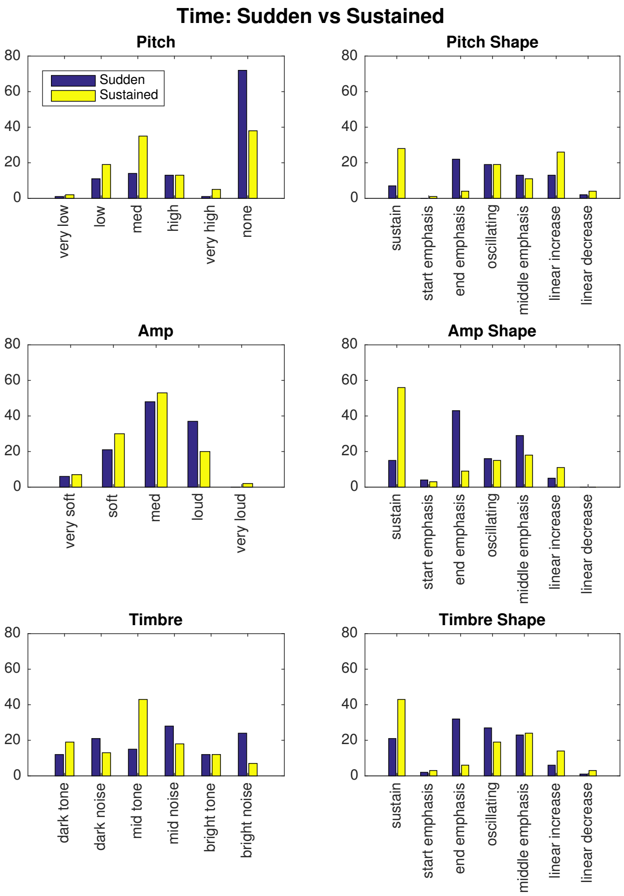

Figure 3: Qualitative label counts for the Effort Factor Time comparing Sudden vs. Sustained

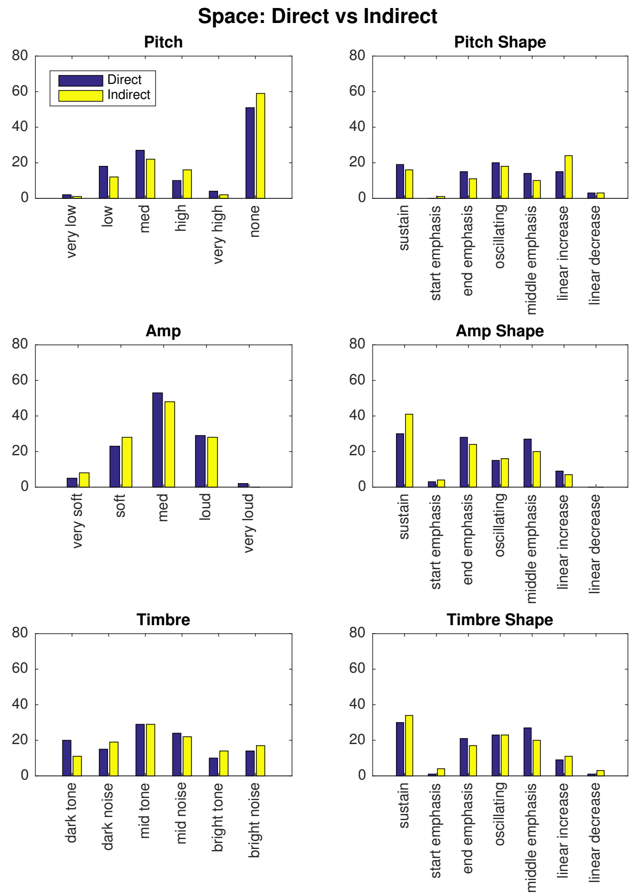

Figure 4: Qualitative label counts for the Effort Factor Space comparing Direct vs. Indirect

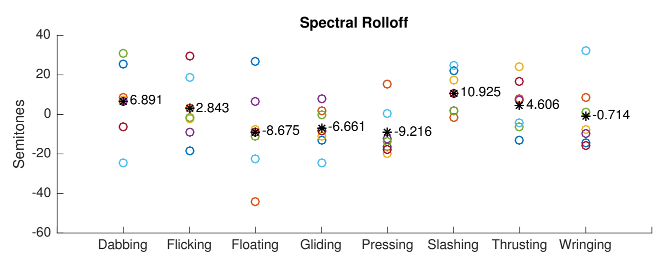

Figure 5: Spectral Roll-off for the eight Basic Effort Actions.

(which we call the *recording mean*) which are subject to further analyses.

In order to remove possible effects due to the different vocal range or musical styles of the participants – for example a female vocalist might have a higher pitched voice than a male – we subtracted from each recording mean the mean of that feature taken across all recording means for that participant. We can then compare these adjusted recording means for all participants, and we can also compare all participants' recording means between different BEAs. For example, the top of Figure 5 shows that the mean spectral rolloff for Dabbing is higher than the mean for Flicking.

As with the qualitative analysis, the recordings, each of which is associated with one of the BEAs, can be reorganized into the Effort Factors: Space Effort, Weight Effort, and Time Effort. For example, Figure 6 shows that the mean spectral rolloff for movements whose Time Effort is Sudden is higher than for movements whose Time Effort is Sustained.

In order to test whether these differences are meaningful we conducted T-tests between the two values of each Effort Factor (i.e. between Direct and Indirect Space Effort, between Strong and Light Weight Effort, and between Sudden and Sustained Time Effort). In the following sections we will discuss the significant differences we found.

4.2.1 Spectral Brightness, Roll-off, and Centroid . Spectral roll-off, spectral brightness, and spectral centroid are features which tell us about the distribution of energy in the spectrum of a sound. Roll-off calculates the frequency below which 85% of the energy lies, brightness calculates the percentage of energy above 1500 Hz, and centroid calculates the centroid of the magnitude spectrum. We converted the roll-off and centroid recording means from Hz

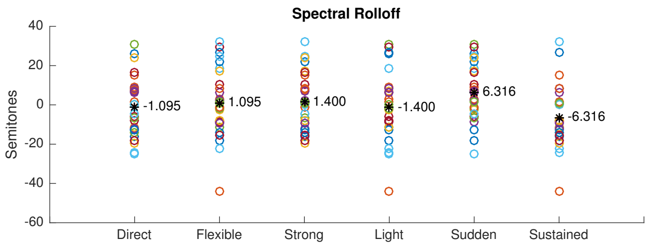

Figure 6: Spectral Rolloff for the Space, Weight, and Time Effort Factors.

into semitones to match the logarithmic scaling of frequency in human perception.

For these features, a higher result corresponds to a "brighter" sound with more energy in the high frequencies. We found that for these features, Sudden movements had sounds with higher spectral brightness, roll-off, and centroid than did sounds for Sustained movements. In other words, Sudden movements led to brighter sounds, and Sustained movements led to darker sounds. (Brightness: $t(54) = 3.2204, p = 0.0022$; Roll-off: $t(54) = 2.74995, p = 0.0081$; Centroid $t(54) = 3.1439, p = 0.0027$.) Figure 6 shows this effect for spectral roll-off.

4.2.2 Zero-Crossing Rate and Spectral Flatness. The time-domain zero-crossing rate is used in speech processing to distinguish voiced from unvoiced sound, and is considered a measurement of the "noisiness" of a signal [19]. For a periodic sound the zero-crossing rate is proportional to frequency, so we convert it to logarithmic spacing (semitones) to match human perception. Spectral flatness is also a measure of the amount of noise in a signal, since the magnitude spectrum of white noise is flat. For both of these features we found that sounds for Sudden movements had higher values (i.e. more noise) than did sounds for Sustained movements. (Zero-crossing rate: $t(54) = 2.4024, p = 0.0198$; Spectral flatness: $t(54) = 2.7636, p = 0.0078$.) Figure 7 shows this effect for zero-crossing rate.

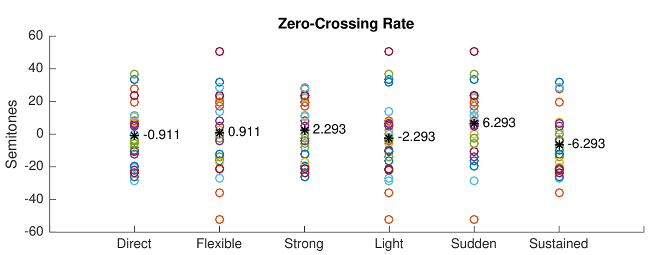

Figure 7: Zero-crossing rate for the Space, Weight, and Time Effort Factors.

4.2.3 Amplitude Envelope and Peak. The amplitude envelope is roughly analogous to the changing loudness of a sound. When we compared the recording means for amplitude envelopes we did not find any significant results. However, when we compared the peak values (in dB) of the amplitude envelopes we found a relationship with Weight Effort. Movements with Strong Weight led to sounds with higher envelope peaks than did movements with Light Weight ($t(54) = 2.2876, p = 0.0261$).

4.2.4 Entropy of Envelope and Flux Envelopes. We plotted the average of all amplitude envelopes for each Effort Factor and noticed that the average amplitude envelopes for Direct, Strong, and Sudden movements seemed to have sharp peaks near the end, whereas their complements had smoother envelopes (see Figure 8). In order to quantify this "peakiness", we calculated the entropy of all amplitude envelopes. A smoother envelope will have higher entropy, whereas an envelope that has strong peaks will have lower entropy. T-tests find an effect for Time Effort, where Sudden movements have lower entropy (more peaky) than Sustained movements ($t(54) = -4.0025, p = 1.9253e-4$). Figure 9 shows this effect for amplitude envelope entropy.

Spectral flux is a measure of the amount of change between successive spectra in a sound, and is related to how much the timbre changes. Plots of average spectral flux over time showed a similar "peakiness", and calculating the entropy of spectral flux also found that Sudden movements led to stronger peaks in spectral flux than did Sustained movements ($t(54) = -3.7444, p = 4.4045e-4$).

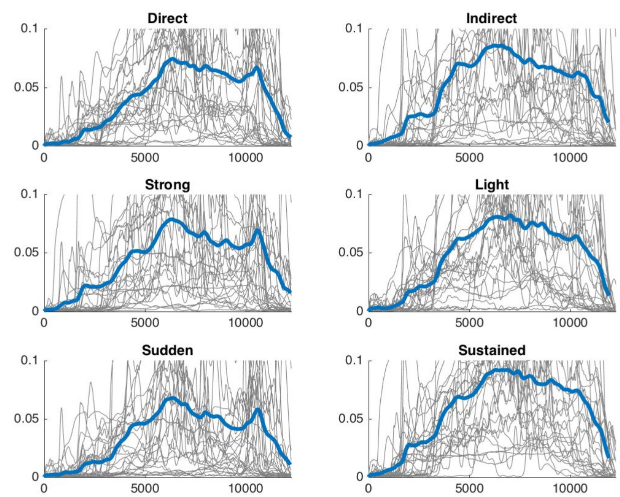

Figure 8: Average Amplitude Envelopes for each Effort Factor).

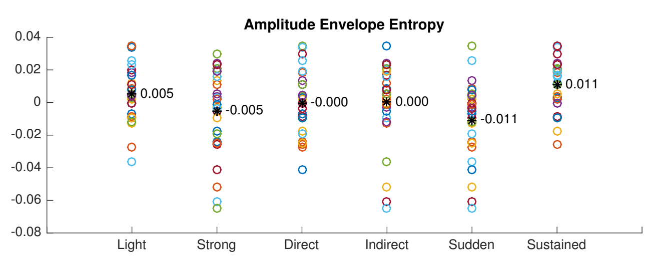

Figure 9: Amplitude Envelope Entropies. Lower values corresponds to stronger "peak" envelopes).

Table 3: P values for significant findings comparing Effort Factors to sound qualities

Feature	Effort Factor	P value	Result Summary
Amp. Envelope Entropy	Time	< 0.001	Sudden sounds tend to contain strong peaks
Spectral Flux Entropy	Time	< 0.001	Sudden sounds tend to contain peaks of intense change
Brightness	Time	< 0.01	Sudden sounds tend to be brighter
Spectral Centroid (Log Hz)	Time	< 0.01	Sudden sounds tend to contain higher frequencies
Spectral Flatness	Time	< 0.01	Sudden sounds tend to be noisier
Spectral Roll-off (Log Hz)	Time	< 0.01	Sudden sounds tend to contain more high frequencies
Zero-crossing Rate (Log Hz)	Time	0.0198	Sudden sounds may be noisier
Envelope Peak (dB)	Weight	0.0261	Strong sounds may have louder peak values

5 DISCUSSION

The significant findings from the quantitative analysis (which are summarized in Table 3) support some of the relationships we found in the qualitative analysis.

For Time Effort, both qualitative and quantitative analyses find that:

- Sudden movements are associated with brighter sounds, whereas sounds for Sustained have darker timbres.
- Sudden movements are associated with noisier sounds, whereas sounds for Sustained movements are more pitched.
- The sounds for Sustained movements have smooth amplitude envelopes and smoothly varying timbre, whereas the sounds for Sudden movements tend to have moments of strong emphasis both in amplitude and timbre.

The qualitative finding that sounds for Sudden movements are louder than sounds for Sustained movements was not confirmed by our quantitative analysis, which looked for significant differences between the mean amplitude envelopes. Neither was it contradicted.

For Weight Effort, both qualitative and quantitative analyses find that:

- Strong movements are associated with sounds that are louder and have higher peak amplitudes, whereas sounds for Light movements are quieter and with smaller peak amplitudes.

The qualitative analysis suggested that sounds for Light movements tend to be brighter and more pitched, compared to those for Strong movements which are darker and more noisy. The qualitative comparisons of brightness (via spectral brightness, roll-off, and centroid) and noisiness (via zero-crossing rate and spectral slope) neither confirmed nor contradicted this.

Qualitative analyses also suggested that Strong movements have more end-emphasis, while Light movements have more sustained amplitude envelopes and more middle-emphasis. We did not conduct a quantitative analysis that would test this finding.

It is interesting to note that none of these findings contradict what we might expect. For example, our common sense would agree that a Sudden movement should correspond to a sound with strong peaks in loudness, while a Sustained movements have smoothly varying sounds. After all, our experience is that in many acoustic instruments (e.g. bowed strings) Sudden movements, characterized by sharp accelerations, *cause* sudden changes in sound.

Regarding Weight Effort, the one qualitative finding which was confirmed by the quantitative analysis, i.e. that sounds for movement with Strong Weight have higher peak amplitude, also agrees with common experience. If we strike a drum lightly, it will not be as loud as if we strike it "strongly", transferring greater energy into the vibrating instrument.

These sound-producing gestures (bowing a violin, and striking a drum) and others can be categorized according to Godøy's three basic motion-effort types: Impulsive, Sustained, and Iterative. He claims that these correspond to three basic dynamical categories of sound, and that this correspondence is part of the basis for our *cross-modal* perception of movement and sound [6][7]. Our findings could be described using Godøy's terms. For example, movements with Sudden Time Effort are Impulsive, as are the sounds that accompany them, while Sustained Time Effort movements and their sounds are Sustained (both in motion-effort and in dynamic topological sound category). Our participants did make some sounds that might be considered Iterative, which were labeled as 'oscillating' for pitch, amplitude, and timbre shapes. But there did not seem to be distinctive differences between Effort Factors in the number of sounds labelled 'oscillating'.

We did not find any results with respect to Space Effort. Perhaps this is because the concept of space does not map very precisely onto the qualities of sounds we studied. Spatial placement of sound is a creative dimension of music composition, however our study focused on monophonic and spatially static sound. Spatial metaphors are often used in describing music and our experience of listening to music [8]. For example, we might say that one melody ascends a scale to land on the tonic (a direct movement), while another "meanders" up and down (an indirect motion). The qualitative labelling of pitch shape (see Figure 4) seems inconclusive on this point, and our quantitative analysis did not measure the shape of pitch variations.

The height metaphor commonly used to describe pitch may be one reason that sound and motion studies frequently find correspondences between vertical position and pitch. We designed our stimulus to be a primarily horizontal movement to avoid tempting our participants to focus on height and pitch, so that we might evince other relationships. Eitan and Granot found that in imagined movement responses to sound, pitch contour correlated to both vertical motion and asymmetrically to horizontal movement [4]. (In our stimulus the horizontal movement is symmetrical.) In the LBMS framework, so-called "spatial affinities" explicate relationships between Effort and another part of the system "Space". These indicate that a stimulus with movements that cross over the body midline (the vertical line of symmetry in humans) might better probe perceptions of Space Effort. More focused investigation into how Space Effort maps onto sound could be fruitful.

6 CONCLUSION AND FUTURE WORK

In this study we investigated correspondences between qualities of movement and features of sound, through qualitative and quantitative analyses of vocalizations performed by trained musicians in response to animations of simulated robotic movement. We found evidence of a number of specific relationships between movement quality – as described by the Time Effort, Weight Effort, and Space Effort of the Basic Effort Actions (BEAs) in the Laban Effort system – and the amplitude, timbre, and pitch of the resulting sounds.

We can use these findings to synthesize new sounds according to the values of parameters for Time, Weight, and Space Efforts. For example, based on our results, a sound to accompany a Wringing movement should be noisy (Indirect Space Effort), louder in amplitude (Strong Weight Effort), and have sustained amplitude envelope (Sustained Time Effort). A sound for a Dabbing movement will be more tonal (Direct Space Effort), quieter (Light Weight Effort), and contain a fast attack and short amplitude envelope (Sudden Time Effort).

Sounds generated in this way will be used in a second study to investigate people's ability to distinguish different movement qualities in robotic movement, and to determine whether adding sounds created according to the relationships found here can improve people's perception of movement quality in platforms whose dynamic capacity is more limited than human bodies.

Our study used a simple out-then-in movement of the arms in the coronal plane as stimuli. This movement lasted four seconds (corresponding to Godøy's *meso* timescale). Generating longer and more complicated multi-stage movements with controllable movement qualities is an area of ongoing research. And further investigation into qualitative movement-to-sound correspondences may be needed to consider how sounds can be varied to accompany longer and more complicated movements with distinct spatial features

ACKNOWLEDGMENTS

The study was made possible in part by the University of Virginia Data Science Institute, Office of the President, and the University of Virginia Office of Graduate and Postdoctoral Affairs. The authors also wish to thank Catherine Maguire, CMA, for movement training in the Laban Effort Factors.

REFERENCES

[1] Lin Bai, Jon Bellona, Luke Dahl, and Amy LaViers. 2016. Design of Perceptually Meaningful Quality in Robotic Motion. In *Workshop on Artistically Skilled Robots, IEEE/RSJ International Conference on Intelligent Robots and Systems.*

[2] Luke Dahl. 2016. Comparing the Timing of Movement Events for Air-Drumming Gestures. In *Music, mind, and embodiment*. Vol. 9617. Springer International Publishing Switzerland, 3–21.

[3] Alfred Effenberg, Joachim Melzer, Andreas Weber, and Arno Zinke. 2005. MotionLab Sonify: A Framework for the Sonification of Human Motion Data. In *Proceedings of the Ninth International Conference on Information Visualisation (IV '05)*. IEEE Computer Society, Washington, DC, USA, 17–23. DOI: http://dx.doi.org/10.1109/IV.2005.84

[4] Zohar Eitan and Roni Y. Granot. 2006. How Music Moves: Musical Parameters and Listeners' Images of Motion. *Music Perception* 23, 3 (2006), 221–248.

[5] Jules Françoise, Sarah Fdili Alaoui, Thecla Schiphorst, and Frederic Bevilacqua. Vocalizing dance movement for interactive sonification of laban effort factors. In *Proceedings of the 2014 conference on Designing interactive systems* (2014). ACM Press, 1079–1082. DOI: http://dx.doi.org/10.1145/2598510.2598582

[6] Rolf Inge Godøy. 2010. Gestural affordances of musical sound. In *Musical gestures: sound, movement, and meaning*, Rolf Inge Godøy and Marc Leman (Eds.). Routledge, New York, Chapter 5, 103–125.

[7] Rolf Inge Godøy, Minho Song, Kristian Nymoen, Mari Romarheim Haugen, and Alexander Refsum Jensenius. 2016. Exploring Sound-Motion Similarity in Musical Experience. *Journal of New Music Research* 0, 0 (2016), 1–13. DOI: http://dx.doi.org/10.1080/09298215.2016.1184689 arXiv:http://dx.doi.org/10.1080/09298215.2016.1184689

[8] Mark L Johnson and Steve Larson. 2003. Something in the Way She Moves – Metaphors of Musical Motion. *Metaphor and symbol* 18, 2 (2003), 63–84.

[9] Heather Knight and Reid Simmons. 2014. Expressive motion with x, y and theta: Laban effort features for mobile robots. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 267 – 273.

[10] Ksenia Kolykhalova, Paolo Alborno, Antonio Camurri, and Gualtiero Volpe. 2016. A serious games platform for validating sonification of human full-body movement qualities. In *Proceedings of the 3rd International Symposium on Movement and Computing*. ACM, 39.

[11] Rudolf Laban and F. C. Lawrence. 1974. *Effort: Economy of Human Movement* (2nd ed.). Macdonald and Evans.

[12] Olivier Lartillot and Petri Toiviainen. 2007. A Matlab toolbox for musical feature extraction from audio. In *International Conference on Digital Audio Effects*. 237–244.

[13] Amy LaViers, Lin Bai, Masoud Bashiri, Gerald Heddy, and Yu Sheng. 2015. Abstractions for Design-by-Humans of Heterogeneous Behaviors. In *Dance Notations and Robot Motion*. Springer Tracts in Advanced Robotics (STAR), 237 – 262. DOI: http://dx.doi.org/10.1007/978-3-319-25739-6_11

[14] Amy LaViers and Magnus Egerstedt. 2012. Style-based robotic motion. In *Proceedings of the 2012 American Control Conference*. Montreal, QC.

[15] Toru Nakata, Taketoshi Mori, and Tomomasa Sato. 2002. Analysis of impression of robot bodily expression. *Journal of Robotics and Mechatronics* 14, 1 (2002), 27–36.

[16] Luiz Naveda, Isabel C Martínez, Javier Dameśon, Alejandro Pereira Ghiena, Romina Herrera, and Manuel Alejandro Ordás. 2015. Cross-Cultural Comparisons of Unconstrained Body Responses to Argentinian and Afro-Brazilian Music. In *International Symposium on Computer Music Multidisciplinary Research*. Springer International Publishing Switzerland, 464–482.

[17] Kristian Nymoen, Rolf Inge Godøy, Alexander Refsum Jensenius, and Jim Torresen. 2012. Analyzing correspondence between sound objects and body motion. *ACM Transactions on Applied Perception (TAP)* 10, 2 (2012). DOI: http://dx.doi.org/10.1145/2465780.2465783

[18] Álvaro Sarasúa and Enric Guaus. 2014. Beat tracking from conducting gestural data: a multi-subject study. In *Proceedings of the 2014 International Workshop on Movement and Computing*. ACM, 118.

[19] Eric Scheirer and Malcolm Slaney. 1997. Construction and evaluation of a robust multifeature speech/music discriminator. In *Acoustics, Speech, and Signal Processing, 1997. ICASSP-97., 1997 IEEE International Conference on*, Vol. 2. IEEE, 1331–1334.

[20] Hugo Scurto, Guillaume Lemaitre, Jules Françoise, Frédéric Voisin, Frédéric Bevilacqua, and Patrick Susini. 2015. Combining gestures and vocalizations to imitate sounds. *The Journal of the Acoustical Society of America* 138, 3 (2015).

[21] Megha Sharma, Dale Hildebrandt, Gem Newman, James E. Young, and Rasit Eskicioglu. 2013. Communicating affect via flight path exploring use of the laban effort system for designing affective locomotion paths. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 293 – 300.

[22] Karen Studd and Laura Cox. 2013. *Everybody is a Body*. Dog Ear Publishing, Indianapolis, IN.

[23] Federico Visi, Esther Coorevits, Rodrigo Schramm, and Eduardo Miranda. 2016. Analysis of mimed violin performance movements of Neophytes: patterns, periodicities, commonalities and individualities. In *Music, mind, and embodiment*. Vol. 9617. Springer International Publishing Switzerland, 88–108.