Effect of facemasks on speech intelligibility

R. Vazquez-Amos, J Thomas, S. Dance, H. Gohil, J. Telford, J. Green, P. Mapp

School of the Built Environment and Architecture
London South Bank University
London, UK SE1 0AA

ABSTRACT

With the advent of COVID, the wearing of face covering has been obligatory in both medical and everyday life. This paper describes three experiments undertaken to establish the effect of face coverings on speech sound power, speech directivity and speech intelligibility. The experiments used two different approaches: acoustic measurements and word scores. The face coverings assessed were a 'standard blue' surgical mask, a typical fabric mask, a prototype clear mask and a plastic transparent visor. The study showed that non-native English speakers had by far the most difficulty in comprehending English language speech when face coverings were worn during phonetically-balanced word list tests. All the masks were found to noticeably affect speech intelligibility, with the surgical mask having the least detrimental effect. The results are also compared to objective measurements of their physical acoustic characteristics to establish their performance.

1. INTRODUCTION

Legislation across the world mandated wearing of face coverings for most of the adult population to reduce the spread of COVID19. Although face masks and face shields are tested and marked in terms of their performance for particle size transmission, only limited research on their effect on speech transmission and speech intelligibility has been undertaken. Typically, this involves measuring the insertion loss, in order to gauge transmission loss and estimate source directivity on a Head and Torso simulator. Ruhale and Ruhale (3 masks) [1], Pörschmann *et al* (6 masks) [2] both found that masks produce attenuation above 1kHz. Corey *et al* (12 masks) [3] agreed that mask attenuation was significant and varied between different types of mask (1 male talker), as well as finding that speech directivity varied significantly. Magee *et al* tested mask attenuation (3 masks) [4], finding high frequency attenuation had no significant effect on speech intelligibility. This was based on the ASSIDS method, using 50 words and 22 sentences, based on five human talkers and five listeners using a virtual acoustic system using headphones under perfect recording and listening conditions. Bottalico *et al* [5] (3 masks) used electroacoustics: a HATS loudspeaker produced speech shaped sound which was binaurally convolved with the impulse response from two separate rooms (RT_{mid} 0.4 seconds and RT_{mid} 3.1 seconds). Forty listeners using headphones under perfect conditions with +3dB SNR listened to eight lists of 50 CNC words. Findings were high frequency mask attenuation, > 2 kHz, overall attenuation was least for surgical mask 2.3 dB 12% reduction in speech intelligibility (SI), and highest for N95 mask 4.2 dB attenuation resulting in 13% reduction in SI.

Mapp [6] (20 masks) used a HATS loudspeaker to test masks, RT_{mid} 0. 3 seconds, in terms of sound attenuation typically 3-5 dB at the key communication frequency of 4 kHz whereas heavier/stiffer medical grade masks produced 9-12 dB, the visor produced a 12 dB gain around 1 kHz before rolling off at higher frequencies. He also showed a significant attenuation effect of wearing medical mask with a visor, 14 dB at 4 kHz. SI tests were undertaken in a room RT_{mid} 0.5 seconds at a socially distant 2m. STIPA (60 dBA signal) was measured to be 0.75 with an ambient noise level of 23 dBA and 0.54 at 54 dBA (NR50) without a mask. The average STIPA of the 19 conditions tested under quiet conditions was 0.71 (worst case 0.65), under the noisy condition 0.45 (worst case 0.40). Mapp speculated that a STIPA of 0.66 was needed for quality communication of complex information such as medical advice. He suggested that hearing impairment would reduce communication in a hospital setting to unintelligible due to the lack of visual cues. He also showed the effect of mask wearing on speech perceived by hearing impaired listeners and the associated loss of consonant sounds [7]. Bannwart *et al* [8] suggested 20% increase in speech intelligibility was likely due to lip reading. In summary, facemasks act as a low pass filter and as such would significantly affect speech intelligibility where the primary intelligibility components are in the higher frequencies at 1 kHz and above.

2. HYPOTHESIS

Tens of billions of face masks have been manufactured over the past two years with little research into their real effect on communication. The key difference between the described and previous work is that the airflow from breathing, necessary in speech production, is not replicated by a Head and Torso Simulator. In addition, communication is more difficult for people not communicating in their first language. This paper addresses these shortcomings by testing common face masks/visors under realistic conditions, simulated in laboratory settings, see figure 1, to establish the real effect of face masks on speech intelligibility, speech directivity, and speech level and compare them to HATS based results. These experiments only became possible in October 2021 with the reduction in COVID19 restrictions in the UK.

Figure 1- Native speaker and HATS wearing Surgical, Visor and Prototype Mask

3. MEASURED INSERTION LOSS (Real and Simulated)

A head and torso simulator was used to measure the insertion loss of a blue surgical mask and a prototype transparent mask. The measurements were undertaken in a listening room, RT_{mid}=0.3s, using a calibrated pink noise signal to simulate speech transmission. Insertion losses are given in Figure 2.
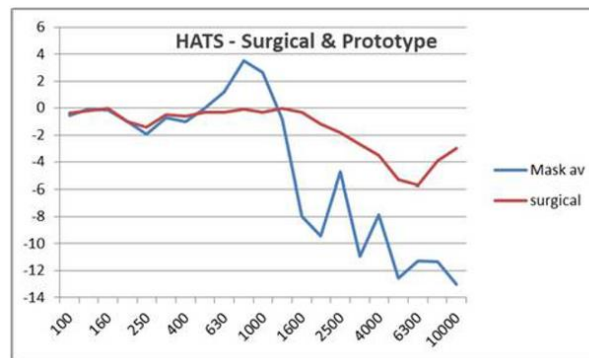

Figure 2 - HATS comparison of insertion loss of blue surgical mask and a prototype mask

Figure 2 shows the 1/3 octave band insertion losses of two masks using a Head and Torso simulator. These results were similar to already published data on surgical facemasks [1-3, 5-6]. However, the prototype mask shows a peak at around 800 Hz and greater attenuation at higher frequencies, as found by Ruhale and Ruhale for a face shield [1]. The additional peaks are unusual and not seen before and are likely to be an artefact of resonance in the material and geometry used in the prototype mask.

A corresponding experiment involving four native English speakers, speaking Harvard sentences for 20 seconds in an anechoic chamber, was undertaken using the same facemasks in order to establish the insertion losses for real talkers. Figure 3 shows the averaged 1/3 octave band insertion loss as measured directly in front of each talker (1.5m) and for comparison the HATS based values for the surgical mask.
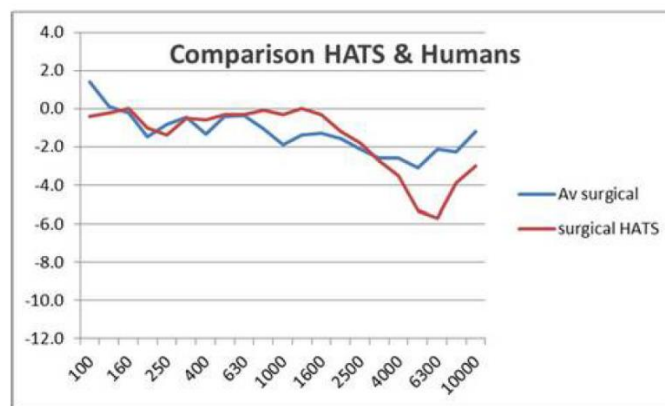

Figure 3 - Insertion loss for the Surgical mask based on 4 native English speakers and HATS data.

The experiment was repeated for the prototype facemask. The averaged insertion losses are compared to the HATS simulation results as shown in figure 4.
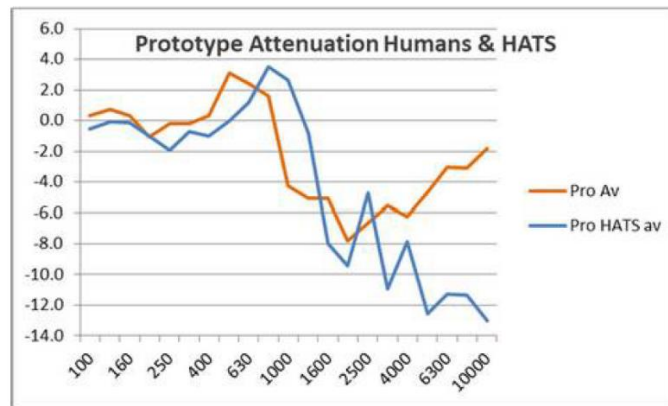

Figure 4 - Prototype mask averaged insertion loss for 4 native English speakers and HATS data

Figure 4 clearly shows that the 4 participants produced a similar trend in the attenuation with a 3-3.5 dB boost shifted down to 500 Hz compared to 800 Hz for HATS simulator test, important for vowels and speech power. At higher frequencies there is a distinct difference in performance with the HATS showing a roll off above 2 kHz of 2 dB per octave whereas the opposite is true for the participants with a 2 dB per octave increase. This is important, as intelligibility is critical at these frequencies and should show up in real intelligibility scores.

4. SPEECH INTELLIGIBILTY

To establish the effect of face coverings on speech intelligibility, a listening test experiment was undertaken. All 11 participants were young adults with good hearing acuity <15 dBHL across 500-8000 Hz using screening audiometry (Amplivox 850 Mark 4). The listening panel consisted of seven native English and three non-Native English speakers, see figure 5. One female and one male native English speaker spoke ten sentences from a list of 720 Harvard sentences for each test wearing a surgical mask, a transparent visor, and a prototype transparent mask as well as a control (no mask) condition. The participants were seated 3.5-4m from the speaker and were given 30 seconds to write down each short sentence under two noise conditions: noise (N) and no noise (NN). An additional variable, lip reading/no visual (NV) was introduced for the prototype mask readings.
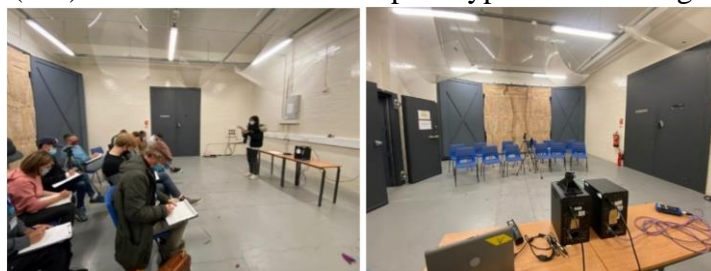

Figure 5 - Speech intelligibility panel in semi reverberant conditions

The experiment was undertaken in a room (202m^3 T20$_{1\,kHz}$ 1.18 sec) under a combination of two noise conditions No Noise LAeq, $_{2\,minutes}$ 36.5 dBA and pink noise 55.4 dBA, see figure 5. The noise condition was measured using a Class 1 NTi XL2 sound meter at the central participant position, 3.5m distant from the talker, and is representative of a typical hospital ward condition. L$_{Aeq,\,16\,hr}$ 54.7 dB [10]. The pink noise was generated by an NTi Minirator Pro connected to a Yahama HS50 loudspeaker located next to the human talker. The human talkers spoke consistently at a nominal 55 dBA as measured at the central location, 3.5m distant, with a small deviation of 1.3 dBA recorded on the sound meter whilst the nine experiments took place. Reverberation times were measured using the impulsive method using a Norsonic Nor140 Class 1 sound level meter in accordance with ISO 3382-2 [11].
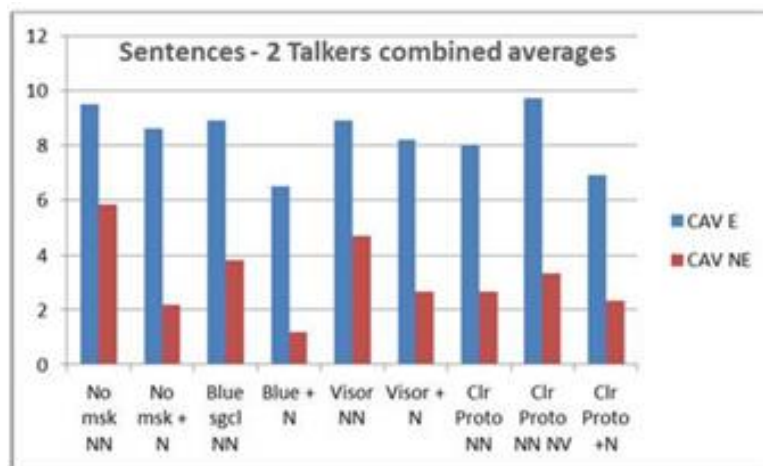

Figure 6 - Shows intelligibility scores for native (CAV E) and non-native (CAV NE) English speakers under Noise (N) and No Noise (NN) condition for 3 facemasks: Surgical, Visor, Prototype.

Analysis of figure 6 shows non-native English (NE) speakers having greater difficulty in understanding sentences in semi-reverberant conditions noise ort no noise, mask or no-mask, with a score 62.3% below that of native speakers. Diving down and looking at the non-noise instance, non-natives had a 55.6% lower intelligibility scores compared to native speakers. However, there was a 72.6% reduction when representative hospital noise levels were added.

Studying the mask specific results. For natives, under no noise, conditions the intelligibility scores for the visor and prototype masks increased by 2.1%, agreeing with the insertion loss results, see figure 2. This was probably due to the mid-frequency boost provided by the plastic; however, when noise was added the intelligibility score fell by 2.5% relative to the control condition. For non-native speakers the best mask was the visor, with a reduction in intelligibility of 20.3% compared to the control condition. Under noisy conditions the visor was the most intelligible mask with a marginal reduction of 2.4% for native speakers, whereas intelligibility increased by 23.8% for non-natives compared to the control condition. The prototype mask intelligibility was marginally worse for natives, 3.8%, and 9.5% better for non-natives. For the surgical mask the no noise condition showed a reduction in intelligibility score of 7.4% for natives and 35.6% for non-natives. When noise was introduced, these reductions increased to 22.9% and 47.6%.

A final noise condition, prototype mask experiment, with the aim to eliminate the effect of lipreading under no-noise conditions was conducted. This involved not looking at the talker and hence focusing on the words when writing them down. This increased the intelligibility scores for both natives and non-natives by 21.3% and 26.9% respectively, thus demonstrating task focus performance. Overall, the visor had the least effect on intelligibility for native speakers and the prototype mask for non-natives.

To understand the effect of the talker, an analysis was undertaken of male/female talker intelligibility scores, see figure 7.
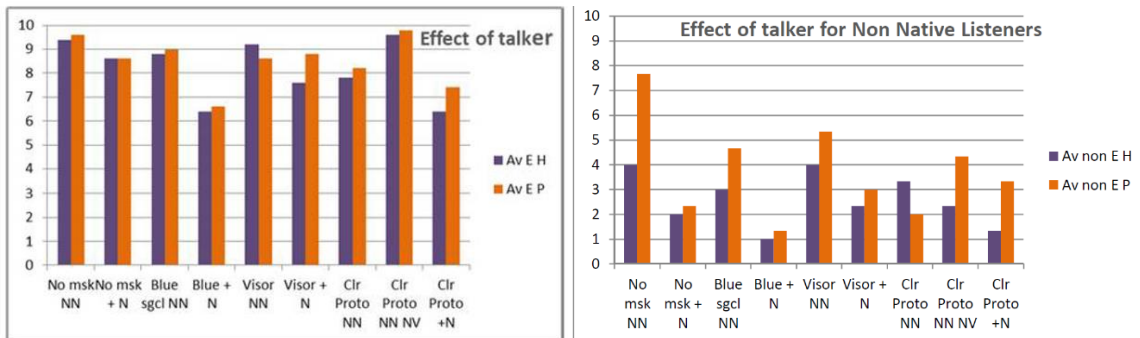

Figure 7 - Effect of the talker on intelligibility score for native/non-native English speakers

Figure 7 shows that the effect of the talker (both native English speakers) had little difference but overall the male talker (P) was slightly more intelligible. Focusing on the non-native listeners the effect of the talker was more significant, see figure 7, with the male talker preferred, even though the inferred speech level (SNR) was generally slightly lower, see figure 8.
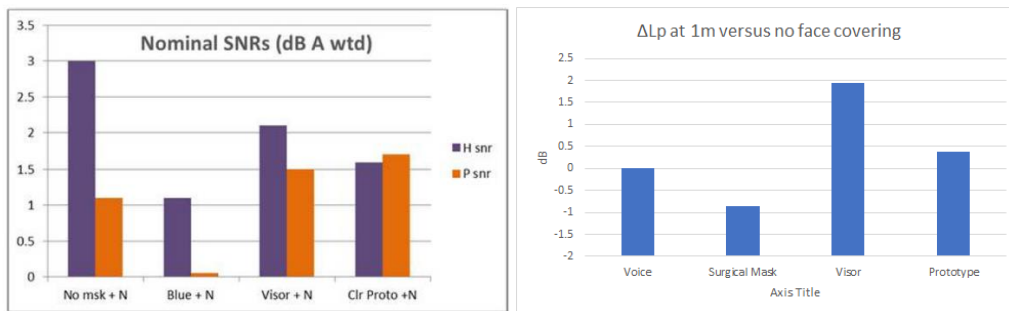

Figure 8 - Semi-Reverberant Signal-Noise (2 talkers, 3 masks) and b) free-field normalised level, dBA.

5. SPEECH DIRECTIVITY WITH MASKS

The effect of face coverings on speech directivity and speech level was established through a series of experiments undertaken in an anechoic chamber. The same Harvard word sentence lists were used by 2 males and 2 females sequentially wearing two face masks (blue surgical and transparent prototype), a visor and the control condition. Measurements were made simultaneously at 15^0 angles using 13, NTi XL2, Class 1 sound level meters. Microphones on XLR extension cables were placed 1.5m from the talker at seated head height, 1.2m, see Figure 9. The sentences spoken were 20 seconds long in accordance with ISO 24504 [13].

Figure 9 - Speech directivity measured at 15o intervals in an anechoic chamber.Averaged speech directivity (dBA) is given in figure 10 based on 4 native English talkers
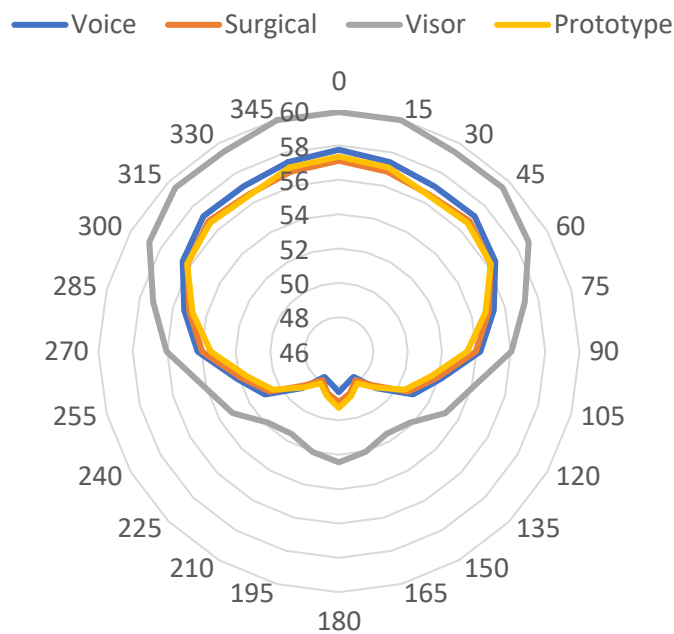

..

Figure 10 - Averaged speech directivity (4 participants) for face coverings

From figure 10 the natural voice, surgical and prototype masks produced very similar speech levels and directivity in terms of overall sound level (dBA) and were in broad agreement with ANSI 3.5 [14] at 60.5 dBA at 1m for normal voice effort. This compares to the measured 59.5 dBA with 0.3/0.6 dBA attenuation for the prototype and surgical masks. However, the visor amplified the speech signal by 2.2 dB, see figure 8b, probably due to the thickness of the plastic acting as a resonator, agreeing with the HATS results, see figure 2. Secondly, there was significantly less directivity, i.e. speech increased by 4.1 dB directly behind the talker. This would make the talker believe they are talking more loudly. This effect was mentioned

by all talkers when wearing the visor. However, this effect is offset by the additional resonance acting as an amplifier, which could explain the visor faring well, see figure 6.

6. CONCLUSIONS

Multiple experiments have been undertaken to determine the effect of facemasks on speech directivity, power and intelligibility. We have found that a HATS simulator does not replicate real speech. Multiple talkers were used to establish the insertion loss of 4 types of masks which showed that plastic material creates a distinctive hump in the frequency response which benefited intelligibility, based on the intelligibility score. Native English speakers understood 95% of sentences with no noise, which reduced to 90.5% with hospital noise levels. For non-native speakers this was 59% of sentences reduced to 36% when noise was introduced. When masks were added, understanding was reduced to 83% (no-noise) and 53.3% (noisy) for natives and 21% (no noise) and 15% (noisy) for non-natives. The plastic visor was the best mask for native speakers under noisy conditions and the prototype plastic mask the best for non-natives. The plastic visor also created the least directional speech radiation pattern. It should be noted these were young adults with good levels of hearing acuity. Hence, there is a definite need to further research this area to enable clear communication as IEC 60268-16 does advise an increased STI minimum for listeners for whom English is an additional language, along with increased STI minimums for older members of society who are likely to have decreased hearing ability.

7. REFERENCES

1. Ruhale R., Ruhale L. Using noise control principles when evaluating the acoustic impacts of face coverings during the coronavirus pandemic, *Inter-Noise*, Washington, 2021
2. Pörschmann C., Lübeck T., Arend J. Impact of face masks on voice radiation, *JASA* 148, 3663 (2020)
3. Corey R., Jones U., Singer A. Acoustic effects of medical, cloth, and transparent face masks on speech signals, *JASA* 148, 2371 (2020)
4. Magee M., Lewis C., Noffs G. Effects of face masks on acoustic analysis and speech perception: Implications for peri-pandemic, JASA 148, 3562 (2020)
5. Bottalico P., Murgia S., Puglisi G. Effect of masks on speech intelligibility in auralized classrooms, JASA 148, 2878 (202)
6. Mapp P. Acoustic and communication effects of face masks, *Institute of Acoustics Bulletin*, Jan/Feb 2021
7. Mapp P. Masking Speech, face coverings' effects on speech perception and intelligibility. Sound and Communications, October 2020
8. Bannwart A., Dell'Aringa A., Satico Adachi E. Lip reading role in the hearing aid fitting process, *Rev. Bras. Otorrinolaringoly*. 73(1), 101–105 (2007).
9. Morales L, Dance S., Leembruggen G. Preliminary validation of the revised STI male scale for the English language, J. *Audio Eng. Society*, 62(7), 493-506, 2014
10. Shield B., Shiers N., Glanville R. The acoustic environment of inpatient hospital wards in the United Kingdom, *JASA* 140, 2213 (2016)
11. ISO 3382-2: 2008- Acoustics — Measurement of room acoustic parameters — Part 2: Reverberation time in ordinary rooms, Geneva, Switzerland.
12. IEC 60268-16:2020 - Sound system equipment - Part 16: Objective rating of speech intelligibility by speech transmission index, Geneva, Switzerland.
13. ISO 24504:2014 – Ergonomics — Accessible design — Sound pressure levels of spoken announcements for products and public address systems, Geneva, Switzerland
14. ANSI 3.5-1997.- Methods for calculating the speech intelligibility index