Chaotic component obscured by strong periodicity in voice production system

Chao Tao and Jack J. Jiang[*]

*Department of Surgery, Division of Otolaryngology Head and Neck Surgery, University of Wisconsin Medical School, Madison,
Wisconsin 53792-7375, USA*

The effect of glottal aerodynamics in producing the nonlinear characteristics of voice is investigated by comparing the outputs of the asymmetric composite model and the two-mass model. The two-mass model assumes the glottal airflow to be laminar, nonviscous, and incompressible. In this model, when the asymmetric factor is decreased from 0.65 to 0.35, only 1:1 and 1:2 modes are detectable. However, with the same parameters, four vibratory modes (1:1, 1:2, 2:4, 2:6) are found in the asymmetric composite model using the Navier-Stokes equations to describe the complex aerodynamics in the glottis. Moreover, the amplitude of the waveform is modulated by a small-amplitude noiselike series. The nonlinear detection method reveals that this noiselike modulation is not random, but rather it is deterministic chaos. This result agrees with the phenomenon often seen in voice, in which the voice signal is strongly periodic but modulated by a small-amplitude chaotic component. The only difference between the two-mass model and the composite model is in their descriptions of glottal airflow. Therefore, the complex aerodynamic characteristics of glottal airflow could be important in generating the nonlinear dynamic behavior of voice production, including bifurcation and a small-amplitude chaotic component obscured by strong periodicity.

PACS number(s): 43.70.+i, 05.45.−a, 87.19.−j

I. INTRODUCTION

The bifurcation and chaos phenomena related with nonlinear systems are widely observed in physical and biological systems, such as electronic circuits [1], superconductors [2], neurology [3], and cardiology [4]. Nonlinear dynamic theory provides deeper insight into these complex natural phenomena. The voice of a vertebrate is generated by a highly nonlinear oscillator, comprised of the vocal folds and glottal airflow. Rich nonlinear phenomena, including subharmonics, diplophonic, double period doubling bifurcation, chaos, and spatiotemporal chaos, have been observed in birdsong [5–7], human voice [8], the cries of newborn infants [9], excised larynx experiments [10], and the calls of mammals [11,12]. Knowledge of the nonlinear components of voice is expected to be useful for understanding voice production [13], improving the realistic synthesis of voice [14], and providing potential clinical applications in assessment and treatment of voice diseases [15]. Therefore, the nonlinear dynamics in voice production have recently received considerable attention.

Several models have been used to investigate the generation of nonlinear dynamics in voice [13,14,16–21]. Irregular vibration patterns were initially detected in a two-mass model with a sufficiently large left-right tension imbalance [13]. Bifurcations and chaos were then predicted by a symmetric vocal fold model with tissue parameters obviously deviating from the normal values [16]. Irregular vocal fold oscillations were also found in the continuous model with soft tissue [20] and pathological phonation [21]. These studies suggest that abnormalities in vocal tissue are important causes of irregular vocal fold vibration. The introduction of random perturbations into a vocal fold model has been found

to decrease the threshold value necessary for a system to produce irregular vocal fold vibration [17]. Recently, by considering the source-filter acoustic interaction, delayed feedback due to reflected sound in the vocal tract was shown to contribute to extremely rich dynamics, such as period doubling bifurcation and nonperiodic oscillation [18]. These studies explain the cause of rough voice production and irregular vibrations of vocal folds that are typical of many vocal diseases.

Glottal airflow plays an important role in voice production because vocal fold vibrations are generated by the fluid-tissue interaction. It is known that real fluid is very complex, and various dynamical behaviors can be found in real fluid, including asymmetry, vortices, turbulence [22,23], bifurcation [24], and spatiotemporal chaos [25,26]. These phenomena are also components of glottal airflow [27–35]. These complexities of glottal airflow could be responsible for variation in the phonation threshold pressure [36] and for the asymmetric vibration of the vocal fold [37]. However, this causal linkage was not considered in most of the previous studies of irregular vocal fold vibration because the airflow was often simplified in accordance with Bernoulli's law and the laminar, nonviscous, and incompressible assumption. Few works have studied the influence of the complex spatiotemporal behaviors of glottal airflow on the vocal fold vibration, particularly irregular vibrations.

In this study, we simulated and analyzed the contribution of the complexity of glottal airflow on irregular vocal fold vibration. The Navier-Stokes (NS) equations were used to predict the spatiotemporal behavior of glottal aerodynamics, and the tissue mechanics of the vocal folds were simplified to an asymmetric two-mass model. The interaction of the vocal fold and the glottal airflow generates self-oscillation. Using this model, we studied the various oscillatory patterns of vocal folds. To investigate the effect of airflow on vocal fold vibration, the dynamic behaviors of the vocal folds when driven by an airflow description based on the NS equa-

*jjjiang@wisc.edu

tions and when driven by an airflow description based on Bernoulli's law were compared. The nonlinear detection method proposed by Barahona and Poon [4,38,39], which has the ability to detect the nonlinearity in a short, strong periodic series, was used to analyze the nonlinear characteristics of vocal fold vibration.

II. MODEL AND ANALYSIS METHOD

A. Asymmetric composite vocal-fold model

An asymmetric composite model, developed from the composite model [37], was employed to predict the vocal fold vibrations induced by complex airflow. The asymmetric two-mass model [13,40] was used to describe the tissue mechanics of vocal folds with unilateral laryngeal paralysis,

$$\frac{d\mathbf{x}_\alpha}{dt} = \mathbf{v}(\mathbf{x}_\alpha, \boldsymbol{\mu}_{t\alpha}) + \mathbf{F}_\alpha. \tag{1}$$

The displacement vector $\mathbf{x}_\alpha$ is written as $\mathbf{x}_\alpha = [x_{1\alpha}, \ v_{1\alpha} \ x_{2\alpha} \ v_{2\alpha}]^T$. The velocity vectors $\mathbf{v}$, relating with $\mathbf{x}_\alpha$ and the tissue parameter vector $\boldsymbol{\mu}_{t\alpha}$, can be described as

$$\mathbf{v}(\mathbf{x}_\alpha, \boldsymbol{\mu}_{t\alpha}) = \begin{pmatrix} v_{1\alpha} \\ \dfrac{1}{m_{1\alpha}}\left(-r_{1\alpha}v_{1\alpha} - k_{1\alpha}x_{1\alpha} - \Theta(-a_{1\alpha})c_{1\alpha}\dfrac{a_{1\alpha}}{2L} - k_{c\alpha}(x_{1\alpha} - x_{2\alpha})\right) \\ v_{2\alpha} \\ \dfrac{1}{m_{2\alpha}}\left(-r_{2\alpha}v_{2\alpha} - k_{2\alpha}x_{2\alpha} - \Theta(-a_{2\alpha})c_{2\alpha}\dfrac{a_{2\alpha}}{2L} - k_{c\alpha}(x_{2\alpha} - x_{1\alpha})\right) \end{pmatrix}, \tag{2}$$

where $x_{i\alpha}$ is the displacement of mass $m_{i\alpha}$, $r_{i\alpha}$ is the damping constant, $k_{i\alpha}$ is the spring constant, L is the length of the glottis, $c_{i\alpha}=3k_{i\alpha}$ is the additional spring constant during collision, and $k_{c\alpha}$ is the coupling constant between the upper and lower masses. The index i denotes the inferior ($i=1$) and superior masses ($i=2$), while the index α represents right ($\alpha=r$) and left ($\alpha=l$) masses. The term $a_{1\alpha}=a_{01}+(x_{1r}+x_{1l})L$ is the lower glottal area and $a_{2\alpha}=a_{02}+(x_{2r}+x_{2l})L$ is the upper glottal area, where a_{0i} refers to the prephonatory glottal areas. The term $\Theta(-a_{i\alpha})c_{i\alpha}a_{i\alpha}/2L$ represents collision with

$$\Theta(x) = \begin{cases} \tanh(50x/x_0), & x > 0 \\ 0, & x \le 0 \end{cases}. \tag{3}$$

Unilateral laryngeal paralysis is modeled as follows:

$$m_{il}=m_{ir}/Q, \quad k_{il}=Qk_{ir}, \quad k_{cl}=Qk_{cr}, \quad c_{il}=Qc_{ir}, \tag{4}$$

where Q is the asymmetric factor. The other parameters of the left vocal fold are identical with those of the right vocal fold. Table I presents the lumped-mass parameters used in the simulation.

The intraglottal airflow force vector $\mathbf{F}_\alpha$ applied to the vocal fold surface provides the driving force for the vocal fold vibration and can be defined as

$$\mathbf{F}_\alpha = \left[0, (L/m_{1\alpha})\int_0^{d_{1\alpha}} p_\alpha(y)dy, 0, 0\right]^T, \tag{5}$$

where $d_{1\alpha}$ is the thickness of mass $m_{1\alpha}$ and $p_\alpha(y)$ refers to the pressure on the vocal glottis wall. The NS equations are used to describe the spatiotemporal pattern of glottal airflow, that is,

$$\nabla \cdot \mathbf{V} = \mathbf{0}, \tag{6}$$

$$\rho\frac{\partial \mathbf{V}}{\partial t} + \rho\mathbf{V} \cdot \nabla\mathbf{V} = \mu\nabla^2\mathbf{V} - \nabla p, \tag{7}$$

where $\rho=1.2$ kg/m^3 is the flow density, $\mu=1.8 \times 10^{-5}$ N s/m^2 is the viscosity, $\mathbf{V}$ is the velocity vector, and p is the pressure of flow. The glottal length is 0.3 cm. The prephonatory superior and inferior glottal widths are both 0.178 57 mm. A 2-cm inlet duct and an 8-cm outlet duct represent the trachea and the vocal tract, respectively. Their widths are both 2 cm. Figure 1 illustrates the structure of the airflow channel around the glottis. The no-slip condition was applied to both the tracheal wall and the vocal tract wall. The specified pressure condition was set at the inlet of the subglottal tube to drive the model, and the zero pressure boundary condition was set at the outlet of the supraglottal tube. The moving wall condition was applied to the fluid-solid interaction boundary. Using the given boundary condition,

TABLE I. Parameter values.

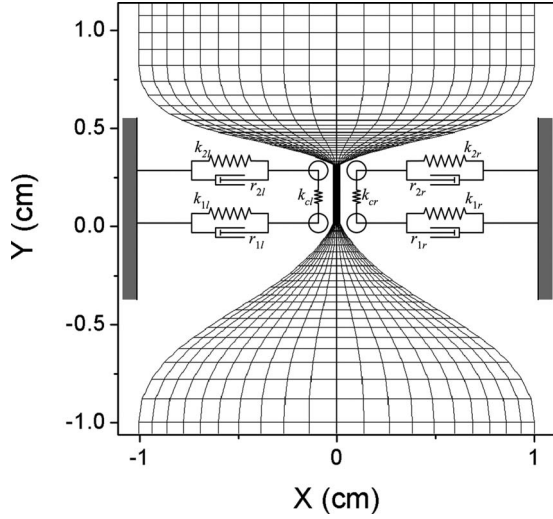| Parameter		Value
$m_{1\alpha}$	Lower mass	0.125 g
$m_{2\alpha}$	Upper mass	0.025 g
$k_{1\alpha}$	Lower spring constant	80 kdyn/cm
$k_{2\alpha}$	Upper spring constant	8 kdyn/cm
$k_{c\alpha}$	Coupling spring constant	25 kdyn/cm
$r_{1\alpha}$	Damping constant	20 g/s
$r_{2\alpha}$	Damping constant	20 g/s

FIG. 1. The system diagram of the composite vocal fold model.

the flow pressure p can be solved using the NS equations.

In our simulation, the NS equations and the two-mass vibration model were fully coupled and simultaneously solved in every iteration step. The values of the glottal displacements and the velocities solved from the two-mass model were applied to the glottal wall to define the boundary condition. Simultaneously, the airflow pressure solved from NS equations was applied as the driving force for the two-mass model. By alternating between solving the NS equations and solving the asymmetric two-mass equation, a self-oscillating solution was obtained from the iterative process. The NS equations were solved by using the FLOTRAN CFD analysis of the finite-element software ANSYS. A custom program was developed in ANSYS Parametric Design Language to solve the two-mass equation with the fourth-order Runge-Kutta routine and manage the airflow-tissue interaction. In the time domain, both the NS equations and the two-mass equation were integrated with a time step of 50 μs. Before the above model was used in this study, we halved the grid spacing, reduced the time step and the convergence criteria by a factor of 10, and reran the simulation. The solution is essentially the same as before. The tests verified that this model is independent of the grid size, time step, and convergence criteria [37,41,42]. The aerodynamic outputs of this model were also compared with previous static model simulations and experimental measurements [27–30] in a symmetric situation ($Q=1$). It was found that glottal aerodynamics, such as the pressure distribution, Coanda effect, and others, can be correctly predicted by this model. More details can be found in our previous paper [37].

B. Nonlinear dynamics detection method

It has been suggested that the strength of nonlinearity in the data can be measured by comparing nonlinear and linear predictability [43], and this concept has been applied to quantify the amount of nonlinearity in the vocalizations of macaque screams, piglet screams, and dog barks [44]. Based on a similar concept, the nonlinear dynamics detection

method proposed by Barahona and Poon [38] was used to analyze the nonlinearity of the model output in this study. This nonlinearity detection method is superior when applied to short time series, even when the series is heavily contaminated with noise or in the presence of strong periodicity. As a result, it has been successfully employed to detect chaos in the beat-to-beat interval series of electrocardiograms [4] and in human speech [45]. This method is based on the comparison of the one-step-ahead prediction error of the linear and the nonlinear model. Its procedure is as follows:

For time series x_n ($n=1,\ldots,N$), a discrete Volterra-Wiener-Korenberg series of degree d and memory κ is used as a model to calculate the predicted time series x_n',

$$
\begin{aligned}
x_n' = a_0 + a_1 x_{n-1} + a_2 x_{n-2} + \cdots + a_\kappa x_{n-\kappa} + a_{\kappa+1} x_{n-1}^2 \\
+ a_{\kappa+2} x_{n-1} x_{n-2} + \cdots + a_{M-1} x_{n-\kappa}^d.
\end{aligned}
\tag{8}
$$

The coefficient a_m is estimated through a Gram-Schmidt procedure, and the standard deviations of the one-step-ahead prediction error are calculated,

$$
\varepsilon(\kappa,d)^2 \equiv \frac{\sum_{n=1}^{N} [x_n'(\kappa,d) - x_n]^2}{\sum_{n=1}^{N} (x_n - \bar{x})^2} \quad \text{with } \bar{x} = \frac{1}{N}\sum_{n=1}^{N} x_n.
\tag{9}
$$

For each data series, the best linear model ($d=1$) and the best nonlinear model ($d>1$) are obtained, respectively, by searching for κ^{lin} and κ^{nl}, which minimizes

$$
C(r) = \log \varepsilon(r) + r/N,
\tag{10}
$$

where $r \in [1,M]$ is the number of polynomial terms of the truncated Volterra expansions from a certain pair $\{\kappa,d\}$. The minimum C values corresponding with the linear and nonlinear model are C^{lin} and C^{nl}, respectively. If the best nonlinear model is significantly more predictive than the best linear model ($C^{\text{lin}} > C^{\text{nl}}$), it indicates that the original series is nonlinear. Otherwise, it may be inferred that the original series is not chaotic or the chaotic component is too weak to be detected.

III. RESULTS

During phonation, the high subglottal pressure pushes the left and right vocal folds apart, and air escapes from the lungs through the glottis. High-velocity airflow decreases the glottal airflow pressure, which causes the vocal folds to close. The glottal airflow provides the energy for the vocal fold vibration, so it has an important effect on vocal fold vibration and voice production. Figure 2 presents the spatiotemporal pattern of the downstream glottal airflow with $Q=0.64$ and $P_s=1.0$ kPa. The velocity value is detected at the line $y=1.0$ as shown in Fig. 1. It can be seen that the airflow shows complex behavior in both the time domain and the space domain. In the time domain, the airflow shows a series of pulses because of the glottis's periodic opening and closing. However, this periodicity of the airflow pulse in the time domain is not strict due to the disturbing effects of
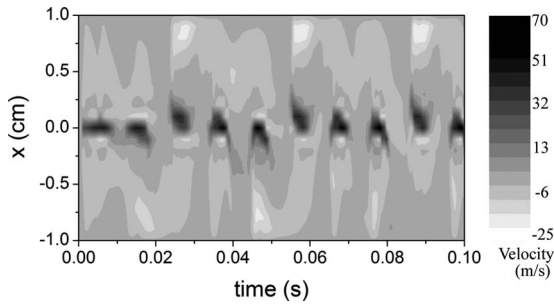
FIG. 2. The spatiotemporal pattern of airflow in the vocal tract.

vortices and turbulence. In the space domain, the airflow is significantly asymmetric due to the tension asymmetry of the vocal folds and the Coanda effect [27,30,32,37,46]. Previous studies have revealed that external noise could affect the onset of chaos [47,48]. Therefore, the complex spatiotemporal behavior of the aerodynamics could also influence vocal fold vibration and voice production.

Figure 3 plots the maxima of $x_{1\alpha}$ as a function of the asymmetric factor Q, which is predicted by the composite model with $P_s = 1.0$ kPa. Rich dynamic behaviors can be found in the composite vocal fold model. The dotted lines 1, 2, 3, and 4 correspond to four typical asymmetric situations, $Q = 0.64$, 0.50, 0.45, and 0.36, respectively. Their corresponding wave forms are illustrated in Figs. 4(a)–4(d).

In the following discussion, the different vibration patterns are labeled with the ratio of the number of the maxima of x_{1l} and x_{1r} during one total cycle [13]. An m:n mode (m and n are integrated numbers) denotes that there are m maxima of x_{1l} and n maxima of x_{1r} during one total cycle. At $Q = 0.64$, the waveforms of x_{1r} and x_{1l} obtained from the composite model are approximately periodic. Moreover, the number of the maxima of x_{1r} is the same as the number of the maxima of x_{1l}, as shown in Fig. 4(a). So we can consider the vibration mode to be 1:1 in this situation [13]. However, the periodicity of the composite model is not strict and the
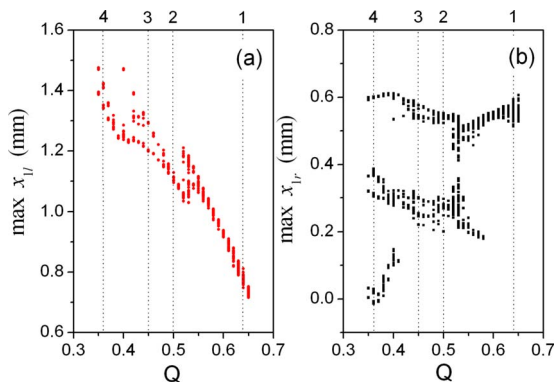

FIG. 3. (Color online) The maxima of $x_{1\alpha}$ as a function of the asymmetric factor Q obtained from the composite vocal fold model, where the subglottal pressure is 1.0 kPa. (a) The left vocal fold; (b) the right vocal fold. Both vocal folds exhibit rich dynamic behavior. The dotted lines 1, 2, 3, and 4 correspond to 1:1 mode ($Q = 0.64$), 1:2 mode ($Q = 0.5$), 2:4 mode ($Q = 0.45$), and 2:6 mode ($Q = 0.36$), respectively.
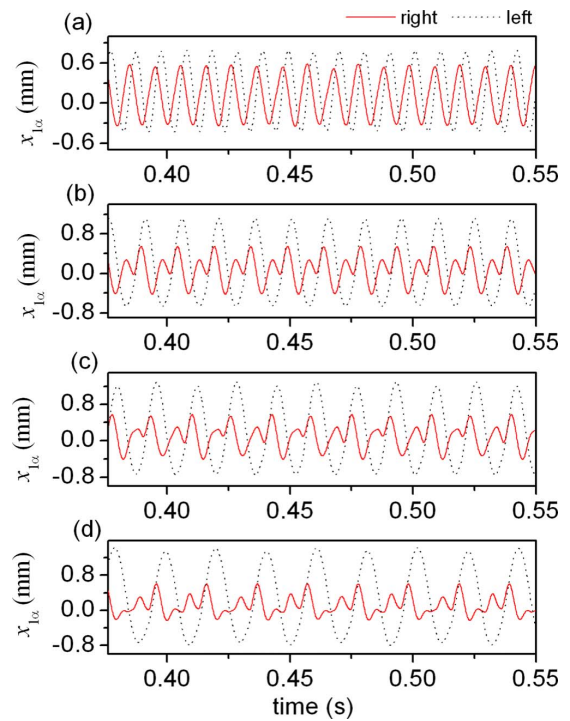

FIG. 4. (Color online) The waveforms of vocal fold vibration obtained from the composite model. (a) $Q = 0.64$ (1:1 mode); (b) $Q = 0.50$ (1:2 mode); (c) $Q = 0.45$ (2:4 mode); (d) $Q = 0.36$ (2:6 mode), where m:n mode is defined as the different vibration patterns with the ratio of the number of the maxima of x_{1l} and x_{1r} during one total cycle [13].

amplitude of oscillation fluctuates slightly. The aperiodic characteristic of the waveform can be better observed in the next-maxima plots, as shown in Figs. 5(a) and 5(b), where the y axis is the amplitude value of the ith peak and the x axis is the amplitude value of the next [$(i+1)$th] peak. It shows that the trajectory obtained from the composite model is scattered within a small domain, rather than being made up of one or several discrete points. This kind of attractor usually indicates a random or chaotic series. By applying Barahona and Poon's nonlinear dynamics detection method to analyze the dynamic characteristics of the maxima series, we are able to classify it as random or chaotic. It was found that the maxima series is predictable. Furthermore, the nonlinear polynomial is significantly more predictable than the linear polynomial. Figs. 5(c) and 5(d) presents these results, which suggest that the maxima series is not a random noise, but deterministic chaos [38].

At $Q = 0.50$, the peak number of the waveform of x_{1r} is almost twice the peak number of x_{1l}, which indicates a 1:2 vibration mode in this situation. Figures 6(a) and 6(b) illustrate the trajectory points of the left and right vocal folds in the next-maxima space. Similar to the situation at $Q = 0.64$, the trajectory is not comprised of discrete points, but is scattered within a certain domain. The trajectory points of the left vocal fold are distributed around (1.10,1.10). However, in contrast to the situation at $Q = 0.64$, the trajectory points of the right vocal fold are grouped in two separate islands. The trajectory evolves in these disconnected domains and alternatively between different points in each of them, as indi-
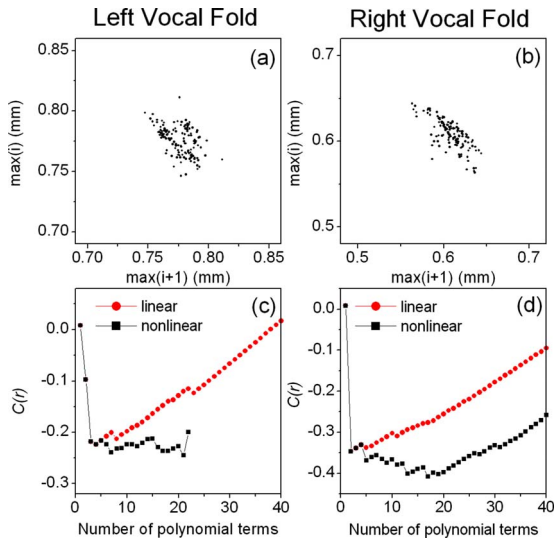
FIG. 5. (Color online) Detection nonlinear dynamics of the maximum series of $x_{1\alpha}$ from the composite model with $Q=0.64$. (a), (b) The next-maximum maps of x_{1l} and x_{1r}. (c), (d) The $C(r)$ value displayed as a function of the number of polynomial terms. In comparison with the linear polynomial, the more predictive nonlinear polynomial indicates the presence of deterministic chaos in the maximum series.

cated by the arrows. When the island in the next-maxima map of the left vocal fold is visited one time, the two islands in the next-maxima map of the right vocal fold are each visited one time. The ratio of the number of the maxima in the left and right vocal folds is 1:2 in one total cycle. Figures 6(c) and 6(d) present the analysis results of the nonlinear dynamics detection method on the maxima series. The better predictability of the nonlinear polynomial suggests that the maxima series of the left and right vocal folds are both es-
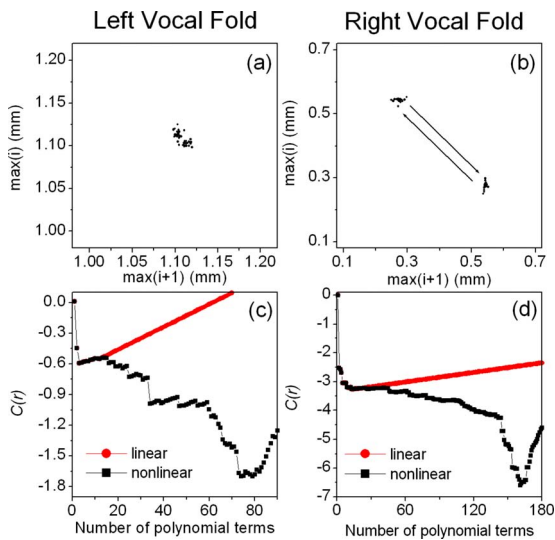

FIG. 6. (Color online) Detection nonlinear dynamics of the maximum series of $x_{1\alpha}$ from the composite model with $Q=0.50$. (a), (b) The next-maximum maps of x_{1l} and x_{1r}. (c), (d) The $C(r)$ value displayed as a function of the number of polynomial terms. The deterministic chaos is presented in the maximum series.
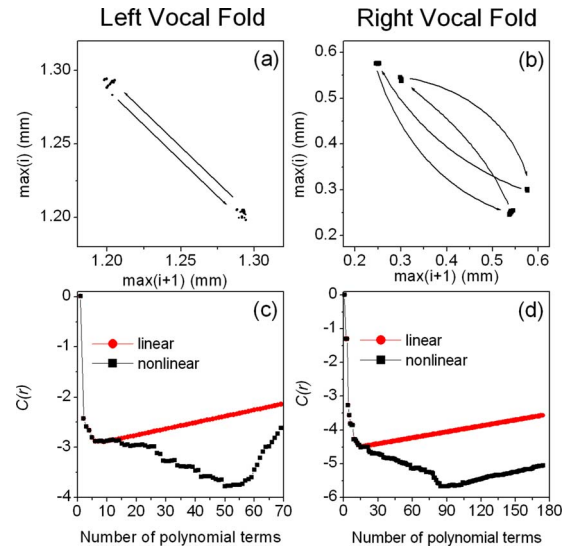

FIG. 7. (Color online) Detection nonlinear dynamics of the maximum series of $x_{1\alpha}$ from the composite model with $Q=0.45$. (a), (b) The next-maximum maps of x_{1l} and x_{1r}. (c), (d) The $C(r)$ value displayed as a function of the number of polynomial terms. The deterministic chaos is presented in the maximum series.

sentially deterministic chaos. In addition, it is noticeable that $C(r)$ of the nonlinear polynomial quickly increases after it has reached a minimum value. This could be a consequence of numerical error due to a large number of high-order nonlinear terms.

More complex dynamic behaviors can also be observed in the composite vocal fold model with the decreasing of asymmetric factor Q. At $Q=0.45$, there are two peaks in the waveform of x_{1l} and four peaks in the waveform of x_{1r} during one total cycle. In the next-maxima plots, the trajectory points of the left vocal fold are assembled in two separate islands, but the trajectory points of the right vocal fold are separated into four disconnected domains, as shown in Figs. 7(a) and 7(b). These separate islands are periodically visited. During one total cycle, two islands in the next-maxima map of the left vocal fold and four islands of the right vocal fold are visited, indicating a 2:4 mode at $Q=0.45$. Under this strong periodicity, the deterministic chaos can also be detected. Figures 7(c) and 7(d) illustrate the results of the nonlinear dynamics detection method. It was found that the nonlinear polynomial is much more predictable than the linear polynomial, so the maxima series is essentially deterministic chaos. Similarly, at $Q=0.36$, the vocal fold vibration is in 2:6 mode and the maxima series was noted to be deterministic chaos as well.

In addition, Steinecke and Herzel found that two different modes could coexist in an asymmetric two-mass model [13]. Moreover, the borderline between the basin attractions of both modes is smooth; therefore, rather weak perturbations may induce abrupt jumps to other regimes. In this study, besides the four typical 1:1, 1:2, 2:4, and 2:6 modes, we also found a transient mode in which the system switches from one mode to another mode. For example, at $Q=0.56$, the system sometimes showed a 1:1 mode, and sometimes showed a 1:2 mode. The transient mode might be the result of jumps between two coexisting modes due to aerodynamic perturbations.
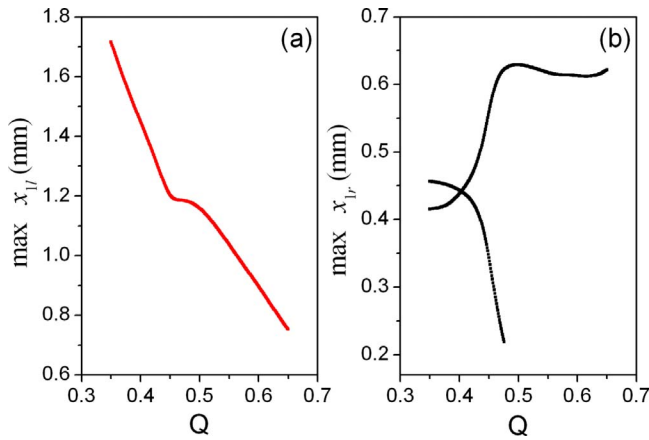
FIG. 8. (Color online) The maxima of $x_{1\alpha}$ as a function of the asymmetric factor Q obtained from the asymmetric two-mass model, where the subglottal pressure is 1.0 kPa ($P_s=0.01$). (a) The left vocal fold; (b) the right vocal fold. When $0.476 < Q \leq 0.65$, the left and right vocal folds show a 1:1 mode. When $0.35 \leq Q < 0.476$, the left and right vocal folds show a 1:2 mode.

The above simulation based on the asymmetric composite model showed rich dynamic behaviors. In order to investigate the role of glottal airflow in generating these dynamic behaviors, we also simulated the vocal fold vibration using the asymmetric two-mass model, where the glottal aerodynamics is simplified to Bernoulli's relationship. Figure 8 plots the maxima of $x_{1\alpha}$ as a function of the asymmetric factor Q, which is calculated from the asymmetric two-mass model with $P_s=1.0$ kPa. Just two kinds of dynamic behaviors can be found in the asymmetric two-mass model when the asymmetric factor Q is decreased from 0.65 to 0.35. When $0.476 < Q \leq 0.65$, the left and right vocal folds show 1:1 mode, as shown in Fig. 9(a). When Q is decreased below 0.476, the ratio of the left maxima number to the right maxima number shifts to 1:2, as shown in Fig. 9(b).

In addition, we can see that the waveform predicted by the two-mass model is strictly periodic, which also diverges from the results predicted by the composite model. The strict
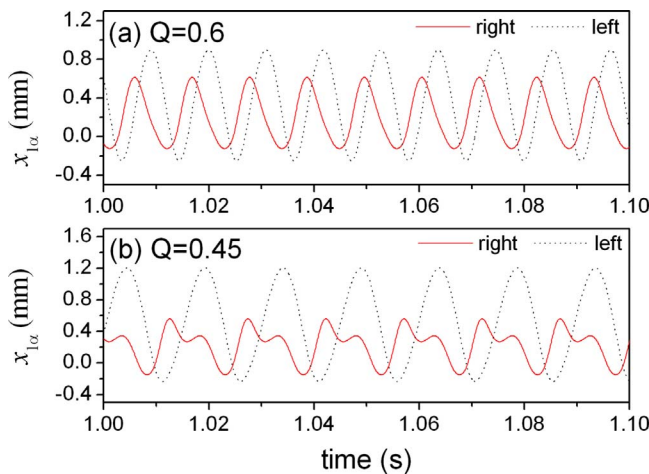

FIG. 9. (Color online) The waveforms of vocal fold vibrations obtained from the asymmetric two-mass model. (a) $Q=0.6$ (1:1 mode); (b) $Q=0.45$ (1:2 mode).
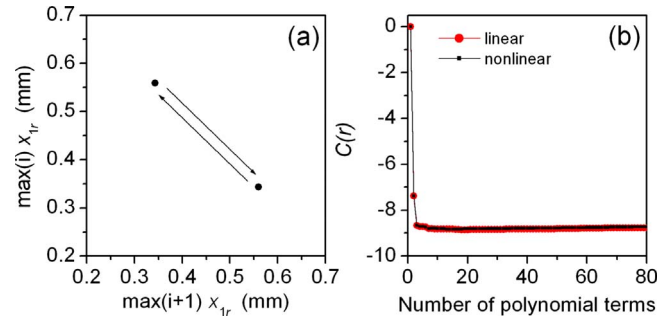

FIG. 10. (Color online) Nonlinear analysis on the maximum series of x_{1r} from the asymmetric two-mass model with $Q=0.45$. (a) The next-maximum maps; (b) the nonlinear detection.

periodicity can easily be displayed by the next-maxima plots. Figure 10(a) presents the next-maxima plots of the right vocal fold at $Q=0.45$. It can be observed that the trajectories are concentrated in two individual points, but are not separated into two distinct domains. The trajectory alternates between the two points. The vocal folds follow a strict double periodic oscillation. The nonlinear dynamics detection method shows that the nonlinear polynomial and linear polynomial are similarly predictive [see Fig. 10(b)], which confirms that the double period series from the two-mass model is regular and periodic, rather than chaotic.

The only difference between the two-mass model and the composite model is in their descriptions of glottal airflow. In the two-mass model, the glottal airflow is assumed to be laminar, nonviscous, and incompressible. The various dynamical behaviors of glottal airflow are ignored. However, in a composite model, the NS equations include a more vivid and accurate description of the airflow than does Bernoulli's law. The NS equations can predict complex spatiotemporal behaviors of airflow. We believe that the different dynamic behaviors predicted by the two different models are caused by this discrepancy between glottal airflow descriptions. The rich dynamic behaviors of the composite model could be related to the aerodynamic characteristics of glottal airflow. Therefore, the above comparison indicates that the complexity of glottal airflow could be one possible reason for the chaotic component in the voice production system.

IV. DISCUSSION AND CONCLUSION

Previous voice analyses have shown that the voices produced by mammals and birds have significant nonlinear characteristics. In the past, these nonlinearities in voice were usually explained by vocal fold asymmetry [13], the source-filter acoustic interaction [19], delayed feedback due to reflected sound in the vocal tract [18], turbulence noise disturbance [17], abnormal biomechanical parameters of vocal fold tissue [16], and extremely high subglottal pressure [10,16]. These theories successfully explain the rough voice that is often related to the pathological conditions of the vocal folds.

However, more general nonlinear phenomena have not been thoroughly explained by these previous studies. It is known that although a normal voice is strongly periodic, small undulations in amplitude still exist. Traditionally, this

disturbance of the voice amplitude is treated as a random noise and described by the shimmer parameters. Recently, nonlinear dynamic analysis of the speech signal shows that these nonlinear characteristics can be found even in normal voice with strong periodicity [8,44,45]. These findings indicate that the small-amplitude disturbances in a periodically healthy voice may not be random noise, and may instead be a deterministic signal. This concept is still arguable [49] because the strong periodic characteristics of voice signal often obscure these random undulations, thereby decreasing the reliability of the analysis results of nonlinear dynamic methods. Moreover, the mechanism of production of the random undulation in a periodic voice signal is not completely clear.

Titze has successfully explained the low-frequency voice perturbation as the results of the random nature of motor neuron firing [50,51]. It is known that the sustained muscle contractions are made up of a large number of individual twitches of groups of muscle fibers. A single twitch usually lasts only a fraction of a second. The interval between two twitches varies with the neural signals to fire the twitch. The neurological fluctuations are perceived as unsteadiness in pitch and loudness [50,51]. Besides the random neuron firing, this study indicates that the glottal airflow could be another source of voice perturbation. We found that the output of a composite model is strongly periodic. However, its amplitude value is disturbed by a noiselike disturbance; moreover, the amplitude of this noiselike disturbance is much smaller than the amplitude of the vocal fold vibration. For example, at $Q=0.64$, the noiselike disturbance amplitude makes up about 1% and 8% of the peak-to-peak amplitude of the left and right vocal folds, respectively. The nonlinear detection method confirms that the noiselike disturbance is not random, but rather it represents deterministic chaos. This simulation suggests that the vocal fold vibration can be considered a periodic signal modulated by a small-amplitude chaotic series. This result is reflected in a similar phenomenon often seen in voice. Although this phenomenon cannot be predicted by the asymmetric two-mass model, it can be observed in the composite model. Therefore, we can say that the small-amplitude chaotic component present in strong periodic voice production systems might be related to the complexities of glottal aerodynamics.

In addition, we also found that in a two-mass model that ignores the complexity of glottal airflow, only two vibratory modes (1:1 mode and 1:2 mode) can be observed within the asymmetric factor range $0.35 \leq Q \leq 0.65$. However, using a composite model that considers complex glottal airflow, the simulation showed more modes within the same asymmetric factor range, including 1:1 mode, 1:2 mode, 2:4 mode, and 2:6 mode. It is known that the vocal fold tissue and glottal airflow interact with each other during voice production. This suggests that the effect of the airflow force on the vocal fold dynamics is not the corruption of observations by errors that are independent of the dynamics, but a feedback process wherein the system is perturbed by a small random amount at each time step, just like the effect of dynamic noise on a dynamical system [47,48,55]. Previous studies found that the dynamic disturbance could significantly shift the structure of the bifurcation diagram [17,47,48,52–55]. Consequently, this rich dynamic phenomenon in the composite model could also be related to the complex aerodynamics of glottal airflow.

In a previous study, it was reported that the disrupting effect of noise becomes more significant when the system is close to the bifurcation point corresponding to the onset of sustained oscillations [55]. In this study, our findings based on 1.0 kPa subglottal pressure are obtained well inside the domain of sustained oscillations. When the subglottal pressure is close to the onset value of sustained oscillation, the lower subglottal pressure brings the vocal fold system near to its bifurcation point, which could create a significant chaos-noise effect [55]. On the other hand, the lower subglottal pressure indicates smaller aerodynamic noise and more regular glottal airflow. As a result, it can be expected that subglottal pressure is an important parameter that influences the chaotic component of a voice production system, which might be of interest for further study.

In summary, this study found that glottal aerodynamics could greatly contribute to the generation of the nonlinear dynamic behaviors of voice production, such as bifurcation and a small-amplitude chaotic component obscured by strong periodicity. These results might be useful in understanding the dynamic mechanisms of voice production and in improving the quality of speech synthesis.

ACKNOWLEDGMENTS

[1] A. Cumming and P. S. Linsay, Phys. Rev. Lett. **60**, 2719 (1988).

[2] J. P. Crutchfield, B. A. Huberman, and N. H. Packard, Appl. Phys. Lett. **37**, 750 (1980).

[3] R. Hornero, P. Espino, A. Alonso, and M. Lopez, IEEE Eng. Med. Biol. Mag. **18**, 73 (1999).

[4] C. S. Poon and C. K. Merrill, Nature **389**, 492 (1997).

[5] R. Laje and G. B. Mindlin, Phys. Rev. Lett. **91**, 258104 (2003).

[6] M. A. Trevisan, J. M. Mendez, and G. B. Mindlin, Phys. Rev. E **73**, 061911 (2006).

[7] M. A. Trevisan, B. Cooper, F. Goller, and G. B. Mindlin, Phys. Rev. E **75**, 031908 (2007).

[8] A. Kumar and S. K. Mullick, J. Acoust. Soc. Am. **100**, 615 (1996).

[9] W. Mende, H. Herzel, and K. Wermke, Phys. Lett. A **145**, 418 (1990).

[10] Y. Zhang and J. J. Jiang, Phys. Rev. E **72**, 035201 (2005).

[11] I. Tokuda, T. Riede, J. Neubauer, M. J. Owren, and H. Herzel, J. Acoust. Soc. Am. **111**, 2908 (2002).

[12] A. Facchini, H. Kantz, and E. Tiezzi, Phys. Rev. E **72**, 021915 (2005).

[13] I. Steinecke and H. Herzel, J. Acoust. Soc. Am. **97**, 1874 (1995).

[14] M. A. Trevisan, M. C. Eguia, and G. B. Mindlin, Phys. Rev. E **63**, 026216 (2001).

[15] C. Tao, Y. Zhang, G. Du, and J. J. Jiang, Phys. Rev. E **69**, 036204 (2004).

[16] J. J. Jiang, Y. Zhang, and J. Stern, J. Acoust. Soc. Am. **110**, 2120 (2001).

[17] J. J. Jiang and Y. Zhang, J. Acoust. Soc. Am. **112**, 2127 (2002).

[18] R. Laje, T. Gardner, and G. B. Mindlin, Phys. Rev. E **64**, 056201 (2001).

[19] R. Laje and G. B. Mindlin, Phys. Rev. E **72**, 036218 (2005).

[20] D. A. Berry, H. Herzel, I. R. Titze, and K. Krischer, J. Acoust. Soc. Am. **95**, 3595 (1994).

[21] J. Neubauer, P. Mergell, U. Eysholdt, and H. Herzel, J. Acoust. Soc. Am. **110**, 3179 (2001).

[22] A. Aitta, G. Ahlers, and D. S. Cannell, Phys. Rev. Lett. **54**, 673 (1985).

[23] M. Torralba, A. A. Castrejón-Pita, G. Hernández, G. Huelsz, J. A. del Río, and J. Ortín, Phys. Rev. E **75**, 056307 (2007).

[24] F. Jousse, R. Farr, D. R. Link, M. J. Fuerstman, and P. Garstecki, Phys. Rev. E **74**, 036311 (2006).

[25] H.-W. Xi, X.-J. Li, and J. D. Gunton, Phys. Rev. Lett. **78**, 1046 (1997).

[26] M. Das, B. Chakrabarti, C. Dasgupta, S. Ramaswamy, and A. K. Sood, Phys. Rev. E **71**, 021707 (2005).

[27] X. Pelorson, A. Hirschberg, R. R. van Hassel, and A. P. J. Wijnands, J. Acoust. Soc. Am. **96**, 3416 (1994).

[28] R. C. Scherer, D. Shinwari, K. J. DeWitt, C. Zhang, B. R. Kucinschi, and A. A. Afjeh, J. Acoust. Soc. Am. **109**, 1616 (2001).

[29] F. Alipour and R. C. Scherer, J. Acoust. Soc. Am. **112**, 2996 (2002).

[30] G. C. Hofmans, G. Groot, M. Ranucci, G. Graziani, and A. Hirschberg, J. Acoust. Soc. Am. **113**, 1658 (2003).

[31] B. R. Kucinschi, R. C. Scherer, K. J. DeWitt, and T. T. M. Ng, J. Acoust. Soc. Am. **119**, 3011 (2006).

[32] D. Shinwari, R. C. Scherer, K. J. DeWitt, and A. A. Afjeh, J. Acoust. Soc. Am. **113**, 487 (2003).

[33] W. Zhao, C. Zhang, S. H. Frankel, and L. Mongeau, J. Acoust. Soc. Am. **112**, 2134 (2002).

[34] C. Zhang, W. Zhao, S. H. Frankel, and L. Mongeau, J. Acoust. Soc. Am. **112**, 2147 (2002).

[35] M. Krane, M. Berry, and T. Wei, J. Acoust. Soc. Am. **122**, 3659 (2007).

[36] M. P. de Vries, H. K. Schutte, A. E. P. Veldman, and G. J. Verkerke, J. Acoust. Soc. Am. **111**, 1847 (2002).

[37] C. Tao, Y. Zhang, D. G. Hottinger, and J. J. Jiang, J. Acoust. Soc. Am. **122**, 2270 (2007).

[38] M. Barahona and C. S. Poon, Nature **381**, 215 (1996).

[39] C. S. Poon and M. Barahona, Science **98**, 7107 (2001).

[40] K. Ishizaka and J. L. Flanagan, Bell Syst. Tech. J. **51**, 1233 (1972).

[41] X. Y. Luo and T. J. Pedley, J. Fluids Struct. **9**, 149 (1995).

[42] X. Y. Luo and T. J. Pedley, J. Fluid Mech. **314**, 191 (1996).

[43] J. D. Farmer and J. J. Sidorowich, Phys. Rev. Lett. **59**, 845 (1987).

[44] I. Tokuda, T. Riede, J. Neubauer, M. J. Owren, and H. Herzel, J. Acoust. Soc. Am. **111**, 2908 (2002).

[45] C. Tao, J. Mu, X. Xu, and G. Du, Acoust. Sci. & Tech. **25**, 51 (2004).

[46] B. D. Erath and M. W. Plesniak, J. Acoust. Soc. Am. **120**, 1000 (2006).

[47] J. P. Crutchfield and B. A. Huberman, Phys. Lett. **77A**, 407 (1980).

[48] J. P. Crutchfield, M. Nauenberg, and J. Rudnick, Phys. Rev. Lett. **46**, 933 (1981).

[49] M. Banbrook, S. McLaughlin, and I. Mann, IEEE Trans. Speech Audio Process. **7**, 1 (1999).

[50] Ingo R. Titze, J. Speech Hear. Res. **34**, 460 (1991).

[51] Ingo R. Titze, *Principles of Voice Production* (NCVS, Iowa City, IA, 1994).

[52] S. K. Hwang, J. B. Gao, and J. M. Liu, Phys. Rev. E **61**, 5162 (2000).

[53] M. Diestelhorst, R. Hegger, L. Jaeger, H. Kantz, and R. P. Kapsch, Phys. Rev. Lett. **82**, 2274 (1999).

[54] R. Wackerbauer and S. Kobayashi, Phys. Rev. E **75**, 066209 (2007).

[55] D. Gonze, J. Halloy, and A. Goldbeter, Proc. Natl. Acad. Sci. U.S.A. **99**, 673 (2002).