DYNAMIC B-MODE ULTRASOUND IMAGING OF VOCAL FOLD VIBRATION DURING PHONATION

CHEN-GIA TSAI,* JENG-HORNG CHEN,† YIO-WHA SHAU,‡ and TZU-YU HSIAO§

*Graduate Institutes of Musicology, National Taiwan University, Taipei, Taiwan; †Applied Mechanics, National Taiwan University, Taipei, Taiwan and Industrial Technology Research Institute, Hsin-Tsu, Taiwan; ‡Department of Systems and Naval Mechatronic Engineering, National Cheng Kung University, Tainan, Taiwan; and §Department of Otolaryngology, National Taiwan University Hospital and National Taiwan University, College of Medicine, Taipei, Taiwan

(Received 18 February 2009; revised 25 May 2009; in final form 4 June 2009)

Abstract—We used B-mode imaging to study the vibratory phenomena of the vocal folds. The presence of multilayered structures of the vocal folds in the B-mode image was verified by using freshly excised human larynges in vitro. To capture images of vocal fold vibration, a special treatment was used to reconstruct the aliasing B-mode motion pictures of vocal fold vibration. Echo-particle image velocimetry (Echo-PIV) analysis was then applied to trace the tissue particles in the motion pictures. The vibratory behavior of the body (vocal ligament and muscle) of the vocal folds was revealed. Further analysis showed a quasi-longitudinal wave along the body of the vocal folds in the coronal plane. This is, to the best of our knowledge, the first time that vocal fold vibration physiology has been studied using B-mode imaging and Echo-PIV. (E-mail: tyhsiao@ntu.edu.tw)

KEY WORDS: Vocal fold, Vibration, Phonation, Ultrasound image, Particle image velocimetry, B-mode.

INTRODUCTION

In recent decades, medical ultrasound (US) imaging has become a very powerful tool for diagnostic radiology, especially in scanning the head and neck regions (Chevallier et al. 2002). B-mode US imaging has been used with some success in the laryngeal area to identify mass and cystic lesions at the vocal folds (Rubin et al. 2004; Schadel et al. 2003) and to detect vocal fold paralysis (Vats et al. 2004; Ooi et al. 1995). Given its noninvasive nature and minimal disturbance to normal voice production, medical US imaging should be an ideal tool for studying vocal fold vibration. However, only a few studies have used dynamic US imaging to investigate the vibrating vocal folds (Hsiao et al. 2001, 2002; Shau et al. 2001).

The vibration of the vocal folds propelled by the glottal airflow periodically interrupts the airflow, causing a pulse train of transient sounds at periodic intervals, which forms the source for human voice (Hirano 1975). The unique multilayered structure of the vocal folds plays an important role in the process. The human vocal fold is divided into the “cover” and the “body” layers, each with different mechanical properties (Hirano 1974). Vocal fold vibration is confined mainly to the pliable cover, which consists of the mucosa and the superficial layer of the lamina propria, which is also known as Reinke’s space (Saito et al. 1983). Underneath the cover, the vocal ligament and the vocal muscle form the body of the vocal folds, which may also be involved in vibration. The vibrating wave is mostly observed in the cover layer and the wave travels vertically from the lower to the upper margin of the vocal fold (Titze 1988). The body motion of the vocal folds cannot be investigated by endoscopy. Previous measurements of the body dynamics of the vocal folds were exclusively performed in vitro (Saito et al. 1983). The frame rate (fs) of dynamic B-mode US (<50 Hz) is always lower than the vibration frequency (f) of the vocal folds (>70 Hz). Under this condition, real-time dynamic motion pictures of vocal fold vibration cannot be obtained using B-mode imaging. If the phonation frequency is tuned to be close to an integer multiple of fs, a slow motion montage of vocal fold vibration may be obtained with B-mode imaging. Such motion pictures may provide valuable data for analyzing vocal fold vibration especially in the body. This unique method

Address correspondence to: Tzu-Yu Hsiao, M.D., Ph.D., Department of Otolaryngology, National Taiwan University Hospital, #7, Chung-San S. Road, Taipei, Taiwan. E-mail: tyhsiao@ntu.edu.tw

Video Clips cited in this article can be found online at: http://www.umbjournal.org.

doi:10.1016/j.ultrasmedbio.2009.06.002

1812
of obtaining aliasing B-mode motion pictures has not been described before.

This article reports our experiences in obtaining and analyzing dynamic B-mode US motion pictures of human vocal fold vibration during phonation. We first sought to verify the multilayered structure of the vocal folds in the US images in vitro. Subsequently, reconstructed aliasing B-mode motion pictures of vocal fold vibration were successfully recorded in healthy subjects during steady modal phonation. Speckles in these images were processed by the particle image velocimetry (PIV) software, which revealed vibratory patterns of the vocal fold body under normal phonation. This is the first time that vocal fold vibration physiology has been investigated using B-mode imaging and echo-particle image velocimetry (Echo-PIV). We further observed a quasi-longitudinal wave in the coronal plane of the vocal fold body.

MATERIALS AND METHODS

A medical high-resolution US scanner (HDI-5000; ATL, Bothell, WA, USA) with a linear-array transducer (CL10-5 25 mm, ATL) was used to record B-mode US images throughout this study. US examinations were performed by the senior author (T.Y.H.). The default settings in tissue specific presets (small parts/superficial or Mskel/S/K) were chosen for the entire examinations. The other parameters set for the B-mode imaging were 2.0 cm depth, high frame rate, medium line density and medium persistence. Regular aqueous coupling gel was used in both the excised larynges and normal subjects as an acoustic coupling medium for US imaging. This study was approved by the Institutional Review Board of National Taiwan University Hospital, Taipei, Taiwan (Approval No. 200705035R). All enrolled subjects gave written informed consent before the experiment.

Verification of the multilayered structure of the vocal folds in vitro with B-mode imaging

Seven freshly excised larynges obtained from patients with hypopharyngeal carcinoma (all male, aged from 62 to 75) and who received total laryngectomy were studied within 6 h of surgical excision. The specimens were carefully examined to make sure at least one side of the vocal fold was free of tumor and that the structure of the vocal fold was grossly normal so that US imaging can verify the multilayered structures in the tissues. The posterior glottis was split vertically at the midline to expose the larynx for easy manipulation. A small vertical piece of calcified cartilage in the middle of thyroid lamina, about 0.3 to 0.5 cm in width, was removed subperichondrally with a rongeur without disturbing the remaining laryngeal framework. This thyroid lamina window was made to improve transmission of ultrasonic signals. The US scanner was applied at the midline of the thyroid cartilage lamina and along the coronal plane and programmed to scan through the pre-made cartilage window. The locations of the vocal fold, the laryngeal ventricle and the ventricular fold were verified in real-time B-mode imaging by placing micro-forceps on the respective mucosa. After identifying the general contours of the hemi-larynx in the coronal plane of the B-mode image, the epithelium of the vocal fold, Reinke’s space, the vocal ligament and the vocal muscle were removed layer by layer by microdissection to verify the corresponding presentations of each structure. The stripped tissues were used for histologic confirmation of their components.

Dynamic B-mode US imaging of vocal fold vibration during phonation

Three male subjects, 20 to 37 years of age and without any history of voice disorders, participated in the in vivo vocal fold vibration experiment. The loudness of phonation was 75 to 80 dB SPL 20 cm from the mouth (Type 2235; Brüel and Kjær, Nærum, Denmark). Loudness measurement was performed only before US image recording as better imaging could be obtained in loud voice. The US scan probe was placed at the midline in the coronal plane of the right thyroid cartilage lamina during sustained loud modal phonation. The US scanning direction was approximately parallel to the propagation direction of the vocal fold's mucosal waves to record the vertically conducting waves (Fig. 1). The US imaging frame rate was 40 Hz. Dynamic B-mode US images were recorded at a phonation frequency of 121 to 125 Hz. This phonation frequency was 1 to 5 Hz (p) higher...
than 3 (an integer multiple, N) times the frame rate of US imaging. Next, we obtained slow motion reconstruction montage of vocal fold vibration and observed that the strobe-like vibration frequency in the motion pictures of the dynamic B-mode imaging was equal to \( f_s \). The frame number per visualized vibration period, as denoted by \( K \), was equal to \( f_s/p \).

**Figure 2** schematically describes this mechanism of wave propagation on a string. A continuous sinusoidal wave was traveling along a string in the \(-x\) direction and at the same time the US beam scanned in the \(+x\) direction. The US beam scanned line-by-line across the viewing window as the string vibrated and the reconstruction of each B-mode frame took approximately \( N \) periods of string vibration to be completed. Two consecutive B-mode frames are shown by the cross-sections in black planes in Fig. 2. Consequently, these two B-mode frames showed a sinusoidal wave traveling in the \(-x\) direction at a much slower velocity.

**Echo particle image velocimetry (Echo-PIV)**

Optical PIV is a flow velocity field mapping technique that measures particle displacements (\( D_x \) and \( D_y \)) from statistical correlations of two consecutive images used in fluid flows. When the time delay between these two images (\( D_t \)) is sufficiently small, \( D_x/D_t \) and \( D_y/D_t \) give good approximation of the two-dimensional (2-D) velocity fields. The mathematical theories and fundamental techniques of optical PIV were described in earlier studies (Adrain 1991; Westerweel 1997). Several sophisticated techniques have also been developed since then to increase the processing speed (Raffel et al. 1998).

![Fig. 2. Schematic of the formation of aliasing ultrasound (US) motion pictures. A continuous sinusoidal wave traveled in the \(-x\) direction on a string. The US beam scanned in the \(+x\) direction with the frame rate \( f_s \). The sinusoidal wave frequency \( f \) is related to \( f_s \) by \( f=Nf_s+p \), with \( N=3 \) and \( p>0 \). The two black planes represented two consecutive B-mode frames. Comparing these two frames yielded a sinusoidal wave propagation in the \(-x\) direction.](image-url)

**RESULTS**

**Presentation of the multilayered structure of the vocal folds in vitro with B-mode imaging**

An example B-mode image of the hemi-larynx in the coronal section is shown in Fig. 3a. The thick white line indicates the air-mucosa interface where a strong acoustic reflection occurred. Underneath this white line, we observed a hypoechoic area about 2 mm in depth throughout the vocal fold region. This was verified as Reinke’s space area by layered microdissection and further histologic confirmation. The relatively bright regions beneath Reinke’s space represent the vocal ligament and the vocal muscle layers (Fig. 3b). There was no clear boundary between these two layers in the B-mode images.

**Reconstruction of aliasing dynamic B-mode US imaging of vocal fold vibration during phonation**

Strobe-like slow motion pictures of vocal fold vibrations were obtained during phonation in real-time dynamic B-mode US imaging (Video 1). A cycle of visualized vocal fold vibration contained 18 images (K). The phonation frequency was approximately 122 Hz (\( N=3 \)), about the average frequency of male phonation. One of the US images in a visualized vibration cycle and its inverse image with the indicated anatomic structures are shown in Fig. 4. The major limitation of this technique is that the subject must be able to maintain a steady phonation frequency throughout the recording. The mucosal wave appeared to propagate from the lower margin to the upper margin of the vocal fold. The cover was the most pliable part and with the greatest excursion during vibration. The speckles at the air-mucosa interface traveled in a circular fashion with a diameter of approximately 2 to 3 mm. This vibration pattern was consistent with previous measurements of the medial surface of the
vocal fold (Berry et al. 2001). The speckles in the body also moved along with the vibration.

Dynamic B-mode imaging data analysis with Echo-PIV

The US image in Fig. 5a was overlapped with the instantaneous velocity fields generated by PIV analysis of vibratory movement of the tissue speckles. This series analysis covered more than half of a vocal fold vibration cycle. The speckles in the B-mode images were the results of interference effects between overlapping echoes. Such an occurrence is related to a great number of randomly distributed structures that scatter within a resolution cell (Burckhardt 1978). A speckle is essentially a deterministic artifact during US imaging of a tissue (Burckhardt 1978; Wells and Halliwell 1981). This deterministic property justified the use of speckles as tissue tracers in echo-PIV. Due to its water-rich and hypoechoic characteristics, the tissue speckles in Reinke’s space in the US images were not clear enough for PIV processing. Therefore, PIV analysis focused on the vocal ligament and the vocal

Fig. 3. B-mode images of an excised human hemi-larynx in a coronal plane. (a) Intact. (b) Inverse image of (a) with the indicated anatomic structures. (c) Reinke’s space was excised. The vocal ligament was exposed to the lumen of the larynx. (d) The excised larynx corresponding to the image in (c) showed that the mucosa and the Reinke’s space were dissected. (e) Reinke’s space and the vocal ligament were excised. The vocal muscle was exposed to the lumen of the larynx. (f) The excised larynx corresponding to the image in (e) showed the exposed vocal muscle.
muscle regions. The maximal tissue velocity was found in the region immediately underneath Reinke’s space and primarily in the vertical direction. The velocity fields suggested a periodic, wavy motion of tissues in the vocal ligament and the muscle layers. A vertical phase difference in these two layers was observed in every image. A visualized half-wavelength was estimated as the distance between two tissue particles with locally maximal opposite vibration in the vertical direction. The vertical vibratory movement from the lower margin to the upper margin of the vocal folds indicated the wave propagation in the vocal ligament and the muscle layers. Velocity decreased as the depth increased throughout the vocal ligament and the vocal muscle layers. The trajectory distance pattern as calculated from the velocity fields is summarized in Fig. 5b. This pattern shows that the vocal ligament and muscle vibrated more in the vertical direction than in the horizontal direction during normal modal phonation. The maximal wave excursion calculated from Echo-PIV was approximately 0.62 mm vertically and 0.12 mm horizontally. Finally, a quasi-longitudinal wave was observed along the body of the vocal folds in the coronal plane.

**DISCUSSION**

In the present study, the vibratory phenomena of the human vocal fold body during normal phonation were unveiled for the first time using a novel treatment to reconstruct aliasing B-mode US motion pictures. Although several sophisticated studies have shown B-mode US imaging of the larynx, including vocal fold movements (Raghavendra et al. 1987; Miles 1989; Garel et al. 1990; Chevallier et al. 2002), there have been very few reports that used US imaging to study vocal fold vibration (Hsiao et al. 2001, 2002; Shau et al. 2001). The two main reasons for this deficiency are the lack of US imaging data on the multilayered structures of the vocal folds and the fact that the frame rate of the US imaging is far below the vibration frequency of the vocal folds. Vocal fold vibration critically depends on its unique layered anatomic structures (Hirano 1974). Therefore, before vocal fold vibration could be reliably studied using US imaging, the multilayered structures of the vocal folds must be identified using US images. We used excised larynges to verify the identities of the layered structures in the B-mode images (Fig. 3). These findings have not been demonstrated before. All of the excised larynges in this study were harvested from elderly male patients with hypopharyngeal carcinoma, which contained heavily calcified thyroid cartilages. A window on the calcified thyroid cartilage lamina was made to improve the transmission of US acoustic energy. Normally, calcification does not substantially hinder US energy transmission and this allows for satisfactory examination in most female and young male subjects (Chevallier et al. 2002). The white line in the B-mode image represents the interface between the air in the lumen and the mucosa and this observation was verified in a previous study (Hsiao et al. 2001). Most of the acoustic energy is reflected at this interface and the signals beyond this line are mainly artifacts. From our study, we observed that the hypoechoic area representing Reinke’s space was well correlated with its high water content. The vocal ligament together with the vocal muscle could also be identified underneath Reinke’s space in the images. However, the boundary between the vocal ligament and the muscle was not clear in the B-mode images.

The frame rate of US imaging is far below the vibration frequency of the vocal folds. However, if a subject phonated a steady voice with a fundamental frequency 1 to 5 Hz higher than the integer multiple of the frame rate of the B-mode imaging, a montage of dynamic vocal fold vibration images can be successfully obtained in slow-motion. This method, which has never been described before, can also in theory be applied to adult female and pediatric subjects.
Although particle image velocimetry (PIV) has been widely used in experimental fluid mechanics for decades, no attempt has been made to combine this technique with medical US imaging until the 2000s. In the pioneering work by Crapper et al. (2000), PIV was applied to obtain accurate, quantitative, 2-D vector maps of the flow field within the sediment suspension. To date, echo-PIV has been applied exclusively to study hemodynamics of blood flow in human intracranial aneurysms (Hölscher et al. 2006) and in porcine hearts (Sengupta et al. 2007). Unlike previous echo-PIV experiments using US contrast agents (Zheng et al. 2005; Hölscher et al. 2006; Sengupta et al. 2007; Vennemann et al. 2007), the speckles in the vocal folds served as tissue tracers for the first time in our present study.

In the aliasing dynamic B-mode imaging of vocal fold vibration, the vibratory wave of the air-mucosa interface of the vocal fold could be visualized very clearly. The speckles moved in a circular fashion with a diameter of approximately 2 to 3 mm. Echo-PIV was used to investigate the vibratory motion in the vocal ligament and the muscle layers during normal phonation. This technique revealed a predominance of vibratory motion in the vertical direction with a vertical phase difference between the lower and upper margins of the vocal fold. The maximum trajectory distance occurred in the region immediately underneath Reinke’s space, and the vibration amplitude attenuated as the depth increased. It is likely that the vocal ligament together with the superficial layer of the vocal muscle played a role in vocal fold vibration during normal phonation. However, in the present study, the boundary between them could not be verified in the B-mode images.

In previous studies, lead particles were introduced into regions of the vocal ligament and the vocal muscle to observe tissue trajectories in these regions during vibration in experiments involving excised larynges (Saito et al. 1983). This present study was the first time that the vibratory behaviors of the vocal ligament and vocal muscle were visualized in human subjects under normal phonation. These US data will be very useful in future biomechanical studies of the vocal fold vibration and can even be applied to study voice disorders in patients with suspected vocal fold paralysis. In previous theoretical models of vocal fold vibration, the vocal fold body (vocal ligament and vocal muscle) was commonly treated as a mass moving in the horizontal plane and the mucosal waves were often shown to propagate along the vocal fold body (Hirano 1974; Titze 1988; Story and Titze 1995). Our present study demonstrated that at least the vocal ligament was involved in the vibration along the vertical direction. There also appeared to be a vertical phase difference in the vibration wave of the vocal ligament. In other words, the vibratory behavior of the vocal ligament may be described as a quasi-longitudinal wave in the coronal plane. This wave propagated in the same direction as the mucosal wave. These surprising results showed some discrepancy with the currently popular models of vocal fold vibration. Further studies must be conducted to build a better model of vocal fold vibration, including one that more aptly describes the vertical vibratory movement of the vocal fold body.

This study represents the first attempt to obtain a dynamic B-mode imaging of vocal fold vibration. Although the fastest frame rate of the best US imaging equipment is still far slower than the rate of vocal fold vibration during phonation, it is still useful and informative to obtain aliasing motion pictures of vocal fold vibration with B-mode imaging. In the image acquisition process, the subject should be able to maintain a steady phonation frequency, which is close to an integer multiple of the frame rate of the US imaging machine. It is also
crucial that the US scanner be appropriately positioned in the coronal plane to scan the vibration of the vocal folds. Future researchers may need a little practice in acquiring the best images since the vocal fold position changes during phonation.

In summary, in this study the multilayered structures of the vocal folds were verified in B-mode images. Aliasing B-mode motion pictures of the vocal fold vibration were subsequently acquired. These US data were analyzed with the Echo-PIV software. The results revealed the vibratory behavior of the vocal fold body, notably a quasi-longitudinal wave observed in the coronal plane of the vocal ligament that was in the same direction as the mucosal wave. We demonstrated in this article that dynamic B-mode imaging can be used to investigate quasi-periodic tissue vibration with a vibration frequency higher than the frame rate of US imaging. Thus, for the first time, the vibratory pattern of the vocal fold body was observed in human subjects under normal phonation. As US imaging causes very little interference to the vocal folds during normal phonation, it is a promising tool for studying the physiology of the vocal folds in phonation function. Further studies must be conducted to clarify the role of the vocal fold body in the vibration of the vocal folds.

Acknowledgments—This research was supported by the National Science Council, Taiwan (Grant NSC 96-2314-B-002-081-MY3).

REFERENCES


Hirano M. Morphological structure of the vocal cord as a vibrator and its variations. Folia Phoniatr 1974;26:89–94.

Hirano M. Phonosurgery: Basic and clinical investigations. Otologia (Fukuoka) 1975;21(Suppl. 1):239–262.


SUPPLEMENTARY DATA

Supplementary data associated with this article can be found in the online version, at doi:10.1016/j.ultrasmedbio.2009.06.002.