Voice Acoustics and Vocal Effort in Mask-Wearing Healthcare Professionals: A Comparison Pre- and Post-Workday

*,†Victoria S. McKenna, *Tulsi H. Patel, *Courtney L. Kendall, ‡Rebecca J. Howell, and ‡Renee L. Gustin, *, †, and ‡

Summary: Objective. We evaluated voice acoustics and self-perceptual ratings in healthcare workers required to wear face masks throughout their workday.

Methods. Eighteen subjects (11 cisgender female, 7 cisgender male; M = 33.72 years, SD = 8.30) completed self-perceptual ratings and acoustic recordings before and after a typical workday. Chosen measures were specific to vocal effort, dysphonia, and laryngeal tension. Mixed effects models were calculated to determine the impact of session, mask type, sex, and their interactions on the set of perceptual and acoustic measures.

Results. The subjects self-reported a significant increase in vocal effort following the workday. These perceptual changes coincided with an increase in vocal intensity and harmonics-to-noise ratio, but decrease in relative fundamental frequency offset 10. As expected, men and women differed in measures related to fundamental frequency and vocal tract length.

Conclusion. Healthcare professionals wearing masks reported greater vocal symptoms post-workday compared to pre-workday. These symptoms coincided with acoustic changes previously related to vocal effort; however, the degree of change was considered mild. Further research is needed to determine whether vocal hygiene strategies may reduce vocal symptoms in mask-wearing workers.

Key Words: Acoustics—Adults—Effort.

INTRODUCTION

Occupational voice users comprise a group of professionals that must verbally communicate with others to meet the demands of their jobs.[1] These professions include singers, teachers, therapists, coaches, and clergy, among others. With increased voice demands, occupational voice users report greater vocal symptoms and exhibit a higher prevalence for voice-related problems compared to the general public.[2] Healthcare professionals represent a subset of occupational voice users who often endure added challenges of communicating in the presence of background noise (eg, diagnostic equipment), and of interacting with those with existing communication difficulties (eg, critically ill, hearing loss).

With the onset of the COVID-19 pandemic, many healthcare professionals are now required to wear face masks (eg, simple surgical masks, N95 respirators) to protect themselves from infection and to prevent potential disease transmission to others.[3] Because these masks are positioned over the nose and mouth, it follows that face masks negatively impact both acoustic and visual information conveyed during speech. Specifically, masks attenuate frequency information $\geq$ 2 kHz in the acoustic signal[4] which, in turn, may alter the phonemic content of the message.[5] Face masks also prohibit the use of visual cues to assist in message comprehension as they can no longer see facial expressions or use lip-reading strategies. Prior research has shown that communication exchange with face masks can be especially difficult in noisy situations[6] and/or when the communication partner has hearing loss.[7] Some masks can lead to a reduction in speech intelligibility by up to 17% using a simulated intensive care unit.[8]

Despite the clear impacts of mask-based communication on the *listener*, these prior studies fail to consider the *speaker* who has to overcome these challenges to effectively communicate. Recently, Robiero et al.[9] interviewed over 400 mask-wearing adults about their vocal health and symptoms, finding a significant increase in vocal effort, difficulty coordinating speaking and breathing, and reduction in auditory feedback when comparing speaker perceptions with and without a face mask. However, this interview study lacked objective acoustic data to link self-perceived voice problems with quantitative vocal changes. Therefore, we investigated voice acoustics and perceptual measures in mask-wearing healthcare professionals who have the added challenge of a visual and acoustic barrier during their existing occupational vocal load. Our overarching hypothesis was that mask-wearing occupational voice users would show evidence of vocal fatigue—quantified as acoustic manifestations of increased vocal effort and laryngeal muscle tension—over a single workday.

Acoustic measures of vocal effort and laryngeal tension

The impact of vocal fatigue on communication in occupational voice users has often been studied relative to the risk for developing acute and/or chronic voice problems. To

Accepted for publication April 27, 2021.
Conflict of interest: None
Presentations: Portions of this paper were presented at the Boston Speech Motor Control Symposium, June 2021.
Funding: National Center for Advancing Translational Sciences of the NIH, #2UL1TR001425-05A1
From the *Department of Communication Sciences and Disorders, University of Cincinnati; †Department of Biomedical Engineering, University of Cincinnati; and the ‡Department of Otolaryngology-Head & Neck Surgery, University of Cincinnati.
Address correspondence and reprint requests to Victoria S. McKenna, Department of Communication Sciences and Disorders, University of Cincinnati 3225 Eden Ave, Cincinnati, Ohio 45267. E-mail: mckennvs@ucmail.uc.edu
Journal of Voice, Vol. ■■, No. ■■, pp. ■■-■■
0892-1997

https://doi.org/10.1016/j.jvoice.2021.04.016

characterize fatigue, many research efforts have turned toward identifying acoustic changes following a typical work day/week or specified vocal load (eg, lecture, exercise class). Thus far, there is consistent evidence that mean fundamental frequency (f_o) and vocal intensity increase following occupation-related vocal loading[10−13] and structured vocal loading paradigms,[14−19] in which the duration, quality, and/or intensity of voice are manipulated.[20]

Remacle and colleagues proposed that an increase in mean f_o may be a normal physiological adaptation to vocal load rather than an indicator of maladaptive vocal patterns.[21] The authors also proposed that individual variability (intrinsic and extrinsic factors) may be the reason why some studies report dysphonic vocal changes (acoustic, perceptual) following occupational or experimental vocal loading, whereas others do not.[21] For example, cepstral peak prominence (CPP)—a promising acoustic measure related to the presence and severity of dysphonia[22,23]—has been shown to significantly decrease with loading,[24] or to not change at all.[14,25,26]

In both field and experimental studies, perceptual ratings of vocal effort are employed to track vocal symptoms. Vocal effort, defined as perceived vocal exertion to a perceived vocal demand,[27] is quantified via speaker ratings immediately following the vocal demand. Recent research characterized modulated vocal effort in adults with typical voices as a way to investigate acoustical and physiological manifestations of effort without confounding factors.[28−30] McKenna and Stepp[29] examined a large set of acoustical measures, determining a significant increase in vocal intensity (measured in dB SPL) as self- and listener-perceived vocal effort increased. Results also showed a significant reduction in both harmonics-to-noise ratio (HNR) and low-to-high spectral ratio (L/H ratio) for both speaker and listener-perceptual models. However, there was no change in CPP values, though previous research on modulated effort by Rosenthal et al.[30] reported a significant *increase* in CPP and its standard deviation (SD) with rising levels of self-reported effort.

A subsequent analysis of the same participants from McKenna and Stepp[29] further revealed that increased subglottal pressure estimates, medial supraglottic compression, and increased activation of suprahyoid muscles were significantly related to the self-perception of vocal effort.[28] The authors suggested that the physiological basis of vocal effort may, in fact, overlap with the symptoms of those with increased laryngeal tension and/or hyperfunctional vocal behavior. This hypothesis falls in line with the research by Remacle et al.,[21] which found that a subset (9%) of their participants exhibited acoustic responses to occupational loading that were consistent with increased glottal adduction (ie, increased HNR with simultaneous reduction in shimmer). Remacle and colleagues argued that this pattern was consistent with symptoms of hyperfunction and laryngeal tension, though laryngeal visualization and manual palpation techniques were not undertaken.

With that in mind, an investigation into the acoustic manifestation of laryngeal tension may be prudent in occupational loading studies. Relative fundamental frequency (RFF) is an acoustic indicator of laryngeal tension and potentially relevant clinical measure for those with hyperfunctional voice disorders. Unlike more traditional acoustic measures that are extracted during vowel steady-state (eg, HNR), RFF is extracted from vowel-transition states in and out of a voiceless obstruent (eg, /ifi/). The last ten voicing cycles of the initial vowel (known as *offset cycles*) and the first ten cycles of the following vowel (*onset cycles*) are theorized to be influenced by increases in laryngeal tension needed for effective devoicing.[31] The voicing cycles closest to the voiceless consonant (offset cycle 10, onset cycle 1) are most sensitive to changes in laryngeal tension. RFF values for these cycles are not only correlated with self-and/or listener-perceptual ratings of vocal effort[29,32] but also normalize in individuals with hyperfunctional voices after a successful course of voice therapy.[31]

Excessive laryngeal tension may also manifest as an elevated laryngeal position during phonation and/or rest.[33−35] Although manual laryngeal palpation is the standard clinical technique for assessing perilaryngeal muscle tension and overall laryngeal height,[36,37] estimates of vocal tract length (VTL) can be ascertained from formant acoustics.[38] Specifically, the resonant features (F2, F3, F4) of the vocal tract change based on the length of the vocal tract, as measured from the vocal folds to the opening of the lips. With our hypothesis that vocal fatigue may result in increased vocal effort and elevated suprahyoid laryngeal muscle tension, we would suspect a shortening in the acoustic VTL estimate at the end of the workday when compared to the beginning.

Purpose

Our study objectives were to: (1) summarize the frequency of self-perceived communication problems and vocal symptoms, (2) compare voice acoustics before and after occupational vocal loading, and (3) assess self-perceptual ratings of dyspnea and vocal effort before and after occupational vocal loading. We proposed the following hypotheses:

1. Healthcare professionals would report that face masks negatively impact their voice and communication in the workplace.
2. Healthcare professionals would show changes in voice acoustics post-workday that would be consistent with documented acoustic outcomes of vocal effort and laryngeal tension (eg, reduced RFF values; increased vocal intensity and mean f_o; shortening of acoustic VTL).
3. Healthcare professionals would exhibit a significant increase in self-reported vocal effort and dyspnea post-workday.

METHODS

Participants

Twenty-one adult English speaking healthcare workers wearing face masks > 6 hours/day were enrolled and

TABLE 1.
Participant Demographics

ID	Sex	Age (yrs)	Occupation	Mask Worn at Work	Workday Duration (hrs)	Relevant History
01	F	39	Physician	N95	6.50	GERD
02	F	36	Physician	N95	7.50	–
03	F	49	PT	N95	8.00	–
04	F	24	RT	N95	12.50	–
05	F	24	SLP	N95	10.50	–
06	F	43	Medical Admin.	Simple	7.00	–
07	F	35	Nurse	Simple	13.25	–
08	F	24	RT	Simple	12.00	Childhood Asthma
09	F	27	SLP	Simple	7.00	Trained singing
10	F	25	SLP	Simple	7.50	GERD; trained singing
11	F	37	SLP	Simple	7.75	–
12	M	44	Physician	N95	8.00	–
13	M	41	Physician	N95	8.00	Childhood Asthma
14	M	45	Physician	Simple	8.75	GERD, Asthma
15	M	35	PT	Simple	10.25	–
16	M	27	PT	Simple	9.25	Exercise induced Asthma
17	M	23	RT	Simple	12.50	–
18	M	29	SLP	Simple	7.75	GERD, Asthma

Abbreviations: SLP, speech-language pathologist; PT, physical therapist; RT, respiratory therapist; Admin, administrator.

completed the study from July to September of 2020. Three participants were later excluded for wearing two different masks throughout the workday (eg, wearing an N95 for 50% of the day and then a simple mask for 50%). The remaining 18 participants (11 cisgender female, 7 cisgender male, M = 33.72 years, SD = 8.30) were used for analysis. Participants reported no history of neurological disorder/disease or head/neck cancer and were free from speech, language, voice, and respiratory illness on the day of testing. Participants did not smoke or vape. Some participants had a diagnosis of gastroesophageal reflux disease (GERD) and/or a history of asthma (eg, childhood, exercise induced). To include a wide-representation of individuals for the potential generalization of this work, participants were not excluded based on these diagnoses.

On the day of evaluation, participants reported working an average of 9.1 hours/day (range = 6.5−13.25 hours; SD = 2.11) and reported using verbal communication throughout their workday. A complete description of the participant demographics is shown in Table 1. All procedures described in this study were first approved by the Institutional Review Board at the University of Cincinnati and all participants underwent the informed consent process.

Protocol

Participants completed two sessions over the course of the study. The first session occurred prior to the beginning of a typical work shift ("pre-workday") and the second session occurred immediately following the workday ("post-workday"). During each session, acoustic and perceptual data were collected from participants when seated in a quiet space. Between sessions, participants went about their usual workplace activities while wearing their usual work-required face mask. Eleven participants wore simple, disposable masks throughout their workday and seven wore N95 masks, of which, six wore an additional simple mask over it. Work settings varied and included outpatient clinics (n = 9), outpatient rehabilitation centers (n = 4), inpatient acute-care (n = 3), and skilled nursing facilities (n = 2).

During the first session, participants provided information about their work schedule and answered Likert rating scale questions to further understand their perceptions on their speech, voice, and communication while wearing face masks. The Likert ratings were from 1 to 5 with 1 representing "Never" and 5 representing "Always." Difficulty with mask-based communication and voice was operationally defined as a rating $\geq$ 4 "Almost Always" on the scale.

Acoustic recordings were made with a headset microphone (C555L) and handheld recorder (Zoom H4N). The microphone was placed 45° from the midline of the lips and 8.5 cm from the center of the mouth. A sound level meter (Extech; dB A) held at the level of the microphone was used to calibrate pure tones played from a phone app (*Pitch Generator & Frequency* [Android] or *Frequency* [iOS]; 500 Hz). Tones were played at varying intensities at the lips and measurements of known intensity levels (dB SPL) were taken at the microphone. Acoustic data were digitized at 44.1 kHz and 16 bits.

Participants completed all recordings at their typical pitch and loudness without a face mask. Speech tasks included: the first paragraph of the Rainbow passage, vowel-voiceless consonant-vowel (VCV) utterances (eg, /ifi/, /ɑfɑ/, /ufu/), and words with target corner vowels /i/, /ɑ/, and /u/ in both the sustained vowel, single word, and sentence context (eg, "eee", "heed", "I wish he would heed my advice"). See Appendix A for a list of all speech tasks.

Upon completion of the recordings, participants made perceptual ratings of dyspnea and vocal effort. Dyspnea ratings were made on the modified Borg scale for Dyspnea, adapted from Borg's original category-ratio scale for perceived exertion.[39,40] The modified scale has a breathing-specific prompt: "How much difficulty is your breathing causing you right now?" The scale ranges from 0 to 10, in which a rating of 0 represents no difficulty in breathing and a value of 10 represents maximal difficulty breathing. Next, a vocal effort rating, defined for each participant as an "exertion of the voice or how hard you have to try to make a voice," was completed by placing a single mark on a 100-mm visual-analog scale (VAS).[24,41–43] The left side of the scale was anchored as "No Effort" and the right side as the "Most Effort."[26,29]

Data processing

Spectral and cepstral measurements

The second two sentences of the Rainbow passage were isolated and pauses greater than 150-ms were removed. All recordings were processed using the Rainbow protocol in the Analysis of Dysphonia and Speech (ADSV; ver 4.0) software.

Prior to analysis, pitch settings were adjusted for each participant's sex (ie, 60−300Hz for male; 90−500Hz for female) and vocalic detection was activated to limit extraction to voiced segments. Otherwise, standard software settings and processing was employed. In brief, the acoustic signal was down sampled to 22,050 Hz from the original 44.1 kHz acquisition rate. L/H Ratio and its SD (L/H SD) were determined from the spectrum of the acoustic signal with a cut-off frequency of 4000 Hz, as these spectral measures are correlated with self-perceptual metrics of effort and dysphonia.[29] A fast Fourier transform (FFT) of the log power spectrum (Hamming sliding window at 1024 frames, 75% overlap) created a smoothed cepstrum (averaged over 7 points). Measures of CPP, defined as the amplitude of the rahmonic to the regressed linear fit of the signal, was determined for the quefrency corresponding to the predicted f_o of the signal. From here values of CPP (dB), CPP SD (dB), CPP f_o (Hz), and CPP f_o SD (Hz) were extracted for each passage. CPP f_o was our estimate of mean f_o for this study.

Relative fundamental frequency (RFF)

RFF values were determined for each VCV utterance (eg, /ifi/) using a semi-automated algorithm[44] in MATLAB (The MathWorks; Natick, MA). A series of 20 values were extracted for each VCV utterance to characterize the change in instantaneous f_o during devoicing (vowel to voiceless consonant) and the re-initiation of voicing (voiceless consonant to vowel). Each value (Hz) was then normalized to a vowel steady-state reference value, and converted to semitones (ST). Values for voicing offset cycle 10 ("offset 10") and voicing onset cycle 1 ("onset 1") were retained for further analysis, as these two cycles are furthest from the reference cycle of the value and shown to be sensitive to changes in laryngeal tension and effort.[29,32,45] Offset 10 and onset 1 values were averaged across multiple productions into a single value for each participant for each study session. On average, 6.83 offset 10 values and 6.78 onset 1 values were available across participants.

Vowel acoustics: HNR (dB), Vocal intensity (dB SPL), formant-based estimates of VTL

The mid-sections of vowels were manually segmented, and acoustic information was extracted using Praat software.[46] In all cases, the onset and offset of the vowel and transitions into other phonemes were not included. Using the voice report function, HNR values were extracted for each vowel segment and averaged together into a single value for each participant. Vocal intensity (dB SPL) was also extracted using Praat functionality and then calibrated to the known vocal intensity captured during the calibration procedure with the Sound Pressure Level meter.

VTL was estimated from the third and fourth formants (F3 and F4) for each target vowel, as higher formants are less likely to be impacted by speech articulation.[47] First, F3 and F4 values were each averaged separately and then the averaged value was used to estimate VTL (see Eq. 1). In this equation, c represents the speed of sound in air, n is the formant number, and f_n is the formant value in Hz. Finally, the new VTL (cm) estimate for F3 and F4 was averaged into a single value for statistical analysis.

$$f_n = \frac{(2n-1) \times \mathrm{c}}{4 \times \mathrm{VTL}} \tag{1}$$

Prior to experimental data extraction, three researchers (V.M., T.P., and C.K.) underwent reliability training, yielding excellent inter-rater reliability (intraclass correlation coefficient [ICC] > .95). Next, experimental data sets were divided amongst the researchers. At least three weeks after initial data extraction, inter- and intra-rater reliability analyses were completed. One participant was randomly selected for blinded data extraction by all researchers resulting in inter-rater reliability ICC(2,1) = .94−.99, depending on the acoustic measure. For intra-rater reliability, one subject was randomly selected for each researcher for blinded re-extraction. Intra-rater reliability values ranged from ICC = .87 −.99, across all researchers and measures. Therefore, inter- and intra-rater reliability were considered good-to-excellent across all manually extracted measures.[48]

Perceptual measures

Data were acquired on hand-written data sheets. VAS ratings were manually measured to the nearest mm and all data were transferred to an excel document. Three weeks following the last participant, the first author blindly re-extracted all measurements. All dyspnea ratings were identical and effort ratings were within 1-mm of the original rating. In the cases where a 1-mm discrepancy arose, the original rating was kept and used for analysis.

Statistical analysis

Mixed-effect models were performed to analyze each acoustic and perceptual measure separately. Participant was set as a random effect, whereas session (pre-, post-workday), sex (male, female), and mask type (N95, Simple) were included as fixed effects. Two-way interactions were examined between fixed effects; however, three-way interactions were not investigated as the present study was not powered to do so. Significance level for each effect was $P < 0.05$. Tukey's *post hoc* analyses were performed to evaluate pairwise comparisons for significant interaction effects. Tukey's test adjusts for family-wise error at the time of the comparisons to reduce Type I error. Therefore, significance level of $P < 0.05$ was set for *post hoc* comparisons as well. Statistical analyses were completed in Minitab Statistical Software (ver. 19) whereas reliability metrics were calculated in *R* (ver. 4.0.2).

RESULTS

Summary statistics

Participants rated questions pertaining to their personal communication and vocal health a 3 ("Sometimes;" median = 3). The highest average rating of 4.06 ("Almost Always"), was reported in response to "I find I have to use more effort to talk while wearing a mask." The frequency that participants employed communication strategies (eg, hand gestures, speaking less, removing masks) was rarely-to-sometimes. Participants scored the lowest when asked about removing their face masks to communicate, with a 2.22 average rating, indicating that they rarely removed their masks at work. The questions and summary statistics can be found in Table 2.

TABLE 2.
Likert Ratings on Perceptions of Mask-Based Communication and Voice

	Rating			
Statement	Mean (SD)	Mdn.	Min	Max
Wearing a mask makes it difficult for people to hear me	3.39(0.70)	3	2	5
Wearing a mask causes me to need to repeat myself more	3.33(0.69)	3	2	5
I have trouble understanding people when they are wearing a mask	2.89(0.68)	3	2	4
I find I have to use more effort to talk while wearing a mask	4.06(1.00)	4	2	5
Wearing a mask is negatively impacting my communication with others at my job	2.89(1.18)	3	1	5
I feel that my voice is more tired at the end of the day when I wear a mask	2.94(1.35)	2	1	5
Wearing a mask makes it harder for me to catch my breath when talking	3.33(0.97)	3	2	5
I find myself using more hand gestures so I can be understood while wearing a mask	3.11(1.18)	3	1	5
I find I tend to speak a lot less while wearing a mask	2.50(1.20)	2	1	5
I find myself removing my mask to communicate when I am at work	2.22(0.81)	2	1	3

Note. Ratings were made on a 5-point Likert scale with the following designations:
1: Never 2: Rarely 3: Sometimes 4: Almost Always 5: Always.

Statistical results

Each mixed-effects model underwent an assessment of normality and homoscedasticity of residual distributions, yielding appropriate model fits for all data.

Acoustic

The main effect of session was significant for vocal intensity (P = 0.017) with greater intensities exhibited post-workday (M = 84.67 dB SPL) compared to pre-workday (M= 83.00 dB SPL). Likewise, a significant increase in HNR (P = 0.003) was noted post-workday (M = 19.53) compared to pre-workday (M = 18.53 dB). RFF offset 10, on the other hand, significantly decreased (P = 0.001) at the end of the workday (pre-workday: M = -0.74 ST; post-workday: M = -1.01 ST). Otherwise, there was no effect of session on any acoustic measure.

A significant impact of sex was found for CPP f_o, CPP f_o SD, and CPP SD with females exhibiting higher values than males. As expected, acoustic VTL was longer in males compared to females ($P < 0.001$).

The impact of mask type was only significant for L/H SD (P = 0.036); participants who wore N95 masks exhibited larger L/H SD (M = 8.51 dB) compared to those who wore simple masks throughout the day (M = 7.89 dB). Furthermore, there was a significant interaction effect of mask type × sex (P = 0.025). *Post hoc* analysis revealed that males who wore N95s exhibited larger L/H SD than males who wore simple masks (P_{adj} = .043). There were no other significant comparisons.

Two-way interaction effects were found for RFF offset 10. Specifically, session × type and session × sex yielded P values of 0.004 and 0.002, respectively. Tukey's pairwise comparisons determined that participants who wore N95 masks exhibited a significant reduction in RFF offset 10 post-workday compared to pre-workday (P_{adj} = 0.001), with no other significant comparisons. *Post hoc* analysis showed significantly lower RFF offset 10 values post-workday when compared to pre-workday recordings for male participants (P_{adj} = 0.001). Once again, no other pairwise differences were found. Figure 1 provides interaction plots for RFF offset 10.

CPP, L/H ratio, and RFF onset 1 models yielded no significant findings for any main or interaction effect. See Table 3 for summary (Mean, SD) of data for each fixed factor with significantly different pairs designated with asterisks.

Perceptual

Self-perceptual ratings were unavailable for one male participant, leaving a total of 17 datasets for analysis. The impact of session was significant for vocal effort (P = 0.012), with an increase in effort ratings post- (M = 27 mm) compared to pre-workday (M = 10 mm), but not significant for self-ratings of dyspnea (P = 0.059). There were no effects of mask type or sex, as well as no interaction effects found for either perceptual rating.

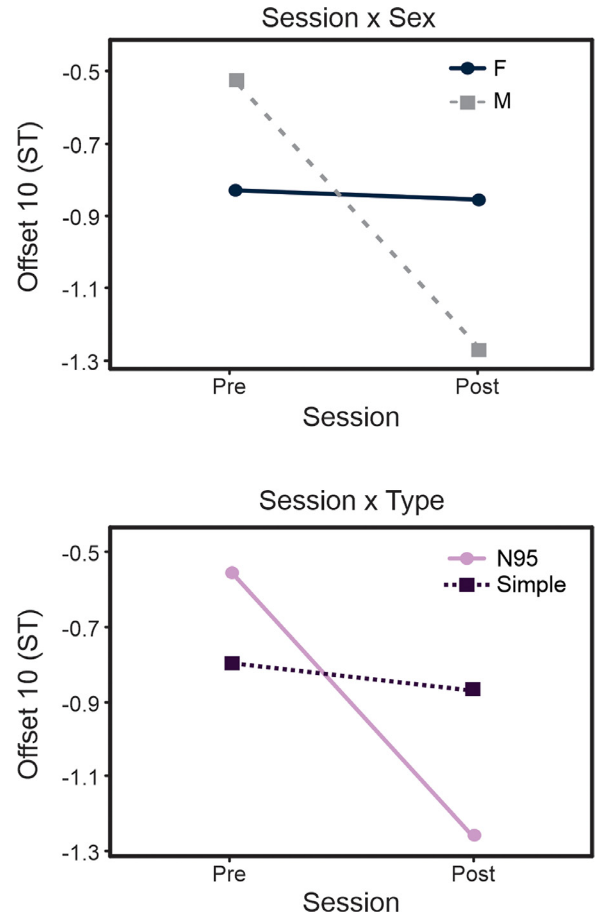

FIGURE 1. Interaction plots for RFF Offset 10.

DISCUSSION

We completed an investigation into the communication experiences, acoustics, and self-perceptual vocal ratings of mask-wearing healthcare professionals. Our overarching hypothesis was that mask wearers would experience communication and voice difficulties along with acoustical changes consistent with previous research on vocal effort and laryngeal tension.

It was first hypothesized that participants would experience increased communication and vocal difficulties while wearing masks in the workplace (defined as a rating ≥ 4, on a 5-point Likert scale). This was only

TABLE 3.
Mean (SD) for Acoustical Measures by Session, Mask Type, and Sex

	Session		Mask Type		Sex	
Measure	Pre (n = 18)	Post (n = 18)	N95 (n = 7)	Simple (n = 11)	Male (n = 7)	Female (n = 11)
CPP (dB)	7.41(1.22)	7.65(1.02)	7.43(0.73)	7.59(1.32)	7.11(1.13)	7.80(1.05)
CPP SD (dB)	3.72(0.70)	3.85(0.88)	3..89(0.84)	3.73(0.76)	3.17(0.90)*	4.19(0.34)*
L/H Ratio (dB)	40.35(2.93)	40.81(2.65)	41.02(2.96)	40.30(2.67)	42.00(2.97)	39.67(2.27)
L/H SD (dB)	8.22(1.53)	8.04(0.82)	8.51(1.64)*	7.89(0.78)*	8.32(1.69)	8.01(0.79)
Offset 10 (ST)	-0.74(0.81)*	-1.01(0.80)*	-1.03(0.61)	-0.78(0.91)	-1.00(0.74)	-0.81(0.85)
Onset 1 (ST)	2.09(1.29)	1.95(0.64)	1.75(0.68)	2.23(1.18)	2.46(0.54)	1.82(1.12)
CPP f_o (Hz)	201.00(64.40)	204.57(60.23)	212.28(60.86)	196.75(62.58)	128.70(17.03)*	249.93(21.61)*
CPP f_o SD (Hz)	86.84(42.13)	81.81(42.72)	94.98(35.38)	77.55(45.19)	39.32(18.49)*	112.96(25.08)*
Intensity (dB SPL)	83.00(3.95)*	84.67(3.89)*	83.96(3.41)	83.75(4.34)	85.02(4.55)	83.08(3.41)
HNR (dB)	18.53(3.67)*	19.53(3.55)*	19.10(4.25)	18.99(3.20)	17.02(2.75)	20.31(3.57)
VTL (cm)	16.35(0.97)	16.24(0.91)	16.12(0.96)	16.41(0.92)	17.22(0.64)*	15.71(0.55)*

* Significantly different at $P < 0.05$.Abbreviations: CPP, cepstral peak prominence; L/H, low-to-high. RFF, relative fundamental Frequency; f_o, fundamental frequency; SPL, sound pressure level; HNR, harmonics-to-noise ratio; VTL, vocal tract length; SD, standard deviation; ST, semitone.

partially supported since participants reported that they only “sometimes” (score of 3) experienced difficulties with shortness of breath, communication, and being understood. However, the frequency at which healthcare professionals experienced vocal effort (M = 4.06, “almost always”), confirmed this hypothesis.

Similiar to the Likert scale ratings, participants reported a significant increase in instantaneous ratings of perceived vocal effort on the 100 mm VAS post-workday (change from 10 mm to 27 mm). Compared to previous reports, average post-workday ratings of perceived effort would be considered “mild,” as they are quite similar to self-cued productions of “mild increases” in vocal effort which resulted in an average rating of 26 mm.[29] If we combine the information from the Likert scores and self-perceptual ratings, our results suggest that healthcare professionals almost always experienced mild symptoms throughout their workday. The cumulative impact of mild vocal symptoms on long-term vocal health outcomes is not known and requires further investigation.

Conversely, there was no significant impact of session on dyspnea values (change from 0.21 to 0.68) indicating only a “very, very slight (just noticeable)” feeling of dyspnea via the Modified Borg rating descriptors of rating point 0.5. However, participants reported that they “sometimes” felt masks made it difficult for them to catch their breath, but it seems that those feelings subside once the mask is removed.

Despite the number of acoustical parameters investigated, only a few acoustic outcomes changed from pre- to post-workday. The post-workday increase in vocal intensity aligns with previous reports on self-modulated vocal effort as well as acoustic changes following vocal loading.[14,19,29,49] The increase in intensity may occur as a compensatory vocal response to having difficulty being heard. The statement “Wearing a mask makes it more difficult for people to hear me” achieved an average rating of 3.39, indicating “sometimes” to “almost always” experiencing difficulty with vocal volume. HNR also increased post-workday, which is likely directly related to increased vocal intensity and is congruent with previous work on workday fatigue in elementary school teachers.[21] Remacle and colleagues[21] proposed that an increase in HNR may be indicative of hyperfunctional behavior; however, our study did not visualize laryngeal function pre- or post-workday and would benefit from additional investigations linking HNR to glottal behavior.

Our study is the first to find RFF offset 10 changes following occupational loading, as other loading paradigms and long-term tracking have not yielded any differences.[24,50] RFF offset 10 values decreased from -0.74 ST to -1.01 ST, similar to prior reports by McKenna and Stepp[29] of moderate levels of vocal effort (M = -1.07 ST). Further investigation into our *post hoc* analysis showed that the reduction was driven by male participants and those wearing N95 masks. It may be that N95 masks create a challenging vocal barrier that leads to maladaptive tension patterns. Yet, our investigation into acoustic VTLdid not change and does not support the addition of suprahyoid tension and/or changes to laryngeal height. Further investigation into the presentation of laryngeal tension in mask-wearing workers is needed.

The lack of additional acoustic findings is consistent with the body of work on occupational and experimental vocal loading in which acoustic changes are quite variable. For example, Sandage et al.[26] reported an increase in self-reported vocal effort and vocal symptoms (loss of voice, hoarseness and pain in 75% to 100% of participants), but no statistically significant changes in CPP or mean f_o, pre and post sorority recruitment event. Our results were similar with no changes to CPP or the CPP-based estimate of mean f_o, (ie, CPP f_o). Still, the lack of change in CPP f_o post-workday is surprising because of the consistent evidence for increased mean f_o following vocal loading.[14–19] It may be that masks present a specific challenge in which the vocal response does not manifest as a change in mean f_o. For example, it is possible that masks reduce lip and jaw movement as well as dampen specific resonant frequencies, making the compensatory vocal responses unique to that specific barrier. That may have also contributed to the changes in L/H SD observed across specific mask types, with more variability in energy production distribution across the spectrum in N95 wearers. It would be beneficial to elucidate the effects of mask-wearing on articulatory movements and specific resonant frequencies during speech.

Limitations and future directions

Healthcare professionals have medical knowledge that may have influenced their overall health behavior and vocal outcomes. More specifically, inclusion of SLPs—who have substantial knowledge on vocal health and hygiene programs—could have impacted our results by either being more sensitive to their vocal behavior and/or by staving off vocal fatigue by implementing hygiene programs. Yet, we think this is unlikely to have affected the outcomes of the current study because a visual inspection of the SLP data yielded no substantial differences between SLPs and other healthcare professionals. These results align with previous findings that SLPs can still experience vocal fatigue and do not necessarily implement hygiene programs.[51]

There are several additional directions for this work. First, the impacts of participant sex and mask type were only investigated in a relatively small number of subjects due to our analysis of unbalanced sub-groups. An enrollment of 26 participants per group would allow for detection of a large effect size (Cohen’s d = 0.80) via a simple t-test comparison of independent means (α = 0.05, power = 0.80). Second, participants reported that they verbally communicated throughout their workday, yet no monitoring or tracking of vocal dose was employed. Next steps could include ambulatory monitoring[52,53] and/or calculation of vocal dose[54,55] over the course of the day, as well as enrolling a control group comparison of maskless healthcare professionals (eg, those using telehealth during the pandemic). Third, there may be additional environmental (ie, amount of background noise) and vocational factors that could contribute to the presence and severity of vocal symptoms. For example, audiologists and otolaryngologists frequently encounter patients with hearing loss, which could be an added communication challenge over the workday. An examination of these additional factors may identify specific professionals at higher risk for developing vocal problems. Further, investigations into the impact of emotional and/or cognitive stress and effort should be considered. Finally, a longitudinal study on longer-term communication and vocal health outcomes is necessary in a larger cohort of mask-wearing workers.

CONCLUSION

This study provided information on vocal measures in mask-wearing healthcare workers after a single workday. We found a significant increase in self-rated vocal effort post-workday with further report of feeling increased vocal effort at a frequency of "almost always." These findings were consistent with post-workday acoustic changes (eg, reduced RFF offset 10, increased HNR and vocal intensity). Still, the degree of these changes was considered relatively mild. Additional research is needed to determine whether these changes are greater than those that exist for occupational voice users who are not required to wear masks throughout their day. Future investigations should also examine the longer-term impacts of mask use on vocal outcomes and whether vocal health strategies improve symptoms of vocal effort in mask-wearing healthcare workers.

ACKNOWLEDGMENTS

The project was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health, under Award Number 2UL1TR001425-05A1. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. The authors would like to thank Teresa Hollenkamp, M.A., CCC-SLP for her assistance with participant recruitment.

REFERENCES

1. Titze IR, Lemke J, Montequin D. Populations in the U.S. workforce who rely on voice as a primary tool of trade: a preliminary report. *J Voice*. 1997;11:254–259.
2. Roy N, Merrill RM, Thibeault S, et al. Prevalence of voice disorders in teachers and the general population. *J Speech Lang Hear Res*. 2004;47:281–293.
3. Bartoszko JJ, Farooqi MAM, Alhazzani W, et al. Medical masks vs N95 respirators for preventing COVID-19 in healthcare workers: A systematic review and meta-analysis of randomized trials. *Influenza Other Respir Viruses*. 2020;14:365–373.
4. Corey RM, Jones U, Singer AC. Acoustic effects of medical, cloth, and transparent face masks on speech signals. *J Acoust Soc Am*. 2020;148:2371–2375.
5. Monson BB, Hunter EJ, Lotto AJ, et al. The perceptual significance of high-frequency energy in the human voice. *Front Psychol*. 2014;5:1–11.
6. Mendel LL, Gardino JA, Atcherson SR. Speech understanding using surgical masks: A problem in health care? *J Am Acad Audiol*. 2008;19:686–695.
7. Champion J, Holt R. Dental care for children and young people who have a hearing impairment. *Br Dent J*. 2000;189:155–159.
8. Radonovich LJ, Yanke R, Cheng J, et al. Diminished speech intelligibility associated with certain types of respirators worn by healthcare workers. *J Occup Environ Hyg*. 2009;7:63–70.
9. Ribeiro VV, Dassie-Leite AP, Pereira EC, et al. Effect of wearing a face mask on vocal self-perception during a pandemic. *J Voice*. 2020 S0892199720303568. https://doi.org/10.1016/j.jvoice.2020.09.006. Published online.
10. Dallaston K, Rumbach AF. Vocal performance of group fitness instructors before and after instruction: changes in acoustic measures and self-ratings. *J Voice*. 2016;30:127.e1–127.e8.
11. Laukkanen A-M, Ilomäki I, Leppänen K, et al. Acoustic measures and self-reports of vocal fatigue by female teachers. *J Voice*. 2008;22:283–289.
12. Lehto L, Laaksonen L, Vilkman E, et al. Occupational voice complaints and objective acoustic measurements—do they correlate? *Logoped Phoniatr Vocol*. 2006;31:147–152.
13. Rantala L, Vilkman E, Bloigu R. Voice changes during work. *J Voice*. 2002;16:344–355.
14. Anand S, Bottalico P, Gray C. Vocal fatigue in prospective vocal professionals. *J Voice*. 2019. Published onlinein press.
15. Boominathan P, Anitha R, Shenbagavalli M, et al. Voice characteristics and recovery patterns in indian adult males after vocal loading. *J India Inst Speech Hear*. 2010;29:220–231.
16. Remacle A, Finck C, Roche A, et al. Vocal impact of a prolonged reading task at two intensity levels: objective measurements and subjective self-ratings. *J Voice*. 2012;26:e177–e186.
17. Stemple JC, Stanley J, Lee L. Objective measures of voice production in normal subjects following prolonged voice use. *J Voice*. 1995;9:127–133.
18. Vintturi J, Alku P, Lauri E-R, et al. The effects of post-loading rest on acoustic parameters with special reference to gender and ergonomic factors. *Folia Phoniatr Logop*. 2001;53:338–350.
19. Whitling S, Rydell R, Lyberg Åhlander V. Design of a clinical vocal loading test with long-time measurement of voice. *J Voice*. 2015;29:261.e13–261.e27.
20. Fujiki RB, Sivasankar MP. A review of vocal loading tasks in the voice literature. *J Voice*. 2017;31:388.e33–388.e39.
21. Remacle A, Garnier M, Gerber S, et al. Vocal change patterns during a teaching day: Inter- and intra-subject variability. *J Voice*. 2018;32:57–63.
22. Awan SN, Roy N, Dromey C. Estimating dysphonia severity in continuous speech: Application of a multi-parameter spectral/cepstral model. *Clin Linguist Phon*. 2009;23:825–841.
23. Heman-Ackah YD, Sataloff RT, Laureyns G, et al. Quantifying the cepstral peak prominence, a measure of dysphonia. *J Voice*. 2014;28:783–788.
24. Fujiki RB, Chapleau A, Sundarrajan A, et al. The interaction of surface hydration and vocal loading on voice measures. *J Voice*. 2017;31:211–217.
25. Gorham Rowan M, Berndt A. The effect of a vocal loading task on vocal function before and after 24 hours of thickened liquid use. *J Speech Pathol Ther*. 2016;1:1–5.
26. Sandage MJ, Hamby HA, Barnett LA, et al. Vocal function differences before and after sorority recruitment. *J Voice*. 2020. Published onlinein press.
27. Hunter EJ, Cantor-Cutiva LC, van Leer E, et al. Toward a consensus description of vocal effort, vocal load, vocal loading, and vocal fatigue. *J Speech Lang Hear Res*. 2020;63:509–532.

28. McKenna VS, Diaz-Cadiz ME, Shembel AC, et al. The relationship between physiological mechanisms and the self-perception of vocal effort. *JSLHR*. 2019;62:815–834.
29. McKenna VS, Stepp CE. The relationship between acoustical and perceptual measures of vocal effort. *J Acoust Soc Am*. 2018;144:1643–1658.
30. Rosenthal AL, Lowell SY, Colton RH. Aerodynamic and acoustic features of vocal effort. *J Voice*. 2014;28:144–153.
31. Stepp CE, Merchant GR, Heaton JT, et al. Effects of voice therapy on relative fundamental frequency during voicing offset and onset in patients with vocal hyperfunction. *J Speech Lang Hear Res*. 2011;54:1260–1266. https://doi.org/10.1044/1092-4388(2011/10-0274.
32. Eadie TL, Stepp CE. Acoustic correlate of vocal effort in spasmodic dysphonia. *Ann Otol Rhinol Laryngol*. 2013;122:169–176.
33. Lowell SY, Kelley RT, Colton RH, et al. Position of the hyoid and larynx in people with muscle tension dysphonia. *The Laryngoscope*. 2012;122:370–377. https://doi.org/10.1002/lary.22482.
34. Roy N, Ferguson N. Formant frequency changes following manual circumalaryngeal therapy for functional dysphonia: evidence of laryngeal lowering? *J Med Speech-Lang Pathol*. 2001;9:169–175.
35. Morrison M. Pattern recognition in muscle misuse voice disorders: how I do it. *J Voice Off J Voice Found*. 1997;11:108–114. https://doi.org/10.1016/s0892-1997(97)80031-8.
36. Khoddami SM, Ansari NN, Jalaie S. Review on laryngeal palpation methods in muscle tension dysphonia: Validity and reliability issues. *J Voice Off J Voice Found*. 2015;29:459–468. https://doi.org/10.1016/j.jvoice.2014.09.023.
37. Roy N. Assessment and treatment of musculoskeletal tension in hyperfunctional voice disorders. *Int J Speech Lang Pathol*. 2008;10:195–209.
38. Lammert AC, Narayanan SS. On short-time estimation of vocal tract length from formant frequencies. Larson CR, ed *PLOS ONE*. 2015;10:e0132193. https://doi.org/10.1371/journal.pone.0132193.
39. Borg G. Psychophysical bases of perceived exertion. *Med Sci Sports Exerc*. 1982;14:377–381.
40. Borg G. Psychophysical scaling with applications in physical work and the perception of exertion. *Scand J Work Environ Health*. 1990;16:55–58.
41. Sandage MJ, Connor NP, Pascoe DD. Voice function differences following resting breathing versus submaximal exercise. *J Voice*. 2013;27:572–578.
42. Sundarrajan A, Fujiki RB, Loerch SE, et al. Vocal loading and environmental humidity effects in older adults. *J Voice*. 2017;31:707–713.
43. Tanner K, Fujiki RB, Dromey C, et al. Laryngeal desiccation challenge and nebulized isotonic saline in healthy male singers and nonsingers: Effects on acoustic, aerodynamic, and self-perceived effort and dryness measures. *J Voice*. 2016;30:670–676.
44. Lien Y-AS, Heller Murray ES, Calabrese CR, et al. Validation of an algorithm for semi-automated estimation of voice relative fundamental frequency. *Ann Otol Rhinol Laryngol*. 2017;126:712–716.
45. McKenna VS, Heller Murray ES, Lien Y-AS, et al. The relationship between relative fundamental frequency and a kinematic estimate of laryngeal stiffness in healthy adults. *JSLHR*. 2016;59:1283–1294.
46. Boersma P. Praat, a system for doing phonetics by computer. *Glot Int*. 2001;5:341–345.
47. Wakita H. Normalization of vowels by vocal-tract length and its application to vowel identification. *IEEE Trans Acoust Speech Signal Process*. 1977;25:183–192. https://doi.org/10.1109/TASSP.1977.1162929.
48. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15:155–163.
49. Vilkman E, Lauri E-R, Alku P, et al. Effects of prolonged oral reading on F0, SPL, subglottal pressure and amplitude characteristics of glottal flow waveforms. *J Voice*. 1999;13:303–312.
50. Heller Murray ES, Hands GL, Calabrese CR, et al. Effects of adventitious acute vocal trauma: Relative fundamental frequency and listener perception. *J Voice*. 2016;30:177–185.
51. Joseph BE, Joseph AM, Jacob TM. Vocal fatigue—do young speech-language pathologists practice what they preach? *J Voice*. 2020;34(4):647.e1–647.e5.
52. Lyberg-Åhlander V, Pelegrín García D, Whitling S, et al. Teachers' voice use in teaching environments: A field study using ambulatory phonation monitor. *J Voice*. 2014;28:841.e5–841.e15.
53. Van Stan JH, Mehta DD, Zeitels SM, et al. Average ambulatory measures of sound pressure level, fundamental frequency, and vocal dose do not differ between adult females with phonotraumatic lesions and matched control subjects. *Ann Otol Rhinol Laryngol*. 2015;124:864–874.
54. Švec J, Popolo P, Titze I. Measurement of vocal doses in speech: experimental procedure and signal processing. *Logoped Phoniatr Vocol*. 2003;28:181–192.
55. Titze IR, Švec JG, Popolo PS. Vocal dose measures: quantifying accumulated vibration exposure in vocal fold tissues. *J Speech Lang Hear Res*. 2003;46:919–932.

APPENDIX A. SPEECH TASKS

1. Sustained vowels x 5 seconds in duration: /i/, /ɑ/, /u/
2. Rainbow Passage
3. Single words and sentences:
   Heed Heed Heed
   I wish he would heed my advice.
   Who'd Who'd Who'd
   I asked myself, "Who'd do that?"
   Hod Hod Hod
   A brick hod is a three-sided box.
   My father hid food to feed the cat on Tuesday morning.
   The cat happened to see the food, my father had hid in her pod, at noon time.
   The fat cat was hot from her sleep in the noon sun beams.