Vocal Vibrations: A Multisensory Experience of the Voice

Charles J. Holbrow*
MIT Media Lab
E14-333B, 75 Amherst Street
Cambridge, MA 02139
charles@media.mit.edu

Elena Jessop
MIT Media Lab
E14-333A, 75 Amherst Street
Cambridge, MA 02139
ejessop@media.mit.edu

Rébecca Kleinberger
MIT Media Lab
E14-333B, 75 Amherst Street
Cambridge, MA 02139
rebklein@mit.edu

ABSTRACT

Vocal Vibrations is a new project by the Opera of the Future group at the MIT Media Lab that seeks to engage the public in thoughtful singing and vocalizing, while exploring the relationship between human physiology and the resonant vibrations of the voice. This paper describes the motivations, the technical implementation, and the experience design of the Vocal Vibrations public installation. This installation consists of a space for reflective listening to a vocal composition (the Chapel) and an interactive space for personal vocal exploration (the Cocoon). In the interactive experience, the participant also experiences a tangible exteriorization of his voice by holding the ORB, a handheld device that translates his voice and singing into tactile vibrations. This installation encourages visitors to explore the physicality and expressivity of their voices in a rich musical context.

Keywords

Voice, Vibrations, Expressive Interfaces, Tod Machover, Public Installations, Tactile Interfaces

1. INTRODUCTION : VOCAL VIBRATIONS

We all have an intimate experience with our own voice. The voice is infinitely expressive and individually defining. However, many people do not pay close attention to their voice, do not feel comfortable using their voice musically in public, or imagine they could participate in a rich musical experience through their voice. To address this, we are developing techniques to engage the public in the regular practice of thoughtful singing and vocalizing, both as an individual experience and as part of a community. In Vocal Vibrations, we aim to guide participants to experience and explore their own voices through an interactive musical installation.

The Vocal Vibrations installation includes two contrasting but connected spaces: one public, and one private. On arrival, participants enter the public space we call "The Chapel." In this space, the acoustics, layout and visual aesthetics work together with a 10 channel surround sound vocal composition arranged specially for the acoustics of the space to create an environment for meditation, relaxing or harmonizing. One by one, participants from the Chapel are invited to enter the private space we call "The Cocoon," and engage in a solo vocal experiment. A participant alone in the Cocoon listens to a musical piece that they are instructed to follow with their voice. While they vocalize, they experience an exteriorization of their voice through a handheld device, the Oral Resonant Ball (the ORB), that transforms the physical vibrations of their voice into vibrations in their hands.

Figure 1: Vocal Vibrations installation

In this installation, we also seek to bring attention to the nature of the voice as a physical instrument. The act of singing and vocalizing creates vibrations throughout the body. However, people are generally not aware of or focused on these vibrations. The awareness of vocally-produced vibrations can be a source of meditative focus, as well as a way for everyone from novices to trained singers to understand their instrument better. The focus provided by tactile and physical feedback can help to give intimate, while still objective, access to the voice. The ORB can help the user to become more aware of the variation and range of vocal expression by externalizing the vibration of his or her voice. In exploring the relationships between human physiology and the resonant vibrations of the voice, we seek to tackle questions related to the voice and its connection with the body as well as its influence on mental and physical health.

2. RELATED RESEARCH ON THE VOICE

2.1 Voice, Body, Mind, and Vibration

One of our principles in designing a compelling and unique experience in this installation was inspired by the way people experience their voices in an everyday context. Most of us do not pay attention to the complex physical processes involved in producing a vocal signal, particularly one that is expressive or emotional. Additionally, the use of our voice is a goal-directed activity. All the complex psychomotor subprocesses are activated without conscious separation [13].

*The three co-authors are presented in alphabetical order, their contributions to the project are comparable

NIME'14, June 30 – July 03, 2014, Goldsmiths, University of London, UK.

Yet, neurological research supports the idea that the brain dissociates voice from speech when processing vocal information [23]. By comparing the auditory cortical response to voice in self-produced voiced sounds and in tape-recorded voice, it has been shown that the brain's response to self-produced voiced sounds is weaker [8]. This result suggests that during vocal production, there is an attenuation of the sensitivity of the auditory cortex and that the brain modulates its activity as a function of the expected acoustic feedback.

Because the voice requires a perfect psychomotor synchronization between many physical processes (such as the breath, the tongue, the vocal tract muscles, the tension of the vocal folds, and the lips), the study of the voice can reveal details about a person's health and mental state [1]. Mental and emotional states are often apparent through the voice because the physical procedure is closely shaped by emotions. Kenneth Stevens [21] describes those correlates in terms of vocal modification in situations of strong arousal. For example, in the case of stress, variations in muscle contraction and breathing patterns have a direct influence on the sound of the voice. Max Little implemented a biologically and mechanically based model that has been used in different clinical contexts from the evaluation of speech pathology, breath and lung problems, neuronal malfunctions, or muscular control difficulties, to the detection of early stages of Parkinson's disease [16].

Not only can studying the voice reveal information about physical, mental and emotional states, but using the voice can also affect those states. In the subclinical domain, several studies have focused on the links between singing and the physiological signs of wellbeing (heart rate, blood pressure, and stress level) [3, 6, 18]. Those studies generally agree on the fundamental importance of breath control, induced by the use of the voice, as an important connection between singing and physiology.

However, very little work has been done on the effects of the vibrations produced in the body by singing, or on the relaxation and meditation potential of the voice. Many studies have shown that meditation training (especially mindfulness meditation) may be an effective component in treating various disorders such as stress, anxiety, and chronic pain [10, 4]. Despite the voice being a major part of several meditation traditions, the effects of the voice in meditation are mostly unexplored. In one study, medical imaging has shown that cerebral blood flow changes during meditation that incorporates chanting on resonant tones, in ways that cannot be explained solely by breath control [11]. The Vocal Vibrations installation is inspired by the power of the voice and meditation to affect our physiology and emotions.

2.2 Prior Vocal and Vibrational Experiences

Our work in Vocal Vibrations is part of a tradition of interactive vocal experiences. For example, in Messa di Voce, Hidden Worlds, and RE:MARK [15], Levin and Lieberman have incorporated graphics shaped by vocal production into public installations. In these experiences, the amplitude and spectral content of visitors' voices were used to affect projected graphics. Another public installation focusing on the voice is Oliver's Singing Tree [19], with which visitors interacted through singing into a microphone. The "pitch, noisiness, brightness, volume, and formant frequencies" of their voices were measured, and these parameters were used in real time to control a music generation engine and a video generation system. All of these installations have a strong playful component, with the goal of an interesting vocal experience. Focusing carefully on subtle variations of sound has been a component of new music compositions, such as Lucier's "I Am Sitting in a Room" and Chowning's "Phoné" [17, 2] but not of an interactive vocal installation.

Prior work also explores the possibilities of rich aesthetic experiences centered around vibration. Skinscape [7] is a tool for composition in the tactile modality. Inspired by knowing whether "the skin [is] capable of understanding and ultimately appreciating complex aesthetic information," this work tackles the relatively uncharted field of tactile composition. Our work to create vibration experiences derived from and driven by the voice is inspired by this research.

3. THE VOCAL VIBRATIONS INSTALLATION

The Vocal Vibrations installation consists of two connected spaces that encourage participants to experience and explore the singing voice, especially their own voices, in thoughtful ways. When a participant first arrives at the Vocal Vibrations installation, she enters a communal space designed for close listening. In this space, which we call the Chapel, the audio is a precomposed electroacoustic composition by Tod Machover based on recordings of voices. At any time, participants can vocally follow along with this composition by singing a D. Headphones playing a pure D note, sung at different octaves, are also available in the space to help the participant to find a note that fits comfortably in her range. Each participant is then approached individually by an assistant who brings him or her to the cocoon and prepares him or her for the interactive vocal experience. This assistant instructs the participant to vocalize on the same D, and helps the participant to put on a microphone and headphones. After being invited to sit, the participant is given the vibrating ORB to hold, and left alone in the space. The participant then has a six-minute solo experience guided by a specially designed and shaped musical composition. At the end of the solo experience, the user returns to the Chapel, where she is free to stay and listen as long as she wishes, as well as to vocally improvise along with the music if she desires. All of the musical content in this installation is new material composed by Tod Machover, based on vocal recordings of soprano Sara Heaton and the Blue Heron Choir.

3.1 The Chapel: Focused Listening

When visitors first arrive at Vocal Vibrations, they enter the outer chamber, the Chapel, intended for a quiet, meditative experience. Here, singing voices surround visitors and gently envelop them in sound. Visitors can remain in this space for as long or short as they desire, choosing to join in through humming or vocalizing or simply to listen. The composition in the Chapel has been assembled from many layers of pre-recorded solo and choral vocal material, designed such that a D is almost always present in the score. The installation is made up of 10 high-fidelity speakers on tripods that have been configured in a large oval within the room. Each speaker plays different elements of the composition providing a dynamic surround experience. The musical material used in the composition comes from a broad spectrum of vocal traditions including Tuvan throat singing, early Renaissance choral music, and solo soprano explorations.

3.2 The Cocoon: Interactive Vocal Experience

In the second portion of the installation, a private environment, the Cocoon, allows individual visitors to explore new vocal sounds and the vibrations generated by their own voice.

From the Chapel, each participant is guided by an assistant into the interactive experience. A short "training" session follows, in which the participant is encouraged to take the first step into producing vocal sounds. The participant is asked to hold the D and is given the simple guidance to explore a range of vocal variations on a single pitch, such as different vowels, sounds, rhythms, textures, and timbres. We seek to free participants to experiment with a wide range of sounds.

We designed the system such that there is no "right" or "wrong" way to use it. Whatever sound a participant produces does not degrade the musicality of the experience. People are invited to vocally follow the music they hear, but they barely hear their own voice through the headphones, which enables them to feel more comfortable trying new sounds.

3.3 The ORB

For this project, we have built the Oral Resonance Ball (ORB), a voice-activated device that allows people to feel the physical vibrations of their voices in their hands. This device provides awareness of the physical processes involved in vocal production by giving feedback about and enhancing the vibrations produced in a person's body. Fingertips contain more sensor receptors than our vocal vibrating chamber [12, 14]; thus, the same vibrational signal sent into the hands will be felt differently and with more detail than when sent into the body. We have found that the hands can detect many variations in vibration caused by amplitude, frequency, and timbre. Additionally, research on the Tactaid (a tactile hearing aid that codes sound information via a set of vibrators resting on the forearm) has shown that vibration enhances lipreading performance in hearing impaired individuals [5].

Figure 2: The ORB

Holding the ORB in one's hands while vocalizing can give one access in another medium to detailed elements from the voice that often remain latent in one's everyday experience of voice. Additionally, making the vibration of the voice something that can be experienced externally is intended to connect people to their voice in a new way. We offer users a tool to exteriorize their voice and experience another form of connection with it, as well as to engage with their voice as one engages with an external instrument.

3.3.1 Hardware

In designing the ORB, we explored several prototypes of sizes, materials, thicknesses, and shapes. The final shape is the result of a collaboration with the Bold Design company located in Paris. The ORB is an almost ovoid porcelain shell measuring about 10 by 9 by 9 centimeters, with five transducers attached on the inside wall. The materials and precise settings are chosen to maximize the tactile feeling of vibration while minimizing any audible resonance from the device. The object can be held in the hands in any orientation and angle, with different positions varying the perceived vibrational effects on the hands and fingers. Because ceramic-type materials have the microstructural property of presenting no directional atomic order, the material offers the beneficial properties of smoothly and blending the vibration from one transducer to another while keeping certain localized effects. Those localized effects enable the signal to naturally vibrate more at certain points on the object given the frequency spectrum of the signal.

3.3.2 Behavior

The control system for the ORB consists of a Max/MSP patch that processes the vocal input in real time, filters the signal to prevent feedback and to optimize it to the tactile sense and the hardware characteristics, before sending the resulting signal to each of the 5 localized channels based on a set of control parameters. The skin's response to stimuli is not linear. When coding the behaviors of the ORB, we have had to take into account that the signal sent to the ORB is subjected to three serial, non-linear sources of physical alterations before being perceived: the transducers, the material of the shell, and the skin of the user's fingers and palm. The nonlinearity of the transducers is resolved by tuning them through applying a different gain to each of the five signals. Additionally, the range of the vibrotactile frequency response to which skin is sensitive is 20 - 1000 Hz. This range is much narrower than the auditory frequency range our ears can detect (20 - 20,000 Hz). The frequencies of signals sent to the ORB should differ from an audio signal in order to be perceived through touch.

The user's vocal signal is the primary input used for the ORB. In addition, the behaviors of the ORB are also subtly shaped by parameters of the individual's vocal explorations. A variety of computational features are extracted from the voice and used as inputs to the Expressive Performance Extension System, a tool designed for flexible mapping of input data streams to output control parameters [22, 9]. This system allows users to obtain raw input data, extract expressive features, define desired qualities of vocal and physical performance, perform pattern recognition to identify those qualities, and manually map information about these high-level expressive parameter spaces to output control parameters of an interactive experience. In Vocal Vibrations, a combination of analyzed vocal parameters are used to control the spatial vibration patterns of the ORB in real time. These interactive modifications are fairly subtle so as to keep the ORB's vibration feeling directly connected to the participant's voice.

4. EVALUATION

Whether participants' experiences in this installation were meditative or explorative, these experiences were always personal. Most participants reacted strongly to this novel exploration of their own voices. To understand the potential impact of this type of vocal practice we collected data from a subset of participants. We measured psychophysical signals (electrodermal activity, heart rate, breathing rate) and bio-acoustic signals (vocally generated vibrations at six points on the face) from 35 participants. The objective of this study is to investigate how vocal sounds, vocal vibrations, singing experience, and psychophysiological signals interact. We also conducted a 15 minute long interview and had a survey questionnaire. The data is currently being analyzed and will be published in August 2014 in R Kleinberger's thesis, work on which is undergoing. We present here some preliminary results from the survey.

During the subsequent interview and survey questionnaires, participants reflected on their experience: 83% of the participants reacted to the experience positively or very positively, and 17% found it neutral. 0% judged the experience to have been negative. People often reacted in one of two quite opposite ways. About two-thirds of the people reported that the experience helped them to calm down, relax or meditate, which is confirmed by a decreasing EDA activity with little disturbance and supported by testimonies such as, "It reminded me when, as a little girl, I would sit on my mother's lap and have my ear against her chest," and "I want one of those ORBs to help me relax when I am at work or before going to bed." A third reported having felt more exploratory or free (among which 72% reported having discovered vocal sounds they never produced before). Overall, 83% of the population reported having felt that they were part of the musical experience.

5. CONCLUSIONS AND FUTURE DIRECTIONS

In this paper, we have described the Vocal Vibrations project and the first public installation developed for the project, including a space for careful, meditative listening and a space for personal vocal exploration. This initial installation premiered at Le Laboratoire in Paris in March 2014, and will remain installed there through September 2014, after which it will move to The Lab Cambridge (USA).

We also seek to expand our explorations of the vibrations tied to the voice and methods for transforming a participant's experience of those vibrations. In this effort, tools built for the deaf community can also be an interesting source of inspiration, such as the Tadoma method of "tactile lip reading" [20], where a deaf person uses their hand to pick up vibrations and movement from the speaker's lips, jaw, cheek, and throat. This use of alternative senses to get as close as possible to the physical process of voice production is inspirational because it also brings people closer to the emotion and liveness of the voice.

Throughout our work with Vocal Vibrations, we hope to encourage people to explore and pay thoughtful attention to the range of their vocal sounds and vibrations, to have rich musical experiences centered on their voices, and to experience their voices in a new way.

6. ACKNOWLEDGMENTS

We would like to thank all of the people who collaborated on this project and especially The Opera of the Future group at the MIT Media Lab, Le Laboratoire, Bold Design, the Dalai Lama Center for Ethics and Transformative Values at MIT, Neri Oxman, the Blue Heron Choirunder the direction of Scott Metcalfe, Sara Heaton, and the Alash Ensemble.

7. REFERENCES

[1] C. Adnene, B. Lamia, and M. Mounir. Analysis of pathological voices by speech processing. In *Signal Processing and Its Applications, 2003. Proceedings. Seventh International Symposium on*, volume 1, pages 365–367 vol.1, July 2003.

[2] J. Chowning. Phoné (1980-81). 1981.

[3] S. Clift, Hancox, et al. Choral singing and psychological wellbeing: Quantitative and qualitative findings from English choirs in a cross-national survey. *Journal of Applied Arts and Health*, 1(1):19–34, Jan. 2010.

[4] R. J. Davidson et al. Alterations in brain and immune function produced by mindfulness meditation. *Psychosomatic medicine*, 65(4):564–570, 2003.

[5] K. L. Galvin, G. Mavrias, et al. A comparison of tactaid ii and tactaid 7 use by adults with a profound hearing impairment. *Ear and hearing*, 20(6):471, 1999.

[6] C. Grape, M. Sandgren, et al. Does singing promote well-being?: An empirical study of professional and amateur singers during a singing lesson. *Integrative physiological and behavioral science : the official journal of the Pavlovian Society*, 38(1):65–74, 2003.

[7] E. Gunther. *Skinscape : A Tool for Composition in the Tactile Modality by*. PhD thesis, MIT, 2001.

[8] J. F. Houde, S. S. Nagarajan, K. Sekihara, and M. M. Merzenich. Modulation of the auditory cortex during speech: an MEG study. *Journal of cognitive neuroscience*, 14(8):1125–38, Nov. 2002.

[9] E. N. Jessop. Capturing the body live: A framework for technological recognition and extension of physical expression in performance. (in press).

[10] J. Kabat-Zinn, L. Lipworth, and R. Burney. The clinical use of mindfulness meditation for the self-regulation of chronic pain. *Journal of Behavioral Medicine*, 8(2):163–190, 1985.

[11] B. G. Kalyani, G. Venkatasubramanian, et al. Neurohemodynamic correlates of om chanting: a pilot functional magnetic resonance imaging study. *International journal of yoga*, 4(1):3, 2011.

[12] T. Kitamura. Measurement of vibration velocity pattern of facial surface during phonation using scanning vibrometer. *Acoustical Science and Technology*, 33(2):126–128, 2012.

[13] P. Ladefoged. A Figure of Speech: A Festschrift for John Laver edited by William J. Hardcastle, Janet Mackenzie Beck. pages 1–14, 1992.

[14] S. Levänen and D. Hamdorf. Feeling vibrations: enhanced tactile sensitivity in congenitally deaf humans. *Neuroscience letters*, 301(1):75–77, 2001.

[15] G. Levin and Z. Lieberman. In-situ speech visualization in real-time interactive installation and performance. In *NPAR*, volume 4, pages 7–14, 2004.

[16] M. A. Little. *Biomechanically Informed Nonlinear Speech Signal Processing*. PhD thesis, University of Oxford, 2006.

[17] A. Lucier. I am sitting in a room. 2000.

[18] V. Müller and U. Lindenberger. Cardiac and respiratory patterns synchronize between persons during choir singing. *PloS one*, 6(9):e24893, Jan. 2011.

[19] W. D. Oliver. *The Singing Tree: a novel interactive musical experience*. PhD thesis, Massachusetts Institute of Technology, 1997.

[20] C. Reed, W. Rabinowitz, et al. Research on the tadoma method of speech communication. *The Journal of the Acoustical society of America*, 77:247, 1985.

[21] K. N. Stevens. *Acoustic phonetics*, volume 30. The MIT press, 2000.

[22] P. A. Torpey and E. N. Jessop. Disembodied performance. In *CHI'09 Extended Abstracts on Human Factors in Computing Systems*, pages 3685–3690. ACM, 2009.

[23] K. von Kriegstein, E. Eger, A. Kleinschmidt, and A. L. Giraud. Modulation of neural responses to speech by directing attention to voices or verbal content. *Brain research. Cognitive brain research*, 17(1):48–55, June 2003.