REVIEW ARTICLE

VOICE AS A DIAGNOSTIC TOOL

ABSTRACT

Detection of diseases using voice analysis is a current research topic in medical engineering. It is a reliable, efficient, economic, and easy to use method. It also helps to detect the disease at its earlier stage. Voice analysis is done with the help of tools such as electroglottography, acoustic analysis, and spectral analysis. Voice analysis has been used in diagnosing conditions such as Parkinson's disease, Alzheimer's disease, mild cognitive impairment, and attention deficit hyperactivity disorder. This review provides an overview of voice analysis, tools used for the analysis of voice, applications of voice analysis in diagnosing medical conditions and potential prospects in the field of dentistry,

Keywords: Voice analysis, acoustic analysis, Parkinson's disease, Alzheimer's disease, cognitive impairment, Oral cancer.

Author:

[1] Adette Mariyam Alexander
[2] Afnan Rahim
[3] Alifna M M
[4] Ammu P Paul
[5] Subramaniam R
[6] Jesline Merly James
[7] Suneesh Kuruvilla

[1,2,3,4]House Surgeon
Indira Gandhi Institute of Dental Sciences
Nellikuzhi P. O., Kothamangalam 686 691, Kerala

[5]Professor and Head
Department of Public Health Dentistry
Indira Gandhi Institute of Dental Sciences
Nellikuzhi P. O., Kothamangalam 686 691, Kerala

[6,7]Senior Lecturer
Department of Public Health Dentistry
Indira Gandhi Institute of Dental Sciences
Nellikuzhi P. O., Kothamangalam 686 691, Kerala

Corresponding author
Dr. Subramaniam R
Professor and Head
Department of Public Health Dentistry
Indira Gandhi Institute of Dental Sciences
Nellikuzhi P.O., Kothamangalam 686 691, Kerala
Email: subbds@gmail.com

J Odontol Res 2021;9(1)18-23.

INTRODUCTION

Voice or vocalization refers to the sound that humans and other vertebrates produce using the lungs and vocal folds. Everyone has a unique voice as their fingerprint. It helps in defining personality, mood, and health. The disorders of voice mainly involve troubles with pitch, loudness, and quality.[1]

Speech recognition software has been used to detect spoken words and turn them into text or respond to commands. In today's world, as researchers are been developing, they are been looking at the voice acoustic qualities to find out about the medical or even the emotional state of the speaker. Interest has been put into the emerging software which detects persons with neurological disorders.[2]

Voice analysis software

Detection of diseases using voice analysis is a current research topic in medical engineering. It is a reliable, efficient, economic, and easy to use method. It also helps to detect the disease at its earlier stage.

Various features are extracted using Digital Signal Processing (DSP) techniques. These features contain information on the health of the voice tract and of organs responsible for speech. These features represent the voice and may be used to discriminate the voice of healthy and unhealthy persons. The spectrum analysis, glottal waveform analysis are some forms used to extract the voice features. These features are then classified using various classification techniques like vector quantization, dynamic time wrapping, support vector machine, Gaussian mixture model, and artificial neural network.[3]

Mechanism

The act of speaking is a complex phenomenon as it involves the movement of the tongue, lips and jaws. There is coordination and movement of many structures at the same time. So thoughts have been given into adapting already existing advanced speech recognition technologies and put them to work for this task.

Using databases of human speakers featuring voices from around the world, Moro Velazquez, a research scientist at the Center for Language and Speech Processing (CLSP), created algorithms that search for irregularities in how they speak. This approach gives a precise diagnosis in the range of 90-95 percent. The main obstacle is consistency. When using certain databases, his diagnosis rate has been closer to 65-70 percent.[3]

Tools used for voice analysis

a. Electroglottography (E.G.G.)

Electroglottography (EGG) is a non-invasive technique that registers laryngeal behavior indirectly by measuring the change in electrical impedance across the throat during the act of speaking.[4] The method was first developed by Fabre (1957) and influential contributions are credited to Fourcin (1971 with Abberton)[5] and Frokjaer-Jensen (1968 with Thorvaldsen).[4,6]

Limitations of EGG

• Difficulties in obtaining signals from women and children than men because of the smaller size of the vocal fold, wider angle of the thyroid cartilage.

• Signal of rapid movement of vocal fold tends to superimpose on the signals produced by slower movements of other structures.

• High chances of signal distortion.

Despite these problems, EEG has established itself as a valuable tool for evaluating laryngeal behavior. The EGG is superior to all other methods in that it is completely non-invasive (it exerts no influence at all on the articulation and production of sounds).[6]

b. Acoustic analysis

The acoustic speech signal is the system's output and is rich in information about pitch, loudness, and quality. The signal is complex and is therefore broken down for analysis into the dimensions of frequency, amplitude, and time. Acoustic measurements are made from signals recorded on high-intensity equipment in a quiet environment and with the use of standard instructions.[7]

Three parameters, namely Peak Slope, Normalized Amplitude Quotient, and Cepstral Peak

Prominence, are used to assess voice quality changes in patients with vocal fold nodules. Peak Slope (PS) allows differentiation of breathy, modal and tense voice, Normalized Amplitude Quotient (NAQ) is used to separate the types of phonation effectively and Cepstral Peak Prominence (CPP) enables the early detection of dysphonia.[4]

Data should be interpreted in light of subject age and gender, interactions between parameters, speech sample selected for analysis, and several tokens. It is also important to know whether the samples were consistent with typical voice use or maximum or best performance.[7,8]

c. Spectral analysis [9]

Spectral analysis is used to find out how acoustic energy is distributed across frequency. It is typically used in phonetics to discover the spectral properties of the vowels and consonants of a language, comparing the productions of different speakers, or finding characteristics that point forward to speech perception or back to articulation. Earlier, the calculations were time-consuming.From 1950 onwards, this was done by the spectrograph, that burnt a spectrogram onto paper as a permanent record. Nowadays, a suitable computer program will calculate speech spectra in seconds. There are two methods for spectral analysis: the fast Fourier transform (FFT) and Linear Prediction (LP).[10,11] FFT finds the energy distribution in the actual speech sound, whereas LP estimates the vocal tract filter that shaped that speech. The advantage of FFT is easier setup, the disadvantage is difficulty identifying formants by speakers with higher-pitched voices. LP has better success with high-pitched voices, but the settings need to be carefully tuned for each speaker.[9–11]

Application in medical diagnosis

In medical diagnosis, voice analysis has made concrete progress. Voice analysis software is developed which can detect neurological diseases such as Parkinson's disease, attention deficit disorder. Researches are in progress for apps that can tell whether you are tied or in fact in depression.The technique that analyzes vocal patterns in Parkinson's disease helps for early diagnosis than is currently possible.[2]

A cross-sectional, cohort study conducted in Brazil reported the use of laryngeal electromyography in a large number of patients with Parkinson's Disease and vocal complaints grouped according to Parkinson's Disease severity. The patterns observed suggested that laryngeal electromyography is a valuable diagnostic tool for Parkinson's Disease even at early phases of the disease.[12]

Max Little, a mathematician and MIT research fellow who works in voice analysis developed vice analysis technology for Parkinson's disease. A healthy person's voice is strong and stable and for someone with Parkinson's, a tremor emerges. Using machine learning algorithm for detection of tremor and weakness in voice, using machine learning algorithm to detect tremor and weakness, Little developed a model to identify the vocal quality of Parkinson patient with an accuracy of around 99%.[2]

According to the present research analysis, sound vibrations, which were earlier employed as therapeutic healing for mental health related conditions, are now being applied into various disease diagnostics including cardiovascular and psychiatric ailments. After successful execution of clinical trials, Health Insurance Portability and Accountability Act (HIPAA)- compliant vocal biomarkers are being tried as effective non-invasive and safe alternatives to currently available disease diagnostic systems such as CT scan, MRI, and X-ray. Prominent voice-based companies such as IBM Corporation, Cogito Corporation, Audio Profiling, Sonde Health, and Beyond Verbal are focusing on sampling data for providing accurate results.[13]

Now, machine learning-based voice recognition technology, such as that developed for Amazon's voice home assistant, Alexa, are being utilized to identify voice patterns that are specific to different neurological diseases. And, as with Alexa, people can give voice samples from the comfort of their home. The biotech companies are working to bring to market technology to monitor a patient's health in the clinic or remotely, using smartphone apps or other wearables, with samples of voice. The hope of

the future is to use voice data to create non-invasive, inexpensive ways to track changes in symptoms and response to medication, tools called voice biomarkers.[14]

Little describes the usage of the telephone for voice analysis as "technologically convenient'' as around 75% of people around the world have access to phones. But the major disadvantage of this software is that it cuts out human subjectivity. Over this, the advantage of any kind of algorithm is that the entire process is repeatable and entirely objective. The stress anyhow this technology has to be used in a clinical context because you needed to have access to care.[2]

A growing body of work is focussing on the ongoing clinical validation of speech-based measures in a variety of clinical contexts. Speech has been demonstrated to have diagnostic validity for Alzheimer's disease (AD) and mild cognitive impairment (MCI) in studies using machine-learning classification models to differentiate individuals with AD/MCI from healthy individuals based on speech samples[15-18]. Additionally, speech analysis can detect individuals with depression[18-21], schizophrenia[18,22,23] autism spectrum disorder[18], and Parkinson's disease, and can differentiate the subtypes of primary progressive aphasia and frontotemporal dementia[18,24]. Classification models provide diagnostic validity for speech measures and could be used to develop tools for disease screening and diagnosis.

Langner a mathematician at Charite Hospital, Berlin developed an analytical model deep speech pattern analysis, which defines six different features of voice: loudness, tempo, rhythm, timbre, articulation, and melody. His current research focuses on diagnosing attention deficit hyperactivity disorder. Based on voice analysis he found differences in utterances of children with and without Attention Deficit Hyperactivity Disorder (ADHD), including fluctuation in speech loudness and melody. But he also says speech analysis should not be used as a standalone diagnostic tool, rather should be a tool to support a doctor.[2] Few other studies have indicated potential use of voice analysis for diagnosis of ADHD.[25,26]

DENTAL PROSPECTS

As the voice analysis software program is gaining great momentum in the field of medical science, the same can also be applied to dental science. It could be of significance for diseases or disorders of tongue and carcinoma involving the oropharynx region. Speech-language pathologists routinely assess oro-motor structure and function to identify contributing factors underlying a client's speech or swallowing disorders. A new examination involves 1hour speech-like tongue exercises (rapid syllable repetitions) that affect dysarthric speech.[27]

Oral squamous cell carcinoma and its treatment impair speech intelligibility by alteration of the vocal tract. It can be done by the means of an automatic, standardized speech recognition system.[28] Studies show the significance of voice analysis software to be an important tool in the diagnosis of laryngeal carcinoma. Laryngeal carcinoma that forms on the vocal cord often causes hoarseness or change in voice. This might lead to them being found at very early stages.[28]

Acoustic analysis can also be employed in making necessary modifications of dental prostheses, especially complete and partial removable dental prostheses to enhance speech clarity.[29,30]

CONCLUSION

Voice as a diagnostic tool is a current research trend in non-invasive diagnostic methods. The focus is on the early diagnosis of diseases enabling prevention. Voice signal analysis can be to help track the changes in medical conditions over time. Several related software is under research. Recognizing the importance of voice analysis, it will become more pervasive. This sort of technology paves way for future research.

REFERENCES

1. What Is Voice? What Is Speech? What Is Language? [Internet]. NIDCD. 2015 [cited 2021 May 5]. Available from: https://www.nidcd.nih.gov/health/what-is-voice-speech-language

2. Welle (www.dw.com) D. Voice analysis: an “objective” diagnostic tool based on flawed algorithms? | DW | 05.11.2013 [Internet]. DW.COM. [cited 2021 May 5]. Available from: https://www.dw.com/en/voice-analysis-an-objective-diagnostic-tool-based-on-flawed-algorithms/a-17187057

3. Saloni NA, Sharma RK, Gupta AK. Disease detection using voice analysis: a review. IJMEI. 2014;6(3):189.

4. Electroglottographic and acoustic analysis of voice in children with vocal nodules | Elsevier Enhanced Reader [Internet]. [cited 2021 May 12]. Available from: https://reader.elsevier.com/reader/sd/pii/S0165587619301570?token=DEFBB5FAB578AA3CC042D3C8FEE4505B7E4C7D0EC0C3149119406001F8C692AFBFA1867A5D2F86D6E2F3E43B43463BB0&originRegion=eu-west-1&originCreation=20210512150926

5. Fourcin AJ, Abberton E. First applications of a new laryngograph. Med Biol Illus. 1971 Jul;21(3):172–82.

6. EGG principles [Internet]. [cited 2021 May 5]. Available from: https://www2.ims.uni-stuttgart.de/EGG/frmst2.htm

7. Niebudek-Bogusz E, Fiszer M, Kotylo P, Sliwinska-Kowalska M. Diagnostic value of voice acoustic analysis in assessment of occupational voice pathologies in teachers. LogopedPhoniatrVocol. 2006;31(3):100–6.

8. Nasser R, Abolfazl S. An Introduction to Speech Sciences (Acoustic Analysis of Speech). Iranian Rehabilitation Journal. :10.

9. Spectral analysis [Internet]. Welcome to SWPhonetics. 2012 [cited 2021 May 11]. Available from: https://swphonetics.com/praat/tutorials/spectral-analysis/

10. Paez NFG, Garzon JSE. Performance evaluation of software for the spectral analysis of speech signals in a MIPS based architecture. Ingeniería y Desarrollo. 34(2):309–32.

11. Singh L. Speech Signal Analysis using FFT and LPC. 2015;4(4):3.

12. Zarzur AP, Duprat A de C, Cataldo BO, Ciampi D, Fonoff E. Laryngeal electromyography as a diagnostic tool for Parkinson’s disease. Laryngoscope. 2014 Mar;124(3):725–9.

13. Can Vocal Biomarkers Truly Emerge as a Medical Diagnostic Alternative? | Medgadget [Internet]. [cited 2021 May 11]. Available from: https://www.medgadget.com/2019/01/can-vocal-biomarkers-truly-emerge-as-a-medical-diagnostic-alternative.html

14. Talk About a Revolution: The Future of Voice Biomarkers in the Neurology Clinic - Memory and Brain Wellness Center [Internet]. [cited 2021 May 11]. Available from: http://depts.washington.edu/mbwc/news/article/talk-about-a-revolution-the-future-of-voice-biomarkers-in-the-neurology-clinic

15. Fraser KC, Meltzer JA, Rudzicz F. Linguistic Features Identify Alzheimer’s Disease in Narrative Speech. J Alzheimers Dis. 2016;49(2):407–22.

16. Konig A, Satt A, Sorin A, Hoory R, Derreumaux A, David R, et al. Use of Speech Analyses within a Mobile Application for the Assessment of Cognitive Impairment in Elderly People. Curr Alzheimer Res. 2018;15(2):120–9.

17. Themistocleous C, Eckerström M, Kokkinakis D. Voice quality and speech fluency distinguish individuals with Mild Cognitive Impairment from Healthy Controls. PLoS One. 2020;15(7):e0236009.

18. Robin J, Harrison JE, Kaufman LD, Rudzicz F, Simpson W, Yancheva M. Evaluation of Speech-Based Digital Biomarkers: Review and Recommendations. DIB. 2020;4(3):99–108.

19. Corcoran CM, Carrillo F, Fernández-Slezak D, Bedi G, Klim C, Javitt DC, et al. Prediction of psychosis across protocols and risk cohorts

using automated language analysis. World Psychiatry. 2018 Feb;17(1):67–75.

20. Taguchi T, Tachikawa H, Nemoto K, Suzuki M, Nagano T, Tachibana R, et al. Major depressive disorder discrimination using vocal acoustic features. J Affect Disord. 2018 Jan 1;225:214–20.

21. Bedi G, Carrillo F, Cecchi GA, Slezak DF, Sigman M, Mota NB, et al. Automated analysis of free speech predicts psychosis onset in high-risk youths. npj Schizophrenia. 2015 Aug 26;1(1):1–7.

22. Mota NB, Vasconcelos NAP, Lemos N, Pieretti AC, Kinouchi O, Cecchi GA, et al. Speech graphs provide a quantitative measure of thought disorder in psychosis. PLoS One. 2012;7(4):e34928.

23. Mota NB, Copelli M, Ribeiro S. Thought disorder measured as random speech structure classifies negative symptoms and schizophrenia diagnosis 6 months in advance. NPJ Schizophr. 2017;3:18.

24. Fraser KC, Meltzer JA, Graham NL, Leonard C, Hirst G, Black SE, et al. Automated classification of primary progressive aphasia subtypes from narrative speech transcripts. Cortex. 2014 Jun;55:43–60.

25. Garcia-Real T, Diaz-Roman TM, Garcia-Martinez V, Vieiro-Iglesias P. Clinical and acoustic vocal profile in children with attention deficit hyperactivity disorder. J Voice. 2013 Nov;27(6):787.e11-18.

26. Barona-Lleo L, Fernandez S. Hyperfunctional Voice Disorder in Children With Attention Deficit Hyperactivity Disorder (ADHD). A Phenotypic Characteristic? J Voice. 2016 Jan;30(1):114–9.

27. Solomon NP. ASSESSMENT OF TONGUE WEAKNESS AND FATIGUE. Int J Orofacial Myology. 2004 Nov;30:8–19.

28. Stelzle F, Knipfer C, Schuster M, Bocklet T, Nöth E, Adler W, et al. Factors influencing relative speech intelligibility in patients with oral squamous cell carcinoma: a prospective study using automatic, computer-based speech analysis. Int J Oral Maxillofac Surg. 2013 Nov;42(11):1377–84.

29. Adaki R, Meshram S, Adaki S. Acoustic Analysis and Speech Intelligibility in Patients Wearing Conventional Dentures and Rugae Incorporated Dentures. J Indian Prosthodont Soc. 2013 Dec;13(4):413–20.

30. Zakkula S. Evaluation of Palatal Plate Thickness of Maxillary Prosthesis on Phonation- A Comparative Clinical Study. JCDR [Internet]. 2014 [cited 2021 May 11]; Available from: http://www.jcdr.net/article_fulltext.asp?issn=0973-709x&year=2014&volume=8&issue=4&page=ZC11&issn=0973-709x&id=4224