

# SCIENTIFIC REPORTS



OPEN

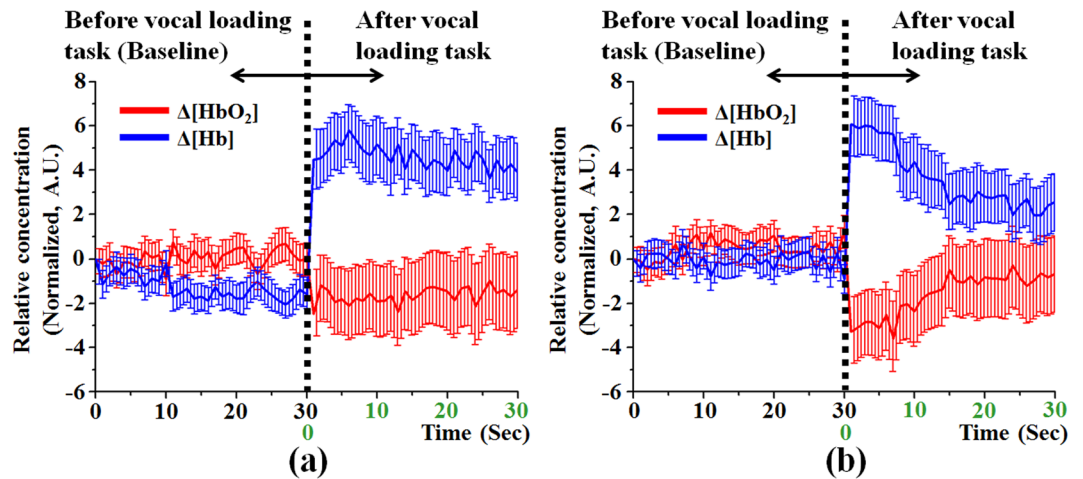
## A feasibility study on non-invasive oxidative metabolism detection and acoustic assessment of human vocal cords by using optical technique

Tzu-Chieh Lin<sup>2</sup>, Jung-Chih Chen<sup>1</sup>, Chih-Hsien Liu<sup>3</sup>, Chia-Yen Lee<sup>4</sup>, Yung-An Tsou<sup>2</sup> & Ching-Cheng Chuang<sup>1</sup>

Voice disorder such as vocal fatigue is a common and complex multifaceted clinical problem that presents a significant impact on quality of life. In this study, the functional near-infrared diffuse optical technique (fNIRS-DOT) was proposed as a novel approach for human vocal cords oxidative metabolism detection and acoustic assessment simultaneously to provide a multidimensional assessment of voice disorder. A totally of 60 healthy subjects included 30 male and 30 female adults of age-matched were recruited and performed a vocal loading task to trigger a mild inflammation of the vocal cords in this study. In the results of oxidative metabolism, the vocal cords expressed hypoxia after vocal loading task in both male and female groups that could provide relevant information on the relationship between tissue oxygen consumption and supply for vocal cords diagnosis. Additionally, the results of optical acoustic assessment revealed the worse/changes voice quality after vocal loading task. Therefore, integration of non-invasive oxidative metabolism detection and acoustic assessment by using optical technique could provide more relevant information for diagnosis of voice disorders. The multi-functional vocal cords detection system could provide a good feasibility for clinical applications such as diagnosis and therapeutic monitoring of voice disorder.

Voice disorder such as the vocal fatigue is a complex clinical phenomenon that is recognized as potentially debilitating presents a significant challenge to clinical practice<sup>1,2</sup>. Professional voice users such as teachers, singers and actors are particularly susceptible to vocal fatigue<sup>3-6</sup>. Certainly, vocal fatigue can seriously affect the social and occupational functioning of such professions. Although vocal fatigue has been defined as a feeling of vocal tiredness and weak voice after the voice is overused or abused<sup>7-9</sup>. Unfortunately, the definition, critical identifying features, and fundamental mechanisms of vocal fatigue remain either unclear. Therefore, vocal fatigue is an important vocal health issue that needs to be further investigated. Numerous studies have adopted the vocal loading test to induce vocal fatigue in voice-healthy participants, and then to track the changes in vocal function that has potential worth to investigate the progression of fatigue as it affects the normal voice<sup>10-17</sup>. The several current clinical detection methods such as auditory perceptual analysis, vocal cords imaging (laryngoscope and videolaryngostroboscopy), aerodynamic analysis, acoustic analysis, and self-evaluation have been widely used for clinical applications of the assessment of voice problems<sup>18-23</sup>. These methods have greatly increased our knowledge about the mechanisms of vocal fatigue. However, vocal fatigue is most likely multifaceted physiological and biomechanical mechanisms. The mechanisms include neuromuscular fatigue, increased vocal fold viscosity, reduced blood circulation, non-muscular tissue strain, and respiratory muscle fatigue. Therefore, the more multidimensional assessment of vocal fatigue is very important for clinical applications.

<sup>1</sup>Institute of Biomedical Engineering, National Chiao Tung University, Hsinchu, 30010, Taiwan. <sup>2</sup>Department of Otolaryngology, China Medical University Hospital, Taichung, 40447, Taiwan. <sup>3</sup>Department of Otolaryngology, Hsinchu Cathay General Hospital, Hsinchu, 30060, Taiwan. <sup>4</sup>Department of Electrical Engineering, National United University, Miaoli, 36063, Taiwan. Correspondence and requests for materials should be addressed to C.-C.C. (email: [ccchuang@nctu.edu.tw](mailto:ccchuang@nctu.edu.tw))



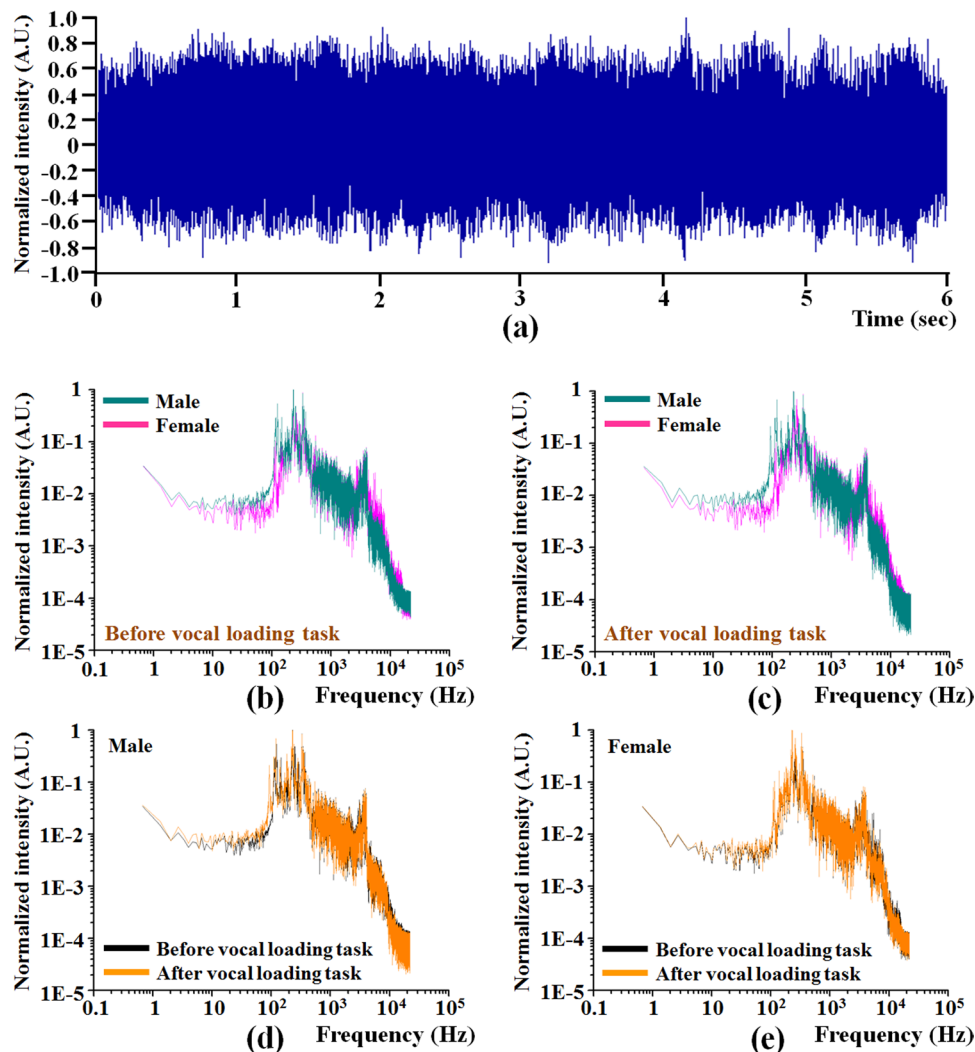
**Figure 1.** Temporal tracings of the concentration changes  $\Delta[\text{HbO}_2]$  and  $\Delta[\text{Hb}]$  before and after vocal loading task. (a) male group and (b) female group.

Functional near-infrared spectroscopy (fNIRS) is a non-invasive technique that provides continuous recording of local changes in tissue oxygenation and perfusion with detection of oxy- and deoxy-hemoglobin concentration changes with at least dual-wavelength (around 800 nm) near-infrared illumination<sup>24,25</sup>. Additionally, fNIRS has several benefits such as less expensive, non-ionizing radiation, real-time measurement, long-time monitoring, easy operation, and completely patient-oriented measurement with high temporal resolution ( $\sim$ ms). The photons in the spectral window of 600–1000 nm wavelength (also known as “optical window” or “therapeutic window”) can penetrate several centimeters into human tissue. Therefore, fNIRS has generated a lot of scientific interest and has been applied in various deep tissue applications such as imaging of the brain, breast, limb, muscle, and joint<sup>24–31</sup>. According to the type of source modulation, fNIRS can be classified into three modes as continuous wave (CW), frequency domain (FD), and time domain (TD)<sup>25,31</sup>. In the CW system, the constant illumination is used for intensity measurement. For this reason, the CW system is solely used to detect changes in absorption coefficient that cannot be used to determine the absolute value of the concentration of oxygenated and deoxygenated hemoglobin. Contrarily, the information of absorption and scattering coefficients of the medium can be determined by using FD and TD system to measure the intensity and phase or time delay of the received light. Therefore, the absolute value of the tissue optical parameters can be estimated to obtain the absolute value of the concentration of oxygenated and deoxygenated hemoglobin. However, in the clinical studies, the analysis of statistically significant difference before and after some specific test is more important than quantification of absolute value. The CW system can be used to detect the concentration changes of oxygenated ( $\Delta[\text{HbO}_2]$ ) and deoxygenated hemoglobin ( $\Delta[\text{Hb}]$ ). Additionally, the CW system provides higher time resolution, relatively low-cost, and easy transportability. Therefore, the CW based system was developed for detection of vocal cords in this study. Besides, the optical detection method such as fiber-optic acoustic sensing is able to measure vibrations from the surface of the skin during vocalization, providing intensity of time series and frequency information that can be processed for acoustic analysis. The fiber-optic acoustic sensing also called optical microphone has many benefits such as high sensitivity, anti-RFI, anti-EMI, high safety and credibility for medical diagnosis<sup>32–35</sup>. Consequently, the optical detection methods can provide not only oxidative metabolism detection but also acoustic assessment for vocal cords diagnosis.

In this study, the multi-functional vocal cords detection system was developed by using the functional near-infrared diffuse optical technique. The techniques of fNIRS and optical microphone were integrated into a novel system that provided the information of oxidative metabolism detection and acoustic assessment simultaneously. A totally 60-min vocal loading task was performed to induce vocal fatigue. Therefore, the oxygenic and acoustic information before and after vocal loading task were obtained and analyzed for investigating of vocal fatigue. The results indicated that the mild inflammation after vocal loading task induced vocal cords hypoxia. Additionally, the results of optical acoustic assessment revealed the worse voice quality after vocal loading task. Our results implied that the integration of oxygen consumption detection and acoustic assessment by using fNIRS diffuse optical technique could provide more multidimensional information for clinical applications of voice disorder.

## Results

Figure 1 shows the mean values of temporal tracings of the concentration changes  $\Delta[\text{HbO}_2]$  and  $\Delta[\text{Hb}]$  before and after vocal loading task. Obviously, the temporal tracings of  $\Delta[\text{HbO}_2]$  and  $\Delta[\text{Hb}]$  were significantly different between male and female groups, especially after vocal loading task. In both groups, the  $\Delta[\text{HbO}_2]$  was lower after vocal loading task than before vocal loading task. Contrarily, the  $\Delta[\text{Hb}]$  was higher after vocal loading task than before vocal loading task. Additionally, the  $\Delta[\text{HbO}_2]$  was higher than the  $\Delta[\text{Hb}]$  before vocal loading task and the  $\Delta[\text{HbO}_2]$  was lower than the  $\Delta[\text{Hb}]$  after vocal loading task. The results indicated that the mild inflammation may induce vocal cord hypoxia. Besides, it is noteworthy that the  $\Delta[\text{HbO}_2]$  and  $\Delta[\text{Hb}]$  maintained a stable trend after vocal loading task in male group (as shown in Fig. 1(a)). In the female group, the  $\Delta[\text{HbO}_2]$  slowly increased

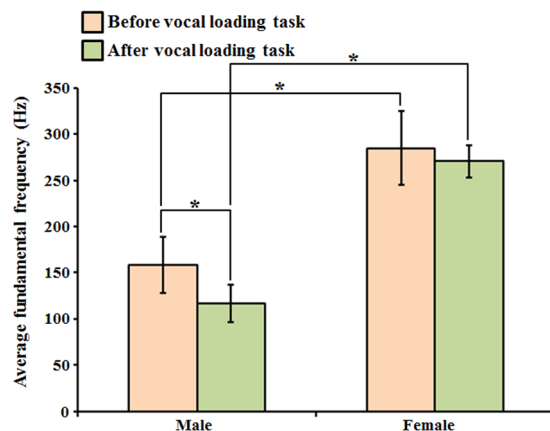


**Figure 2.** The distribution of time series optical signals and the frequency spectrum with logarithmic scale. (a) the time series optical signals of the tissue surface vibration during subjects voiced the sustained vowel / $\hat{a}$ /; (b) the frequency response in the step before vocal loading task of male and female groups; (c) the frequency response in the step after vocal loading task of male and female groups; (d) comparison of frequency response before and after vocal loading task in male group; (e) comparison of frequency response before and after vocal loading task in female group.

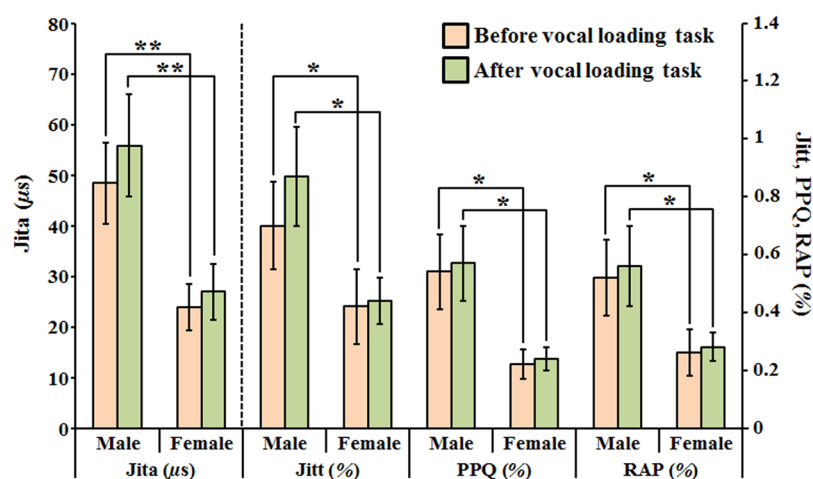
and the  $\Delta[\text{Hb}]$  slowly decreased to a stable state after vocal loading task (as shown in Fig. 1(b)). Compared with female group, the male group presented higher  $\Delta[\text{Hb}]$  after vocal loading task. The results suggested that the vocal cord hypoxia after vocal loading task was stronger in male group than in the female group.

In this study, the time series optical signals of the tissue surface vibration during subjects voiced the sustained vowel / $\hat{a}$ / were detected for acoustic analysis (as shown in Fig. 2(a)). Figure 2(b) to (e) show the frequency spectrum with a logarithmic scale that was processed from the time series optical signals by using the algorithm of fast Fourier transform (FFT). Figure 2(b) and (c) shows the comparison of frequency spectrum between male and female groups before and after vocal loading task. Although the male and female groups showed a rather similar distribution of the frequency spectrum, the response spectrum was significantly different in regions with the frequency below 1 kHz. Figure 2(d) shows the comparison of frequency spectrum before and after vocal loading task in male group. Similarly, Fig. 2(e) shows the comparison of frequency spectrum before and after vocal loading task in female group. Obviously, the vocal loading task caused the variation in regions with low frequency. The results suggested that the vocal loading task could indeed trigger a mild inflammation of the vocal cords and make the voice hoarse and deeper.

Figure 3 shows the variation of average fundamental frequency before and after vocal loading task in the male and female groups. The average fundamental frequency in male group was lower than in female group. Additionally, the result indicated that the average fundamental frequency was significantly different between the male and female groups whether before or after vocal loading task. In the male group, the result of average fundamental frequency was significantly different before and after vocal loading task. However, the average fundamental frequency was not significantly different before and after vocal loading task in female group. Nevertheless, the



**Figure 3.** The analysis of average fundamental frequency variation before and after vocal loading task in male and female groups. (two sample *t*-test  $*p < 0.01$ ).



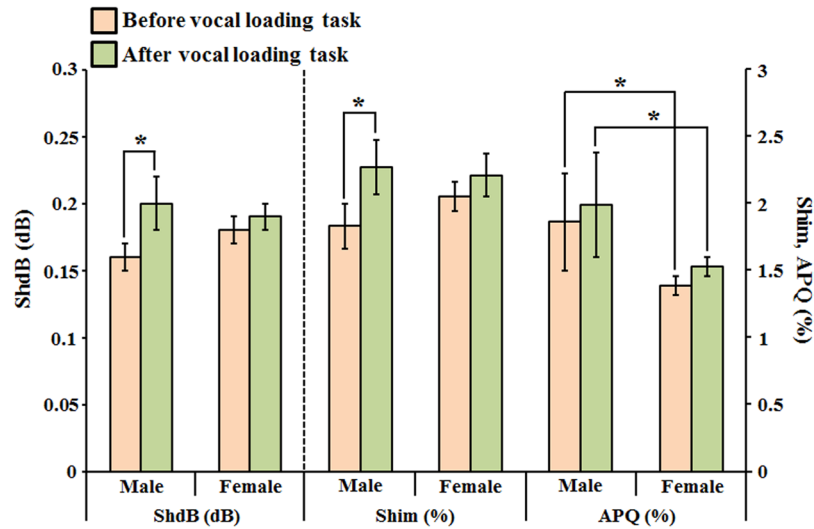
**Figure 4.** The analysis of the variation of parameters related to frequency perturbation before and after vocal loading task in male and female groups. (two sample *t*-test  $*p < 0.01$ ,  $**p < 0.001$ ).

fundamental frequency was lower after vocal loading task in both male and female groups. The result implied the worse voice quality after vocal loading task.

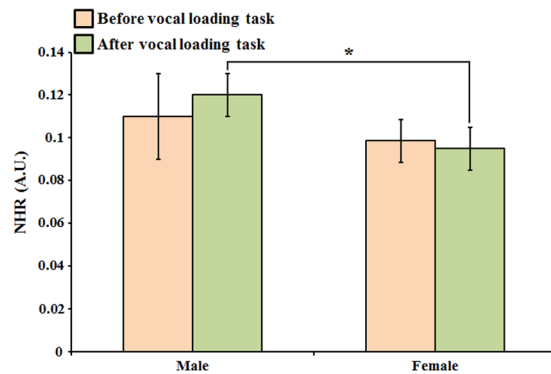
In this study, the acoustic parameters related to frequency perturbation were calculated to evaluate the differences between male and female groups. Additionally, the differences before and after vocal loading task were also analyzed in male and female groups, respectively. Figure 4 shows the variation of the parameters related to frequency perturbation. The results indicated that all the values of parameters in the male group were higher than in the female group whether before or after vocal loading task. Obviously, all the parameters were significantly different between the male and female groups. However, the result presented no significant difference before and after vocal loading task in the male group or in the female group. It is still noteworthy that all the value of parameters increased after vocal loading task. The increase in frequency perturbation represents the worse voice quality.

Likewise, the acoustic parameters related to amplitude perturbation were also calculated to evaluate the differences between male and female groups after and before vocal loading task. Figure 5 shows the variation of the parameters related to amplitude perturbation. The result indicated that all the parameters were not significantly different between male and female groups except the parameter of APQ. Besides, there were only ShdB and Shim presented significantly different before and after vocal loading task in the male group. It is also noteworthy that all the value of parameters of amplitude perturbation increased after vocal loading task.

Figure 6 shows the variation of noise-to-harmonic ratio (NHR) before and after vocal loading task in the male and female groups. The result presented significantly different between the male and female groups after vocal loading task. The value of NHR was not significantly different before and after vocal loading task in the male group or in the female group. Besides, the value of NHR increased after vocal loading task in the male group, and slightly decreased after vocal loading task in the female group.



**Figure 5.** The analysis of the variation of parameters related to amplitude perturbation before and after vocal loading task in male and female groups. (two sample *t*-test \* $p < 0.01$ ).

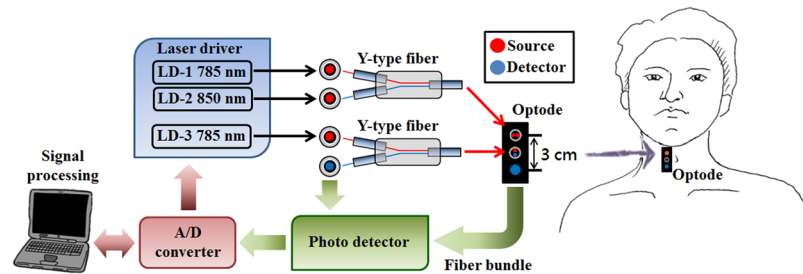


**Figure 6.** The analysis of NHR variation before and after vocal loading task in male and female groups. (two sample *t*-test, \* $p < 0.01$ ).

## Discussion

Acoustic analysis is the non-invasive and most used approach to voice quality assessment of voice disorder in clinical research and application<sup>36–40</sup>. Unlike previous studies of acoustic analysis that used the traditional microphone to record voice sample, the present study developed and used an optical technique (as an optical microphone) to record voice sample for acoustic analysis. This method was used to receive the vibration signals transmitted from the vocal cords to the surface of the tissue. The greatest benefit of the optical microphone is to avoid ambient sound and electromagnetic interference.

The spectrum analysis of signals in Fig. 2 showed that the optical method can indeed reconstruct the voice signal correctly. Additionally, the results of acoustic analysis of optical signal were also used to prove the feasibility of this method. In the Fig. 3, the result of average fundamental frequency in the male group was lower than in the female group that showed significant gender differences. Our result corresponds with several previous studies that also revealed lower fundamental frequency in healthy male group and significant gender differences<sup>41–46</sup>. The results implied that the pitch is affected by the size of vocal cords. The size (length) of vocal cords is larger in the adult male group than in the female group<sup>47,48</sup>. This is the reason why adult male voices are usually deeper and present lower fundamental frequency than female voices. In fact, the decreased fundamental frequency might not really reveal the poor voice quality but it is really correlated with the swelling of vocal cords and decreased mucosa wave because of the swelling of Ranke's space. There might be more swelling even some transient inflammatory response to vocal cords and cause decreased mucosa wave. Chronic vocal trauma even causes fluid accumulation in Ranke's space or higher collagen formation and decreased thus decreased vocal vibration and cause decreased fundamental frequency. Our results also showed that the fundamental frequency was significantly lower after vocal loading task in male group. Although the fundamental frequency was not significantly lower after vocal loading task in the female group, the fundamental frequency still decreased after vocal loading task. This finding implies that the voice quality changes (the voice sound “hoarser” and “deeper” after vocal loading task) in the male group more than in the female group after the same strength of the vocal loading task.



**Figure 7.** Multi-functional vocal cords detection system. LD-1 and LD-2 were used to calculate the concentration changes in oxygenated hemoglobin ( $\Delta[\text{HbO}_2]$ ) and deoxygenated hemoglobin ( $\Delta[\text{Hb}]$ ) for oxygen consumption evaluation; LD-3 was used to reconstruct the optical acoustic signals for acoustic analysis.

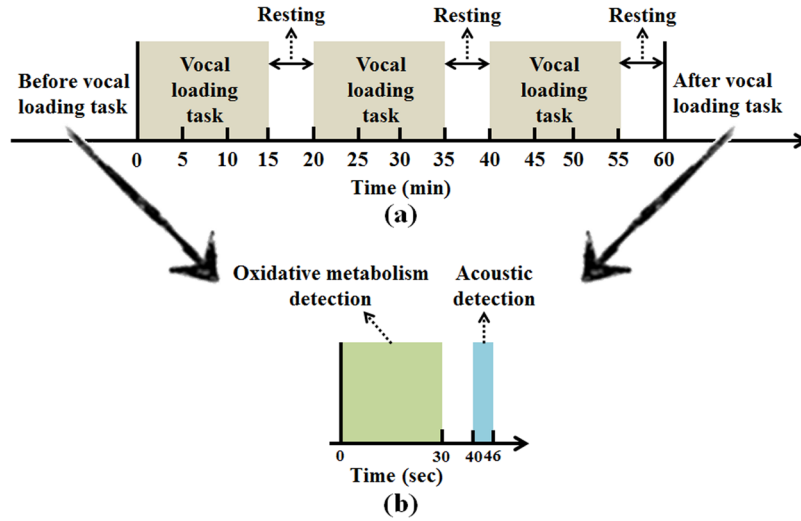
The increased frequency perturbation is responsible for hoarse, harsh or rough voice quality. Our results showed that all the parameters of frequency perturbation presented significant gender differences (as shown in Fig. 4). The parameters of frequency perturbation were higher in the male group than in the female group. These results implied that the adult male voices are rougher than female voices. Although the parameters of frequency perturbation were not significantly different after vocal loading task in all groups, the parameters still increased after vocal loading task. The results implied the worse vocal stability after vocal loading task. Same as frequency perturbation, the amplitude perturbation was used to evaluate the vocal stability. In the Fig. 5, the parameters of amplitude perturbation were not significantly different between male and female groups except APQ. The results are consistent with the previous study that used traditional voice assessment devices<sup>42</sup>. The ShdB and Shim were significantly higher after vocal loading task in the male group. It is noteworthy that all the parameters increased after vocal loading task. The results also implied the worse vocal stability after vocal loading task. The result of the NHR presented a significant difference between male and female groups after vocal loading task. This finding also implies that the voice quality decreases in the male group more than in the female group after the vocal loading task.

The acoustic analysis may be influenced by several factors include age, gender, vowel phonation, and vocal intensity<sup>41–46</sup>. Besides, acoustic analysis cannot be simply compared in different speech analysis software<sup>42</sup>. As mentioned before, voice disorder is a complex multifaceted clinical problem. The multidimensional assessment is very important for clinical applications of voice disorder. Additionally, oxidative metabolism may play a very important factor in vocal fatigue. Therefore, the oxidative metabolism was simultaneously detected before and after vocal loading task in this study. According to the result of previous study, hypoxia of vocal cords was induced by vocal overuse<sup>49</sup>. Our results demonstrated the vocal cord hypoxia after vocal loading task both in male and female groups (as shown in Fig. 1). It is noteworthy that the oxygen recovery was faster in the female group. This result was consistent with our results of the optical acoustic analysis that suggested the voice quality decreases in the male group more than in the female group after vocal loading task. Besides, this finding also implied that the recovery rate of oxygen consumption may relate to the size/structure of vocal cords. Therefore, integration of oxidative metabolism detection and acoustic assessment by using optical technique may provide more information for assisted diagnosis of vocal disorders.

## Methods

**Participants.** In this study, 60 healthy adults (30 males and 30 females), age between 20 to 25 years old (the mean age was  $22 \pm 2.15$  years old) without throat or vocal cord diseases at least one month, were recruited from National Chiao Tung University, Taiwan. All participants provided written informed consent. The study was in accordance with the latest version of the Declaration of Helsinki, and approved by the Institutional Review Board (IRB) in Research Ethics Committee for Human Subject Protection, National Chiao Tung University, Taiwan.

**Experimental Setup.** In this study, the non-invasive oxidative metabolism detection and acoustic assessment were integrated into a novel system by using the fNIRS-DOT, that we called it a multi-functional vocal cords detection system. Therefore, the information of oxidative metabolism and acoustics of the vocal cords could be captured simultaneously. Figure 7 shows the scheme of multi-functional vocal cords detection system. The system setup included a pair of laser diodes (785 nm and 850 nm) (WSLR-785/850–050 m-M-PD, Wavespectrum, Beijing, China), an independent laser diode (785 nm) and two silicon photo detectors (PDA100A, Thorlabs, Newton, New Jersey, U.S.). In function of the oxidative metabolism detection, the bifurcated fiber bundles, also called Y-type fibers (BFY400HS02, Thorlabs, New Jersey, US) was used to introduce laser light from the two laser heads into one optical fiber as the same emission point. The source-detector separation is 3 cm that can provide adequate depth of penetration (the penetration depth into tissue is approximately 1.5 cm or larger<sup>26–29</sup>) for oxidative metabolism detection of the vocal cords. In function of the acoustic assessment, the Y-type fiber was used to transmit and receive the optical signal as an optical microphone to detect the signal of the tissue surface vibration. All the backscattered optical signals from vocal cords and tissue surface vibration were detected with two silicon photo detectors (Si-PD), respectively. The amplified analog signals from the Si-PD were converted to digital data by using an A/D converter (NI myRIO-1900, National Instruments, Austin, Texas, U.S.) with 12-bit resolution. The A/D converter was also used to regulate the optimum laser diode drive current with the technique



**Figure 8.** The protocols of vocal loading task and data acquisition. (a) the protocol of vocal loading task; (b) the protocol of the oxidative metabolism detection and optical acoustic data acquisition.

of Time-Division Multiplexing (TDM) to obtain stable and accurate measurements. The optical power of laser diode was regulated at 5 mW. The optical fibers were fixed on a black flexible probe that was brought in close contact with the skin to prevent noise from the environmental and surface backscattering.

**Experimental Protocol and Data Acquisition.** Figure 8 shows the protocols of vocal loading task and optical signals measurement. During the experiment, subjects sat on a comfortable chair in a silent room. A totally 60-min vocal loading task consists of three 15-min intervals and 5-min vocal resting state between each task (as shown in Fig. 8(a)). During the vocal loading task, subjects were asked to read aloud at a level approximately 90 dB and a digital sound level meter was used to verify. According to previous study, this protocol could trigger a mild inflammation of the vocal cords<sup>13,14</sup>. Therefore, the optical measurements were performed before and after vocal loading task according to the following three steps: 1) oxidative metabolism detection for 30-sec; 2) acoustic detection after a 10-sec separation for 6-sec (as shown in Fig. 8(b)). In the step of oxidative metabolism detection, subjects kept in the vocal resting state for measurement. In the step of acoustic detection, subjects were asked to voice the sustained vowel /a/ at a flat tone and comfortable pitch that without breaks for at least 6 seconds.

### Data Analysis

**Oxidative metabolism detection.** In this function, the backscattered optical signals of the dual-wavelength (785 nm and 850 nm) from human tissue were received for calculating the concentration changes in  $\Delta[\text{HbO}_2]$  and  $\Delta[\text{Hb}]$  with technique of continuous-wave and a sampling rate of 50 Hz. According to the modified Beer-Lambert Law (MBLL)<sup>25,50,51</sup>, the optical density (OD) can be defined as follows:

$$OD_\lambda = -\log_{10}\left(\frac{I}{I_0}\right) = \varepsilon_\lambda CLB_\lambda + G_\lambda \tag{1}$$

where  $I_0$  and  $I$  are the intensities of incident light and detected light, respectively; The  $OD_\lambda$  is the optical density for wavelength  $\lambda$  that means the attenuation of near-infrared light intensity in tissue;  $\varepsilon_\lambda$  is the extinction coefficient of the chromophore;  $C$  is the concentration of the chromophore;  $L$  is the distance between the light entry and exit the tissue;  $B_\lambda$  is a pathlength factor that related to tissue scattering;  $G_\lambda$  is defined as a geometric factor that related to tissue geometry. Assuming that  $G_\lambda$  remains constant during a measurement, the change in optical density  $\Delta OD_\lambda$  can be obtained as follows:

$$\Delta OD_\lambda = OD_\lambda(t) - OD_\lambda(t_0) = -\log_{10}\left(\frac{I(t)}{I(t_0)}\right) = \varepsilon_\lambda \Delta CLB_\lambda \tag{2}$$

where  $OD_\lambda(t_0)$  and  $OD_\lambda(t)$  are the initial and instantaneous values of the optical density from the tissue, respectively;  $I(t_0)$  and  $I(t)$  are the measured intensities at initial and instantaneous time;  $\Delta C$  is the change in concentration of the chromophore. In human tissue, changes in concentration were dominated with the chromophore of  $\text{HbO}_2$  and  $\text{Hb}$ . Therefore, the change in optical density can be obtained as follows:

$$\Delta OD_\lambda = (\varepsilon_\lambda^{\text{HbO}_2} \cdot \Delta[\text{HbO}_2] + \varepsilon_\lambda^{\text{Hb}} \cdot \Delta[\text{Hb}]) \cdot LB_\lambda \tag{3}$$

The concentration changes in  $\Delta[\text{HbO}_2]$  and  $\Delta[\text{Hb}]$  could be obtained by solving Eq. (3) with the optical signals of dual-wavelength (785 nm and 850 nm). Assuming that  $L$  and  $B_\lambda$  remains constant during a measurement, the change in optical density can be obtained as follows:

$$\Delta OD_{785} = (\varepsilon_{785}^{\text{HbO}_2} \cdot \Delta[\text{HbO}_2] + \varepsilon_{785}^{\text{Hb}} \cdot \Delta[\text{Hb}]) \quad (4)$$

$$\Delta OD_{850} = (\varepsilon_{850}^{\text{HbO}_2} \cdot \Delta[\text{HbO}_2] + \varepsilon_{850}^{\text{Hb}} \cdot \Delta[\text{Hb}]) \quad (5)$$

Finally, the description can be rewritten from Eq. (4) and Eq. (5) as follows:

$$\Delta[\text{HbO}_2] = \frac{\varepsilon_{785}^{\text{Hb}} \cdot \Delta OD_{850} - \varepsilon_{850}^{\text{Hb}} \cdot \Delta OD_{785}}{\varepsilon_{785}^{\text{Hb}} \cdot \varepsilon_{850}^{\text{HbO}_2} - \varepsilon_{850}^{\text{Hb}} \cdot \varepsilon_{785}^{\text{HbO}_2}} \quad (6)$$

$$\Delta[\text{Hb}] = \frac{\varepsilon_{850}^{\text{HbO}_2} \cdot \Delta OD_{850} - \varepsilon_{785}^{\text{HbO}_2} \cdot \Delta OD_{785}}{\varepsilon_{785}^{\text{Hb}} \cdot \varepsilon_{850}^{\text{HbO}_2} - \varepsilon_{850}^{\text{Hb}} \cdot \varepsilon_{785}^{\text{HbO}_2}} \quad (7)$$

In this study, the concentration changes of  $\Delta[\text{HbO}_2]$  and  $\Delta[\text{Hb}]$  were used to analyze the oxidative metabolism of vocal cords before and after vocal loading task.

**Acoustic assessment with optical microphone.** In this function, the backscattered optical signal of the independent laser diode (785 nm) from the tissue surface vibration was detected with a sampling rate of 50 kHz during subjects voiced the sustained vowel /a/. The sustained vowels /a/, /u/, /o/, and /i/ are usually used in clinical acoustic assessments. However, the sustained vowel /a/ would enhance measurement reliability<sup>43</sup>. The spectrum distribution was calculated with the algorithm of fast Fourier transform (FFT). Additionally, the software of Multi-Dimensional Voice Program (MDVP) was utilized to analyze voice parameters included average fundamental frequency, frequency perturbation (Jita, Jitt, PPQ, and RAP), amplitude perturbation (ShdB, Shim, and APQ), and noise-to-harmonic ratio (NHR)<sup>52–55</sup>. The choice of these voice parameters were designated by the clinician. The description of each of the extracted parameter was listed as follows: Average fundamental frequency ( $f_0$ , Hz):  $f_0$  is an average value of all extracted period-to-period fundamental frequency value that is calculated from the extracted period-to-period pitch data; Absolute jitter (Jita,  $\mu\text{sec}$ ): To calculate the period-to-period variability of the pitch period that dependent on the average fundamental frequency. Therefore, Jita is typically related to hoarse voices; Jitter percent (Jitt, %): Jitter percent is a relative measure of a very short-term variability of the pitch period. It also related to hoarse voices. The higher value of Jita and Jitt means unstable of voice quality; Pitch perturbation quotient (PPQ, %): PPQ is a relative measure of pitch perturbation in short term cycle (smoothing cycle of 5 periods) of voice analysis. Hoarse and breathy voices may cause PPQ to increase; Relative average perturbation (RAP, %): RAP is a relative measure of average pitch perturbation in a short-term cycle (smoothing cycle of 3 periods) of voice analysis. Hoarse and breathy voices also may cause PPQ to increase; Shimmer in dB (ShdB, dB): To calculate in dB of the very short term variability of the peak-to-peak amplitude of the voice that is very sensitive to amplitude variations between consecutive pitch periods. It is also related to hoarse and breathy voices; Shimmer percent (Shim, %): Shim is a relative measure of a very short-term variability of the peak-to-peak amplitude of the voice. It is typically related to hoarse and breathy voices; Amplitude perturbation quotient (APQ, %): APQ is a relative measure of a short term (smoothing cycle-to-cycle of 11 periods) irregularity of the peak-to-peak amplitude of the voice. Hoarse and breathy voices may also cause APQ to increase; Noise-to-harmonics ratio (NHR, A.U.): NHR is defined as the ratio of the energy of noise and the energy of harmonic spectral during the range of 70–4200 Hz. The lower value of NHR indicates the better voice quality. Contrarily, the higher value of NHR is interpreted as more spectral noise that may be caused by the frequency and amplitude variations, turbulent noise, subharmonic components or voice breaks.

In the both two functions, the average data was obtained for group-level analysis to reduce the effects of individual differences. The results were expressed as the mean  $\pm$  SD. The significant differences analysis between healthy male and female group was made with a two-sample  $t$ -test. Additionally, the significant difference analysis before and after vocal loading task was also made with a two-sample  $t$ -test. Analyses were performed with software of LabVIEW (Version 2017, National Instruments, Austin, Texas, U.S.), MATLAB (Version 7.11.0.584 R2010b, MathWorks Inc., Natick, MA, U. S.) and Multi-Dimensional Voice Program (Version MDVP model 5150, Lincoln Park, NJ, U.S.). A more stringent  $p$  value of  $<0.01$  was considered as statistically significant in two-sample  $t$ -test.

## Conclusion

In this study, the multi-functional vocal cords detection system was developed by using optical technique. This system can provide a novel and non-invasive detection approach for human vocal fold oxidative metabolism detection and acoustic assessment simultaneously. Our results demonstrated that the physiological analysis of vocal cords (include oxidative metabolism detection and acoustic assessment) before and after vocal loading task could be successfully measured by using an optical method. Therefore, this optical method could be a potential tool for clinical application of voice disorder. Although there are still several limitations<sup>31</sup>, fNIRS could provide relevant information on key mechanisms of oxidative metabolism of different organs by performing a clear and reliable hypothesis of the test protocol. In the future study, the system will be optimized for more clinical applications such as postoperative prognosis and monitoring. We hope that the proposed method can provide more clinical information to help ENT physician to develop the treatment strategy and therapeutic monitoring of vocal disorders.



## References

- Nanjundeswaran, C., Jacobson, B. H., Gartner-Schmidt, J. & Verdolini, A. K. Vocal fatigue index (VFI): Development and validation. *J. Voice*. **29**, 433–440 (2015).
- Caratya, M.-J. & Montacié, C. Vocal fatigue induced by prolonged oral reading: Analysis and detection. *Comput. Speech. Lang.* **28**, 453–466 (2014).
- Costa, V. D., Prada, E., Roberts, A. & Cohen, S. Voice disorders in primary school teachers and barriers to care. *J. Voice*. **26**, 69–76 (2015).
- Charn, T. C. & Mok, P. K. H. Voice problems amongst primary school teachers in Singapore. *J. Voice*. **26**, 141–147 (2012).
- Chen, S. H., Chiang, S.-C., Chung, Y.-M., Hsiao, L.-C. & Hsiao, T.-Y. Risk factors and effects of voice problems for teachers. *J. Voice*. **24**, 183–190 (2010).
- Williams, N. R. Occupational groups at risk of voice disorders: a review of the literature. *Occup. Med.* **53**, 456–460 (2003).
- Vilkman, E. Occupational safety and health aspects of voice and speech professions. *Folia. Phoniatr. Logop.* **56**, 220–253 (2004).
- Yiu, E. M.-L. *et al.* Quantitative high-speed laryngoscopic analysis of vocal fold vibration in fatigued voice of young karaoke singers. *J. Voice*. **27**, 753–761 (2013).
- Schwartz, S. R. *et al.* Clinical practice guideline: hoarseness (dysphonia). *Otolaryngol Head Neck Surg.* **141**, S1–S31 (2009).
- Hanschmann, H., Gaigl, C. & Berger, R. Preliminary results of a computer-assisted vocal load test with 10-min test duration. *Eur Arch Otorhinolaryngol.* **268**, 309–313 (2011).
- Ben-David, A., Finck, C., Roche, A. & Morsomme, D. Vocal impact of a prolonged reading task at two intensity levels: objective measurements and subjective self-ratings. *J. Voice*. **26**, e177–e186 (2012).
- Dehqan, A. & Scherer, R. C. Acoustic analysis of voice: Iranian teachers. *J. Voice*. **27**, e17–e21 (2013).
- Echternach, M., Nusseck, M., Dippold, S., Spahn, C. & Richter, B. Fundamental frequency, sound pressure level and vocal dose of a vocal loading test in comparison to a real teaching situation. *Eur Arch Otorhinolaryngol.* **271**, 3263–3268 (2014).
- Ingle, J. W. *et al.* Role of steroids in acute phonotrauma: A basic science investigation. *Laryngoscope.* **124**, 921–927 (2014).
- Whitling, S., Rydell, R. & Lyberg Åhlander, V. Design of a clinical vocal loading test with long-time measurement of voice. *J. Voice*. **29**, e13–e27 (2015).
- Ben-David, B. M. & Icht, M. Voice changes in real speaking situations during a day, with and without vocal loading: Assessing call center operators. *J. Voice*. **30**, e1–e11 (2016).
- Gorham-Rowan, M., Berndt, A., Carter, M. & Morris, R. The effect of a vocal loading task on vocal function before and after 24 hours of thickened liquid use. *J. Speech Pathol. Ther.* **1**, 1–5 (2016).
- Welham, N. V. & Maclagan, M. A. Vocal fatigue: current knowledge and future directions. *J. Voice*. **17**, 21–30 (2003).
- Gillivan-Murphy, P., Drinnan, M. J., O'Dwyer, T. P., Ridha, H. & Carding, P. The effectiveness of a voice treatment approach for teachers with self-reported voice problems. *J. Voice*. **20**, 423–431 (2006).
- Chen, S. H., Hsiao, T. Y., Hsiao, L. C., Chung, Y. M. & Chiang, S. C. Outcome of resonant voice therapy for female teachers with voice disorders: perceptual, physiological, acoustic, aerodynamic, and functional measurements. *J. Voice*. **21**, 415–425 (2007).
- Boucher, V. J. & Ayad, T. Physiological attributes of vocal fatigue and their acoustic effects: a synthesis of findings for a criterion-based prevention of acquired voice disorders. *J. Voice*. **24**, 324–336 (2010).
- Jiang, J. J. & Maytag, A. L. Aerodynamic measures of glottal function: what extra can they tell us and how do they guide management? *Curr. Opin. Otolaryngol Head Neck Surg.* **22**, 450–454 (2014).
- Behlau, M., Madazio, G. & Oliveira, G. Functional dysphonia: strategies to improve patient outcomes. *Patient Related Outcome Measures.* **6**, 243–253 (2015).
- Bale, G., Elwell, C. E. & Tachtsidis, I. From Jöbsis to the present day: a review of clinical near-infrared spectroscopy measurements of cerebral cytochrome-c-oxidase. *J. Biomed. Opt.* **21**(091307), 1–18 (2016).
- Scholkmann, F. *et al.* A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *Neuroimage.* **85**, 6–27 (2014).
- Kravari, M., Angelopoulos, E., Vasileiadis, I., Gerovasili, V. & Nanas, S. Monitoring tissue oxygenation during exercise with near infrared spectroscopy in diseased populations-A brief review. *Int. J. Ind. Ergonom.* **40**, 223–227 (2010).
- Hamaoka, T., McCully, K. K., Niwayama, M. & Chance, B. Phil. The use of muscle near-infrared spectroscopy in sport, health and medical sciences: recent developments. *Trans. R. Soc.* **28**, 4591–4604 (2011).
- Binzoni, T. & Spinelli, L. Near-infrared photons: a non-invasive probe for studying bone blood flow regulation in humans. *J. Physiol Anthropol.* **34**, 1–6 (2015).
- Koga, S. *et al.* Effects of increased skin blood flow on muscle oxygenation/deoxygenation: comparison of time-resolved and continuous-wave near-infrared spectroscopy signals. *Eur. J. Appl. Physiol.* **115**, 335–343 (2015).
- Seong, M. *et al.* Simultaneous blood flow and blood oxygenation measurements using a combination of diffuse speckle contrast analysis and near-infrared spectroscopy. *J. Biomed Opt.* **21**(027001), 1–6 (2016).
- Grassia, B. & Quaresima, V. Near-infrared spectroscopy and skeletal muscle oxidative function *in vivo* in health and disease: a review from an exercise physiology perspective. *J. Biomed Opt.* **21**(091313), 1–20 (2016).
- Bucaroa, J. A. & Lagakos, N. Lightweight fiber optic microphones and accelerometers. *Rev. Sci. Instrum.* **72**, 2816–2821 (2001).
- Kadirvel, K. *et al.* 42<sup>nd</sup> AIAA 2004-1310 (2004).
- NessAiver, M. S., Stone, M., Parthasarathy, V., Kahana, Y. & Paritsky, A. Recording high quality speech during tagged cine-MRI studies using a fiber optic microphone. *J. Magn Reson Imaging.* **23**, 92–97 (2006).
- Teixeira, J. G. V., Leite, I. T., Silva, S. & Frazão, O. Advanced fiber-optic acoustic sensors. *Photonic sensors.* **4**, 198–208 (2014).
- Gillespie, A. I., Dastolfo, C., Magid, N. & Gartner-Schmidt, J. Acoustic analysis of four common voice diagnoses: moving toward disorder-specific assessment. *J. Voice*. **28**, 582–588 (2014).
- Teixeira, J. P. & Fernandes, P. O. Acoustic analysis of vocal dysphonia. *Procedia Computer Science.* **65**, 466–473 (2015).
- Batra, K., Bhasin, S. & Singh, A. Acoustic analysis of voice samples to differentiate healthy and asthmatic persons. *IJECS.* **4**, 13161–13164 (2015).
- Lin, F. C., Chen, S. H., Chen, S. C., Wang, C. T. & Kuo, Y. C. Correlation between acoustic measurements and self-reported voice disorders among female teachers. *J. Voice*. **30**, 460–465 (2016).
- Gillespie, A. I., Gartner-Schmidt, J., Lewandowski, A. & Awan, S. N. An examination of pre- and posttreatment acoustic versus auditory perceptual analyses of voice across four common voice disorders. *J. Voice*. Accepted for publication April 19, 2017. Article in press (2017).
- Ting, H. N., Chia, S. Y., Abdul Hamid, B. & Mukari, S. Z. Acoustic characteristics of vowels by normal Malaysian Malay young adults. *J. Voice*. **25**, e305–e309 (2011).
- Ting, H. N., Chia, S. Y., Kim, K. S., Sim, S. L. & Abdul Hamid, B. Vocal fundamental frequency and perturbation measurements of vowels by normal Malaysian Chinese adults. *J. Voice*. **25**, e311–e317 (2011).
- Brockmann, M., Drinnan, M. J., Storck, C. & Carding, P. N. Reliable jitter and shimmer measurements in voice clinics: the relevance of vowel, gender, vocal intensity, and fundamental frequency effects in a typical clinical task. *J. Voice*. **25**, 44–53 (2011).
- Yamauchi, A. *et al.* Age- and gender-related difference of vocal fold vibration and glottal configuration in normal speakers: analysis with glottal area waveform. *J. Voice*. **28**, 525–531 (2014).
- Yamauchi, A. *et al.* Quantitative analysis of digital videokymography: a preliminary study on age- and gender-related difference of vocal fold vibration in normal speakers. *J. Voice*. **29**, 109–119 (2015).

46. Lovato, A. *et al.* Multi-dimensional voice program (MDVP) vs Praat for assessing euphonic subjects: A preliminary study on the gender-discriminating power of acoustic analysis software. *J. Voice*. **30**, 765.e1–765.e5 (2015).
47. Su, M. C. *et al.* Measurement of adult vocal fold length. *J. Laryngol Otol*. **116**, 447–449 (2002).
48. Borkowska, B. & Pawlowski, B. Female voice frequency in the context of dominance and attractiveness perception. *Anim. Behav.* **82**, 55–59 (2011).
49. Fang, R., Jiang, J. J., Smith, B. L. & Wu, D. Expression of hypoxia inducible factor-1 $\alpha$  and vascular endothelia growth factor in vocal polyps. *Laryngoscope*. **123**, 2184–2188 (2013).
50. Boas, D. A. *et al.* The accuracy of near infrared spectroscopy and imaging during focal changes in cerebral hemodynamics. *Neuroimage*. **13**, 76–90 (2001).
51. Kocsis, L., Herman, P. & Eke, A. The modified Beer-Lambert law revisited. *Phys. Med. Biol.* **51**, N91–N98 (2006).
52. Kent, R. D., Vorperian, H. K., Kent, J. F. & Duffy, J. R. Voice dysfunction in dysarthria: application of the Multi-Dimensional Voice Program. *J. Commun. Disord.* **36**, 281–306 (2003).
53. Nicastri, M., Chiarella, G., Gallo, L. V., Catalano, M. & Cassandro, E. Multidimensional Voice Program (MDVP) and amplitude variation parameters in euphonic adult subjects. Normative study. *Acta. Otorhinolaryngol Ital.* **24**, 337–341 (2004).
54. Godino-Llorente, J. I. *et al.* Acoustic analysis of voice using WPCVox: a comparative study with Multi Dimensional Voice Program. *Eur. Arch. Otorhinolaryngol.* **265**, 465–476 (2008).
55. Maryn, Y., Corthals, P., De Bodt, M., Van Cauwenberge, P. & Deliyski, D. Perturbation measures of voice: a comparative study between Multi-Dimensional Voice Program and Praat. *Folia. Phoniatr. Logop.* **61**, 217–226 (2009).

## Acknowledgements

This work was supported in part by the Taiwan National Science Council under Grant Nos. MOST 104-2221-E-009 -192 -MY3, and a grant from Ministry of Education, Aim for the Top University Plan in National Chiao-Tung University 106W970.

## Author Contributions

C.-C. Chuang and T.-C. Lin wrote the main manuscript text; J.-C. Chen and C.-H. Liu recruited the participants and collected the data; C.-C. Chuang and C.-Y. Lee developed the system and algorithm for signal processing and prepared the figures; C.-C. Chuang drew the Figure. 7; Y.-A. Tsou collected the literatures to provide discussion and clinical advice; C.-C. Chuang also performed the statistical analyses. All authors have reviewed and agreed to all of the content in the manuscript.

## Additional Information

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017