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

Courtney L. Kendall, B.S.,[1] Tulsi H. Patel, B.S.,[1] & Victoria S. McKenna, Ph.D.[1,2]

[1]Department of Communication Sciences and Disorders, University of Cincinnati
[2]Department of Biomedical Engineering, University of Cincinnati

Voice & Swallow Mechanics Lab

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.[1–5]
- 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.[6] As high frequency spectral information is critical for speech intelligibility,[7] 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 masked and unmasked communication in working healthcare professionals.
- Aim 2: To investigate whether additional factors (type of mask, sex) impacted these measures.

Materials & Methods

Participants:

- 21 healthcare professionals (8 M, 13 F) aged 23-49 years
- Speakers of Standard 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 sound level meter
- Participants were then instructed to read aloud a series of vowels, words, and sentences with corner target vowels (/i/, /u/, /ɑ/) in both masked and unmasked conditions- see **Table** below for all speech tasks
- Participants then made perceptual ratings of their dyspnea and vocal effort following masked and unmasked readings using the modified Borg Scale[8] and 100-mm visual-analog scale (VAS), respectively

Sustained vowels × 5 seconds in duration
- /i/, /u/, /ɑ/

Single words repeated × 3
- Heed, who'd, hod

Sentences (bolded word was targeted for analysis)
- 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
- /ɑfɑ/, /ifi/, /ufu/

Data Extraction:

Vowel Analysis
- Manually extracted using Praat, with pitch settings modified for sex of participant (female 90-500 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/, /ɑ/) 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 extraction range modified based on sex (female 90-500 Hz; male 30-60 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 Fundamental Frequency (RFF)
- Extracted the ten voicing cycles preceding (offset cycles) and following (onset cycles) the voiceless consonant /f/ in each production of VCV using MATLAB[9]
  - 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

Self-perceptual Analysis
- VAS ratings were measured and reported in millimeter units by one rater, then re-checked by a blinded second rater
- Original measurement was used for analysis as all re-checked effort ratings were within 1 mm of the original measurement
- All ratings of dyspnea were deemed 100% reliable and identical to the original transfer

Statistical Analysis:

- Completed separate mixed-effects analyses of variance (ANOVA) models for each acoustical and perceptual measure
  - Fixed effects: condition (masked, unmasked), mask type (N95, simple), and sex (male, female), and their two- and three-way interactions
  - Significance was set to $\alpha<.05$
- *Post hoc* analyses completed via Tukey's simultaneous tests for differences ($p_{adj}<.05$)

Results

Measure	Unmasked Mean (SD)	Masked Mean (SD)	*p*	*d*	Effect Size Interpretation
HNR (dB)	18.55 (3.69)	20.04 (4.05)	.002	.39	Small
CPP (dB)	7.37 (1.24)	8.04 (0.99)	.001	.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	.61	Medium
RFF offset 10 (ST)	-0.77 (0.90)	-0.96 (1.05)	.034	.19	Small
VAI	0.91 (0.06)	0.87 (0.06)	.039	.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

Note. HNR=harmonics-to-noise ratio; CPP=cepstral peak prominence; L/H Ratio=Low-to-high spectral ratio; RFF=relative fundamental frequency; VAI=vowel articulation index. Effect size calculation= $(Mean_1-Mean_2)/SD_{pooled}$

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.
- RFF- significant reduction in 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.

On the other hand, we found increases in:
- Vocal effort- values increased from 11.6 mm to 41.2 mm, indicative of a moderate amount of vocal effort when compared to other work.[10]
- Dyspnea- participants reported increases from 0.18 (unmasked) to 1.45 (masked), indicating a perceptual increase from feeling "nothing at all" to "very slight dyspnea." This significant increase is likely not 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.

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 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 2UL1TR001425-05A1. We would like to acknowledge Renee Gustin, M.S., CCC-SLP and Rebecca Howell, M.D., for their contributions. This work is currently in revision at The Laryngoscope.

[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(3):254-259. **[2]** Mori MC, Francis DO, Song PC. Identifying Occupations at Risk for Laryngeal Disorders Requiring Specialty Voice Care. *Otolaryngol Neck Surg*. 2017;157(4):670-675. **[3]** 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(2):281-293. **[4]** Roy N, Merrill RM, Gray SD, Smith EM. Voice Disorders in the General Population: Prevalence, Risk Factors, and Occupational Impact: *The Laryngoscope*. 2005;115(11):1988-1995. **[5]** Williams NR. Occupational groups at risk of voice disorders: a review of the literature. *Occup Med*. 2003;53(7):456-460. **[6]** Goldin A, Weinstein BE, Shiman N. How do medical masks degrade speech perception? 2020;27(5):8-9. **[7]** 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. **[8]** Borg G. Psychophysical bases of perceived exertion. *Med Sci Sports Exerc*. 1982;14(5):377-281. **[9]** 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. **[10]** McKenna VS, Stepp CE. The relationship between acoustical and perceptual measures of vocal effort. J Acoust Soc Am. 2018;144:1643–1658. **[11]** Eadie TL, Stepp CE. Acoustic Correlate of Vocal Effort in Spasmodic Dysphonia. *Ann Otol Rhinol Laryngol*. 2013;122(3):169-176. **[12]** Awan SN, Giovinco A, Owens J. Effects of Vocal Intensity and Vowel Type on Cepstral Analysis of Voice. *J Voice*. 2012;26(5):670.E15-670.E20.