DOI: 10.61782/fa.2023.1188

forumacusticum 2023

CAN DIFFERENT VOICE ACOUSTIC PARAMETERS PREDICT BETTER VOICE PATHOLOGY TYPES?

Lady Catherine Cantor-Cutiva [1-2*] **Eric J. Hunter** [1-2] **Aishwarya Ramani** [1]
[1] Department of Communicative Sciences and Disorders, Michigan State University, MI, USA
[2] Department of Communicative Sciences and Disorders, University of Iowa, Iowa, USA

ABSTRACT

This study aimed to identify the cluster of parameters that effectively distinguish diverse types of voice pathologies using voice samples from the Perceptual Voice Qualities Database (PVQD) with 296 participants. We used Principal Component Analysis (PCA) as a data analysis technique for identifying the parameters that better differentiate voice pathologies. Voice acoustic analysis was conducted using custom MATLAB scripts and PRAAT, while PCA analysis utilized SPSS. Two main results were found. First, Principal Component Analysis showed that Cepstral Peak Prominence Smoothed, shimmer, Harmonics-to-Noise Ratio, and Pitch Period Entropy were key components with high explanatory power in identifying pathological voices from the normal controls. Secondly, Generalized Estimating Equations with a multinomial distribution demonstrated a small yet statistically significant increase in Harmonics-to-Noise Ratio and Pitch Period Entropy in healthy voices, whereas shimmer showed a smaller effect. In conclusion, our analysis highlights that Harmonics-to-Noise Ratio, Pitch Period Entropy, and shimmer are parameters that show strong diagnostic properties in distinguishing pathological voices from normal ones. Integration of these analysis techniques into future large data or machine learning analysis protocols holds promise for further advancements in the field.

Keywords: *voice, voice disorders, principal component analysis*

[*] ***Corresponding author***: *lcantorcutiva@uiowa.edu*

1. INTRODUCTION AND METHODS

This section will provide an overview of the Perceptual Voice Qualities Database (PVQD), the importance of voice acoustic analysis, the significance of Principal Component Analysis (PCA) and Generalized Estimating Equations (GEEs) in analyzing the data, and the specific voice acoustic parameters used in the study.

The PVQD is a database funded by the Voice Foundation and includes voice samples from 296 participants, consisting of 195 females and 101 males, with an age range between 14 and 93 years old. The database is freely available and has been utilized in numerous studies related to voice perception and analysis [1]. However, to the best of the authors' knowledge, no previous studies have examined the diagnostic properties of voice acoustic parameters for identifying diverse types of voice pathologies using this database.

Voice acoustic analysis refers to the use of computerized techniques to measure and analyze the acoustic properties of the human voice, including pitch, intensity, and spectral characteristics. Therefore, these parameters can provide objective information about various aspects of vocal performance, such as vocal quality, loudness, and articulation. Some common acoustic parameters include fundamental frequency (fo), jitter, shimmer, harmonics-to-noise ratio (HNR), cepstral peak prominence smoothed (CPPS), and pitch period entropy (PPE). fo corresponds to the number of cycles per second of the vocal folds during phonation, jitter represents the instabilities in the oscillating pattern of the vocal folds, shimmer quantifies the cycle-to-cycle alterations in amplitude, HNR is derived from the signal-to-noise estimates from the autocorrelation of each cycle, and PPE quantifies the impaired control of fundamental frequency using a logarithmic scale [2-4].

Principal Component Analysis (PCA) is a statistical technique used to identify patterns and relationships among variables in a data set. It is a dimensionality reduction method that transforms a large number of variables into a smaller number of components, while still retaining most of the variability present in the original data set [5]. PCA is commonly used for exploratory data analysis and predictive modeling, allowing for dimensionality reduction while retaining as much variation as possible. This method provides the weights for each variable needed to get a new variable that best explains the variation in the whole dataset in a certain sense. This new variable, including the defining weights, is called the principal component (PC) [6].

Generalized estimating equations (GEEs) are a statistical methodology used for analyzing data that has a correlated structure, such as repeated measurements or clustered data. GEEs are a type of regression analysis that considers the correlation among observations within the same individual or cluster [7].

Custom MATLAB and PRAAT scripts were used to calculate eleven voice acoustic parameters: jitter, shimmer, HNR, CPP, CPPS, Alpha ratio, PPE, voice Sound Pressure Levels (SPL), fo, SPL standard deviation (SD), and fo SD. To streamline voice assessment using acoustic parameters, a two-step approach was used. First, PCA was conducted to identify the principal components with high explanatory power in differentiating pathological voices from normal controls, potentially speeding up voice assessment. Second, GEEs were performed to evaluate the association between voice acoustic parameters and disordered voices, further enhancing our ability to identify and assess several types of pathologies using acoustic parameters.

2. RESULTS AND CONCLUSION

Figure 1 shows the results of the variance analysis for the PCA showing that the Alpha ratio and CPP explained around 50% of the variance of perceptually identified voice disorders. Moreover, CPPs, shimmer, HNR, PPE, and jitter are components with high explaining values to identify pathological voices from the normal controls.

As shown in Figure 2, four principal components (PC) were calculated. PC#1 had a statistically significant lower mean for normal voice compared with other voice pathologies, such as adductory dysphonic dysphonia (Add. SD) and vocal fold paresis (VFP). PC#2 mean was statistically significantly smaller for VFP compared with other voice pathologies such as Add.SD, lesions, Parkinson's disease, and vocal fold scarring (VFS).

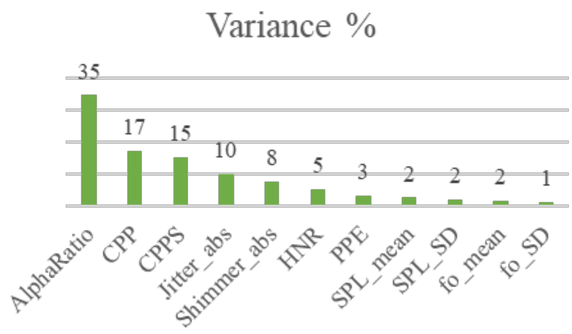

Figure 1. Explained variance of voice acoustic parameters.

PC#3 was statistically significantly bigger for VFS compared with other voice pathologies, such as lesions and VFP. PC#4 was statistically significantly bigger for Parkinson's disease compared with other voice pathologies and normal voice.

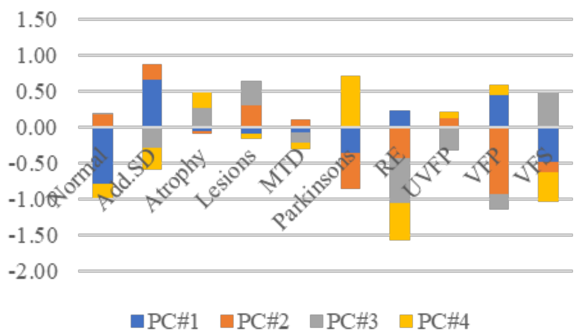

Figure 2. Principal Component Analysis per voice pathology.

Table 1 shows the results of the GEEs with a multinomial distribution for HNR. The GEE revealed that the largest decrease in the HNR was observed in Adductory Spasmodic Dysphonia (Beta= -11.56) and Vocal Fold Paresis (Beta= -8.57).

Table 1. Generalized Estimating Equations for assessing the association between Harmonics-to-Noise Ratio and voice pathologies.

Voice Acoustic Parameter	HNR		
	Beta	**SE**	**P-value**
Intercept	21.32	0.73	0.00
Male	Reference Category		
Female	0.14	0.87	0.88
Normal	Reference Category		
Add. SD	-11.56	2.08	0.00
Atrophy	-5.46	0.92	0.00
Lesions	-5.08	0.93	0.00
MTD	-6.09	0.88	0.00
Parkinson's	-2.79	1.42	0.05
RE	-3.91	1.02	0.00
UVFP	-7.12	1.09	0.00
VFP	-8.57	2.78	0.00
VFS	-1.97	2.09	0.35

The GEEs with multinomial distribution results for PPE are presented in Table 2, where it is shown that the most significant increase in PPE was found in Adductory Spasmodic Dysphonia (Beta=0.32) and Unilateral Vocal Fold Paralysis (Beta=0.12).

Table 2. Generalized Estimating Equations for assessing the association between Pitch Period Entropy and Voice Pathologies.

Voice Acoustic Parameter	PPE		
	Beta	**SE**	**P-value**
Intercept	0.45	0.01	0.00
Male	Reference Category		
Female	-0.01	0.02	0.40
Normal	Reference Category		
Add. SD	0.32	0.01	0.00
Atrophy	0.11	0.02	0.00
Lesions	0.09	0.02	0.00
MTD	0.10	0.02	0.00
Parkinson's	0.09	0.02	0.00
RE	0.01	0.02	0.63
UVFP	0.12	0.02	0.00
VFP	0.09	0.04	0.02
VFS	0.02	0.03	0.63

The results of the GEEs with a multinomial distribution for Shimmer are presented in Table 3. The analysis revealed that the largest increase in Shimmer was observed in Adductory Spasmodic Dysphonia (Beta=0.05) and RE (Beta=0.04), indicating that these conditions are associated with a higher perturbation in amplitude compared to the other pathologies evaluated.

Table 3. Generalized Estimating Equations for assessing the association between shimmer and voice pathologies.

Voice Acoustic Parameter	Shimmer		
	Beta	**SE**	**P-value**
Intercept	0.07	0.00	0.00
Male	Reference Category		
Female	-0.01	0.00	0.08
Normal	Reference Category		
Add. SD	0.05	0.01	0.00
Atrophy	0.02	0.00	0.00
Lesions	0.02	0.01	0.00
MTD	0.02	0.00	0.00
Parkinson's	0.01	0.00	0.07
RE	0.04	0.00	0.00
UVFP	0.03	0.01	0.00
VFP	0.03	0.01	0.00
VFS	0.01	0.01	0.31

The findings suggest that PC#1, which includes five parameters (CPPS, shimmer, HNR, PPE, and jitter) can effectively differentiate pathological voices from normal controls.

The results of the GEEs with a multinomial distribution show that the analyzed acoustic parameters, including HNR, PPE, and Shimmer, are associated with specific voice disorders. In particular, Adductory Spasmodic Dysphonia is associated with a significant decrease in HNR, an increase in PPE, and higher values of Shimmer compared to other pathologies evaluated.
Vocal Fold Paralysis was associated with a significant decrease in HNR, while Unilateral Vocal Fold Paralysis showed a significant increase in PPE. Additionally, RE was associated with higher values of Shimmer compared to other pathologies evaluated.

3. ACKNOWLEDGMENTS

The authors would like to express their gratitude to the authors of the PVQD. Without their cooperation and willingness to share their experiences, this research would not have been possible.

This research was partly supported by the NIDCD of the NIH under Award Number R01DC012315. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

4. REFERENCES

[1] Walden, P. R. (2022). Perceptual voice qualities database (PVQD): database characteristics. Journal of Voice, 36(6), 875-e15.

[2] Little, M., McSharry, P., Hunter, E., Spielman, J., & Ramig, L. (2008). Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. Nature Precedings, 1-1.

[3] Abdurrahman, G., & Sintawati, M. (2020, May). Implementation of xgboost for classification of Parkinson's disease. In Journal of Physics: Conference Series (Vol. 1538, No. 1, p. 012024). IOP Publishing.

[4] Hillenbrand, J. M. (2011). Acoustic analysis of voice: a tutorial. Perspectives on Speech Science and Orofacial Disorders, 21(2), 31-43.

[5] Abdi, H., & Williams, L. J. (2010). Principal component analysis. Wiley interdisciplinary reviews: computational statistics, 2(4), 433-459.

[6] Bro, R., & Smilde, A. K. (2014). Principal component analysis. Analytical methods, 6(9), 2812-2831.

[7] Hardin, J. W., & Hilbe, J. M. (2012). Generalized estimating equations. CRC press.