Updating signal typing in voice: Addition of type 4 signals

Alicia Sprecher, Aleksandra Olszewski, and Jack J. Jiang

Department of Surgery, Division of Otolaryngology–Head and Neck Surgery, University of Wisconsin School of Medicine and Public Health, Madison, WI 53792

Yu Zhang^{a)}

Key Laboratory of Underwater Acoustic Communication and Marine Information, Xiamen University, Xiamen, Fujian 361005, People's Republic of China

(Received 11 September 2009; revised 22 March 2010; accepted 25 March 2010)

The addition of a fourth type of voice to Titze's voice classification scheme is proposed. This fourth voice type is characterized by primarily stochastic noise behavior and is therefore unsuitable for both perturbation and correlation dimension analysis. Forty voice samples were classified into the proposed four types using narrowband spectrograms. Acoustic, perceptual, and correlation dimension analyses were completed for all voice samples. Perturbation measures tended to increase with voice type. Based on reliability cutoffs, the type 1 and type 2 voices were considered suitable for perturbation analysis. Measures of unreliability were higher for type 3 and 4 voices. Correlation dimension analyses increased significantly with signal type as indicated by a one-way analysis of variance. Notably, correlation dimension analysis could not quantify the type 4 voices. The proposed fourth voice type represents a subset of voices dominated by noise behavior. Current measures capable of evaluating type 4 voices provide only qualitative data (spectrograms, perceptual analysis, and an infinite correlation dimension). Type 4 voices are highly complex and the development of objective measures capable of analyzing these voices remains a topic of future investigation.

© 2010 Acoustical Society of America. [DOI: 10.1121/1.3397477]

PACS number(s): 43.70.Jt, 43.70.Dn, 43.70.Ep, 43.25.Ts [AL]

Pages: 3710–3716

I. INTRODUCTION

In a 1995 summary statement from the National Center for Voice and Speech Workshop on acoustic voice analysis, Titze (1995) proposed the classification of voices into three types. In his system, he defined type 1 voices as nearly periodic and suggested that only this type of voice was suitable for perturbation analysis. Type 2 voices were defined as containing strong modulations or subharmonics that approached the fundamental frequency in energy. Finally, type 3 voices were those that were irregular or aperiodic in nature. Titze (1995) recommended spectrographic and perceptual analysis methods for type 2 and type 3 voices. Today, Titze's (1995) recommendations continue to be employed to determine the suitability of voice signals for perturbation analysis (Bielamowicz *et al.*, 1996; Vieira *et al.*, 2002; Shaw and Deliyski, 2008; Zhang and Jiang, 2008). The task of analyzing type 2 and 3 signals has fallen to spectrograms, perceptual analysis, and, more recently, nonlinear dynamic analysis. Further investigation has identified a wide range of samples classified as type 3. While some type 3 signals have a finite dimension, some are more stochastic in nature with infinite dimensionality (Little *et al.*, 2007). Because the presence of stochastic behavior has a major impact on the ability to complete such nonlinear analyses as correlation dimension (D_2), we propose a change in Titze's (1995) classification scheme where type 3 voices that are stochastic in nature are distinguished

as type 4 voices (Zhang *et al.*, 2005a; Little *et al.*, 2007; Shaw and Deliyski, 2008; Zhang and Jiang, 2008).

In recent years, nonlinear dynamics measurement has received considerable attention as a system of voice analysis. Research indicates several nonlinear attributes of voice, including nonlinear pressure and flow in the glottis, stress-strain curves of vocal fold tissue, and behavior of vocal fold collisions (Herzel *et al.*, 1995). Furthermore, studies have demonstrated low dimensionality even in strongly periodic voices (Tao and Jiang, 2008). Unlike perturbation analysis which is dependent on the accurate determination of fundamental frequency, nonlinear analysis can be completed without establishing a fundamental frequency (Ma and Yiu, 2005; Zhang *et al.*, 2005b). Accordingly, nonlinear dynamic analyses, such as correlation dimension (D_2), have been shown to be reliable for use on all three of Titze's (1995) voice types (Zhang *et al.*, 2005b; Little *et al.*, 2007; Zhang and Jiang, 2008). Despite the improvements afforded by nonlinear analysis, many of its parameters are negatively impacted by infinite dimensionality (Little *et al.*, 2007). This behavior is commonly interpreted as phonation breathiness and is especially common in disordered voices (Herzel *et al.*, 1994; Vieira *et al.*, 2002; Zhang *et al.*, 2005a, 2005b). Because of the turbulence surrounding the airflow jet from which voice is produced, all voices are associated with noise that has an infinite dimension (Jiang and Zhang, 2002; Krane, 2005). In most voices, D_2 analysis is possible because low-dimension characteristics dominate. However, in some voices, stochastic characteristics override low-dimensional patterns, obscuring periodicities and preventing analysis (Jiang *et al.*, 2003; Zhang *et al.*, 2004; Jiang *et al.*, 2006; Little *et al.*, 2007).

^{a)}Author to whom correspondence should be addressed. Electronic mail: yuzhang@xmu.edu.cn

TABLE I. Subject Characteristics for each signal type group. Displayed are the mean age, age range, and gender ratio.

Signal type	Age	Gender (M:F)
Type 1	32.6(22–44)	1:9
Type 2	51.8(26–73)	6:4
Type 3	55.5(16–80)	3:7
Type 4	54.6(25–85)	7:3

In voices containing high levels of stochastic noise characteristics, neither perturbation nor D_2 analysis methods will generate reliable results. Several studies have attempted to separate broadband noise from low-dimension chaos and others have noted the value of such a division (Behnman *et al.*, 1998; Goode *et al.*, 2001; Zhang *et al.*, 2004; Little *et al.*, 2007). As the application of many current nonlinear dynamic analysis methods is severely limited by stochastic noise, the identification and exclusion of these voices from such methods is necessary. Accordingly, we propose the addition of a fourth type of voice to Titze's (1995) classification system. We define a type 4 voice as one predominated by stochastic noise characteristics. To better describe this fourth voice type we categorized 40 human voice samples as type 1, 2, 3, or 4 using spectrograms. Using these groups we compared the results of perturbation, perceptual, and correlation dimension analyses.

II. METHODS

A. Voice selection

Approximately 150 sustained phonations were collected from the Disordered Voice Database Model 4337 (KayPENTAX, Lincoln Park, NJ). The samples were approximately 1 s sections from the interior portion of recorded phonations. Summary characteristics for the final 40 samples were recorded and are reported in Table I.

B. Spectrogram analysis

Narrowband spectrograms were generated for approximately 100 randomly selected abnormal samples and 50 randomly selected normal samples using the PRAAT software version 5.1.02 (Boersma and Weenink, 2009). A window length of 50 ms, a time step of 0.002 s, a frequency step of 5 Hz, and a dynamic range of 40 dB were employed. A Hamming window shape was used to create the spectrogram. Voice typing was completed by one author (A. Sprecher) with consultation from the remaining authors. Classification was based on the recommendations outlined by Titze (1995) with some modifications. After all signals were typed, ten samples were chosen as the most representative of each voice type.

Type 1 signals were periodic without strong modulations or subharmonics. The spectrograms for type 1 signals showed multiple clearly defined harmonics, although the number and spacing of harmonics varied between samples. Harmonics were nearly straight. A small amount of noise

was permitted between the harmonics provided their strength was greatly reduced in comparison with the intensity of the harmonic.

Type 2 signals contained strong modulations, bifurcations, or subharmonics while the signal was generally periodic in nature. The spectrograms for these samples contained multiple harmonics and the inter-harmonic noise formed clearly defined subharmonics with intensities nearing the strength of the fundamental frequency. These subharmonics often appeared intermittently in the 1 s segment; however, even the very brief presence of a subharmonic resulted in a type 2 classification. Rarely, harmonics showed evidence of bifurcation which led to a type 2 classification; however, bifurcations generally preceded transient periods of chaos leading to a type 3 classification. Signals with strong modulations in the harmonics, where the harmonics appeared wavy, were also classified as type 2; however, these samples were not included as they were not very common and the cutoff for strong modulations is somewhat subjective.

Titze (1995) defined type 3 signals as aperiodic; for this study we define type 3 signals as chaotic with a finite dimension. The spectrograms of type 3 signals show a smearing of energy across harmonics. While the fundamental frequency and even one or two harmonics were visible in many of the signals classified as type 3, higher frequency harmonics were obscured by diffuse energy. The type 3 signals showed band limited characteristics, with the energy concentrated at lower frequencies (generally below 1500 Hz).

Our newly defined type 4 signals had an infinite dimension. The spectrograms of these signals could exhibit a fundamental frequency, although many did not. In contrast to the type 3 signals, type 4 signals showed a smearing of energy across a broader range of frequencies, as is characteristic of broadband white noise.

C. Perturbation analysis

Perturbation analysis was conducted using the TF32 software (Milenkovic, 2001, Madison, WI). Measures of percent jitter, percent shimmer, and signal-to-noise ratio (SNR) were calculated for each signal. Jitter quantifies the cycle-to-cycle variation in the signal fundamental frequency while shimmer measures the cycle-to-cycle variation in signal amplitude (Milenkovic, 1987). SNR is a measure of the relative dominance of the harmonic signal over turbulent noise (Milenkovic, 1987). The reliability of these measures was assessed using TF32 values of "trk," which quantifies the number of dramatic fluctuations in pitch and "err," which indicates discrepancies in the calculated fundamental frequency likely due to voice breaks present in the sample (Milenkovic, 2001). Although there is no established cutoff for these measures, we generally assume an err greater than 10 indicates a sample that is ill-suited for perturbation analysis. trk and err were used to indicate our level of confidence with the perturbation results.

D. Correlation dimension analysis

D_2 is commonly applied to describe periodic, strongly aperiodic, and chaotic voices, particularly in conjunction

FIG. 1. Waveforms of four of the voice samples used in the analysis. Panels (A), (B), (C), and (D) are samples of type 1, 2, 3, and 4 voices, respectively. The waveforms in panels (A) and (D) are from female voices while those in panels (B) and (C) are from male voices.

with perturbation analysis. Detailed descriptions of this method can be found in numerous publications and therefore will be described only briefly here (Jiang and Zhang, 2002; Zhang *et al.*, 2005a, 2005b; Zhang and Jiang, 2008).

D_2 was used to quantify the nonlinearity of the voice signals in this study. This measure specifies the number of degrees of freedom necessary to describe a system where a system with greater complexity requires more degrees of freedom to characterize its dynamic state (Jiang *et al.*, 2006). The algorithms employed in this study to calculate D_2 have previously been applied successfully to excised larynx pho-

nations (Jiang *et al.*, 2003, 2006) and human voices (Titze *et al.*, 1993; Jiang and Zhang, 2002; Zhang *et al.*, 2004; Zhang and Jiang, 2008). Briefly, the time delay technique was used to create a phase space: $\mathbf{X}_i = \{x(t_i), x(t_i - \tau), \dots, x(t_i - (m - 1)\tau)\}$, where m is the embedding dimension and τ is the time delay (Packard *et al.*, 1980). Takens' (1981) embedding theorem was used to define m , such that when $m > 2D + 1$ (D is the Hausdorff dimension), the reconstructed phase space and the original phase space are topographically equivalent. The time delay was determined according to the mutual information method as proposed by Fraser and Swinney (1986). A correlation integral $C(r)$ was calculated using the improved algorithm proposed by Theiler (1986), where r is the radius around $\mathbf{X}_i$. $C(r)$ uses the distance between points in the reconstructed phase space and determines the number of these distances that are less than the radius r . The function exhibits power law behavior as described by $C(r) \propto r^{D_2} e^{-m\tau K_2}$, which reveals the geometric scaling property of the attractor (Grassberger and Procaccia, 1983). Using r to define the scaling region, curves of $\log_2 C(r)$ versus $\log_2 r$ were generated for each embedding dimension m . The value of D_2 was calculated at the point where these curves converged. A standard deviation of D_2 was computed and remained less than 5% to ensure reliable calculations.

FIG. 2. Spectrograms generated from voice data. Panels (A), (B), (C), and (D) were classified as type 1, 2, 3, and 4, respectively. The spectrograms in panels (A) and (D) are from female voices while those in panels (B) and (C) are from male voices.

FIG. 3. Results of nonlinear and perturbation analysis. Error bars represent standard error. None of the type 4 voices converged using nonlinear analysis. Each group contains ten voice samples. A one-way ANOVA on ranks detected significance for D_2 ($H=11.435$, 2 d.f., $p=0.003$).

E. Perceptual analysis

All of the voices were subjectively rated. Each voice sample was independently rated by three doctors with experience in otolaryngology and speech science. Raters were blinded to the voice type and pathology. Voice samples were randomly assorted into groups of ten. Each listener was permitted to listen to a whole group before being asked to rate the individual samples. Raters could replay a sample as many times as necessary to determine a rating. A six-point scale was used to evaluate overall grade and breathiness. Each voice received a score between 0 and 5 where 0 =normal, 1=slight, 2=mild, 3=moderate, 4=moderate/severe, and 5=severe. The three raters' scores were averaged and an overall mean score for each voice type was determined. According to a Spearman rank order correlation test, the scores had a correlation coefficient of 0.867 and $p < 0.001$.

F. Statistical analysis

Correlation dimension, perturbation, and perceptual data were analyzed using SIGMASTAT 11.0 (Systat Software, San Jose, CA). The voice type groups were compared using a one-way analysis of variance (ANOVA) on ranks. *Post-hoc* Dunn's tests were used for pairwise comparisons. A significance level of $p < 0.05$ was used throughout. *k*-means clustering was completed with MATLAB (Mathworks, Natick, MA) and used to define the four groups based on percent jitter, shimmer, SNR, err, trk, grade, and breathiness. D_2 was not included because not all voices produced a D_2 value. These groupings were compared with the spectrogram-based signal typing using individual Mann-Whitney rank sum tests.

III. RESULTS

Based on the described spectrogram patterns we selected ten signal of each type. Shown in Figs. 1 and 2 are sample waveforms and spectrograms of the four types of voice samples. Evident in the waveforms is an increasing complexity of the signal for type 3 and 4 voices. The spectrogram of the type 1 sample shows clearly defined harmonics and a fundamental frequency of approximately 200 Hz. In contrast

FIG. 4. Mean and standard error for the unreliability measures. Each group contained ten samples. A one-way ANOVA on ranks detected significance for both trk ($H=30.116$, 3 d.f., $p < 0.001$) and err ($H=32.484$, 3 d.f., $p < 0.001$).

to the type 2 sample, there is no evidence of subharmonics. In the type 2 sample, a fundamental frequency of approximately 140 Hz is present along with a $F/3$ subharmonic as indicated in Fig. 2(B). A fundamental frequency is visible in the type 3 sample; however, most of the harmonics are obscured by low frequency chaos. Finally, the type 4 sample is characterized by diffuse energy spanning the range of frequencies displayed.

Perturbation analysis indicated an increasing level of aperiodicity (Fig. 3), across types. Both percent jitter and percent shimmer increased with each voice type. Similarly, SNR decreased from type 1 through type 4 voices, indicating that the evidence of harmonics decreased as signal type increased. Unreliability measures are displayed in Fig. 4. The one-way ANOVA found significance for both trk and err indicating fundamental frequency variability ($H=30.116$, 3 d.f., $p < 0.001$ and $H=32.484$, 3 Degrees of Freedom (d.f.), $p < 0.001$, respectively). Pairwise comparisons for both trk and err grouped type 1 and type 2 signals together and showed significant differences between these and the group of type 3 and 4 samples ($p < 0.05$). Differences in trk and err between the type 3 and 4 signals were not significant ($p > 0.05$). Typically, an err value of less than 10 is used as the

FIG. 5. Mean perceptual ratings for each voice type. Mean overall score and mean breathiness score are given for each type with error bars representing standard error ($n=10$). An ANOVA on ranks detected significance for both scores ($H=30.767$, 3 d.f., $p < 0.001$ for overall grade and $H=31.172$, 3 d.f., $p < 0.001$ for breathiness). Pairwise comparisons found significance for type 1 and type 4, type 2 and type 4, and type 1 and type 3.

FIG. 6. Mean perturbation measures for groups as classified by spectrogram typing and by cluster analysis. Each spectrogram group contained 10 samples, while cluster groups contained 9, 10, 13, and 8 samples for type 1–4 groups, respectively. Error bars represent standard error.

cutoff point for suitability of perturbation analysis (Milenkovic, 2001). Using this cutoff point, only the type 1 and type 2 signals were appropriate for perturbation analysis. Perturbation measures were not compared statistically, because the unreliability measures suggested values for type 3 and 4 voices were not reliable.

D_2 results corresponded well with perturbation analysis results. The D_2 value varied significantly between voice types with an increasing trend, with the notable exception of the type 4 samples in which the curves did not converge and D_2 could not be calculated (Fig. 3, $H=11.435$, 2 d.f., $p=0.003$). The standard deviation of D_2 increased steadily with each voice type indicating increasing difficulty determining this value. The one-way ANOVA on ranks detected significant difference for the standard deviation of D_2 ($H=7.101$, 2 d.f., $p=0.029$). Dunn's tests revealed the only significant differences as between the type 1 and 3 samples for D_2 and its standard deviation ($p<0.05$).

Results of perceptual analysis are given in Fig. 5. Perceptual ratings of overall grade and breathiness increased with each signal type. An ANOVA on ranks detected significance for both overall grade and breathiness ($H=30.767$, 3 d.f., $p<0.001$ and $H=31.172$, 3 d.f., $p<0.001$, respectively). Pairwise comparisons for both measures detected significant differences between type 1 and type 4, type 2 and type 4, and type 1 and type 3 voices ($p<0.05$).

As shown in Figs. 6–8, the groups defined by the k -means clustering did not differ substantially from the signal typing groups determined through spectrogram analysis. Only 6 of the 40 samples were classified into groups inconsistent with their signal type. In comparing the means found by the perturbation and perceptual measures, the means of groups defined by clustering did not differ significantly from

those of groups defined by our signal typing procedures.

IV. DISCUSSION

In this study we define a new voice signal type. The newly proposed type 4 voice signal is characterized by a predominance of infinite dimensionality, making it unsuitable for both perturbation and correlation dimension analysis. Voice samples were classified on the basis of distinctive spectrogram patterns. Next we applied perturbation, perceptual, and correlation dimension analysis to all voice samples. All measures indicated increasing disorder from the type 1 voices through the type 4 voices.

Our perturbation analysis suggests that type 1 and potentially type 2 voices were suitable for perturbation analysis. Although high err values in type 3 and 4 voices led us to discard these parameters from statistical analysis, perturbation parameters suggesting worsening voice as voice type increased. Type 1 and type 2 voices produced err values below the cutoff of 10 and maintained similarly low values for trk. Both trk and err increased significantly in type 3 and type 4 voices. Although our group does not generally specify a cutoff value for trk, high values indicate numerous variations in pitch and increased likelihood of errors when computing the fundamental frequency (F_0) (Milenkovic, 2001). Perturbation measures detect variations in F_0 ; consequently, an accurate determination of F_0 is essential to perturbation analysis (Ma and Yiu, 2005; Little et al., 2007). The samples in which pitch cannot be reliably estimated or in cases where the pitch jumps as a result of bifurcations may therefore be unsuitable for perturbation analysis. Using the TF32 pitch tracking indicators, perturbation analysis was valid for the

FIG. 7. Mean unreliability measures as classified by spectrogram typing and by cluster analysis. Each spectrogram group contained 10 samples, while cluster groups contained 9, 10, 13, and 8 samples for type 1–4 groups, respectively. Error bars represent standard error.

FIG. 8. Mean perceptual measures as classified by spectrogram typing and by cluster analysis. Each spectrogram group contained 10 samples, while cluster groups contained 9, 10, 13, and 8 samples for type 1–4 groups, respectively. Error bars represent standard error.

type 1 and some of the type 2 samples. Perturbation analysis was not valid for the type 3 and 4 samples.

As expected, correlation dimension analysis was able to generate results for type 1, 2, and 3 voices, once the noisy voice samples were categorized as a separate type 4 group. The calculation of D_2 does not require a determination of fundamental frequency; therefore, it is unaffected by modulations in pitch or tracking errors (Zhang *et al.*, 2005b). D_2 increased with each signal type indicating increasing system complexity (Jiang *et al.*, 2006; Zhang and Jiang, 2008). Notably, the type 4 voices could not be quantified using correlation dimension. This result indicates the presence of stochastic noise behavior (Jiang *et al.*, 2003). Combining our correlation dimension results with the spectrogram information clearly indicates a predominance of noise in the type 4 samples.

Perceptual measurements concurred with the results of both perturbation and correlation dimension analyses. Audible voice disorder increased with each voice type as did the degree of breathiness. As seen in previous studies, the presence of phonation breathiness precluded complexity computation by correlation dimension analysis.

As seen in the cluster analysis, the voice signals split naturally into the four groups we defined. Using only the perturbation and perceptual measurements, the data split into our previously defined voice type groups in 34 out of 40 cases. Moreover, the mean values for each of the parameters used were not significantly different ($p > 0.05$) from those seen in the spectrogram-defined voice type groups (Figs. 6–8).

In this investigation we selected voice samples based on their types in order to present the concept of a type 4 signal. In future studies we plan to investigate possible associations between voice pathology and voice type such that clinicians can more accurately use voice type to facilitate diagnoses. Determining a general and accurate signal typing system with the newly proposed four types will allow better selection of analysis methods to reduce error in voice evaluation.

As we have demonstrated, type 4 voices appear unsuitable for both perturbation and correlation dimension analyses. Commonly used nonlinear methods, such as correlation dimension or Lyapunov exponents, require that low-dimensional behavior of the sample not be obscured by noise with an infinite dimension (Little *et al.*, 2007). Some measures can establish that the necessary conditions for chaos

are met with a greater resistance to noise contamination; however, these measures confirm only the potential for chaos, not its presence (Poon and Barahona, 2001). Spectrographic and perceptual analyses remain valid for these voice types; however, they are unable to produce objective quantitative data (Shrivastav and Sapienza, 2003). Although an infinite D_2 provides information about a sample's predominantly infinite dimensional nature, such data are also qualitative. Future research should focus on describing noisy voices using objective measures.

V. CONCLUSION

We defined a fourth type of voice, predominated by stochastic noise behavior and unsuitable for perturbation and correlation dimension analysis methods. Using voice samples representative of each voice type, we determined that perturbation analysis could be applied to type 1 signals. Based on pitch tracking measures, some type 2 voices, those without bifurcations, could also be analyzed using perturbation. The large number of type 2 signals that could be analyzed with perturbation is inconsistent with Titze's (1995) recommendations and may be attributed to our specific signal typing procedures. Future investigations using type 2 voices should consider pitch tracking measures before including these voices. Infinite dimension, type 4 signals could not be quantified using correlation dimension. As expected, high typed voices were perceived as increasingly dysphonic and breathy. Future research will determine which disorders are most likely to produce type 4 signals and will endeavor to determine the best way to evaluate such voices.

ACKNOWLEDGMENTS

This study was supported by NIH Grant No. RO1 DC006019 from the National Institute of Deafness and other Communication Disorders. The authors would like to thank Fan Zhang, Ling Ying Chai, and Grace Choi for their assistance in completing perceptual analysis of the samples.

- Behrman, A., Agresti, C., Blumstein, E., and Lee, N. (1998). "Microphone and electroglottographic data from dysphonic patients: Type 1, 2 and 3 signals," *J. Voice* **12**, 249–260.
- Bielamowicz, S., Kreiman, J., Gerratt, B. R., Dauer, M. S., and Berke, G. S. (1996). "Comparison of voice analysis systems for perturbation measurement," *J. Speech Hear. Res.* **39**, 126–134.
- Boersma, P., and Weenink, D. (2009). Praat: Doing phonetics by computer

- (Version 5.1.02) [Computer program], retrieved March 9, 2009, from <http://www.praat.org>, as recommended by the creators.
- Fraser, A. M., and Swinney, H. L. (1986). "Independent coordinates for strange attractors from mutual information," *Phys. Rev. A* **33**, 1134–1140.
- Goode, B., Cary, J., Doxas, I., and Horton, W. (2001). "Differentiating between colored random noise and deterministic chaos with the root mean squared deviation," *J. Geophys. Res.* **106**, 21277–21288.
- Grassberger, P., and Procaccia, I. (1983). "Measuring the strangeness of strange attractors," *Physica D* **9**, 189–208.
- Herzel, H., Berry, D., Titze, I., and Steinecke, I. (1995). "Nonlinear dynamics of the voice: Signal analysis and biomechanical modeling," *Chaos* **5**, 30–34.
- 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. J., and Zhang, Y. (2002). "Chaotic vibration induced by turbulent noise in a two-mass model of vocal folds," *J. Acoust. Soc. Am.* **112**, 2127–2133.
- Jiang, J. J., Zhang, Y., and Ford, C. N. (2003). "Nonlinear dynamics of phonations in excised larynx experiments," *J. Acoust. Soc. Am.* **114**, 2198–2205.
- Jiang, J. J., Zhang, Y., and McGilligan, C. (2006). "Chaos in voice, from modeling to measurement," *J. Voice* **20**, 2–17.
- Krane, M. H. (2005). "Aeroacoustic production of low-frequency unvoiced speech sounds," *J. Acoust. Soc. Am.* **118**, 410–427.
- Little, M. A., McSharry, P. E., Roberts, S. J., Costello, D. A., and Moroz, I. M. (2007). "Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection," *Biomed. Eng. Online* **6**, 23.
- Ma, E. P., and Yiu, E. M. (2005). "Suitability of acoustic perturbation measures in analysing periodic and nearly periodic voice signals," *Folia Phoniatr Logop* **57**, 38–47.
- Milenkovic, P. (1987). "Least mean square measures of voice perturbation," *J. Speech Hear. Res.* **30**, 529–538.
- Milenkovic, P. (2001). *TF32 User's Manual*, Madison, WI.
- Packard, N. H., Crutchfield, J. P., Farmer, J. D., and Shaw, R. S. (1980). "Geometry from a time-series," *Phys. Rev. Lett.* **45**, 712–716.
- Poon, C. S., and Barahona, M. (2001). "Titration of chaos with added noise," *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7107–7112.
- Shaw, H. S., and Deliyski, D. D. (2008). "Mucosal wave: A normophonic study across visualization techniques," *J. Voice* **22**, 23–33.
- Shrivastav, R., and Sapienza, C. M. (2003). "Objective measures of breathy voice quality obtained using an auditory model," *J. Acoust. Soc. Am.* **114**, 2217–2224.
- Takens, F. (1981). "Detecting strange attractors in turbulence," in *Lecture Notes in Mathematics*, edited by D. Rand and L. Young (Springer-Verlag, Berlin), pp. 366–381.
- Tao, C., and Jiang, J. J. (2008). "Chaotic component obscured by strong periodicity in voice production system," *Phys. Rev. E* **77**, 061922.
- Theiler, J. (1986). "Spurious dimension from correlation algorithms applied to limited time-series data," *Phys. Rev. A* **34**, 2427–2432.
- Titze, I. (1995). "Workshop on acoustic voice analysis: Summary statement," National Center for Voice and Speech, Denver, CO.
- Titze, I., Baken, R., and Herzel, H. (1993). "Evidence of chaos in vocal fold vibration," in *Vocal Fold Physiology: Frontiers in Basic Science*, edited by I. Titze (Singular, San Diego, CA), pp. 143–188.
- 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.
- Zhang, Y., and Jiang, J. J. (2008). "Acoustic analyses of sustained and running voices from patients with laryngeal pathologies," *J. Voice* **22**, 1–9.
- Zhang, Y., Jiang, J. J., Biazzo, L., and Jorgensen, M. (2005a). "Perturbation and nonlinear dynamic analyses of voices from patients with unilateral laryngeal paralysis," *J. Voice* **19**, 519–528.
- Zhang, Y., Jiang, J. J., Wallace, S. M., and Zhou, L. (2005b). "Comparison of nonlinear dynamic methods and perturbation methods for voice analysis," *J. Acoust. Soc. Am.* **118**, 2551–2560.
- Zhang, Y., McGilligan, C., Zhou, L., Vig, M., and Jiang, J. J. (2004). "Nonlinear dynamic analysis of voices before and after surgical excision of vocal polyps," *J. Acoust. Soc. Am.* **115**, 2270–2277.