How do masks affect the way we speak?

Alexandra Lilia Nastevski, Bingqing Yu, Suyuan Liu, Michelle Kamigaki-Baron,
Gillian De Boer, Bryan Gick

COVID-19 has made mask-wearing a fact of face-to-face communication, and so the present study endeavoured to test whether and to what extent people change their audiovisual speech-related behaviour in response to wearing a mask. Mask-wearing obscures the visible speech articulators and has been shown to dampen power distributions and degrade transmissions in high-frequency bands (Bandaru et al. 2020; Corey et al. 2020; Magee et al. 2020; Pörschmann et al. 2020). Masked speakers report heightened self-perception of unintelligibility and increased vocal effort (Ribeiro et al. 2020; Bandaru et al. 2020). Masked speech also correlates with increased spectral tilt (Magee 2020). This behaviour is consistent with hyperarticulation and secondary cue exaggeration as speakers compensate to increase their intelligibility, as in noisy environments (Haggard et al. 1970; Raphael 1972; Ohde 1984; Lindblom 1990; Denes 1995). Interacting with masked speakers also hinders the perception efficiency of the listeners (Yi et al. 2021), as well as their ability to judge their interlocutor's emotions (Carbon 2020). We found evidence of this compensation in the acoustic signal, but not in visual-based behaviour.

In this study, participants were asked to direct experimenters through a map task (Anderson et al. 1991) with another participant over Zoom. Both their audible and visible communicative behaviours were recorded. Data were analysed using a within-participant design comparing Masked and Unmasked conditions. Acoustic data were segmented and annotated using Montreal Forced Aligner (McAuliffe et al. 2017) and analysed using VoiceSauce (Shue et al. 2011). We used OpenFace 2.0 image processing software (Baltrusaitis et al. 2018) to analyze head and eyebrow movements in visual data. We present the results of a repeated measures ANOVA for behavioural (Human) and mechanical (Mech) trials in Table 1, with significant variances in shaded cells.

Table 1. Summary of effects of mask-wearing on parameters for vowels /ɑ/ and /i/

Vowel	H1-H2*	H1-A2*	H1-A3*	F0	F1	F2	F3	CPP
Hum /ɑ/	+0.2 dB	-0.1 dB	+0.4 dB	+8 Hz	+0.6 Hz	+9.3 Hz	-20.7 Hz	+0.3 dB
Mech /ɑ/	+0.2 dB	**-2.3 dB**	**-2.1 dB**	**-18.4 Hz**	-12.3 Hz	-64.6 Hz	-2.2 Hz	**-1 dB**
Hum /i/	+0.3 dB	**+2.8 dB**	**+1.8 dB**	**+7.6 Hz**	+10.6 Hz	**-78.2 Hz**	**-64.9 Hz**	+0.2 dB
Mech /i/	+0.6 dB	+0.7 dB	-0.03 dB	**-22.1 Hz**	+11.6 Hz	-36.8 Hz	+14 Hz	**-0.6 dB**

Our main acoustic findings were that (i) mask wearing is a significant predictor for perturbation of spectral slope (H1-A2*, H1-A3*), fundamental frequency (F0), and Cepstral Peak Prominence (CPP), and (ii) speakers will compensate for the pure mechanical perturbation effect of masked speech for lower frequency /ɑ/, but for higher frequency /i/, speakers will exhibit divergent behaviour whether an isolated mechanical effect exists or not. Our visual findings, interestingly, show that speakers do not compensate for the lack of lower face visual cues by adopting more upper face (eyebrow), head or hand movements. There was little difference between the frequency of eyebrow, head or hand movements across conditions. This also revealed that the use of hand movements is strongly speaker-dependent.

While the scope of this study only focused on /ɑ/ and /i/, future work will incorporate the effect of masked speech on consonants and larger prosodic units, as well as the effects of adaptation, and how speaker subjectivity might lead to anticipatory overcompensation effects. Though there is no significance found in the visual section, future work could examine whether the articulators gestures are exaggerated under the mask. This study could expand our current understanding on phonetic adaptation and processes related to intelligibility. This could hold implications for communication in industries like medicine, sanitation, or construction, where face coverings are commonplace.

References

Anderson, A., Bader, M., Bard, E., Boyle, E., Doherty, G. M., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C., Thompson, H. S. and Weinert, R. (1991). The HCRC Map Task Corpus. *Language and Speech 34*, 351-366.

Baltrusaitis, T., Zadeh, A., Lim, Y. C., & Morency, L. (2018). OpenFace 2.0: Facial behavior analysis toolkit. Paper presented at the 59-66.

Bandaru, S., Augustine, A., Lepcha, A., Sebastian, S., Gowri, M., Philip, A., & Mammen, M. (2020). The effects of N95 mask and face shield on speech perception among healthcare workers in the coronavirus disease 2019 pandemic scenario. *The Journal of Laryngology & Otology 134*(10), 895-898.

Carbon, C. C. (2020). Wearing face masks strongly confuses counterparts in reading emotions. *Frontiers in Psychology 11*, 2526.

Corey, R. M., Jones, U., & Singer, A. C. (2020). Acoustic effects of medical, cloth, and transparent face masks on speech signals. *The Journal of the Acoustical Society of America 148*(4), 2371-2375.

Denes, P. (1955). Effect of duration on the perception of voicing. *The Journal of the Acoustical Society of America 27*(4), 761-764.

Haggard, M., Ambler, S., & Callow, M. (1970). Pitch as a voicing cue. *The Journal of the Acoustical Society of America 47*(2B), 613-617.

Lindblom, Björn & Hardcastle, W. & Marchal, Alain. (1990). Explaining Phonetic Variation: A Sketch of the H&H Theory.

Magee, M., Lewis, C., Noffs, G., Reece, H., Chan, J. C., Zaga, C. J., ... & Vogel, A. P. (2020). Effects of face masks on acoustic analysis and speech perception: Implications for peri-pandemic protocols. *The Journal of the Acoustical Society of America 148*(6), 3562-3568.

McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M., Sonderegger, M. (2017). *Montreal Forced Aligner* [Computer program]. Version 0.9.0, retrieved from http://montrealcorpustools.github.io/Montreal-Forced-Aligner/

Ohde, R. N. (1984). Fundamental frequency as an acoustic correlate of stop consonant voicing. The Journal of the Acoustical Society of America, 75(1), 224-230.

Pörschmann, C., Lübeck, T., & Arend, J. M. (2020). Impact of face masks on voice radiation. *The Journal of the Acoustical Society of America 148*(6), 3663-3670.

Raphael, L. J. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristic of word‐final consonants in American English. *The Journal of the Acoustical Society of America 51*(4B), 1296-1303.

Ribeiro, V. V., Dassie-Leite, A. P., Pereira, E. C., Santos, A. D. N., Martins, P., & de Alencar Irineu, R. (2020). Effect of wearing a face mask on vocal self-perception during a pandemic. *Journal of Voice.* (Preprint). DOI:10.1016/j.jvoice.2020.09.006

Shue, Y.-L., Keating, P., Vicenik, C., Yu, K. (2011). VoiceSauce: A program for voice analysis. *Proceedings of the ICPhS XVII*, pp. 1846-1849. Retrieved from http://www.phonetics.ucla.edu/voicesauce/

Yi, H., Pingsterhaus, A., & Song, W. (2021). The adverse effect of wearing a face mask during the COVID-19 pandemic and benefits of wearing transparent face masks and using clear speech on speech intelligibility. (Preprint.) DOI:10.31234/osf.io/z9c4k