Published in final edited form as:

J Voice. 2012 January ; 26(1): 1–7. doi:10.1016/j.jvoice.2010.10.009.

The Effect of Segment Selection on Acoustic Analysis

Seong Hee Choi, JiYeoun Lee, Alicia J. Sprecher, and Jack J. Jiang

Division of Otolaryngology, Department of Surgery, University of Wisconsin School of Medicine and Public Health, Madison, WI, U.S.A

Abstract

Objective/Hypothesis—Acoustic analysis is a commonly used method for quantitatively measuring vocal fold function. Voice signals are analyzed by selecting a waveform segment and using various algorithms to arrive at parameters such as jitter, shimmer, and signal-to-noise ratio (SNR). Accurate and reliable methods for selecting a representative vowel segment have not been established.

Study Design—Prospective repeated measures experiment

Methods—We applied a moving window method by isolating consecutive, overlapping segments of the raw voice signal from onset through offset. Ten normal voice signals were analyzed using acoustic measures calculated from the moving window. The location and value of minimum perturbation/maximum SNR was compared across individuals. The moving window method was compared with data from the whole vowel excluding onset and offset, the mid-vowel and the visually selected steadiest portion of the voice signal.

Results—Results showed that the steadiest portion of the waveforms, as defined by minimum perturbation and maximum SNR values, was not consistent across individuals. Perturbation and nonlinear dynamic values differed significantly based on what segment of the waveform was used. Other commonly used segments selection methods resulted in significantly higher perturbation values and significantly lower SNR values than those determined by the moving window method ($p < 0.001$).

Conclusions—The selection of a sample for acoustic analysis can introduce significant inconsistencies into the analysis procedure. The moving window technique may provide more accurate and reliable acoustic measures by objectively identifying the steadiest segment of the voice sample.

Keywords

moving window; window selection; acoustic analysis

Introduction

The characteristics of the speech waveform have been used in research to promote a better understanding of normal and pathological voicing, and to evaluate treatment efficacy.^{1,2} To

Correspondence: Jack J. Jiang M.D., Ph.D., Address: 5745 Medical Sciences Center, 1300 University Ave, Madison, WI 53706, Tel: (608) 265-9854, FAX: (608) 265-2139, jjjiang@wisc.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

provide valuable data on voicing, these acoustic measures must be sensitive to small changes in the speech waveform while generating consistent values for repeated measures.

Currently, acoustic analysis is performed by selecting a particular segment from each voice signal and analyzing the selected segment using defined acoustic algorithms. Titze (1995) suggested that only periodic or nearly periodic voice signals should be analyzed using acoustic measures.³ Therefore, the selection of a stable vowel segment is essential to analysis; however, there is no standard guidance of how to choose a voice segment.

Numerous methods exist to select voice segments. Human perception is commonly used to determine a steady segment of voice.⁴⁻⁶ This method introduces inconsistencies based on variability between judges; moreover, samples are generally selected with minimum amplitude variation with little or no regard for frequency variation. Other researchers use the mid-portion of vowels in their studies⁷⁻⁹; however, such segmentation does not consider changes in the sample stability over time. Feijoo & Hernández (1990) used a 40-ms window and moved in 20-ms interval to determine the accurate pitch period for short-term perturbation measurement in normal and glottal cancer patients.¹⁰ Karnell (1991) selected samples by editing out the beginning and end of each phonation.¹¹ Several other authors have used similar methods to exclude the variability of onset and offset which are generally associated with a rapidly changing fundamental frequency and amplitude leading to increased and unreliable jitter and shimmer values.¹²⁻¹⁴ In many studies the exact means of segment selection is unclear and even where the procedures are well documented they lead to results that are either subjective and irreproducible or unrepresentative of the waveform.

No voice is ever completely stationary. Dynamic changes in stability will always be found upon examination with acoustic methods and periodicity (as expressed by the perturbation values of the waveform) will fluctuate with time as well. Even normal voice samples produce a range of perturbation values.^{7,15} Therefore, it can be expected that inconsistencies in the selection of a sample will produce different perturbation measurements depending upon the exact location of the sample taken from the waveform. To limit these concerns, the point of minimum perturbation and maximum signal to noise ratio (SNR) can be identified objectively in the given voice signals and used to measure aperiodicity.

In this study, a moving window was used to identify the most stable portion (optimal values for the parameters of interest) of voice signals collected from 10 normal subjects. Perturbation parameters and the nonlinear dynamic measure of correlation dimension were calculated. The moving window was used to characterize the changing acoustic parameters within a single voice sample. Results of the moving window selection technique were compared to those obtained by selecting a fixed portion or visually selecting the stable portion of the voice signal.

Materials and Methods

Participants

Five healthy males, aged 21 to 23 years (mean of 21.8) and five healthy females aged 20 to 22 years (mean of 20.8) participated in the study. Subject participation was approved by the IRB of the University of Wisconsin Madison. All participants were nonsmoking native speakers of American English. They reported normal hearing ability, no laryngeal or airway infection and good general health. The subjects were judged to present normal voice and language skills as determined by a certified speech-language pathologist.

Acoustic Recording

Participants were asked to produce sustained /a/ vowel phonation with comfortable pitch and loudness for 3 seconds. 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. A DAT-recorder (SONY TCD-8, Japan) was used with software of the Computerized Speech Lab, Model 4300 (Kay Elemetrics, Lincoln Park, NJ). Voices were recorded with a sampling rate of 44.1kHz.

Data Analysis

Moving window—Samples 500ms in length were used to calculate acoustic parameters. The calculation of acoustic parameters was repeated for each new voice segment as the window was moved forward by 25ms increments from the onset of phonation through its offset as shown in Figure 1. A moving window length of 500ms was selected based on previous data that suggested signal lengths of 500 to 825ms to keep variance in the parameters below 5% of the minimum. After measuring all of the perturbation and nonlinear dynamic (D_2) values in each of the moving window segments, optimal parameter values (minimum jitter, shimmer and D_2 , maximum SNR) were determined and used to identify the most stable portion of the voice. The minimums for percent jitter, percent shimmer and D_2 as well as the maximum signal-to-noise ratio (SNR) were recorded for each voice.

Whole Vowel—Vowels were cut to exclude 500ms at the front and 500ms at the end of the signal. This process removed the onset and offset segments of the sample. The remaining 2 second segment was used to calculate acoustic parameters. The 500ms of onset and 500ms of offset phonation were saved for additional acoustic analysis.

Mid-Vowel—A segment of 1 second phonation was captured from the middle of the phonation (between 1 second after onset and 2 seconds after onset). This segment was used to complete acoustic analysis.

Visually Selected—Voice samples were presented to a researcher blinded to the stability findings of the moving window data. This individual selected a one second segment subjectively considered to be the most stable portion of the phonation.

Acoustic analyses

Perturbation analyses—Following the vowel segmentation, percent jitter, percent shimmer and SNR were measured using CSpeech (P. Milenkovic, 1992, Madison, WI).

Nonlinear Dynamic Analyses—The nonlinear dynamic measure of correlation dimension (D_2) was calculated for each vowel segment. D_2 estimates the number of degrees of freedom necessary to describe a system; a system with greater complexity will produce a higher D_2 value. D_2 was calculated using software developed by the Laryngeal Physiology Laboratory at the University of Wisconsin. Detailed descriptions of the algorithms can be found in numerous publications and will therefore not be described in depth here.¹⁶⁻¹⁸ Briefly, nonlinear dynamic analysis, signals were down sampled to 25kHz using Goldwave software version 5.1 (Goldwave Inc, St. John's NL, Canada). A phase space, $\mathbf{X}_i = \{x(t_i), x(t_i - \tau), \dots, x(t_i - (m-1)\tau)\}$, was created using a time delay technique with m indicating the embedding dimension and τ denoting the time delay.¹⁹ Using Takens' embedding theorem we defined m and D such that $m > 2D + 1$ where D is the Hausdorff dimension.²⁰ Using this method, the reconstructed phase space and the original phase space are topographically equivalent. The time delay was determined using the mutual information method proposed by Fraser.²¹ Using Thelie's improved algorithm, a correlation integral, $C(r)$, was calculated

where r was the radius around $\mathbf{X}_i$.²² $C(r)$ determines the number of distances between points in the reconstructed phase space that are less than r . The function exhibits power law behavior as described by: $C(r) \propto r^{D_2} e^{-mrK_2}$, which reveals the geometric scaling property of the attractor.²³ Using r to define the scaling region, curves of $\log_2 C(r)$ versus $\log_2 r$ were generated for each embedding dimension, m . D_2 was calculated at the point where these curves converged. A standard deviation measure is generated by the program which evaluates the quality of the D_2 estimation and was below 5% for all calculations. Zhang et al (2005) used this method previously and found stable estimations of D_2 with sampling rates as low as 2kHz and signal lengths as short as 20ms.¹⁷

Statistical analyses

Statistical analyses were conducted using Sigma Stat 3.0 (Jandel Scientific, SanRafael, CA). One-way repeated measure ANOVA on ranks were performed to test differences among vowel segments selection methods. Comparisons were calculated for jitter, shimmer, SNR and D_2 values. Multiple pairwise comparisons were conducted with the Turkey method. An alpha of 0.05 was employed for all comparisons.

Results

The mean, standard deviation and range of the perturbation parameters and D_2 as calculated by the moving window are shown in Table 1. As shown in Figure 2, these parameters varied dramatically during the vowel phonation. The locations of the optimal value for each parameter varied across individuals. The minimum percent jitter values were located in windows with an initial time point between 0.275 seconds and 2.05 seconds. Six individuals had minimum percent jitters in windows starting in the first second of the recording while three achieved minimum perturbation values in windows starting between one and two seconds after onset and one individual had a minimum percent jitter in the window between 2 and 2.5 seconds. The minimum percent shimmer windows started between 0.075 seconds and 1.825 seconds after onset with three individuals reaching minimums in the first second and seven having minimum percent shimmer in windows started between the one and two seconds after onset. Maximum SNR values mirrored those of percent jitter with maximum values found in windows stating between 0.25 and 2.05 seconds after onset. Six values were located in windows starting in the first second, three occurred in windows between one and two seconds after onset and two reached maximums in window starting between 2 and 2.5 seconds after onset. For D_2 values, the minimums occurred in windows starting 0.325 to 2.15 seconds after onset. Five voices reached minimums in the windows starting in the first seconds while four achieved minimums between one and two seconds after onset and one voice reached a minimum between 2 and 2.5 seconds after onset. The minimums for jitter, shimmer and D_2 and the maximum for SNR were generally, but not always closely distributed within one voice sample (Figure 3).

Perturbation and D_2 maximums and SNR minimums were more consistently concentrated within the first second of phonation. Of the ten subjects, six had a maximum percent jitter on onset, seven had a maximum percent shimmer on onset, eight showed a minimum SNR at onset and six had a maximum D_2 during voice onset (Figure 3).

The one way repeated measure ANOVA on ranks showed significant variability in the percent jitter, shimmer, SNR and D_2 values between the selection methods ($p < 0.001$, Table 2; Figure 5). The results of the multiple pairwise comparisons are shown in Table 3. For percent jitter, percent shimmer and D_2 , significant differences existed between the minimum value as calculated by the moving window method and all other segment selection methods with the moving window method producing lower values ($p < 0.05$). Comparisons between the other segment selection methods were not significant. In SNR, the maximum value was

significantly different from the whole vowel and mid-vowel selection procedures, but was not significantly different from the visually selected vowel segment ($p=0.005$, $p=0.005$ and $p=0.980$, respectively). For SNR, the visually selected vowel segment produced values that were significantly different from those generated by the whole vowel and mid-vowel methods ($p=0.012$ and $p=0.012$, respectively). SNR was higher when calculated using the moving window or visual selection methods.

Voice onset and offset segments were characterized by higher perturbation and D_2 values and a lower SNR; however, as shown in Table 4, there were no significant differences between voice onset and offset ($p>0.05$).

Discussion

In order to determine the effects segment selection on acoustic measurements, methods for segment selection which have been used in previous acoustic literature were compared. The whole vowel, mid-vowel and visually selected vowel segments generated higher perturbation and D_2 measures. SNR values from visually selected segments were not significantly different from those determined using the moving window method; however, whole vowel and mid-vowel segments tended to have lower SNR values.

The impact of segment selection was demonstrated by Linville *et al.* (1990) who compared perturbation calculated from a fixed location with perturbation calculated from the most stable-appearing portion of the production regardless of its location.²⁴ They found that percent jitter values were greatly reduced when the steadiest segment of vowel production was measured. In contrast to their results, we found no significant differences between the visually selected segments and those selected based on their location within the sample except in SNR; however, this difference may be attributed to the fact that our samples were from normal individuals while the Linville *et al.* study used elderly subjects whose voices may show a greater variability over time.

The moving window provided insight into the behavior of the acoustic parameters over the course of a single voice sample. As seen in Figure 2, acoustic measures vary dramatically within a single voice file, and these variations may be more dramatic in disordered voices. Moreover, the minimum values for perturbation measures and D_2 and the maximum values for SNR varied in location within the voice signal across the individuals. As the steadiest portion of vowels varied between individual subjects, the selection of a vowel segment based on objective steadiness is recommended over fixed point selection. Moreover, as abnormal voices tend to show a greater level of variability, consistent and objective segment selection methods are of even greater importance.²⁵

The location of the minimum values of percent jitter, percent shimmer and D_2 , and the maximum value of SNR were closely distributed within a single voice signal; however, in some samples, these locations varied dramatically. In these cases, the selection of a single voice segment for analysis would not retrieve the best value for all parameters. This is a major challenge when using visually selected segments. As seen in Figure 3, female 5 (F5) had a minimum shimmer value located more than one second from their minimum jitter, D_2 and maximum SNR. Visual selection focuses more heavily amplitude variability; therefore, the chosen segment may have contained the minimum shimmer point, but reported less optimal values for the other parameters. Rather than selecting a single, representative sample, the moving window method allows a user to find the optimal value for each parameter within the sample, regardless of its location.

Generally, the optimal values were found towards the front half of the signal. In many samples, the highest perturbation and D_2 values were located at voice onset; however, the

signal rapidly stabilized such that optimum values for all parameters were almost entirely located within the two seconds of onset. These findings suggest that a short recording length may be adequate to collect stable voice samples.

The acoustic analysis of voice onset and offset has received little attention. Voice onset and offset segments contain changes in the glottal aerodynamic and biomechanical properties resulting in extremely complex and chaotic patterns.²⁶ We found no significant difference in perturbation and nonlinear dynamic values at onset as compared to offset; however, both measures were elevated in these regions indicating the importance of excluding voice onset and offset during acoustic analysis.

The moving window method is useful as it is both receptive to variability in stability and repeatable across trials. While neither the minimum or maximum values of the parameters could be considered typical for a voice, both can be consistently selected. This selection method can improve patient tracking by limiting variability introduced during data analysis. Moreover, if moving window methods were used consistently, comparisons across studies would be feasible. According to Titze, perturbation values above 5% are not reliable; therefore, we selected optimal parameters values to increase the probability that all perturbation measures would be below this cutoff.

Although the moving window technique appears energy intensive, a number of advances could make it amenable to the clinic. We have developed a MATLAB (The MathWorks, Natick, Massachusetts) program to automatically cut wav files into sequential segments of a defined length. The program generates a folder of the cut samples which can easily be run as a batch in CSPEECH or TF32 (P. Milenkovic, 2005, Madison, WI). The data can then be imported into a program such as Excel to identify optimal values for each parameter. The integration of these steps could improve the ease with which moving window can be used. Currently the calculation of D_2 is time consuming and must be improved before it can be implemented clinically.

Although our small sample size is a limitation in this study, the data suggests that the moving window technique can provide more accurate and reliable acoustic measures by objectively identifying the steadiest segment. Future research is needed to generate more information from a larger subset of normal voices and extend this research to pathological voices. In pathologic voices, which show greater variability within and between subjects, the use of the moving window can help to identify the most stable portion of the voice sample to apply acoustic measures reliably and objectively, and can extend acoustic analysis to more disordered voices by avoiding transient periods that may include subharmonics, strong modulations, bifurcations or chaos.

Conclusion

In this study, we applied a moving window to select the steadiest 0.5 second segment from the normal voice signals. The moving window provides evidence of the effects of voice sample segments on acoustic analysis by demonstrating variability of the parameters and their minimums both within and between normal voices. Significant differences were observed for percent jitter, percent shimmer and D_2 when comparing the moving window selected segment to traditionally used segment selection methods. SNR showed differences between the fixed point segment selection procedures and the subjectively and objectively chosen vowel segments. This variability indicates a need to standardize vowel segment selection procedures. The moving window method is responsive to the changing levels of stability in the voice, repeatable between studies and can potentially be used to improve the reliability of acoustic measures.

Acknowledgments

This study was supported by grants R01DC05522 from the National Institute on Deafness and Other Communication Disorders.

References

1. Laver J, Hiller S, MacKenzie J, Rooney E. An acoustic screening system for detection of laryngeal pathology. *J Phonetics*. 1986; 517:24.
2. Kasuya H, Ogawa S, Kikuchi Y. An acoustic analysis of pathological voice and its application to the evaluation of laryngeal pathology. *Speech Communication*. 1986; (5):171–81.
3. Titze, IR. Workshop on acoustic voice analysis: Summary Statement. Denver, CO: National Center for Voice and Speech; 1995. p. 1-36.
4. MacCallum JK, Olszewski AE, Zhang Y, Jiang JJ. Effects of low-pass filtering on acoustic analysis of voice. *J Voice*. In Press.
5. Jiang JJ, Zhang Y, MacCallum J, Sprecher A, Zhou L. Objective acoustic analysis of pathological voices from patients with vocal nodules and polyps. *Folia Phoniatri Logop*. 2009; 61(6):342–9. [PubMed: 19864916]
6. 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.
7. Glaze LE, Bless DM, Susser RD. Acoustic analysis of vowel and loudness differences in children's voice. *J Voice*. 1990; 4(1):37–44.
8. Gelfer MP. Fundamental frequency, intensity, and vowel selection: Effects on measures of phonatory stability. *J Speech Hear Res*. 1995; 38(6):1189–1198. [PubMed: 8747813]
9. Bielamowicz S, Kreiman J, Gerratt BR, Dauer MS, Berke GS. Comparison of voice analysis systems for perturbation measurement. *J Speech Hear Res*. 1996 Feb; 39(1):126–34. [PubMed: 8820704]
10. Feijoo S, Hernandez C. Short-term stability measures for the evaluation of vocal quality. *J Speech Hear Res*. 1990; 33:324–334. [PubMed: 2359273]
11. Karnell MP. Laryngeal perturbation analysis: minimum length of analysis window. *J Speech Hear Res*. 1991 Jun; 34(3):544–8. [PubMed: 2072678]
12. 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]
13. Munoz J, Mendoza E, Fresneda MD, Carballo G, Lopez P. Acoustic and perceptual indicators of normal and pathological voice. *Folia Phoniatri et logopaedica*. 2003; 55:102–144. [PubMed: 12697983]
14. 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 Dec; 15(4):529–42. [PubMed: 11792029]
15. 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]
16. Zhang Y, Jiang JJ, Biazzo L, Jorgensen M. Perturbation and nonlinear dynamic analyses of voices from patients with unilateral laryngeal paralysis. *J Voice*. 2005a; 19(4):519–528. [PubMed: 16301098]
17. Zhang Y, Jiang JJ, Wallace SM, Zhou L. Comparison of nonlinear dynamic methods and perturbation methods for voice analysis. *J Acoust Soc Am*. 2005b; 118:2551–2560. [PubMed: 16266175]
18. Zhang Y, Jiang JJ. Acoustic analyses of sustained and running voices from patients with laryngeal pathologies. *J Voice*. 2008; 22:1–9. [PubMed: 16978835]
19. Packard NH, Crutchfield JP, Farmer JD, Shaw RS. Geometry from a time-series. *Physical Review Letters*. 1980; 45:712–716.

20. Takens, F. Detecting strange attractors in turbulence. In: Rand, D.; Younge, L., editors. Lecture Notes in Mathematics. Springer-Verlag; Berlin: 1981. p. 366-381.
21. Fraser AM, Swinney HL. Independent coordinates for strange attractors from mutual information. *Phys Rev A*. 1986; 33:1134–1140. [PubMed: 9896728]
22. Theiler J. Spurious dimension from correlation algorithms applied to limited time-series data. *Phys Rev A*. 1986; 34:2427–2432. [PubMed: 9897530]
23. Grassberger P, Procaccia I. Measuring the strangeness of strange attractors. *Physica D*. 1983; 9:189–208.
24. Linville SE, Korabic EW, Rosera M. Intraproduction variability in jitter measures from elderly speakers. *J Voice*. 1990; 4(1):45–51.
25. Boyanov B, Hadjitodorov S. Acoustic analysis of pathological voices. A voice analysis system for the screening of laryngeal diseases. *IEEE Eng Med Bio Mag*. 1997; 16(4):74–82. [PubMed: 9241523]
26. 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]

Figure 1.

A 500ms moving windows was shifted by 25 ms from voice onset through offset during a sustained /a/ phonation. The top window shows the whole 3 second sample while the bottom window shows one 0.5 second window used for analysis.

Figure 2.

Demonstration of the variation of perturbation and nonlinear dynamic values from voice onset to offset in a 3 second recoding from one normal voice using a moving window calculation. Panels A, B, C and D show percent jitter, percent shimmer, SNR and D2, respectively.

Figure 3.

Shown in the top panel are the beginnings of the windows generating the optimal values for each parameter (Minimum jitter, shimmer and D_2 and maximum SNR). The beginnings of windows containing parameter maximums (and SNR minimums) are shown in the bottom panel. The points for each subject are connected to provide an indication of the variability of local stability for each parameter; however, these lines should not be interpreted to indicate an inter-relatedness of the parameters at hand.

Figure 4.

Box plots of percent jitter, percent shimmer, SNR and $D2$ calculated using the whole vowel, mid-vowel, optimal values determined using the moving window and the visually selected most stable segment. The top and bottom of the box represent the 75th and 25th the percentile, respectively. The midline indicates the median; the whiskers correspond to the minimum and maximum values and black dots indicate outliers. Stars indicate segmentation methods that are significantly different from the group according to pairwise comparisons.

Table 1

Mean (M), standard deviation (SD) and range of perturbation and D2 values calculated using the moving window from vowel onset through offset.

	Jitter (%)		Shimmer (%)		SNR (dB)		D ₂	
	M(SD)	Range	M(SD)	Range	M(SD)	Range	M(SD)	Range
M1	0.60(0.10)	0.34	3.72(0.48)	2.22	18.31(0.80)	2.80	1.34(0.16)	1.593
M2	0.80(0.29)	10.30	4.38(0.85)	31.77	16.35(0.70)	16.00	1.41(0.36)	2.469
M3	0.22(0.05)	3.69	1.80(0.41)	11.96	25.08(0.76)	14.10	1.29(0.49)	2.509
M4	0.60(0.56)	2.89	5.40(5.13)	15.08	13.90(5.13)	9.00	2.20(2.33)	4.385
M5	0.39(0.05)	3.26	1.88(0.24)	2.06	22.72(1.64)	4.80	1.26(0.35)	3.098
F1	0.58(0.18)	0.73	2.31(0.44)	2.06	24.41(2.33)	6.30	1.42(0.42)	3.532
F2	0.52(0.10)	0.75	4.10(0.60)	3.50	19.06(0.48)	2.00	1.38(0.21)	2.055
F3	0.23(0.04)	0.27	1.35(0.16)	1.15	27.08(0.86)	8.00	1.22(0.18)	1.371
F4	0.44(0.15)	0.40	1.28(0.36)	3.55	26.97(1.46)	9.40	1.16(0.12)	0.700
F5	0.29(1.17)	0.17	2.26(0.04)	1.81	23.99(0.60)	3.90	1.33(0.24)	1.869
Average	0.467(0.27)	2.28	2.85(0.871)	7.52	21.62(1.41)	7.63	1.40(0.49)	2.360

Note. Range indicates the difference between minimum and maximum perturbation (percent jitter, percent shimmer and SNR) and D2 values.

Table 2

Comparisons of vowel signals excluding voice onset and offset (Whole vowel), mid-vowel, visually selected and moving window for percent jitter, percent shimmer, SNR and D_2 in 10 normal subjects.

	Whole Vowel		Mid-Vowel		Moving Window		Visually selected		P-values
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
Jitter (%)	0.453	0.180	0.409	0.143	0.299	0.110	0.430	0.178	$P<0.001^*$
Shimmer (%)	2.719	1.275	2.700	1.374	1.999	0.831	2.643	1.279	$P<0.001^*$
SNR (dB)	21.90	4.530	21.90	4.76	23.45	4.11	23.29	5.455	$P<0.001^*$
D_2	1.303	0.203	1.3535	0.429	1.053	0.141	1.384	0.331	$P<0.001^*$

Table 3

Multiple pairwise comparisons of acoustic measures among the different vowel segment selection methods.

	Whole Vowel	Mid-Vowel	Visually Selected	Moving Window
Percent Jitter				
Whole Vowel	-	P=0.355	P=0.817	P<0.001 *
Mid-Vowel	-	-	P=0.854	P=0.002 *
Visually Selected	-	-	-	P<0.001 *
Percent Shimmer				
Whole Vowel	-	P=0.999	P=0.953	P<0.001 *
Mid-Vowel	-	-	P=0.979	P<0.001 *
Visually Selected	-	-	-	P<0.001 *
SNR				
Whole Vowel	-	P=1.000	P=0.012 *	P=0.005 *
Mid-Vowel	-	-	P=0.012 *	P=0.005 *
Visually Selected	-	-	-	P=0.980
D2				
Whole Vowel	-	P=0.907	P=0.670	P=0.012 *
Mid-Vowel	-	-	P=0.966	P=0.002 *
Visually Selected	-	-	-	P<0.001 *

* Denotes significant comparison

Table 4

Comparison of voice onset and offset values of percent jitter, percent shimmer, SNR and D_2 for 10 normal subjects.

	Voice Onset		Voice Offset		p-value
	Mean	SD	Mean	SD	
Jitter (%)	1.58	1.50	1.05	0.84	P=0.327
Shimmer (%)	7.17	4.85	4.73	3.76	P=0.224
SNR (dB)	16.64	3.94	20.5	5.25	P=0.060
D_2	2.42	0.72	2.25	1.22	P=0.716