Published in final edited form as:

Ann Otol Rhinol Laryngol. 2011 March ; 120(3): 155–161.

Objective Methods of Sample Selection in Acoustic Analysis of Voice

Aleksandra E. Olszewski, Lisa Shen, B.S., and Jack J. Jiang, M.D., Ph.D.

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

Abstract

Objectives—In acoustic voice analysis, reproducible methods of sample selection are not defined, impeding research study generalizability and clinical assessment of treatment efficacy. Because perturbation results differ along a single signal, this study sought to establish objective methods of sample selection using a moving window to determine the most stable regions of phonation.

Methods—Voice signals obtained from 21 patients affected by laryngeal conditions associated with Parkinson's disease were analyzed to study jitter, shimmer, signal-to-noise ratio (SNR), and correlation dimension (D_2) parameters when various sample selection procedures are used. Objectively selected voice samples were chosen based upon 5%, 10%, and 20% variance from a signal's minimum perturbation value. The stability of these samples, defined by standard deviations of the acoustic measurements, was compared to the stability of unselected samples and subjectively selected samples.

Results—A significant decrease in standard deviation values of acoustic parameters was found when comparing the objectively selected samples (particularly those selected with 5% and 10% variance) to the subjectively selected and unselected samples.

Conclusions—These results suggest that the development of an objective sample selection method may have significant effects on the stability and reliability of acoustic voice measurements.

Keywords

acoustic analysis; moving window; window selection; idiopathic Parkinson's voice; disordered voice

INTRODUCTION

Acoustic and perturbation parameters such as F_0 , jitter, shimmer, SNR, and D_2 , are well-established as noninvasive and objective methods for quantitatively assessing hoarseness, grade, and various physiological aspects of the vocal folds. Such parameters allow researchers to more fully describe normal and pathological voices, while providing clinicians with a method to temporally assess treatment efficacy.^{1,2,3} The value of acoustic parameters depends on the ability to measure data consistently and accurately such that nuances in a given speech waveform have significance when compared to other waveforms. Prior to acoustic analysis in clinical or research settings, recorded samples are cut by

Correspondence and reprint requests should be addressed to: Dr. Jack J. Jiang, Professor, University of Wisconsin School of Medicine and Public Health, Medical Sciences Center, Room 5745, 1300 University Avenue, Madison, WI 53706-1532, Tel: (608) 265-7888, Fax: (608) 265-2139, jiang@surgey.wisc.edu.

selection of a particular segment of phonation. Despite the importance of sensitivity and repeatability in acoustic analysis, the methods are yet to be standardized in terms of length and location of sample selection.

As explained by Titze (1995), because acoustic analysis should only be applied to periodic or nearly periodic signals, it depends upon selection of stable segments of phonation.¹ Thus, ensuring that the most stable vowel sample is always selected will allow researchers to obtain stable and consistent results because acoustic and perturbation results vary depending on the location along the waveform of a signal (Figures 1, 2, 3).² Most researchers today acknowledge the instability of the beginning and final segments of voice samples, respectively termed onset and offset.^{3,4,5,6} These waveform segments are extremely complex due to the changes in aerodynamic and muscular parameters present during vocal onset and offset. Such changes result in elevated jitter and shimmer values due to the quickly changing fundamental frequency, potentially leading to false positive diagnoses and impeding treatment plans.^{6,7,8} Consequently, the middle and often most stable regions of the voice signals are usually selected when performing acoustic analysis.^{9,10,11}

Various procedures for sample selection exist today, many of which depend on subjective perception of stable signals.^{9,10,11} Unfortunately, such selection procedures may produce varying results due to the inconsistencies between judges. Although it is accepted that voice onset and offset are excluded, the specific methodology of sample selection is often not specified in studies. Aside from leading to subjective and irreproducible results, such ill-defined methods may result in selection of samples that are not representative of a given voice.

A voice signal, even one deemed normal, is never stable.^{12,13} Because a single sample may produce a large range of perturbation values, each possible region selected for analysis may produce a different mean jitter or shimmer. Furthermore, outliers in a temporal progression of measured acoustic parameters often cannot be spotted within the original waveform signal during subjective sample selection, so they may be selected as part of the sample. However, during analysis these outliers jump abruptly outside of a steady range of values, and they have a great capacity for skewing the resulting measurements.^{12,13}

Some objective selection methods depend on choosing a segment from the center or another defined point in time along the sample. However, any given region may contain outliers or unstable segments. Due to the instability of each voice signal, basing selection on a sample's location over time rather than on its relative perturbation values does not select the region most suited for acoustic analysis or one which is necessarily comparable between groups. Combining the need for stable segments in acoustic analysis¹ with the importance of consistent sample selection, we propose objective selection using moving window analysis to determine the regions of lowest perturbation. That is, always choosing the regions of lowest perturbation in a signal allows researchers and physicians to choose samples that are more amenable to acoustic analysis.¹ To measure aperiodicity reproducibly, such that researchers can compare perturbation values between subjects and clinicians can accurately evaluate treatment efficacy, the sample should always be selected from the same location (the objectively determined minimum value of perturbation).

This study employed moving window analysis to consistently select the most stable and least disordered regions of the voice signal. Perturbation parameters including percent jitter, percent shimmer, and signal-to-noise ratio (SNR), as well as the nonlinear dynamic parameter correlation dimension (D_2), were used to establish this region. That is, the most stable region was defined as that which presented minimum percent jitter, percent shimmer, and D_2 with the maximum SNR. This region was found with respect to each perturbation

value, and tended to be in the same region within a signal (towards the second half of the recording) (Figures 1,2,3). The deviations of these acoustic parameters within a certain range of the most stable segment (5%, 10%, and 20%) defined the objective criteria for selecting voice samples. In comparison with a subjective method, these objective methods of sample selection are expected to show increased sample stability and measurement consistency (as determined by smaller standard deviations for each parameter).

MATERIALS AND METHODS

Subjects

The patients involved in this study were individuals diagnosed with Parkinson's disease selected by their neurologist as having typical characteristics of idiopathic Parkinson's disease. Compared to normal voices, this disordered voice type presents a greater variability within and between phonations and thus a greater need for standardized sample selection. Furthermore, a major goal of the study was to improve consistency in treatment efficacy, necessitating the use of disordered voice samples. The 21 patients (12 male and 9 female) who participated all complained of deterioration of voice, speech, or both, during the course of idiopathic Parkinson's disease, but none of the subjects had other indications of pharyngeal or laryngeal disease. The average age of the patients was 66.4 years with a standard deviation of 6.55. Subject participation was approved by the Institutional Review Board of the University of Wisconsin-Madison.

Acoustic Recording

Voice recordings were made in a double-walled, sound-attenuated room. Subjects were asked to sustain the vowel/a/at a comfortable pitch and intensity, as steadily as possible for at least 1 second. Voice recording was completed using a hand-held microphone (AKG-c410, Kay Elemetrics) positioned 10cm from the subject's mouth at a 45 degree angle. For each subject up to three separate voice samples were recorded. Voice signals were digitized with a 12-bit analog-to-digital converter (National Instruments AT-MIO-E-2) at a sampling rate of 20 kHz in LabVIEW 4.0 (National Instruments).

Data selection and analysis

The *CSpeech software*, version 4.0, was used for traditional acoustic analysis of the voice signals (Milenkovic and Read, Madison, Wisconsin).¹⁴ Acoustic parameters are used to describe a voice as normal versus aperiodic. In this study, F_0 , SNR, and perturbation parameters of percent jitter and percent shimmer, were collected. F_0 is a quantification of the vocal fold vibratory frequency. SNR, also called harmonics-to-noise ratio (HNR), is the ratio between the total energy of the signal and the energy of the noise components. Higher SNR values suggest higher harmonic components of the signal over other noise components and are essentially associated with normal voices. That is, SNR inversely correlates with the "noisiness and roughness" of a voice.¹ Jitter (percent jitter) measures the percent variation in cycle-to-cycle frequency of a voice, while shimmer (percent shimmer) measures the percent variation in cycle-to-cycle amplitude of a signal. Measuring jitter and shimmer allows detection of changes in mass, tension, biochemistry and neuromuscular vocal fold control; low jitter and shimmer values are associated with an ability to maintain periodic vibration. An increase in jitter or shimmer implies a less periodic voice and is associated with perceived roughness or hoarseness.¹

Nonlinear dynamic calculations were based on the numerical algorithms described for past studies analyzing pathological human voices and excised larynx phonations.^{16,17,18} For a detailed description of the methodology, please refer to these studies. To create a reconstructed phase space, a voice signal was plotted against itself at some time delay.

Because periodic signals are depicted by a closed trajectory while aperiodic signals are irregular, the reconstructed phase space can be used to qualitatively illustrate the dynamic behavior of a signal.¹⁵ From the reconstructed phase space, D_2 was calculated. D_2 refers to the number of degrees of freedom needed to describe a system. Since a disordered system with a higher dimension needs more state variables to describe it, the D_2 values for such systems will be substantially higher as compared to those of normal systems.

The time delay technique was applied to reconstruct an m -dimensional delay-coordinate 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.¹⁵ To determine m , the embedding theorem was used.¹⁹ When $m > 2D+1$, where D is the Hausdorff dimension, the reconstructed phase space topologically mimics the original phase space. To estimate the proper τ , the mutual information method proposed by Fraser and Swinney was applied.²⁰ Then, Theiler's improved algorithm was used to determine the correlation integral $C(r)$, where r is the radius around $\mathbf{X}_i$.²¹ Exhibiting power law behavior described by: $C(r) \propto r^{D_2} e^{-m\tau K_2}$, which reveals the attractor's geometrical scaling property, $C(r)$ is a measure of the number of distances between points in the reconstructed phase space that are smaller than the radius r . After calculating $C(r)$, D_2 was estimated in the scaling region of r with the increase of m . Curves of $\log_2 C(r)$ versus $\log_2 r$ were generated for each embedding dimension, m , and the curve fit was used to derive D_2 as well as its standard deviation. Standard deviation of the values was less than 5%.

Each voice signal was divided into overlapping segments of 500ms by using square windows. The time shift between two consecutive voice segments was 25ms. This window and overlap size allowed for a high resolution of the time domain and a more accurate assessment of the lowest perturbation segment. The mean measurements of acoustic parameters (percent jitter, percent shimmer, SNR, and D_2) were then determined for voice samples in five different groups (Figures 1,2,3). Group 1 represents the original unselected voice samples, in which the entire length of the sample was used for analysis. The mean percent jitter value of each unselected voice sample was obtained by averaging the measured percent jitter values of the 500ms segments throughout the entire sample. The same was done for percent shimmer, SNR, and D_2 . Group 5 represented a sample that was selected subjectively. Selection criteria was that the sample be located in the middle of the voice signal and appear to be stationary. This group represents the control group, as subjective selection is the procedure most commonly used by clinicians today.^{3,4,5,6} For selected samples, we simultaneously applied the minimum percent jitter, percent shimmer, and D_2 values, and maximum SNR values to define the most stable 500ms voice segment. Groups 2, 3, and 4 include samples that were selected objectively based upon calculations of variance (5%, 10%, and 20% variance, respectively) from the most stable voice segment (Figures 1, 2, 3).

Statistical Analysis

Standard deviations were used as an indication of the stability of voice samples. Data was analyzed using SigmaStat 3.0.1 (Systat Software, San Jose, CA). A Dunn one-way analysis of variance (ANOVA) was performed on the standard deviations of percent jitter, percent shimmer, SNR, and D_2 for the 5 groups of samples. A significance level of $p < 0.05$ was used for all groups.

RESULTS

Sample Length

Samples selected with the most stringent criteria (objectively selected samples with 5% variance) yielded the shortest samples. Sample length increased for samples selected with

10% variance, 20% variance, and subjective methods (Figures 1,2,3). The ranges of sample length for samples selected with 5% variance, 10% variance, 20% variance, and subjective selection methods were 500ms–825ms, 550ms–950ms, 550ms–1100ms, and 625ms–1400ms, respectively. The unselected samples were the longest, as they resulted in samples that ran the entire length of the waveform (1000ms–2700ms).

Jitter

Analysis of the percent jitter results for each of the five groups confirmed that sample selection, and especially objective sample selection, significantly lowered the deviance of the resulting measurements (Figure 4). The standard deviations for percent jitter increased dramatically as the selection criteria decreased from objective selection of samples with no more than 5% variance to subjective perceptual selection. As expected, the unselected samples showed the greatest standard deviation values since this group included the range of percent jitter values for the entire length of the wave signal. A box plot distribution of the standard deviations (Figure 4) illustrates the great range of deviance in the unselected samples (Group 1) as compared to objectively and subjectively selected samples (Groups 2–5). A comparison of the five groups using Dunn's method showed that there was a significant difference in the values of standard deviation between the group of unselected samples (Group 1) and each group of selected samples (Groups 2–5) ($p < 0.01$). Furthermore, results showed that the standard deviations of subjectively selected samples (Group 5) were significantly larger as compared to the standard deviations of objectively selected samples with 5% and 10% variance (Groups 2 and 3, respectively) ($p < 0.05$). Since lower standard deviation values are correlated with overall stability of the sample, the objectively selected samples with 5 and 10% variance appear most stable. However, no significant difference in standard deviation was found between the subjectively selected samples (Group 5) and the objectively selected samples with 20% variance (Group 4) ($p > 0.05$).

Shimmer

Analysis of percent shimmer results, as illustrated in Figure 5, confirmed that the unselected samples (Group 1) had significantly greater standard deviation values as compared to objectively selected samples (Groups 2–4) ($p < 0.01$). Furthermore, the standard deviations of the subjectively selected samples (Group 5) were significantly larger as compared to the standard deviations of the objectively selected samples with 5% and 10% variance (Groups 2 and 3, respectively) ($p < 0.05$). However, the standard deviation of subjectively selected samples (Group 5) was not significantly different from those of the unselected group (Group 1) or the objectively selected samples with 20% variance (Group 4) ($p > 0.05$). This indicates that neither the subjective sample selection method nor the objective method at 20% variance yielded significantly more stable results. Results shows that the objective selection process significantly increased the stability of the sample, and the groups selected with 5% and 10% variability are especially stable as compared to the other groups.

Signal-to-noise ratio (SNR)

Analysis of SNR variance showed a significant difference in the values of standard deviation between the unselected samples (Group 1) and subjectively selected samples (Group 5) (Figure 6) ($p < 0.05$). Furthermore, the test showed that the standard deviations of the unselected samples (Group 1) were significantly larger as compared to the standard deviations of the objectively selected samples with 5% and 10% variance (Groups 2 and 3, respectively) ($p < 0.01$). However, they were not significantly different from the samples with 20% variance (Group 4) ($p > 0.05$). There was also no significant difference in standard deviation values between the subjectively selected samples (Group 5) and the objectively

selected samples (Groups 2, 3, and 4). These results show an increased stability for signals selected at 5% and 10% variance but not for those selected at 20% variance.

D₂

Analysis of the D₂ values showed that standard deviations of the unselected group (Group 1) were significantly greater than those of the selected groups (Groups 2–5) (Figure 7) ($p < 0.05$). Furthermore, standard deviations of D₂ values of the subjectively selected samples (Group 5) were significantly greater than those of the samples selected objectively at 5% and 10% variance (Groups 2 and 3, respectively) ($p < 0.05$). However, no significant difference in standard deviation values was found between the subjectively selected samples (Group 5) and the objectively selected samples of 20% variance (Group 4), indicating that a criteria of 20% variability may not be sufficient for obtaining results with better reliability than those of current practices.

DISCUSSION

Acoustic parameters are often used to characterize the voice. Although researchers agree that onset and offset of phonation must be cut out of a signal prior to analysis, the exact position and length of selected samples differs among studies, causing incomparable perturbation values between studies.^{22,23,24,25} This study tested an objective method for sample selection to obtain maximally reliable acoustic analysis. Past studies that have compared acoustic and perturbation measures taken from objective and subjective selection methods showed reduced percent jitter with selection of steady vowel samples, emphasizing the need for a standard sample selection method.²⁶ Looking at the variance of acoustic and perturbation values rather than comparing the values themselves, our results provide evidence of the importance of a sample selection process, as indicated by the significant decrease in variance when employing sample selection, especially objective selection.

Due to the complex and chaotic patterns inherent to vocal onset and offset, these segments are often excluded in analysis although few studies have considered the various components affecting this increased instability.²⁷ Our results show that unselected voices contained large variability in acoustic and perturbation values as a result of the onset and offset included as part of the voice sample. This is in accordance with past studies, and supports the need for researchers to characterize these unstable segments.

With sample selection of voices, instability and variance were significantly reduced. When comparing the performance of subjective and objective methods, we found that there were no significant differences between the standard deviations of percent shimmer of unselected and subjectively selected groups (Figure 5). On the other hand, objective methods did show a significant difference in stability, indicating that in the case of percent shimmer, the subjective method may not be useful for obtaining stable samples, while the objective method is valuable for increasing sample stability. In the cases of percent jitter and D₂, any sample selection, whether subjective or objective, significantly decreased measurement standard deviation (Figures 4, 7). However, the measurement stability significantly increased for signals cut at 5% and 10% variance, as compared to signals selected at 20% variance and by subjective cutting. Furthermore, according to the results of SNR stability, samples cut at 5% and 10% variability showed significant differences as compared to unselected samples, whereas samples cut at 20% variance did not differ from unselected signals (Figure 6). It is noteworthy that SNR results did not show increased stability from objective selection methods as compared to subjective methods, perhaps due to the greater proportion of the signal included in analysis, resulting from the relative stability of the overall signal's SNR.

Disordered voices show greater variability than normal voices, stressing the need for a consistent and objective selection method. Accordingly, in our findings, the minimum value of perturbation varied in location in each voice signal across individuals, indicating that vowel selection based on minimum perturbation values is preferable to fixed point selection, the objective method most commonly applied today. However, the minimum values of percent jitter, percent shimmer, and D₂, and maximum SNR values were found in similar regions within a given signal (Figures 1, 2, 3), indicating that the region of greatest stability overlapped for each of these parameters. Typically, the minimum values were found towards the second half of the signal for Parkinson's patients, although this location may vary for normal voices and voices with different disorders, stressing the need to apply this study with a variety of voice types.

In general, results indicate that using 5% or 10% variance for objectively selecting samples, the most stable and consistent sections will be selected for analysis, prospectively resulting in a more reliable method of acoustic analysis to track the vocal fold physiology of Parkinson's patients. The objective selection criteria of 20% variance was not found to be an improved method of sample selection as it often did not show a significant difference in standard deviation values when compared with the subjective method that is applied in current clinical practice. It is important to consider that samples with higher stability are compromised by shorter lengths. In some samples that were selected with 5% variance, only 500ms of the signal could be selected, and studies have shown that a minimum signal length of 1 second is necessary for formant estimation.²⁸ Thus, the selection procedure using 5% variance may not be applicable in all circumstances. In this case, the selection method using 10% variance as the limiting criteria may be the most beneficial due to its significantly stable results and sufficiently longer sample lengths.

In the future, this study will be applied to other types of voices as well, such as normal voices and those affected by laryngitis, nodules, or polyps. The true benefits of objective moving window sample selection will be determined after comparisons are made between voices of various types of laryngeal pathologies. For example, with normal voices, it is likely that objective selection will not be as beneficial, due to the decreased likelihood that outlying perturbation values will be included in any given voice segment. Future studies may also consider treatment evaluations over time using differing sample selection methods to determine if treatment recommendations are affected by selection methods. Lastly, applying variances between 10% and 20% may offer a more specific evaluation of various degrees of objective selection. In this study, using the tested subject population of Parkinson's patients, objective methods using a moving window to select stable segments provided significantly more stable and consistent acoustic and perturbation values, indicating that such methods may offer a standardized way to ensure acoustic analysis in the clinic is reliable and repeatable.

CONCLUSIONS

In this study, we used moving window analysis to select the most stable 500ms segments from voice recordings of patients with idiopathic Parkinson's disorder. Compared to both unselected samples and subjectively selected samples, objective (moving window) analysis provided significantly decreased variance of values of percent jitter, percent shimmer, and D₂. For SNR, moving window analysis increased signal stability from that of the unselected sample but not compared to the stability of the subjectively selected sample. Our results indicate that moving window analysis avoids variance due to the fluctuating stability of a given voice signal and provides a method of acoustic analysis that is repeatable between studies. However, stringent signal selection also decreases the signal length being analyzed. It is important to find selection criteria that maximize stability while avoiding inefficient

signal length. We conclude that selecting samples that vary from the most stable segment by 10% provided the best results in terms of both sample length and stability.

Acknowledgments

The study was supported by NIH grant number 1-R01DC05522-01 from the National Institute of Deafness and other Communication Disorders.

References

1. Titze, IR. Workshop on acoustic voice analysis: Summary Statement. Denver, CO: National Center for Voice and Speech; 1995. p. 1-36.
2. Kent RD, Vorperian HK, Kent JF, Duffy JR. Voice dysfunction in dysarthria: application of the multi-dimensional voice program. *Journal of Communication Disorders*. 2003; 36:281–306. [PubMed: 12837587]
3. Bielamowicz S, Kreiman J, Gerratt BR, Dauer MS, Berke GS. Comparison of voice analysis systems for perturbation measurement. *Journal of Speech and Hearing Research*. 1996; 39(1):126–134. [PubMed: 8820704]
4. Karnell MP. Laryngeal perturbation analysis: minimum length of analysis window. *Journal of Speech and Hearing Research*. 1991; 34(3):544–548. [PubMed: 2072678]
5. Jones TM, Trabold M, Plante F, Cheetham BMG, Earis JE. Objective assessment of hoarseness by measuring jitter. *Clinical Otolaryngology*. 2001; 26(1):29–32. [PubMed: 11298163]
6. Yu P, Ouaknine M, Revis J, Giovanni A. Objective voice analysis for dysphonic patients: a multiparametric protocol including acoustic and aerodynamic measurements. *J Voice*. 2001; 15(4): 529–42. [PubMed: 11792029]
7. Brockmann MB, Storck C, Carding PN, Drinnan MJ. Voice loudness and gender effects on jitter and shimmer in healthy adults. *J Speech Lang Hear Res*. 2008; 51:1152–1160. [PubMed: 18664710]
8. Muñoz J, Mendoza E, Fresneda MD, Carballo G, Lopez P. Acoustic and perceptual indicators of normal and pathological voice. *Folia Phoniatrica et logopaedica*. 2003; 55:102–144. [PubMed: 12697983]
9. MacCallum JK, Olszewski AE, Zhang Y, Jiang JJ. Effects of low-pass filtering on acoustic analysis of voice. *J Voice*. 2010 Mar 24. [In press]. [Epub ahead of print].
10. Jiang JJ, Zhang Y, MacCallum J, Sprecher A, Zhou L. Objective acoustic analysis of pathological voices from patients with vocal nodules and polyps. *Folia Phoniatr Logop*. 2009; 61(6):342–9. [PubMed: 19864916]
11. Scherer RC, Gould WJ, Titze IR, Meyers AD, Sataloff RT. Preliminary evaluation of selected acoustic and glottographic measures for clinical phonatory function analyses. *J Voice*. 1988; 2(3): 230–244.
12. Ferrand CT. Effects of practice with and without knowledge of results on jitter and shimmer levels in normally speaking women. *J Voice*. 1995; 9(4):419–423. [PubMed: 8574308]
13. Glaze LE, Bless DM, Susser RD. Acoustic analysis of vowel and loudness differences in children's voice. *J Voice*. 1990; 4(1):37–44.
14. Milenkovic P. Least mean square measures of voice perturbation. *J Speech Hear Res*. 1987; 30:529–538. [PubMed: 2961933]
15. Packard NH, Crutchfield JP, Farmer JD, Shaw RS. Geometry from a time series. *Phys Rev Lett*. 1980; 45:712.
16. Jiang JJ, Zhang Y. Chaotic vibration induced by turbulent noise in a two-mass model of vocal folds. *J Acoust Soc Am*. 2002; 112:2127–2133. [PubMed: 12430824]
17. Zhang Y, Jiang JJ. Nonlinear dynamic analysis of signal typing of pathological human voices. *Electron Lett*. 2003; 39:1021–1023.
18. Jiang JJ, Zhang Y, Ford CN. Nonlinear dynamics of phonations in excised larynx experiments. *J Acoust Soc Am*. 2003; 114:2198–2205. [PubMed: 14587617]
19. Takens, F. Detecting strange attractors in turbulence. In: Rand, DA.; Young, LS., editors. *Lecture Notes in Mathematics*. Berlin: Springer-Verlag; 1981. p. 366-381.

20. Fraser AM, Swinney HL. Independent coordinates for strange attractors from mutual information. *Phys Rev A*. 1986; 33:1134–1140. [PubMed: 9896728]
21. Theiler J. Spurious dimension from correlation algorithms applied to limited time series data. *Phys Rev A*. 1986; 34:2427–2432. [PubMed: 9897530]
22. Vieira MN, McInnes FR, Jack MA. On the influence of laryngeal pathologies on acoustic and electrographic jitter measures. *Journal of the Acoustical Society of America*. 2002; 111:1045–1055. [PubMed: 11863161]
23. Robinson JL, Mandel S, Sataloff RT. Objective Voice Measures in Nonsinging Patients With Unilateral Superior Laryngeal Nerve Paresis. *J Voice*. 2005; 19:665–667. [PubMed: 16129578]
24. Lim JY, Choi JN, Kim KM, Choi HS. Voice analysis of patients with diverse types of Reinke's edema and clinical use of electroglottographic measurements. *Acta Oto-Laryngologica*. 2006; 126:62–69. [PubMed: 16308257]
25. Petrović-Lazić M, Babac S, Vuković M, Kosanović R, Ivanković Z. Acoustic Voice Analysis of Patients with Vocal Fold Polyp. *J Voice*. [In press].
26. Linville SE, Korabic EW, Rosera M. Intraproduction variability in jitter measures from elderly speakers. *J Voice*. 1990; 4(1):45–51.
27. Regner MF, Tao C, Zhuang P, Jiang JJ. Onset and offset phonation threshold flow in excised canine larynges. *Laryngoscope*. 2008; 118(7):1313–1317. [PubMed: 18401267]
28. Stassen H. Affective State and Voice: The Specific Properties of Overtone Distribution. *Methods of Information in Medicine*. 1991; 30(1):44–52. [PubMed: 2005833]

Figure 1.

Time variance of a voice signal with respect to percent jitter. The different samples 2–4 are marked with arrows to show where the original voice signal was cut. Blue arrows represent sample 2 (5% variance), red arrows represent sample 3 (10% variance), and black arrows represent sample 4 (20% variance). Sample 1 is uncut and represented by the entire length of voice signal.

Figure 2.

Time variance of a voice signal with respect to percent shimmer. Arrows define sample selection for samples 2–4, with color designation same as in Figure 1. Sample 1 is uncut and represented by the entire length of the voice signal.

Figure 3.

Time variance of a voice signal with respect to SNR. Sample 1 is uncut and represented by the entire length of the voice signal. In the special case of SNR, samples 3 and 4 are also represented by the entire length of the voice signal. Blue arrows represent sample 2 (5% variance).

Figure 4.
A box plot distribution comparing the standard deviations of percent jitter from various sample selection methods.

Figure 5.

A box plot distribution comparing the standard deviations of percent shimmer from various sample selection methods.

Figure 6. A box plot distribution comparing the standard deviations of SNR from various sample selection methods.

Figure 7.
A box plot distribution comparing standard deviations of D_2 from various sample selection methods.