Effect of FFP2/3 Masks on Voice Range Profile Measurement and Voice Acoustics in Routine Voice Diagnostics

Guan-Yuh Ho\textsuperscript{a}  Ines Kristina Kansy\textsuperscript{a}  Katharina Anna Klavacs\textsuperscript{b}  Matthias Leonhard\textsuperscript{a}  Berit Schneider-Stickler\textsuperscript{a, b}

\textsuperscript{a}Division of Phoniatrics-Logopedics, Department of Otorhinolaryngology, Medical University Hospital Vienna, Vienna, Austria; \textsuperscript{b}Medical Center of Communication, Vienna, Austria

\textbf{Keywords}
Laryngology · COVID-19 pandemic · Coronavirus · Vocal damping by facial masks · Voice diagnostics · Dysphonia Severity Index · Vowel formants · Maximum phonation time · Phonetogram

\textbf{Abstract}
\textbf{Introduction:} Voice diagnostics including voice range profile (VRP) measurement and acoustic voice analysis is essential in laryngology and phoniatrics. Due to COVID-19 pandemic, wearing of 2 or 3 filtering face piece (FFP2/3) masks is recommended when high-risk aerosol-generating procedures like singing and speaking are being performed. Goal of this study was to compare VRP parameters when performed without and with FFP2/3 masks. Further, formant analysis for sustained vowels, singer’s formant, and analysis of reading standard text samples were performed without/with FFP2/3 masks. \textbf{Methods:} Twenty subjects (6 males and 14 females) were enrolled in this study with an average age of 36 ± 16 years (mean ± SD). Fourteen patients were rated as euphonic/not hoarse and 6 patients as mildly hoarse. All subjects underwent the VRP measurements, vowel, and text recordings without/with FFP2/3 mask using the software DIVAS by XION medical (Berlin, Germany). Voice range of singing voice, equivalent of voice extension measure (eVEM), fundamental frequency (F0), sound pressure level (SPL) of soft speaking and shouting were calculated and analyzed. Maximum phonation time (MPT) and jitter-% were included for Dysphonia Severity Index (DSI) measurement. Analyses of singer’s formant were performed. Spectral analyses of sustained vowels /a:/, /i:/, and /u:/ (first = F1 and second = F2 formants), intensity of long-term average spectrum, and alpha-ratio were calculated using the freeware praat. \textbf{Results:} For all subjects, the mean values of routine voice parameters without/with mask were analyzed: no significant differences were found in results of singing voice range, eVEM, SPL, and frequency of soft speaking/shouting, except significantly lower mean SPL of shouting with FFP2/3 mask, in particular that of the female subjects ($p = 0.002$). Results of MPT, jitter, and DSI without/with FFP2/3 mask showed no significant differences. Further mean values analyzed without/with mask were ratio singer’s formant/loud singing, with lower ratio with FFP2/3 mask ($p = 0.001$), and F1 and F2 of /a:/, /i:/, /u:/, with no significant differences of the results, with the exception of F2 of /i:/ with lower value with FFP2/3 mask.

This is an Open Access article licensed under the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC) (http://www.karger.com/Services/OpenAccessLicense), applicable to the online version of the article only. Usage and distribution for commercial purposes requires written permission.

Correspondence to:
Berit Schneider-Stickler, berit.schneider-stickler@meduniwien.ac.at
With the exceptions mentioned, the $t$ test revealed no significant differences for each of the routine parameters tested in the recordings without and with wearing a FFP2/3 mask. **Conclusion:** It can be concluded that VRP measurements including DSI performed with FFP2/3 masks provide reliable data in clinical routine with respect to voice condition/constitution. Spectral analyses of sustained vowel, text, and singer’s formant will be affected by wearing FFP2/3 masks.

© 2022 The Author(s).
Published by S. Karger AG, Basel

### Introduction

Since 2001, the basic protocol for functional assessment of voice pathology of the Committee of Phoniatrics of the European Laryngological Society has become standard for clinical routine voice diagnostics in most laryngology departments and voice centers. Its basic set of "truncus communis" for the assessment of common dysphonia considers (a) perception, (b) videostroboscopy, (c) acoustics, (d) aerodynamics/efficiency, and (e) subjective rating by patient [1].

The COVID-19 pandemic still continues and has a major impact on daily life and challenges clinical routine work. Airborne transmission by respiratory droplets or aerosols has been identified as the most important transmission pathway for the severe acute respiratory syndrome coronavirus (SARS-CoV-2) [2–4]. Aerosol-generating procedures, such as endoscopy of nose, mouth, or larynx and voice diagnostics, are considered as high risk of transmitting coronavirus (COVID-19). Among healthcare providers, laryngologists, phoniatricians, speech language pathologists, and speech therapists are at elevated risk of COVID-19 infection due to a high potential of viral transmission by aerosols and airborne droplets produced during speaking, singing, sneezing, and coughing [5]. Inhaling small airborne droplets is the most probable route of infection, in addition to more widely recognized transmission via larger respiratory droplets and direct contact with infected people or contaminated surfaces [6]. The infection-control pyramid considers physical elimination of the pathogen, distancing to separate people and pathogen, administrative controls, and personal protective equipment (PPE) [7]. Any preventive measure (such as wearing masks, ventilation, social distancing) aims at reduction of inhaled particle concentration and thus should reduce the infection probability and spreading of COVID-19. The WHO recommends wearing of protective class 2 or 3 filtering face piece (FFP2/3) masks, if high-risk aerosol-generating procedures are required and cannot be avoided [8, 9].

In 2020, the Union of the European Phoniatricians published a position paper on recommendations for phoniatricians and ENT surgeons on how to provide and/or run healthcare services for voice, swallowing, speech and language, or pediatric audiology proposing the use of PPE by both the medical staff and the patients [10]. Asadi et al. [11] pointed out a higher rate of particle emission during loud vocalization, ranging from approximately 1 to 50 particles per second (with 0.06 to 3 particles per cm$^3$) for low to high amplitudes.

Several studies revealed that singing, especially choir singing, poses a major risk during COVID-19 pandemic due to high infection rates [12–15]. Due to the high risk of person-to-person infection with SARS-CoV-2 during phonation (speaking and singing), the COVID-19 pandemic poses major challenges for healthcare workers involved with voice diagnostics. Routine voice diagnostic measurements cannot be performed by distance tele-health, as recommended by the Union of the European Phoniatricians (UEP) for speech and language assessment and therapy [10], because there is a lack of instrument-based standards.

Standardized office-based instrumental measurements of voice acoustics remain mandatory for correct diagnosis of dysphonia. Wearing face masks seems to affect speech intelligibility and acoustic voice measurements. In classroom communication, the use of fabric masks yielded a significantly greater reduction in speech intelligibility compared to the other masks [16]. Several common types of face masks (N95, surgical, and cloth) have been tested for acoustic measures of timing, frequency, perturbation, power spectral density, speech intelligibility, and word and sentence accuracy by Magee et al. [17]. The data indicated that face masks change the speech signal, but specific acoustic features remain largely unaffected (e.g., measures of voice quality) irrespective of the mask type. Maryn et al. [18] confirmed the influence of facial masks on speech sound properties, although the impact differed between mask types. The authors included 26 speech sound properties related to voice production (fundamental frequency, sound intensity [1] level), voice quality (jitter percent, shimmer percent, harmonics-to-noise ratio, smoothed cepstral peak prominence, Acoustic Voice Quality Index), articulation, and resonance parameters in the analysis.

So far, recent studies focus mainly on perturbation measurement, formant analysis, speech intelligibility, maximum phonation time (MPT), and spectral analysis.
when comparing voice production without and with wearing a facial mask. As for personal protection, it still remains unclear whether the recommended FFP2/3 masks in voice centers alter the outcome of routine voice range profile (VRP) measurements and voice acoustics in clinical diagnostics. One routine voice diagnostic measurement is the VRP, also called phonetogram [19, 20], which displays sound pressure level (SPL) versus fundamental frequency [21] for singing and speaking tasks. So far, it has not been investigated whether these measurements performed without and with wearing a facial mask provide similar results.

The goal of this study was the comparison of VRP measurements for speaking and singing voice without and with wearing either a FFP2/3 mask. As several commercially available VRP measurement devices offer Dysphonia Severity Index (DSI) measurements [22] which uses highest frequency (F0-high in Hz) and lowest intensity (in dB) from the VRP, the MPT (in second) and jitter (%) were included additionally for DSI measurement. Further, spectral analyses of sustained vowels (first and second formants/F1 and F2), analyses of singer’s formant as well as assessment of long-term average spectrum (LTAS) and alpha-ratio (α-ratio) measurement were considered in this study.

Subjects and Methods

The study has been approved by the Ethic Committee of the Medical University of Vienna (EK 1536/2021). For this study, a total of 20 subjects were enrolled at either the ENT outpatient department of the Medical University Hospital Vienna or the Medical Center of Communication Vienna. Ten out of the 20 subjects were interested internal department staffs, who were recruited by flyers. The other subjects were patients with functional voice disorders.

As a result of the pandemic-related legal requirements, patients were required to wear either FFP2 or, in individual cases, FFP3 masks, in hospitals or at private doctor visits. During the study, all subjects wore their own CE certified face mask. All patients were tested negative for SARS-CoV-2 infection prior to the measurements, and the correct fitting of the face mask was checked by the examiner. Both speech therapists (I.K. and K.K.) involved in the measurements wore PPE including FFP3 masks during voice recordings and VRP measurements. For the voice recording, a XION headset with incorporated microphone with 30 cm mouth-microphone distance was used. All recordings were taken using the software DiVAS by XION medical (Berlin, Germany), which is the standard software for voice diagnostics at the department of phoniatrics and logopedics and Medical Center of Communication. For all subjects, the recording was performed in the same order (1) recording standard text “The North Wind and the Sun” in German, (2) recording of sustained vowels /a:/, /i:/, and /u:/, and (3) recording of singing and speaking (numbers in German) voices. The voice recordings were automatically saved as separate sound files in wave format to the computer where the software DiVAS by XION medical (Berlin, Germany) was installed. To avoid bias, the order of recording and measuring without and with wearing a mask was randomly chosen (shown in Fig. 1, 2). All recordings were performed at recording room setup for routine voice diagnostics.

All subjects underwent the VRP measurements without and with wearing their FFP2/FFP3 masks using the software DiVAS. For automated computation of DSI, MPT (in second) and jitter (in %) of a sustained /a:/ vowel were also measured. For further analysis of VRP parameters, the values of each subject without and with mask recorded by DiVAS were exported to Excel table.

To measure the VRP, first, the voice ranges of soft and loud singing were recorded. All subjects had to sing over the entire vocal pitch range, at the beginning softly and then as loud as possible.
For both soft and loud singing, each subject began at its middle pitch and went up to the highest possible pitch. Afterward, each subject went down to its lowest possible pitch.

To display the special energy components of the singer’s formant level, the SPL in the spectral range of 2–4 kHz was simultaneously recorded along with the loudest sung tones. Afterward, the ratio (in %) between singer’s formant in dB and loud singing in dB was calculated.

Recordings of the speaking voice from soft speaking to shouting were taken for the assessment of the fundamental frequency (F0) and SPL of the levels of increase. During the voice recordings, all subjects had to count numbers (German language from 21 upward), in the orders: soft speaking, conversational loudness, classroom loudness, and shouting voice. For the automated calculation of DSI, MPT and jitter-% at comfortable pitch and loudness were measured using the software DiVAS. For the measurement of MPT, all subjects were asked to take a deep breath and hold vowel /a:/ as long as possible. To measure jitter-%, vowel /a:/ of each subject was recorded for 3 s; afterward, 2 s of the midportion of the voice recording was marked for automated analysis by DiVAS.

Considering the logarithmic character of the frequency scale (measured in Hz) of the VRP, the use of a semitone (st) scale has been preferred with A1 = 55 Hz = st 1 (lower range limit of VRP) and f3 = 1,319 Hz = st 57 (upper range of VRP). The F0 of soft speaking at lowest SPL and shouting voice at highest SPL as well as the frequency of soft singing at lowest SPL and loud singing at highest SPL was allocated to the appropriate st-level, and the voice range of speaking as well as singing voice, as st-range, was calculated accordingly.

For quantification of size and extent of the singing voice range, an algorithm has been developed by one of the authors (ML) using pixel data equivalent of the vocal extent measure (eVEM in px) by Caffier et al. [23]. The data of each singing VRP was exported from the DiVAS application, and the SPL (dB) levels were plotted corresponding with the frequencies on the st scale as vertices of the area representing the voice range on a defined image of 1,275 × 800 pixels. The number of pixels of the area was calculated and used for quantification of the singing voice ranges with and without FFP2 masks as an eVEM in pixel for each subject.

Additionally, the voice recordings of the sustained vowels /a:/, /i:/, and /u:/ without and with mask, which were performed using the software DiVAS, were taken for 3–4 s at comfortable pitch and loudness. Afterward, the recordings were exported and 2 s of the midportion of each vowel recording was analyzed for formants F1 and F2 in computer under Windows 10 environment using the freeware program praat (version 6.1.53) (https://www.fon.hum.uva.nl/praat/download_mac.html). The values of F1 and F2 of all subjects were then entered in an Excel table for further evaluation.

To assess the LTAS, the voice recordings of the standard text “The North Wind and the Sun” in German without and with mask were performed at comfortable pitch and loudness for all subjects. Later, the recordings were exported and trimmed using praat, where the unvoiced segments between each sentence were removed and then saved as new files in wave format. For analyzing the LTAS of the trimmed standard text recordings, the bandwidth of 400 Hz was used. In praat, the following steps were performed: first, under “Analyze spectrum,” the menu option “To LTAS” was selected, and then the bandwidth was set to 400 Hz. After applying the setting, the selected voice spectrum was converted to LTAS object, which was then queried to get the mean LTAS value (in dB) within the frequency range of 0–1,000 Hz and 1,000 Hz–6,000 Hz as described by Sundberg and Nordenberg. [24]. Each value of LTAS below and above the frequency of 1,000 Hz was entered in an Excel table, and afterward, the α-ratio, which is defined as ratio between the sound energies expressed as dB above and below 1,000 Hz: α = IHF/I LF, was calculated [24, 25].

In summary, from the VRP and voice recordings, the following parameters have been considered for analysis: range of singing voice (in st-range), evEM (in pixel), SPL of soft speaking and shouting voices (in dB), st-level above A1 = 55 Hz of soft speaking and shouting voice (st-level), MPT (s), jitter (%), DSI, vowel formants F1 and F2 (Hz) for /a:/, /i:/, and /u:/, singer’s formant, LTAS, and α-ratio. Due to differences in frequency of speaking and singing voice between male and female, the parameters were analyzed separately for each gender as well.

For statistical comparison of the data of small dependent samples (2 dependent groups), the t test was applied with a significance level of 5% (p < 0.05). All the statistical analyses were performed in computer with Windows 10 using IBM SSPS Statistics version 26, 64-bit version.

Results

Descriptive Data of Patients

From the included 20 subjects, 6 were male and 14 were female with an average age of 36 ± 16 years (mean ± standard deviation [SD]) with age ranges from 19 years to 75 years. The 10 subjects of the department were judged with euphonic voices. The 10 patients with functional voice disorders had either none (n = 4) or mild (n = 6) degrees of hoarseness at the time of examination. To sum up, 14 were rated as euphonic and 6 as mildly hoarse.

Results of Singing Voice Analysis

The results of the analysis for singing voice are shown in Table 1, which contains the results of the average singing voice range and eVEM without and with FFP2/3 mask. Comparing the results without and with wearing a FFP2/3 mask of all subjects and of each gender separately, no significant differences were found (p < 0.05).

Results of Speaking Voice Analysis

In this study, the speaking voice without and with FFP2/3 mask was analyzed and the results are listed in Table 1. For soft speaking, the results of the mean SPL of all subjects and of each gender without and with FFP2/3 mask did not show significant differences (p < 0.05).

Comparing the results of mean SPL for the shouting voice of all subjects without and with FFP2/3 mask, it showed significantly lower values when wearing FFP2/3 mask (p = 0.005). Comparing the value of mean SPL of shouting voice separately by gender, only female subjects...
showed significantly lower value when wearing FFP2/3 mask ($p = 0.002$). Nevertheless, from the clinical point of view, it is of no importance, as both values are higher than 90 dB indicating a normal voice condition/constitution [26].

When comparing the values of frequency of soft speaking and shouting voice as well as voice range of speaking voice of all subjects and by gender without and with FFP2/3 mask, no significant differences were found ($p < 0.05$). In this study, MPT, jitter, and DSI of the 20 subjects

**Table 1. Mean statistics±SD of VRP, vowel formants, and LTAS**

| Parameters                        | Without FFP2/3 mask | With FFP2/3 mask | $p$ value |
|-----------------------------------|---------------------|------------------|-----------|
| VRP                               |                     |                  |           |
| Voice range of singing voice (st-range) |                     |                  |           |
| Male                              | 26 ±9               | 30 ±6            | 0.157     |
| Female                            | 27 ±8               | 27 ±7            | 0.885     |
| All                               | 27 ±8               | 28 ±7            | 0.329     |
| eVEM (pixel)                      |                     |                  |           |
| Male                              | 156,734 ±48,188     | 154,401 ±43,445  | 0.723     |
| Female                            | 100,145 ±56,985     | 89,428 ±52,815   | 0.052     |
| All                               | 117,122 ±59,505     | 108,920 ±57,780  | 0.053     |
| SPL soft speaking (dB)            |                     |                  |           |
| Male                              | 46.3 ±1.6           | 45.3 ±1.9        | 0.203     |
| Female                            | 52.2 ±4.8           | 51.4 ±4.2        | 0.344     |
| All                               | 50.5 ±4.9           | 49.6 ±4.6        | 0.163     |
| F0 of soft speaking (st-level)    |                     |                  |           |
| Male                              | 10 ±4               | 11 ±3            | 0.286     |
| Female                            | 23 ±2               | 23 ±2            | 0.390     |
| All                               | 19 ±6               | 19 ±6            | 0.206     |
| SPL shouting (dB)                 |                     |                  |           |
| Male                              | 101.7 ±8.6          | 101.5 ±5.6       | 0.921     |
| Female                            | 101.5 ±3.2          | 96.4 ±6.5        | 0.002     |
| All                               | 101.6 ±5.1          | 97.9 ±6.5        | 0.005     |
| F0 of shouting (st-level)         |                     |                  |           |
| Male                              | 25 ±4               | 27 ±3            | 0.165     |
| Female                            | 32 ±4               | 33 ±3            | 0.355     |
| All                               | 30 ±5               | 31 ±4            | 0.206     |
| Voice range of speaking voice (st-range) |                 |                  |           |
| Male                              | 15 ±7               | 16 ±5            | 0.474     |
| Female                            | 10 ±4               | 10 ±4            | 0.716     |
| All                               | 11 ±6               | 12 ±5            | 0.527     |
| MPT (s)                           |                     |                  |           |
| Male                              | 20 ±2               | 19 ±6            | 0.801     |
| Female                            | 15 ±5               | 15 ±4            | 0.600     |
| All                               | 17 ±5               | 16 ±5            | 0.368     |
| Jitter (%)                        |                     |                  |           |
| Male                              | 1.1 ±0.5            | 1.3 ±0.6         | 0.485     |
| Female                            | 1.5 ±1.9            | 1.4 ±1.3         | 0.637     |
| All                               | 1.4 ±1.6            | 1.4 ±1.1         | 0.951     |
| DSI                               |                     |                  |           |
| Male                              | 5.1 ±1.2            | 4.7 ±1.6         | 0.337     |
| Female                            | 5.2 ±4.5            | 4.7 ±4.2         | 0.238     |
| All                               | 5.1 ±3.8            | 4.7 ±3.6         | 0.127     |
| Ratio singer’s formant/SPL loud singing (%) |           |                  |           |
| Male                              | 81.9 ±5.7           | 79.7 ±5.9        | 0.125     |
| Female                            | 78.9 ±4.5           | 76.3 ±3.3        | 0.004     |
| All                               | 79.8 ±5.0           | 77.3 ±4.4        | 0.001     |
were also measured and the results are shown in Table 1. Comparing the mean MPT, jitter, and DSI of all subjects and by gender without and with FFP2/3 mask, it did not show significant differences of the values of these three measurements ($p < 0.05$).

### Results of Singer’s Formant and Vowel Formants

Within the VRP measurement, the singer’s formant had been evaluated. As shown in Table 1 comparing the average ratio between singer’s formant and loud singing of all subjects without and with FFP2/3 mask, significant

| Parameters | Without FFP2/3 mask | With FFP2/3 mask | $p$ value |
|------------|---------------------|-----------------|-----------|
|            | mean ±SD            | mean ±SD        |           |
| Vowel formant |                    |                 |           |
| /a:/        |                    |                 |           |
| F1 (Hz)     | Male 633.5 41.3     | Male 611.2 38.2 | 0.228     |
|            | Female 691.2 155.2  | Female 691.6 138.5 | 0.994    |
|            | All 673.9 132.9     | All 676.4 122.2 | 0.858     |
| F2 (Hz)     | Male 1,163.9 98.9   | Male 1,123.4 68.0 | 0.059   |
|            | Female 1,355.5 160.7 | Female 1,448.5 315.6 | 0.306  |
|            | All 1,298.0 168.4   | All 1,350.9 304.5 | 0.406    |
| /i:/        |                    |                 |           |
| F1 (Hz)     | Male 296.4 24.9     | Male 292.5 25.8 | 0.818     |
|            | Female 351.5 42.8   | Female 356.3 60.8 | 0.751   |
|            | All 335.0 45.7      | All 337.1 60.0 | 0.850     |
| F2 (Hz)     | Male 2,205.9 121.4  | Male 2,094.0 84.0 | 0.120   |
|            | Female 2,550.7 165.3 | Female 2,461.7 136.2 | 0.027  |
|            | All 2,447.2 221.1   | All 2,351.4 210.8 | 0.005   |
| /u:/        |                    |                 |           |
| F1 (Hz)     | Male 345.8 38.1     | Male 369.5 47.1 | 0.299     |
|            | Female 387.4 39.8   | Female 400.0 42.3 | 0.309   |
|            | All 374.9 43.0      | All 390.9 44.9 | 0.132     |
| F2 (Hz)     | Male 905.1 294.5    | Male 1,072.3 515.1 | 0.453   |
|            | Female 988.9 409.7  | Female 954.9 174.7 | 0.771  |
|            | All 963.8 373.2     | All 990.1 306.2 | 0.796     |
| LTAS        |                    |                 |           |
| Intensity   | LF (dB)             |                 |           |
|            | Male 12.6 1.8       | Male 14.4 2.7   | 0.056     |
|            | Female 13.9 3.9     | Female 14.4 3.1 | 0.311     |
|            | All 13.5 3.4        | All 14.4 2.9   | 0.046     |
| HF (dB)     | Male −5.7 2.6       | Male −8.6 3.8  | 0.021     |
|            | Female −5.0 3.6     | Female −8.6 2.8 | 0.000     |
|            | All −5.2 3.3        | All −8.6 3.0  | 0.000     |
| α ratio     | l (dB)$_{HF}$/l (dB)$_{LF}$ |         |           |
|            | Male −0.5 0.3       | Male −0.7 0.5  | 0.120     |
|            | Female −0.5 0.4     | Female −0.6 0.3 | 0.019     |
|            | All −0.5 0.4        | All −0.7 0.3   | 0.003     |

SD, standard deviation; st, semitone; MPT, maximum phonation time; DSI, dysphonia severity index; SPL, sound pressure level; LTAS, long-term average spectrum.
Fig. 3. Vowel formants F1 and F2: male subjects (a), female subjects (b), all subjects (c).
difference was found, with lower ratio with FFP2/3 mask \((p = 0.001)\). When comparing the results without and with FFP2/3 mask by gender, significant difference was applied only to female subjects with lower ratio with FFP2/3 mask \((p = 0.004)\).

The results of the analysis of formants \(F1\) and \(F2\) of the vowels /a:/, /i:/, /u:/ without and with FFP2/3 mask are shown in Table 1. Comparing the results of mean frequency of \(F1\) of all subjects and of each gender without and with FFP2/3 mask, no significant differences were found \((p < 0.05)\). The assessment of the mean frequency of \(F2\) of vowels /a:/ and /u:/ did not show significant differences when compared without and with FFP2/3 mask \((p < 0.05)\). Significant difference in mean frequency was only found in \(F2\) of vowel /i:/ of all subjects with lower value with FFP2/3 mask \((p = 0.005)\). Comparing the mean frequency of \(F2\) of vowel /i:/ without and with FFP2/3 mask by gender, significant difference was only observed in female subjects with lower value with FFP2/3 mask \((p = 0.027)\). The data of \(F1\) and \(F2\) for /a:/, /i:/, /u:/ without and with FFP2/3 mask are illustrated in Figure 3.

**Results of LTAS of the Voice of Reading Text**

The results of the analysis of LTAS of the sound file of the reading text “The North Wind and the Sun” are listed in Table 1. Comparing the mean value of LTAS of all subjects from the frequency range 0–1,000 Hz of all subjects without and with FFP2/3 mask, it showed a significantly higher value with FFP2/3 mask \((p = 0.046)\). When comparing the values separately by gender, higher LTAS values were observed in both male and female subjects with FFP2/3 mask and that of the male subjects were on the verge of significance \((p = 0.056)\).

When comparing the mean value of LTAS from frequency range 1,000–6,000 Hz of all subjects and by gender without and with FFP2/3 mask, it showed significant difference in terms of decrease of mean value with FFP2/3 mask \((p = 0.000)\). After assessment of the \(\alpha\) measurement, it showed significant higher mean \(\alpha\)-ratio of all subjects with FFP2/3 mask \((p = 0.003)\). When comparing the mean \(\alpha\)-ratio by gender, only that of the female subjects was significantly higher with FFP2/3 mask \((p = 0.019)\).

**Discussion**

In anticipation of a possible next wave of the COVID-19 pandemic, this study should investigate the VRP measures and acoustic voice measurements as main methods of routine voice diagnostics performed without wearing a FFP2/3 mask compared to those with wearing it. Under the circumstances of the COVID-19 pandemic and facing pandemic-related legal infection-control directives such as social distancing and the necessity of PPE, wearing of face masks has become globally usual and omnipresent in hospitals. PPEs are able to limit transmission of virus between healthcare professionals and patients in healthcare settings and are recommended as key control measures in infection prevention and viral spread [27].

However, does covering of nose and mouth with FFP2/3 face mask have impact on daily routine voice diagnostics? In this study, voice range of singing and speaking voice measured by VRP did not show significant differences between measurements without and with FFP2/3 mask. Further, no significant differences were observed in parameters SPL and \(F0\) of soft speaking when measured without and with FFP2/3 mask. Neither did the routine diagnostics parameters DSI and its components reveal significant differences between measurement conditions without and with wearing FFP2/3 masks. The only exception was the SPL of shouting voice, which was significantly lower with mask in female subjects, and due to this, it had statistical impact in all subjects. It can hardly be explained by the higher \(F0\) in women than in male subjects (about 1 octave higher), but with higher \(F0\) and wider differences between the harmonics, the FFP2/3 masks might impact the female voice more than the male one. Nevertheless, both values of SPL of shouting voices without and with FFP2/3 mask were higher than 90 dB, meaning normal vocal constitution/condition.

The assessment of sustained vowels /a:/, /i:/, and /u:/ without and with mask regarding \(F1\) and \(F2\) did not show significant differences either, with the exception of \(F2\) of the vowel /i:/ in female subjects (Table 1). Since \(F2\) is highest in vowel /i:/, and that of the female subjects is higher, it was therefore significantly affected when wearing FFP2/3 mask. This in turn confirmed the attenuating effect of mask for higher spectral energy. The analysis of the eVEM without and with mask did not show significant differences. Significant differences were only found in ratio between singer’s formant, a prominent spectrum envelope peak near 3 kHz in sung voice [28], and SPL of loud singing without and with mask. The significantly lower ratio between singer’s formant and SPL of loud singing of subjects with mask was mainly due to the attenuating effect of FFP2/3 mask on voice signals above 1 kHz, since masks had limited influence on voice signals below 1 kHz [29].
by gender, the significantly lower ratio was only seen in female subjects and not in male subjects.

Assessment of the speaking voice of the reading text “The North Wind and the Sun” showed significant attenuation effect of FFP2/3 mask primarily in HF range. Since the LTAS is affected by vocal loudness [15] and due to the attenuating effect of mask in HF range, as seen in values of LTAS, which are presented in Table 1, the I in HF range was significantly lower with FFP2/3 mask compared to that without FFP2/3 mask, whereas the I in lower frequency range was on the verge of significance. When assessing the LTAS separately by gender, no significant differences were found.

With mask, the significantly lower I of LTAS in HF range and slightly higher I in low frequency (LF) range in turn resulted in higher a-ratio compared to that without mask. The significant differences in a-ratio were only seen in female subjects. Thus, the outcome of the assessment of LTAS and a-ratio complies with the outcome of ratio of singer’s formant/SPL loud singing and confirmed the attenuating effect of FFP2/3 mask in HF range. Therefore, the damping effect of FFP2/3 mask explained why subjects wearing FFP2/3 mask have perceived poorer speech intelligibility.

As shown in this study, it can be concluded that wearing FFP2/3 mask did not affect the outcome of the routine voice diagnostics and reliable data regarding VRP and sustained vowels can still be assessed wearing a FFP2/3 mask. Further, in this study, it showed that masks did not affect voice constitution (SPL with mask >90 dB) [26] and frequency range of voice diagnostics. However, due to the attenuating effect of FFP2/3 masks in HF range, when performing voice diagnostic of vowels with mask, it is recommended to focus on vowels /a:/ and /u:/, and for performing voice diagnostic of singing voice, the singer’s formant should not be taken into consideration either. Summing up, since FFP2/3 masks did barely impact on routine voice diagnostic results, their use by the patients during measurement can be recommended in accordance with public infection-control directives in order to minimize the spread of the COVID-19 infection.

**Conclusion**

In this study, wearing of FFP2/3 face masks did not affect the outcome and interpretation of routine VRP measurements. The impact of FFP2/3 on spectral analysis of sustained vowels, text, and singer’s formant needs to be considered in clinical voice diagnostics. With regard to support public health and infection control during the COVID-19 pandemic, patients should be instructed to use FFP2/3 masks during examination.

**Statement of Ethics**

This study protocol was reviewed and approved by the Ethics Committee of the Medical University of Vienna, approval number 1356/2021. All procedures performed in the studies comply with the guidelines for human studies, and the research was conducted ethically in accordance with the World Medical Association Declaration of Helsinki. Informed consent was obtained from all participants.

**Conflict of Interest Statement**

The authors have no conflicts of interest to declare.

**Funding Sources**

No funding was required for this study.

**Author Contributions**

Guan-Yuh Ho: data analysis, literature research, and preparation of the manuscript. Ines Kristina Kansy and Katharina Anna Klavacs: recruitment of patients, voice recordings, and VRP measurements. Matthias Leonhard: statistical analysis, eVEM, and manuscript proofreading. Berit Schneider-Stickler: literature research, study design, preparation of the manuscript, and study supervision.

**Data Availability Statement**

All data generated or analyzed during this study are included in this article. Further inquiries can be directed to the corresponding author.

**References**

1. Dejonckere PH, Bradley P, Clemente P, Cornut G, Crevier-Buchman L, Friedrich G, et al. A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. Guideline elaborated by the committee on phoniatrics of the European Laryngological Society (ELS). Eur Arch Otorhinolaryngol. 2001;258(2):77–82.
2. Cheng Y, Ma N, Witt C, Rapp S, Wild P, Andreae MO, et al. High efficacy of face masks explained by characteristic regimes of airborne SARS-CoV-2 virus abundance. medRxiv. 2021. [Preprint].
3 World Health Organization. Transmission of SARS-CoV-2: implications for infection prevention precautions. World Health Organization; 2020.
4 Thamboo A, Lea J, Sommer DD, Sowerby L. Airborne aerosol generation during endonasal procedures in the era of COVID-19: risks and recommendations. Otolaryngol Head Neck Surg. 2020;163(3):465–70.
5 Workman AD, Jafari A, Welling DB, Varvares MA, Gray ST, Holbrook EH, et al. Airborne aerosol generation during endonasal procedures in the context of COVID-19: an evidence based review and recommendations. Otolaryngol Head Neck Surg. 2020;163(3):465–70.
6 Morawska L, Tang JW, Bahnfleth W, Bluysen PM, Boerstra A, Buonanno G, et al. How can airborne transmission of COVID-19 indoors be minimised? Environ Int. 2020;142:105832.
7 CDC and NIOSH. Hierarchy of controls [Internet]. 2021. [cited 2021 Sep 30]. Available from: https://www.cdc.gov/niosh/topics/hierarchy/default.html.
8 Lepelletier D, Grandbastien B, Romano-Bertrand S, Aho S, Chidiac C, Gehanno JF, et al. What face mask for what use in the context of COVID-19 pandemic? The French guidelines. J Hosp Infect. 2020;105(3):414–8.
9 World Health Organization. Mask use in the context of COVID-19 interim guidance. World Health Organization; 2020.
10 Geneid A, Nawka T, Schindler A, Oguz H, Chrobok V, Calcinoni O, et al. Union of the European phoniatricians’ position statement on the exit strategy of phoniatric and laryngological services: staying safe and getting back to normal after the peak of coronavirus disease 2019 (issued on 25th May 2020). J Laryngol Otol. 2020;134(8):661–4.
11 Asadi S, Wexler AS, Cappa CD, Barreda S, Bouvier NM, Ristenpart WD. Aerosol emission and superemission during human speech increase with voice loudness. Sci Rep. 2019;9(1):2348.
12 Echternach M, Gantner S, Peters G, Westphalen C, Benthaus T, Jakubass B, et al. Impulse dispersion of aerosols during singing and speaking: a potential COVID-19 transmission pathway. Am J Respir Crit Care Med. 2020;202(11):1584–7.
13 Daflorn H, Balmer K, Breton J, Singing together, yet apart: the experience of UK choir members and facilitators during the COVID-19 pandemic. Front Psychol. 2021;12:624474.
14 Charlotte N. High rate of SARS-CoV-2 transmission due to choir practice in France at the beginning of the COVID-19 pandemic. J Voice. 2020;20892-1997(20)30452-5. Epub ahead of print.
15 Hamner L, Dubbel P, Capron I, Ross A, Jordan A, Lee J, et al. High SARS-CoV-2 attack rate following exposure at a choir practice: Skagit County, Washington, March 2020. MMWR Morb Mortal Wkly Rep. 2020;69(19):606–10.
16 Bottalico P, Murgia S, Puglisi GE, Astolfi A, Kirk KL. Effect of masks on speech intelligibility in auralized classrooms. J Acoust Soc Am. 2020;148(5):2878.
17 Magee M, Lewis C, Noffs G, Reece H, Chan JCS, Zaga CJ, et al. Effects of face masks on acoustic analysis and speech perception: implications for peri-pandemic protocols. J Acoust Soc Am. 2020;148(5):2878.
18 Maryn Y, Wuyts FL, Zaraowski A. Are acoustic markers of voice and speech signals affected by nose-and-mouth-covering respiratory protective masks? J Voice. 2021;S0892-1997(21)00037-0. Epub ahead of print.
19 Damste PH. The phonetogram. Pract Otorhinolaryngol. 1970;32(3):185–7.
20 Schutte HK, Seidner W. Recommendation by the Union of European Phoniatricians (UEP): standardizing voice area measurement/phonetography. Folia Phoniatr. 1983;35(6):286–8.
21 Titze IR. Acoustic interpretation of the voice range profile (phonetogram). J Speech Hear Res. 1992;35(1):21–34.
22 Wuyts FL, De Bodt MS, Molenberghs G, Remacle M, Heylen L, Millet B, et al. The dysphonia severity index: an objective measure of vocal quality based on a multiparameter approach. J Speech Lang Hear Res. 2000;43(3):796–809.
23 Caffier PP, Möller A, Forbes E, Müller C, Freymann ML, Nawka T. The vocal extent measure: development of a novel parameter in voice diagnostics and initial clinical experience. Biomed Res Int. 2018;2018:3836714.
24 Sundberg J, Nordenberg M. Effects of vocal loudness variation on spectrum balance as reflected by the alpha measure of long-term-average spectra of speech. J Acoust Soc Am. 2006;120(1):453–7.
25 Frokjaer-Jensen B, Prytz S. Registration of voice quality. Brüel Kjaer Tech Rev. 1976;3:1–17.
26 Schneider B, Enne R, Cecon M, Diendorfer-Radner G, Wittels P, Bigenzahn W, et al. Effects of vocal constitution and autonomic stress-related reactivity on vocal endurance in female student teachers. J Voice. 2006;20(2):242–50.
27 Gordon C, Thompson A. Use of personal protective equipment during the COVID-19 pandemic. Br J Nurs. 2020;29(13):748–52.
28 Sundberg J. Level and center frequency of the singer’s formant. J Voice. 2001;15(2):176–86.
29 Corey RM, Jones U, Singer AC. Acoustic effects of medical, cloth, and transparent face masks on speech signals. J Acoust Soc Am. 2020;148(4):2371.