INTERSPEECH 2020: Show & Tell Contribution
October 25–29, 2020, Shanghai, China

Smart Tube: A Biofeedback System for Vocal Training and Therapy through Tube Phonation

Naoko Kawamura[1], Tatsuya Kitamura[2], Kenta Hamada[2]

[1]Himeji Dokkyo University, Japan
[2]Konan University, Japan

kawamura@gm.himeji-du.ac.jp, t-kitamu@konan-u.ac.jp

Abstract

Tube phonation, or straw phonation, is a frequently used vocal training technique to improve the efficiency of the vocal mechanism by repeatedly producing a speech sound into a tube or straw. Use of the straw results in a semi-occluded vocal tract in order to maximize the interaction between the vocal fold vibration and the vocal tract. This method requires a voice trainer or therapist to raise the trainee or patient's awareness of the vibrations around his or her mouth, guiding him/her to maximize the vibrations, which results in efficient phonation. A major problem with this process is that the trainer cannot monitor the trainee/patient's vibratory state in a quantitative manner. This study proposes the use of Smart Tube, a straw with an attached acceleration sensor and LED strip that can measure vibrations and provide corresponding feedback through LED lights in real-time. The biofeedback system was implemented using a microcontroller board, Arduino Uno, to minimize cost. Possible system function enhancements include Bluetooth compatibility with personal computers and/or smartphones. Smart Tube can facilitate improved phonation for trainees/patients by providing quantitative visual feedback.

Index Terms: semi-occluded vocal tract exercises (SOVTE), voice rehabilitation, Arduino Uno, acceleration sensor, LED strip

1. Introduction

Tube phonation is a semi-occluded vocal tract exercise (SOVTE) for voice training and therapy wherein the vocal tract is partly constricted to maximize interaction between vocal fold vibrations and vocal tract resonance [1]. It is widely used in the treatment of patients and the warm-ups for singers. In tube phonation, a trainee must maintain a comfortable level of phonation through a tube or straw for repeated intervals lasting 5 to 10 seconds. At the beginning of the exercise, the trainee is instructed to phonate at an almost constant pitch. She/he is then guided to use a pitch that glides gradually. This is done to strengthen and balance laryngeal musculature and to regulate the glottal airflow caused by muscular effort.

Theoretical studies [2, 3] indicate that the artificial lengthening and semi-occlusion of the vocal tract lowers the first formant frequency and increases impedance at the mouth. Under such condition, the fundamental frequency can be more easily produced at around the first formant frequency (formant tuning) and it provides a better impedance match between the glottis and the epilaryngeal areas of the vocal tract. Such interaction can increase vocal intensity, efficacy, and economy. In addition, previous studies showed the positive effects of exercises involving changes to the supraglottal shape (e.g., elevation of the velum, enlargement of the lower pharyngeal cavity, and lowering of the larynx) [4], vocal fold vibration (i.e., an increase in its amplitude and a decrease in its closed quotient) [4, 5], and vocal functions (i.e., mean airflow increase and phonation threshold pressure (PTP) decrease [6]). These results clearly demonstrate that tube phonation is an effective approach to vocal training and therapy.

When the interaction between the vocal fold vibration and the vocal tract resonance increases during an exercise and the trainee achieves effective phonation, s/he feels strong vibrations around the lips, indicating that acoustic energy is focused around this area. The aforementioned feeling is thus used as a target of the exercise [3]. Vibratory sensations, however, differ from person to person and cannot be directly observed or measured in an objective manner by a voice therapist or trainer. Kawamura *et al.* [7] thus developed computer software that provided visual biofeedback on vibration amplitude, but it proved to be too complicated and costly to be used clinically. This study aims to develop Smart Tube, a simple and low-cost biofeedback system that communicates the intensity of vibration around a trainee's mouth by means of LED lights.

2. Smart Tube system

2.1. System description

Smart Tube provides biofeedback on the amplitude and frequency of the vibrations on the wall of the tube which are caused by the phonation. It does this through a full color LED strip with three individual LEDs attached to the end of the straw. Figures 1 and 2 show the proposed system and a trial run, respectively. Smart Tube makes use of a plastic straw measuring 10 mm in diameter and 210 mm in length. The vibration of the straw is measured by an on-board accelerometer attached to the end of the straw in contact with the subject's mouth. While Kawamura *et al.* [7] measured the vibrations around the mouth directly on the subject's upper lip, Smart Tube measures vibrations on the straw itself to avoid the cumbersome process of attaching the acceleration sensor to the subject's face. To maintain adequate hygiene, the acceleration sensor and LED strip are adhered to plastic clips to enable easy replacement. Smart Tube is controlled through a small and inexpensive microcontroller board, Arduino Uno, which makes it handy and cost-effective.

The acceleration sensor (Analog Devices, ADXL335) picks up acceleration perpendicular to the wall of the straw (z-axis of the sensor). The measured signal is amplified through a high-pass filter with a cutoff frequency of 50 Hz to eliminate the DC signal and any vibrations due to hand movement. The analog signal is then input to Arduino Uno at a sampling frequency of approximately 8.9 Hz (the speed of A/D conversion is not strictly constant in Arduino Uno) with 10-bit resolution. The lighting pattern of the LED strip is controlled in real-time depending on the amplitude and pitch frequency of the acceler-

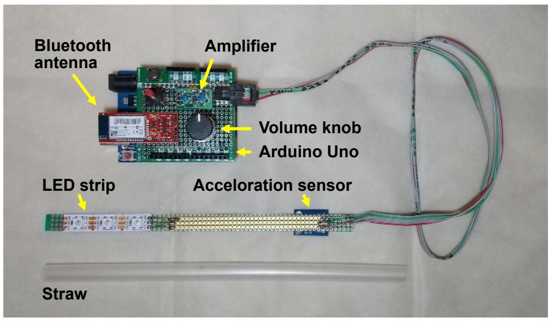

Figure 1: *Smart Tube system.*

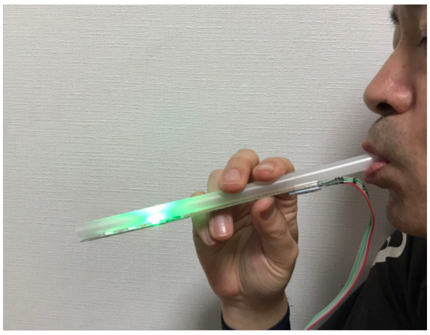

Figure 2: *Smart Tube trial run.*

ation signal calculated at a frame size of 640 points (approximately 72 ms) in accordance with the specifications of the ATmega328P, the microcontroller of Arduino Uno.

Smart Tube currently has two modes: the (1) Amplitude mode and (2) Amplitude and Pitch mode. The Amplitude mode only provides five-level feedback on the vibration (acceleration) amplitude in decibels (dB) indicated by the number of lit LEDs and their color, as shown in Figure 3. Level 1 indicates a lack of amplitude; levels 2 and 3 indicate moderate amplitude; while levels 4 and 5 indicate excessive amplitude. The threshold for each level was determined heuristically.

The Amplitude and Pitch mode is used when the trainee raises and lowers her/his vocal pitch to stretch the intralaryngeal muscles. The number of lit LEDs represents the vibration amplitude. None or one of the three LEDs will light up when amplitude is absent; two or all three LEDs will light up for moderate amplitude; and all LEDs will blink when amplitude is excessive. A specific color was assigned to each vibration frequency. For example, red and orange lights represent the C and D musical scale, respectively. The system enables the voice therapist or trainer to quantitatively measure the trainee's phonatory efficiency, which is otherwise contingent upon the trainer's subjective judgement.

2.2. Link-up with a computer

As a function enhancement, Smart Tube can be connected to a computer or smartphone either via a USB cable or wirelessly through a Bluetooth modem. A Windows PC can also receive acceleration data from Arduino Uno through a COM port and can be set up to display the values on screen, which may well reduce the monotony of vocal training and keep trainees motivated. PC games can be incorporated into the system using gamification principles [8] to attract more trainees and engage

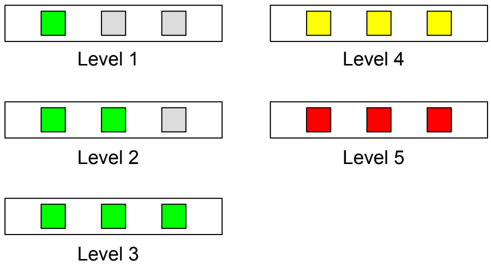

Figure 3: *LED lighting patterns on Amplitude mode.*

them through positive reinforcement.

3. Conclusions

Smart Tube offers simple and easy-to-understand visual feedback of the amplitude and frequency of a trainee's voice during tube phonation. Smart Tube captures the acceleration of the vibrations of a plastic straw when the trainee phonates through it. The amplitude and frequency of the acceleration data are represented by the color and/or number of LED lights. Computer or smartphone integration is also possible for Smart Tube for enhanced management and monitoring of vocal training or therapy.

4. Acknowledgements

The authors would like to thank Prof. Michiko Hashi of Prefectural University of Hiroshima for her help. This study was supported by Hyogo Science and Technology Association, Kawai Foundation for Sound Technology and Music, and Hyogo Foundation for Science and Technology, and The Mikiya Science and Technology Foundation.

5. References

[1] I. R. Ttize and K. V. Abbott, *Vocology: The Science and Practice of Voice Habilitation*. Salt Lake City: National Center for Voice and Speech, 2012.

[2] B. Story, A.-M. Laukkanen, and I. R. Titze, "Acoustic impedance of an artificially lengthened and constricted vocal tract," *Journal of Voice*, vol. 14, no. 4, pp. 455–469, 2000.

[3] I. R. Titze and A.-M. Laukkanen, "Can vocal economy in phonation be increased with an artificially lengthened vocal tract? a computer modeling study," *Logopedics Phoniatrics Vocology*, vol. 32, no. 4, pp. 147–156, 2007.

[4] M. Guzman, A.-M. Laukkanen, P. Krupa, J. Horáček, J. G. Švec, and A. Geneid, "Vocal tract and glottal function during and after vocal exercising with resonance tube and straw," *Journal of Voice*, vol. 27, pp. 523.e19–523.e34, 2013.

[5] K. Minami, H. Maruyama, and T. Haji, "Vocal fold vibration changes with resonance tube method," *The Japan Journal of Logopedics and Phoniatrics*, vol. 56, no. 2, pp. 180–185, 2015.

[6] J. Kang, C. Xuea, D. Piotrowski, T. Gong, Y. Zhang, and J. J. Jiang, "Lingering effects of straw phonation exercises on aerodynamic, electroglottographic, and acoustic parameters," *Journal of Voice*, vol. 33, pp. 810.e5–810.e11, 2019.

[7] N. Kawamura, T. Kitamura, and O. Shiromoto, "Effects of biofeedback system based on skin vibration during tube phonation," *The Japan Journal of Logopedics and Phoniatrics,*, vol. 59, no. 4, pp. 334–341, 2018.

[8] K. Werbach and D. Hunter, *For the Win: How Game Thinking Can Revolutionize Your Business*. Philadelphia: Wharton School Press, 2012.