Vocal Musical Expression with a Tactile Resonating Device and its Psychophysiological Effects

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

ABSTRACT

This paper presents an experiment to investigate how new types of vocal practices can affect psychophysiological activity. We know that health can influence the voice, but can a certain use of the voice influence health through modification of mental and physical state? This study took place in the setting of the Vocal Vibrations installation. For the experiment, participants engage in a multisensory vocal exercise with a limited set of guidance to obtain a wide spectrum of vocal performances across participants. We compare characteristics of those vocal practices to the participant's heart rate, breathing rate, electrodermal activity and mental states. We obtained significant results suggesting that it might be possible to correlate psychophysiological states with characteristics of the vocal practice if we also take into account biographical information.

Author Keywords

Voice, Vibrations, Welbeing, EDA, Physiology

CCS Concepts

•**Human-centered computing** → *Haptic devices;* •**Applied computing** → *Sound and music computing;* •**Hardware** → *Sensors and actuators;*

1. INTRODUCTION

Vocal Vibrations [12] is an interactive voice-based multisensory installation first exhibited in Paris then in Cambridge (MA, USA). The experience seeks to engage the public in thoughtful singing and vocalizing, while exploring the relationship between physiology and the vibrations of the voice. Part of the installation consists of an interactive individual experience during which a participant sits alone is a small room called *the cocoon*, is given headphones and a earset microphone as well as small connected ceramic object called the orb. The orb is an handheld device that translates voices and singing into tactile vibrations. During the experience, the headphones play an original 6 minutes musical piece composed of voice samples from different traditions to transport the listener into a journey of many different possible use of the voice. While listening, the user is asked to freely vocalize along. Because of the music, users can barely hear their voice and only experience the tactile tangible exteriorization of their voice by holding the orb. After observing the public success of the Vocal Vibrations installation and how strongly people seemed to find benefits from the interactive experience, we decided to conduct a study to assess whether we can find significant correlation between people's vocal experience and what they experienced physiologically.

Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). Copyright remains with the author(s).

NIME'18, June 3-6, 2018, Blacksburg, Virginia, USA.

Figure 1: Synchronized view of the bioacoustic and psychophysical signals during a participant's vocal performance

2. MOTIVATION AND BACKGROUND

The voice has often been used to gather information about someone's physical of mental health. Bioacoustics is the domain that investigates how living bodies produce, disperse and react to sounds. Monitoring and listening to the sounds produced in the body is a common diagnosis method. In the case of diagnosis from the recording of voiced sounds, different methods can cover a wide range of diseases from lung diseases [20] to the origin of chest pain [3] to cardiac anomalies [27] and even very early stages of Parkinson disease [25]. Acoustic voice parameters have also shown to be useful in psychiatric research to help diagnose depression [29] or anxiety disorder [28].

In addition, vibration therapies have shown many different medical effects depending on frequency, mode of exposure, length of application, amplitude and acceleration. Several of those methods use frequencies that are in the human vocal range (50 to 1200Hz). For instance, vibrations from 50 to 150Hz have beneficial effects on chronic pain [24]. 100Hz vibrations have shown to reduce tooth pain [9]. In biological, vibrations around 300Hz are used to enhance DNA synthesis [22]. And lower frequencies from 10 to 45Hz have positive effects on fluid [31] and blood [23] circulation and thus help reduce the risk of edema formation [17]. Vibrations around 15-30Hz have also been used to help Parkinson patients with motor impairment [16].

One of the main motivations for this project comes from the ancestral tradition of chanting and mantra and the role played by voice in many forms of meditation and health related practices [32]. 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 [14, 6]. In addition, speech perception skills have been shown to be significantly higher in adults who practiced meditation [19].

But despite the voice being a major part of several meditation traditions and a flashlight to diagnose diseases, the physical and physiological effects of voice practices are mostly unexplored and 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 as suggested by studies showing that medical imaging has shown that cerebral blood flow changes during meditation that incorporates chanting in ways that cannot be explained solely by breath control [15]. It has also been show that group singing and participation in a choir fosters wellbeing and mental health [7, 4] but in those studies, researchers haven't examined the particular correlations between the characteristics of the vocal practices and health.

In our case, and in concert with the Vocal Vibrations project we were interested in understanding whether a certain type of vocal practice can influence people's relaxation and attention state. The experiment was executed by comparing people's performance during the experience to measures of their psychophysical state. To characterize people's psychophysical activity, we measure three signals: the Electro Dermal Activity (EDA), heart rate and breathing rate. To characterize the parameters of people's vocal performance, we use the hexauscultation sensor described in section 3 as well as audio vocal parameters.

The main contributions of this work are to state the needs and potential to more holistic research on the voice, the internal vibrations produced by the voice and their possible effects on physical and mental health. In addition the research aims to bridge different domains that study the voice (subclinical, meditation, vibration therapy, interactive installation) and to propose the design of new types of hardware and software tools to understand and modify people's use and perspective on their own voice.

3. EXPERIMENT DESIGN

3.1 Participants

The institutional review board approved this study, which was registered as protocol no. 1403006219. The sample comprised 39 adults of both sexes, aged from 21 to 83 years, the average age is 41. There were two groups: a control group consisting of 3 adults, and a study group comprising 36 adults. The study was performed while the installation was closed for the public.

Each participant reads and signs a consent form then is greeted with an introduction speech detailing the purpose of the study. They are told that, in the regular Vocal Vibrations experience, participants wear a microphone, a headphone with which they hear a 6:24 minutes musical composition. The music is based on the note D and is composed to bring the listener in a journey of exploration of different sound and vocal textures. During those 6:24 minutes, they are asked to really explore and play with their own voice. They are recommended to hold the orb with both hands and told that it will vibrate in reaction to their voice.

3.2 Sensors

The sensors useed for this study are: an optical heart rate sensor on the left pinky; an EDA sensor on the right hand, applied with gel; a custom-made hexauscultation mask that senses vocal vibrations at 6 different points on the face (right and left fronthead, nose and neck) [18]. The mask is used to measure the Vibration Coefficient (VibCoeff).

As explained later, the EDA, heart sensors and breathing sensors are often used in the literature to gain near real time access to psychophysiological activity as EDA, breathing rate, heart rate and Heart Rate Variability (HRV) are affected by sympathetic and parasympathetic nervous systems and give insights about arousal level [8]. Later on, the valence is measured though the interview and questionnaire.

3.3 Interview and Survey

Immediately after the experiment, the participant undergoes an open-ended interview during which they were still wearing the sensors. They were asked to talk about their experience through different questions such as *How would you describe the experience?* 2) *Can you recall other experiences you had that were similar to this?*, etc. After the interview, we give participants a questionnaire that starts with questions directly related with how they felt during the experience as well as self report on their level arousal and valence. We reserve biographical questions for the end.

4. DATA ANALYSIS

Video: Each video is manually analyzed independently by two researchers for better reliability. Hands and body motions are observed to see if any motion artifacts might have occurred in the EDA signal. **Survey questions, processing and interpretation**: To remove individual bias from the survey, we rescale the results of all numerical questions between 0 and 5 for each participant individually.

EDA: measurement and interpretation: We used an Affectiva Q sensor. Data sample rate is set up at 8Hz. The sensor is placed at the participant's fingertips rather than on their wrist, where the signal is weaker. Once the sensor is in place, it needs time for the signal to reach the right level. Because of this, we setup the EDA sensor 5 minutes before the beginning of the experience. To preprocess the EDA signal, the raw signal is separated between phasic and tonic signal using the LedaLab Matlab extension [1, 2]. The tonic part corresponds to the background variation while the phasic part is the rapid component resulting from sympathetic neuronal activity. From this, we extract the number of peaks above the median threshold in the phasic signal. We watched the video of the experience to exclude potential motion artifacts.

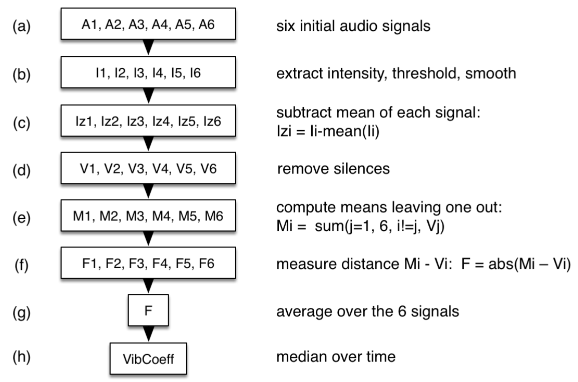

Figure 2: Method used to compute the VibCoeff parameter

Because each individual has a different EDA level and reactivity, we need an individual baseline for each participant. We extract this baseline from the interview time. This baseline is not perfect and the result might be influenced by the fact that the people are asked to interact to the experimenter. But we still consider that it gives good enough information about the individual variation of EDA outside the experimental context and to know if certain participants are likely to be hyper or hypo-responders [26].

From those considerations we define a certain number of metrics that will be used to assess the results. EDA_N_A (aligned normalized EDA) is the raw EDA signal normalized over the experience and from which we subtract the initial value in order to align the signal to the initial state to see the evolution trend brought by the experience. EDAT_N_A (aligned normalized tonic EDA) is the tonic component of the EDA passed through the same process. EDAP_N (normalized phasic EDA) is the phasic component of the EDA normalized throughout the experiment. EDA_peaks is the number of peaks above the median threshold from EDAP_N normalized by the number of peaks measured the same way on the participant's baseline signal. For the 36 participants, the EDA_peaks value goes from 8 to 140. The higher the number, the higher the arousal was during the interaction .

One has to be careful when interpreting EDA signals. It has been shown that it is linked with sympathetic nervous system arousal and thus can give insight about the arousal level of people [5, 11]. But in our type of study, one challenge comes from the specificity problem described by Hedman [10] as follow: "we may be unable to determine what factor(s) resulted in a specific physiological response". Arousal is a measure of how responsive people are to stimuli; it involves the activation of the reticular activating system (RSA) and the endocrine system. To understand the psychophysiological meaning of arousal, it is common to couple it with the measure of valence. The knowledge of valence enables to label arousal values to emotional states. Indeed, as 100 percent of the participants reported feeling positively after the experience, we can consider that the peaks in the EDA signals do not correspond to stress but to high level of engagement, while a decreasing EDA would not correspond to bordome but to state of calm. In our case, understanding the EDA value as "how responsive people are to stimuli" is as relevant as linking it to specific emotions. Indeed, one possible objective is to focus on the task as a self-learning task and use the EDA as a measure of flow [30].

Interpretation of the Vibration signal: we took the same approach as in research measuring the level of nasalization by computing the ratio between amplitude signals measured at the nose and neck skin [21, 13]. In our case we measure to what extent people were exploring diverse vibration patterns as opposed to staying on the same sound. Figure 2 describes the key steps of the process. From the 6 audio measures, we extract the intensity levels (step a) and we threshold and smooth the signal by moving average (step b). Each signal is then detrended (we subtract the mean value of the time-series signal) resulting in a zero-mean signal (step c). The result of the comparison should not be influenced by the voiceless parts of the experience. So we only crop the voiced signal for comparison (step d). To compare them, one possible method would be to compute the mean intensity signal, then to compute the variance of the vector [V0(t), V1(t), V2(t), V3(t), V4(t), V5(t), V6(t)], and to use the average of this value over time as a measure of vibration richness. In our case, we were more interested in how much the signals change relative position. To measure this we use a Jackknife method. We compute 6 times the mean of the signals leaving one out (step e). Mi = sum(j=1, 6, i!=j, Vj). Then we measure the distance between each signal and its associated mean F(t) = abs(Mi(t) - Vi(t)) (step f). We compute the average of the 6 distance signals (step g) and use the median as a measure of vibration pattern richness that we call VibCoeff (step h). We the normalize the result. Figure 3 shows examples of vibration patterns with a high (above) and low (below) VibCoeff over time.

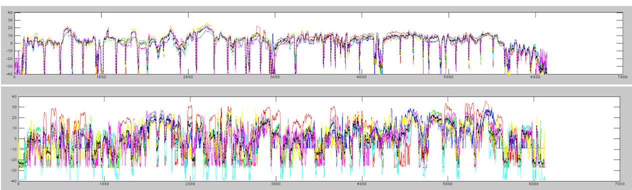

Figure 3: Example of vibrations pattern with a low VibCoeff = 2.58 (top) and with a high VibCoeff = 8.08 (bottom)

5. RESULTS

The objective of this study is to assess whether performance characteristics have an implication on the EDA results. The null hypothesis can be phrased as: H0 = there is no link between the EDA_peaks and VibCoeff. We assess the validity of our hypothesis through the p-value that gives us the probability of observing the given sample result under the assumption that the null hypothesis is true. When observing the relation between EDA_peacks and VibCoeff throughout the whole population, no trend can be extracted. Figure 4 shows the LR curve for all the participants in black. On the graph, no pattern is apparent. The results are also shown in Figure 5. The p-value we obtain is very high (>0.1) which means there is no presumption against the null hypothesis according to statistical significance testing.

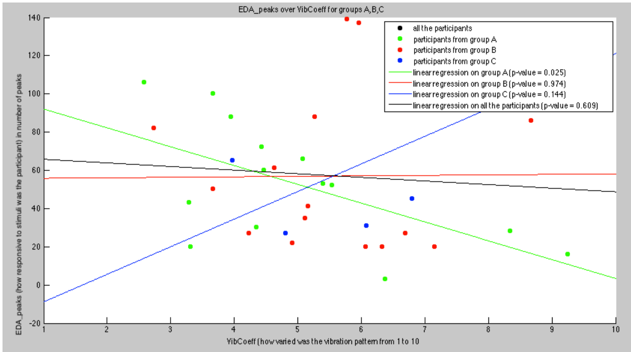

Figure 4: Plots for all participants in black, group A in green, group B in red, group C in blue. The x axis represents EDA_peaks in number of peaks; The y axis is the VibCoeff from 1 to 10

However when subjets are divided by their answer to the question "do you like your voice", the results become more interesting. We call group A the people who answered: "I like my voice quite a lot", B the group of people who answered: "I like it a bit" and C the group who answered "I do not like it at all". Figure 4 and 5 show the linear regression for each groups. For group A, the p-value we obtain is <0.05 which gives us a strong presumption against the null hypothesis. According to statistical significance testing we consider this result as significant which suggests that for people who "like" their voice, a vocal practice with by rich vibration patterns will result in low arousal. .

		Estimate	SE	tStat	pValue
Linear Regression for all participants	(Intercept)	67.694	20.875	3.2428	0.0026536
	x1	-1.8928	3.6619	-0.5169	0.60857
	R-Squared: 0.0078, Adjusted R-Squared: -0.0214				
	F-statistic vs. constant model 0.267, p-value = 0.609				
Linear Regression for group A	(Intercept)	101.92	20.625	4.9416	0.00034117
	x1	-9.8472	3.8697	-2.5447	0.025716
	R-Squared: 0.35, Adjusted R-Squared: 0.296				
	F-statistic vs. constant model 6.48, p-value = 0.0257				
Linear Regression for group B	(Intercept)	-23.309	53.671	-0.43429	0.68218
	x1	14.447	8.3432	1.7316	0.14389
	R-Squared: 0.375, Adjusted R-Squared: 0.25				
	F-statistic vs. constant model 3, p-value = 0.144				
Linear Regression for group C	(Intercept)	55.596	44.332	1.2541	0.23189
	x1	0.25555	7.8165	0.032694	0.97442
	R-Squared: 8.22e-05, Adjusted R-Squared: -0.0768				
	F-statistic vs. constant model 0.00107, p-value = 0.974				

Figure 5: Results of Linear Regression for all and by group

Figure 4 shows the linear regression curve on group C in blue. The results of this RL are shown in Figure 5. The p-value we obtain is slightly >0.1 but still inferior to the value obtained for all participants. This suggests that, for people who are not at all at ease with their voice, a vocal practice characterized by a rich vibration pattern will result in higher arousal. The results have to be tempered as this group contains only seven participants. Finally for group B, the result from LR shows a very high p- value, which means there is no presumption against the null hypothesis according to statistical significance testing.

6. DISCUSSION

The question motivating this study was whether we could significantly show that specific vocal practices could affect psychophysical activity. Our experimental results suggest that given certain knowledge on biographical data, we can predict how people would react to a certain set of vocal exercises. By letting people free in their vocal experience we have been able to assess the performances according to certain criteria of vibration patterns. We have described how the correlation of these practices and biographical data can predict the influence that certain practices can have on people's state of relaxation or engagement. This study would require further experimentation to verify that imposing a specific vocal practice would not bring other perturbations.

Our results suggest that for people who have an harmonious relationship with their voice, rich vibration patterns lead to low EDA, which can be interpreted as: the more they focus on the exercise and contemplate their voice as an instrument through the orb, the more they are driven into a relaxing state. When people from group A tried various vibrations patterns, it helped them calm down and put their mind at rest. When people from group A stayed on a static vibration pattern, the exercise didn't help them be very focused and calm. For the group C on the other hand, the exercise consisting of maintening a sustained sound seems to help them relax more than trying rich sound patterns.

Those explorations suggest that we might be able to correlate physiological state through EDA with characteristics of vocal practices if we also take into account biographical information. This study is a step toward designing Vocal Self Reflective Experiences that are personalised and designed to lead people into certain chosen states (meditation, relaxation or exploration).

7. REFERENCES

[1] M. Benedek et al. A continuous measure of phasic electrodermal activity. *Journal of neuroscience methods*, 190(1):80–91, 2010.

[2] M. Benedek et al. Decomposition of skin conductance data by means of nonnegative deconvolution. *Psychophysiology*, 47(4):647–658, 2010.

[3] W. E. Cayley Jr. Diagnosing the cause of chest pain. *Am Fam Physician*, 72(10):2012–21, 2005.

[4] S. Clift and I. Morrison. Group singing fosters mental health and wellbeing: findings from the east kent singing for health network project. *Mental Health and Social Inclusion*, 15(2):88–97, 2011.

[5] H. D. Critchley. Book review: electrodermal responses: what happens in the brain. *The Neuroscientist*, 8(2):132–142, 2002.

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

[7] G. A. Dingle et al. to be heard: The social and mental health benefits of choir singing for disadvantaged adults. *Psychology of Music*, 41(4):405–421, 2013.

[8] L. Feldman Barrett and J. A. Russell. Independence and bipolarity in the structure of current affect. *Journal of personality and social psychology*, 1998.

[9] P. Hansson et al. Influence of naloxone on relief of acute oro-facial pain by transcutaneous electrical nerve stimulation (tens) or vibration. *Pain*, 1986.

[10] E. Hedman. *Observational Psychophysiology for Empathic Design*. PhD thesis, MIT, 2014.

[11] E. Hedman et al. icalm: Measuring electrodermal activity in almost any setting. In *Affective Computing and Intelligent Interaction and Workshops. ACII 2009. 3rd International Conference on*. IEEE, 2009.

[12] C. Holbrow, E. N. Jessop, and R. Kleinberger. Vocal vibrations: A multisensory experience of the voice. In *NIME*, pages 431–434, 2014.

[13] Y. Horii. An accelerometric measure as a physical correlate of perceived hypernasality in speech. *Journal of speech and hearing research*, 26(3):476, 1983.

[14] 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.

[15] B. G. Kalyani et al. Neurohemodynamic correlates of 'om'chanting: a pilot functional magnetic resonance imaging study. *International journal of yoga*, 2011.

[16] L. K. King et al. Short-term effects of vibration therapy on motor impairments in parkinson's disease. *NeuroRehabilitation*, 25(4):297–306, 2009.

[17] G. Kitay et al. Efficacy of combined local mechanical vibrations, continuous passive motion and thermotherapy in the management of osteoarthritis of the knee. *Osteoarthritis and Cartilage*, 2009.

[18] R. Kleinberger. V 3: an interactive real-time visualization of vocal vibrations. In *ACM SIGGRAPH 2015 Posters*, page 13. ACM, 2015.

[19] U. A. Kumar et al. Effects of meditation on temporal processing and speech perceptual skills in younger and older adults. *Asian Journal of Neuroscience*, 2013.

[20] R. Laennec. *De l'auscultation médiate: traité du diagnostic des maladies des poumons et coeur*. 1819.

[21] R. Lippmann. Detecting nasalization using a low-cost miniature accelerometer. *J Speech Hear Res*, 1981.

[22] J. Liu et al. Biosynthetic response of cultured articular chondrocytes to mechanical vibration. *Research in experimental medicine*, 2001.

[23] E. B. Lohman III et al. The effect of whole body vibration on lower extremity skin blood flow in normal subjects. *Medical Science Monitor*, 2007.

[24] T. Lundeberg. Vibratory stimulation for the alleviation of chronic pain. *Acta Physiologica Scandinavica. Supplementum*, 523:1–51, 1983.

[25] I. Moroz. *Biomechanically informed nonlinear speech signal processing*. PhD thesis, Oxford University, 2007.

[26] B. I. Mpr et al. A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments. 2013.

[27] R. L. Murphy Jr et al. Visual lung-sound characterization by time-expanded wave-form analysis. *New England Journal of Medicine*, 1977.

[28] E. Nerrière et al. Voice disorders and mental health in teachers: a cross-sectional nationwide study. *BMC Public Health*, 9(1):370, 2009.

[29] Å. Nilsonne et al. Measuring the rate of change of voice fundamental frequency in fluent speech during mental depression. *The Journal of the Acoustical Society of America*, 83(2):716–728, 1988.

[30] C. Peifer. Psychophysiological correlates of flow-experience. In *Advances in flow research*. 2012.

[31] J. M. Stewart et al. Plantar vibration improves leg fluid flow in perimenopausal women. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 288(3):R623–R629, 2005.

[32] D. B. Wolf and N. Abell. Examining the effects of meditation techniques on psychosocial functioning. *Research on Social Work Practice*, 13(1):27–42, 2003.