MuSS-Bits: Sensor-Display Blocks for Deaf People to Explore Musical Sounds

Benjamin Petry[1] bpetry@acm.org
Thavishi Illandara[2] thavishi@ahlab.org
Suranga Nanayakkara[1] suranga@sutd.edu.sg
[1] Singapore University of Technology and Design, Singapore
[2] University of Moratuwa, Sri Lanka

Figure 1. MuSS-Bits sense sound from (a) an instrument, (b) a digital device, or (c) the environment and translate it to (d) visual and (e) vibrotactile feedback.

ABSTRACT

Hearing loss makes learning a musical instrument a challenging task. Prior work suggests that a universal sensory substitution system that works uniformly across all deaf users may not exist given the diversity within the deaf community. In this paper, we present Music Sensory Substitution (MuSS) Bits, wireless sensor-display pairs that enable exploration of musical sound as well as customization of visual and vibrotactile feedback to cater to individual requirements and preferences. MuSS-Bits are portable, easy to deploy on the user's body, on an instrument, or in the environment, and provide real-time feedback. We review existing music sensory substitution systems, discuss the design space for MuSS-Bits, present details of a prototypical implementation and illustrate interaction possibilities including initial user reactions.

Author Keywords

Music; Sensory Substitution; Learning; Assistive Technologies; Deaf

ACM Classification Keywords

K.4.2 Social Issues: Assistive technologies for persons with disabilities

INTRODUCTION

Making music has, compared to listening to music, certain advantages as it is *"more powerful..., transformative..., and a way to express yourself"* (Machover, 2008). While listening to music is basically the *interpretation* of sensations, making music requires the user to actively *create* content. Hence, it requires a closed feedback loop to compare and evaluate the created sound with the intended sound. This becomes a challenging task for those with hearing disabilities who are nevertheless interested in learning to play an instrument. Due to limited access to the auditory channel, deaf people have less information to evaluate their performance.

"It is obvious that not all hearing impaired people will be musical in its fullest sense. But, then neither are all hearing folk. What is needed is the opportunity to experiment in order to discover what musical abilities lie dormant in us." (Fawkes, 2006). Prior work developed educational approaches for teaching music to deaf children (Hagedorn, 1992; Hash, 2003; May, 1961) and initiatives, such as "Music and the Deaf" (MATD, 2015), aim to encourage deaf people to make music through seminars, concerts and workshops. Moreover, the HCI community explored assistive technology using visual (Fourney and Fels, 2009; Mori and Fels, 2009; Pouris and Fels, 2012; Zhou et al., 2012) and vibrotactile (Karam et al., 2009a; Karam et al., 2008; La Versa et al., 2014; Nanayakkara et al., 2012; Nanayakkara et al., 2009; Palmer, 2016) sensory substitution systems to bridge the feedback loop gap for musical activities. We propose that a music-making sensory substitution system has to provide the opportunity to explore sound and to customize feedback. Exploration is an important part of learning (Jr, 2011; Medina, 2011) and music-learning in particular. Given the diversity within the people with hearing difficulties (Clark, 1981), feedback should be customizable by the user to cater to individual requirements and preferences (Shinohara and Tenenberg, 2009).

In this paper, we present Music Sensory Substitution (MuSS) Bits, wireless sensor-display pairs. MuSS-Bits were developed to support (1) *exploration of sound* through real-time feedback from various audio sources, such as an instrument (see Figure 1a), a digital device (see Figure 1b) or from the environment (see Figure 1c); and to support (2) *customization of feedback* through spatial deployment of Display-Bits (on the body of the user, on an instrument or in the environment), as well as

OzCHI '16, November 29-December 02, 2016, Launceston, TAS, Australia

DOI: http://dx.doi.org/10.1145/3010915.3010939

the selection and calibration of the output-modality, such as visual (see Figure 1d) and vibrotactile (see Figure 1e) feedback. Though MuSS-Bits can potentially provide information about e.g. pitch, or melody, we focused on rhythm as a starting point, since rhythm is part of the very first exercises when teaching music to deaf children (Fawkes, 2006).

The main contributions of this paper are as follows:

- We present a **literature review and comparison of existing music sensory substitution systems** for deaf people. We analyzed these systems among three dimensions: support for (1) sound exploration and (2) feedback customization as well as (3) the feedback modality used.
- We discuss **design goals** for music sensory substitution systems, derived from the areas of (1) hearing loss, (2) music-making and (3) learning.
- We provide technical details of the **MuSS-Bits implementation**. MuSS-Bits are implemented as standalone, tangible bits (see Figure 1) that communicate peer-to-peer.
- We envision four **possible interaction scenarios** and report on actual interactions of deaf users with MuSS-Bits.

RELATED WORK

Deafness does not prevent people from music-listening or music-making. Deaf musicians, such as percussionist Evelyn Glennie (Glennie, 2003), opera singer Janine Roebuck (Roebuck, 2015) or rapper and songwriter Sean Forbes[1], are extraordinary individuals that made music their profession. Nevertheless, a deaf person's interest in music depends on his or her affiliation with the hearing or deaf culture (Darrow, 1993). Darrow found that deaf students reported to feel uncomfortable in mixed music classes with their hearing peers as they were punished for sounding not right (Darrow, 1993).

There is a growing community that works on teaching music to deaf people. "Music and the Deaf" (MATD, 2015) offers seminars, workshops and concerts for deaf people and provides teaching material for music instructors. Hagedorn (Hagedorn, 1992) describes an approach aiming to improve sound perception by using Sanders' hierarchy of auditory processing. She describes incremental musical activities for deaf people to train their residual hearing in passing each auditory processing level of the hierarchy.

A set of instructions for teaching music to hearing impaired children and teenagers is given by William G Fawkes (Fawkes, 2006) and consists of three stages. The first stage addresses the stabilization of a steady-beat. A two-beat rhythm is introduced to the children through walking, swaying or hand clapping movements and later through the use of the voice (singing vowels). In the second stage, basic music notation is introduced and trained using tambourines, triangles, shakers and small drums. The last stage is concerned with tuned instruments, such as recorders, melodica and glockenspiel, as well as tone notation. Fawkes noted that the students always maximize the use of their residual hearing throughout the exercises. Hearing aids are one way to improve residual hearing. However, hearing aids and also cochlea implants are known to distort music which affects its enjoyment (Chasin, 2003; Drennan and Rubinstein, 2008; Fourney, 2012; Galvin et al., 2009; Limb, 2011).

MUSIC SENSORY SUBSTITUTION SYSTEMS

"Sensory substitution devices (SSDs) convey information that is normally perceived by one sense, using an alternative sense." (Levy-Tzedek et al., 2012). Sensory substitution applications for deaf people include sound type detection (Ho-Ching et al., 2003; Matthews et al., 2005), direction cueing (Jain et al., 2015; Tan et al., 2003), speech processing (e.g. Tadoma method (Reed et al., 1982)) and enhancement of musical activities. One of the main challenges of sensory substitution systems is to design an intuitive mapping (Karam et al., 2009b). For example, there is little overlap of the brain's processing of visual and audio information (Bertini et al., 2010). Processing of the audio and vibrotactile sensory channels are more overlapping. However, the skin's perceivable frequency range is quite small (up to 1000 Hz) with a peak frequency at 250 Hz, compared to the auditory channel, that can sense frequencies between 20 Hz and 20 kHz (Stumpf, 1883; Verillo, 1991).

We reviewed existing music sensory substitution systems to inform our design of the MuSS-Bits (see Table 1). To avoid confusion about terminology, the terms *frequency* and *intensity* are used with reference to vibrotactile feedback, whereas *pitch* and *loudness* are used to refer to auditory feedback. Furthermore, the terms x-, y- and z-axis describe the horizontal, vertical and depth representation of an object on a screen.

We compared these systems among three dimensions: support of (1) sound exploration, (2) feedback customization and (3) the modalities supported. As seen from Table 1, most systems support the exploration of sound partially and some even to a full extent. We found 2 systems that provide partial customization, but no system providing full customization. Moreover, most systems use either visual or vibrotactile feedback.

We further analyzed the mappings of these systems. We found 2 main types of time representations across all systems: (1) along an axis, such as the x-axis of a screen or (2) as instantaneous events. The first allows the user to see past and future events of sound and music in particular. The second representation displays only present information, which is closer to the human listening process that does not have access to past or future events (except for auditory memory). All vibrotactile approaches applied the instantaneous time representation. While vibrotactile approaches map pitch, loudness, timbre and time in very similar ways, visual approaches use very different mappings. For example, 2 systems use color for pitch, but 3 systems use it for timbre. These differences may be attributed to the little overlap of visual and auditory processing as previously

---

[1] http://www.deafandloud.com/

Music Sensory Substitution System		**Year(s)**	**Exploration**	**Customization**	**Feedback Modality**	**Focus**	**Display Technology**	**Mapping (Selected Musical Elements)**			
								Pitch	*Loudness*	*Timbre*	*Time*
EnAct (Mori and Fels, 2009; Rashid et al., 2006; Vy et al., 2008)	🦻	2006 - 2009	○	○	👁	🎧	Screen	*	*	*	*
Piano Roll View (Fourney and Fels, 2009; Isaacson, 2005; MAM, 2016)	🦻	2005 - 2009	◐	○	👁	🎧	Screen	Y-axis		Color	X-axis
Part Motion View (Fourney and Fels, 2009; MAM, 2016)	🦻	2009	◐	○	👁	🎧	Screen	Y-axis		Color	X-axis
Tonal Compass View (Fourney and Fels, 2009; MAM, 2016)	🦻	2009	◐	○	👁	🎧	Screen	Angle	Size		Instantaneous
Multimedia Visualizer (Fourney and Fels, 2009; iTunes, 2016)	🦻	2009	◐	◐	👁	🎧	Screen	Mostly arbitrary	Mostly arbitrary		Instantaneous
Motion Pixels of Music (Fourney and Fels, 2009)	🦻	2009	◐	○	👁	🎧	Screen	Angle			In-/outwards movements
MusicViz (Pouris and Fels, 2012)	🦻	2012	◐	○	👁	🎧	Screen	Y-axis	Size + Brightness	Color + Shape	Z-axis
MOGAT (Zhou et al., 2012)	🦻	2012	●	○	👁	🎸	Mobile Phone	Y-axis			Instantaneous
Movies from Music (Mitroo et al., 1979)		1979	○	○	👁	🎧	Screen	Color + Brightness			Distance from center
Seen Music (Kim et al., 2015)	🦻	2015	◐	◐	👁	🎸	Tangible Objects	Color			Instantaneous
Spectrogram (Isaacson, 2005)		2005	●	○	👁	🎧	Screen	Y-axis			X-axis
CAMLS for hearing-impaired (Yang et al., 2007)	🦻	2007	●	○	👁	🎸	Screen	Written text, Notation			Position inside the notation
Seeing Sound (Ferguson et al., 2005)		2005	●	○	👁	🎸	Screen	Angle	Height		Instantaneous
Music that Moves (Music that Moves, 2016)	🦻	2016	●	○	👁 ☜	🎧	Mobile Phone		#Objects/ Intensity		Instantaneous
Haptic Chair (Nanayakkara et al., 2012; Nanayakkara et al., 2009)	🦻	2009 - 2012	◐	○	👁 ☜	🎧	Screen + Surface Transducers	Y-axis + Size/ Frequency	Color + Brightness/ Intensity		Instantaneous
MUVIB (La Versa et al., 2014)	🦻	2014	◐	○	☜	🎧	ERM Motor		Intensity		Instantaneous
Tactilicious Flute Display (Birnbaum and Wanderley, 2007; Marshall and Wanderley, 2006)		2006 - 2007	◐	○	☜	🎧	Voice Coil	Frequency	Intensity		Instantaneous
Tac-Tile Sounds (Palmer, 2016)	🦻	1994	◐	○	☜	🎧 🎸	Speaker	Frequency	Intensity		Instantaneous
Emoti-Chair (Frequency Model) (Karam et al., 2010; Karam et al., 2009a; Karam et al., 2008)	🦻	2008 - 2010	◐	○	☜	🎧	Voice Coil	Frequency + spatial location	Intensity		Instantaneous
Emoti-Chair (Track Model) (Karam et al., 2010; Karam et al., 2009a; Karam et al., 2008)	🦻	2008 - 2010	◐	○	☜	🎧	Voice Coil	Frequency + Spatial location**	Intensity	Spatial location	Instantaneous
Emoti-Chair (Control Model) (Karam et al., 2010; Karam et al., 2009a; Karam et al., 2008)	🦻	2008 - 2010	◐	○	☜	🎧	Voice Coil	Frequency	Intensity		Instantaneous
VibroChord (Branje and Fels, 2014)		2014	◐	○	☜	🎸	Voice Coil	Frequency + Spatial location	Intensity		Instantaneous
Mobile Music Touch (Huang et al., 2010)		2010	◐	○	☜	🎸	ERM Motor	Spatial location			Instantaneous
MuSS-Bits	🦻		●	◐	👁 ☜	🎸	LED + ERM Motor		Brightness/ Intensity		Instantaneous

* conveys emotions of music (happiness, sadness, fear and anger) instead of certain musical elements

** the track model maps the instrument to a spatial location and the pitch for each instrument is located around this location

🦻 designed or evaluated with deaf or hard of hearing individuals

Feedback Modality: 👁 Visual Feedback ☜ Haptic Feedback Focus: 🎧 Music-Listening 🎸 Music-Making

Evaluation Criteria:	● Fully supported	◐ Partially supported	○ Not supported
Exploration	The system supports **different audio sources** (e.g. instruments or digital devices) and provides **real-time feedback.**	The system supports either different audio sources or provides real-time feedback.	The system does not support both.
Customization	Feedback, the **mapping** and **presentation** (e.g. spatial location or modality calibration), can be customized by the user.	Either the mapping or the feedback presentation can be customized by the user.	Both are fixed and cannot be customized by the user.

Table 1. Overview of visual and vibrotactile sensory substitution systems for music.

mentioned. Further, such little overlap makes it challenging to find an intuitive mapping. A multimodal approach might support visual mappings and can improve the overall sensing resolution.

It is clear that various proposed sensory substitution systems to support music-making has not become the *de facto* standard among the deaf community. The reason behind this could be, as Russ Palmer, a deaf blind international music therapist, suggests: *"a music system needs to be **adaptable** and **simple** to use otherwise the user will end up storing it away in a cupboard, which will not be used again."*[2]

DESIGN GOALS

Sound exploration provides an understanding of the relationship between action (such as hitting a drum) and feedback (such as vibration on the wrist). Feedback customization enables a user to create a feedback that is meaningful to him or her. In this section, we derive basic design goals for music sensory substitution systems for deaf users from the areas of (1) hearing loss, (2) music-making and (3) learning.

Learning through Exploration: Exploration is an important part of learning (Jr, 2011; Medina, 2011). Hence, a music-making system, which involves learning, has to provide an opportunity for exploration of sound. This includes being able to capture audio from different audio sources, such as instruments, digital devices or from the environment, and to provide real-time feedback.

Self-Learning: Learning through exploration aims to support the development of a conceptual model of sound in a deaf user. Self-Learning goes further and enables a deaf user to independently learn a musical instrument, for example through online tutorials[3] [4] [5]. However, online tutorials are less accessible for a person with hearing loss, since visual subtitles and visible interactions with the instrument convey only limited information. Additional feedback about the sound could enhance the self-learning process.

Support for Rhythm: Establishing a steady-beat is one of the first exercises in teaching musical instruments to deaf children, which then is followed by the introduction to rhythm (Fawkes, 2006). This makes rhythm support a fundamental requirement of a music sensory substitution system.

Free Limb Movement: Music-making is an activity that often requires a musician to move his or her body (hands, legs, mouth etc.) in a specific way, depending on the instrument being played. Hence, a music sensory substitution system should not restrict the freedom of limb movements.

Simple to Operate: As mentioned by Russ Palmer, it is important to make the interaction with the system intuitive and simple. Since we suppose this system to be used mainly by non-expert users who start learning an instrument, this becomes a very important requirement.

Customization of Feedback: Since there is a significant diversity of hearing conditions and personal preferences within the deaf community, a music sensory substitution system should allow for customization. We differentiate two ways of feedback customization: (1) customization of the audio-to-modality mapping and (2) customization of the properties of the feedback presentation. The later one includes the selection of the output modality, calibrating it to a comfortable level and choosing the spatial location (on body, on an instrument, or in the environment) to receive the feedback.

MUSS-BITS

MuSS-Bits consist of sensor-display pairs. The Sensor-Bit captures sound from various sound sources and transmits it to the Display-Bit, which translates it into vibrotactile and visual feedback. The Sensor- and Display-Bits communicate peer-to-peer via WiFi to ensure fast audio transmission for real-time feedback. We used ESP8266-12F modules to establish the WiFi communication. The ESP8266-12F comes with an on-board processor running at 80Mhz which was used for processing and translation of the audio information. The ESP module has an on-board analog-digital converter (ADC) with a sampling rate of 200 Hz, but we decided to use an external ADC (AD7991 - 12bit resolution and I2C communication) with a sampling rate of 140 kHz for a better audio signal.

MuSS-Bits are powered by a 3.7V Polymer Lithium Ion Battery (400mAh). The Sensor-Bits have a power consumption of 80mA giving them approx. 5 hours of continuous operation. The Display-Bits consume between 90mA (talking into the microphone) up to 260mA (constant maximum signal) resulting in approx. 1.5h up to 4h of continuous operation. MuSS-Bits are enclosed in rectangular shaped plastic cases (5cm x 3cm x 5cm). To identify a pair, we color coded the casings in the same color. The weight of the current prototype (Sensor-Bit is 65g; Display-Bit is 70g) is comparable to commercial wearables, such as 42mm Apple Watch[6]: 50g and Samsung Gear S[7]: 67g - 84g. Thus, MuSS-Bits can be attached to the user's body and still allow free limb movement as well the selection of the spatial location.

Sensor-Bits

The Sensor-Bits embody 2 audio sensors: (1) an omnidirectional in-air microphone (BOB-09964) and (2) a 3.5mm audio jack to allow sound input from a computer or electrical instrument. The processing unit automatically selects the input source as follows: if a signal is present at the audio jack, the audio jack is used,

---

[2] http://www.russpalmer.com/feeling.html

[3] http://freedrumlessons.com/drum-lessons/

[4] https://www.youtube.com/user/guitarlessonscom

[5] https://www.youtube.com/user/PianoLessonsForKids

[6] https://support.apple.com/kb/SP735?locale=de_DE

[7] http://www.samsung.com/ae/consumer/mobile-devices/wearables/gear/SM-R7500ZWAXSG

otherwise the microphone's signal is processed and sent to the Display-Bit. This makes the audio sensor selection implicit and intuitive for the user.

Display-Bits

The Display-Bits contain a vibrotactile display (ERM motor; model 307-103 from precisionmicrodrives) and a single pixel display (a RGB LED). We used an ERM motor, since it is lightweight (4.6 g), can be driven with a DC signal, does not need an amplifier, and has a frequency range that includes 250 Hz (the skin's most sensitive frequency). The used ERM motor has a lag time of 8ms (time until the motor reaches 0.08G), a rise time of 28ms (time until the motor reaches half of its maximum amplitude) and a maximum amplitude of 7G. Thus, it is responsive and allows the perception of vibrations even through garments. The visual brightness of the LED and the vibration motor's amplitude can be controlled independently by 2 potentiometers to allow the calibration of the modalities intensity to the user's preference.

Designing an intuitive mapping from audio to visual/vibrotactile feedback is a challenging task. Thus, we decided to use a simple mapping as an initial start which supports rhythm perception. Informed by previous work (Branje and Fels, 2014; Karam et al., 2010; La Versa et al., 2014; Nanayakkara et al., 2009; Pouris and Fels, 2012), MuSS-Bits translate loudness to the brightness of the LED as well as to intensity of the ERM motor. This was sufficient to convey rhythm information.

Communication

Real-time feedback is a necessary requirement for music-making. In fact, in music ensemble performances a widely accepted delay threshold is a maximum of 30ms (Kurtisi et al., 2006; Schuett, 2002). Hence, we used WiFi for communication facilitated by ESP8266-12F modules in combination with the User Datagram Protocol (UDP) to keep the transmission delay as small as possible. However, UDP can be unreliable and packets can get lost. We found that the ESP8266-12F can receive up to 500 packets per second without packet loss, which was sufficient for our set-up.

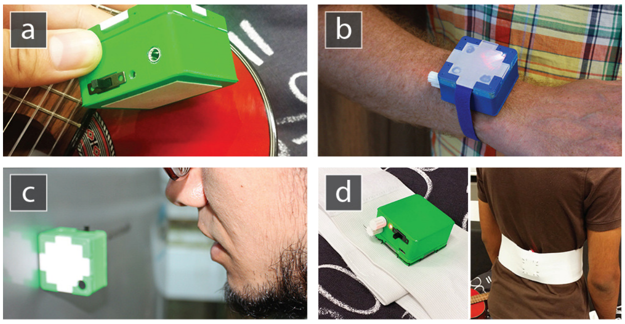

Figure 2. Possible attachment modes for MuSS-Bits: (a) adhesive tape, (b) Velcro band, (c) magnetic and (d) sewing.

Attachment Modes

The attachment mechanism of MuSS-Bits is important, for easy exploration of various audio sources and customization of the spatial location of the received feedback. We designed the MuSS-Bits casing the way that it allows different attachment configurations, such as adhesive tape to mount the MuSS-Bits to a flat surface, such as the corpus of guitar (see Figure 2a); Velcro band to attach the bits to uneven surfaces, such as arms (see Figure 2b); magnets to securely attach the bits to magnetic surfaces (see Figure 2c); or small holes in the top part of the bits to sew the bits into garments, such as a back brace (see Figure 2d).

HARDWARE DESIGN SPACE

In this section, we present the design space of possible hardware implementations for a music sensory substitution system that would enable the previously described design goals.

General Hardware Architecture

The flow of the audio information in a music sensory substitution system that enables exploration and customization is as follows:

(1) audio is captured by a sensor
(2) audio data is processed (e.g. FFT)
(3) processed audio data is transmitted from the sensor to the display part
(4) processed audio data is translated into vibrotactile and visual information under consideration of the user's input (e.g. calibration of the modality)
(5) feedback is presented through visual and vibrotactile displays (e.g. LED or motor)

Based on this flow we derived a general hardware architecture that consists of: (1) audio input, (2) visual and vibrotactile displays, (3) processing units, (4) communication unit, (5) user input and (6) power supply. An overview of the architecture and the flow of information is given in Figure 3.

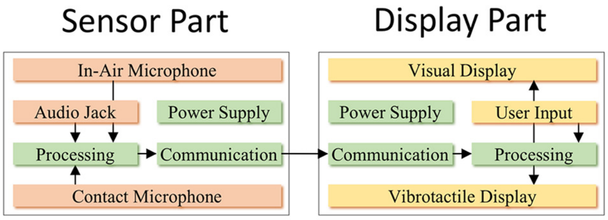

Figure 3. General hardware architecture. Power supply, processing unit and communication interface are present in sensor and display part. Orange components are specific to the sensing and yellow components to the display part. The arrows indicate the flow of information between components.

Audio Input

Automatic selection of the audio source depending on the type of sound input is a desirable feature to simplify the user interaction. Specialized microphones can be used to capture various types of audio inputs. Omnidirectional in-air microphones could be used to capture most of the audible sounds. In contrast, contact microphones could be used to sense sound directly from an object they are attached to, such as the corpus of a guitar, without capturing in-air sounds. Audio output from a computer or an electrical instrument can be embodied through an audio jack for a good quality audio signal.

Visual Display

Visual feedback can be presented through various technologies, such as large screens, monitors, head-mounted displays, mobile displays or single pixel displays. They vary in expressivity, portability and wearability. Large screens provide more space for detailed visualizations whereas single pixel displays can convey only limited information. Most of the reviewed music sensory substitution systems use a screen as a visual display (see Table 1). Nevertheless, point lights have been shown to be useful for communication purposes (Harrison et al., 2012). Single pixel displays allow bringing the visual feedback close to the area of action minimizing the distraction from the `focusing on the instrument' strategy that novice deaf musicians typically employ.

Vibrotactile Display

Vibrotactile displays used in previous works (see Table 1) involve voice coils, speakers and ERM motors. Voice-coil actuators work the same way as speakers. They are very responsive, support a large frequency spectrum, but usually come with a high power consumption and can be quite heavy. LRA pancake motors are very light and power efficient, but have a moderate response (about 40ms) and a low frequency span (typically around 20Hz). ERM motors can be responsive (8ms - 28ms), have a limited bandwidth (typically 0 - 250Hz), are lightweight and consume a moderate amount of power. However, the vibration frequency and intensity cannot be controlled independently.

User Input

The user needs the opportunity to select the feedback modalities and calibrate the intensity of the feedback. The calibration requires a non-discrete input, such as provided by sliders or potentiometers. The selection of a modality is an ON-OFF switching operation and can be implemented by a push button, switches or gesture input such as shaking. The selection of the modality can be combined with the calibration input by interpreting the smallest calibration value as indication to turn off the corresponding modality.

Processing

The processing of the audio signal and following translation into visual and vibrotactile information requires a processing unit. It is important to consider the required tasks from this unit. The processing unit of the sensor part could just forward the pure audio signal, perform onboard basic signal processing, such as Fast Fourier Transform, or even derive musical elements, such as timbre, melody or tempo, from the audio signal. The processing unit of the display part has to translate the received information (raw or processed audio) into the corresponding modality considering the user's input.

Communication

As mentioned, real-time feedback is of high importance for music-making systems and therefore the communication between sensor and display part has to be fast. Wiring sensor and display part would provide the fastest transmission, but hinder easy deployment and redeployment. Also entangled wires might restrict moving them freely. Wireless approaches, such as Bluetooth, RF or WiFi will be more promising if the transmission delay is minimized. RF and Bluetooth (BLE) solutions are energy efficient but have limited bandwidth for data transmission. WiFi has the largest bandwidth and allows real-time streaming of audio data, despite the higher power consumption. The communication mesh between sensors and display part could be one-to-one, one-to-many or many-to-many and allows to combine audio information from different audio sources. MuSS-Bits implement a one-to-one mapping as a starting point to make the relationship between input and output direct and easier to understand.

INTERACTING WITH MUSS-BITS

Interaction Possibilities

With the design goals and the MuSS-Bits prototype, we implemented a number of interaction possibilities. Each of the following scenarios serves as an example to showcase the coherent set of possible uses of a music sensory substitution system for music learning and MuSS-Bits in particular.

Learning an Instrument

MuSS-Bits support instructed as well as self-learning (see Figure 4). A deaf user could use one MuSS-Bit pair to receive feedback from his or her own instrument about sound he or she created. The user can receive guided feedback by attaching a second pair of MuSS-Bits to a teacher's instrument or to a computer that runs a video tutorial. The second pair will provide a ground-truth to compare and compensate for the gap in the feedback loop.

Figure 4. Learning an instrument with MuSS-Bits (a) co-located and through (b) self-learning.

Attaching Display-Bits to the body will provide a stronger vibrotactile feedback. In previous studies, we saw that deaf children playing an instrument visually concentrate on the area of action (e.g. fingers pulling the guitar strings). Augmenting the instrument with Display-Bits with visual feedback can complement this strategy.

Collaborative Music Sessions

MuSS-Bits can be used in collaborative music-making sessions with multiple deaf musicians. Sensor-Bits can be placed near a source that generates a metronome and each musician could attach a Display-Bit to his or her body to perceive the steady-beat. This can reduce the need for permanent visual contact of the performers during a performance to stay synced.

Auditory Explorer

Exploration provides a powerful way of learning, especially when there is no prior knowledge. For example, babies learn mostly by exploration of their world and receiving visual, auditory, haptic, gustatory and olfactory feedback. Similarly, congenitally profoundly deaf children could use MuSS-Bits to develop an understanding of sounds as MuSS-Bits are portable, easily deployable and support real-time feedback of real-world sounds. Users can explore questions related to sound such as 'What is sound?', 'What actions create what kind of sound?' or 'Which actions create similar sounds?'.

Augmenting Space with Feedback

MuSS-Bits can be attached to each other easily with the built-in magnets. This allows users to create composite Display- or Sensor-Bits. A deaf user can, for example, combine a set of Display-Bits to create a seat mat to receive vibrotactile feedback while playing a piano. In a collaborative music session, performers can join their Display-Bits to create a shared visual display representing each performer as one pixel.

Figure 5. Deaf users interacting with MuSS-Bits: (a) blowing into microphone; (b) tapping on Sensor-Bit surface (c) participants playing a counting game; (d) Display-Bit worn around the wrist, (e) the upper arm and (f) held on the hand; (g) collaborative music-making; (h) sharing of a Display-Bit; (i) following an online tutorial.

Initial User Interaction with MuSS-Bits

To understand how users will interact with MuSS-Bits, we gave the prototype to 4 deaf participants (4 male; 12 to 17 years; severe to profoundly hearing loss) from a residential Deaf School (group 1) and to 7 deaf participants (3 female, 4 male; 17 to 25 years; mild to moderate hearing loss) from a local music-making group (group 2). In general, participants from group 1 were trying more exploratory activities whereas group 2 was more concerned about musical features, such as keeping a steady beat and being able to differentiate between instrument and voice. Figure 5 shows instances of the groups' interactions with MuSS-Bits.

Exploration of Sound: In both groups, we saw instances of exploration, such as blowing or speaking into the microphone, hand clapping, playing sound from a mobile phone and the use of a bass drum. We also observed one participant of group 2 tapping a rhythm on the Sensor-Bits surface making the Display-Bit light up and vibrate. He was turning the Sensor-Bit itself to a beat generator. We saw another participant (group 1) holding the vibrating Display-Bit in one hand and using his other hand to imitate the rhythm that was played by his friend on a bass drum. In group 2, one participant mentioned that the most important functionality of these kind of devices is to allow them following the beat, which is in according with Fawkes (Fawkes, 2006) and our design goal to support rhythm and beat in particular.

Customization of Sound: The Display-Bits were situated on the wrist, upper arm, table, or were held in hand. Participants from group 1 found that attaching the Display-Bit to the upper arm is less distractive than the wrist when playing an instrument. One participants of group 1 said, after calibrating the vibrotactile feedback: *"Vibrations are smoother [now]."*

Further Interaction with MuSS-Bits: We saw that participants in group 1 sometimes used one single Display-Bit to share feedback. Furthermore, they started to play a counting game with the MuSS-Bits: one participant was playing a number of beats on a bass drum, another participant, facing a wall and wearing a Display-Bit, was trying to guess the number of beats. We observed that he was able to count the beat with no mistake. Participants of group 2 mentioned, after watching a drumming instruction video, that there was no difference in feedback between the voice of the instructor and the sound when the instructor was playing the drums. This suggests a possibility of representation for timbre. We also observed that the participants in both groups mainly used the velcro-band attachment mechanism. The sewing and magnetic mechanism were not used at all. Participants of group 2 think, if a new member joins their music group, this device could support him or her to catch up. Additionally, the teachers from the Deaf School felt that MuSS-Bits can be a great tool for the school's conducted music sessions.

LIMITATIONS, CONCLUSION AND FUTURE WORK

In this work, we presented MuSS-Bits, a music sensory substitution system that consists of sensor-display pairs. It is intended to allow deaf users to explore and develop a conceptual model of musical sounds, as well as provides the possibility to customize the feedback to cater to individual requirements and preferences. MuSS-Bits have the potential to support deaf users in instructed, as well as self-learning of an instrument and could aid collaborative music sessions. User interaction sessions indicated that actual usage of MuSS-Bits includes exploration of different sounds, such as voice, hand clapping and bass drum, as well as trying different body sites for feedback, such as wrist, hand and upper arm.

Although the user reactions were quite encouraging, our system has limitations. While the size of MuSS-Bits is sufficient to make it portable and wearable, we envision a smaller form factor for the future. Operating time of the device could be extended by harvesting energy from the user's body movement (Saha et al., 2008). MuSS-Bits enable the user to calibrate and select the desired modality as well as customize the spatial location of feedback. However, they do not support customization of the audio-to-modality mapping. An advanced mapping option between features of music and output modality will provide further customization possibilities. At the same time, it opens up questions such as: 'How one would design an intuitive user interface for the creation of music-to-modality mappings, when the users do not have a conceptual model of sound and music in particular?' We consider this as future work.

We are also interested in how MuSS-Bits will affect the musical learning over a longer period of time. We plan to conduct a long term study with a group of deaf participants to understand, how exploration of sound and customization of feedback supports them in learning an instrument and understanding sounds in general. We believe that MuSS-Bits are a first step towards customizable exploratory music-to-sensory substitution systems for deaf people.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of Dr. Reijntjes School for the Deaf, Sri Lanka and ExtraOrdinary Horizons, Singapore. This work is partially funded by the SUTD President's Graduate Fellowship.

REFERENCES

Bertini, C., Leo, F., Avenanti, A., and Ladavas, E. Independent mechanisms for ventriloquism and multisensory integration as revealed by theta-burst stimulation. European Journal of Neuroscience 31, 10, (2010), 1791–1799.

Birnbaum, D.M. and Wanderley, M.M. A systematic approach to musical vibrotactile feedback. In Proc. ICMC 2007. (2007), 397–404.

Branje, C. and Fels, D.I. Playing vibrotactile music: A comparison between the Vibrochord and a piano keyboard. International Journal of Human-Computer Studies 72, 4, (2014), 431–439.

Chasin, M. Music and hearing aids. The Hearing Journal 56, 7, (2003), 36.

Clark, J.G. Uses and abuses of hearing loss classification. Asha 23, 7, (1981), 493–500.

Darrow, A.-A. The Role of Music in Deaf Culture: Implications for Music Educators. Journal of Research in Music Education 41, 2, (1993), 93–110.

Drennan, W.R. and Rubinstein, J.T. Music perception in cochlear implant users and its relationship with psychophysical capabilities. Journal of rehabilitation research and development 45, 5, (2008), 779–789.

Fawkes, W.G. The Teaching of Music to Hearing Impaired Children and Teenagers. (2006).

Ferguson, S., Moere, A.V., and Cabrera, D. Seeing sound: real-time sound visualisation in visual feedback loops used for training musicians. In Proc. IV 2005. (2005), 97–102.

Fourney, D. Can Computer Representations of Music Enhance Enjoyment for Individuals Who Are Hard of Hearing? In: K. Miesenberger, A. Karshmer, P. Penaz and W. Zagler, eds., Computers Helping People with Special Needs. Springer Berlin Heidelberg, (2012), 535–542.

Fourney, D.W. and Fels, D.I. Creating access to music through visualization. Science and Technology for Humanity (TIC-STH), 2009 IEEE Toronto International Conference (2009), 939–944.

Galvin, J.J., Fu, Q.-J., and Shannon, R.V. Melodic Contour Identification and Music Perception by Cochlear Implant Users. Annals of the New York Academy of Sciences 1169, 1, (2009), 518–533.

Glennie, E. How to truly listen. TED (2003). http://goo.gl/iRWrKX

Hagedorn, V.S. Musical Learning for Hearing Impaired Children. Research Perspectives in Music Education, (1992).

Harrison, C., Horstman, J., Hsieh, G., and Hudson, S. Unlocking the Expressivity of Point Lights. In Proc. CHI 2012, ACM (2012), 1683–1692.

Hash, P.M. Teaching Instrumental Music to Deaf and Hard of Hearing Students. Research and Issues in Music Education 1, 1, (2003).

Ho-Ching, F.W., Mankoff, J., and Landay, J.A. Can You See What I Hear?: The Design and Evaluation of a Peripheral Sound Display for the Deaf. In Proc. CHI 2003, ACM (2003), 161–168.

Huang, K., Starner, T., Do, E., et al. Mobile Music Touch: Mobile Tactile Stimulation for Passive Learning. In Proc. CHI 2010, ACM (2010), 791–800.

Isaacson, E.J. What You See Is What You Get: on Visualizing Music. ISMIR, Citeseer (2005), 389–395.

iTunes. iTunes - Everything you need to be entertained. https://www.apple.com/itunes/. Accessed: March 23, 2016.

Jain, D., Findlater, L., Gilkeson, J., et al. Head-Mounted Display Visualizations to Support Sound Awareness for the Deaf and Hard of Hearing. In Proc. CHI 2015, ACM (2015), 241–250.

Jr, T.G.R. Active Exploration. In: S. Goldstein and J.A. Naglieri, eds., Encyclopedia of Child Behavior and Development. Springer US, (2011), 26–27.

Karam, M., Branje, C., Nespoli, G., Thompson, N., Russo, F.A., and Fels, D.I. The Emoti-chair: In Proc. CHI EA 2010, ACM (2010), 3069–3074.

Karam, M., Nespoli, G., Russo, F., and Fels, D.I. Modelling Perceptual Elements of Music in a Vibrotactile Display for Deaf Users: A Field Study. In Proc. ACHI 2009. (2009a), 249–254.

Karam, M., Russo, F.A., and Fels, D.I. Designing the Model Human Cochlea: An Ambient Crossmodal

Audio-Tactile Display. IEEE Transactions on Haptics 2, 3, (2009b), 160–169.

Karam, M., Russo, F., Branje, C., Price, E., and Fels, D.I. Towards a Model Human Cochlea: Sensory Substitution for Crossmodal Audio-tactile Displays. In Proc. GI 2008, Canadian Information Processing Society (2008), 267–274.

Kim, J., Ananthanarayan, S., and Yeh, T. Seen Music: Ambient Music Data Visualization for Children with Hearing Impairments. In Proc. IDC 2015, ACM (2015), 426–429.

Kurtisi, Z., Gu, X., and Wolf, L. Enabling Network-centric Music Performance in Wide-area Networks. Commun. ACM 49, 11, (2006), 52–54.

La Versa, B., Peruzzi, I., Diamanti, L., and Zemolin, M. MUVIB: Music and Vibration. In Proc. ISWC 2014 Adjunct Program, ACM (2014), 65–70.

Levy-Tzedek, S., Hanassy, S., Abboud, S., Maidenbaum, S., and Amedi, A. Fast, accurate reaching movements with a visual-to-auditory sensory substitution device. Restorative neurology and neuroscience 30, 4, (2012), 313–323.

Limb, C. Building the musical muscle. TED (2011). http://goo.gl/rvDR1x

Machover, T. Tod Machover + Dan Ellsey: Inventing instruments that unlock new music. TED (2008). http://goo.gl/h2hzfj

MAM. Music Animation Machine — “Music Worth Watching.” http://www.musanim.com/. Accessed: March 19, 2016.

Marshall, M.T. and Wanderley, M.M. Vibrotactile Feedback in Digital Musical Instruments. In Proc. NIME 2006, IRCAM — Centre Pompidou (2006), 226–229.

MATD. Music and the Deaf - Enriching lives through music. http://matd.org.uk/. Accessed: May 1, 2015.

Matthews, T., Fong, J., and Mankoff, J. Visualizing Non-speech Sounds for the Deaf. In Proc. ASSETS 2005, ACM (2005), 52–59.

May, E. Music for Deaf Children. Music Educators Journal 47, 3, (1961), 39–42.

Medina, J. Brain Rules: 12 Principles for Surviving and Thriving at Work, Home, and School (Large Print 16pt). ReadHowYouWant.com, (2011).

Mitroo, J.B., Herman, N., and Badler, N.I. Movies from Music: Visualizing Musical Compositions. In Proc. SIGGRAPH 1979, ACM (1979), 218–225.

Mori, J. and Fels, D.I. Seeing the music can animated lyrics provide access to the emotional content in music for people who are deaf or hard of hearing? Science and Technology for Humanity (TIC-STH), 2009 IEEE Toronto International Conference (2009), 951–956.

Music that Moves. https://goo.gl/MTakSg. Accessed: March 24, 2016.

Nanayakkara, S.C., Wyse, L., Ong, S.H., and Taylor, E.A. Enhancing Musical Experience for the Hearing-Impaired Using Visual and Haptic Displays. Human–Computer Interaction 28, 2, (2012), 115–160.

Nanayakkara, S., Taylor, E., Wyse, L., and Ong, S.H. An Enhanced Musical Experience for the Deaf: Design and Evaluation of a Music Display and a Haptic Chair. In Proc. CHI 2009, ACM (2009), 337–346.

Palmer, R. http://www.russpalmer.com/tactile.html. Accessed: March 20, 2016.

Pouris, M. and Fels, D.I. Creating an Entertaining and Informative Music Visualization. In: K. Miesenberger, A. Karshmer, P. Penaz and W. Zagler, eds., Computers Helping People with Special Needs. Springer Berlin Heidelberg, (2012), 451–458.

Rashid, R., Aitken, J., and Fels, D.I. Expressing Emotions Using Animated Text Captions. In: K. Miesenberger, J. Klaus, W.L. Zagler and A.I. Karshmer, eds., Computers Helping People with Special Needs. Springer Berlin Heidelberg, (2006), 24–31.

Reed, C.M., Doherty, M.J., Braida, L.D., and Durlach, N.I. Analytic Study of the Tadoma MethodFurther Experiments with Inexperienced Observers. Journal of Speech, Language, and Hearing Research 25, 2, (1982), 216–223.

Roebuck, J. http://goo.gl/ECM8Dj. Accessed: August 26, 2015.

Saha, C.R., O’Donnell, T., Wang, N., and McCloskey, P. Electromagnetic generator for harvesting energy from human motion. Sensors and Actuators A: Physical 147, 1, (2008), 248–253.

Schuett, N. The effects of latency on ensemble performance. Bachelor Thesis, CCRMA Department of Music, Stanford University, (2002).

Shinohara, K. and Tenenberg, J. A Blind Person’s Interactions with Technology. Commun. ACM 52, 8, (2009), 58–66.

Stumpf, C. Tonpsychologie. Leipzig: Hirzel 1, (1883).

Tan, H.Z., Gray, R., Young, J.J., and Traylor, R. A haptic back display for attentional and directional cueing. Haptics-e 3, 1, (2003), 1–20.

Verillo, R.T. Vibration sensing in humans. Music Perception 9, 3, (1991), 281–302.

Vy, Q.V., Mori, J.A., Fourney, D.W., and Fels, D.I. EnACT: A Software Tool for Creating Animated Text Captions. In: K. Miesenberger, J. Klaus, W. Zagler and A. Karshmer, eds., Computers Helping People with Special Needs. Springer Berlin Heidelberg, (2008), 609–616.

Yang, H.-J., Lay, Y.-L., Liou, Y.-C., Tsao, W.-Y., and Lin, C.-K. Development and evaluation of computer-aided music-learning system for the hearing impaired. Journal of Computer Assisted Learning 23, 6, (2007), 466–476.

Zhou, Y., Sim, K.C., Tan, P., and Wang, Y. MOGAT: Mobile Games with Auditory Training for Children with Cochlear Implants. In Proc. MM 2012, ACM (2012), 429–438.