Acoustics'08 Paris
June 29-July 4, 2008
www.acoustics08-paris.org

A numerical analysis of fluctuations in pressure wave within the larynx using two-dimensional asymmetrical vocal folds model

Hideyuki Nomura[a], Tomoo Kamakura[a] and Tetsuo Funada[b]

[a]Dept. of Electronic Eng., Univ. of Electro-Communications, 1-5-1, Chofugaoka, 182-8585 Chofu-shi, Japan
[b]Div. of Electronic Eng. and Computer Sci., Kanazawa Univ., Kakuma-machi, 920-1192 Kanazawa-shi, Japan
nomu@ee.uec.ac.jp

In this study, we numerically simulate the speech production on the basis of asymmetrical vocal fold (VF) model as a pathological VF model, and consider asymmetrical VF vibrations and also fluctuations in pressure wave within the larynx. The simulation show that the left and right VFs asymmetrically vibrate with a phase difference in the asymmetrical model. We estimate the fundamental frequency, amplitude and waveform fluctuations in pressure wave within the larynx. The fluctuations obtained by numerical experiments in the symmetrical VF model are in rough agreement with those obtained by real speech signal. However, with increasing asymmetry, the fluctuations of numerical experiment are out of the range of the real speech data. This result suggests that the degree of VF asymmetry can be detected by estimating fluctuations in speech wave.

1 Introduction

Voice disorders render a comfortable social life very difficult, since a voice is one of the most important tools for human communication. A better understanding of the dynamics of the speech production may contribute to the improvement of quality of life.

An analysis of pathological voice production contributes to the development of medical examination technique against voice disorders and rehabilitation of the recovery of voice. Tanabe *et al.*[1], Isshiki *et al.* [2] and Aomatsu *et al.*[3] studied pathological vocal fold (VF) vibrations by means of the observations of canine VFs and numerical simulations of asymmetrical VF models. Left and right VFs synchronously and symmetrically vibrate for non-pathological VFs. The vibration of pathological VFs, such as the VFs with a polyp, which have different properties in the left and right VFs, would indicate different behavior from that of normal VFs.

In this study, we numerically simulate the speech production on the basis of a pathological VF model as an application of our proposed glottal source model [4], and consider asymmetrical VF vibrations and also fluctuations in pressure wave within the larynx. First, pathological VFs are model as asymmetrical VFs with asymmetries of geometrical or mechanical parameters. Then, the speech production on the asymmetrical VF model is numerically simulated, and fluctuations in pressure wave are estimated. Finally, we compare the fluctuations obtained from the numerical experiment and some measured data for real speech.

2 Speech production model

2.1 Larynx and vocal tract models

A two-dimensional larynx model in the coronal (z-x) plane is shown in Fig. 1(a). We assume that a configuration is uniform in the sagittal (y) direction in order to simply a problem. The vocal tract attached to the larynx is approximated by a uniform rigid duct. The detailed description of the shapes has been presented in previous papers [5, 6].

2.2 Mechanical VF model

The present mechanical model of vibrating VF is based on the distributed parameter model proposed by Ikeda *et al.* [7]. The VF can be divided into two tissue layers with different mechanical properties: a cover layer and a body layer. The cover layer is assumed to be an elastic cover with the effective mass of the VF. In order to take the mechanical properties of the VF into account, the elastic cover is supported by an array of small mechanical elements having nonlinear spring and damper [4].

Figure 1(b) shows a proposed VF model. In order to simplify the analysis, we restrict VF vibrations to the lateral direction x. The vibration of the VF is governed by the equation of motion consisting of the fluid force, i.e., the pressure and the viscous stress of the flow, the restoring force of the spring, the viscous drag force of the damper, and the shear force of the elastic cover [4]. The geometrical property of the VF is characterized by the thickness of the VF (T_{VF}), the effective depth of the VF vibrational area (D_{VF}) and the half gap distance from the midline of symmetrical larynx ($G_{\mathrm{VF}}/2$). The mechanical property of the VF is characterized by the Young's modulus of spring element and the elastic cover (E_{VF}), the volume density of the elastic cover (ρ_{VF}) and the viscosity of the damper element (ν_{VF}).

2.3 Glottal flow model

The glottal flow is assumed to be an unsteady two-dimensional compressible viscous fluid. The fluid motion is analyzed on the basis of boundary fitted coordinates along the surfaces of larynx. In order to consider a moving boundary problem arising as a result of the shape change of the VF, we employ a finite difference method based on the ALE (arbitrary Lagrangian-Eulerian) method [8].

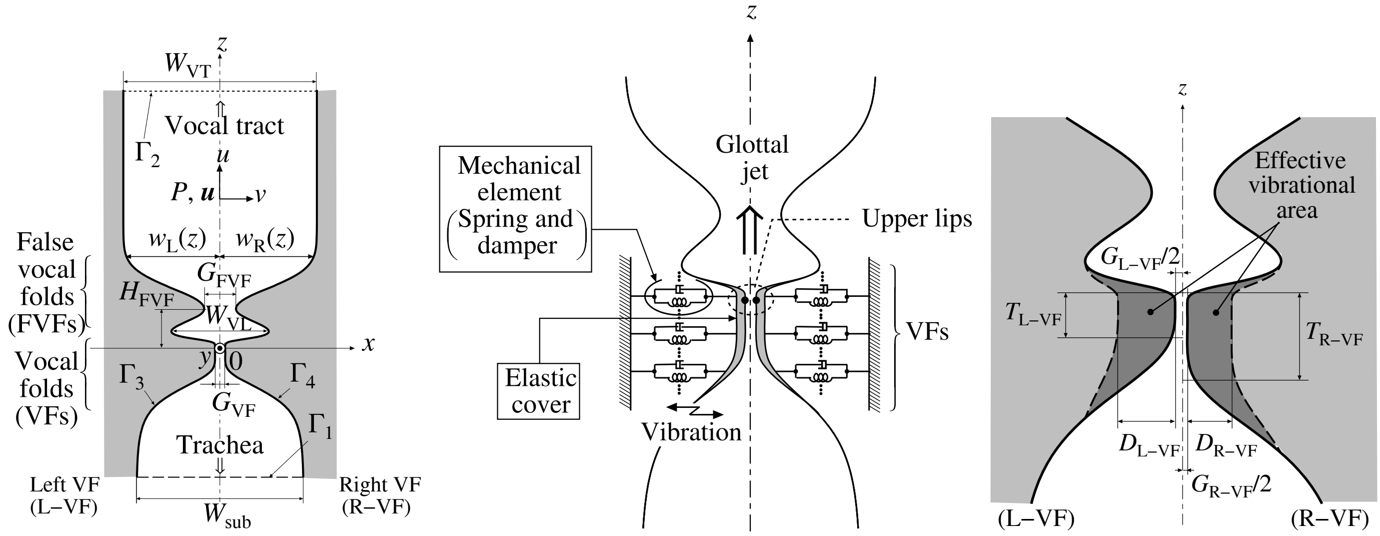
2.4 Asymmetrical VF model

We consider effects of geometrical and mechanical asymmetries of VFs on the speech production. Figure 1(c) shows schematic presentation of (geometrical) asymmetrical VF model. The asymmetries of the left VF (L-VF) and right VF (R-VF) are controlled by the following asymmetrical parameters:

$$\left.\begin{aligned} \alpha_D &= D_{\mathrm{L-VF}}/D_{\mathrm{R-VF}}, \\ \alpha_G &= G_{\mathrm{L-VF}}/G_{\mathrm{R-VF}}, \\ \alpha_T &= T_{\mathrm{L-VF}}/T_{\mathrm{R-VF}}, \end{aligned}\right\} \tag{1}$$

$$\left.\begin{aligned} \alpha_E &= E_{\mathrm{L-VF}}/E_{\mathrm{R-VF}}, \\ \alpha_\rho &= \rho_{\mathrm{L-VF}}/\rho_{\mathrm{R-VF}}, \\ \alpha_\nu &= \nu_{\mathrm{L-VF}}/\nu_{\mathrm{R-VF}}. \end{aligned}\right\} \tag{2}$$

α_D, α_G and α_T are geometrical asymmetrical parameters, and α_E, α_ρ and α_ν are mechanical those. The subscripts L-VF and R-VF denote the parameters of the left and right VFs, respectively. In the present study, keeping the parameters of the R-VF at the normal values showed, the parameters of the L-VF are varied.



(a) Two-dimensional larynx and vocal tract models in the coronal (z-x) plane

(b) Mechanical vocal fold model

(c) Asymmetric vocal folds

Figure 1: Analytical models of speech production.

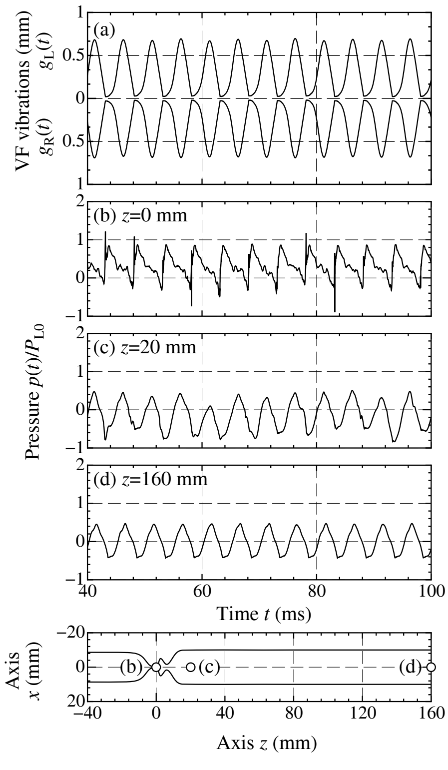


Figure 2: Example of sound generation based on symmetrical VF model at $P_{L0} = 800$ Pa. VF vibrations $g_L(t)$ and $g_R(t)$ denote the distance of the left and right VF surfaces from the midline of the larynx (the z axis), respectively, at $z \simeq 0$ mm in (a).

3 Results and discussion

3.1 Simulation of speech production

Figure 2 shows an example of simulation result in the symmetrical VF model. In the present simulations, lung pressure P_{L0} is set to a constant value of 800 Pa, which corresponds to the value for an ordinary conversation level. VF vibrations $g_L(t)$ and $g_R(t)$ denote the distance from the midline of the larynx (the z axis) to the surfaces of the L-VF and R-VF, respectively. The vibrations are measured at the location of upper lips at $z \simeq 0$ mm. Normalized pressure waves $p(t)/P_{L0}$ are measured at different distances ($z = 0$, 20 and 160 mm) from the glottis, where t denotes the time from the beginning of application of lung pressure. A symmetrical vibration between the L-VF and R-VF is observed. As an interesting phenomenon, the pressure wave indicates fluctuations, i.e., the fundamental frequency f_0, the amplitude (peak-to-peak) P_{p-p} and the waveform, caused by unsteady vortex motions within the larynx [6].

3.2 Effects of VF asymmetry on fluctuation

Pressure waveforms in the larynx indicated fundamental frequency, amplitude and waveform fluctuations. In order to quantitatively evaluate the relationship between the fluctuations and VF asymmetries, the following three fluctuation measures are estimated [9, 10]:

- The coefficient of variation of fundamental frequency (CV of f_0): the measure of fundamental frequency fluctuation,
- The CV of amplitude (CV of P_{p-p}): the measure of amplitude fluctuation,
- The harmonic-to-noise ratio (HNR) [11]: the measure of waveform fluctuation.

The CV is the ratio of the standard deviation, σ, of sequence to the average value, m, of sequence, that is, CV= σ/m. f_0 and P_{p-p} sequences are extracted at each pitch period from a pressure wave as follows (see Fig. 3).

1. Extract the fundamental wave from the pressure wave by fundamental wave filtering [12].
2. Segment by detecting the zero crossing of the fundamental wave.

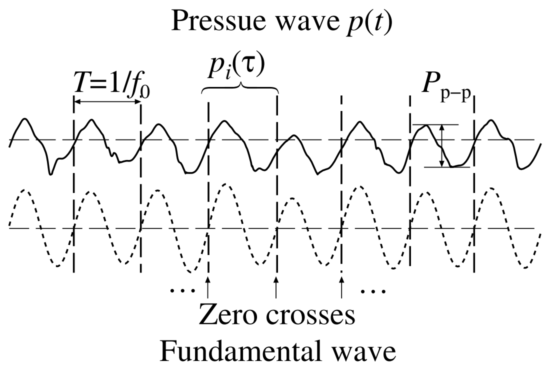


Figure 3: Segmentation into pitch periods by detecting zero crossing of fundamental wave [12].

3. Extract f_0 (reciprocal number of segment duration T) and $P_{\mathrm{p-p}}$ from individual segments.

The HNR is a measure of the cycle-to-cycle similarity of a waveform, and is sensitive to waveform aperiodicity [11]. The HNR is defined as the ratio between the energy of the periodic (harmonic) component H to the energy of the aperiodic (noise) component N in a wave. The energy of the periodic component H is consistent with the energy of the average wave that is determined by taking the average of a succession of period sequence. The energy of the aperiodic component N is the mean energy of the difference between individual periods and the average wave [9, 10, 11].
A original pressure wave $p(t)$ can be considered as the concatenation of the waves $p_i(\tau)$ from each pitch period, where $i = 1, 2, \cdots, M$, and M is the number of samples (see Fig. 3). Examples of extracted $p_i(\tau)$ are shown in Fig. 4(a) by thin gray lines.
The average wave $p_{\mathrm{A}}(\tau)$ is defined as

$$p_{\mathrm{A}}(\tau) = \frac{1}{M}\sum_{i=1}^{M} p_i(\tau), \tag{3}$$

and is shown in Fig. 4(a) by the thick black line as an example. For the calculating $p_{\mathrm{A}}(\tau)$, we assume that $p_i(\tau)$ is equal to zero in the interval between T_i and $T_{\max}$, where T_i and $T_{\max}$ are the duration of the i-th period and the maximum period, respectively. The energy of the periodic component H is defined as

$$H = M\int_0^{T_{\max}} \{p_{\mathrm{A}}(\tau)\}^2 \,\mathrm{d}\tau. \tag{4}$$

The noise wave in the i-th pitch period is equal to $p_i(\tau) - p_{\mathrm{A}}(\tau)$ (see Fig. 4(b)). The energy of the aperiodic component N is defined as

$$N = \sum_{i=1}^{M}\int_0^{T_i} \{p_i(\tau) - p_{\mathrm{A}}(\tau)\}^2 \,\mathrm{d}\tau. \tag{5}$$

Then, HNR$= 10\log(H/N)$ (dB) [10].
Figures 5 and 6 show the relationships between the geometrical and mechanical asymmetries and the fluctuation measures in pressure waves at different distances from the glottis, respectively. The asymmetrical parameters except for that indicated in the lateral axis are kept constant at unity. Closed circles, open circles and squares denote measured values at $z = 0$, 20 and 160 mm, respectively. Solid, dotted and dashed curves denote the fitted curves to the measured data obtained using a squared function of asymmetrical parameter.
The variations of the CVs of f_0 and $P_{\mathrm{p-p}}$ with the depth and density asymmetries indicate concave profiles and that of HNR indicates a convex profile. For most cases, the fluctuations at $z = 160$ mm are smallest. This trend clearly appears in HNR.
For reference, the gray zones denote distributions of fluctuation measures of real speech data for normal utterances of Japanese five vowels. The speech data of three adult males were recorded at the mouth in an anechoic room. We estimated the fluctuation measures for the data with an almost constant amplitude in duration of 1 or 2 s. The CV of f_0 and HNR at $z = 160$ mm for the symmetrical condition ($\alpha_T = \alpha_D = \alpha_G = 1$) obtained from the numerical experiment are in the range of the distribution of the real data.

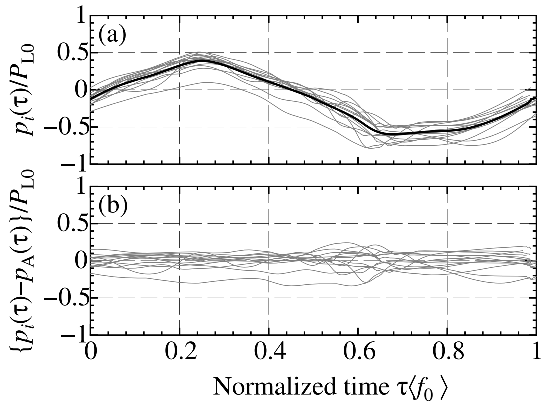
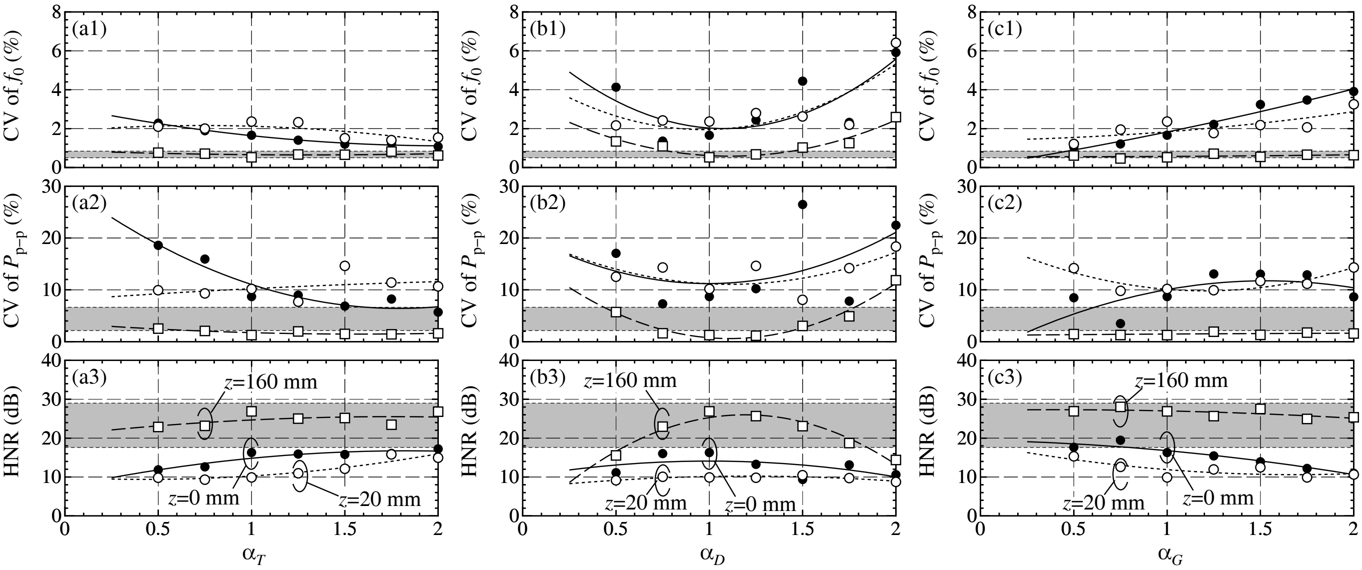


Figure 4: Periodic (harmonic) and aperiodic (noise) pressure waves. The pressure wave $p_i(\tau)$ in each period (thin gray line in (a)) is extracted from the original pressure wave $p(t)$ by segmentation using the zero crossing of the fundamental wave [12]. The average wave (thick black line in (a)) corresponds to the harmonic wave. The noise wave $p_i(\tau) - p_{\mathrm{A}}(\tau)$ in each period is indicated by thin gray line in (b).

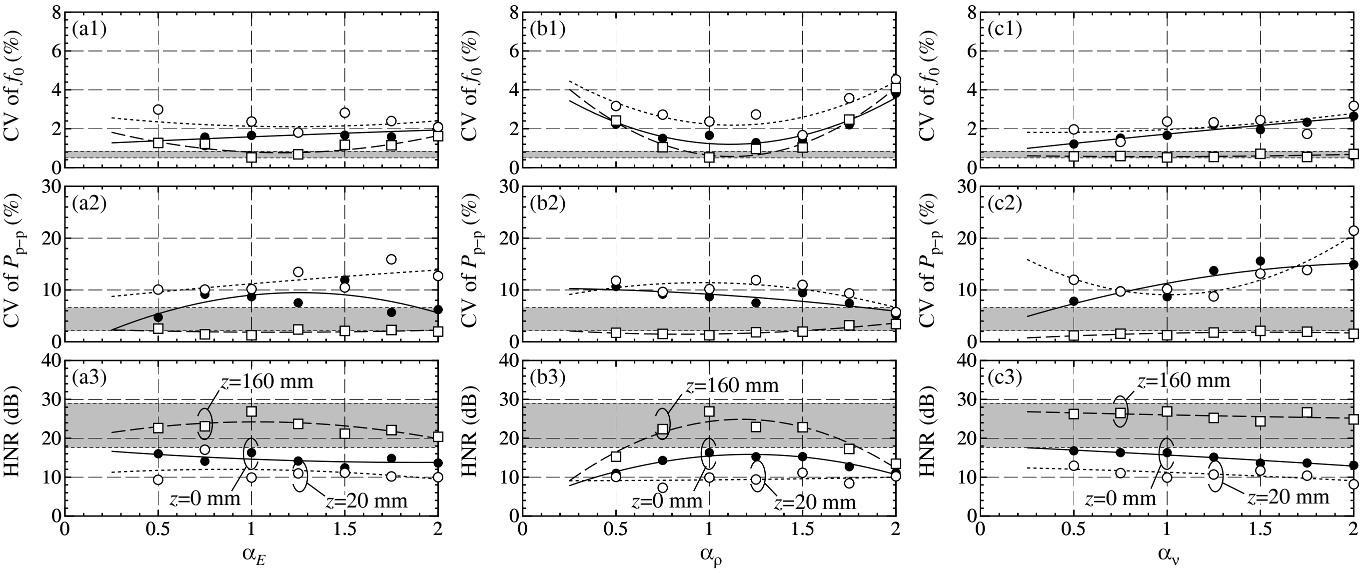
3.3 Discussion

In this paper, we estimated three fluctuation measures, the CVs of f_0 and $P_{\mathrm{p-p}}$ and HNR. The fluctuation of magnitude of the CV of f_0 and the profile of the CV of $P_{\mathrm{p-p}}$ showed no clear trends with measured location. On the other hand, trends in the magnitude and the profile of the HNR with measured location could be observed. Therefore, the HNR in speech wave is usefully for the estimation of fluctuation of speech.
The fluctuations obtained by numerical experiments in the symmetrical model are in rough agreement with that obtained by real speech data. However, with increasing the asymmetry, for example, $\alpha_\rho < 0.5$ or $\alpha_\rho > 1.75$ in HNR, the results of numerical experiment are out of the range of real speech data. This suggests that the degree of VF asymmetry can be detected by estimating fluctuations in speech wave. In reality, the estimation



(a) Thickness asymmetry (b) Depth asymmetry (c) Gap asymmetry

Figure 5: Effects of geometrical asymmetries of VFs on pressure wave fluctuations . Closed circles, open circles and squares denote the measured data, and the solid, dotted and dashed curves denote the fitting curves to the measured data obtained using a square function of the asymmetrical parameter. Gray zones denote distributions of fluctuations of real speech data for normal utterances of Japanese five vowels measured at the mouth.



(a) Young's modulus asymmetry (b) Density asymmetry (c) Viscosity asymmetry

Figure 6: Effects of mechanical asymmetries of VFs on pressure wave fluctuations. The symbol marks and curves are the same as in Fig. 5.

of HNR in speech wave was examined as an index of the degree of hoarseness [11].
The causes of fluctuations in speech wave is not only mechanical and geometrical asymmetries of VF, but also neurological and aerodynamic effects [13]. In addition, in mechanical and geometrical asymmetries, fluctuated wave is caused by not only one type of asymmetry, but also an interaction of multi type of asymmetries. More detailed work of estimations of fluctuation for real speech data and numerical simulation is necessary in order to apply estimations of fluctuation in speech wave for the support technique for a diagnose of pathological voice.

4 Conclusion

In this paper, we numerically simulated the speech production on the basis of a pathological VF model, and considered asymmetrical VF vibrations and also fluctuations in pressure wave within the larynx. Pathological VFs were model as asymmetrical VFs with asymmetries of geometrical or mechanical parameters. Speech production was numerically simulated by alternately solving the glottal flow and VF vibration.
We estimated the fundamental frequency, amplitude and waveform fluctuations in pressure wave within the larynx, and compared those with the fluctuations obtained from real speech signals. The fluctuations obtained numerical experiments in the symmetrical VF model are in rough agreement with those obtained by real speech signal. However, with increasing asymmetry, the fluctuations of numerical experiment were out of the range of the real speech data. This result suggests that the degree of VF asymmetry can be detected by estimating fluctuations in speech wave.
The causes of fluctuations in speech wave is not only mechanical and geometrical asymmetries of VF, but also neurological and aerodynamic effects. In addition, the fluctuations are caused by not only one type of asymmetry, but also an interaction of multi type of asymmetries. More detailed work of estimations of fluctuation for real speech data and numerical simulation is necessary in order to apply estimations of fluctuation in speech wave for the support technique for a diagnose of pathological voice.

5 Acknowledgments

The present research was supported in part by a Grant-in-Aid for Young Scientists (B) (No. 18700171) from the Japan Society for the Promotion of Science.

References

[1] M. Tanabe, N. Isshiki and K. Kitajima, "Vibratory pattern of the vocal cord unilateral paralysis of the cricothyroid muscle," *Acta Otolaryng.*, *74*, 339–345 (1972).

[2] N. Isshiki, M. Tanabe, K. Ishizaka and D. Broad, "Clinical significance of asymmetrical vocal cord tension," *Ann. Otol. Rhinol. Laryngol.*, *86*, 58–66 (1977).

[3] T. Aomatsu, Y. Matsuzaki and T. Ikeda, "Numerical analysis of pathological voice production using asymmetrical distributed-parameter model for vocal fold: 2nd Report, Numerical simulation and analysis of asymmetrical vocal fold vibration," *Trans. Jpn. Soc. Mech. Eng. B*, *68*, 799–805 (2002) (in Japanese).

[4] H. Nomura and T. Funada, "Sound generation by unsteady flow ejecting from the vibrating glottis based on a distributed parameter model of the vocal cords," *Acoust. Sci. & Tech.*, *28*, 392–402 (2007).

[5] H. Nomura and T. Funada, "Numerical simulation of unsteady flow through the rigid glottis," *Acoust. Sci. & Tech.*, *27*, 154–162 (2006).

[6] H. Nomura and T. Funada, "Effects of the false vocal folds on sound generation by an unsteady glottal jet thorough rigid wall model of the larynx," *Acoust. Sci. & Tech.*, *28*, 403–412 (2007).

[7] T. Ikeda, Y. Matsuzaki and T. Aomatsu, "A numerical analysis of phonation using a two-dimensional flexible channel model of the vocal cords," *J. Biomech. Eng.*, *123*, 571–579 (2001).

[8] C.W. Hirt, A.A. Amsden and J.L. Cook, "An arbitrary Lagrangian-Eulerian computing method for all flow speeds," *J. Comput. Phys.*, *14*, 227–253 (1974).

[9] H. Nomura and T. Funada, "Fluctuations in pressure wave within the larynx," *Acoust. Sci. & Tech.*, (Accepted).

[10] H. Nomura and T. Funada, "Effect of unsteady glottal flow on the speech production process," *Proc. Interspeech 2007*, Antwerp, Belgium, pp. 1390–1393 (2007).

[11] E. Yumoto and W.J. Gould, "Harmonic-to-noise ratio as an index of the degree of hoarseness," *J. Acoust. Soc. Am.*, *71*, 1544–1550 (1982).

[12] H. Ohmura and K. Tanaka, "Fine pitch contour extraction by voice fundamental wave filtering method," *J. Acoust. Soc. Jpn.* (*J*), *51*, 509–518 (1995) (in Japanese).

[13] I.R. Titze, *Principles of Voice Production*, 2nd Ed. (NCVS, IA, 2000), Chap. 11, pp. 316–321.