Acoustic characteristics of phonation in “wet voice” conditions

Shanmugam Murugappan^{a)}

Department of Otolaryngology, Head and Neck Surgery, University of Cincinnati Medical Center,
231 Albert B. Sabin Way, Cincinnati, Ohio 45267-0528

Suzanne Boyce

Department of Communication Science and Disorders, French East Building, University of Cincinnati,
3202 Eden Avenue, Cincinnati, Ohio 45267-0379

Sid Khosla

Department of Otolaryngology, Head and Neck Surgery, University of Cincinnati Medical Center,
231 Albert B. Sabin Way, Cincinnati, Ohio 45267-0528

Lisa Kelchner

Department of Communication Science and Disorders, French East Building, University of Cincinnati,
3202 Eden Avenue, Cincinnati, Ohio 45267-0379

Ephraim Gutmark

Department of Aerospace Engineering and Engineering Mechanics and Department of Otolaryngology,
Head and Neck Surgery, University of Cincinnati Medical Center, 231 Albert B. Sabin Way, Cincinnati,
Ohio 45267-0528

(Received 11 May 2009; revised 11 January 2010; accepted 13 January 2010)

A perceptible change in phonation characteristics after a swallow has long been considered evidence that food and/or drink material has entered the laryngeal vestibule and is on the surface of the vocal folds as they vibrate. The current paper investigates the acoustic characteristics of phonation when liquid material is present on the vocal folds, using *ex vivo* porcine larynges as a model. Consistent with instrumental examinations of swallowing disorders or dysphagia in humans, three liquids of different Varibar viscosity (“thin liquid,” “nectar,” and “honey”) were studied at constant volume. The presence of materials on the folds during phonation was generally found to suppress the higher frequency harmonics and generate intermittent additional frequencies in the low and high end of the acoustic spectrum. Perturbation measures showed a higher percentage of jitter and shimmer when liquid material was present on the folds during phonation, but they were unable to differentiate statistically between the three fluid conditions. The finite correlation dimension and positive Lyapunov exponent measures indicated that the presence of materials on the vocal folds excited a chaotic system. Further, these measures were able to reliably differentiate between the baseline and different types of liquid on the vocal folds.

© 2010 Acoustical Society of America. [DOI: 10.1121/1.3308478]

PACS number(s): 43.70.Gr, 43.60.Hj, 43.40.Ga, 43.60.Wy [AL]

Pages: 2578–2589

I. INTRODUCTION

A. Background

Because a number of medical conditions affect the vibratory characteristics of the larynx, acoustic differences between normal and abnormal phonation have been of great interest to scientists and clinicians, as aids to diagnosis and treatment. Most of this work has been focused on conditions that result from normal versus abnormal physiology related to laryngeal disease. For instance, there is a long tradition of comparing normal voice production with that from patients with vocal folds polyps, cysts and nodules, vocal folds edema, spasmodic dysphonia, Parkinson disease, and surgical scarring, among others (Klingholtz, 1990, Kreiman *et al.*, 1993, Herzel *et al.*, 1994, Giovanni *et al.*, 1999a, Giovanni *et al.*, 1999b, Vieira *et al.*, 2002, Jiang *et al.*, 2006, Zhang

et al., 2005, Zhang and Jiang, 2004). One type of abnormal voice production that has not been well studied is the case of vocal fold vibration when foreign material is present in the laryngeal area. This phenomenon occurs frequently in patients with poor coordination of airway protection and swallowing, because solid and/or liquid food may enter the laryngeal vestibule during the pharyngeal phase on its way to the esophagus. It also is commonly found in patients who fail to clear mucosal secretions from the surface of the larynx by coughing, primarily because they have diminished sensory feedback due to disease. While foreign material of this kind does not prevent voice production, secretions, or food material passing through the glottis into the trachea (commonly referred to as aspiration) and thence to the lungs means that patients with these problems are at high risk for respiratory diseases. Thus, detection of foreign materials within the larynx and on the true vocal folds is potentially a useful clinical indicator of swallowing problems.

Patients with suspected swallowing disorders, who show a change in voice quality, are often described as demonstrat-

^{a)} Author to whom correspondence should be addressed. Electronic mail: mugam@uc.edu

ing a wet or “gurgly” voice (Logemann 1998, Murray *et al.*, 1996, Warms and Richards, 2000). While this phenomenon is widely known among clinicians and often assumed to indicate aspiration risk when observed during a clinical evaluation or meal, few studies have addressed the phenomenon from an objective and/or quantitative point of view. In one of the two studies readily available in the literature, Ryu *et al.* (2004) compared phonation in 23 patients who had shown in an earlier videofluorographic study that they were at high versus low risk for laryngeal penetration of food materials. They found that patients in the high risk group showed some statistically significant differences in acoustic indices immediately after swallowing, when compared to low risk patients. However, they did not have independent evidence that the phonation behavior was linked simultaneously to the presence of foreign materials on the larynx. In a later study, Groves-Wright (2007) collected simultaneous recordings of phonation and digital fluorographic images of the vocal tract, including the larynx, from 78 patients. She compared patients *with* observable materials in the laryngeal vestibule to patients *without* such observable materials, and found statistically significant differences in acoustic measures [e.g., jitter, shimmer, noise-to-harmonic ratio (NHR)]. Thus, there are strong reasons to believe that foreign materials present in the larynx can induce acoustical changes in phonation. Further, the acoustical characteristics of phonation under such conditions may provide a clinical marker of swallowing dysfunction and disease.

In a real patient, it is impossible to fully control the amounts and total viscosity of the materials on the larynx, both because each swallow is different, and because the materials may mix with secretions already present. Thus, it is difficult to determine the relationship between amount and viscosity of materials, and the acoustical characteristics of voice production when materials are present. An excised larynx model, on the other hand, allows easy control of these factors. Excised larynx models have already contributed significantly to our understanding of laryngeal physiology and voice production (Berry *et al.*, 1996, Alipour *et al.*, 1997, Švec *et al.*, 1999, Jiang *et al.*, 2003, Ayache *et al.*, 2004, Khosla *et al.*, 2007).

B. Review of acoustic measures

Ryu *et al.* (2004) and Groves-Wright (2007) employed the classic measures of jitter, shimmer, NHR, relative average perturbation, voice turbulence index, and spectral tilt. In recent years, non-linear methods of analysis have been gaining popularity in the study of vocal folds acoustics (Mende *et al.*, 1990, Herzel *et al.*, 1994, Titze, 1994a, Steinecke and Herzel, 1995, Narayanan and Alwan, 1995, MacAuslan *et al.*, 1997, Giovanni *et al.*, 1999a, Giovanni *et al.*, 1999b, Švec *et al.*, 1999, Jiang *et al.*, 2003, Zhang and Jiang, 2004, Zhang *et al.*, 2005, Jiang *et al.*, 2006, Zhang and Jiang, 2008). This emphasis on non-linear methods is due to recognition that phonation involves a number of non-linear mechanisms, such as flow-pressure relations, source-tract interaction, flow turbulence, and stress-strain relations within the vocal fold tissue. Similar effects apply to the vocal folds

in the course of collision. The presence of material on the larynx during phonation also fits the model of a non-linear system because the presence of foreign materials introduces the behavior of a dynamic load, which is subject to a large number of variables with high sensitivity and low predictability. For instance, fluid droplets due to secretions might remain in place or be flung far from their previous location as a result of small variations in the composition and tension (or motion) of the vocal folds tissue, the local viscosity of the fluid itself, temperatures differences at different points within the larynx, and variations in force exerted by the glottal airstream. Further, loss of the fluid onto other anatomical structures (e.g., aryepiglottic folds and posterior pharyngeal wall) or down into the larynx, subglottic region, and trachea (evaporation, impingement of the particles on each other, and fluid-dynamic instabilities, among other factors) may combine to affect the acoustic output.

The purpose of the current study is to conduct a series of experiments on an excised larynx model, with the aim of (1) methodically characterizing the acoustics and vibratory behavior of the larynx under controlled conditions due to differing viscosities of materials present during phonation, and (2) determining which acoustic measures are most highly correlated with changes in the vibratory behavior of the vocal folds due to foreign materials present during phonation.

II. METHODS

A. Experiment

Most previous *ex vivo* studies have used canine larynges, which resemble the human vocal anatomy tissue composition and angle relative to the trachea, but show some differences in size and vibratory characteristics. Recent data suggest that porcine larynges are anatomically similar to human larynges in most aspects, and show acoustical and vibratory behavior similar to that of humans (Jiang *et al.*, 2001), but differ in having a 40° angle from vocal process (higher) to anterior commissure (lower) relative to the trachea (Alipour and Jaiswal, 2009). For purposes of this study, which is focused on acoustical and vibratory characteristics, we chose the porcine model as most appropriate. It should be noted that the more angled anatomy of the porcine model may result in greater pooling and longer persistence of liquid on the vocal folds at the anterior commissure, relative to behavior in the human larynx.

Six excised porcine larynges were obtained within 4 h after sacrifice. Cartilage and soft tissue above the true vocal folds was removed. A “collar” or lip of a few millimeters (ranging between 1 mm and 4 mm) around the larynx was left in place. After the extraneous tissue and muscles were dissected, removed, and cleaned, the larynges were immediately placed in a normal saline solution (0.9% NaCl). The tracheas were 8–10 cm long. The average glottal length measured from the anterior commissure to the vocal process was 20.8 mm, and it varied from 19.6–22.9 mm between larynges. For all larynges, the inferior 4 cm of the trachea was placed over a rigid tube (inner diameter of 1/2 in. and outer diameter of 5/8 in.). The outer trachea wall was clamped to the tube to prevent air leaks. A stitch was used to adduct the

FIG. 1. (Color online) Photograph of the experimental setup.

vocal processes. Khosla placed one suture through both vocal processes at the same level with the aid of magnification. The stitch was tied with the minimal tension needed to have a prephonatory width between the vocal processes of 0 mm. Special care was taken to position the suture symmetrically in both the anterior-posterior and inferior-superior direction. The posterior glottis was completely closed with a suture. The larynx was fixed in space by using a square mounting apparatus that had four double prong pins on each side. Each pin was inserted into the thyroid cartilage (see also Fig. 1). Symmetry was monitored via video camera placed above the glottal exit. The flow that exited the rigid tube and entered the trachea was supplied by a blower, which could produce a maximum flow rate of 2500 cc/s at 35 psi. Pressure regulator, thermocouple, electronic pressure gauges, mass flow meter, and an electronic control valve were used to regulate the air upstream. The air was moistened using a humidifier with thermostat control (Conchatherm III, Hudson Respiratory Care Inc., Temecula, California).

Acoustic recordings were obtained by a Bruel and Kjaer freefield $\frac{1}{2}$ in. microphone (Model:4189), placed 12 cm above the glottal exit in such a way that it did not interfere with laryngeal airflow. The microphone has a sensitivity of 4 mV/Pa, a frequency bandwidth of 4 Hz–100 kHz (± 2 dB), and dynamic sensitivity of 28–164 dB. The response of this microphone is completely flat from 10 Hz–50 kHz and has 3% distortion for SPL > 164 dB. The axis of the microphone was at 45° inward to the plane of the glottal exit. Of the various microphone settings, only the filter setting was changed to a low cut-off of 0.1 Hz and a high cut-off of 11 kHz. These settings were maintained for all the experiments. Static and dynamic subglottal pressure was recorded by a Honeywell pressure transducer (Model FPG1 WB) with sensitivity in the range of ± 24 cm H₂O relative to atmospheric pressure, a frequency response up to 2 kHz, and an accuracy of 0.1%. This pressure transducer was placed flushed on the walls of the rigid tube at a distance of 15 cm, below the vocal folds. Both the microphone and pressure transducers were calibrated before the start of every experiment. The EGG, acoustic and subglottal pressure signals recorded during phonation were each sampled at 20 kHz using a 16 bit NI data acquisition card (Model PCI 6259). A high-speed video camera (High Speed Phantom Version 7.1, sampled at 8 kHz)

was placed approximately 30 cm above the glottal exit to visualize the vocal fold vibration and fluid motion on the laryngeal surface.

The experiments described in this paper used liquid of standardized viscosities as supplied by a manufacturer of Varibar barium (E-Z, EM, Inc.). The liquids used differed respectively by an order of magnitude: (1) Varibar “thin liquid,” with a viscosity of 4 cP, (2) Varibar “nectar,” with a viscosity of 300 cP, and (3) Varibar “improved honey,” with a viscosity of 3000 cP. Our aims in using barium were: (a) because it can be obtained in standardized viscosity form; (b) it is opaque and easy to visualize by visible wavelength camera; and (c) it is routinely used during videofluorographic (VFSS) studies of human swallowing function. Although there may be reasons to suspect that VFSS tests using barium-mixed liquids are not fully representative of natural swallowing behaviors (Steele and Van Lieshout, 2005), barium as a contrast agent is always required for radiographic diagnostic imaging of swallow function (Logemann, 1993, 1998). We reasoned that using barium in our *ex vivo* studies would be maximally comparable to this commonly used clinical test.

The same sequence of events was followed for each of the six larynges. There were four trials. Condition 1 consisted of a baseline trial (phonation with no fluid on the folds). Conditions 2, 3, and 4 involved the presence of materials on the larynx. The latter conditions will be referred to as the “viscosity” conditions. For each trial, a premeasured amount of liquid (0.3 ml) was placed on the larynx at a distance of approximately 1 cm above the glottal exit before the start of phonation by locating a syringe above the horizontal and vertical midline of the folds. Again for each trial, the larynx was set into vibration by increasing the subglottal pressure until the larynx began to sustain phonation. This pressure, i.e., the minimum pressure required to set the vocal folds into vibration, is commonly called phonation threshold pressure (PTP) (Titze, 1994b). Because it reflects various aspects of the physiology of the vocal folds (i.e., scar tissue, edema, bowing, among others) it is an important clinical parameter that appears to correlate well with speech system dysfunctions (Titze, 1988; Steinecke and Herzel, 1995; Lucero and Koenig, 2005, 2007; Tao and Jiang, 2008). We recorded the pressure at this point for each experimental condition.

The larynx was kept in vibration at PTP for 2 s. Before each experiment began, the larynx was cleaned of any remaining materials. The larynx was then removed from the apparatus and soaked in saline solution for 10 min before the start of the next condition. Images from the baseline condition were used to verify that the larynx was configured, within the apparatus, in the same manner for each of the conditions.

Observation of the high-speed images revealed that, during the 2-s period of the phonation experiment, each of the fluids (honey, nectar, and thin liquid) remained within the “collar” of tissue surrounding the vocal folds. All of the fluids were observed to slip gradually through the glottis, as the vocal folds opened and closed. We assumed evaporation to be negligible over this 2-s period.

B. Acoustic experiments

The first 500 ms of the acoustic recordings were extracted, and the following measurements were recorded: (1) average phonation frequency at PTP, (2) % jitter, (3) % shimmer, (4) correlation dimension (D_2), and (5) maximum Lyapunov exponent (L). In addition, we examined the signals using (6) conventional (CN) spectrograms, (7) time-corrected (TC) spectrograms, and (8) phase plots. % jitter, % shimmer, and average phonation frequency were calculated using the TF32 voice analysis software (Milenkovic, 2006). Jitter is defined as cycle to cycle variation in pitch period during voicing, and shimmer is variation between amplitudes in waveform cycles. In this study, we report both of these measures as percentages. The algorithm used to compute instantaneous pitch is that described by Milenkovic (1987).

Multibody dynamic systems such as those described here—i.e., liquid material lying on the surface of a vibrating vocal fold—are known to give rise to additional, non-harmonic frequencies. These frequencies are not well represented by conventional spectrograms, which tend to “smear” them. Accordingly, we evaluated the acoustic time series (x) with both conventional and time-corrected spectrograms (Fulop and Fitz, 2006). The later technique is a variant of the spectrogram reassignment method described in Fulop and Fitz (2006), as follows.

Given a spectrogram S_x , $S_x = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_x(s, \xi) W_h(t-s, \nu - \xi) ds d\xi$, which is defined as the two-dimensional convolution of the Wigner–Ville distribution (WVD), W_x and W_h , $W_x(t, \nu) = \int_{-\infty}^{\infty} x(t + (\tau/2)) x^*(t - (\tau/2)) e^{-j2\pi\nu\tau} d\tau$, and similarly, W_h is the WVD of the window function; superscript $*$ denotes the complex conjugate.

Each point in time (t) and frequency (ν) space in the spectrogram (t, ν) is reassigned to another point (t', ν') in the time-corrected spectrogram. Mathematically,

$$\hat{t}(x; t, \nu) = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s W_h(t-s, \nu-\xi) W_x(s, \xi) ds d\xi}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_h(t-s, \nu-\xi) W_x(s, \xi) ds d\xi} \hat{\nu}(x; t, \nu)$$

$$= \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \xi W_h(t-s, \nu-\xi) W_x(s, \xi) ds d\xi}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_h(t-s, \nu-\xi) W_x(s, \xi) ds d\xi}.$$

The reassigned spectrogram $S_x^{(r)}$ at any point (t', ν') is given as

$$S_x^{(r)}(t', \nu'; h) = \int_{-\infty}^{\infty} S_x(t, \nu; h) \delta(t' - \hat{t}(x; t, \nu)) \delta(\nu' - \hat{\nu}(x; t, \nu)) dt d\nu.$$

The reassignment can be thought of as a two step process: (1) smoothing using a window function, which dampens the oscillations that arises due to the cross interference terms; and (2) refocus the contributions that survived after smoothing. This method introduces a weighing function to t' and ν' ,

such that the new time frequency coordinate better reflects the distribution of energy in the analyzed signal.

In order to allow better comparison between the conventional and TC spectrogram, the variables: (1) type of window (Hamming), (2) window size (256 samples), (3) overlap of the signal (50%), and (4) number of points of Fast Fourier Transform (FFT) (1024) were kept constant. Further, we observed from the high-speed video images that the liquid placed on the larynx tended to follow a cycle of collecting in pools, dispersing, and then collecting again. We suspected that this action of the material on the larynx might produce signals that were intermittent in nature.

C. Nonlinear methods

To investigate the nonlinear dynamics of the acoustic time series data $x(t_i)$, we employed three variables: (1) phase space methods, (2) correlation dimension, and (3) Lyapunov exponents (Heath, 2000; Kantz and Schreiber, 2000; Sprott, 2003). An m -dimensional phase space could be reconstructed as $y_i = \{x(t_i), x(t_i - \tau), \dots, x(t_i - (m-1)\tau)\}$, where $x(t_i)$ represents a single sample at time instant t_i . The signal $x(t)$ is acquired at equal intervals, $dt = 1/\text{fs}$, where fs is the sampling frequency (fs = 20 kHz in this study). These m -dimensional vectors, y_i , trace a trajectory in time, which eventually settle down on a path known as an attractor (Heath, 2000; Kantz and Schreiber, 2000; Sprott, 2003). The delay time τ and embedding dimension m are two fundamental parameters in the phase space reconstruction. Takens (1988) showed that when $m > 2D + 1$, where D = Hausdorff dimension, the reconstructed phase space using the lagged coordinates is topologically equivalent to the original phase space. In this study, we employed the Takens method to estimate m . The time delay τ was identified according to the method of Fraser and Swinney (1986) as the first minimum corresponding to the least value of the mutual information S . Mathematically, $S = -\sum_{i,j} P_{ij}(\tau) \ln P_{ij}(\tau) / P_i P_j$, where P_i is the probability of finding a time series value in the i th interval and P_{ij} is the joint probability that (a) an observation falls on the i th interval and (b) after a delay τ , the same observation falls on the j th interval.

1. Correlation Dimension

The measure of correlation dimension D_2 is commonly used to study invariant characteristics of non-linear dynamical systems (Theiler, 1990). The higher the value of D_2 , the more complex the dynamical system, and the more degrees of freedom are needed to describe the dynamics of the system. In contrast, lower D_2 values require lesser states to describe the system. Mathematically, the correlation sum for a finite data set such as a time series is given as

$$C(N, \varepsilon) = \frac{2}{N(N-1)} \sum_{i=1}^N \sum_{j=i+1}^N \Theta(\varepsilon - \|x_i - x_j\|).$$

In this case, Θ is the Heaviside function, $\Theta(x) = 0$ if $x \leq 0$, and $\Theta(x) = 1$ for $x > 0$. The Theiler correction was applied to $C(N, \varepsilon)$ to discard data points that are strongly correlated in

TABLE I. *Post hoc* Tukey HSD Q values (critical Q at an alpha of 0.05 is 4.9).

	Q value (correlation dim)	Q value (Lyapunov exponent)	Q value (PTP)
Baseline versus thin liquid	5.85 ^a	2.8	0.4
Baseline versus nectar	8.7 ^a	8.6 ^a	0.75
Baseline versus honey	12.7 ^a	9.3 ^a	8.12 ^a
Nectar versus thin liquid	2.8	5.8 ^a	0.4
Nectar versus honey	4.0	0.7	7.37 ^a
Honey versus thin liquid	6.9 ^a	6.4 ^a	7.78 ^a

^aStatistically significant differences.

the phase space (Theiler, 1986, 1987). The modified formula is given as

$$C(N, \varepsilon, W) = \frac{2}{(N - W)(N + 1 - W)} \sum_{i=1}^N \sum_{j=i+1}^N \Theta(\varepsilon - \|x_i - x_j\|).$$

In this case, W is the number of closely spaced data points around x_i that are correlated. The correlation dimension D_2 is computed as the slope of logarithmic correlation sum to logarithmic distance ε ,

$$d(N, \varepsilon) = \frac{\partial \ln C(N, \varepsilon, W)}{\partial \ln \varepsilon}$$

$$D_2 = \lim(\varepsilon \rightarrow 0) \lim(N \rightarrow \infty) d(N, \varepsilon)$$

In the current study, C was plotted as a function of ε for varying values of m . D_2 was estimated as the slope of the linear fit in the region where all the curves collapse for dimensional values greater than m (embedding dimension).

2. Lyapunov exponent

The Lyapunov exponent is another parameter that is used to describe the stability properties (i.e., sensitivity) of the dynamical system. A positive Lyapunov exponent L is an indication of the presence of chaos. The Lyapunov exponent expresses the relationship between positive and negative deviations from a trajectory in phase space. When deviations are damped over time, $L < 0$. When deviations grow over time, $L > 0$. A negative L is associated with stability of the dynamical system, while a positive L is associated with instability of the dynamical system (Wolfe *et al.*, 1985). In this study we used the Wolfe *et al.* (1985) method to compute the largest positive Lyapunov exponent L . The algorithm involves computing Euclidean distances between two neighboring points in the attractor over an “evolution time” T . The ratio of initial and final separation distances $d(0)$ and $d(T)$ is used to compute L , using the formula

$$L = \frac{1}{T} \left\langle \ln \frac{d(T)}{d(0)} \right\rangle,$$

where $\langle \rangle$ denotes average over several data points. The averaging procedure is performed over small separation distances d and along the most unstable direction (Wolfe *et al.*, 1985).

D. Statistical methods

A statistical significance level of $p=0.05$ was arbitrarily chosen in this study. A one-way analysis of variance (ANOVA) test was performed separately for each of the six different parameters (phonation threshold pressure, phonation frequency, % jitter, % shimmer, Lyapunov exponent, correlation dimension) across the four groups (baseline, thin liquid, nectar, and honey). ANOVA test is used to check the null hypothesis that different groups come from the same distribution (Dawson and Trapp, 2004). The ANOVA test assumes that data samples be drawn from normally distributed populations and that variances between the groups are similar. These two conditions were reasonably satisfied by the test data for the different computed parameters. A p value less than 0.05 indicates that a difference exist among different groups. For those parameters where the ANOVA results (from the parameter mean) were significant ($p < 0.05$), a pairwise comparison was made among the groups using Tukey HSD *post hoc* analysis. A pairwise comparison involves comparing each of the possible pairs among the four groups (six possible pairs for four different groups, see also Table I). Stoline (1981) indicates that the Tukey HSD method is the most powerful and accurate method to make pairwise *post hoc* comparisons. MATLAB 7.6 (Mathworks) was used for statistical analysis.

III. RESULTS

A. Subglottal pressure experiments

The first set of experiments was conducted to determine a phonation threshold, i.e., the subglottal pressure required to initiate sustained phonation for each of the four conditions: (1) baseline, (2) thin liquid, (3) nectar, and (4) honey consistency. Thus, each larynx was set into vibration under each of the four conditions, and the subglottal pressure at which sustained vibration occurred was recorded. We expected that the addition of liquid would increase the aerodynamic load of the vocal folds, and thus, the subglottal pressure necessary for sustaining phonation. As Table II shows, the subglottal pressure for each of the six larynges did in fact increase under load. This difference was statistically significant with a p value of 0.00002 (see also Table III). Further, the pressure difference was related to the viscosity of the material; thus, the required pressure in the honey condition (3000 cP) was higher than the required pressure in the nectar condition (300 cP), and the pressure in the nectar condition was higher than the thin liquid (4 cP) condition.

TABLE II. PTP in cm of H₂O/phonation frequency (Hz) for baseline, “thin liquid” (4 cP), “nectar” (300 cP), and “honey”(3000 cP). Error in PTP $\pm$ 0.1%.

Viscosity	Larynx 1	Larynx 2	Larynx 3	Larynx 4	Larynx 5	Larynx 6
Baseline	6.4/278	6.9/260	6.3/266	7.2/269	7.8/274	6.1/274
4 cP	6.6/278	7/262	6.35/268	7.3/267	7.8/274	6.2/273
300 cP	6.6/282	7.1/268	6.4/269	7.35/269	8/271	6.45/272
3000 cP	8.5/283	9.0/268	8.9/268	9.1/272	10.2/272	8.0/277

B. Acoustic Experiments

It is interesting to consider what measures were consistent across conditions and larynges. Average phonation frequency (F_0) at PTP for the baseline, thin liquid, nectar, and honey conditions was consistent across the three conditions and the six larynges, falling within a range of 266–283 Hz (see also Table II). The difference between the F_0 for the different viscosity condition was statistically insignificant ($p=0.76$) at a p value of 0.05 (see also Table III) [for the baseline condition, the average F_0 was 270 Hz (range = 260–278 Hz), for the thin liquid condition, it was 270.3 Hz (range=262–278 Hz), for the nectar condition, it was 272(range=268–282 Hz), and for the honey condition, it was 273(range=268–283 Hz)]. Interestingly, this suggests that although the extra material on the vocal folds increased the subglottal pressure required to initiate phonation, it did not materially affect the rate of vocal fold vibration. Note that in a study of the effect of hydration, Ayache *et al.* (2004) found that adding viscous fluid to the surface of the vocal folds lowered the vibratory frequency. However, in the Ayache *et al.* study, as opposed to the current study, a very small amount of viscous fluid was swabbed onto the free margin of the fold. Further, the comparison condition involved vocal folds with no fluid applied. As such, the Ayache *et al.* study is less applicable than the current study to the conditions obtaining for patients with swallowing disorders. For instance, the methodology of the Ayache *et al.* study makes it possible that the fluid in question was absorbed by the dehydrated vocal folds internally, rather than resting on the surface during phonation.

Overall, as the experiments progressed, less and less materials remained on the vocal folds. Accordingly, the time trace of the subglottal pressure showed major differences between the baseline and the “viscosity” conditions. An example comparison is shown in Figs. 2(a) and 2(b), which contrast time traces of the subglottal pressure during phonation for the baseline no fluid condition [seen in Fig. 2(a)] and the highest viscosity honey condition [seen in Fig. 2(b)] for Larynx 1. As the figure shows, when we compare the baseline versus honey condition, there is a clear difference in the pressure peak-to-peak amplitude and pitch-pitch variation

between cycles, with the baseline condition exhibiting a periodic signal with nearly constant peak-to-peak amplitude, and the honey condition showing wide variation in peak-to-peak amplitude. The honey condition also shows irregular kinks/humps that vary with time [see Fig. 2(b)]. Further, the average baseline PTP for this larynx was higher in the honey condition (6.4 cm of H₂O in the baseline condition, and 8.5 cm in the honey condition), while the average peak-peak amplitude of the signal was lower (2.89 cm H₂O versus 2.1 cm H₂O, respectively). In other words, as noted above, the mean PTP increased when fluid was added, and the increase was proportional to the change in viscosity (see also Table II). In contrast, the mean peak-peak amplitude of subglottal pressure decreased proportionately as viscosity increased during phonation.

Note that in all cases, the PTP and F_0 values were consistent with those produced by normal human females using their habitual speaking voice. Figure 3 shows an FFT spectrum of the signal calculated across the entire 2-s experiment for Larynx 1, in each of the three baseline, thin liquid, nectar, and honey conditions. Distinct peaks can be observed at the average fundamental (278 Hz) and its harmonics. Consistent with our earlier observation (see also Table II), the change in average fundamental frequency between the three viscosity conditions was negligible.

Each of the effects described above are consistent with the characteristics of an unsteady dynamical system consisting of the combined vocal fold vibration and motion of fluid on the surface of the vocal folds. To further characterize the behavior of the system, we plotted the spectral density over time of the acoustic signals for each combination of condition and larynx, using CN and TC versions. Figure 4 shows an example of these measures for Larynx 1. The baseline condition is shown in Figs. 4(a) and 4(b), thin liquid condition in Figs. 4(c) and 4(d), the nectar condition in Figs. 4(e) and 4(f), and the honey condition in Figs. 4(g) and 4(h).

Both CN and TC spectrograms show distinct lines at the fundamental (first harmonic) and second harmonic ($2 \times F_0$) in the time frequency plots for all cases. The conventional spectrogram shows clear evidence of energy smearing (or broadening) at the fundamental (278 Hz) and second harmonic

TABLE III. p , F values from one-way ANOVA test for different metrics.

	PTP	Phonation frequency	% jitter	% shimmer	L	D_2
p value	0.00002 ^a	0.76	0.14	0.07	0.004 ^a	0.001 ^a
F value, $df=3, 20$	15.08	0.39	2.44	3.52	10.2	14.3

^aStatistically significant differences ($p < 0.05$).

FIG. 2. (Color online) Acoustic time traces of phonation (a) baseline, no fluid on folds, and (b) honey condition for Larynx 1.

(556 Hz) for all cases. Some resolution is added by the TC spectrograms, such that we observe the presence of secondary frequencies between the fundamental and second harmonic. The presence of higher amplitude intermittent low frequencies (lower than F_0) can also be clearly observed for the case with honey in both the conventional and TC spectrograms. Both CN and TC spectrograms show intermittent low frequencies for the load conditions, and a distinct dampening in the higher harmonics (834, 1112, 1390, 1668 Hz), but this dampening is much greater in the honey condition than in the nectar condition.

The presence of frequencies lower than the fundamental (i.e., the phonation frequency), the generation of high frequencies, the spectral broadening, and intermittent nature of these elements are classic signs of a non-linear system. Figures 5(a)–5(d) show the phase portraits of the baseline and the three viscosity conditions. The phase reconstruction is the plot of the acoustic phonation signal in delayed coordinates (Heath, 2000, Kantz and Schreiber, 2000, Sprott,

FIG. 3. Fast Fourier transform of the acoustic signal for baseline, thin liquid, nectar, and honey condition for Larynx 1.

FIG. 4. (Color online) (a,c,e,g) Time corrected and (b,d,f,h) conventional acoustic spectrogram for (a,b) baseline, no fluid on folds, (c,d) thin liquid on folds, (e,f) nectar on folds, and (g,h) honey on vocal folds for Larynx 1.

2003). For the baseline condition, the signal is periodic, and shows the presence of a well-defined, i.e., regular, closed loop trajectory [see Fig. 5(a)]. The greatest contrast is shown by the honey condition, indicating an irregular pattern confirming the characteristics observed in the CN and TC spectrogram plots [see Fig. 5(d)]. The thin liquid and nectar conditions are intermediate in nature, showing a closed loop trajectory with less complex behavior, as compared with the honey condition. Note that the baseline condition and the nectar and thin liquid conditions differ in the attractor trajectory, which is seen in the thickness of the difference between the inner and outer boundaries of the loop. This in part arises due to the large variation in amplitudes between cycles in the thin liquid and nectar condition (also seen by large shimmer values, Fig. 6).

For purposes of quantifying the behavior observed in the phase plot, we computed the mean and standard error of the

FIG. 5. (Color online) Phase portraits for (a) baseline, no fluid on folds, (b) thin liquid condition, (c) nectar condition, and (d) honey condition for Larynx 1.

correlation dimension (D_2), and the maximal Lyapunov exponent (L). In addition, we computed the conventional measures of % jitter and % shimmer. These are plotted in Fig. 6 for the six larynges, and for each of the four conditions. The vertical bars represent the standard error computed over all six larynges.

FIG. 6. (Color online) Bar plot of correlation dimension, Lyapunov exponent*10, % jitter, % shimmer, PTP, and F_0 for the baseline, thin liquid, nectar, and honey condition. The vertical bars represent the standard error computed over all six larynges for each parameter.

In all of the cases, we observed a positive value for L , which suggests that the system exhibits chaotic behavior. However, a more stringent test to distinguish a random data from that of a chaotic data is the surrogate analysis (Theiler *et al.*, 1992). This test was performed in the current study. The surrogate data has the same power spectra as the original data set. A surrogate test involves subjecting both the original and surrogate data to the same type of non-linear analysis. This test has already been used in experimental data (Grassberger, 1986, Gober *et al.*, 1992, Kurths and Herzog, 1987, Narayanan and Alwan, 1995). If the results from the surrogate data are significantly different from those of the original data, then the null hypothesis that the original data arises from a random process can be rejected. One method of generating surrogate data is to take the Fourier transform of the original data, randomize phases, and then inverse Fourier transform to obtain the surrogate data set. This method was adopted in our study. We performed this analysis on two of the fluid conditions in Larynx 1. A plot of the correlation dimension and maximal Lyapunov exponent as a function of the embedding dimension is shown in Fig. 7. It can be seen that with increasing embedding dimension, the original data sets converged to a fixed correlation dimension and

FIG. 7. Plot of correlation dimension and maximum Lyapunov exponent as a function of the embedding dimension, m for original and surrogate acoustic data for thin liquid and honey condition for Larynx 1.

TABLE IV. Range of % jitter, shimmer, correlation dimension, and maximal Lyapunov exponent for all the six larynges.

	Baseline	Thin liquid	Nectar	Honey
% jitter	0.38–0.58	2.6–4	1.5–6.7	3.8–11.39
% shimmer	3.6–6.4	17.6–18.9	14–29	9–26
Correlation dimension, D_2	2.28–2.4	3.0–3.34	3.4–3.7	3.9–4.17
Maximal Lyapunov exponent, L	0.05–0.1	0.19–0.32	0.58–0.595	0.625–0.637

Lyapunov exponent, whereas the surrogate data sets does not achieve a fixed value. The value of the Lyapunov exponent for the surrogate data sets is found to decrease with increasing embedding dimension, but never converges to a fixed value; whereas the value of correlation dimension steadily increases with increasing m . This test indicates that the surrogate data sets represent characteristics of a stochastic system. Additionally, the convergence of D_2 and L from the fluid conditions to a fixed value signifies that the data arose from a chaotic process.

Table IV shows mean and range for (1) % jitter, (2) % shimmer, (3) Lyapunov exponent, and (4) correlation dimension for the baseline and different viscosity conditions in all the six larynges. The p value and F statistic from the ANOVA test is summarized in Table III. For the non-linear metrics (correlation dimension and maximal Lyapunov exponent), p values were less than 0.05, indicating that there was a statistical difference between the baseline and three different viscosity conditions. In contrast, the classical measures showed a non-significant difference ($p > 0.05$). Furthermore, no differences were identified in the phonation frequency between the different conditions, whereas a statistically significant difference was observed in the phonation threshold pressure between the different conditions.

The above finding suggests that the classic perturbation measures of jitter and shimmer cannot reliably identify differences between the baseline and different load conditions. Note that increased jitter and shimmer are sometimes found, but do not seem to be a consistent indicator of material on the vocal folds [this finding is consistent with the study of Ramig *et al.* (1990), which showed increased shimmer in some, but not all amyotrophic lateral sclerosis patients who presented with a wet sounding voice].

A *post hoc* analysis was performed using the Tukey HSD method to identify differences between the different conditions for PTP and the two nonlinear metrics, D_2 and L . The studentized range statistic Q is given in Table I for each different comparison of the baseline and viscosity conditions. To reach statistical significance at the 0.05 level, the necessary Q value for alpha is 4.9 (degrees of freedom=6 and number of treatments=4); that is, values of $Q > 4.9$ indicate that there is a statistical difference. As Table I shows, all comparisons between baseline and viscosity conditions were significant for D_2 . In contrast, the Lyapunov exponent was unable to detect a significant difference between baseline and thin liquid conditions, but was able to identify differences between nectar and thin liquid. D_2 did not detect reliable differences between thin liquid and nectar, or nectar and honey conditions. All three measures (PTP, D_2 and L) were

able to identify differences between thin liquid and honey conditions.

It should be noted that the computation of correlation dimension and Lyapunov exponent in finite experimental data has several limitations. Some of these limitations include the presence of noise, stationarity of the signal, finite signal length, and influence/presence of auto-correlation effects. In this study, we attempted to obtain reliable estimates and validate them in the following manner. For instance, external noise perturbations were kept to a minimum during the measurements (this can be seen in Fig. 4, where the acoustic amplitude of the fundamental was at least 15 times higher the background noise level). The operational definition of stationarity used here was that proposed by Theiler (1991). In the current study, the spectral plot of the baseline and load conditions show amplitude at (frequencies $<$ fundamental), which are at least an order of magnitude lower than the dominant frequencies (see Fig. 3). Narayanan and Alwan (1995) and Herzel (1993) imposed durational constraints on voice sounds in their non-linear analysis. The rationale behind this approach was that a large number of pitch cycles could then provide stationary time series segments. In their data, a duration of as little as 100 ms was found to be sufficient for the purpose. In the current study, we computed D_2 and L in four equal segments (500 ms) over the entire 2 s phonation period. The variations between D_2 and L between segments in all cases were within 3%. Using the formula given by Eckmann and Ruelle (1992), the upper bound for D_2 for a signal length of $N=5000$ (time segment*sampling frequency=500 ms*10 000) is 7.4. For both the baseline and viscous load conditions in all the six larynges, the value of D_2 was less than 7.4. This indicates that N was long enough for reasonable estimates of D_2 . To minimize the possibility that the auto-correlation effects might result in spurious dimension estimates due to loss of scaling behavior in the correlation integral (Theiler, 1987), we employed the mutual information method for computation of τ (Narayanan and Alwan, 1995).

IV. DISCUSSION AND CONCLUSION

Overall, the results of this study document the following three characteristics of vocal fold vibration under liquid load.

- (1) Phonation with liquid material on the vocal folds shows characteristics of irregular and aperiodic phonation.
- (2) The non-linear measures tested were more sensitive to differences between viscosity conditions than were the classic measures of % jitter and % shimmer.

- (3) Differences across conditions were consistent with the prediction that the effect of increasing liquid viscosity on the vocal folds is to increase the PTP, % jitter, and % shimmer, and reduce the peak-peak amplitude of subglottal pressure.

There are a number of similarities between our results and the results of studies on abnormal phonation in pathological voice conditions involving edema, vocal folds polyps, nodules, and cysts, among others (Herzel *et al.*, 1994, Berry *et al.*, 1996, Titze, 1994a, Giovanni *et al.*, 1999a, Giovanni *et al.*, 1999b, Švec *et al.*, 1999, Jiang *et al.*, 2003, Zhang and Jiang, 2004, Zhang *et al.*, 2005, Jiang *et al.*, 2006, Zhang and Jiang, 2008). As in the Jiang *et al.* (2003) excised experiments, we found that the acoustic signal showed signs of irregular phonation, and that non-linear methods worked better than classic perturbation methods for characterizing baseline versus viscosity conditions. Similarly, Jiang *et al.* (2003), Zhang *et al.* (2005), and Zhang and Jiang (2008) found that perturbation metrics such as jitter and shimmer were less effective than non-linear measures at characterizing differences between pathological voices, or at documenting change as a result of clinical intervention. In general, jitter and shimmer as measures are dependent on accurate extraction of the pitch period. One possible explanation for the relative insensitivity of these measures to our data may be explained by the fact that the viscosity conditions showed a number of characteristics that are challenging for typical pitch extraction algorithms. These characteristics include the generation of intermittent low and high frequencies in the proximity of the phonation frequency and its harmonics.

A similar explanation may apply to the pathological voices examined by Herzel *et al.*, 1994, Berry *et al.*, 1996, Titze, 1994a, Giovanni *et al.*, 1999a, Giovanni *et al.*, 1999b, Švec *et al.*, 1999, Jiang *et al.*, 2003, Zhang and Jiang, 2004, Zhang *et al.*, 2005, Jiang *et al.*, 2006, and Zhang and Jiang, 2008.

In addition, we found that the non-linear measures were particularly sensitive to differences in viscosity. For example, both correlation dimension and the Lyapunov exponent were able to identify difference between thin liquid and honey. Additionally, the Lyapunov exponent was able to detect differences between thin liquid and nectar, whereas correlation dimension did not (it should be pointed out that neither metric was sensitive to the difference between honey versus nectar). This result suggests that non-linear analysis of voice after swallowing in patients who show signs of swallowing disorder may be useful in assessing the presence of foreign materials. It might also be useful in assessing difference between different viscosity liquids present on the vocal folds (thick versus thin mucous/secretions or food material) in patients who fail to clear secretions or patients who present with a high risk of aspiration on a certain viscosity food material.

Our finding that liquid load on the vocal folds increases the subglottal pressure necessary for phonation (PTP) is not unexpected, simply because it takes a higher pressure to move a greater mass. Additionally, the presence of liquid on the vocal folds reduces the peak to peak subglottal pressure

amplitude. We speculate that the presence of fluid on the folds dampens vibrations and mucosal waves on the surface. Higher viscosity fluids can be thought of inducing greater damping force on the vibrations, which could lead to higher PTP and lower peak-peak subglottal pressure amplitude in all the six larynges. Larger variation in subglottal pressure amplitude was also observed with the presence of liquid, which could be possibly attributed to dynamic motion of the liquid and its subsequent loss into the glottis opening. The finding of irregular and aperiodic components in the phonatory signal is also not surprising. The dampened vibrations generated due to presence of liquid on the vocal folds surface act to decrease the amplitude of sound production. In addition, the fluid may come into contact with the glottal airflow as a consequence of its motion on the surface of the vocal folds. Contact of this nature may well generate a multitude of effects related to fluid evaporation and fluid inertial behaviors (e.g., pooling on the surface of the larynx and into adjacent structures, spilling below the vocal folds, impinging on the walls of the supraglottal structures, and breaking up/atomization into smaller fluid particles). Various aeroacoustic mechanisms can cause generation of sound when the air flow is occasionally interrupted or comes in contact/impinges with the liquid during glottal opening. For example, the interaction of glottal air flow with liquid/food material on the vocal fold surface presumably will generate a fluctuating pressure, which in turn produces a new dipole source or modification of the rate of change in airflow Q at glottis, dQ/dt (Zhao *et al.*, 2002). Such effects may explain the intermittent acoustic events that were observed in the TC spectrograms [see Figs. 4(c)–4(h)].

This study focuses on only a few of the variables involved in evaluating vocal fold phonation under load. For instance, fluid mass affects the total load on the vocal fold. While in the present study, we employed a constant fluid volume for liquids of different viscosity, the behavior of different volumes is another dimension of variation, which arises commonly in clinical swallow studies. Many patients with swallowing difficulties also present with laryngeal pathologies, and it will be important to determine how these different pathologies affect the behavior of the material on the vocal folds, and their acoustic consequences. Additional work in the area of spatio-temporal vocal fold characteristics and liquid-phonatory airflow-vocal folds vibratory structure coupling will also be required to understand the complex interaction and its effect on non-linear metrics. Future work should also involve comparing these results with clinical endoscopic evaluations (flexible endoscopic swallowing study) of larynx in patients with swallowing disorders. Studies should also include comparing these data to other *ex vivo* tissue models (e.g., canine) and evaluate the effect that arises due to geometrical differences (e.g., sloping glottis).

The overall direction of these studies will be twofold. In the one case, the analysis of acoustic signals produced by vocal folds under different load conditions may provide a supplementary tool for use in the differentiation and monitoring of phonatory behavior during diagnosis and treatment of swallowing disorders. In the other case, if we can establish accurate identification of swallowed material on the vocal

folds during phonation via acoustic analysis, then this may help in estimating the threat of these conditions on lower airway protection and respiratory health.

ACKNOWLEDGMENTS

The authors would like to thank Raghav Lakhamraju for his help with experimental setup and data collection. They would also like to thank the two anonymous reviewers and the associate editor, Dr. Anders Lofquist, for taking extra care in providing valuable comments and suggestions that have enabled us to improve the quality of the article. The authors would also like to thank Kathy Groves-Wright and Bernice Klaben for inspiring this study, and Joel MacAuslan from STAR Analytical Services for providing insightful comments. They are grateful for the support received from the University of Cincinnati Research Council to conduct this work.

- Alipour, F., and Jaiswal, S. (2009). "Glottal airflow resistance in excised pig, sheep, and cow larynges," *J. Voice* **23**, 40–50.
- Alipour, F., Scherer, R. C., and Finnegan, E. (1997). "Pressure-flow relationships during phonation as a function of adduction," *J. Voice* **11**, 187–194.
- Ayache, S., Ouaknine, M., Dejonkere, P., Prindee, P., and Giovanni, A. (2004). "Experimental study of the effects of surface mucus viscosity on the glottic cycle," *J. Voice* **18**, 107–115.
- Berry, D. A., Herzel, H., Titze, I. R., and Story, B. H. (1996). "Bifurcations in excised larynx experiments," *J. Voice* **10**, 129–138.
- Dawson, B., and Trapp, R. G. (2004). *Basic and Clinical Statistics*, 4th ed. (McGraw-Hill, New York).
- Eckmann, J. P., and Ruelle, D. (1992). "Fundamental limitations for estimating dimensions and Lyapunov exponents in dynamic systems," *Physica D* **56**, 185–187.
- Fraser, A. M., and Swinney, H. L. (1986). "Independent coordinates for strange attractors from mutual information," *Phys. Rev. A* **33**, 1134–1140.
- Fulop, S. A., and Fitz, K. (2006). "Algorithms for computing the time-corrected instantaneous frequency (reassigned) spectrogram, with applications," *J. Acoust. Soc. Am.* **119**, 360–371.
- Giovanni, A., Ouaknine, M., Guelfucci, B., Yu, P., Zanaret, M., and Triglia, J.-M. (1999a). "Nonlinear behavior of vocal folds vibrations: The role of coupling between vocal folds," *J. Voice* **13**, 465–476.
- Giovanni, A., Ouaknine, M., and Triglia, J.-M. (1999b). "Determination of largest Lyapunov exponent of the vocal signal: Application to unilateral laryngeal analysis," *J. Voice* **13**, 341–354.
- Gober, M., Herzel, H., and Graf, H.-F. (1992). "Dimension analysis of El Nino/southern oscillation time series," *Ann. Geophys.* **10**, 729–734.
- Grassberger, P. (1986). "Do climatic attractors exist?," *Nature (London)* **323**, 609–612.
- Groves-Wright, K. (2007). "Acoustics and perception of wet vocal quality in Identifying penetration/aspiration during swallowing," Ph.D. thesis, Communication Sciences and Disorders, University of Cincinnati, Ohio.
- Heath, R. (2000). *Nonlinear Dynamics: Techniques and Applications in Psychology* (Lawrence Erlbaum Associates, Publishers, Mahwah, New Jersey).
- Herzel, H. (1993). "Bifurcations and chaos in voice signals," *Appl. Mech. Rev.* **46**, 399–413.
- Herzel, H., Berry, D., Titze, I. R., and Saleh, M. (1994). "Analysis of vocal disorders with methods from nonlinear dynamics," *J. Speech Hear. Res.* **37**, 1008–1019.
- Jiang, J., Zhang, Y., and Ford, C. N. (2003). "Non-linear dynamics of phonation in excised larynx experiments," *J. Acoust. Soc. Am.* **114**, 2198–2205.
- Jiang, J., Zhang, Y., and McGilligan, C. (2006). "Chaos in voice, from modeling to measurement," *J. Voice* **20**, 2–17.
- Jiang, J. J., Raviv, J. R., and Hanson, D. G. (2001). "Comparison of the phonation related structures among pig, dog, white-tailed deer, and human larynges," *Ann. Otol. Rhinol. Laryngol.* **110**, 1120–1125.
- Kantz, H., and Schreiber, T. (2000). *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, England).
- Khosla, S., Murugappan, S., Gutmark, E., and Scherer, R. (2007). "Vortical flow field during phonation in an excised canine larynx model," *Ann. Otol. Rhinol. Laryngol.* **116**, 217–228.
- Klingholtz, F. (1990). "Acoustic recognition of voice disorders: A comparative study of running speech versus sustained vowels," *J. Acoust. Soc. Am.* **87**, 2218–2224.
- Kreiman, J., Gerratt, J. B., Kempster, G. B., Erman, A., and Berke, G., (1993). "Perceptual evaluation of voice quality review, tutorial, and a framework for future research," *J. Speech Hear. Res.* **36**, 21–40.
- Kurths, J., and Herzel, H. (1987). "An attractor in solar time series," *Physica D* **25**, 165–172.
- Logemann, J. A. (1993). *Manual for the Videofluorographic Study of Swallowing*, 2nd ed. (Pro:Ed, Austin, Texas).
- Logemann, J. A. (1998). *Evaluation and Treatment of Swallowing Disorders*, 2nd ed. (Pro:Ed, Austin, Texas).
- Lucero, J. C., and Koenig, L. L. (2005). "Phonation thresholds as a function of laryngeal size in a two-mass model of the vocal folds," *J. Acoust. Soc. Am.* **118**, 2798–2801.
- Lucero, J. C., and Koenig, L. L. (2007). "On the relation between the phonation threshold lung pressure and the oscillation frequency of the vocal folds," *J. Acoust. Soc. Am.* **121**, 3280–3283.
- MacAuslan, J., Chenausky, K., Juda, J., and Manev, I. (1997). "Identifying laryngeal attractor changes from periodicity diagrams," Proceedings of the Nonlinear Signal and Image Processing Conference, Mackinac, MI.
- Mende, W., Herzel, H., and Wermke, K., (1990). "Bifurcations and chaos in newborn cries," *Phys. Lett.* **145**, 418–424.
- Milenkovic, P. (1987). "Least mean square measures of voice perturbation," *J. Speech Hear. Res.* **30**, 529–538.
- Milenkovic, P. H. (2006). TF32Voice Reference Manual, Department of Electrical and Computer Engineering, University of Wisconsin-Madison, <http://userpages.chorus.net/cspeech/> (Last viewed May, 2009).
- Murray, J., Langmore, S. E., Ginsberg, S., and Dostie, A. (1996). "The significance of oropharyngeal secretions and swallowing frequency in predicting aspiration," *Dysphagia* **11**, 99–103.
- Narayanan, S., and Alwan, A. (1995). "A nonlinear dynamical systems analysis of fricative consonants," *J. Acoust. Soc. Am.* **97**, 2511–2524.
- Ramig, L. O., Scherer, R. C., Klasner, E. R., Titze, I. R., and Horii, Y. (1990). "Acoustic analysis of voice in amyotrophic lateral sclerosis: A longitudinal case study," *J. Speech Hear. Disord.* **55**, 2–14.
- Ryu, J. S., Park, S. R., and Choi, K. H. (2004). "Prediction of laryngeal aspiration using voice analysis," *Am. J. Phys. Med. Rehabil.* **83**, 753–757.
- Sprott, J. C. (2003). *Chaos and Time Series Analysis* (Oxford University Press, New York).
- Steele, C. M., and Van Lieshout, P. H. H. M. (2005). "Does barium influence tongue behaviors during swallowing?," *Am. J. Speech Lang. Pathol.* **14**, 27–39.
- Steinecke, I., and Herzel, H. (1995). "Bifurcations in an asymmetric vocal fold model," *J. Acoust. Soc. Am.* **97**, 1874–1884.
- Stoline, M. R. (1981). "The status of multiple comparisons: Simultaneous estimation of all pairwise comparisons in one-way ANOVA designs," *Am. Stat.* **35**, 134–141.
- Švec, J. G., Schutte, H. K., and Miller, D. G. (1999). "On pitch jumps between chest and falsetto registers in voice: Data from living and excised human larynges," *J. Acoust. Soc. Am.* **106**, 1523–1531.
- Takens, F., (1988). "Limit capacity and Hausdorff dimension of dynamically defined cantor sets," *Dynamical Systems* (Springer, Berlin), Vol. **1331**, pp. 196–212.
- Tao, C., and Jiang, J. J. (2008). "The phonation critical condition in rectangular glottis with wide prephonatory gaps," *J. Acoust. Soc. Am.* **123**, 1637–1641.
- Theiler, J. (1986). "Spurious dimension from correlation algorithms applied to limited time series data," *Phys. Rev. A* **34**, 2427–2432.
- Theiler, J. (1987). "Efficient algorithm for estimating the correlation dimension from a set of discrete points," *Phys. Rev. A* **36**, 4456–4462.
- Theiler, J. (1990). "Estimating fractal dimension," *J. Opt. Soc. Am. A* **7**, 1055–1073.
- Theiler, J. (1991). "Some comments on correlation dimension of $1/f^\alpha$ noise," *Phys. Lett.* **155**, 480–493.
- Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., and Farmer, J. D. (1992). "Testing for nonlinearity in time series: The method of surrogate data," *Physica D* **58**, 77–94.
- Titze, I. R. (1988). "The physics of small-amplitude oscillation of the vocal folds," *J. Acoust. Soc. Am.* **83**, 1536–1552.
- Titze, I. R. (1994a). "Summary statement," Proceedings of the Workshop on

- Acoustic Voice Analysis, NCVS, Denver, CO, Feb. 17 and 18.
- Titze, I. R. (1994b). *Principles of Voice Production* (Prentice-Hall, Englewood Cliffs, NJ).
- Vieira, M. N., McInnes, F. R., and Jack, M. A. (2002). "On the influence of laryngeal pathologies on acoustic and electroglottographic jitter measures," *J. Acoust. Soc. Am.* **111**, 1045–1055.
- Warms, T., and Richards, J. (2000). "Wet voice as a predictor of penetration and aspiration in oropharyngeal dysphagia," *Dysphagia* **15**, 84–88.
- Wolf, A., Swift, J. B., Swinney, H. L., and Vastano, J. A. (1985). "Determining Lyapunov exponents from a time series," *Physica D*, **16**, 285–317.
- Zhang, Y., and Jiang, J. (2004). "Chaotic vibrations of a vocal folds model with a unilateral polyp," *J. Acoust. Soc. Am.* **115**, 1266–1269.
- Zhang, Y., and Jiang, J. (2008). "Acoustic analysis of sustained and running voices with laryngeal pathologies," *J. Voice* **22**, 1–9.
- Zhang, Y., Jiang, J., and Rahn, D., III (2005). "Studying vocal folds vibrations in Parkinson's disease with a nonlinear model," *Chaos* **15**, 033903.
- Zhao, W., Zhang, C., Frankel, S. H., and Mongeau, L. (2002). "Computational aeroacoustics of phonation, Part 1: Computational methods and sound generation mechanisms," *J. Acoust. Soc. Am.* **112**, 2134–2146.