

Assessing the Reliability of a Mobile Phone Recorder in Acoustic Voice Analysis: A Cross-sectional Study

Anagha A Joshi¹, Nishigandha V Nehete², Pushkaraj A Kulkarni³, Vaishnavi R Sangle⁴, Renuka A Bradoo⁵, Varun Dave⁶

Received on: 25 September 2022; Accepted on: 28 May 2023; Published on: 30 June 2023

ABSTRACT

Aim: To study the variability in acoustic voice parameters captured by different recording devices using the Dr Speech™ software program.

Materials and methods: A total of 60 participants were evaluated in a cross-sectional study for a period of 4 months from September to December 2019. A total of 25 of these participants had no voice abnormality, whereas 35 had a known voice disorder. Voices of study participants were recorded using an inbuilt mobile phone and the microphone provided by Dr Speech™ software company. The voice recorded by mobile phone recorder was converted in 16-bit, 44100 Hz, .wav format using Audacity™ software as it was the requirement to analyze voice in Dr Speech™ software. Both sets of voice recordings were then analyzed using the Dr Speech™ software.

Results: Voice recordings were made using the microphone provided by Dr Speech software company and compared with the recordings made using the mobile phone recorder. We observed no statistically significant difference between voice parameters from both recordings except in a few parameters like shimmer and harmonic to noise ratio (HNR). We conclude that there is a correlation between the acoustic voice parameters of voices recorded by the microphone of Dr Speech™ software company and the in-built recorder of a mobile phone except few. Using the mobile phone recorder, the voice parameters of study participants with normal voices were compared with those with voice disorders, and the difference was found to be statistically significant. This suggested that mobile phones could differentiate between normal and diseased voices.

Conclusion: The result of this cross-sectional study indicated that voices recorded using a mobile phone could be effectively used for voice analysis in certain difficult conditions.

Clinical significance: Patients can record their voices on their personal mobile phones and send the recording as an attachment to the treating laryngologist. This will be convenient and cost-saving, with no requirement for a patient to travel to the healthcare center repeatedly; it also helps in reducing the risk of transmission of any infection, especially coronavirus disease 2019 (COVID-19).

Keywords: Coronavirus disease 2019 infection, Dr Speech software, Microphone, Mobile phone recorder, Voice analysis.

International Journal of Phonosurgery & Laryngology (2023): 10.5005/jp-journals-10023-1243

INTRODUCTION

Acoustic voice analysis is a noninvasive test that provides quantitative data that is highly descriptive of vocal fold vibration, making it a very useful tool for voice assessment. However, voice analysis equipment is not always available in remote areas, and it could be costly and difficult for patients to travel long distances for repeated follow-ups. It would be convenient for the treating doctor as well as for the patient if voice data could be effectively recorded using an Android or iOS smartphone and transferred to a voice clinic where acoustic analysis equipment is installed. The voice surgeon and speech and language pathologist could ask patients to record their voices with the in-built mobile phone recorder in .wav format and e-mail the file to the treating hospital. The file could then be analyzed using software like Dr Speech™ software for acoustic voice analysis. Furthermore, considering the current worldwide crisis of the coronavirus disease 2019 (COVID-19) pandemic, it would reduce unnecessary travel and face-to-face doctor-patient interaction during repeated follow-ups, thus minimizing the risk of transmission of any infection from patient to doctor or vice versa.

The healthcare industry is getting digitalized by increasing the use of smartphones, and smartphone user is estimated to be 6.2 billion worldwide.¹ A smartphone is easily accessible for patients in even remote areas. However, there are certain limitations to using mobile phones for voice analysis, such as background noise, resilient

¹⁻⁶Department of ENT and Head & Neck Surgery, Lokmanya Tilak Municipal Medical College and General Hospital (LTMGH), Mumbai, Maharashtra, India

Corresponding Author: Nishigandha V Nehete, Department of ENT & Head and Neck Surgery, Lokmanya Tilak Municipal Medical College and General Hospital (LTMGH), Mumbai, Maharashtra, India, Phone: +91 9960623514, e-mail: nishigandha62@gmail.com

How to cite this article: Joshi AA, Nehete NV, Kulkarni PA, *et al.* Assessing the Reliability of a Mobile Phone Recorder in Acoustic Voice Analysis: A Cross-sectional Study. *Int J Phonosurg Laryngol* 2023;13(1):5-8.

Source of support: Nil

Conflict of interest: None

measurement conditions, and different technical specifications of different mobile phones.

AIMS AND OBJECTIVES

The objective of this study is to analyze and compare subject variability in acoustic voice parameters using two different recording devices, namely, the table-mounted microphone provided by Dr Speech software company and the inbuilt recorder of a smartphone Redmi Note 5 Pro™ (Android v4.4.1) mono

speaker in .wav format and analyzing voice parameters of both the recordings using Dr Speech™ software.

MATERIALS AND METHODS

Study Design

- Type of study: Cross-sectional
- Sample size (n): 60
- Participants with no voice abnormality: 25
- Participants with voice disorders: 35
- Duration of study: 4 months
- Period of study: September–December 2019

Method of Voice Recording

The voice of each study participant was simultaneously recorded with the inbuilt mobile phone voice recorder of a Redmi Note 5 Pro™ (Android v4.4.1) mono speaker and the C01U Pro Professional Universal Serial Bus (USB) studio condenser microphone of Tiger Dr Speech™ software. The Samson C01U USB condenser microphone is a USB output microphone. It also has an analog-to-digital converter built in. The output is of the USB type at 16 bits, 48 kHz, meant to be plugged directly into the USB port on a computer.

C01U Specification

- Frequency response: 20~18000 Hz.
- Polar pattern: Hyper-cardioid.
- Element type: Electret condenser type.
- SPL: 136 dB max.

The Redmi Note 5 Pro™ (Android v4.4.1) phone with recorder specification of standard quality is MP3, 100 kbit/second, 45 MB/hour was chosen for the study as this was the phone being used by the chief investigator for personal use as well. However, any other phone could be used for voice recording with similar results.

Both the recording devices were kept at approximately 15 cm in a sound-treated audiometry room. Everyone was made to say the vowel “e” for 4 seconds.

Audacity Software

The voice recorded in the smartphone was converted to 16 bits 44100 Hz monotype .wav format (which is required for analysis by Dr Speech™ software) using the software “Audacity”™.²

“Audacity”™ is a software which can import and export, edit, record, mix tracks, and add effects to sounds in all formats viz

.wav, .AIFF, .MP3, .OGG, etc. This software also has some advanced features like editing amplitude, customizing the spectrogram, and analyzing the frequency of the sound.

Voice recorded by the table-mounted microphone was analyzed directly by Dr Speech™ software without any conversion.

Acoustic Voice Analysis of Both the Recorded Voices by Dr Speech™ Software

Acoustic voice analysis was done by Dr Speech™ software. The mid-portion of all the recordings was selected to analyze voice parameters.

Data were collected and archived using Microsoft Excel. The student t-test was used to compare voices recorded by both devices, and calculations were done via Statistical Package for the Social Sciences software™. The student unpaired t-test was used to compare acoustic voice analysis for voice disorders and normal voice recordings by both recordings of study participants.

RESULTS

Comparison between All Parameters of Voice Recorded by Microphone Provided by Dr Speech™ Company and by Mobile Phone Recorder

Voice parameters of all study participants were recorded by the microphone and mobile phone recorder. These were then compared with each other. The student’s paired t-test was applied to all voice parameters, and no statistically significant difference ($p > 0.05$) was seen except in shimmer ($p < 0.01$) and HNR ($p < 0.01$) in subject with normal voice and shimmer ($p < 0.005$) and HNR ($p < 0.01$) (Tables 1 and 2). We, therefore, conclude that there was a correlation or similarity between the acoustic voice parameters of voice recorded using the microphone of Dr Speech™ software and the in-built recorder of a mobile phone except for two parameters, but the range was within normal limits.

Assessing Differences in Voice Parameters of Study Participants with No Voice Abnormality and with Voice Disorders Recorded by Both the Dr Speech™ Software Microphone and the Mobile Phone

Voice parameters of study participants with a normal voice and participants with voice disorders were recorded by using the microphone of Dr Speech™ software and were grouped into voice recorded by microphone in normal subject (MiN) and voice recorded by microphone in diseased subject (MiD), respectively.

Table 1: Comparison between voice recorded by microphone of Dr Speech™ company and mobile phone recorder for all parameters in normal subjects (n = 25)

SR	Parameters	Dr Speech™ microphone (mean ± SD)	Mobile phone recorder (mean ± SD)	t-test value, significance, and p-value
1	Habitual F0	192.68 ± 49.07	193.95 ± 47.57	1.1, nonsignificant (NS), $p = 0.30$
2	Jitter	0.22 ± 0.09	0.17 ± 0.05	4.2, significant (S), $p = 0.1$
3	Shimmer	3.37 ± 1.61	2.69 ± 1.50	5.5, S, $p < 0.01$
4	F0 tremor	2.69 ± 2.21	2.57 ± 2.00	0.8, NS, $p = 0.43$
5	Mean F0	192.78 ± 48.94	193.02 ± 48.93	1.2, NS, $p = 0.24$
6	SD F0	1.80 ± 0.98	1.70 ± 0.81	1.5, NS, $p = 0.15$
7	Maximum F0	200.47 ± 54.50	197.08 ± 50.19	1.2, NS, $p = 0.25$
8	Minimum F0	187.73 ± 47.29	188.63 ± 47.76	1.5, NS, $p = 0.17$
9	NNE	-10.26 ± 3.66	-10.25 ± 5.87	0.1, NS, $p = 0.9$
10	HNR	17.56 ± 3.97	23.56 ± 3.83	8.9, S, $p < 0.01$



Voice parameters recorded using a mobile phone recorder were grouped into voice recorded by mobile phone in normal subject (MoN) and voice recorded by mobile phone in diseased subject (MoD). Student unpaired *t*-test was applied between groups MiN and MiD (Table 3), and student unpaired *t*-test was applied to voice parameters recorded using a mobile phone between normal and diseased groups (MoN and MoD) (Table 4). There was a statistically significant difference ($p < 0.05$) between the voice parameters of study participants with a normal voice and participants with voice disorders in both groups. This suggested

that mobile phones could differentiate between normal and diseased voices.

DISCUSSION

Analysis of acoustic voice parameters like habitual fundamental frequency (F0), F0 tremors, mean F0, standard deviation (SD) F0, maximum F0, minimum F0, normalized noise energy (NNE), and amplitude tremors of the same study participant recorded by both modalities, that is, the microphone of Dr Speech software™

Table 2: Comparison between voice recorded by microphone of Dr Speech™ company and mobile phone recorder for all parameters in subjects with voice disorders ($n = 35$)

SR	Parameters	Dr. Speech microphone™ (mean ± SD)	Mobile phone recorder (mean ± SD)	<i>t</i> -test value, significance, and <i>p</i> -value
1	Habitual F0	238.76 ± 60.32	238.75 ± 59.94	0.1, NS, $p = 0.9$
2	Jitter	0.50 ± 0.52	0.46 ± 0.56	0.8, NS, $p = 0.4$
3	Shimmer	3.89 ± 2.69	3.30 ± 2.21	3.1, S, $p = 0.005$
4	F0 tremor	3.89 ± 3.89	3.68 ± 3.12	0.9, NS, $p = 0.39$
5	Mean F0	238.57 ± 60.16	238.18 ± 60.05	0.46, NS, $p = 0.65$
6	SD F0	2.79 ± 2.17	2.80 ± 2.34	0.1, NS, $p = 0.9$
7	Maximum F0	245.98 ± 61.85	246.11 ± 62.38	0.2, NS, $p = 0.9$
8	Minimum F0	230.47 ± 59.46	230.37 ± 58.28	0.1, NS, $p = 0.9$
9	NNE	-8.34 ± 5.72	-7.91 ± 5.33	1.1, NS, $p = 0.3$
10	HNR	19.71 ± 6.72	21.40 ± 6.69	2.3, S, $p = 0.03$

Table 3: Difference in voice parameter of normal and diseased subject recorded by microphone of Dr Speech™ software

SR	Parameters	Normal (MiN) ($n = 25$)	Diseased (MiD) ($n = 35$)	Difference (normal-diseased)	<i>t</i> -test value, significance, and <i>p</i> -value
1	Habitual F0 Hz	192.68 ± 49.07	238.76 ± 60.32	-46.09 ± 14.14	3.3, S, $p = 0.002$
2	Jitter	0.22 ± 0.09	0.50 ± 0.52	0.28 ± 0.09	3.1, S, $p = 0.003$
3	Shimmer	3.37 ± 1.61	3.89 ± 2.69	-0.53 ± 0.56	1.0, NS, $p = 0.035$
4	F0 tremor	2.69 ± 2.21	3.89 ± 3.89	-1.21 ± 0.79	1.5, NS, $p = 0.013$
5	Mean F0	192.78 ± 48.94	238.57 ± 60.16	-45.79 ± 14.11	3.3, S, $p = 0.002$
6	SD F0	1.80 ± 0.98	2.79 ± 2.17	-0.99 ± 0.42	2.4, S, $p = 0.02$
7	Maximum F0	194.92 ± 51.66	245.98 ± 61.85	-51.06 ± 14.69	3.5, S, $p = 0.001$
8	Minimum F0	187.73 ± 47.29	230.47 ± 59.46	-42.74 ± 13.79	3.1, S, $p = 0.003$
9	NNE	-10.26 ± 3.66	-8.34 ± 5.72	-1.92 ± 1.21	1.6, NS, $p = 0.012$
10	HNR	17.56 ± 3.97	19.71 ± 6.72	-2.15 ± 1.38	1.6, NS, $p = 0.011$

Table 4: Difference in voice parameter of normal and diseased subject recorded by mobile phone recorder

SR	Parameters	Normal (MoN) ($n = 25$)	Diseased (MoD) ($n = 35$)	Difference (normal-diseased)	<i>t</i> -test value, significance, and <i>p</i> -value
1	Habitual F0 Hz	193.95 ± 47.57	238.75 ± 59.94	-44.80 ± 13.89	3.2, S, $p = 0.002$
2	Jitter	0.17 ± 0.05	0.46 ± 0.56	-0.29 ± 0.09	3.0, S, $p = 0.004$
3	Shimmer	2.69 ± 1.50	3.30 ± 2.21	-0.61 ± 0.48	1.3, NS, $p = 0.021$
4	F0 tremor	2.57 ± 2.00	3.68 ± 3.12	-0.11 ± 0.66	1.7, NS, $p = 0.010$
5	Mean F0	193.02 ± 48.93	238.18 ± 60.05	-45.16 ± 14.10	3.2, S, $p = 0.002$
6	SD F0	1.70 ± 0.81	2.80 ± 2.34	-1.10 ± 0.43	2.6, S, $p = 0.012$
7	Maximum F0	246.85 ± 290.02	246.11 ± 62.38	-0.74 ± 60.67	0.1, NS, $p = 0.019$
8	Minimum F0	230.37 ± 247.0	230.37 ± 58.28	-0.00 ± 50.48	0.0, NS, $p = 0.04$
9	NNE	-10.25 ± 5.87	-7.91 ± 5.33	-2.34 ± 1.48	1.6, NS, $p = 0.017$
10	HNR	23.57 ± 3.83	21.40 ± 6.69	2.16 ± 1.36	1.6, NS, $p = 0.012$

S, significant; NS, not significant; Habitual F0, habitual fundamental frequency; F0 tremor, Fundamental Frequency Tremor; Mean F0, mean fundamental frequency; SD F0, standard deviation fundamental frequency; MAX F0, maximum fundamental frequency; MIN F0, minimum fundamental frequency; NNE, normalized noise energy; HNR, harmonic to noise ratio; MiN, voice recorded by microphone in normal subject; MiD, voice recorded by microphone in diseased subject; MoN, voice recorded by mobile phone in normal subject; MoD, voice recorded by mobile phone in diseased subject

and mobile phone recorder, did not show statistically significant difference except shimmer and HNR which were statistically significant but still they were within normal range.

The frequency and amplitude of periodic voice signals are measured accurately through sustained vowels. They are represented by acoustic measures mentioned above by F0, SD of F0, jitter%, shimmer%, and HNR.

Our results are consistent with those reported by Manfredi et al. in their study comparing two smartphone recordings of voice against a clinic microphone. Voice analysis was done between different smartphones and clinic microphones in a soundproof room using Praat software. They concluded that smartphones could be used to analyze unhealthy voices.³

A total of 10 men and women with normal voices were studied by Grillo et al., using different smartphones like iPhone 5, iPhone 6, and Samsung Galaxy S5, and head-mounted microphones. Two different sounds produced at comfortable fundamental frequency and intensity were analyzed using various software and three different apps. There was no significant difference noted amongst various recording devices.⁴

Uloza et al. compared voice recorded by Samsung Galaxy Note 3 and a head-mounted microphone stating strong correlations between the two.⁵

Fava and colleagues, in 2016, compared voices recorded in soft habitual and loud tones using type 1 SLM, SLM app on iPhone 5, and RadioShack SLM. The results were significantly different for different recording devices. This study did not consider within-subject variability.⁶

Our study showed that there was no statistically significant difference in most of the parameters. Hence our study proved that the microphone of Dr Speech company and the recorder of a smartphone could be utilized to record the voice and to analyze voice parameters in certain difficult situations where patients cannot reach doctors. It could also be postulated that it can differentiate normal from diseased voices.

CONCLUSION

Acoustic voice analysis can be easily and reliably done on the voice recorded by a mobile phone recorder. The in-built recorder of a mobile phone can effectively differentiate between diseased and normal voice and can be used to share voice data without the need for actual physical consultation, especially during the COVID-19 pandemic. However, more research with a larger sample size is required before smartphones can be used widely in acoustic voice analysis.⁷

Clinical Significance

- Voice data can be easily shared between doctors and patients.
- It will reduce face-to-face doctor-patient interaction, eliminating the risk of transmission of any disease, especially when patients are unable to follow up due to COVID-19 infection.

- Since the smartphone recorder can differentiate between diseased and normal voices, it can be helpful for preoperative and postoperative voice comparison without the physical follow-up of patients.
- When a voice specialist is unavailable to patients staying in remote areas.
- When an appointment with a voice specialist is delayed due to logistic reasons.

LIMITATIONS

There are certain limitations which need to be considered before mobile phones can be used routinely for voice recording:

- Background noise could create discrepancies in the voice recorded by patients.
- There could be variability of recording between different models of phones.
- Few patients could be technically incapable of making a proper recording of their voice and sending the same to the doctor for analysis.

Most of these limitations can be corrected by proper patient education and instructions. With advancements in technology, mobile phones would be expected to have standardized recording software, eliminating variability within different mobile devices.

DISCLOSURE

- Written, informed, and valid consent was taken from all patients.
- Permission from the ethics committee of the institution was taken before commencing the study.

REFERENCES

1. <https://www.statista.com/forecasts/1143723/smartphone-users-in-the-world/>Date: 2023 Date accessed: March 21, 2023(Published 2023)
2. <http://audacity.en.softonic.com>
3. Manfredi C, Lebacqz J, Cantarella G, et al. Smartphones offer new opportunities in clinical voice research. *J Voice* 2017;31(1):111.e1-111.e7. DOI: 10.1016/j.jvoice.2015.12.020
4. Grillo EU, Brosious JN, Sorrell SL, et al. Influence of smartphones and software on acoustic voice measures. *Int J Telerehabil* 2016;8(2):9-14. DOI: 10.5195/ijt.2016.6202
5. Uloza V, Padervinskis E, Vegiene A, et al. Exploring the feasibility of smart phone microphone for measurement of acoustic voice parameters and voice pathology screening. *Eur Arch Otorhinolaryngol* 2015;272(11):3391-3399. DOI: 10.1007/s00405-015-3708-4
6. Fava G, Oliveira G, Baglione M, et al. The use of sound level meter apps in the clinical setting. *Am J Speech Lang Pathol* 2016;25(1):14-28. DOI: 10.1044/2015_AJSLP-13-0137
7. Munnings AJ. The current state and future possibilities of mobile phone "voice analyser" applications, in relation to otorhinolaryngology. *J Voice* 2020;34(4):527-532. DOI: 10.1016/j.jvoice.2018.12.018