A METAPHOR-BASED TECHNICAL FRAMEWORK FOR MUSICAL SONIFICATION IN MOVEMENT REHABILITATION

Prithvi Ravi Kantan [1], *Erika G. Spaich* [2], *Sofia Dahl* [1]

Department of Architecture, Design and Media Technology [1]
Department of Health Science and Technology [2]
Aalborg University, Denmark
prka,sof@create.aau.dk, espaich@hst.aau.dk

ABSTRACT

Interactive sonification has increasingly shown potential as a means of biofeedback to aid motor learning in movement rehabilitation. However, this application domain faces challenges related to the design of meaningful, task-relevant mappings as well as aesthetic qualities of the sonic feedback. A recent mapping design approach is that of using conceptual metaphors based on image schemata and embodied music cognition. In this work, we developed a framework to facilitate the design and real-time exploration of rehabilitation-tailored mappings rooted in a specific set of music-based conceptual metaphors. The outcome was a prototype system integrating wireless inertial measurement, flexible real-time mapping control and physical modelling-based musical sonification. We focus on the technical details of the system, and demonstrate mappings that we created through it for two exercises. These will be iteratively honed and evaluated in upcoming user-centered studies. We believe our framework can be a useful tool in musical sonification design for motor learning applications.

1. INTRODUCTION

In recent decades, the field of sonification has found application in the medical domain, most notably in diagnostics and rehabilitation [1]. In the latter, real-time sonification of human movement patterns, or *movement sonification*, has been increasingly explored as a form of *biomechanical biofeedback* for the enhancement of motor learning and control [2, 3, 4, 5]. Certain affordances of the auditory medium have been found to be advantageous, such as the increased bandwidth to convey bodily information with excellent temporal resolution [1]. Movement sonification can enrich the otherwise scarce auditory information from human movement by presenting kinematic or kinetic variables as auditory patterns that can aid motor learning [4, 6]. Practical real-life systems targeting physiotherapy have increasingly been developed and tested, aided by advances in sensor technology and digital computing power/tools [7, 8, 9].

So far, we have only glimpsed the potential of sonification as a rehabilitative tool. Auditory feedback has successfully been applied in motor learning of fast, repetitive tasks as reviewed in [2]. A recent review of sonification in physiotherapy found that it regularly brings about improvements in motor control, movement quality, spatial awareness and execution of complex movements, while encouraging movement and providing data relevant for physiotherapists [8]. Another review found positive short-term effects of sonified feedback in gait training [9], an area of critical importance in rehabilitation. Similar findings exist in balance training, where the portability advantage of auditory feedback devices over traditional visual systems has also been highlighted [7].

Despite this promise, most reviews have pointed out crucial shortcomings in the present state of research. These include the lack of rigorous effect evaluation through randomized-controlled testing with patients, a general absence of effect retention tests [2, 8, 9], as well as a more fundamental issue - the design of movement-sound mappings, which tends to be approached in an ad-hoc manner [10]. Difficulties in finding suitable sonic representations for different data types are not new, in fact, they have plagued the field of auditory display since its inception [1, 11, 12, 13]. But it is of great importance in rehabilitation applications, as the ability to improve one's movement patterns directly depends on receiving meaningful and relevant task-specific information [5, 14]. The type of information provided plays a key role in determining the extent of sonification utility, usability and sustained motor improvements [2, 5, 9, 14], and is thus a central design consideration. Also, movement sonification aesthetics has often received little attention, leading to auditory displays that are fatiguing and annoying to listen to for long periods [10, 11]. Display aesthetics must be interwoven into the functionality of rehabilitation applications [8, 9], especially considering the target group. Musical sounds have shown promise as a means to enhance motivation and aesthetic experiences [15, 16]. Overall, sonification designers need to explore how different sounds may be attributed to movement data to maximize motor control and learning effects [8], which essentially makes design an iterative process with extensive prototyping prior to implementation [14, 17]. This need calls for technological tools for the exploration and evaluation of sonification mappings.

We here present a framework to address the mapping and aesthetics issues in sonification for movement rehabilitation. We employed a generalizable design philosophy rooted in embodied conceptual metaphors [18, 19] and music-based interaction [15, 16], creating a practical system to design, explore and evaluate movement-sound mappings. We first motivate our approach based on present knowledge, followed by an outline of our key design tenets and an in-depth technical description of the system. Lastly, we demonstrate a set of mappings created using the system for selected training exercises.

2. RELATED RESEARCH

The goal of rehabilitation in several patient groups (e.g. stroke, traumatic brain injury) is motor re-learning by leveraging neural plasticity [14]. This process goes through several distinct phases, and feedback (intrinsic and/or extrinsic) plays a significant role throughout [5]. Movement sonification is a form of augmented extrinsic feedback [2], which can assist during both the initial cognitive phase of goal-oriented motor learning and the subsequent associative phase [5]. In the former, it can be used to highlight the already-present proprioceptive signals, thereby `educating the attention' of the learner and building sensorimotor knowledge essential for task execution [14]. In the associative phase, it can integrate with intrinsic feedback sources below the level of consciousness, creating multisensory associations that are more effective for motor learning than unimodal ones [4]. For these benefits to be realized in the context of a motor skill, task-relevant variables suitable for sonification must first be identified. This is challenging, as human movements are complex with several degrees of freedom, sonifying all of which would lead to a cognitively overwhelming feedback stimulus. As reviewed in [10, 14], this can be economized in goal-oriented training by focusing the sonification design on providing high-level information on the *end effector* of a repetitive movement (e.g. finger in a pointing task), which during desirable performance always reproduces certain patterns irrespective of the underlying interacting motor mechanisms.

User comprehension of a movement sonification hinges greatly on how the movement was converted to sound, and it is here that *mapping* and *display aesthetics* are relevant. Current literature still lacks guidelines for mapping design [2, 10, 18], although there are some general recommendations, such as limiting simultaneous data streams [12], conveying numeric values through pitch [12], temporal information through rhythmic patterning [2, 12] and mapping data streams to distinct timbres [2, 12]. Others have suggested an approach based on embodied cognition [13, 14], wherein intuitive *naturalistic mappings* are designed according to embodied associations built by humans over a lifetime of sensorimotor experience [20]. An example could be a heartbeat sonification with heart rate mapped to the tempo/repetition rate of a recorded heartbeat sound rather than just the pitch of a sine wave.

Embodied associations are formalized in terms of image schemata - blocks of thought derived from frequently-encountered sensorimotor structures [18, 19]. Image schemata structure perception and cognition through *conceptual metaphors*, a cognitive process whereby image schemata in a familiar domain of thought are used to make sense of an abstract domain [19]. In other words the brain uses one concept as a filter/extractor for another [20]. These principles can be integrated with the *Embodied Sonification Listening Model* proposed by Roddy and Bridges [19] to guide sonification design. This would dictate that a data phenomenon (e.g. stock value) is represented by a *sonic complex* (e.g. sine wave), whilst a measurement of the data phenomenon (e.g. daily price) is represented by a *dimension* of the sonic complex (e.g. sine wave pitch). The data-sound mapping is determined by a conceptual embodied metaphor (*price increase = raise pitch* based on the verticality schema) [19]. Such mappings have been found to be easier to familiarize oneself with, thus allowing for faster and more accurate comprehension of the underlying data [20]. Movement sonification can similarly represent movement phenomena and measurements through sonic complexes and mapped dimensions. With suitable metaphors, the motor learner can then readily apply familiar embodied cognition models [21] to the task of comprehending the abstract target domain of movement kinematics.

Thus framed, the mapping problem becomes one of representing movement phenomena using sonic complexes and dimensions that leverage intrinsic embodied associations. The common use of sine waves and white noise with pitch, loudness and bandwidth mappings in sonification design [10, 12] is problematic not only due to their documented psychoacoustic interactions [11, 12], but also because the signals lack contextual meaning and are aesthetically impoverished [10, 11]. They function simply as `displays of abstract data' rather than as `auditory analogues of movement' [14]. All these issues can be addressed by centering the sonification design around *music*.

Music has been suggested as a sonification medium [15, 16, 22] because it provides a universal aesthetic baseline and sonic grammar that people can effortlessly and rapidly decode even without training (as in everyday music listening) [22]. This, too, is underpinned by embodied mechanisms that enable powerful action-perception mediation [21]. Music and movement are intricately connected [15], and musical properties (melody, harmony and rhythm) are often understood in terms of spatial and physical metaphor [20]. Melody-based sonification has been found to help in structuring and sequencing timed actions, as well as recovering complex target patterns [23]. Interactive musical sonification can serve to motivate, monitor and modify movement [15], and research has shown that it affords similar motor performance to non-musical sonification [16]. Embodied music associations are hence suited to creating conceptual metaphors for sonification, and we integrate these principles as described next.

2.1. Framing the Present Work

The Embodied Sonification Listening Model [19] can be adapted to musical sonification of movement as follows. A movement phenomenon (entire movement or sub-component of a complex movement) is represented by a musical instrument (“sonic complex”), and measurements pertaining to the movement phenomenon are represented by sonic dimensions of the musical instrument (e.g. melody tone height) through conceptual metaphor mappings. In theory, multiple sub-movements within a complex movement (e.g. thigh and trunk rotations in a sit-to-stand transition) can be represented by separate musical sonic complexes (SCs) to form a coherent musical whole with perceptually distinct sonic constituents. But even with feedback designated to play a specific role (e.g. motor learning and/or music-driven motivation, endurance building) and the sonic information design space constrained for a relatively narrow application context, many unresolved design questions remain, namely:

- How many and which movement phenomena are suitable to sonify during a given training type?
- How many and which measurements within a movement phenomenon are suitable to sonify? [14, 10]
- What conceptual metaphors best represent a movement phenomenon through music?
- How to quantitatively specify each metaphor in terms of parameter mapping [1]? (mapping function, polarity, continuous/discrete nature, perceptual scaling)
- What is the impact of factors as culture [12], formal music training [11], auditory/cognitive impairment [17] and auditory display expertise [12] on familiarization time, aesthetic

appeal and usability of the feedback?

What level of exibility and adjustability do the metaphors require in order to cater to individual physical and cognitive variability within a target group [12, 17]?

The appropriate choice of metaphors and SCs is highly training-speci c, and depends upon the goals and task-related information that the patient requires (e.g. if the training goal is movement smoothness, the feedback should intuitively inform on movement intermittencies). As patients exhibit considerable motor variability, the mapping design space is conceivably vast. We argue that the process of designing a suitable set of metaphors for a movement training scenario would greatly bene t from a technological tool to rapidly create, explore and test mapping combinations. This boils down to real-time control over parameter mapping of movement features to music-based SCs. It can contain some form of 2D *mapping matrix* for topology creation (see [24]) along with other mapping controls [1] (polarity, mapping function, etc.). The present work aims to (A) suggest a conceptual metaphor philosophy to connect the physical and sonic domains in a motor rehabilitation context; (B) present the design and implementation of a technical framework for the creation, exploration and evaluation of these metaphors through parameter mapping; (C) showcase a set of example mappings we designed for two training use-cases.

3. PHILOSOPHY AND SYSTEM DESIGN

A technical framework for movement-music metaphor exploration has to encapsulate functionality for body movement capture, parameter mapping and music generation. Our present scope is limited to training activities in the form of sit-to-stand (STS), knee and trunk exercises, all of which are key foci of neurorehabilitation [25, 26]. The movement effectors are the trunk, thigh and shank segments, with the hip and knee joints serving as rotational fulcra that ex and *extend* in speci c patterns during execution. Joint discoordination and decreased movement range commonly typify impaired movement patterns in all these contexts [25, 26], and can be captured by joint/body-segment angular trajectories using wearable inertial measurement units (IMUs) [7]. The biomechanical measurement space of our framework thus purely comprises angular movement features of these body segments and joints, derived from a system of three wireless IMUs, one per body segment. These features or *movement parameters (MPs)* are good candidates for soni cation purposes as per [10, 14] as they directly represent effector activity.

We propose a series of metaphors to link the physical domain to the sonic domain as shown in Fig. 1. Most are musical adaptations of known embodied associations [1, 12, 18, 20, 27, 28, 29]. Movement progress => harmonic progress refers to the resolution of chordal tension as a movement progresses towards completion [29]. We later suggest parameter mappings to realize these metaphors as manipulations of audio synthesizer and effect controls. Ideally, one must balance sonic realism with exible parametric control over sound properties [1], which we achieve by using physically-modelled musical instrument simulations (see Table 1). In other words, MPs are mapped to signal-level *audio parameters (APs)* of physically-modelled SCs and digital effects according to the metaphors in Fig. 1.

To realize a metaphor through parameter mappings, it does not suf ce to simply assign MPs to APs. The precise manner in which the AP responds to changes in the MP has to be speci ed to ensure that the sound behavior *meaningfully* corresponds to the movement phenomenon underlying the MP. Cognitively speaking, the generated musical gestalts must be readily associable with the movement gestalts they represent [10] through promixity in time and perceived likelihood of the movement having caused that sound [21] based on internal models. Due to the complexity of creating compelling metaphorical associations from signal-level MP-AP mappings, our framework contains a con gurable mapping layer. This allows real-time control over not only the mapping topology but also the mapping function shape, polarity and discretization. A set of simultaneous metaphors, which we call a *metaphor structure*, can thus be realized through parameter mapping combinations for a given training type. The remainder of this section contains technical details of system implementation, nally showcasing some example metaphor structures along with their underlying parameter mappings.

Figure 1: A depiction of the movement-music metaphors (as the term is de ned in [18, 20]) employed in the framework design.

3.1. Hardware and Software Architecture

Figure 2: A owchart of the developed system depicting the distribution of functional elements. The body is shown in the sagittal plane, and the inclinations of the the segments ( ) are indicated. The right of the vertical is considered the positive direction.

The framework was built in a distributed architecture [5], wherein sensing, processing and feedback actuation occur at different locations as shown in Fig. 2. This was done to integrate the convenience of wireless inertial sensing technology with the processing power and interface design exibility afforded by a computer. Control is via a software application, which we developed exclusively using free, open-source software tools, namely the JUCE environment[1] in C++ and the FAUST audio programming language[2], enabling the use of essential GUI objects, timer classes and computationally ef cient audio processing libraries. Body segment movement is captured using up to three M5Stack-Grey ESP32 devices with inbuilt MPU9250 IMU chips (triaxial accelerometers, gyroscopes and magnetometers), programmable in the Arduino IDE. These devices are securely mounted to body locations corresponding to the trunk, thigh and shank using silicone casing and elastic straps. Their transmitted data is received and processed in the software at a xed callback rate of 100 Hz as shown in Fig. 2. The real-time audio callback is separate, and produces the nal sonic output at the audio device sampling rate and bit-depth (48 kHz/24-bit in our case). The full source code of the system is licensed under GNU GPL[3] 3.0

No.	Sonic Complex	Synthesis Method	Sonic Dimensions
1	Djembe	Physical Model	Tone Height Strike Sharpness
2	Marimba	Physical Model	Tone Height Strike Sharpness
3	Singing Voice	Physical Model	Tone Height Vowel Shape Fricative Noise
4	Piano Chords	Subtractive	Tone Height
5	Guitar	Physical Model	Tone Height Stiffness Dynamics
6	Warning Bell	Physical Model	-

Table 1: A summary of sonic complexes provided in the framework and their mappable sonic dimensions, aside from a trigger control. There are two separate dimensions `detune' and `frequency warping' for providing negative feedback.

3.2. Movement Measurement

3.2.1. Data Capture

The M5Stack devices periodically capture and transmit instantaneous inertial sensor readings as Open Sound Control packets over a local WiFi network to separate remote UDP ports. This occurs at a rate of 125 Hz (higher than the 100 Hz software callback rate to compensate for UDP packet drops). Depending on the training, not all three devices may be necessary. E.g. for trunk training, only one device is suf cient. The software application has a dedicated sub-interface for sensor IP con guration, body segment assignment, bias calibration and connection status monitoring (battery status and connection health). New messages received at local UDP ports are divided into separate streams for each IMU axis. Once connected, each sensor undergoes stationary calibration for 10 seconds to determine axis-wise bias. This is compensated in subsequent data packets, which are then pre- ltered to interpolate over dropped packet intervals. 3-point median lters and 2nd order IIR Butterworth LPFs with a 50 Hz cutoff frequency were empirically chosen.

3.2.2. Movement Analysis

The preprocessed IMU data are converted into an array of MP measurements. This involves estimating body segment inclination along with higher-order descriptors related to smoothness, joint angles and movement progress. The software has a sub-interface to de ne the placement (front/side) and directional polarity of each sensor, as well as its angular movement range of interest in each direction in the horizontal plane. This provides versatility in terms of possible use-cases, permissible sensor placements and feedback sensitivity to angular position changes. Joint angle ranges and hyperextension thresholds are similarly con gurable. Optionally, movement time series data can be logged for further analysis and these logs can later be streamed by the software in real-time to recreate (and re-sonify) the movement patterns they contain.

3.2.2.1. Angle Estimation

The rst step is estimating body segment inclination from the raw inertial values. This is done using the Madgwick Gradient Descent algorithm [30], a motion tracking method meant for rehabilitation systems. In short, it provides good estimation accuracy (1.7 dynamic RMS error) at a low computational cost without requiring a high data sampling rate like traditional Kalman lters [30], and can be used with or without magnetometer readings. The algorithm employs a quaternion representation, which when computed can be converted to Euler angles about the pitch ( ), roll ( ) and yaw ( ) axes. 6 and 6 are of present interest as they respectively represent anteroposterior (forward-backward) and mediolateral (left-right) inclinations [30]. As 6 (absolute directional facing) is not needed and magnetometer readings are typically subject to distortions, only 6 and 6 are computed and stored for all connected sensors. The algorithm has only one hyperparameter - gradient descent learning rate. We found the optimal value (0.033) recommended in [30] to adequately balance the trade-off between gyroscopic drift and gradient descent overshoot.

3.2.2.2. Higher-Order Parameters

Angular Jerk: Jerk is a simple measure of movement smoothness [31] de ned as the derivative of acceleration (typically linear). We chose to use the norm of angular jerk by double differentiation of angular velocity readings from the gyroscope. Unlike linear accelerometer readings, this has the advantage of not being affected by gravitation, while still capturing movement intermittencies.

Joint Angles, Velocities and Hyperextension: Hip/knee exion and extension angles are derived from the computed inclinations of adjacent body segments. These angles are then compared to con gurable joint hyperextension thresholds, according to which hyperextension ag parameters are set. Respective joint angular velocities are also computed by single differentiation of joint angles.

STS (sit-to-stand) Phase: This parameter represents movement progress during STS training by estimating what phase of the movement the patient is presently in (Steady Sitting, Stand Onset, Seat Off, Steady Standing, Sit Onset, Seat On). This is done by monitoring trunk and thigh 6, and compar-

[1] JUCE - https://juce.com/
[2] FAUST Programming Language - https://faust.grame.fr/
[3] https://github.com/prithviKantanAAU/ICAD2021

Figure 3: A step-by-step ow diagram of the mapping process, indicating how each GUI element of the mapping matrix interface is linked to feedback variable (FV) computation.

ing these to pre-de ned horizontality and verticality criteria as well as the previously known phase estimate in a series of logical operations. Movement phase is coded as a number between 1 and 6, appearing as a cyclic staircase waveform.

3.3. Parameter Mapping and Metaphor Creation

The mapping layer links the MP array to the audio synthesis functionality, which was programmed in FAUST and compiled as a JUCE-controllable optimized C++ class. It has a total of six SCs whose sound triggers and sonic dimensions are accessible through an array of APs (audio parameters) representing instrument triggering, tone height, timbre and FX parameter values as shown in Table 1. This list provides melodic, harmonic and percussive options with signal parameters perceptually correlated with the sonic domain in Fig. 1. Their frequency ranges and spectral balances were adjusted so that they are tonally distinct enough to facilitate perceptual stream segregation if presented concurrently [12].

The details of the real-time mapping layer can be completely speci ed using the *mapping matrix* sub-interface (see left part of Fig 3). The MPs have different value ranges and signal characteristics, as do the APs (e.g. trigger value bounds are 0-1, frequency parameters are 50-1000 Hz). The underlying auditory dimensions of pitch, timbre and dynamics also exhibit varied, non-linear perceptual scaling. To accommodate for these types of variability, an intermediary transformation is needed - the mapping layer converts MP values to separate *feedback variable (FV)* values compatible with each of the APs they are mapped to. This process is based on the scaling of motion data for soni cation described in [32]. Once complete for all mapped parameters, FV values are written to the corresponding FAUST controls.

FV Calculation: This is illustrated in Fig. 3. The APs are considered serially. If no MP (mapping matrix row) is mapped to an AP (column), it is assigned a prede ned default FV value and the next AP is considered. If only one MP is mapped, then it is entirely responsible for the behavior of that AP - thus the extreme MP values will map to extreme AP values. If the MP value exceeds a threshold, it is normalized to the 0-1 range, preampli ed by a mapping gain factor and transformed by the selected mapping function. Next, the FV polarity is inverted if necessary and can undergo 0-5 bit quantization (2-32 discrete levels, 0 for continuous mappings), should the AP input require this. The nal 0-1 ranged FV is normalized to the range of the AP, and mapped to its FAUST control. If multiple MPs are mapped to a single AP, a weighted sum of their individual normalized contributions (by mapping gain) is used. The mapping layer state can be stored and recalled in the form of *mapping presets*.

Tailoring FV Behavior to Various AP Types: Although the APs are treated equally at the mapping matrix level, they pertain to very different sonic properties and, as such, require speci c types of input FV signals to function correctly. Consider APs responsible for sound triggering; these trigger sound whenever a change in the FV occurs. Hence if a continuously-varying MP (e.g. trunk θ) is mapped to this, the sound (e.g. singing voice) will be triggered at the callback rate (100 Hz ≈ 100 notes per second), which is unnatural. However, if a 3-bit quantization is applied to the FV prior to mapping, only *eight notes* (2^3) will be triggered as the trunk traverses its entire θ range - a more usable sonic outcome. But a different AP - such as *vowel shape* would conversely sound very unnatural if quantized like this, and would bene t from a continuous mapping allowing for smooth formant shifts [1].

A one-to-many topology can map trunk θ to voice trigger (quantized FV) and vowel shape (continuous FV). The angle value is audible in the vowel shape, whilst sound triggering occurs with

Metaphor Structure	Effector Mov. - Musical Sound	Mov. Energy - Mus. Dynamics	Effector Pos. - Tone Ht.	Joint 6 - Vowel/String	Undesired Mov. - Undesired Snd.	Intermittency - Glitch
STS as Percussion	Trunk 6 + Djembe Trigger Thigh 6 + Marimba Trigger	Hip Angular Velocity + Djembe Strike Sharpness	Trunk 6 + Djembe Tone Ht. Thigh 6 + Marimba Tone Ht.	N/A	N/A	N/A
STS as Singing Voice	Trunk 6 + Voice Trigger	N/A	Trunk 6 + Voice Tone Ht.	Knee 6 + Vowel Shape	Trunk 6 + Melody Detune Hip Hyperextension + Bell Warning	Trunk Angular Jerk + Freq. Warping
STS Progress as Harmonic Progress	STS Phase + Gtr, Piano Trigger	N/A	STS Phase + Gtr, Piano Tone Ht.	N/A	Trunk 6 + Melody Detune Hip Hyperextension + Bell Warning	Trunk Angular Jerk + Freq. Warping
Knee Extension as Guitar	Knee 6 + Gtr Trigger	Knee Angular Velocity + Gtr Pluck Dynamics	Knee 6 + Gtr Tone Ht.	Knee 6 + Gtr Stiff	Knee Hyperextension + Bell Warning	N/A

Table 2: A summary of how the conceptual metaphors in Fig. 1 (columns) are combined into four example metaphor systems for STS and knee training (rows). The table shows how each metaphor is realized in terms of low-level MP-AP mappings (ref Sec. 3.6 for a more detailed explanation).

changes in the angle. Faster changes cause more rapid sound triggering, conforming to the rate-tempo metaphor in Fig. 1. Adding a voice tone height mapping can create a melodic note sequence; this AP needs an extra computational step (explained next).

3.4. Melody and Harmony Control

Melody- and harmony-based SCs (singing voice, piano chords, guitar) require precise musical frequency information. FVs routed to tone height APs are processed prior to mapping (not shown in Figure 3) to convert their normalized FV values (0-1 range) to musical frequency values according to the chosen key, scale/mode and chord type, all con gurable in a sub-interface.

FV values are quantized to degrees of a diatonic or pentatonic scale, such that the 0-1 range maps to about two octaves for melody instruments, and one octave for chords. This scale degree is converted to an offset in semitones relative to the scale tonic note using look-up tables, yielding a MIDI note number which is converted to a frequency in Hz. For chords, scale degree offsets are added to the root note degree to obtain four note degrees depending on the chosen chord type (e.g. for a seventh chord, the scale degree offsets added are 0 (root note), 2 (third), 4 (fth), 6 (seventh)) and four frequencies are thus obtained. This system makes a wide range of melodic and harmonic patterns available for experimentation.

3.5. Audio Synthesis and Mixing

The FAUST audio object takes the nal FV values as input, and outputs a stereo audio signal to the audio buffer callback of the JUCE component. Within FAUST, the FV values from the control channels are routed as signals to the input parameters of synthesizer functions corresponding to the six SCs, as well as some audio effects. Physical models (see Table 1) were directly used from FAUST standard libraries, whereas the polyphonic Piano Chords SC was synthesized by a custom algorithm using layered pulse waves and time-varying lters. The synthesized SCs are treated like audio tracks in a music mix, and undergo basic post-processing to optimize sound quality. To be precise, each SC is passed through a dynamic range compressor and four-band parametric equalizer with custom settings, followed by equal-power panning and a reverb aux send. The SC stereo pairs are summed in a master section with UI-con gurable track faders and a master limiter. All processing algorithms were written using FAUST library functions.

3.6. Mappings for Training Exercises

Devising mappings for any use-case context context is a matter of (A) listing training goals, key movement phenomena and designating MPs to capture them, and (B) linking these MPs to the APs best suited to communicating them through conceptual metaphors (see

Fig. 1). We did this empirically for two exercises (STS and Knee Extension), which culminated in the four metaphor structures in Table 2. Movement-relevant MPs in each case were designated as follows:

STS: We considered angular trajectories of the trunk and thigh segments as representations of effector movement in the sagittal plane, with the STS Phase MP describing movement progress. The mediolateral trunk angle represents undesirable side-to-side displacement and trunk angular jerk brings out movement intermittencies. We used the hip angle to determine joint hyperextension, and the hip angular velocity to represent movement energy.

Knee Extension: The knee angle captures the effector trajectory in the sagittal plane, in which joint hyperextension can be defined. Knee angular velocity is used to represent movement energy.

As seen in Table 2, we have avoided engaging all eight metaphors in Fig 1 simultaneously, so as not to overload the motor learner with information. The four metaphor structures are demonstrated in the Supplementary Videos [33] The `Movement Rate =) Music Tempo' metaphor is implicit in all four due to the quantized trigger mechanism described earlier. These mapping designs will serve as a starting point in future studies.

4. DISCUSSION AND CONCLUSIONS

We have presented a music-based framework for movement sonification founded in embodied conceptual metaphors [18, 20], designed and built to address current issues related to mapping and aesthetics in movement sonification for rehabilitation. We defined a set of conceptual metaphors encompassing relevant aspects of movement, based on known embodied associations [1, 12, 18, 20, 27, 28, 29], and built a practical prototype system to explore these metaphors.

Our mapping layer and supplementary functionality provide a high level of flexibility for mapping experimentation, and our audio synthesis functionality gives the designer an array of melodic and percussive options to create musical mappings for a chosen exercise. The MPs, mapping controls and low-level APs make it possible to exert precise control over movement-sound relations, and tailor the metaphors to situational needs. However, our prescripted architecture only has a limited set of mapping function shapes, and further customization would be difficult without additional controls, which could make the interface more cumbersome. We also noted that the repetitive nature of certain cyclic training movements leads to repetitive musical sequences which may not be very pleasant for longer listening, considering the importance of novelty and expressiveness in musical biofeedback [15] - something to be addressed in future versions.

With a formal evaluation pending due to the pandemic, key aspects remain unclear at present. Our metaphors account for several movement descriptors and undesired movement characteristics, but studies involving clinicians and patients must ascertain whether they sufficiently represent the movement space to assist motor learning in our application domain. Also, future testing is required to assess how faithfully our MPs and APs represent and capture the corresponding movement phenomena and sonic features listed in Fig. 1.

Such an evaluation can, for example indicate whether it is better to realize STS metaphor structures using joint angles and center-of-mass trajectories [34] rather than body segment angles. Also, we use joint angular velocity to represent `movement energy', but it is possible that electromyographic signals or force/momentum measures may be more appropriate for this [1]. Similar considerations may apply to knee training, and we generally foresee a necessity to iteratively experiment with movement representations to sonify. With this in mind, our framework design is easily scalable; adding new movement measures is a simple matter of expanding the arrays of motion sensors and MPs. Preset mapping combinations can also be defined, stored and recalled.

We plan to use the framework extensively in future studies, starting by creating a larger preliminary set of metaphor structures, and initially evaluating them through expert interviews with clinicians. After any necessary redesign, we will assess how effectively the metaphor structures convey information in an intermodal discrimination study [6] to gauge how well participants can identify distinct movement patterns using different metaphor structures. Finally, we will carry out randomized studies to formally investigate motor learning effects of the most promising mappings in well-defined tasks, similar to [4, 10].

Overall, we expect that our framework can serve as a valuable technological tool to design and explore movement sonification mappings. Its scalable structure allows parameters to be added and modified to create varied metaphors, and the workflow it affords is well-suited to iterative sonic information design [14, 17, 19]. Hopefully, the metaphor-based philosophy and exercise-specific mappings we presented can serve as a starting point for rehabilitative musical sonification design in the future.

5. ACKNOWLEDGMENT

Author PRKA was main responsible for system development and manuscript writing, under the supervision of authors SD and EGS who also assisted in writing. All authors approved of the final manuscript. The work is partially funded by NordForsk's Nordic University Hub, Nordic Sound and Music Computing Network NordicSMC, project number 86892.

6. REFERENCES

[1] T. Hermann, A. Hunt, and J. G. Neuhoff, *The Sonification Handbook*. Logos Verlag Berlin, 2011.

[2] R. Sigrist, G. Rauter, R. Riener, and P. Wolf, "Augmented Visual, Auditory, Haptic, and Multimodal Feedback in Motor Learning: A Review," *Psychonomic Bulletin & Review*, vol. 20, 11 2012.

[3] O. Giggins, U. McCarthy Persson, and B. Caulfield, "Biofeedback in Rehabilitation," *Journal of Neuroengineering and Rehabilitation*, vol. 10, p. 60, 06 2013.

[4] A. O. Effenberg, U. Fehse, G. Schmitz, B. Krueger, and H. Mechling, "Movement Sonification: Effects on Motor Learning Beyond Rhythmic Adjustments," *Frontiers in Neuroscience*, vol. 10, p. 219, 2016.

[5] A. Kos and A. Umek, *Biomechanical Biofeedback*. Springer, 09 2018.

[6] P. M. Vinken, D. Kröger, U. Fehse, G. Schmitz, H. Brock, and A. O. Effenberg, "Auditory Coding of Human Movement Kinematics," *Multisensory Research*, vol. 26, no. 6, pp. 533–552, 2013.