Introduction

• Healthcare professionals represent a subset of individuals who must communicate verbally to meet their work demands, otherwise known as occupational voice users.1 Occupational voice users develop vocal problems at higher rates than non-occupational voice users, often leading to negative socioeconomic and emotional consequences.1-8

• There is a need to understand how the speaker may be impacted during communication exchanges while wearing masks mandated during the COVID-19 pandemic.

• Studies involving artificial speech stimuli have demonstrated that masks attenuate speech frequencies between 2–7 kHz at a range of 3–12 dB SPL, depending on mask type. As high frequency speech information is critical for speech intelligibility,1,9 it is possible that speakers may be making compensatory articulatory adjustments during masked speech.

• To date, an analysis detailing the impacts of mask-wearing on acoustic and perceptual measures has not been undertaken in the health field.

Purpose

• Aim 1: To examine speech acoustics and self-reported vocal symptoms during (female 90–600 Hz; male 60–300 Hz)

• Extracted the vocal intensity level (dB SPL) that was then adjusted to actual intensity levels from the SPL calibration procedure

• Extracted first (F1) and second (F2) formant values using a wide-band spectrogram, which were then averaged for each vowel of interest (i.e., /i/, /u/, /a/) to calculate the Vowel Articulation Index (VAI)

• Used the Voice Report tool to extract measures of mean fo, standard deviation (SD) of fo, jitter, shimmer, and harmonics-to-noise ratio (HNR)

Spectral & Cepstral Analysis

• 2nd and 3rd sentences of the Rainbow Passage were analyzed using the Analysis in Dysphonia in Speech and Voice (ADSV) software

• Cepstral peak prominence range modified based on sex (female 90–500 Hz; male 500–600 Hz)

• Low-to-high spectral ratio (L/H ratio) cut-off frequency was set to 4 kHz

• Cepstral peak prominence (CPP) and it’s SD extracted

Relative Frequency

• Extracted the ten voicing cycles preceding (offset cycles) and following (onset cycles) the voiceless consonant /f/ in each production of VCV using MATLAB

• Specifically examined offset cycle 10 and onset cycle 1

• The instantaneous fo of each cycle was then normalized to each corresponding vowel steady-state fo and converted to semitones

Materials & Methods

Participants:

• 21 healthcare professionals (8 M, 13 F) aged 23-49 years

• Speakers of American English who were non-smokers and non-vapers with no history of speech, language, hearing, or voice problems

• Occupations primarily consisted of speech-language pathologists, physicians, physical therapists, and respiratory therapists

Protocol:

• Participants wore a headset microphone attached to a handheld recorder to acquire acoustic data

• Microphone recordings were calibrated to sound pressure level (dB SPL) using a standards meter

• Participants were then instructed to read aloud a series of vowels, words, and sentences with vowel target vowels (i, u, a) in both masked and unmasked conditions to see below for all task baselines

• Participants then made perceptual ratings of their dysphonia and vocal effort following masked and unmasked readings using the modified Borg Scale10 and 100-mm visual-analog scale (VAS), respectively

Vowel Analysis:

• Manually extracted using Praat, with pitch settings modified for sex of participant (female 90–600 Hz; male 60–300 Hz)

• Extracted the vocal intensity level (dB SPL) that was then adjusted to actual intensity levels from the SPL calibration procedure

• Extracted first (F1) and second (F2) formant values using a wide-band spectrogram, which were then averaged for each vowel of interest (i.e., /i/, /u/, /a/) to calculate the Vowel Articulation Index (VAI)

• Used the Voice Report tool to extract measures of mean fo, standard deviation (SD) of fo, jitter, shimmer, and harmonics-to-noise ratio (HNR)

Spectral & Cepstral Analysis

• 2nd and 3rd sentences of the Rainbow Passage were analyzed using the Analysis in Dysphonia in Speech and Voice (ADSV) software

• Cepstral peak prominence range modified based on sex (female 90–500 Hz; male 500–600 Hz)

• Low-to-high spectral ratio (L/H ratio) cut-off frequency was set to 4 kHz

• Cepstral peak prominence (CPP) and it’s SD extracted

Relative Frequency

• Extracted the ten voicing cycles preceding (offset cycles) and following (onset cycles) the voiceless consonant /f/ in each production of VCV using MATLAB

• Specifically examined offset cycle 10 and onset cycle 1

• The instantaneous fo of each cycle was then normalized to each corresponding vowel steady-state fo and converted to semitones

Discussion

From our analyses we found a reduction in:

• 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, potentially leading to a “mumbling effect.”

• L/H ratio—significant attenuation (>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.

• CPP and HNR—significant reduction during offset cycle 10 during masked speech. From our results, one may surmise that the mask-wearing healthcare providers examined in this study exhibited indirect evidence of increased laryngeal tension and vocal effort during masked speech, which may increase their risk for developing vocal fatigue.

• The on the other hand, we found increases in:

• Vocal effort—increased from 11.6 mm to 41.2 mm, indicative of a moderate amount of vocal effort when compared to other work.10

• Dysphonia—participants reported increases from 0.18 (unmasked) to 1.45 (masked), indicating a perceptual increase from feeling “nothing at all” to “voicing slight dysphonia.”

• The significant effect is likely impacting their speech or communication outcomes.

• CPP and HNR—significant increases during masked speech, which are in accordance with an increase in vocal intensity, as previous literature has shown that CPP and HNR measures depend upon speaker intensity.11,12

We found no significant changes in:

• Vocal intensity—maintained 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. Therefore, in order for the microphone signal to show consistent vocal intensity values, the participants must have increased their vocal intensity during masked speech.

Results

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 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 whether vocal health education may help to offset these effects.

Acknowledgements & References

The project was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health, under Award Number 2UL1TR001435-05A1. We would like to acknowledge Rachel A. Guadn, M.S., CCEP/LP and Rebecca Howell, M.D., for their contributions. This work is currently in revision at The Laryngoscope.