The Laryngoscope

Impact of Face Masks on Speech Acoustics and Vocal Effort in Healthcare Professionals

Victoria S. McKenna, PhD, CCC-SLP; Courtney L. Kendall, BS; Tulsi H. Patel, BS;
Rebecca J. Howell, MD; Renee L. Gustin, MS, CCC-SLP

Objectives/Hypothesis: We investigated speech acoustics and self-reported vocal symptoms in mask-wearing healthcare professionals. We hypothesized that there would be an attenuation of spectral energies and increase in vocal effort during masked speech compared to unmasked speech.

Study Design: Within and between subject quasi-experimental design.

Methods: We prospectively enrolled 21 healthcare providers (13 cisgender female, 8 cisgender male; $M = 32.9$ years; $SD = 7.9$ years) and assessed acoustics and perceptual measures with and without a face mask in place. Measurements included: 1) acoustic Vowel Articulation Index (VAI); 2) cepstral and spectral acoustic measures; 3) traditional vocal measures (e.g., fundamental frequency, intensity); 4) relative fundamental frequency (RFF); and 5) self-reported ratings of vocal effort and dyspnea.

Results: During masked speech, there was a significant reduction in VAI, high-frequency information (>4 kHz), and RFF offset 10, as well as a significant increase in cepstral peak prominence and perceived vocal effort. Further analysis showed that high-frequency attenuation was more pronounced when wearing an N95 mask compared to a simple mask.

Conclusions: Face masks pose an additional barrier to effective communication that primarily impacts spectral characteristics, vowel space measures, and vocal effort. Future work should evaluate how long-term mask use impacts vocal health and may contribute to vocal problems.

Key Words: Speech-language pathology, acoustics, voice disorders.

Level of Evidence: 3

***Laryngoscope*, 00:1–7, 2021**

INTRODUCTION

Healthcare professionals represent a subset of individuals who must communicate verbally to meet their work demands, otherwise known as *occupational voice users*.[1] It is well established that occupational voice users develop vocal problems at higher rates than nonoccupational voice users, often leading to negative socioeconomic and emotional consequences.[1–8] In 2014, a survey of 362 healthcare workers (e.g., physicians, speech-language pathologists) reported that nearly 50% frequently spoke in a loud voice.[9] Furthermore, 15% of those healthcare workers who experienced a voice disorder in the past year attributed their voice issues to a high vocal load. Despite the high demand for communication in the workplace, few studies have examined vocal risk factors pertaining to healthcare professionals.

Over the past year, healthcare professionals have been required to wear face masks to reduce the transmission of COVID-19.[10] By obscuring the nose and mouth of the speaker, face masks pose a unique challenge to effective communication exchange for the listener; this includes inhibiting one's ability to lip-read, see facial expressions, and extract nonverbal pragmatic intent. Wong et al.[11] reported that mask-wearing healthcare professionals have more difficulty establishing patient rapport compared to unmasked professionals.

Although there are several studies that describe the impact of face masks on the listener,[12–14] there is a need to understand how the *speaker* may be impacted during communication exchanges. An analysis of artificial speech stimuli demonstrated that masks attenuate speech frequencies between 2 and 7 kHz at a range of 3–12 dB, depending on mask type.[15] Consistent with this finding, Nguyen et al.[16] reported attenuation of speech frequencies between 1 and 8 kHz with a greater impact from N95 masks compared to surgical masks in a group of 16 speakers. Investigation into specific spectral bands (1 kHz bins up to 10 kHz) revealed that N95 masks attenuate frequencies from 3 to 10 kHz, whereas simple masks do not impact speech frequencies until about 5 kHz.[17] As high-frequency spectral information is critical for speech intelligibility,[18] it is possible that speakers may be making compensatory articulatory adjustments to maintain

From the Department of Communication Sciences and Disorders (V.S.M., C.L.K., T.H.P.), University of Cincinnati, Cincinnati, Ohio, U.S.A.; Department of Biomedical Engineering (V.S.M.), University of Cincinnati, Cincinnati, Ohio, U.S.A.; and the Department of Otolaryngology-Head and Neck Surgery (R.J.H., R.L.G.), University of Cincinnati, Cincinnati, Ohio, U.S.A.

Editor's Note: This Manuscript was accepted for publication on July 08, 2021.

Portions of this work were presented at the Boston Speech Motor Control Symposium in June of 2021.

The project was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health, under Award Number 2UL1TR001425-05A1.

The authors have no other funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Victoria S. McKenna, PhD, CCC-SLP, 3225 Eden Avenue, Cincinnati, OH 45267. E-mail: mckennvs@ucmail.uc.edu

DOI: 10.1002/lary.29763

effective communication. To our knowledge, no study has investigated the effect of masks on speech articulation.

In addition to the paucity of research on speech acoustics, there is a general lack of understanding of how masks impact vocal function. A recent survey of over 400 participants found that mask-wearing adults report increased vocal effort as well as difficulty coordinating speaking and breathing.[19] However, it does not seem that face masks affect vocal perturbation or breathiness correlates (e.g., jitter, shimmer),[17,20] though these studies did not investigate perceived vocal effort and vocal symptoms. To date, an analysis detailing the impacts of mask-wearing on acoustic *and* perceptual measures has not been undertaken in the health field.

The aim of the current study was to examine speech acoustics and self-reported vocal symptoms during masked and unmasked communication in working healthcare professionals. A secondary objective was to investigate whether additional factors (type of mask, sex) impacted these measures. We hypothesized that masked communication would attenuate spectral information[16,17] and subsequently elicit compensatory speaking behaviors, such as increased vocal volume and articulatory space, in order to overcome the acoustic changes. We further hypothesized that participants would report greater amounts of vocal effort and dyspnea along with concurrent increases in voice acoustics indicative of vocal effort during masked communication.

MATERIALS AND METHODS

Participants

Twenty-one healthcare professionals (8 cisgender male, 13 cisgender female) aged 23–49 years (M = 32.9 years; SD = 7.9 years) were prospectively enrolled in the study. Participants were speakers of standard American English, reported no history of neurological disease or head/neck cancer, and were free from speech, language, hearing, and voice problems at the time of the study. Participants were nonsmokers and nonvapers. Some participants (N = 8) reported health histories of gastroesophageal reflux and/or asthma, but were not excluded based on these diagnoses.

Occupations primarily included speech-language pathologists (N = 7), physicians (N = 5), physical therapists (N = 3), and respiratory therapists (N = 3). The types of masks that participants wore at work included simple disposable masks (referred to as "simple masks") and N95 respirators, often with the addition of simple masks over them. Table I provides demographic information about each participant. Participants were recruited between July and September of 2020 and were free from COVID-19 symptoms (e.g., cough, fever) at the time of the study. All methodology described in this study was approved by the Institutional Review Board at the University of Cincinnati. All participants provided informed, written consent prior to participation.

Protocol

Recordings were completed in a quiet room at the participant's place of employment prior to starting their work shift. Participants donned a headset microphone (MicroMic C555L) attached to a handheld recorder (Zoom H4n) set to acquire acoustic data at a sampling rate of 44.1 kHz with a 16-bit resolution. The microphone was placed at a 45° angle from the midline of the lips at a distance of 8.5 cm; the nonstandard microphone distance was to allow for the placement and removal of face masks. Microphone recordings were calibrated to sound pressure level (dB SPL) using a sound level meter (Extech; dB A). To calculate vocal intensity (dB SPL) from the speech recordings, pure tones (500 Hz) were played at varying intensity levels from a mobile phone application ("Frequency Generator" in Android or "Frequency" in iOS) that was held at the lips, while the intensity was measured with the sound level meter held at the microphone.

Participants were instructed to read aloud a series of vowels, words, and sentences with corner vowel targets (/i/, /u/, /ɑ/) in order to calculate a vowel-space metric as an indirect measure of articulatory movements in the oral cavity. Participants also read the first paragraph of the Rainbow Passage to acquire a sample of running speech for cepstral and spectral acoustical analyses.[21] Finally, participants produced 12 repetitions of vowel-voiceless consonant-vowel (VCV) productions (e.g., /ifi/) to calculate relative fundamental frequency (RFF) for characterizing fundamental frequency (f_o) during voicing transitions. See Table II for complete list of speech tasks. All tasks were completed with (masked) and without a face mask (unmasked).

Participants made perceptual ratings of their dyspnea and vocal effort following masked and unmasked readings. The modified Borg Scale[22] for dyspnea is a category-ratio scale from 0 to 10. The prompt asks the participant to rate: "How much difficulty is your breathing causing you right now?" and was slightly modified to be indicative of their breathing difficulty during the speech readings. A self-rating of "0" indicated that the participant's breathing was not causing any difficulty at all, whereas a score of 10 indicated "Maximal" difficulty. Next, participants rated the amount of vocal effort on a 100-mm visual-analog scale (VAS) in which effort was described as "an exertion of the voice or how hard you have to try to make a voice." A rating on the left side of the scale represented "no effort" and would be indicative of a lower score, whereas a rating on the right was anchored as "most effort" and measured as a higher rating.[23,24] Participants were instructed to place a single line on the scale to indicate their perceived vocal effort.

Data Extraction

Vowel analysis. Acoustic measures were manually extracted from vowel segments using Praat software.[25] Praat pitch settings were modified for the self-reported sex of each participant, wherein a female range was 90–500 Hz and a male participant range was set from 60–300 Hz. Three researchers (authors V.S.M., C.L.K., and T.H.P.) were trained to perform acoustic extraction on a sample of acoustic recordings captured with and without a mask. The trained researcher extracted the middle portion of the sustained vowel (~3 seconds in duration), and midsegment of the vowel (~100 msec in duration) during single word and sentence stimuli. The midsegment was identified to exclude vocalic offset or onset behavior. From these segments, the researcher used the Voice Report tool to extract measures of mean f_o, standard deviation (SD) of f_o, jitter, shimmer, and harmonics-to-noise ratio (HNR). Praat was also used to extract the vocal intensity level (dB SPL) that was then adjusted to the actual intensity levels from the SPL calibration procedure described earlier.

First (F1) and second (F2) formant values were extracted using a wide-band spectrogram over the same selected vowel segments. Formant values were then averaged for each vowel of interest (i.e., /i/, /u/, /ɑ/) and Vowel Articulation Index (VAI) was calculated (Eq. 1). The VAI is an estimate of vowel space that minimizes the effects of inter-speaker variability[26] and instead

TABLE I.
Participant Demographic Information.

Subject	Age (Years)	Sex	Mask Type	Occupation	Relevant Health History
01	39	F	N95 + Simple	Physician	GERD
02	36	F	N95 + Simple	Physician	
03	30	F	N95 + Simple	SLP	
04	33	F	N95 + Simple	SLP	Asthma; GERD
05	27	F	N95 + Simple	SLP	
06	31	F	N95 + Simple	SLP	
07	24	F	N95	SLP	
08	25	F	Simple	SLP	GERD
09	24	F	Simple	RT	
10	24	F	Simple	RT	Childhood asthma
11	35	F	Simple	Nurse	
12	43	F	Simple	Medical Admin.	
13	49	F	Simple	PT	
14	41	M	N95 + Simple	Physician	Childhood asthma
15	44	M	N95 + Simple	Physician	
16	45	M	Simple	Physician	Childhood asthma; GERD
17	35	M	Simple	PT	
18	27	M	Simple	PT	Exercise induced asthma
19	26	M	Simple	Medical Sales	
20	29	M	Simple	SLP	Asthma; GERD
21	23	M	Simple	RT	

Admin = administrator; PT = physical therapist; RT = respiratory therapist; SLP = speech-language pathologist.

TABLE II.
Speech Stimuli.

Sustained vowels × 5 sec in duration
/i/, /u/, /ɑ/
Single words repeated × 3 each
Heed, who'd, hod
Sentences (bolded word was analyzed)
I wish he would **heed** my advice.
I asked myself, "**Who'd** do that?"
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.
Rainbow Passage (paragraph 1)
VCV utterances, repeated × 4 each
/ɑfɑ/, /ifi/, /ufu/

normalizes the relationships between vowel space across speakers.[27] A smaller VAI value indicates vowel centralization.[28]

$$\mathrm{VAI} = \frac{F2/i/ + F1/\mathrm{ɑ}/}{F1/i/ + F1/u/ + F2/u/ + F2/\mathrm{ɑ}/} \quad (1)$$

Following experimental data extraction, interrater reliability was calculated for one randomly selected participant, with researchers blinded to other values. An intraclass correlation coefficient (two-way, consistency) found moderate-toexcellent reliability[29] with $M = 0.93$ (range = 0.68–0.99) across all acoustic measures. Intra-rater reliability was completed approximately 1 month after the original acoustic extraction on a randomly selected participant for each researcher, blinded to previous extracted values. Pearson correlations between the original and new extraction values were $M = 0.94$ ($r = 0.90$–0.97), $M = 0.94$ ($r = 0.83$–0.99) and $M = 0.91$ ($r = 0.77$–0.99) for each researcher, respectively.

Spectral and cepstral analysis. The second and third sentences of the Rainbow Passage were analyzed[21] using the Analysis in Dysphonia in Speech and Voice software (version 4.0). The following steps were completed: 1) removal of pauses >150 msec, with most pauses occurring at punctuation boundaries, 2) down-sampling the acoustic signal to 22.5 kHz, and 3) setting the software to the "Rainbow Passage" analysis setting. This set the spectral window size to 1024 with 75% overlap and a cepstral time averaging of 7. The cepstral peak extraction range was modified for each participant based on their self-reported sex, with 90–500 Hz for females and 60–300 Hz for males. Vocalic detection was applied to improve accuracy of extraction. The low-to-high spectral ratio (L/H ratio) cut-off frequency was set to 4 kHz, which is consistent with previous analyses of effort and dysphonia.[23] From here, the cepstral peak prominence (CPP) and its SD (CPP SD) and the L/H ratio and its SD (L/H ratio SD) were automatically calculated for each recording.

Relative fundamental frequency. RFF values were extracted from each VCV production using a semi-automated MATLAB algorithm,[30] which calculated the instantaneous f_o of the voicing cycles during the transition into/out of the voiceless consonant, normalized cycle f_o to the steady-state f_o of the closest vowel, and converted resulting values into semitones (ST; Eq. 2).

We analyzed the RFF values of the voicing cycles closest to the voiceless consonant (offset cycle 10, onset cycle 1). These cycles been shown to be the most sensitive to changes in the physiological mechanisms underlying RFF—including vocal fold abduction, increased laryngeal muscle tension to cease vocal fold vibration, and aerodynamic forces to reinitiate vocal fold oscillation.[31] Offset cycle 10 and onset cycle 1 are significantly lower in people with voice disorders characterized by increased laryngeal tension[31–33] and in those reporting elevated vocal effort.[23,34] As a result of the algorithmic processing, the average number of usable offset and onset RFF instances was 7.0 and 7.5, respectively, out of 12 repetitions. Usable instances were averaged within-participant[35] prior to subsequent analyses, resulting in one offset cycle 10 and one onset cycle 1 value per participant.

$$\mathrm{RFF}\,(\mathrm{ST}) = 12 \times \log_2\left(\frac{f_o}{f_o^{\mathrm{ref}}}\right) \quad (2)$$

Self-perceptual analysis. Self-perceptual ratings were extracted from paper datasheets and transferred to a Microsoft Excel document (Microsoft Office version 2016). The VAS ratings were measured and reported in millimeter units by one rater, then re-checked by a second rater who was blinded to the first rating. Because all re-checked effort ratings were within 1 mm of the original measurement, the original measurement was used for analysis. All category-ratio scale ratings of dyspnea were deemed 100% reliable and identical to the original transfer.

Statistical Analysis

The aims of the study were 1) to investigate acoustic and perceptual measures during masked and unmasked conditions, and 2) to examine the impact of additional factors (mask type, participant sex) on these measures. Therefore, we completed separate mixed-effects analyses of variance models for each acoustical and perceptual measure. The fixed effects were condition (masked, unmasked), mask type (N95, simple), and sex (male, female), as well as their two- and three-way interactions. Participant was input as a random factor. Significance was set to $\alpha < 0.05$. Mixed-effect models were assessed for normality and homoscedasticity, resulting in the appropriate model fit for the data.

Post hoc analyses were completed via Tukey's simultaneous tests, which automatically adjusted for family-wise error at the time of analysis (i.e., to reduce type 1 error). Subsequently, the α criterion of $P_{adj} < .05$ was established for all pairwise comparisons. Cohen's d, calculated as $(\mu_1 - \mu_2)$/pooled SD, was reported to quantify the effect sizes of significant findings. Reliability analyses were completed in *R* software (version 4.0.2), whereas all other analyses were completed in Minitab statistical software (version 19).

RESULTS

Acoustic Measures

There were several significant differences between masked and unmasked conditions (Table III). Our results showed significant increases in HNR ($P = .002$, $d = .39$), CPP ($P = .001$, $d = .59$), and L/H ratio ($P < .001$, $d = 1.18$) during masked speech compared to unmasked. Conversely, L/H SD ($P = .006$, $d = .61$), RFF offset 10 ($P = 0.34$, $d = .19$) and VAI ($P = .039$, $d = .69$) all significantly decreased during the masked condition.

Female participants exhibited significantly greater mean f_o ($P < .001$, $d = 4.04$) and CPP SD ($P = .001$, $d = 1.65$) compared to male participants. There was also a significant interaction effect of condition × sex for CPP ($P < .012$). Post hoc comparisons revealed that male participants had significantly higher CPP values ($P_{adj} = .003$, $d = 1.39$) in the masked condition ($M = 8.13$ dB) compared to the unmasked condition ($M = 6.76$ dB), with no other significant pairwise comparisons between males and females by condition.

There was no impact of the main effect of mask type on any acoustic measure. However, there was a significant interaction effect of condition × mask type for L/H ratio ($P < .001$) and L/H ratio SD ($P = .009$). L/H ratio was significantly greater in the masked condition for those who were wearing N95 masks, compared to the unmasked N95 condition ($P_{adj} < .001$, $d = 1.43$), unmasked simple condition ($P_{adj} = .003$, $d = 2.02$), and masked simple condition ($P_{adj} = .032$, $d = 1.38$). L/H ratio SD was significantly greater in unmasked N95 wearers compared to all other condition/type combinations. Finally, there were no main or interaction effects for f_o SD, jitter, shimmer, vocal intensity, or RFF onset 1.

Perceptual Measures

One participant did not complete self-perceptual ratings, resulting in an analysis of 20 datasets. There was a

TABLE III.
Results of Significant Statistical Findings for Acoustic and Perceptual Measures During Masked and Unmasked Conditions.

Measure	Unmasked Mean (SD)	Masked Mean (SD)	*P* Value	*d*	Effect Size Interpretation
HNR (dB)	18.55 (3.69)	20.04 (4.05)	.002	0.39	Small
CPP (dB)	7.37 (1.24)	8.04 (0.99)	.001	0.59	Medium
L/H ratio (dB)	39.86 (4.01)	44.39 (3.69)	<.001	1.18	Large
L/H ratio SD (dB)	8.47 (1.67)	7.69 (0.70)	.006	0.61	Medium
RFF offset 10 (ST)	−0.77 (0.90)	−0.96 (1.05)	.034	0.19	Small
VAI	0.91 (0.06)	0.87 (0.06)	.039	0.69	Medium
Dyspnea	0.18 (0.33)	1.45(1.08)	.002	1.59	Large
Vocal effort (mm)	11.6 (12.05)	41.2 (21.67)	<.001	1.69	Large

Effect size interpretations are based on criteria from Cohen.[36] Effect size calculation = $(\mathrm{Mean}_1 - \mathrm{Mean}_2)/\mathrm{SD}_{\mathrm{pooled}}$.

CPP = cepstral peak prominence; HNR = harmonics-to-noise ratio; L/H ratio = low-to-high spectral ratio; RFF = relative fundamental frequency; ST = semitones; VAI = vowel articulation index.

significant increase in self-reported vocal effort ($P < .001$, $d = 1.69$) and dyspnea ($P = .002$, $d = 1.59$) during masked communication. There were no impacts of mask type or participant sex as well as no interaction effects found.

DISCUSSION

We hypothesized that VAI would increase during masked speech as a compensatory strategy; however, the opposite effect was found: VAI significantly decreased during masked speech. We theorize that this effect occurred because masks reduce or restrict movement of the lips and jaw (and subsequently lingual excursions) during masked speech. The significant reduction in articulatory range could lead to a "mumbling effect" and therefore impede speech intelligibility and comprehension during communicative exchanges.

We did not find a statistically significant increase in vocal intensity across masked/unmasked conditions. Participants maintained their vocal intensity in both conditions, increasing slightly from 83.0 dB SPL (unmasked) to 83.39 dB SPL (masked). It is important to note that vocal intensity was measured from a microphone placed *outside* of the mask. Previous work using an artificial speech signal through simple and N95 masks showed that masks cause an intensity attenuation of approximately 3–12 dB across various frequencies, depending on the type of mask used.[15] Therefore, in order for the microphone signal to show consistent vocal intensity values, there must have been an increase in vocal intensity during masked speech. These findings provide evidence that speakers are, indeed, speaking louder in order to maintain their same vocal intensity target during masked speech. Speaking loudly for long periods of time has been associated with higher levels of vocal effort and fatigue,[37,38] as well as increased risk for developing phonotraumatic lesions[39] and voice disorders.[40] As such, it is possible that healthcare professionals may be at increased risk for developing vocal issues due to their occupational vocal load in combination with face masks.

We found a significant increase in both CPP and HNR in the masked condition compared to unmasked. These findings are in accordance with suspected elevation in vocal intensity levels during masked speech, as previous literature has shown that CPP and HNR measures depend upon speaker intensity.[41,42] Nguyen et al.[16] also reported significant increases in HNR during masked speech, but the authors hypothesized that the change could be due to the mask's attenuating effects on high frequency content—where aperiodic glottal noise often resides. The vocal parameters of jitter and shimmer (correlates to vocal perturbation and breathiness), did not change in comparisons between masked and unmasked conditions in our study, nor in other literature.[17,20] Therefore, we suspect that the significant increase in CPP and HNR are likely due to an increase in vocal intensity and/or high frequency attenuation, and not reflective of an improvement in vocal quality during masked speech.

As expected, we found a significant attenuation of high-frequency information (>4 kHz), exhibited by an increase in L/H ratio values during masked speech. Interaction analyses showed that participants wearing N95s were driving the increase, with significantly greater L/H ratios (indicating high frequency attenuation) compared to both simple masks and unmasked conditions. Investigation into the comparisons yielded large effect size differences (Cohen's $d > 1.0$), meaning that N95 masks may pose greater communication challenges than simple masks. The study by Nguyen et al.[16] also found a significant attenuation in frequency information from 1 to 8 kHz with greater attenuation noted with N95 masks (5.2 dB), and surgical masks (2.0 dB), compared to unmasked speech. Besides these findings, there were few main effects of mask type, which may have been due to our relatively small subsets of participants for the comparison (N95: N = 9; simple: N = 12). Subsequently, the impact of different face mask types should be investigated further in a larger group of speakers.

In line with our hypothesis that participants would exhibit acoustic effects consistent with vocal effort, we found a significant reduction in RFF offset cycle 10 during masked speech. RFF has shown promise as an acoustic indicator of laryngeal tension; lower values are reported in individuals with hyperfunctional voice disorders[32,33] and adults with healthy voices who were simulating vocal effort and strain.[23,34,43] Although our participants showed lower values during masked speech compared the unmasked, the effect size difference was small. Thus, it is difficult to predict the clinical risk of developing a voice disorder from these small changes. Currently, there is no established clinical guideline for a meaningful change in RFF values, nor is there a cut-off criterion or critical range to indicate a higher risk for developing acute or long-term voice problems. Acoustical measures such as RFF should be interpreted in the context of additional clinical factors such as self-reported effort, peri-laryngeal tension, and fatigue.

In our study, participants reported significantly greater amounts of vocal effort—with a statistically large effect size—in the masked condition. Vocal effort values increased from 11.6 mm to 41.2 mm, which is consistent with levels of "moderate effort" (i.e., 43.8 mm).[23] We suspect that mask wearers were trying to overcome the degradation of the speech signal via increases in vocal intensity and laryngeal tension, which they perceived as increased vocal effort. This suggests that mask-wearing healthcare professionals have an added risk factor for developing vocal problems beyond those presented by their occupational vocal demands. An investigation into how long-term daily mask use impacts vocal symptoms and voice acoustics is needed.

Finally, participants reported a significant increase in perceived dyspnea during masked speech. Although significant with a large effect size, this increase in dyspnea is less likely to be clinically meaningful. The modified Borg Scale for dyspnea provides descriptive anchors at each interval. Using these anchors to understand the degree of change perceived by the speaker, participants reported feeling "nothing at all" (score = 0.18) while unmasked, to "very slight dyspnea" (score = 1.45) while masked. With that in mind, this statistically significant increase is likely not directly impacting their speech or communication function.

Limitations and Future Directions

Our work is limited because we used a simple spectral cut-off frequency of 4 kHz. It may be beneficial to examine more specific frequency bands (e.g., 2–5 kHz, 5–8 kHz) to understand how frequency attenuation may impact speech formants. We showed evidence for a reduction in vowel articulation based on F1 and F2 formants, but additional information on consonant articulation and high frequency phonemic information (e.g., /f/ vs. /s/) is needed. Furthermore, we did not acquire information about the laryngeal source in our study. Electroglottography or neck-surface accelerometry could provide important information about vocal fold contact (timing, duration) during masked speech. A laryngeal microphone could also be used to measure amplitudes that are not affected by the attenuating characteristics of the mask.

Although this work provides information about the effects of wearing masks on vocal function, healthcare professionals experience numerous additional communication challenges, including face-shields, plexiglass barriers, and loud environments. Future work should aim to comprehensively evaluate speech and voice function in these "real-life" work environments, including the cumulative effects of occupational voice demands, speaking environment, and type of personal protective equipment. A longitudinal study would be beneficial to understand how these factors may increase the risk of developing voice problems over time. Finally, an investigation into compensatory techniques (e.g., clear speech) is needed to understand how training may help healthcare professionals improve comprehension and reduce vocal effort during masked communication.

CONCLUSION

Face masks are a barrier to communication that impact speech acoustics and result in speaker's perception of increased vocal effort. Healthcare professionals may be at increased risk for developing vocal issues due to their occupational vocal load in combination with face masks. Further work is needed to understand how long-term mask use may increase the risk for developing voice problems and to investigate whether vocal health education may help to offset these effects.

Acknowledgments

The project was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health, under Award Number 2UL1TR0 01425-05A1. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. The authors thank Teresa Hollenkamp, MA, CCC-SLP, for her assistance with participant recruitment.

BIBLIOGRAPHY

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. Heidel SE, Torgerson JK. Vocal problems among aerobic instructors and aerobic participants. *J Commun Disord* 1993;26:179–191. https://doi.org/10.1016/0021-9924(93)90007-W.
3. Mori MC, Francis DO, Song PC. Identifying occupations at risk for laryngeal disorders requiring specialty voice care. *Otolaryngol Head Neck Surg* 2017;157(4):670–675. https://doi.org/10.1177/0194599817726528.
4. Munier C, Kinsella R. The prevalence and impact of voice problems in primary school teachers. *Occup Med* 2008;58:74–76. https://doi.org/10.1093/occmed/kqm104.
5. Roy N, Merrill RM, Thibeault S, Parsa RA, Gray SD, Smith EM. Prevalence of voice disorders in teachers and the general population. *J Speech Lang Hear Res* 2004;47:281–293.
6. Roy N, Merrill RM, Gray SD, Smith EM. Voice disorders in the general population: prevalence, risk factors, and occupational impact. *Laryngoscope* 2005;115:1988–1995.
7. Sapir S, Keidar A, Mathers-Schmidt B. Vocal attrition in teachers: survey findings. *Int J Lang Commun Disord* 1993;28:177–185. https://doi.org/10.3109/13682829309041465.
8. Williams NR. Occupational groups at risk of voice disorders: a review of the literature. *Occup Med* 2003;53:456–460. https://doi.org/10.1093/occmed/kqg113.
9. Boltežar L, Šereg BM. Voice disorders in occupations with vocal load in Slovenia. *Slov J Public Health* 2014;53:304–310. https://doi.org/10.2478/sjph-2014-0033.
10. Bartoszko JJ, Farooqi MAM, Alhazzani W, Loeb M. Medical masks vs N95 respirators for preventing COVID-19 in healthcare workers: a systematic review and meta-analysis of randomized trials. *Influenza Other Respi Viruses* 2020;14:365–373.
11. Wong CKM, Yip BHK, Mercer S, et al. Effect of facemasks on empathy and relational continuity: a randomised controlled trial in primary care. *BMC Fam Pract* 2013;14:200. https://doi.org/10.1186/1471-2296-14-200.
12. Mendel LL, Gardino JA, Atcherson SR. Speech understanding using surgical masks: a problem in health care? *J Am Acad Audiol* 2008;19:686–695.
13. Champion J, Holt R. Dental care for children and young people who have a hearing impairment. *Br Dent J* 2000;189:155–159.
14. Radonovich LJ, Yanke R, Cheng J, Bender B. Diminished speech intelligibility associated with certain types of respirators worn by healthcare workers. *J Occup Environ Hyg* 2009;7:63–70.
15. Goldin A, Weinstein BE & Shiman N How do medical masks degrade speech perception? 2020;27:8–9.
16. Nguyen DD, McCabe P, Thomas D, et al. Acoustic voice characteristics with and without wearing a facemask. *Sci Rep* 2021;11:5651. https://doi.org/10.1038/s41598-021-85130-8.
17. Magee M, Lewis C, Noffs G, et al. Effects of face masks on acoustic analysis and speech perception: implications for peri-pandemic protocols. *J Acoust Soc Am* 2020;148:3562. https://doi.org/10.1121/10.0002873.
18. Monson BB, Hunter EJ, Lotto AJ, Story BH. The perceptual significance of high-frequency energy in the human voice. *Front Psychol* 2014;5:1–11.
19. Ribeiro VV, Dassie-Leite AP, Pereira EC, Santos ADN, Martins P, Irineu R d A. 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.
20. Fiorella ML, Cavallaro G, Di Nicola V, Quaranta N. Voice differences when wearing and not wearing a surgical mask. *J Voice* 2021; S0892199721000709. https://doi.org/10.1016/j.jvoice.2021.01.026.
21. 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.
22. Borg G. Psychophysical bases of perceived exertion. *Med Sci Sports Exerc* 1982;14:377–281.
23. McKenna VS, Stepp CE. The relationship between acoustical and perceptual measures of vocal effort. *J Acoust Soc Am* 2018;144:1643–1658.
24. Sandage MJ, Hamby HA, Barnett LA, Harris ML, Parker CR, Allison LH. Vocal function differences before and after sorority recruitment. *J Voice* 2020;S0892199718301760. https://doi.org/10.1016/j.jvoice.2020.04.026.
25. Boersma P. Praat, a system for doing phonetics by computer. *Glot Int* 2001; 5:341–345.
26. Sapir S, Ramig L, Spielman J, Fox C. Acoustic metrics of vowel articulation in Parkinsons's disease: vowel space area (VSA) vs. vowel articulation index (VAI). *Models of Analysis of Vocal Emissions for Biomedical Applications: 7th International Workshop*. Florence, Italy: Firenze University Press; 2011:173–175.
27. Caverlé MWJ, Vogel AP. Stability, reliability, and sensitivity of acoustic measures of vowel space: a comparison of vowel space area, formant centralization ratio, and vowel articulation index. *J Acoust Soc Am* 2020;148: 1436–1444. https://doi.org/10.1121/10.0001931.
28. Skodda S, Visser W, Schlegel U. Vowel articulation in Parkinson's disease. *J Voice* 2011;25(4):467–472. https://doi.org/10.1016/j.jvoice.2010.01.009.
29. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15:155–163.
30. 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.
31. Stepp CE, Merchant GR, Heaton JT, Hillman RE. 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. Heller Murray ES, Lien Y-AS, Van Stan JH, et al. Relative fundamental frequency distinguishes between phonotraumatic and non-phonotraumatic vocal hyperfunction. *J Speech Lang Hear Res* 2017;60:1507–1515.
33. Stepp CE, Hillman RE, Heaton JT. The impact of vocal hyperfunction on relative fundamental frequency during voicing offset and onset. *J Speech Lang Hear Res* 2010;53:1220–1226. https://doi.org/10.1044/1092-4388(2010/09-0234).
34. Lien Y-AS, Michener CM, Eadie TL, Stepp CE. Individual monitoring of vocal effort with relative fundamental frequency: relationships with aerodynamics and listener perception. *J Speech Lang Hear Res* 2015;58:566–575. https://doi.org/10.1044/2015_JSLHR-S-14-0194.
35. Eadie TL, Stepp CE. Acoustic correlate of vocal effort in spasmodic dysphonia. *Ann Otol Rhinol Laryngol* 2013;122:169–176.
36. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Second ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
37. Bottalico P. Speech adjustments for room acoustics and their effects on vocal effort. *J Voice* 2017;31(3):392.e1–e12. https://doi.org/10.1016/j.jvoice.2016.10.001.
38. Bottalico P, Graetzer S, Hunter EJ. Effects of speech style, room acoustics, and vocal fatigue on vocal effort. *J Acoust Soc Am* 2016;139(5):2870–2879. https://doi.org/10.1121/1.4950812.
39. Bastian RW, Thomas JP. Do talkativeness and vocal loudness correlate with laryngeal pathology? A study of the vocal overdoer/underdoer continuum. *J Voice* 2016;30(5):557–562. https://doi.org/10.1016/j.jvoice.2015.06.012.
40. Byeon H. The risk factors related to voice disorder in teachers: a systematic review and meta-analysis. *Int J Environ Res Public Health* 2019;16:3675. https://doi.org/10.3390/ijerph16193675.
41. Awan SN, Giovinco A, Owens J. Effects of vocal intensity and vowel type on cepstral analysis of voice. *J Voice* 2012;26(5):670.E15–E20. https://doi.org/10.1016/j.jvoice.2011.12.001.
42. Brockmann-Bauser M, Bohlender JE, Mehta DD. Acoustic pertubation measures improve with increasing vocal intensity in individuals with and without voice disorders. *J Voice* 2018;32(2):162–168. https://doi.org/10.1016/j.jvoice.2017.04.008.
43. McKenna VS, Heller Murray ES, Lien Y-AS, Stepp CE. The relationship between relative fundamental frequency and a kinematic estimate of laryngeal stiffness in healthy adults. *J Speech Lang Hear Res* 2016;59:1283–1294.