POMA Proceedings of Meetings on Acoustics

Volume 42 http://acousticalsociety.org/

179th Meeting of the Acoustical Society of America

*Acoustics Virtually Everywhere*

7-11 December 2020

2004

Interdisciplinary: Paper 4plDc3

# The effect of face masks on the intelligibility of unpredictable sentences

**Melissa Randazzo**
*Department of Communication Sciences & Disorders, Adelphi University, Garden City, NY, 11530; mrandazzo@adelphi.edu*

**Laura L. Koenig and Ryan Priefer**
*Adelphi University, 1 South Ave., Garden City, NY, 11530-4213 ; lkoenig@adelphi.edu, rpriefer@adelphi.edu*

To prevent the transmission of Covid-19 the US Centers for Disease Control recommends using a face mask in public. Converging evidence examining the acoustics of masked speech shows that different styles of masks have differential effects on sound attenuation. The purpose of this study was to compare speech intelligibility in different mask conditions (no mask, disposable surgical mask, cloth mask, and N95 respirator). Four native English speakers recorded unpredictable sentences, which avoid the confound of contextual predictability. Sentences were mixed with multitalker babble to simulate the noise experienced by listeners during activities of daily living. Forty-one listeners heard 24 pseudorandomized unpredictable sentences from the four mask conditions and typed what they heard. We measured intelligibility as the percentage of whole words correctly perceived. Acoustic analysis revealed that all masks filter the signal, with the greatest overall effect for the N95 respirator. A mixed model two-way ANOVA showed a significant effect of mask condition, driven by the low percent words correct for N95 respirators. Our error analysis revealed that listeners more often provided no response for the N95 respirator, but supplied phonetic approximations for surgical and cloth masks.

Published by the Acoustical Society of America

 https://doi.org/10.1121/2.0001374

# 1. INTRODUCTION

## A. How do face masks affect speech intelligibility?

To prevent the transmission of Covid-19 the US Centers for Disease Control [CDC] recommends using a mask while interacting in public[1]. Face masks protect public health by containing respiratory droplets[1]. However, they hinder speech perception by decreasing auditory and visual feedback, filtering the acoustic signal, and interfering with the physical actions of speech production[2]. Converging evidence from studies examining the acoustics of masked speech shows that different styles of masks have differential effects on sound attenuation. Simple cloth face masks attenuate sound by 3-4dB, while N95 masks attenuate sound by 12dB[3]. Transparent masks that maintain the visual aspect of the signal attenuate sound by 8dB, creating an acoustic-visual trade-off[4]. Although sound attenuation from cloth face masks is lower compared to other types of masks on average, the thickness of the weave matters, with heavier weave face masks attenuating sound more than N95 masks[4].

Thus, while face masks covering both the mouth and nose offer protection from disease, mask-wearing alters communication. Extended wearing of facemasks may increase cognitive load in communication[5]. Examination of the impact of mask wearing during everyday activities suggests that masks increase the self-perception of vocal effort, reduce speech intelligibility, and yield difficulty coordinating speech with breathing[2]. Adults with hearing loss report speaking less often, elaborating less, feeling more anxious, and socializing less when communicating with masks[6,7,8,9]. During testing in a second language, participants wearing masks report speaking more slowly and loudly, and hesitating more.

Most of the previous research regarding speech intelligibility in face masks focused on the healthcare setting. Given the various types of masks worn during activities of daily living, further investigation is required to determine their effects on communication in other contexts. Also, to date, speech intelligibility in masks has been studied using predictable sentences or semantically related items that could support comprehensibility via contextual bootstrapping. This could explain the discrepancy between self-reports of communication difficulties[2,5,9] and ceiling effects reported in studies examining speech intelligibility for adults with normal hearing and moderate hearing loss[11,12].

## B. Questions

The current study further examines speech intelligibility through masks using unpredictable sentences that separate intelligibility, i.e., how well listeners can correctly identify what they hear, from comprehensibility, or the ease of comprehending the overall message[13]. We pose the following questions: 1) Which face mask (surgical, cloth, N95 respirator) most impedes the intelligibility of unpredictable sentences compared to an unmasked condition? 2) Which aspects of the speech signal are most attenuated by different types of masks?

# 2. METHODS

## A. Speaker characteristics

Four native adult speakers of American English (two female–F1, F2, two male–M1, M2) produced a list of N=50 unpredictable sentences[13] in four conditions *(C1=No Mask, C2=Surgical Mask, C3=Cloth Mask, C4=N95 Mask)*; see Figure 1.

***Figure 1: Mask types pictured left to right: No mask [C1], Surgical mask [C2], Cloth mask [C3], N95 respirator [C4].***

### B. Recording procedure

Due to COVID-19 restrictions, equipment was shipped to speakers to record in their homes. Speakers were given detailed instructions on equipment use and donning the N95 respirator. Data were recorded by the four speakers using a ZoomH2next recorder. Speakers were asked to find a quiet room and limit external noise as much as possible. For some of our speakers the recordings did have some background traffic noise. Sentences were recorded in .wav format at a 48kHz sampling rate with 16-bit precision.

### C. Stimuli and processing

Thirty sentences were selected from the set of semantically unpredictable sentences[13]. Each sentence is 7 words long, makes sense, and is grammatically correct (e.g., *Animals often wander across woody grassy paths; Old baking books seem cheaper every summer*). Each utterance was labeled in Praat[14]. Using the Praat labels, we extracted each sentence, and an equal duration of multitalker babble noise (Auditec, Inc.). Sentence tokens were amplitude normalized within Praat to 65 dB and noise at 70 dB. These values were chosen based on pilot testing to avoid ceiling or floor effects. Sentences and noise were mixed and 100 ms silence was added at the beginning of each file.

### D. Participants

A total of 45 participants were recruited through the Prolific (www.prolific.com) platform. Participants who did not complete the task, scored below a threshold of 80% on catch trials (no mask, no noise), or scored less than 50% across all conditions were excluded from the final analysis (n=4). The final analysis included 41 participants (n=17 females, age *M*=30.87, *SD*=12.95). Participants were native English speakers with no self-reported history of hearing, learning, or communication impairments. Participants were assigned to listen to sentences spoken by one of the four speakers under the four mask conditions (F1, N=10; F2, N=9; M1, N=11; M2, N=11). This was done to isolate effects of mask condition from speaker differences.

### E. Listening task

Listeners heard 24 pseudorandomized sentences from the four mask conditions (6 different sentences in each condition) and typed what they heard. Six catch trials (no mask, no noise) were included every 5th trial to verify listener attentiveness. Sentences were counterbalanced between conditions and speaker. The listening task was programmed in PsychoPy 2020.2.5[15] and presented via the Pavlovia (www.pavlovia.org) platform. Listeners completed the *Hearing Handicap Inventory*[16] to provide information about their hearing in daily life. To characterize the results of the study, we queried participants regarding their mask use. Listeners received $4.00 for their participation. The study was approved by the Adelphi University IRB.

### F. Scoring

We measured speech intelligibility as the percentage of whole words correctly perceived. Misordering of words was not penalized (e.g. delicate pale → pale delicate). Spelling errors, typos, and morphological errors (e.g. animals → animal) were not penalized. We did not include trials with no responses or responses that were unrelated to the prompt (e.g. “something something”). This accounted for 4% of total trials. We also conducted an initial exploration of whether the nature of errors varied across mask conditions.

## 3. RESULTS

### A. Communication in daily life

Most of the participants reported speaking to someone who is wearing a mask often (60.00%), and the remaining participants sometimes (27.50%) and rarely (12.50%). The majority of the participants wear a face mask in public often (92.50%), and the remaining participants sometimes (5.00%) and never (2.50%).

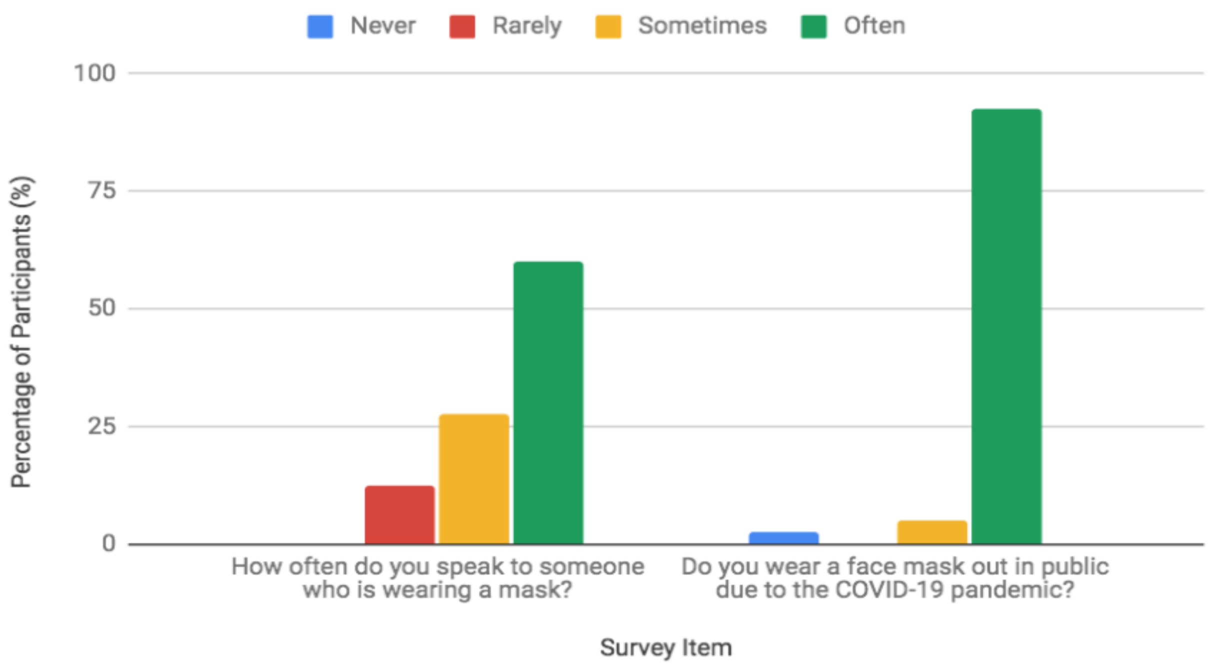


***Figure 2: Participant mask usage***

Result of the *Hearing Handicap Inventory* revealed that none of our participants experienced difficulties with hearing in their daily lives with all participant mean scores below the 16% threshold for no handicap ($M$=1.12, $SD$=3.42)[16].

### B. Acoustic results

Figure 3 presents the long-term averaged spectra for one speaker, showing results up to 8 kHz. These were obtained over the entire recordings (all sentences) for the four conditions. As expected, all masks impose some degree of filtering on the signal. The overall effect is greatest for the N95 mask.

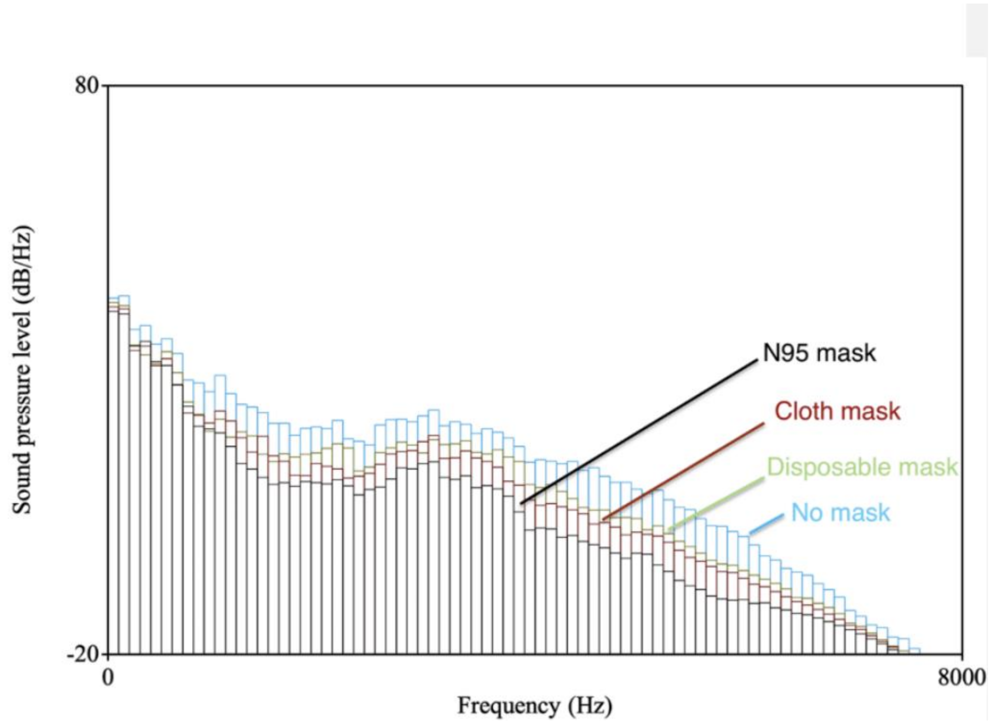


***Figure 3: Long-term average spectra for speaker F2***

Figure 4 provides another representation of the filtering effects of the different masks up to 4 kHz. The lines show the amplitude reduction across frequencies relative to the no-mask condition. This shows that the N95 mask can reduce amplitudes more than 10 dB at around 800 Hz and between about 2–3 kHz. These are in the frequency ranges of the first and second vocal tract formants (F1, F2) for adults[17], suggesting that the N95 mask restricts orofacial movements that convey basic information about vowels and consonant place of articulation.

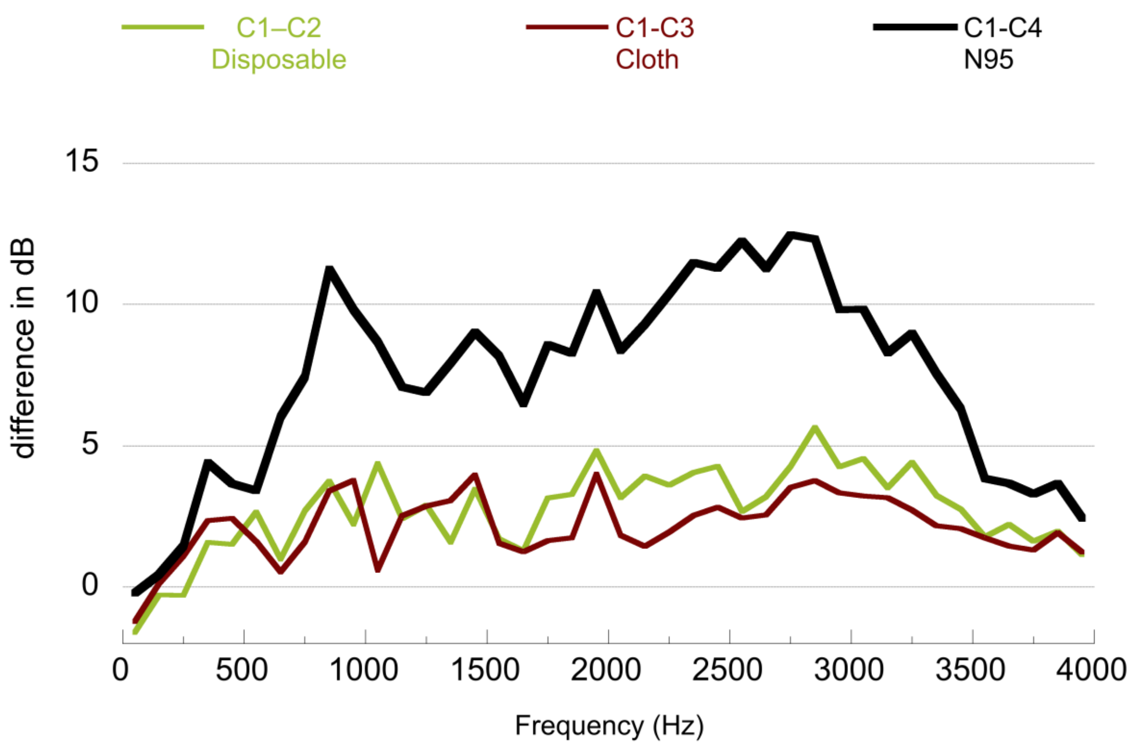

***Figure 4: Long-term spectra for speaker M2, showing differences relative to the no-mask condition [C1].***

## C. Perceptual results

Figure 5 shows percent words correct by condition. A mixed model two-way ANOVA with a within-subjects factor of mask condition and a between-subjects factor of speaker condition showed the effect of mask condition to be significant ($F_{3,\,152} = 14.325$, $p < .001$, $\eta^2 = .274$); in post-hoc (Bonferroni) tests, the no-mask condition differed from the disposable mask (p=.032), cloth mask ($p$=.003) and the N95 mask ($p < .001$). The disposable mask condition differed from the N95 mask (p=.006), and the cloth mask also differed from the N95 mask (p=.032). The disposable mask did not differ from the cloth mask ($p$=1.00).

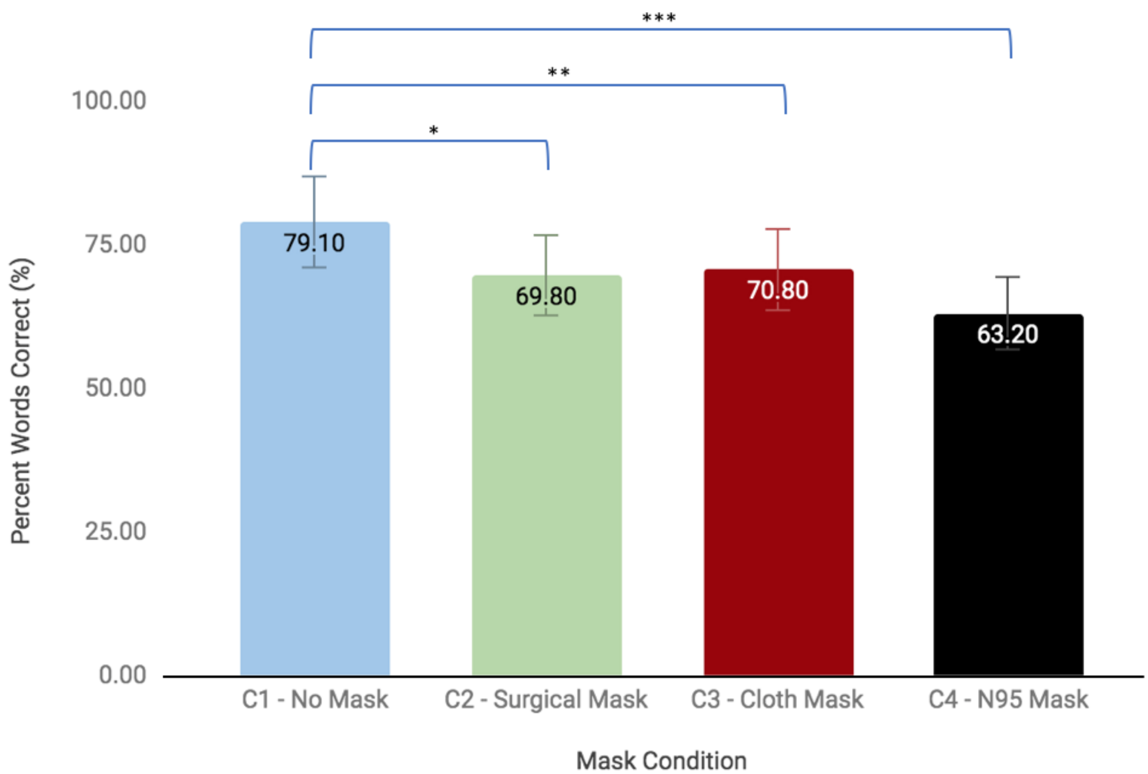


***Figure 5: Percent words correct by condition.*** *Note:* * = $p \leq .05$; ** = $p \leq .01$; *** = $p \leq .001$. Error bars show standard error of the mean.

The mixed model two-way ANOVA did not show a significant main effect of speaker ($F_{3,\,152}$=2.249, $p$ = .098, $\eta^2$=.1.51). Values for the four speakers, averaged across mask conditions, were F1 ($M$= 66.4, $SD$=17.0), F2 ($M$= 64.3, $SD$=17.6), M1 ($M$= 77.1, $SD$=14.0), and M2 ($M$= 73.4, $SD$=14.6). Examination of percent words correct by speaker and condition revealed that the N95 condition had the lowest scores across all speakers. There was no significant interaction between condition and speaker. Figure 6 shows the percent words correct for each mask condition by speaker.

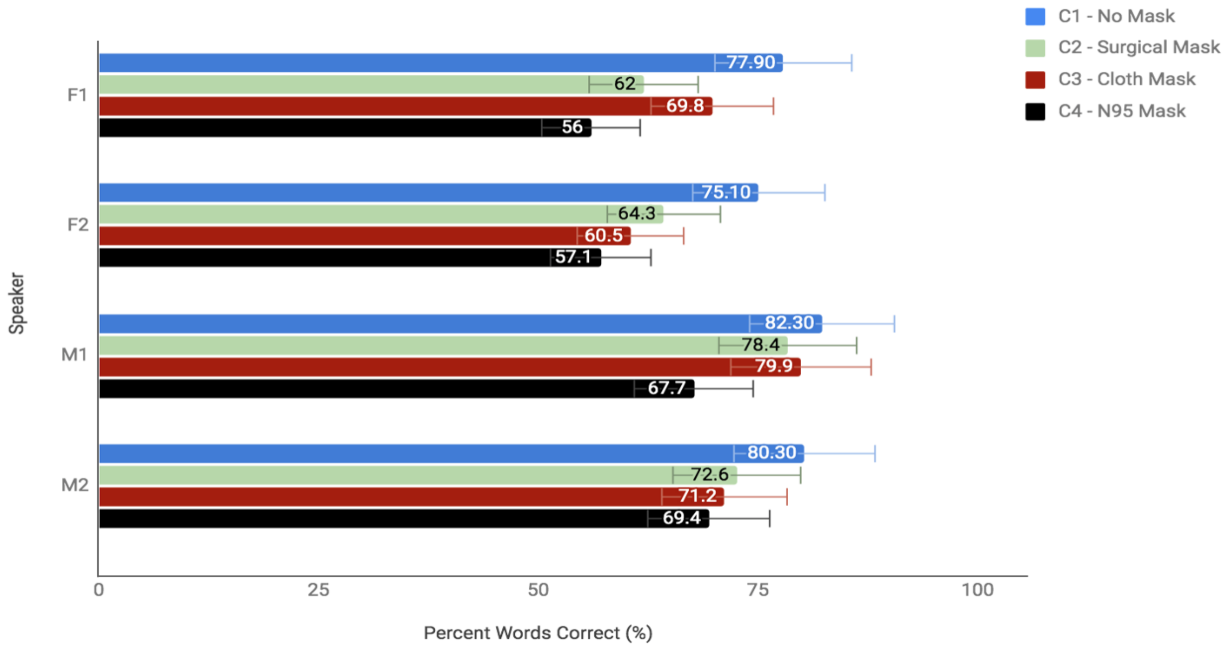


***Figure 6: Percent words correct by speaker and condition.*** Error bars show standard error of the mean.

A preliminary review of error types across the three mask conditions suggests that listeners more frequently supplied phonetic approximations ('giant' → 'diet', 'lonely' → 'only', 'lamb' → 'clams') for the disposal and cloth mask compared to the N95 mask. Percentages of non-responses, i.e. words where listeners did not supply an answer, were 36.90%, 37.70%, and 44.10% for the disposable, cloth, and N95 masks, respectively. A one-way ANOVA examining the percentage of non-responses, i.e. no word provided, showed the effect of mask condition to be significant ($F_{2,585}$=4.725, $p$=.009); in post-hoc (Tukey HSD) tests the N95 mask differed from both the disposable and cloth masks ($p < .04$ for both), whereas the disposal and cloth masks did not differ from each other ($p$=.941). Not surprisingly, the number of blank responses correlated with percent words correct ($r$=-0.816, $p < 0.001$). The difference across mask conditions could mean, however, that the cloth and disposable masks, more than the N95 mask, provided listeners with enough acoustic information that they were willing to guess at individual words.

## 4. DISCUSSION

The purpose of this study was to compare speech intelligibility in different mask conditions. We used unpredictable sentences to avoid the confounds of contextual predictability on our results. Additionally, we used multitalker babble to simulate the noise experienced by listeners during activities of daily living. Overall, intelligibility for all three mask conditions was significantly lower compared to the no mask condition. Previous studies indicate that N95 masks attenuate sound to a greater extent than surgical or cloth masks[3]. In the current study, percent words correctly identified was lowest for N95 masks across all mask conditions and in all four of our speakers. Disposable masks and cloth masks, commonly worn during activities of daily living, did not differ from each other and showed similar effects on intelligibility.

Although survey studies have reported that individuals, particularly those with hearing loss, experience difficulties with communication through masks[2,9] and increased cognitive load[5], some studies of human speech perception to date have failed to find significant differences between masked and unmasked conditions[12]. The comparison across mask conditions using unpredictable sentences in multitalker babble in the current study simulates the multiple variables experienced by individuals when communicating in public, complementing previous work carried out in laboratory conditions or in health-care settings. Taken together, these results suggest that multiple variables are relevant in determining the most appropriate mask for a given situation. There is an inverse relationship between the weave of masks (related to their breathability and comfort) and their ability to block respiratory droplets.[18] These factors also interact with sound attenuation. Therefore, individuals should choose the mask types based on type of mask depending on a range of factors including health risks, environmental noise, communication effort, and context.

Acoustic results show marked suppression of frequencies in the range of the first and second formants (F1, F2) for N95 masks. This suggests that the mask affects the movement of the jaw and lips. Reduced articulatory excursions may be one explanation for why intelligibility is the most reduced in this condition compared to the no mask condition. Our error analysis revealed that listeners more often provided no response for the N95 mask condition, whereas they supplied phonetic approximations for the surgical and cloth mask conditions. Compared to the no-mask condition, all masks affect intelligibility as measured by the percentage of words correct in our task. However, our follow-up analysis suggests that different mask types may differ in how much usable phonetic information is provided. In the current study, a non-response and a phonetic approximation were both scored as errors with the same weight. This may account for the lack of significant differences between mask conditions. Future studies should consider weighting different types of errors to further characterize the impact of different masks on perception.

## 5. LIMITATIONS AND FUTURE DIRECTIONS

This pilot study was conducted online due to the restrictions on face-to-face data collection during the Covid-19 pandemic. Online data collection did not allow us to control for differences in participants' hardware for listening to the sentences. These results should be confirmed under more controlled laboratory conditions. Future studies will examine speech perception through masks in adults with aging-related hearing loss and other populations with communication impairments. More detailed acoustic analyses will be conducted to determine how different masks and other sources of environmental noise (e.g. street traffic, schools) impact production and perception of specific speech sounds. Future studies will also examine the influence of experience and attitudes toward face masks in speech perception.

## 6. CONCLUSIONS

All three types of masks investigated in this study interfered with speech intelligibility. The N95 mask provides the most protection from respiratory droplets but may provide the least usable information to comprehend speech in context. Critical information conveyed in healthcare settings may be supported by giving patients written or visual information. Surgical masks and cloth masks most commonly worn during activities of daily living also interfere with speech perception. Professionals in settings such as schools should consider the trade-off between clear communication and health risks when selecting which masks to use in different contexts.

## ACKNOWLEDGMENTS

The authors would like to thank Erika Henningsen, Kyle Selig, Elizabeth Nicholson, and Steven Nicholson for help with stimulus creation. The authors would also like to thank research assistants compensated by Adelphi University: Lauren Rothberg, Rebekah Guastella, Moriah Rastegar, Danielle Bustamante, and Thomas Moran.

# REFERENCES

[1]COVID-19: Considerations for Wearing Masks. (2020, December 18). Retrieved January 04, 2021, from https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/cloth-face-cover-guidance.html

[2]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*. https://doi.org/10.1016/j.jvoice.2020.09.006

[3]Goldin, A., Weinstein, B., & Shiman, N. (2020). How do medical masks degrade speech perception? *Hearing Review, 27*, 8–9.

[4]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*, 2371–2375. https://doi.org/10.1121/10.0002279

[5]Round, M., & Isherwood, P. (2020). Speech intelligibility in respiratory protective equipment—Implications for verbal communication in critical care. *Trends in Anaesthesia and Critical Care*. https://doi.org/10.1016/j.tacc.2020.08.006

[6]Chodosh, J., Weinstein, B. E., & Blustein, J. (2020). Face masks can be devastating for people with hearing loss [Editorial]. *British Medical Journal, 370*. https://doi.org/10.1136/bmj.m2683

[7]Trecca, E. M., Gelardi, M., & Cassano, M. (2020). COVID-19 and hearing difficulties. *American Journal of Otolaryngology*, *41*, 102496. 10.1016/j.amjoto.2020.102496

[8]Naylor, G., Burke, L., & Holman, J. (2020). Covid-19 lockdown affects hearing disability and handicap in diverse ways: A rapid online survey study. *Ear and Hearing 41*, 1442–1449. doi: 10.1097/AUD.0000000000000948

[9]Saunders, G. H., Jackson, I. R., & Visram, A. S. (2020). Impacts of face coverings on communication: An indirect impact of COVID-19. *International Journal of Audiology*, 1–12. https://doi.org/10.1080/14992027.2020.1851401

[10]Loukina, A., Evanini, K., Mulholland, M., Blood, I., & Zechner, K. (2020). Do face masks introduce bias in speech technologies? The case of automated scoring of speaking proficiency. *arXiv preprint arXiv:2008.07520*. 10.21437/Interspeech.2020-1264

[11]Atcherson, S. R., Mendel, L. L., Baltimore, W. J., Patro, C., Lee, S., Pousson, M., & Spann, M. J. (2017). The effect of conventional and transparent surgical masks on speech understanding in individuals with and without hearing loss. *Journal of the American Academy of Audiology*, 28, 58–67. doi: 10.3766/jaaa. 15151

[12]Mendel, L. L., Gardino, J. A., & Atcherson, S. R. (2008). Speech understanding using surgical masks: A problem in health care? *Journal of the American Academy of Audiology, 19*, 686–695. DOI: 10.3766/jaaa.19.9.4

[13]McHenry, M. A., & Parle, A. M. (2006). Construction of a set of unpredictable sentences for intelligibility testing. *Journal of Medical Speech Language Pathology, 14*, 269–271.

[14]Boersma, P. & Weenink, D. (2020). Praat: doing phonetics by computer [Computer program]. Version 6.1.38. http://www.praat.org/

[15]Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods.* Doi:10.3758/s13428-018-01193-y

[16]Newman, C. W., Weinstein, B. E., Jacobson, G. P., & Hug, G. A. (1990). The Hearing Handicap Inventory for Adults: Psychometric adequacy and audiometric correlates. *Ear and Hearing*, *11*, 430–433.

[17]Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America, 24*, 175–184. https://doi.org/10.1121/1.1906875

[18]Aydin, O., Emon, B., Cheng, S., Hong, L., Chamorro, L. P., & Saif, M. T. A. (2020). Performance of fabrics for home-made masks against the spread of COVID-19 through droplets: A quantitative mechanistic study. *Extreme Mechanics Letters*, *40*, 100924. https://doi.org/10.1016/j.eml.2020.100924