COVID-19

Acoustic voice analysis in the COVID-19 era

L'analisi acustica della voce in era COVID-19

Giada Cavallaro, Vincenzo Di Nicola, Nicola Quaranta, Maria Luisa Fiorella

Otolaryngology Unit, Department of Biomedical Sciences, Neuroscience and Sensory Organs, University of Bari "Aldo Moro", Italy

SUMMARY

Objective. Among the different procedures used by the ENT, acoustic analysis of voice has become widely used for correct diagnosis of dysphonia. The instrumental measurements of acoustic parameters were limited during the COVID-19 pandemic by the common belief that a face mask affects the results of the analysis. The purpose of our study was to investigate the impact of surgical masks on F0, jitter, shimmer and harmonics-to-noise ratio (HNR) in adults.
Methods. The study was carried out on a selected group of 50 healthy subjects. Voice samples were recorded directly in Praat. All subjects were trained to voice a vocal sample of a sustained /a/, at a conversational voice intensity, with no intensity or frequency variation, for the Maximum Phonation Time (MPT), wearing the surgical mask and then without wearing the surgical mask.
Results. None of the variations in acoustic voice analysis detected wearing a surgical mask and not wearing a surgical mask were statistically significant.
Conclusions. Our study demonstrates that the acoustic voice analysis procedure can continue to be performed with the use of a surgical mask for the patient, even during the COVID-19 pandemic.

KEY WORDS: COVID-19, acoustic voice analysis, Praat, dysphonia, surgical mask

RIASSUNTO

Obiettivo. Tra le diverse procedure diagnostiche di competenza otorinolaringoiatrica, l'analisi acustica della voce si rivela utile alla valutazione quantitativa della disfonia ma la misurazione strumentale dei parametri acustici è stata limitata durante la pandemia da COVID-19 essendo la mascherina chirurgica considerata comunemente come da ostacolo nella registrazione della voce. Lo scopo del nostro studio è stato quello di analizzare l'impatto della mascherina chirurgica su parametri vocali quali F0, jitter, shimmer e harmonic-to-noise ratio (HNR).
Metodi. È stato studiato un campione di 50 soggetti eufonici utilizzando il programma Praat. I soggetti sono stati istruiti a fonare la vocale /a/ tenuta ad intensità di voce di conversazione, senza variazioni di intensità o frequenza, per il Tempo Massimo Fonatorio (TMF), con e senza mascherina chirurgica.
Risultati. L'analisi acustica eseguita con le due diverse modalità non ha rivelato differenze statisticamente significative nei parametri vocali considerati.
Conclusioni. Il nostro studio dimostra come la procedura diagnostica di analisi acustica della voce può essere eseguita sul paziente che indossa la mascherina chirurgica.

PAROLE CHIAVI: *COVID-19, analisi acustica della voce, Praat, disfonia*

Introduction

During the ongoing COVID-19 pandemic caused by SARS-CoV-2, the World Health Organization and other public health organisations agree that face masks can limit the spread of respiratory viral diseases [1,2]. Whether masks are useful depends on the mechanisms for transmission for SARS-CoV-2, which are likely an association of contact, droplet and aerosol modes. Surgical face masks have been in use since the early 1900s to help prevent infection of surgical wounds from staff-generated oral and nasal bacteria [3]. Today, appli-

cations have evolved from prevention of patient infection to prevention of employee exposure. However, there is ongoing debate about the use of surgical masks as respiratory protection devices [4]. For ENT specialists, dysphonia examination by laryngoscopy requires unavoidable contact with the upper airway, and any reflex coughing or sneezing during procedures will cause direct contamination to medical staffs and office workers [5,6]. Among the different procedures used by the ENT, acoustic analysis of voice has become widely used for correct diagnosis of dysphonia, but the instrumental measurements of acoustic perturbation was limited during the COVID-19 pandemic by the common belief that a face mask affects the results of the analysis. The purpose of our study was to investigate the impact of surgical mask on F0, jitter, shimmer and harmonics-to-noise ratio (HNR) in adults.

Materials and methods

The study was carried out on a selected group of 50 healthy subjects (20 men and 30 women, mean age 47 years, range 26-69) recruited among hospital staff of the ENT Department of the Polyclinic Hospital in Bari (South Italy).
Participants were approached and informed about the study objectives and significance. All participants who agreed to participate in the study signed an informed consent form, previously approved by the local hospital Ethics Committee.
Inclusion criterion was ability to phonate and sustain a vowel for at least 10 seconds. The participants were excluded if they met any of the following criteria: reporting recent voice problems or a voice disorder history, a condition that might affect the normal voice function, any previous formal voice training or voice therapy, any laryngeal, mouth, or throat abnormality, or any respiratory infection for the last 2 weeks before recording. The subjects who met selection criteria were recruited. The participants were asked to stand in front of a microphone (Samson Meteor Mic - USB Studio Condenser Microphone) at a distance of 20 cm from the lips, in a quiet room (< 30 dB background noise). Voice samples were recorded directly in Praat. All subjects were trained to voice a vocal sample of a sustained /a/, at a conversational voice intensity, always within 55 dB and 65 dB, on average (not including recordings the average intensity of which was out of range), as constant as possible, with no intensity or frequency variation, for the Maximum Phonation Time (MPT), wearing a surgical mask and then without a surgical mask. The vocal parameters analysed with Praat were median pitch, mean pitch, minimum pitch, maximum pitch, number of pulses, number of periods, jitter (local), jitter (rap), jitter (ppq5), jitter (ddp), shimmer (local), shimmer (apq3), shimmer (apq5), shimmer (apq11), shimmer (dda) and mean harmonics-to-noise ratio (HNR).

Results

The results are recorded as average and standard deviation (SD). Results were then submitted to statistical analysis by comparing mean values of each parameter. All parameters were analysed in the same patients during phonation with surgical mask (SM) and without surgical mask (NSM). We used Student's test with p = 0.05 significance level after evaluating the t value in each parameter.
As illustrated in Table I, the acoustic analysis showed that there was not a significant difference (at the 0.05 level) in median pitch values (Mean SM = 187.36; SD SM = 52.36; Mean NSM = 189.38; SD NSM = 55.52; p = 0.8523) and in the mean pitch values (Mean SM = 183.52; SD SM = 51.13; Mean NSM = 185.52; SD NSM = 55.12; p = 0.8513) in the two different situations (wearing surgical mask – not wearing surgical mask) (Tab I).
As can be seen in Table II, differences in HNR values were not significant (Mean SM = 20.91; SD SM = 3.44; Mean NSM = 20.92; SD NSM = 3.47; p = 0.9885). At the same time, significant differences were not noticed in jitter or shimmer values (jitter local Mean SM = 0.327; SD SM = 0.134; Mean NSM = 0.298; SD NSM = 0.124; p = 0.2641; shimmer local Mean SM = 3.34; SD SM = 1.420; Mean NSM = 3.165; SD NSM = 1.572; p = 0.5605) (Tabs. III, IV). In conclusion, none of the variations in acoustic voice analysis detected in the same patients with surgical mask and without surgical mask were statistically significant.

Discussion

Acoustic voice analysis is considered to be a very useful technique for detection of voice disorders that can be detected by analysing several acoustic parameters[7]. Subjective assessment methods, such as auditory perceptual analysis, largely depend on the experience of professionals, and may lead to different results. This requirement encourages the use of objective measurement of voice. Processing of a speech signal is used to yield a set of voice parameters. It allows detection of vocal fold pathologies, or other related pathologies, by comparing patients' data with that of other individuals having normal healthy voices [7]. Voice disorders require often voice therapy and other treatments that are based on an initial assessment to quantify deviation from normal measures and an ongoing evaluation to record the progress. Measuring treatment outcomes is the basic component of evidence-based practice. The objective

Table I. Acoustic analysis of median pitch, mean pitch, minimum pitch and maximum pitch values wearing surgical mask (SM) and not wearing surgical mask (NSM). With significance level at 0.05, values obtained by Student t test (calculated) in the same patients with surgical mask and without surgical mask are not statistically significant.

	Median Pitch (Hz) SM	Median Pitch (Hz) NSM	Mean pitch (Hz) SM	Mean pitch (Hz) NSM	Minimum pitch (Hz) SM	Minimum Pitch (Hz) NSM	Maximum Pitch (Hz) SM	Maximum Pitch (Hz) NSM
Mean	187.36	189.38	183.52	185.52	173.37	181.87	194.52	195.94
Standard Deviation	52.36	55.52	51.13	55.12	54.15	59.05	54.63	56.47
T-test	p = 0.8523	p = 0.8513	p = 0.4549	p = 0.8986				

Table II. Acoustic analysis of the number of pulses, number of periods and of the HN (harmonics-to-noise ratio) values wearing surgical mask (SM) and not wearing surgical mask (NSM). With significance level at 0.05, values obtained by Student t test (calculated) in the same patients with surgical mask and without surgical mask are not statistically significant.

	Number of pulses SM	Number of pulses NSM	Numbers of periods SM	Numbers of periods NSM	Mean HNR (dB) SM	Mean HNR (dB) NSM
Mean	574.18	575.00	573.14	574.00	20.91	20.92
Standard Deviation	157.88	168.76	157.85	168.76	3.44	3.47
T-test	p = 0.9800		p = 0.9791		p = 0.9885	

Table III. Acoustic analysis of jitter values wearing surgical mask (SM) and not wearing surgical mask (NSM). With significance level at 0.05, values obtained by Student t test (calculated) in the same patients with surgical mask and without surgical mask are not statistically significant.

	Jitter local SM (%)	Jitter local NSM (%)	Jitter rap SM (%)	Jitter rap NSM (%)	Jitter ppq5 SM (%)	Jitter ppq5 NSM (%)	Jitter ddp SM (%)	Jitter ddp NSM (%)
Mean	0.327	0.298	0.184	0.159	0.182	0.165	0.535	0.533
Standard Deviation	0.134	0.124	0.084	0.068	0.071	0.062	0.240	0.411
T-test	p = 0.2641		p = 0.1051		p = 0.2052		p = 0.9764	

Table IV. Acoustic analysis of shimmer values wearing surgical mask (SM) and not wearing surgical mask (NSM). With significance level at 0.05, values obtained by Student t test (calculated) in the same patients with surgical mask and without surgical mask are not statistically significant.

	Shimmer local SM (%)	Shimmer local NSM (%)	Shimmer apq3 SM (%)	Shimmer apq3 NSM (%)	Shimmer apq5 SM (%)	Shimmer apq5 NSM (%)	Shimmer apq11 SM (%)	Shimmer apq11 NSM (%)	Shimmer dda SM (%)	Shimmer dda NSM (%)
Mean	3.34	3.165	1.726	1.589	2.008	1.836	2.705	2.689	5.070	4.766
Standard Deviation	1.420	1.572	0.840	0.974	1.061	0.897	1.087	1.194	2.531	2.922
T-test	p = 0.5605		p = 0.4531		p = 0.3835		p = 0.9443		p = 0.5794	

assessment of voice, especially acoustic analysis, has received our attention because of its comparatively low cost, ease of application and quantitative output. Previous studies [8,9] have found that fundamental frequency (F0) can be affected by different factors, i.e., age, vocal fold length and language or ethnological background. Until now, no study has investigated the effects of the use of a surgical mask on acoustic parameters. According to previous studies, one of the most investigated voice acoustic parameters has been voice perturbation [10,11]. Subsequently, we investigated parameters such as F0, jitter, shimmer and harmonics-to-noise ratio (HNR) during phonation wearing surgical mask

and then not wearing surgical mask. The fundamental frequency or mean pitch (F0) of a speech signal refers to the approximate frequency of the (quasi-)periodic structure of voiced speech signals. Jitter (%) is defined as cycle-to-cycle and short-term perturbation in the fundamental frequency of the voice. The shimmer (%) is a cycle-to- cycle, short-term perturbation in the amplitude of voice. Another acoustic parameter (HNR) is influenced by both the shimmer and jitter and referred to as the mean ratio of harmonics to non-harmonics [12].

In accordance with such a high risk of infection, only emergency consultations and procedures should be performed by ENT specialists during the COVID-19 pandemic in areas with confirmed SARS-CoV-2 cases [13]. In China, Cheng et al. noted that the rate of work-related SARS-CoV-2 infection was higher among ENT specialists that in other medical specialties [14]. During the lockdown of the population in Italy, ENT activities were reduced to emergency treatments and those that could not be deferred without constituting a real loss of chance for the patient's recovery or survival. ENT specialists are exposed to SARS-CoV-2 infection because of the necessity to examine the upper respiratory tract. At the same time, they perform procedures that generate aerosolised secretions and often bleeding [15]. In the study by Krajewska et al. [16] ENT units are important for preoperative testing for SARS-CoV-2: this should be performed in all individuals undergoing high-risk procedures. The authors also assert that chest CT should be performed in patients before ENT interventions, because it could be of great value in individuals with negative RT-PCR.

According to Tysome et al., high-risk procedures must be performed using enhanced personal protective equipment [17]. As highlighted by Lescanne et al. [18], during ENT examinations or procedures that not need exposure to projection/aerosolisation of organic material of human origin, the ENT medical team should wear clean outfits as well as single-use gloves in case of contact with a mucosa. If worn properly, a face mask is a disposable device that is used to help block large-particle droplets, sprays, splashes, or splatters that may contain viruses and bacteria. It is used to create a physical barrier between the potential contaminants in the immediate environment and the mouth and nose of the wearer and it is also useful to block saliva and respiratory secretions from the wearer to another [19]. In our study, the surgical masks used were three-ply. This three-ply material is made up of a melt-blown polymer, most commonly polypropylene, placed between non-woven fabric. For examinations and procedures with exposure to projection/aerosolisation of organic material of human origin, protection must be supplemented by wearing a surgical mask, protective goggles, a single-use plastic apron

and single-use gloves. Insofar as an asymptomatic patient may be infectious, the same precautions must be employed whether the patient is ill with, suspected of having, or without any clinical evidence of COVID-19 infection [20]. After the examination, the professional must carefully disrobe in compliance with hygiene rules, with the immediate elimination of gloves, hair cap, mask and gown. The room where the examination is carried out must undergo air renewal as per legislation [20]. Most of these best practice recommendations are not based on scientific data established for the COVID-19 infection, but come from what is known about other viral respiratory infections.

For ENT specialists, voice acoustic analysis is a very valuable technique for voice disorders diagnosis and therapy monitoring [21]. Speech signal processing allows the extraction of a set of voice parameters that may be used to diagnose many pathologies of the vocal cords in individuals by comparison with healthy voice. The parameters obtained by the acoustic analysis have the advantage of describing the voice objectively rather than subjective perceptual analysis, and they represent a useful method to objectify the dysphonia, even in the pandemic period. The use of the surgical mask provides the patient and operator with the right protection necessary to perform this procedure, and at the same time it does not involve important alterations of the vocal parameters to be analysed. Several types of software have been developed for acoustic analysis, namely, Praat [22], LingWAVES [23], Multidimensional Voice Program [24] etc. The current study used Praat (version 6.1.16) for voice analyses, which is a computer software package for speech, phonetic and voice analysis. It was first designed in 1992 by Paul Boersma and David Weenick from the Institute of Phonetic Sciences, University of Amsterdam. Praat can be used on various operating systems and uses the finest algorithms including the most accurate algorithm of pitch analysis, articulatory synthesis and gradual learning algorithm for free variation. We used the inbuilt option of voice report in Praat pulses menu, which includes pitch and perturbation analyses. In particular, the voice samples collected for perturbation measures were analysed by selecting the middle 3 seconds from the sound wave. Each acoustic signal was perceptually examined for instability and visually displayed using Praat with an oscillogram and "Show intensity" and "Show pulses" settings. We acoustically analysed the voice samples recorded by each participant wearing and not wearing the surgical mask in order to find objective voice measurements including the F0, jitter, shimmer, and HNR. The statistical comparison carried out between the parameters extracted with and without surgical mask did not reveal any significant differences that would lead to an avoidance of the procedure for health safety reasons.

Conclusions

Excluding positive COVID-19 cases for which the use of more adequate protective devices is necessary, our study demonstrates that the acoustic voice analysis procedure can continue to be performed with the use of surgical mask for the patient during the COVID-19 pandemic.

References

1. https://www.who.int/news-room/q-a-detail/q-a-on-covid-19-and-masks

2. European Centre for Disease Prevention and Control. Using face masks in the community. Reducing COVID-19 transmission from potentially asymptomatic or pre-asymptomatic people through the use of face masks. Stockholm: ECDC; 8 April 2020.

3. Belkin NL. A century after their introduction, are surgical masks necessary? AORN J 1996;64:602-7. https://doi.org/10.1016/s0001-2092(06)63628-4

4. Long Y, Hu T, Liu L, et al. Effectiveness of N95 respirators versus surgical masks against influenza: a systematic review and meta-analysis. J Evid Based Med 2020;13:93-101. https://doi.org/10.1111/jebm.12381

5. van Doremalen N, Bushmaker T, Morris DH, et al. Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. N Engl J Med 2020;382:1564-7. https://doi.org/10.1056/NEJMc2004973

6. Ong SWX, Tan YK, Chia PY, et al. Air, surface environmental, and personal protective equipment contamination by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) from a symptomatic patient JAMA 2020;323:1610-2. https://doi.org/10.1001/jama.2020.3227

7. Gorris C, Ricci Maccarini A, Vanoni F, et al. Acoustic analysis of normal voice patterns in italian adults by using Praat. J Voice 2019:S0892-1997(19)30034-7. https://doi.org/10.1016/j.jvoice.2019.04.016.

8. Braun A. Fundamental frequency - how speaker-specific it is? In: Braun A, Koster JP, editors. Studies in Forensic Phonetics: BEIPHOL 64. 1995. pp. 9-23.

9. Mennen I, Schaeffler F, Docherty G. Cross-language differences in fundamental frequency range: a comparison of English and German. J Acoust Soc Am 2012;131:2249-60. https://doi.org/10.1121/1.3681950

10. Petrović-Lazić M, Babac S, Vuković M, et al. Acoustic voice analysis of patients with vocal fold polyp. J Voice 2011;25:94-7. https://doi.org/10.1016/j.jvoice.2009.04.002

11. Godino-Llorente JI, Osma-Ruiz V, Sáenz-Lechón N, et al. The effectiveness of the glottal to noise excitation ratio for the screening of voice disorders. J Voice 2010;24:47-56. https://doi.org/10.1016/j.jvoice.2008.04.006

12. Di Nicola V, Fiorella ML, Spinelli DA, et al. Acoustic analysis of voice in patients treated by reconstructive subtotal laryngectomy. Evaluation and critical review. Acta Otorhinolaryngol Ital 2006;26:59-68.

13. Cui C, Yao Q, Zhang D, et al. Approaching otolaryngology patients during the COVID-19 Pandemic. Otolaryngol Head Neck Surg 2020;163:121-31. https://doi.org/10.1177/0194599820926144.

14. Cheng X, Liu J, Li N, et al. Otolaryngology providers must be alert for patients with mild and asymptomatic COVID-19. Otolaryngol Head Neck Surg 2020;162:809-10. https://doi.org/10.1177/0194599820920649

15. Zou L, Ruan F, Huang M, et al. SARS-CoV-2 Viral load in upper respiratory specimens of infected patients. N Engl J Med 2020;382:1177-9. https://doi.org/10.1056/NEJMc2001737

16. Krajewska J, Krajewski W, Zub K, et al. COVID-19 in otolaryngologist practice: a review of current knowledge. Version 2. Eur Arch Otorhinolaryngol 2020;277:1885-97. https://doi.org/10.1007/s00405-020-05968-y.

17. Tysome JR, Bhutta MF. COVID-19: protecting our ENT Workforce. Clin Otolaryngol 2020;45:311-2. https://doi.org/10.1111/coa.13542.

18. Lescanne E, van der Mee-Marquet N, Juvanon JM, et al. Best practice recommendations: ENT consultations during the COVID-19 pandemic. Eur Ann Otorhinolaryngol Head Neck Dis 2020;S1879-7296(20)30126-5. https://doi.org/10.1016/j.anorl.2020.05.007

19. Brewster DJ, Chrimes N, Do TB, et al. Consensus statement: safe airway society principles of airway management and tracheal intubation specific to the COVID-19 adult patient group. Med J Aust 2020;212:472-81. https://doi.org/10.5694/mja2.50598

20. Liang T, Yu L. Handbook of COVID-19 prevention and treatment. Zhejiang: Zhejiang University School of Medicine; 2020.

21. Di Nicola V, Fiorella ML, Luperto P, et al. La valutazione obiettiva della disfonia. Possibilità e limiti [Objective evaluation of dysphonia. Possibilities and limitations]. Acta Otorhinolaryngol Ital 2001;21:10-21.

22. Boersma P, Weenink D. Praat: doing phonetics by computer [Computer program]. Version 6.0.37, 2018; http://www.praat.org/, accessed August 5, 2020

23. Caffier PP, Möller A, Forbes E, et al. The vocal extent measure: development of a novel parameter in voice diagnostics and initial clinical experience. Biomed Res Int 2018;2018:3836714. https://doi.org/10.1155/2018/3836714

24. Lovato A, De Colle W, Giacomelli L, et al. Multi-dimensional voice program (MDVP) vs Praat for assessing euphonic subjects: a preliminary study on the gender-discriminating power of acoustic analysis software. J Voice 2016;30:765.e1-765.e5. https://doi.org/10.1016/j.jvoice.2015.10.012