19th INTERNATIONAL CONGRESS ON ACOUSTICS
MADRID, 2-7 SEPTEMBER 2007

NUMERICAL SIMULATION OF ATTENUATION CHARACTERISTICS OF SOFT-TISSUE CONDUCTED SOUND ORIGINATED FROM VOCAL TRACT

PACS: 43.35.Mr

Otani, Makoto[1]; Hirahara, Tatsuya[1]; Adachi, Seiji[2]
[1]Faculty of Engineering, Toyama Prefectural University; Kurokawa 5180, Imizu, Toyama, 939-0398, Japan; otani@pu-toyama.ac.jp
[2]Fraunhofer Institute for building physics; Nobelstrasse 12, 70569 Stuttgart, Germany;

ABSTRACT

A non-audible murmur (NAM), a very weak speech sound produced without vocal vibration, can be detected by a special NAM microphone attached to the neck, thereby providing a new communication tool for use with functional speech disorders. The NAM microphone is a condenser microphone covered with soft-silicone impression material that provides good impedance matching with the soft tissues of the neck. Because higher-frequency components are suppressed severely, however, the NAM detected with this device can be insufficiently clear. To improve NAM clarity, the mechanism of NAM production as well as the transfer characteristics of the NAM in soft neck tissues must be clarified. We have investigated sound propagation from the vocal tract to the neck surface, using a finite difference method and a head model based on magnetic resonance imaging scans. Numerical results show that, compared to vocal tract transfer functions, soft-tissue-conducted sound attenuates 50 dB at 1 kHz, which consists of 30 dB full-range attenuation due to air-to-soft-tissues transmission loss and -10 dB/octave spectral decay due to a propagation loss in soft tissues. The decay agrees well with the spectral characteristics of the measured NAM.

INTRODUCTION

The ordinary voice is air-conducted sound and is easily audible for listeners located near the speaker. The mechanism of production of the ordinary voice is as follows: First, the vocal cords are vibrated by an air stream from the lungs, generating the periodic glottal flow that is the primary source of sound; the vocal tract is resonated due to the vocal cords vibration. Next, the resonated sound radiates from the mouth as an air-conducted sound. Besides this ordinary voice, there are various other types of vocal sounds, *e.g.*, shouting, laughing, crying and whispering. These vocal sounds are more or less audible for people near the speaker. Additionally, there are much weaker murmured vocalizations that have very small energy and are consequently usually unheard even by listeners very nearby. Such a weakly murmured vocalization, however, can be detected by using a special microphone attached to the surface of the skin close behind the ear. In other words, an “air-conducted” weak murmured voice is inaudible; however, a “body-conducted” one can be audible. Such body-conducted weak murmured voice sounds are named non-audible murmur (NAM).

The NAM can be detected by using a special microphone, called a NAM microphone [1-3]. This device is a condenser microphone covered with soft impression material such as soft silicon and urethane elastomer, which provides good impedance matching with the soft tissue of the neck. The NAM, which is inaudible even for people nearby, can be audible when detected by the NAM microphone. This may enable the development of human-to-machine and human-to-human interfaces whose inputs are inaudible voice, thereby providing a “silent” communication tool. The NAM microphone also will be able to revive the speech communication of those with vocal cord problems caused by laryngeal cancer, nerve disorders and muscle diseases.

The NAM has been recorded in recent years. Characterization of the NAM indicates that the signal exhibits severely suppressed higher-frequency components, and does not have sufficient clarity. In order to improve the clarity of NAM detection, the mechanism of NAM production and the sound transfer characteristics of the NAM in soft neck tissues must both be investigated. To reveal transfer and attenuation characteristics of NAM propagation through soft tissues, we investigated sound propagation from the vocal tract to neck surface, using a finite difference time domain (FDTD) method and a head model constructed on the basis of magnetic resonance imaging scans.

HEAD MODEL

Three-dimensional geometrical data of a human head uttering a vowel /e/ NAM were obtained using phonation-synchronized magnetic resonance imaging scans [4]. A two-dimensional image of the median sagittal plane was extracted (Fig. 1a). Subsequently, for simplicity, a homogeneous head model was generated (Fig. 1b), *i.e.*, the head is approximated as being composed only of soft tissue. The vocal tract was replaced by a simplified model with a rectangular cross-section 30 mm wide. This latter approximation was made because a scanned vocal tract is extremely narrow in the vicinity of the vocal cords, and would require too fine of a grid to be simulated by the FDTD method.

(a) (b)

Fig. 1 – (a) MR image at median saggital plane; (b) Two-dimensional head model

NUMERICAL SIMULATION

Cyber Logic Wave 2000 Pro [5] was used as a two-dimensional FDTD solver. The simulated region is a rectangular area, which includes the head model of the following dimensions: 414 mm (height) x 292 mm (width), with a 2-mm grid. The edges of the region are an infinite boundary, *i.e.* absorptive boundary. Physical parameters of soft tissue were set as follows; density ρ, 1,100 kg/m^3; bulk modulus λ_1, 2,600 MPa; bulk viscosity λ_2, 0.001 Pa·s; shear modulus μ_1, 0.025 MPa; shear viscosity μ_2, 1,500 Pa·s, according to the reference [6]. Note that μ_2 was 100 times larger than the value found in the reference, in order to avoid numerical divergence. Also note that the value of λ_2 is determined from those of other soft materials, because the value of λ_2 for soft tissue is not known. Corresponding parameters of air are as follows; ρ, 1.24 kg/m^3; λ_1, 0.147 MPa; λ_2, 0.13 Pa·s; $\mu_1 = \mu_2 = 0$.

Figure 2 shows the geometry of the simulated region including the head, a sound source, and the receivers. Ideally, the sound source would be located near the vocal cords, as in the real human head; unfortunately, due to a limitation of numerical modelling, it cannot be located there. Instead, the sound source is located in front of the mouth, assuming a reciprocal

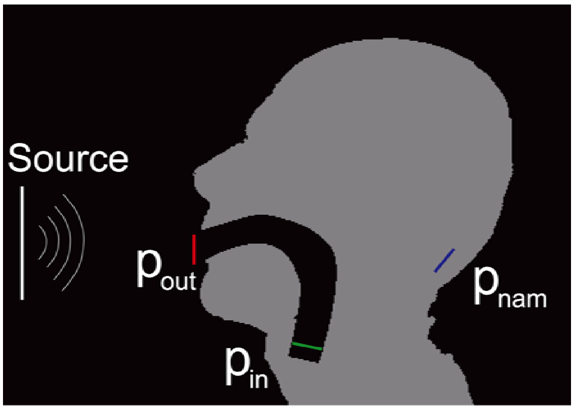

Fig.2 – Simulated region including the head, vibrating plate, and receivers

theorem. The sound source is a pulse-driven vibrating plate. The receivers are located at the inlet and outlet of the vocal tract and at the surface of the neck close behind the ear, *i.e.*, where the NAM microphone would be attached. These receivers are referred to as "**out**", "**in**", and "**nam**", respectively. Note that **nam** is located inside the head in 2-D modelling, although it is located on the neck surface in the real human situation.

RESULTS

Figure 3 shows the sound intensity level observed at each receiver. We approximated the sound intensity level is by dividing the square of sound pressure at each receiver by acoustical impedance at each medium. I_{out}, I_{in}, and I_{nam} denote the sound intensity levels at **out**, **in**, and **nam**, respectively. I_{out} and I_{in} are relatively suppressed at lower frequencies, which reflect the frequency characteristics of the particle velocity pulse used as a sound source signal. I_{nam}, however, is flat. I_{out} is larger at higher frequencies than I_{in} due to characteristics of sound radiation from the mouth. It is also shown that I_{out} and I_{in} have some spectral peaks that are formants. I_{nam} has spectral peaks at 6.3 and 9.7 kHz. These peaks can be attributed to eigenfrequencies of the head [7].

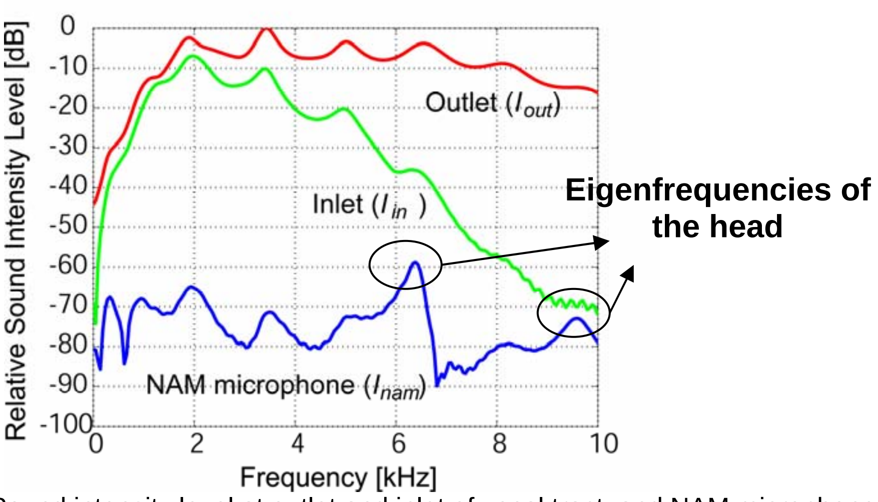

Fig. 3 – Sound intensity level at outlet and inlet of vocal tract, and NAM microphone

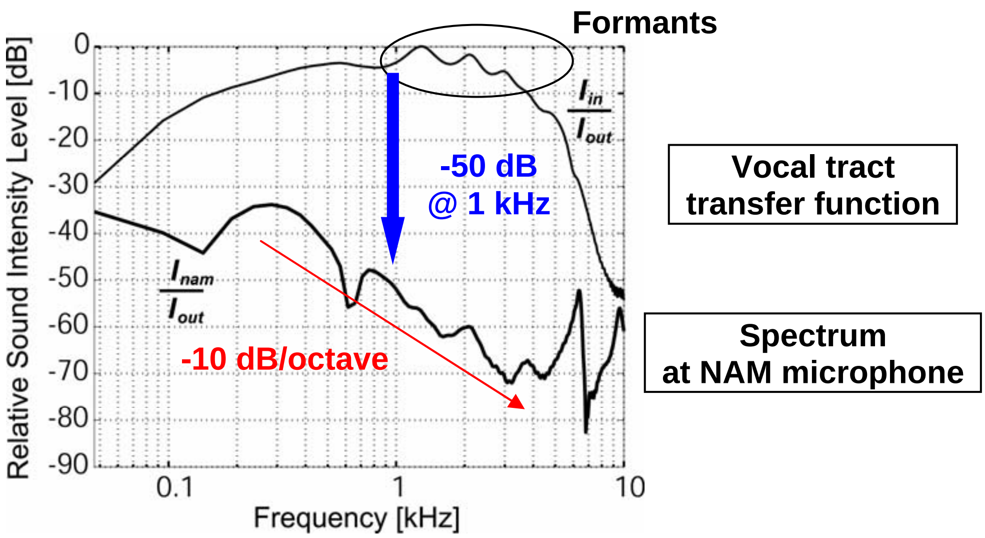

Fig. 4 – Attenuation characteristics at NAM microphone

In order to clarify a transmission loss of sound from air to soft tissue and a propagation loss through soft tissues, I_{in} / I_{out} and I_{nam} / I_{out} were calculated (Fig. 4). I_{in} / I_{out} corresponds to the vocal tract transfer functions, showing spectral peaks, *i.e.* formants, at 0.5, 1.4, 2.1, and 3.0 kHz.

I_{nam} / I_{out} corresponds to the sound intensity level observed at **nam** if a sound radiated from the mouth has flat characteristics. Compared to the vocal tract transfer function, the sound intensity level at **nam** (I_{nam} / I_{out}) is approximately 50 dB lower at 1 kHz. Furthermore, the sound intensity level at **nam** shows -10 dB per octave spectral decay at higher frequencies.

DISCUSSION

Sound energy originating in the vocal tract propagates from the air inside the vocal tract to the soft tissues. Subsequently, it propagates inside the soft tissues and reaches to the NAM microphone, traveling in soft tissues. Within the soft tissues in the current 2D numerical model, the shortest propagation distance is 70 mm. The acoustic impedance of air (Z_{air}) and soft tissues (Z_{tissue}) are approximately 408 and 1,630,000 kg/m^{2}s, respectively. According to these values, transmission loss from air to soft tissues ($4\ Z_{air}\ Z_{tissue} \ /\ (Z_{air} + Z_{tissue})^2$) corresponds to about 30 dB attenuation in sound intensity level at full band. Furthermore, a sound decays during propagation in the soft tissues. It is likely that this propagation loss results in the spectral decay of -10 dB/octave observed in the numerical result. These phenomena can reasonably explain the attenuation characteristics of the NAM detected at the neck surface. The acoustic analysis of the measured NAM shows a spectral decay of -23 dB/octave, which involves glottal sound-source characteristics having a spectral decay of -12 dB/octave [8]. This result indicates that the measured NAM involves -11 dB/octave spectral decay due to propagation loss in the soft-tissue, which is roughly parallel to the spectral decay observed in the current numerical result, *i.e.* -10 dB/octave. However, the frequency dependency of the soft-tissue propagation loss is inconsistent with the spectral decay derived from an absorption coefficient of soft tissues in the ultrasonic range, *i.e.* 0.5–2 dB/cm/MHz [9]. According to this value, there would be virtually no spectral decay due to soft-tissue propagation in the audible range.

CONCLUSIONS

In order to clarify the attenuation characteristics of the NAM detected by a NAM microphone, we numerically simulated a sound that originates in the vocal tract and propagates to the neck surface. Compared to the vocal tract transfer function, the sound detected at the NAM microphone attenuates by 50 dB at 1 kHz and has a spectral decay of -10 dB/octave at higher frequencies. The attenuation of 50 dB at 1 kHz consists of the air-to-soft-tissue transmission loss of 30 dB and the spectral decay of -10 dB/octave due to propagation in the soft tissue, showing good agreement with the acoustic analysis of the measured NAM.

This work is supported by SCOPE-S of the Ministry of Internal Affairs and Communications of Japan.

References

[1] Y. Nakajima *et al.*, “Non-Audible Murmur Recognition Input Interface Using Stethoscopic Microphone Attached to the Skin”, *In Proc. ICASSP*, 708–711, (2003).

[2] Y. Nakajima *et al.*, “Development and Evaluation of Soft Silicone NAM Microphone”, *Tech. Rep. IEICE Speech*, **SP2005-7**, 7–12, (2005).

[3] Y. Nakajima *et al.*, “Non-Audible Murmur (NAM) Recognition”, *IEICE Trans. on information and systems*, **Vol.E89-D**, **No.1**, 1–8, (2006).

[4] Y. Nota, T. Kitamura, K. Honda, H. Takemoto, H. Hirata, Y. Shimada, I. Fujimoto, Y. Shakudo, and S. Masaki, “A bone-conduction system for auditory stimulation in MRI", *Acoust. Sci. & Tech.*, **28**, 33–38, (2007).

[5] http://www.cyberlogic.org/

[6] H.L. Oestreicher, “Field and Impedance of an Oscillating Sphere in a Viscoelastic Medium with an Application to Biophysics”, J. Acoust. Soc. Am., **23**(6), 707–714, (1951).

[7] Y. Fujisaka, S. Nakagawa, T. Ogita, and M. Tonoike, “Analysis of wave propagation for bone-conducted ultrasonic in the heterogeneous human head model”, *Tech. Rep. IEICE*, **103**(608), (2004).

[8] T. Hirahara, S. Shimizu, and M. Otani, “Acoustic characteristics of non-audible murmur”, *In Proc. Japan-China Joint Conference on Acoustics 2007*, (2007.6).

[9] D.E. Goldman and T.F. Hueter, “Tabular data of the velocity and absorption of high-frequency sound in mammalian tissues”, *J. Acoust. Soc. Am.*, **28**(1), 35–37, (1956).