



# Energy Distribution Analysis and Nonlinear Dynamical Analysis of Adductor Spasmodic Dysphonia

Jiantao Wu<sup>1,3</sup>, Ping Yu<sup>2</sup>, Nan Yan<sup>1</sup>, Lan Wang<sup>1</sup>, Xiaohui Yang<sup>3</sup>, Manwa L. Ng<sup>4</sup>

<sup>1</sup> Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences/The Chinese University of Hong Kong, Shenzhen, China

<sup>2</sup> Department of Otorhinolaryngology Head and Neck Surgery, Chinese People's Liberation Army General Hospital, Beijing 100853, China

<sup>3</sup> Changchun University of Science and Technology, Changchun, China

<sup>4</sup> Speech Science Laboratory, University of Hong Kong, China

jt.wu@siat.ac.cn, nan.yan@siat.ac.cn

## Abstract

The present study investigated the voice quality associated with adductor spasmodic dysphonia by means of various acoustical measures. Energy distribution analyses and nonlinear dynamical measurements were used to depict the differences among voices associated with adductor spasmodic dysphonia (ADSD), vocal nodules (VN) and normal voices. Possible acoustical differences between voices associated with unilateral vocal fold paralysis (UVFP) and ADSD, UVFP and normal voice were investigated. Noise-to-harmonic ratio (NHR), harmonic-to-noise ratio (HNR), glottal-to-noise excitation (GNE), empirical mode decomposition excitation ratio (EMD-ER), nonlinear recurrence period density entropy (RPDE), detrended fluctuation analysis (DFA), correlation dimension (D2), and permutation entropy (PE) values were obtained from the sustained vowel /a/ produced by the subjects. Results revealed high specificity of these acoustic measures in distinguishing the voice quality of ADSD, VN, and UVFP voices from normal voices. In addition, combining GNE and D2 measures appears to be effective in distinguishing ADSD from normal and VN voices.

**Index Terms:** energy distribution analysis, nonlinear dynamic analysis, adductor spasmodic dysphonia

## 1. Introduction

Spasmodic dysphonia (SD) is a neurological disorder which manifests itself as a movement disorder affecting the laryngeal muscles [1, 2]. Generally, SD can be classified into three types: adductor, abductor and mixed, among them the adductor type (ADSD) is most common. ADSD is characterized by spasm of adductory laryngeal muscles (such as the lateral cricoarytenoid and the interarytenoid muscles), resulting in the vocal folds forcefully slamming against each other during phonation. Such muscle spasm renders vocal folds difficult to vibrate and generate sounds, and resulting in frequent voice breaks and difficult vocal initiation. Consequently, ADSD patients often exhibit signs of irregular voice breaks, effortful phonation, hoarseness and breathiness [2, 3].

Diagnosis of SD has been a long-standing challenge to practicing speech therapists. They mainly rely on both perceptual judgment of voice quality and first hand visual data

obtained using viewing instrumentation such as laryngoscopy, and stroboscopy. However, regardless of the speech therapists' experience, perceptual judgment is still considered subjective. Meanwhile, laryngeal viewing is invasive and is likely uncomfortable to patients [4, 5, and 6]. With the presence of intermitted voice breaks in ADSD voice, the noise component is obvious, which adversely affects the quality of ADSD voice [7]. Currently, different acoustic parameters have been used to evaluate pathological voices [8, 9]. Attempts have been made to evaluate the voice quality of ADSD voices by using traditional acoustic measures. According to Dejonckere *et al* [8], evaluation of ADSD has significantly benefited from measurements such as jitter, shimmer and number of voice breaks. Moreover, fundamental frequency was also effective in assessing ADSD voice [10, 11]. Despite the validity of traditional acoustic measures in distinguishing ADSD from normal voices as reported in these studies, none quantified the relative proportion of signal and noise, as well as the chaotic feature of ADSD voices. Due to the irregularity in vocal fold vibration in ADSD patients, noise markedly contributes to the associated voice quality.

According to Titze *et al* [12], there are three types of voice signals: (1) nearly-periodic signals, (2) signals with strong sub-harmonics or modulations, and (3) aperiodic signals. Traditional acoustic measures such as jitter, shimmer, and fundamental frequency have been investigated to a great extent and frequently applied in the clinical practice. However, they were applied only to the first voice signals [13-16]. Since these calculations rely heavily on the assumption of the linearity of signals, the use of traditional analyses appears to be of practical limitations [13, 15]. Some new analytic measures have shown potential in reliably evaluating the pathological quality of voices, such as GNE, empirical mode decomposition excitation ratio (EMD-ER), and the nonlinear dynamics techniques [17-21]. GNE and EMD-ER have been used to calculate the SNR of voices of unilateral vocal fold paralysis (UVFP). It has been found that the measures were significant different between pre-operative and post-operative UVFP voices [22]. Used as a new tool to analyze the nonlinear characteristics of different lengths of aperiodic voices signals, D2 has already been successfully used for measuring pathological voices and, accordingly, with greater reliability than traditional perturbation analyses [13-16]. These new

measures offer a huge advantage in studying pathological voices, such as voices associated with vocal nodules, vocal polyps, esophageal phonation and UVFP [15, 16 and 23]. Permutation entropy (PE) is a complex measure that only makes use of the order of original data, not the complex modes of operation to the aperiodic time series [24-26]. According to Bandt et al [25], the robustness of PE is under the nonlinearity distortion of time series and the calculation of PE is often effective. As a nonlinear measure for status characterization, PE has been effectively used in biomedical signals [27].

Previous studies have demonstrated the usefulness of the energy distribution analyses and nonlinear dynamical analysis in distinguishing the pathological from normal voices, in particular nonlinear dynamic measures can be used to analyze aperiodic voices [13-16, 22]. Yet, no study has made use of such acoustical analytic methods to evaluate the ADSD voice, an aperiodic voice that is caused by spasmodic vocal fold behavior. The differences between the ADSD and VN being addressed by using these acoustical analytic methods are also not known. It is not known if patients could equally benefit from the procedure. It was suggested that combing the energy distribution techniques with nonlinear dynamics measures may be effective in contrasting ADSD from normal voices, and these measures may be significantly different in investigating the ADSD voice, VN voice and UVFP voice.

In the present study, some new acoustic measures including energy distribution analyses and nonlinear dynamical techniques in objectively assessing the voice quality of ADSD voices. The nonlinear dynamical measures, including D2, PE, RPDE and DFA were used to evaluate the irregularity and disorder in voice signal [23, 24]. The energy distribution measures, such as HNR, NHR, GNE [17, 18], and EMD-ER [21] were used to evaluate the signal-to-noise ratio (SNR) of different frequency bands in the voice signals.

## 2. Methods

### 2.1. Participants

Fourteen female ADSD patients, 14 female patients with VN and 14 normal female (NF) speakers participated in this study. In addition, ten UVFP patients (6 males, 3 females) and 9 normal male participants were also included for comparison. All speakers were matched with age, with age distributions in the five subject cohorts as follows: ADSD (range: 22-55 years, mean: 39.4 years), VN (range: 21-66 years, mean: 41.2 years), NF speakers (range: 25-54 years, mean: 34.2 years), UVFP (range: 28-50 years, mean: 41.1 years) and normal males (range: 24-57 years, mean: 32.8 years). All subjects were native speakers of Chinese and had no reported history of respiratory, speech, language, and/or hearing problems, except for the patients of different laryngeal pathologies. All subjects were willing to participate in the study and signed informed consents were obtained prior to the experiment.

### 2.2. Speech tasks and Recording procedure

During the experiment, the speakers were instructed to sustain the vowel /a/ three times for as long as they could. To avoid possible tone effect, only high level tone was used, and all the patients were instructed to produce the speech samples at a comfortable level of loudness. During the recording, the microphone was placed approximately 10 cm from the speakers' mouth. Acoustic signals were recorded in a

professional recording booth using an acoustic sensor (B&K 4189), which was sheathed with a mesh screen in order to block the airflow to avoid recording of unnecessary noise. The recorded voice signals were digitized using Praat at a sampling rate of 44 kHz and quantization rate of 16 bits/sample. To avoid effect from phonation initiation and termination, only the medial 80% portion of the sustained vowel segment was used for analysis. All the voice samples were recorded in the Department of Otorhinolaryngology of the People's Liberation Army General Hospital.

### 2.3. Energy distribution analysis

In the present study, four measures were used to quantify the energy distribution in signals.

(a). As the traditionally measure, HNR was used to analysis the extent of noise in the voice signal.

(b). NHR is the reciprocal of each average period of HNR. NHR and HNR were calculated by using Praat [28].

(c). GNE was used to assess the noise in healthy and pathological voice signals which was based on the correlation coefficient Hilbert envelope of different frequency bands. Through the algorithm of GNE measure, we calculated the SNR of speech signal in different frequency bands [17, 18].

(d). As a promising nonlinear approach to time-series analysis, EMD-ER has attracted widespread attention, especially in quantifying the SNR in voice. As proposed by Huang, the voice signal can be decomposed into a limited number of Intrinsic Mode Functions (IMF) [18, 19 and 21]. The energy distribution of each IMF was calculated by using the nonlinear Shannon Entropy [18-22].

### 2.4. Nonlinear dynamic analysis

Nonlinear dynamical analysis methods have been described in detail in the literatures [15, 16, 19, 23, 29, and 30]. In the present study, RPDE, D2, and PE were used to analyze the complexity and irregularity in voice signal. RPDE quantifies the periodicity and uncertainty in measuring the embedded signal [14]. D2 was used to measure the irregularity of the reconstructed phase space. The derivation of D2 needed two input parameters: the embedding dimension 'm' and the time delay 'τ'. M was obtained by the False Nearest Neighbors method [29] and τ was estimated using the C-C methods [30]. D2 and RPDE were manually estimated in the scaling region of the radius r with m. To a time-series signal, as a fast and robust complexity analysis for chaotic time series [31, 32], PE was used to quantified the multi-scale structure. For PE calculation, it only depends on the selection of m (same to embedding dimension) which was also calculated by False Nearest Neighbors (FNN) [19]. In addition to the above-mentioned measures, DFA was used to determine the stochastic self-similarity of turbulent noise in the speech signal. The reliability of all above nonlinear measures was demonstrated in previous studies.

### 2.5. Statistical analysis

All parameters obtained were disposed by using SPSS19.0 software. Since the nonlinear parameters have non-Gaussian population, nonparametric Mann-Whitney rank sum test was used to assess the parameters. The significance probability was preset at 0.05.

TABLE 1. Summary statistics for all measures in ADSD, VN and NF.

measures	ADSD		Vocal Nodules		Normal Female (NF)		NF vs ADSD	NF vs VN	ADSD vs VN
	mean	SD	mean	SD	mean	SD			
mean_NHR	0.2040	0.2261	0.0462	0.0221	0.0154	0.0119	0.0003**	0.0004**	0.0981
mean_HNR	12.1891	5.4384	15.2141	2.6562	20.1689	2.9378	0.0001**	0.0003**	0.1415
GNE_mean	0.7515	0.2294	0.9269	0.0404	0.9634	0.0106	0.0000**	0.0007**	0.0028*
GNE_std	0.2025	0.2136	0.0255	0.0185	0.0124	0.0074	0.0001**	0.0088*	0.0077*
GNE_TKEO_SNR	1.0189	0.2142	1.0675	0.1909	1.0808	0.1690	0.5582	0.4568	0.6623
GNE_SEO_SNR	1.2549	0.2511	1.2918	0.1996	1.3499	0.1922	0.3284	0.5496	0.4969
EMD_ENTO_SNR	2.9419	0.2732	2.9657	0.2296	3.3721	0.0352	0.0000**	0.0001**	0.7290
RPDE	0.5708	0.0914	0.4520	0.0668	0.3347	0.0592	0.0000**	0.0002**	0.0009**
D2	2.5717	0.3587	2.0970	0.6341	1.9459	0.2803	0.0001**	0.8542	0.0088**
PE	0.6003	0.0889	0.6819	0.1017	0.5892	0.0371	0.9268	0.0051*	0.0024*
DFA	0.6881	0.1411	0.7408	0.1194	0.5224	0.0826	0.0024*	0.0001**	0.4907

Entries marked (\*) are significant at the 95% level; Entries marked (\*\*) are significant at the 90% level

TABLE 2. Summary statistics for all measures in UVFP and Normal Male (NM).

measures	UVFP		Normal Male (NM)		Normal male (NM) vs UVFP	UVFP vs ADSD
	mean	SD	mean	SD		
mean_NHR	0.3325	0.3177	0.0289	0.02	0.0094*	0.2748
mean_HNR	8.087	6.6439	17.3	2.2663	0.0094*	0.088
GNE_mean	0.6297	0.3625	0.9394	0.0252	0.0003**	0.1152
GNE_std	0.0532	0.1088	0.0224	0.0105	0.8252	0.0140*
GNE_TKEO_SNR	1.0341	0.3477	1.0672	0.181	0.8738	0.7654
GNE_SEO_SNR	1.3556	0.2425	1.317	0.184	0.3359	0.0038 *
EMD_ENTO_SNR	2.6918	0.0418	3.3796	0.0491	0.0003**	0.1102
RPDE	0.6168	0.0703	0.4461	0.098	0.0029*	0.088
D2	2.616	0.8858	1.7826	0.0883	0.0013*	0.7528
PE	0.7285	0.1114	0.5888	0.0468	0.0092*	0.002*
DFA	0.9865	0.1536	0.6999	0.1103	0.0023*	0.0008**

Entries marked (\*) are significant at the 95% level; Entries marked (\*\*) are significant at the 90% level

### 3. Results

The summary statistics for the energy distribution analysis, glottal vibration analysis and nonlinear dynamical analysis of the voice samples are given in tables 1 and 2. Both mean and standard deviation values of the parameters with significant statistical differences are shown.

Results of Mann-Whitney rank sum test revealed significant differences between the NF and ADSD in energy distribution measures (mean\_NHR ( $p < 0.001$ ), mean\_HNR ( $p < 0.001$ ), GNE\_mean ( $p < 0.001$ ), GNE\_std ( $p < 0.001$ ), EMD\_ENTO\_SNR ( $p < 0.001$ )) and some nonlinear dynamic measures (RPDE ( $p < 0.001$ ), D2 ( $p < 0.001$ ), DFA ( $p < 0.05$ )). This study also found significant differences between the NF and VN in measures (mean\_NHR ( $p < 0.001$ ), mean\_HNR ( $p < 0.001$ ), GNE\_mean ( $p < 0.001$ ), GNE\_std ( $p < 0.05$ ), EMD\_ENTO\_SNR ( $p < 0.001$ ), RPDE ( $p < 0.001$ ), PE ( $p < 0.05$ ) and DFA ( $p < 0.001$ )). Some measures revealed a significant main effect between ADSD and VN (GNE\_mean ( $p < 0.05$ ), GNE\_std ( $p < 0.05$ ), RPDE ( $p < 0.001$ ), D2 ( $p < 0.001$ ) and PE ( $p < 0.005$ )).

In addition, we compared all the parameters between UVFP and NM, finding significant difference in measures [mean\_NHR ( $p < 0.05$ ), mean\_HNR ( $p < 0.005$ ), GNE\_mean ( $p < 0.001$ ), EMD\_ENTO\_SNR ( $p < 0.001$ ), RPDE ( $p < 0.05$ ), D2 ( $p < 0.05$ ), PE ( $p < 0.05$ ) and DFA ( $p < 0.05$ )]. In UVFP and ADSD, the significant difference changes were found in GNE\_std ( $p < 0.05$ ), GNE\_SEO\_SNR ( $p < 0.05$ ), PE ( $p < 0.005$ ) and DFA ( $p < 0.001$ )].

### 4. Discussion

According to previous studies, the energy distribution analysis and nonlinear dynamic analysis appeared to be more effective in analyzing aperiodic voice signals, when compared with traditional perturbation analysis, because the former analyses are independent of the temporal domain [13-16]. Yet, perturbation analysis relies heavily on a linear network, which has been well defined in calculating the linear signals. However, the ADSD voice was notoriously aperiodic and the associated pitch and period information was difficult to extract. As such, perturbation analysis seems inapplicable.

#### 4.1. The comparison between ADSD and NF

The values of mean\_NHR, mean\_HNR, GNE\_mean, GNE\_std, EMD\_ENTO\_SNR, RPDE, D2 and DFA were significantly different between ADSD and normal female (NF) voices, indicating that the ADSD voices were significantly irregular and aperiodic when compared with NF voices. This finding is also supported by previous studies [3, 5, and 10]. Although GNE\_TKEO\_SNR and GNE\_SEO\_SNR parameters in table 1 revealed no significant differences between ADSD voices and NF voices, as indicated by their mean and standard deviations of values, ADSD voices appear to be inferior to normal voice. This is mainly related to the vocal spasm that disrupts normal vocal cord movement [1, 2] in ADSD patients. Such disruption in normal vocal fold movement in ADSD patients caused intermittent voice breaks and uncontrolled movements of vocal fold vibration led to the elevated noise in the voice, as well as and chaos of the voice signals. The previous studies of PE measures indicated that, as PE represented the complexity, it was correlated with irregular, disordered and aperiodicity of time series [24, 27, 31 and 32]. While the PE parameters listed in table 1 show no significant difference between ADSD and NF voices, the mean and standard deviation values of PE associated with ADSD voices were higher than those with normal voices. This indicates that ADSD voices were more complex than normal voices. This is consistent with findings reported previously [3-6].

#### 4.2. The comparison between ADSD and VN

As shown in table 1, the quantitative results from acoustic analysis including GNE\_mean, GNE\_std, RPDE, D2 and PE revealed significant differences between ADSD voices and VN voices. Moreover, the standard deviation values of these parameters associated with ADSD voices were higher than those with VN voices, revealing that the ADSD signals were more chaotic and irregular than VN signals. It follows that these parameters might be effective in distinguishing ADSD voices from VN voices. The discrepancies between ADSD and VN voices may be related to the differential pathological conditions. Recall that ADSD is a neurological disorder of the larynx [1, 2] that disrupts normal vocal fold vibration, and VN is an organic disorder which involves an additional tissue growing on the vocal cord [33]. Besides, the parameters such as mean\_NHR, mean\_HNR, GNE\_mean, GNE\_std, RPDE and DFA were not only significantly different between ADSD voice and NF voice, they were significantly different between ADSD and VN voices as well. Accordingly, the values of mean\_NHR, mean\_HNR, GNE\_mean, GNE\_std, RPDE and DFA could be used to distinguish the pathological voices from normal voices. Although the average of mean\_NHR in ADSD voices (mean = 0.2040) was largely different from that of VN voices (mean = 0.0462), statistically there was no significant difference between ADSD and VN voices. The lack of significance might be related to the use of the Mann-Whitney rank sum test as the statistical analysis test. Original data were not used, instead, ranking of the original data were used to calculate the statistical significance. It follows that, even if the mean value of mean\_NHR has shown significant difference between ADSD and VN, but there was no significant difference between ADSD and VN according to the mean\_NHR parameter.

#### 4.3. The comparison between ADSD and UVFP

According to table 2, significant differences in GNE\_std and GNE\_SEO\_SNR parameters which indicated the Signal-to-Noise-Ratio in different voice categories have been found when comparing between ADSD and UVFP voices. The PE parameter and DFA parameter that represented the chaotic features of voice signals were significantly different between ADSD and UVFP voices. Such results are likely to be related to the different vocal cord movement in ADSD patients (the spasms that cause the disorder of vocal cord movement [1, 2]) and UVFP patients (the damage in one recurrent laryngeal nerve disabling the adductory function of the vocal cords [13, 15, and 35]). Table 2 also shows the irregularity of vocal fold vibration in UVFP patients through the parameters: mean\_NHR, mean\_HNR, GNE\_mean, EMD\_ENTO\_SNR, RPDE, D2, PE and DFA. These findings are consistent with those reported previously [14, 15, and 22]. However, it should be noted that the ADSD patients were different from the UVFP patients in the present study. Nine UVFP patients (three females and six males) and 14 ADSD patients (all were females) were studied in the present investigation. The marked difference in gender allocation and the number of patients between the different vocal pathologies may be factors affecting the present results. And the extreme gender mismatch is the reason why we didn't compare UVFP and NF.

To obtain a more comprehensive description of ADSD voices and to distinguish between ADSD and VN voices using acoustic analyses, future studies should make use of more ADSD and VN speech samples, and perhaps incorporate more nonlinear dynamical measures into the categorization protocol in order to improve the reliability and availability of the protocol. Furthermore, future research should include the inherent characteristics of rough phonation in a more complete scheme and other cardinal vowels, such as the /i/, /ɔ/ and /u/ should also be considered.

### 5. Conclusions

The present study investigated the possible use of energy distribution analyses and nonlinear dynamic measures as tools to evaluate the voice quality of ADSD voices. The mean\_NHR, mean\_HNR, GNE\_mean, GNE\_std, EMD\_ENTO\_SNR, RPDE, D2, and DFA might be effective in evaluating ADSD voices. Results indicated that the parameters, including mean\_NHR, mean\_HNR, GNE\_mean, EMD\_ENTO\_SNR, RPDE, and DFA could be used to effectively distinguish different pathological voices (ADSD, VN, and UVFP) from normal voices. GNE and D2 seem potentially useful in identifying ADSD from other pathological voices, such as VN and UVFP. In conclusion, the acoustical measures proposed can significantly reflect the pathological characteristics of ADSD voices and the difference among ADSD, VN, and UVFP voices.

### 6. Acknowledgements

This study was jointly supported by a grant from National Natural Science Foundation of China (NSFC 61135003, 91420301 and 61401452), Shenzhen Speech Rehabilitation Technology Laboratory and Health and Health Services Research Fund (HHSRF), Food and Health Bureau of Hong Kong (Project Number: 09101191), Shenzhen Fundamental Research Program JCYJ20130401170306806.

## 7. References

- [1] M. Thomas, "Spasmodic dysphonia: Let's look at that again," *Journal of Voice*, vol. 28, no. 6, pp. 694-699, 2014.
- [2] C. L. Ludlow, "Spasmodic Dysphonia: a Laryngeal Control Disorder Specific to Speech," *Journal of neuroscience*, vol. 31, no. 3, pp. 793-797, 2011.
- [3] D. I. Siemons-Lühring, M. M. Jean-Pierre, M. D. D. Frank Müller, and P. D., "Spasmodic dysphonia, perceptual and acoustic analysis: presenting new diagnostic tools," *European Archives of Oto-Rhino-Laryngology*, vol. 266, no. 12, pp. 1915-1922, 2009.
- [4] K. Simonyan, C. L. Ludlow, and A. O. Vortmeyer, "Brainstem pathology in spasmodic dysphonia," *Laryngoscope*, vol. 120, no. 1, pp. 121-124, 2010.
- [5] C. Zhen, "A preliminary study of perceptual diversity in adductor spasmodic dysphonia," *Master of Arts thesis, University of Iowa*, 2011.
- [6] E. M. G. Fabron, V. C. de. C. Marino, and T. de. C. Nobile, et al, "Medical treatment and speech therapy for spasmodic dysphonia: a literature review," *CEFAC, Mai-Jun*, vol. 15, no. 3, pp. 713-724, 2013.
- [7] G. Dursun, S. Boynukalin, O. B. Ozgursoy, and I. Coruh, "Long-term results of different treatment modalities for glottic insufficiency," *American Journal of Otolaryngology*, vol. 29, no. 1, pp. 7-12, 2008.
- [8] P. H. Dejonckere, K. J. Neumann, M. B. J. Moerman, J. P. Martens, A. Giordano, and C. Manfredi, "Tridimensional assessment of adductor spasmodic dysphonia pre- and post-treatment with Botulinum toxin," *European Archives of Oto-Rhino-Laryngology*, vol. 269, no. 4, pp. 1195-1203, 2012.
- [9] C. M. Sapienza, T. Murry, WS. Jr. Brown, "Variations in adductor spasmodic dysphonia acoustic evidence," *Journal of voice*, vol. 12, no. 2, pp. 214-222, 1998.
- [10] V. Hartmann, C. J. Hartmann, H. Hefter, W. Angerstein, "Correlation of self-assessment with expert rating and acoustic analysis for spasmodic dysphonia treatment with botulinum neurotoxin A," *Basal Ganglia*, vol. 3, no. 3, pp. 165-169, 2013.
- [11] C. Watts, R. Whurr, and C. Nye, "Botulinum toxin injections for the treatment of spasmodic dysphonia," *Cochrane Database Systems*, no. 3, CD004327, 2004.
- [12] I. R. Titze, "Workshop on acoustic voice analysis: Summary statement," *National Center for Voice and Speech*, 1995.
- [13] N. Yan, M. L. Ng, D. Wang, L. Zhang, V. Chan, and R. S. Ho, "Nonlinear dynamical analysis of laryngeal, esophageal, and tracheoesophageal speech of Cantonese," *Journal of voice, official journal of the Voice Foundation*, vol. 27, pp. 101-10, 2013.
- [14] N. Yan, M. L. Ng, and L. Wang, "Acoustical analysis of voices produced by Cantonese patients of unilateral vocal fold paralysis acoustical analysis of voices by Cantonese UVFP," *Signal Processing, Communication and Computing, 2013 IEEE International Conference on, Aug*, pp. 1-5, 2013.
- [15] Y. Zhang, J. J. Jiang, L. Biazzo, and M. Jorgensen, "Perturbation and nonlinear dynamic analyses of voices from patients with unilateral laryngeal paralysis," *Journal of Voice*, vol. 19, no. 4, pp. 519-528, 2005.
- [16] Y. Zhang, C. McGilligan, L. Zhou, M. Vig, and J. J. Jiang, "Nonlinear dynamic analysis of voices before and after surgical excision of vocal polyps," *The Journal of the Acoustical Society of America*, vol. 115, no. 5, pp. 2270-2277, 2004.
- [17] D. Michaelis, T. Gramms, and H. W. Strube, "Glottal-to-Noise Excitation Ratio-a New Measure for Describing Pathological Voices," *Acta. Acustica*, vol. 83, no. 4, pp. 700-706, 1997.
- [18] A. Tsanas, M. A. Little, P. E. Mcsharry, and O. R. Lorraine, "Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity," *The Royal Society*, vol. 8, no. 59, pp. 842-855, 2011.
- [19] D. Dimitriadi, A. Potamianos, and P. Maragos, "A Comparison of the Squared Energy, and Teager-Kaiser Operators for Short-Term Energy Estimation in Additive Noise," *Signal Processing, IEEE Transactions*, vol. 57, no. 7, pp. 2569-2581, 2009.
- [20] A. N. Patrick, A. Kounoudes, J. Gudnason, and M. Brookes, "Estimation of Glottal Closure Instants in Voiced Speech Using the DYPSA Algorithm," *Audio, Speech and Language Processing, IEEE Transa*, vol. 15, no. 1, pp. 34-43, 2007.
- [21] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, and Q. Zheng, et al, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *The Royal Society*, vol. 454, no. 1971, pp. 903-995, 1998.
- [22] S. S. Liu, N. Yan, M. L. Ng, L. Wang, and Z. J. Wang, "Multidimensional acoustic analysis for evaluation of voice quality of unilateral vocal fold paralysis," *Information Science and Technology, 2014 4th IEEE International Conference on*, pp. 706-709, 2014.
- [23] C. K. Peng, S. Havlin, H. E. Stanley, and A. L. Goldberger, "Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 5, no. 1, pp. 82-87, 1995.
- [24] C. Bandt and B. Pompe, "Permutation entropy: a natural complexity measure for time series," *Physical review letters*, vol. 88, no. 17, 174102, 2002.
- [25] C. Bandt, "Ordinal time series analysis," *Ecological Modelling*, vol. 182, pp. 229-238, 2005.
- [26] K. Keller and M. Sinn, "Ordinal analysis of time series," *Physica A: Statistical Mechanics and its Applications*, vol. 356, no. 1, pp. 114-120, 2005.
- [27] A. A. Bruzzo, B. Gesierich, M. Santi, C. Tassinari, N. Birbaumer, and G. Rubboli, "Permutation entropy to detect vigilance changes and preictal states from scalp EEG in epileptic patients: A preliminary study," *Neurological sciences*, vol. 29, no. 1, pp. 3-9, 2008.
- [28] L. I. e Juan Godino, P. Gimez-vilda, N. Sienz-lechun, M. V. Blanco, F. Cruz-rolden, and M. A. Ferrer, "Discriminative methods for the detection of voice disorders," *Proceedings of International Conference on Non-Linear Speech Processing (NOLISP '05)*, Barcelona, Spain, vol. 19, no. 4, pp. 158-167, 2005.
- [29] M. B. Kennel, R. Brown, and H. D. I. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction," *Physical Review A*, vol. 45, pp. 3403-3411, 1992.
- [30] H. S. Kim, R. Eykholt, and J. D. Salas, "Nonlinear dynamics, delay times, and embedding windows," *Physica D: Nonlinear Phenomena*, vol. 127, pp. 48-60, 1999.
- [31] M. Matilla-Garcia, M. R. Marin, "A non-parametric independence test using permutation entropy," *Journal of Econometrics*, vol. 144, no. 1, pp. 139-155, 2008.
- [32] U. Nair, B. M. Krishna, V. N. N. Nambhothiri, and V. P. N. Nampoori, "Permutation entropy based real-time chatter detection using audio signal in turning process," *International Journal of Advanced Manufacturing Technology*, vol. 46, pp. 61-68, 2010.
- [33] R. Shah, H. Feldman, and R. Nuss, "A grading scale for pediatric vocal fold nodules," *Otolaryngology Head and Neck Surgery*, vol. 136, no. 2, pp. 193-198, 2007.
- [34] P. Boersma and D. Weenink. (2009). Praat: doing phonetics by computer (Version 5.1.05). Available: <http://www.praat.org/>.
- [35] D. P. Lau, G. A. Lee, S. M. Wong, V. P. Lim, Y. H. Chan, and N. G. Tan, et al, "Injection laryngoplasty with hyaluronic acid for unilateral vocal cord paralysis. randomized controlled trial comparing TWO different particle sizes," *Journal of Voice*, vol. 24, no. 1, pp. 113-118, 2010.