


# Challenging Times for Cochlear Implant Users – Effect of Face Masks on Audiovisual Speech Understanding during the COVID-19 Pandemic

Trends in Hearing  
Volume 26: 1–9  
© The Author(s) 2022  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/23312165221134378  
journals.sagepub.com/home/tia  


Rasmus Sönnichsen<sup>1</sup> , Gerard Llorach Tó<sup>2</sup>, Volker Hohmann<sup>2,3,4</sup>, Sabine Hochmuth<sup>1,\*</sup> and Andreas Radeloff<sup>1,3,4,\*</sup>

## Abstract

Unhindered auditory and visual signals are essential for a sufficient speech understanding of cochlear implant (CI) users. Face masks are an important hygiene measurement against the COVID-19 virus but disrupt these signals. This study determinates the extent and the mechanisms of speech intelligibility alteration in CI users caused by different face masks. The audiovisual German matrix sentence test was used to determine speech reception thresholds (SRT) in noise in different conditions (audiovisual, audio-only, speechreading and masked audiovisual using two different face masks). Thirty-seven CI users and ten normal-hearing listeners (NH) were included. CI users showed a reduction in speech reception threshold of 5.0 dB due to surgical mask and 6.5 dB due to FFP2 mask compared to the audiovisual condition without mask. The greater proportion of reduction in SRT by mask could be accounted for by the loss of the visual signal (up to 4.5 dB). The effect of each mask was significantly larger in CI users who exclusively hear with their CI (surgical: 7.8 dB,  $p = 0.005$  and FFP2: 8.7 dB,  $p = 0.01$ ) compared to NH (surgical: 3.8 dB and FFP2: 5.1 dB). This study confirms that CI users who exclusively rely on their CI for hearing are particularly susceptible. Therefore, visual signals should be made accessible for communication whenever possible, especially when communicating with CI users.

## Keywords

speech intelligibility, hearing, speechreading, speech reception threshold test, personal protective equipment

Received 13 May 2022; Revised 26 September 2022; accepted 30 September 2022

## Introduction

Understanding speech is a highly complex process involving integration of auditory and visual information (Grant et al., 1998). Hearing-impaired listeners are especially dependent on both modalities for sufficient communication since their use of acoustic signals is already compromised by the hearing loss.

Cochlear implants (CI) provide an efficient way to improve hearing in patients with severe to profound sensorineural hearing loss (Lenarz et al., 2012). However, hearing is often compromised in noisy environments, which more realistically reflect daily life situations (Zaltz et al., 2020). The reason for this phenomenon is the limited proficiency of CIs to provide a signal which matches the original acoustic signal in terms of spectral resolution and fine spectro-temporal cues (Fu & Nogaki, 2005).

Due to frequent wearing of face masks during the COVID-19 pandemic it became even more apparent how

alterations of acoustic and visual signals impact daily communication (Brown et al., 2021; Sönnichsen et al., 2022). The influence of different face masks on acoustic speech

<sup>1</sup>Department of Otolaryngology, Head and Neck Surgery, University of Oldenburg, Oldenburg, Germany

<sup>2</sup>Auditory Signal Processing and Hearing Devices, University of Oldenburg, Oldenburg, Germany

<sup>3</sup>Research Center Neurosensory Science, University of Oldenburg, Oldenburg, Germany

<sup>4</sup>Cluster of Excellence “Hearing 4 All”, University of Oldenburg, Oldenburg, Germany

\*equal contribution

### Corresponding author:

Dr. med. Rasmus Sönnichsen, Universitätsklinik für Hals-Nasen-Ohren-Heilkunde, Kopf- und Halschirurgie, Universitätsmedizin Oldenburg, Steinweg 13-17, 26122 Oldenburg  
Email: rasmus.soennichsen@uni-oldenburg.de



signals has been studied thoroughly (Corey et al., 2020; Goldin et al., 2020; Muzzi et al., 2021; Palmiero et al., 2016). However, access to visual signals and integrative processes may be even more relevant for speech intelligibility than unhindered acoustic signals in hearing-impaired and even normal-hearing listeners (NH) (Sönnichsen et al., 2022).

In this study, we examine hindered communication mechanisms by evaluating two different types of face masks in CI users and NH.

## Methods

### Participants

All participants provided their informed written consent and participated on a voluntary basis. The experimental protocol was approved by the institutional review board (Medizinische Ethikkommission) of the University of Oldenburg (AZ 2020-135). All participants were native German speakers with normal or corrected-to-normal vision, without previously known cognitive impairment, at least 18 years of age and naïve with respect to the experimental audiovisual test material.

Forty post-lingually deafened CI users participated in the study. Inclusion criteria were at least 6 months CI experience or a percent correct score of at least 60% in the Freiburg monosyllables test in quiet with the CI at a level of 65 dB

SPL. Three CI users were excluded, one due to a broken processor and two because of exceeding the presentation limit during training in the audiovisual condition (see paragraph on Conditions below for further details). The remaining 37 CI users (22 female, 15 male) had a mean age of  $59.2 \pm 16.2$  years. The average age at implantation was  $55.4 \pm 18.2$  years and the average duration of CI use  $43.4 \pm 41.3$  months (see detailed demographic data in Table 1). In addition, ten NH (five female and five male, mean age of  $23.3 \pm 3.7$  years) with a pure tone average at 0.5, 1, 2, and 4 kHz (PTA4) instead of PTA4 of less than 20 dB HL in both ears participated.

Main analysis focused on comparisons within and between CI users and normal-hearing participants. For a more in-depth analysis CI users were subcategorized post hoc according to the hearing status of the contralateral ear:

1. ‘CI-only’: Bilateral CI users or CI users with a contralateral PTA4 of more than 80 dB HL and without hearing aid usage. (n = 14).
2. ‘Asymmetric hearing loss (AHL)’: CI users with a contralateral PTA4 of more than 25 dB HL (n = 16). Fourteen of them were using a hearing aid on the contralateral ear.
3. ‘Single sided deafness (SSD)’: CI users with normal contralateral hearing with a PTA4 of 25 dB HL or less (n = 7).

### Stimuli

The audiovisual version of the female German matrix sentence test as proposed by Llorach et al. (2022) was used. The matrix sentence test uses 5-word sentences. The sentences follow a fixed structure of 5-word categories: name-verb-number-adjective-noun. In total, there are 50 possible words, 10 for each word category. The audiovisual version of the matrix sentence test combines the audio speech material of the female German matrix sentence test (Wagner et al., 2014) and video recordings of the speaker (Llorach et al., 2020). It enables a realistic and detailed analysis of speech perception and underlying mechanisms with acoustic and visual signals (Llorach et al., 2022).

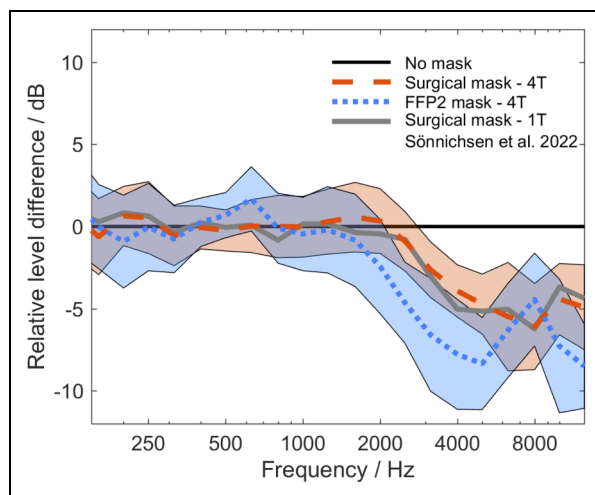
The masked stimuli were created for this experiment. To create the masked sentences, a simple mask-shaped object was added to the video of the audiovisual condition to mask the mouth of the speaker (Figure 1), and the audio signal was filtered according to the attenuation patterns of a surgical mask type II (EN14683) and an FFP2 mask type NR (EN 149:2001 + A1:2009). The attenuation patterns were created by recording 30 matrix-sentences each uttered by two female and two male speakers with and without the two face mask types. A finite impulse response filter was created for each face mask type. The filter parameters were based on the difference in third-octave frequency spectra between the sentences with and without face mask. The

**Table 1.** Demographic Data of CI Users and Normal-Hearing Participants. Data are Represented as Mean  $\pm$  Standard Deviation (SD), if Applicable

Category	CI Users (n = 37)	Normal-Hearing (n = 10)
Gender, No. (%)		
Female	22 (59.5%)	5 (50.0%)
Male	15 (40.5%)	5 (50.0%)
Age (years)		
Mean $\pm$ SD	$59.2 \pm 16.2$	$23.3 \pm 3.7$
Range	19–83	19–32
Age at implantation (years)		
Mean $\pm$ SD	$55.4 \pm 18.2$	
Range	11–82	
Duration of CI use (months)		
Mean $\pm$ SD	$43.4 \pm 41.3$	
Range	2–197	
Contralateral hearing status, No. (%)		
SSD	7 (18.9%)	
AHL	16 (43.2%)	
CI-only	14 (37.8%)	
Manufacturer, No. (%)		
Advanced Bionics	2 (5.4%)	
Cochlear	26 (70.3%)	
MED-EL	9 (24.3%)	



**Figure 1.** Test material example of test conditions with and without mask. For audiovisual conditions a video of the speaker was presented with the auditory speech signal (left). For surgical mask and FFP2 mask conditions a mask-shaped object was added to cover the speaker's mouth (right) and the acoustic signal was modified according to the respective acoustic filter properties. Adapted with permission from Wolters Kluwer Health, Inc. from Sönnichsen et al. (2022), How Face Masks Interfere With Speech Understanding of Normal-Hearing Individuals: Vision Makes the Difference. *Otology & Neurotology*, 43(3), 282–288. <https://doi.org/10.1097/MAO.0000000000003458>. The Creative Commons license does not apply to this content. Use of the material in any format is prohibited without written permission from the publisher, Wolters Kluwer Health, Inc. Please contact [permissions@lww.com](mailto:permissions@lww.com) for further information.



**Figure 2.** Mean spectral differences of four talkers' speech (4 T) spoken with surgical (red dashed line) and FFP2 mask (blue dotted line) normalized to uncovered speech (black line) are shown. Shaded areas refer to standard deviation of spectral differences on single sentence level (red: surgical mask, blue: FFP2 mask). For reasons of comparison, spectral attenuation of a surgical mask with one talker (1 T) from Sönnichsen et al. (2022) is replotted (gray line).

attenuation patterns of the two masks are shown in Figure 2. The surgical mask led to an average attenuation of 4 dB and the FFP2 mask of 6.4 dB in the frequency range between 2.5 and 8 kHz. The original speech material of the female German matrix sentence test was filtered accordingly and set to the original rms-level.

## Experimental Setup

The measurement setup was equal to the setup described in Sönnichsen et al. (2022) and is described in this section. Measurements were conducted in a sound-treated room. Listeners were seated in front of a loudspeaker (8030C studio monitor, Genelec, Iisalmi, Finland) and a 23.8" screen (P2419H, DELL GmbH, Frankfurt, Germany). Loudspeaker and screen were placed 80 cm apart from the listener. Matlab2018 was used to perform the measurements. The audio signals were presented through a soundcard (RME Fireface uc). The video was time aligned to the audio signal as described in Llorach et al. (2022) and reproduced using VLC media player 3.0.3. The audio signals were calibrated to a level of 85 dB SPL (C) using a PCA 322A level meter.

## Conditions

Five conditions were used in this experiment:

- Audiovisual (AV)
- Audio-only (AO)
- Speechreading (SR)
- Masked audiovisual with surgical mask (AVM-S)
- Masked audiovisual with FFP2 mask (AVM-FFP)

Each condition consisted of 20 five-word-sentences. speech reception thresholds (SRT) of 80% intelligibility ( $SRT_{80}$ ) were determined adaptively in test-specific stationary noise, i.e., the noise was presented at a fixed level of 65 dB SPL and the level of the speech signal changed according to the

correct number of words answered by the participant (see Llorach et al., 2022 for details). A single  $SRT_{80}$  value was given for each 20-sentence set according to the adaptive procedure. The initial signal-to-noise ratio (SNR) was set to  $-5$  dB SNR. The presented SNR was limited to 20 dB SNR, at which the measurement was aborted automatically to avoid levels above 85 dB SPL.

In the audio-only condition the speech signal and the noise were presented while the screen remained black. The speechreading condition consisted of video recordings of the speaker combined with a 65 dB SPL noise and no speech signal. This condition was used to measure the ability to speechread. We adopted a definition of speechreading from Strelnikov et al. (2009) who define it as speech perception by solely viewing a speaker's facial expressions and lip movements.

### Experiment Procedure

Listeners were given an instruction sheet containing the  $5 \times 10$  word matrix of the speech material. Listeners conducted two lists of 20 sentences in the AV condition for training. Afterwards, the five conditions described before (AV, AO, SR, AVM-S, AVM-FFP) were conducted in randomized order. After listening to a sentence, listeners were asked to repeat the words they understood. Guessing was permitted and breaks between conditions were taken as needed.

All measurements were performed best-aided, meaning SSD participants with unilateral CI and contralateral NH, AHL participants with unilateral CI and contralateral hearing aid if available, and CI-only participants with bilateral CI or unilateral CI and no contralateral hearing aid. Binaural

testing was chosen for all conditions to more realistically reflect daily life situations compared to monaural testing.

## Results

### Data Analysis and Statistical Tests

Data are described by median values with range, unless stated otherwise.

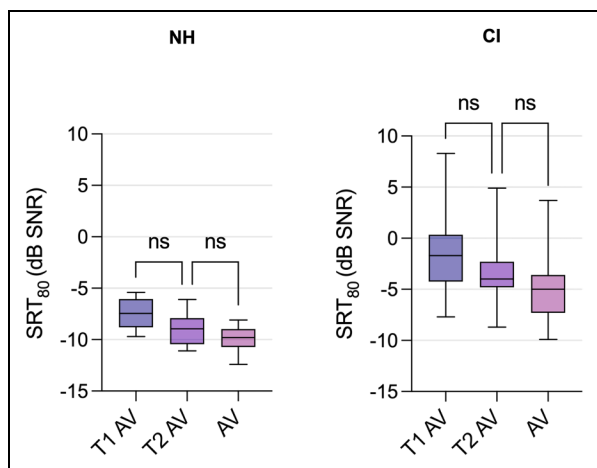
Besides individually measured  $SRT_{80}$  values for the different conditions two parameters were calculated for each participant and used for statistical analysis:

- Visual benefit describing the difference in  $SRT_{80}$  between the audiovisual and the audio-only condition ( $SRT_{80}(AV) - SRT_{80}(AO)$ ). It defines the improvement in speech intelligibility by adding the visual signal to the audio signal.
- Reduction of speech intelligibility solely due to the acoustic attenuation of the face masks described by the difference in SRT between the two masked audiovisual conditions and the audio-only condition ( $SRT_{80}(AVM-S) - SRT_{80}(AO)$  and  $SRT_{80}(AVM-FFP) - SRT_{80}(AO)$ ).
- The overall reduction in speech intelligibility by face masks combines the loss of the visual signal with the individual acoustic attenuation of each mask ( $SRT_{80}(AVM-S) - SRT_{80}(AV)$  and  $SRT_{80}(AVM-FFP) - SRT_{80}(AV)$ ).

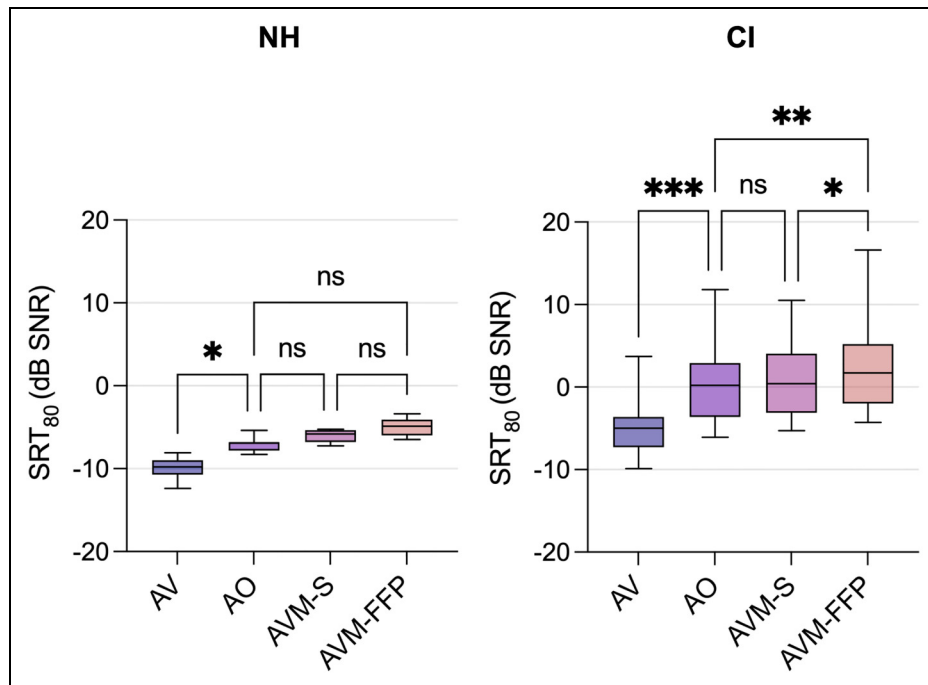
Statistical analysis was conducted using GraphPad Prism (GraphPad Software, Version 9.2.0, San Diego, CA).  $SRT_{80}$  and  $SRT_{80}$ -differences were tested on normal distribution using the Shapiro-Wilk test. If normal distribution was violated, non-parametric test statistics were used. Friedman's tests with Dunn's multiple comparisons were applied on the SRT data for within subjects' analysis separately for NH and CI users (Figure 3 and Figure 4). Between subjects' analysis was performed using the Kruskal-Wallis test with Dunn's multiple comparisons separately on the SRT-differences for visual benefit and the acoustic attenuation of the face masks (Figure 5). Between subjects' analysis of the combined effect of missing visual information and acoustic attenuation due to the face masks was conducted using a separate Brown-Forsythe and Welch ANOVA test with Dunnett's T3 multiple comparisons test. Speechreading scores between NH and CI users overall were compared using the Mann-Whitney-U test and subgroup analysis was conducted using the Kruskal-Wallis test. Linear Regressions and Spearman's rank correlation coefficient ( $r_s$ ) were calculated for speechreading scores vs. visual benefit (Figure 6).

All p-values were adjusted for multiple comparisons and a p-value  $< 0.05$  was considered statistically significant.

No significant differences between the two training lists and between the second training list and the audiovisual condition were detected for CI users or NH (Figure 3), showing that sufficient training was applied prior to the measurements of different conditions.



**Figure 3.** Speech reception threshold of 80% intelligibility ( $SRT_{80}$ ) of audiovisual training test lists (T1 AV, T2 AV) and the audiovisual test condition (AV) are shown for NH (left panel) and CI users (right panel). Box plots with line at median, lower and upper box level at 25% and 75% percentiles and whiskers for minimum and maximum values. ns = not statistically significant.



**Figure 4.** Speech reception threshold values of 80% intelligibility ( $SRT_{80}$ ) for audiovisual (AV), audio-only (AO), surgical mask (AVM-S) and FFP2 mask (AVM-FFP) for NH (left panel) and CI users (right panel). Box plots with line at median, lower and upper box level at 25% and 75% percentiles and whiskers for minimum and maximum values. ns = not statistically significant, significance level: \*  $p \leq 0.05$ , \*\*  $p \leq 0.01$ , \*\*\*  $p \leq 0.001$ .

### Speech Reception Thresholds

Figure 4 shows  $SRT_{80}$ -values and statistical comparisons across conditions for CI users (right panel) and NH (left panel). CI users demonstrated a median  $SRT_{80}$  of 0.2 dB SNR (range:  $-6.1$  to  $11.8$ ) in the audio-only condition (AO), meaning that 80% of words were understood correctly, if the speech was presented 0.2 dB louder than the noise of 65 dB SPL. There was a significant improvement to a median  $SRT_{80}$  of  $-5.0$  dB SNR (range:  $-9.9$  to  $3.7$ ,  $p < 0.001$ ), when visual cues were added to the audio signal. On the other hand, the FFP2 mask led to a significant reduction in  $SRT_{80}$  (median  $SRT_{80}$ : 1.7 dB SNR, range:  $-4.3$  to  $16.6$ ,  $p = 0.003$ ) compared to AO. The surgical mask did not show any significant effect on  $SRT_{80}$  (median  $SRT_{80}$ : 0.4 dB SNR, range:  $-5.3$  to  $10.5$ ,  $p > 0.99$ ) compared to AO.

NH yielded a median  $SRT_{80}$  of  $-6.9$  dB SNR in AO (range:  $-8.3$  to  $-5.4$  dB SNR). Compared to CI users NH showed a less distinct but still significant improvement in AV (median  $SRT_{80}$ :  $-9.8$  dB SNR, range:  $-12.4$  to  $-8.1$  dB SNR,  $p = 0.02$ ) when having visual and acoustic information available. The acoustic attenuation of the masks did not show a statistically significant effect on SRT compared to AO. Median  $SRT_{80}$  were  $-5.9$  dB SNR (range:  $-7.3$  to  $-5.3$  dB SNR,  $p > 0.99$ ) for the surgical mask and  $-4.9$  dB SNR (range:  $-6.5$  to  $-3.4$  dB SNR,  $p = 0.06$ ) for the FFP2-mask.

### Visual Benefit and Acoustic Attenuation

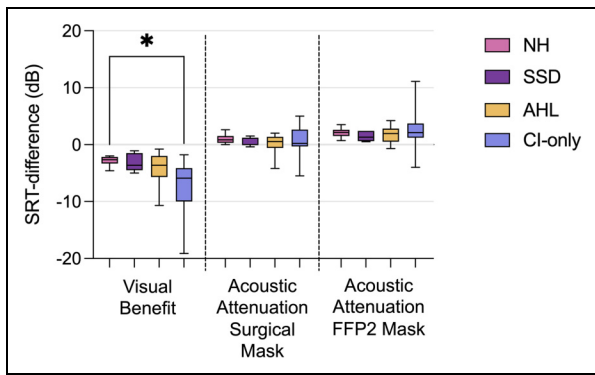
The median visual benefit was  $-4.5$  dB for CI users and  $-2.7$  dB for NH. The acoustic attenuation was less due to the surgical mask (CI users: 0.2 dB, NH: 0.9 dB) than due to the FFP2 mask (CI users: 2.1 dB, NH: 2.2 dB). The combination of visual information loss and acoustic attenuation lead to a reduction in SRT of 5.0 dB and 3.8 dB due to the surgical mask, and of 6.5 dB and 5.1 dB due to the FFP2 mask for CI users and NH, respectively.

Subcategorization according to hearing status of the contralateral ear revealed distinct differences between CI subgroups and NH. Throughout all conditions the SSD group performed better than the AHL group and similar as NH. The AHL group performed better than the CI-only group. The visual benefit was significantly larger for the CI-only group ( $-5.9$  dB) compared to NH ( $-2.7$  dB,  $p = 0.025$ ). For AHL and SSD the visual benefit of  $-3.6$  dB for both groups was in between NH and CI-only, but not significantly different from the other subgroups (see Figure 5, left).

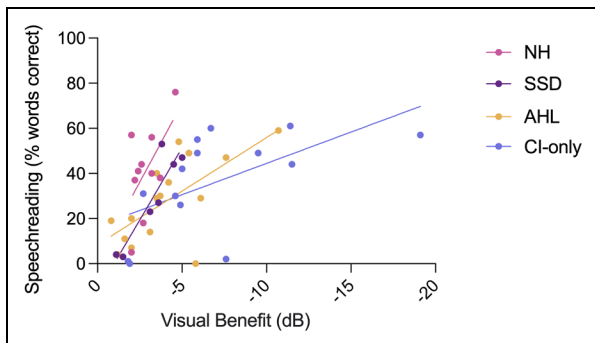
The effect of acoustic attenuation of the masks on speech intelligibility was similar for all CI subgroups and NH, and showed the largest reduction in SRT of 1.3 to 2.2 dB for the FFP2 mask and less than 1 dB for the surgical mask (compare Figure 5 middle for the surgical mask and right for the FFP2 mask).

The combined effect of missing visual information and acoustic attenuation due to the face masks was largest for





**Figure 5.** Visual benefit (left, difference of speech reception threshold of 80% intelligibility ( $SRT_{80}$ ) between audio-only and audiovisual condition), and acoustic attenuation effect of the surgical mask (middle,  $SRT_{80}$ -difference between surgical mask and audio-only condition) and FFP2 mask (right,  $SRT_{80}$ -difference between FFP2 mask and audio-only condition) are shown for subgroups of CI users and normal-hearing listeners (NH). CI user subgroups: single-sided-deafness (SSD), asymmetric hearing loss (AHL) and CI-only. Box plots with line at median, lower and upper box level at 25% and 75% percentiles and whiskers for minimum and maximum values. significance level: \*  $p \leq 0.05$ .



**Figure 6.** Correlation of speechreading scores and visual benefit shown for subgroups of CI users and normal-hearing listeners (NH). CI user subgroups: single-sided-deafness (SSD), asymmetric hearing loss (AHL) and CI-only. Lines depict linear regression lines.

the CI-only group with an SRT reduction of 7.8 dB (range: 2.0 to 13.6) for the surgical mask and 8.7 (range: 3.1 to 15.6) dB for the FFP2 mask. Significant differences were found compared to NH (surgical: 3.8 dB, range: 2.3 to 5.8,  $p=0.005$ ; FFP2: 5.1 dB, range: 2.7 to 6.9,  $p=0.01$ ) and SSD (surgical: 4.1 dB, range: 1.4 to 5.3,  $p=0.007$ ; FFP2: 5.1 dB, range: 2.7 to 6.0,  $p=0.007$ ).

### Speechreading and its Correlation to Visual Benefit

Overall, CI users showed on average slightly lower speechreading scores than NH (CI: 30%, range: 0 to 61%; NH:

40.5%, range: 5 to 76%). However, no significant difference was found ( $p=0.25$ ). A further analysis on subgroups (including NH as one of the groups) did not show significant differences either (CI-only: 43%, range: 0% to 61%; AHL: 29%, range: 0% to 59%; SSD: 27%, range: 3% to 53%;  $p=0.37$ ).

The ability to speechread was significantly correlated with the visual benefit over all groups of listeners ( $r_s=-0.59$ ,  $p<0.001$ ). Better speechreading performance correlated with a larger visual benefit, although large individual differences could be noted. Further analyses for each subgroup and NH revealed significant rank correlations for SSD, AHL and CI-only ( $r_s=-0.86$ ,  $p=0.02$ ;  $r_s=-0.62$ ,  $p=0.01$ ;  $r_s=-0.71$ ,  $p=0.005$ , respectively). Correlations are shown in Figure 6.

## Discussion

Face mask wearing during the COVID-19 pandemic revealed the importance of visual and unfiltered acoustic cues for adequate speech understanding for CI users and NH. Daily communication and important aspects of CI rehabilitation were adversely affected by the pandemic and associated preventive measures such as face masks (Aschendorff et al., 2021; Gordon et al., 2021). However, they provided an effective and therefore necessary tool to reduce virus transmission (Howard et al., 2021).

We demonstrated that the detrimental effect of surgical and FFP2 masks was largely due to the loss of the visual signal and to a smaller proportion by attenuation of high frequency regions in the acoustic signal, if the overall level was kept constant. This applied to CI users and NH. However, the effect was more pronounced for CI users, especially for those who exclusively rely on their CI for hearing. A recent study suggests that our findings are in line with the individual subjective perception of CI users in everyday life situations during the COVID-19 pandemic. Homans and Vroegop (2021) reported that 84% of CI users believed that the loss of lipreading ability alone or in combination with the attenuated speech signal was the reason for limited communication with face masks. Only 14% felt that the attenuated speech signal alone impaired their ability to communicate (Homans & Vroegop, 2021).

### Acoustic Effects

The acoustic attenuation properties of the two mask types were similar as in previous studies (Corey et al., 2020; Mcleod et al., 2022; Sönnichsen et al., 2022). The amount of SRT reduction for NH due to acoustic attenuation was less for the surgical (0.9 dB) and FFP2 (2.2 dB) mask than could be expected according to the importance of spectral weighting to speech intelligibility by means of the speech intelligibility index as presented by Mcleod et al. (2022). Note that possible differences according to level with and without mask were equalized in our study. The spectral attenuation of the FFP2 mask had a significant effect on speech intelligibility for the CI users in our study. In contrast, Vos

et al. (2021) found no significant difference in speech recognition of CI users for acoustic filter properties of N95 masks, which have very similar acoustic properties compared to FFP2 masks (Corey et al., 2020). However, their experimental setup differed from that of our study since speech understanding was only tested in quiet (Vos et al., 2021), which is less challenging for CI users (Fu & Nogaki, 2005). In addition, their cohort appeared to consist of a high percentage of CI users with an above-average performance of >80% in the AzBio sentence test in quiet (Vos et al., 2021), possibly leading to ceiling effects (Brant et al., 2018).

### Speechreading

Our data show that the main proportion of the reduction in speech intelligibility by face masks was due to the loss of the visual signal, which is equivalent to the loss of the possibility to speechread (see Figure 5, left). Speechreading (for a definition please see methods section) is a process that is influenced by many different parameters related to the participant (e.g., hearing loss (Auer & Bernstein, 2007), age (Schreitmüller et al., 2018; Tye-Murray et al., 2016), motivation (Van Tasell & Hawkins, 1981)) or the test material (e.g., talker familiarity (Yakel et al., 2000), complexity of the test material (Auer, 2010; Kaiser et al., 2003)). Previous studies have shown an advantage in speechreading for hearing-impaired listeners and CI users compared to NH (Auer & Bernstein, 2007; Kaiser et al., 2003; Rouger et al., 2007; Schreitmüller et al., 2018). Within our study it appears that either CI users underperformed or NH overperformed in the speechreading condition as both scores were similar. However, it is important to note that the authors' experience is that speechreading scores of both groups are highly individual with a great range of individual results, presumably due to the different factors described previously. CI users in the study of Schreitmüller et al. (2018) showed a small advantage in speechreading scores (CI mean speechreading score: 37.7%) compared to our study in a similar experimental setup. The authors discuss that participants with low scores partially showed impaired motivation (Schreitmüller et al., 2018). It has been shown that manipulation of motivation to guess can influence speechreading test results (Van Tasell & Hawkins, 1981). Likewise, the combination of an already challenging speechreading condition with a persistent noise distraction could have hindered motivation to interpret the visual signals and led to low scores in some of our participants. Hearing-impaired listeners have been shown to be significantly more distracted by noisy environments than NH, which might have led to slightly reduced (but non-significant) speechreading scores of CI users compared to NH in our study (Fu & Nogaki, 2005; Zheng et al., 2017). This effect has been demonstrated for speechreading in babble for NH (Myerson et al., 2016). It implies that measurements without noise might improve speechreading scores in CI users.

In our study a comparison of younger CI users (<50 years,  $n = 8$ ), older CI users (>50 years,  $n = 29$ ) and NH revealed no significant age-dependent differences in speechreading scores although this has been shown before. Schreitmüller et al. (2018) describe an age dependency of visual-only (speechreading) conditions as speechreading scores decrease with increasing age of NH and CI users. Our results do not match these findings. However, they could be possibly confounded by the different sample sizes in our study. Also, the effects of age-related differences in speechreading scores on audiovisual speech understanding seem to be less pronounced (Huysse et al., 2014; Tye-Murray et al., 2016).

### Visual Benefit

Visual benefit which is the difference between AV and AO conditions consists not only of the information obtained from speechreading, but also of additional information derived from the complex process of audiovisual integration (Grant et al., 1998; Stevenson et al., 2017). This process is influenced by the individual's ability to integrate auditory and visual information and depends on many factors such as lexical difficulty or talker characteristics (Kaiser et al., 2003).

Two previous studies using the audiovisual German matrix sentence test found greater visual benefits in young and elderly NH than our study (Gieseler et al., 2020; Llorach et al., 2022). The study by Gieseler et al. (2020) used fluctuating noise (ICRA-250) instead of stationary noise, therefore the visual benefit is not directly comparable to our study. In the study by Llorach et al. (2022), the participants did more training trials and were tested with more lists in total (the training trials and the test conditions involved 23 lists). It is likely that they became more familiar with the speaker's face and therefore improved their audiovisual scores (Lander & Davies, 2008).

### Audiovisual Integration

Due to the experimental design speechreading and audiovisual integration are inseparable in our study. Both influence the visual benefit, but the authors can only estimate the extent to which loss of audiovisual integration alone causes altered communication by face masks. Interestingly, we found that some participants who scored 0% in the SR condition, still benefited from additional visual input (Figure 6). When data from the ten CI users with the lowest SR scores (3.5% words correct, range: 0 to 14) were analyzed we still found a visual benefit of  $-1.9$  dB (range:  $-11.5$  to  $-1.2$ ). These results are consistent with the findings of Grant et al. (1998) who described the same phenomenon for participants with low speechreading performances in their study. This highlights the importance of audiovisual integration despite limited speechreading skills or poor motivation to actively speechread.

Face masks, especially because of the lack of visual information, showed a greater reduction in speech intelligibility for the

CI-only group compared to NH (Figure 5). This is in accordance with a study that found greater integration proficiency in CI users compared to NH. It seems that CI users continue to rely on visual information and audiovisual integration after CI surgery, especially in challenging environments (Rouger et al., 2007). This at least partially explains the greater impact of face masks on speech intelligibility in CI-only users. However, NH and SSD listeners in our study population appear to use similar speech recognition strategies as the reduction in speech intelligibility due to face masks was rather similar.

### Limitations

The performance of AHL listeners was more variable than of NH and SSD participants. As mentioned before, the AHL group was very diverse regarding the hearing status of the contralateral ear. It can be assumed that listeners with weaker hearing abilities on the contralateral ear made more use of audiovisual integration. Since binaural testing was used the hearing status of the contralateral ear was critical for overall speech intelligibility within this experimental setup. Monaural testing could determine the impact of face masks solely on the CI ear, but our free field approach is not suitable to completely exclude the contralateral ear. A possible solution could be streaming the audio signal, so it can only be detected with the CI. However, in this study we were aiming to reflect real-life situations as close as possible with accounting for different hearing status of the contralateral ear and its impact on binaural hearing with CI.

### Conclusion

CI users rely considerably on visual cues for sufficient speech understanding in noise. Therefore, face mask wearing during the COVID-19 pandemic impairs daily communication, particularly by the loss of the visual signal. CI users who rely solely on their CI for hearing are especially vulnerable to this impairment. NH seem less impaired by masks because of a greater utilization of the acoustic signal. However, both groups benefit from visual cues when communicating. These aspects should be considered when communicating while wearing a face mask.


### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Deutsche Forschungsgemeinschaft, (grant number 352015383)

### ORCID iD

Rasmus Sönnichsen  <https://orcid.org/0000-0003-3574-7123>

### References

- Aschendorff, A., Arndt, S., Kröger, S., Wesarg, T., Ketterer, M. C., Kirchem, P., Pixner, S., Hassepaß, F., & Beck, R. (2021). Quality of cochlear implant rehabilitation under COVID-19 conditions. *HNO*, *69*(S1), 1–6. <https://doi.org/10.1007/s00106-020-00923-z>
- Auer, E. T. (2010). Investigating speechreading and deafness. *Journal of the American Academy of Audiology*, *21*(03), 163–168. <https://doi.org/10.3766/jaaa.21.3.4>
- Auer, E. T., & Bernstein, L. E. (2007). Enhanced visual speech perception in individuals with early-onset hearing impairment. *Journal of Speech, Language, and Hearing Research*, *50*(5), 1157–1165. [https://doi.org/10.1044/1092-4388\(2007\)080](https://doi.org/10.1044/1092-4388(2007)080)
- Brant, J. A., Eliades, S. J., Kaufman, H., Chen, J., & Ruckenstein, M. J. (2018). Azbio speech understanding performance in quiet and noise in high performing cochlear implant users. *Otology & Neurotology*, *39*(5), 571–575. <https://doi.org/10.1097/MAO.0000000000001765>
- Brown, V. A., Van Engen, K. J., & Peelle, J. E. (2021). Face mask type affects audiovisual speech intelligibility and subjective listening effort in young and older adults. *Cognitive Research: Principles and Implications*, *6*(1), 49. <https://doi.org/10.1186/s41235-021-00314-0>
- 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. <https://doi.org/10.1121/10.0002279>
- Fu, Q.-J., & Nogaki, G. (2005). Noise susceptibility of cochlear implant users: The role of spectral resolution and smearing. *Journal of the Association for Research in Otolaryngology*, *6*(1), 19–27. <https://doi.org/10.1007/s10162-004-5024-3>
- Gieseler, A., Rosemann, S., Tahden, M., Wagener, K. C., Thiel, C., & Colonius, H. (2020). Linking Audiovisual Integration to Audiovisual Speech Recognition in Noise (OSF Preprints) [Preprint]. Open Science Framework. <https://doi.org/10.31219/osf.io/46caf>
- Goldin, A., Weinstein, B., & Shiman, N. (2020). How do medical masks degrade speech reception? *Hearing Review*, *27*(5), 8–9. <https://hearingreview.com/hearing-loss/health-wellness/how-do-medical-masks-degrade-speech-reception>
- Gordon, K. A., Daien, M. F., Negandhi, J., Blakeman, A., Ganek, H., Papsin, B., & Cushing, S. L. (2021). Exposure to spoken communication in children with cochlear implants during the COVID-19 lockdown. *JAMA Otolaryngology–Head & Neck Surgery*, *147*(4), 368–376. <https://doi.org/10.1001/jamaoto.2020.5496>
- Grant, K. W., Walden, B. E., & Seitz, P. F. (1998). Auditory-visual speech recognition by hearing-impaired subjects: Consonant recognition, sentence recognition, and auditory-visual integration. *The Journal of the Acoustical Society of America*, *103*(5), 2677–2690. <https://doi.org/10.1121/1.422788>
- Homans, N. C., & Vroegop, J. L. (2021). Impact of face masks in public spaces during COVID-19 pandemic on daily life communication of cochlear implant users. *Laryngoscope Investigative Otolaryngology*, *6*(3), 531–539. <https://doi.org/10.1002/liv.2.578>



- Howard, J., Huang, A., Li, Z., Tufekci, Z., Zdimal, V., van der Westhuizen, H.-M., von Delft, A., Price, A., Fridman, L., Tang, L.-H., Tang, V., Watson, G. L., Bax, C. E., Shaikh, R., Questier, F., Hernandez, D., Chu, L. F., Ramirez, C. M., & Rimoin, A. W. (2021). An evidence review of face masks against COVID-19. *Proceedings of the National Academy of Sciences*, *118*(4), e2014564118. <https://doi.org/10.1073/pnas.2014564118>
- Huyse, A., Leybaert, J., & Berthommier, F. (2014). Effects of aging on audio-visual speech integration. *The Journal of the Acoustical Society of America*, *136*(4), 1918–1931. <https://doi.org/10.1121/1.4894685>
- Kaiser, A. R., Kirk, K. I., Lachs, L., & Pisoni, D. B. (2003). Talker and lexical effects on audiovisual word recognition by adults with cochlear implants. *Journal of Speech, Language, and Hearing Research*, *46*(2), 390–404. [https://doi.org/10.1044/1092-4388\(2003\)032](https://doi.org/10.1044/1092-4388(2003)032)
- Lander, K., & Davies, R. (2008). Does face familiarity influence speechreadability? *Quarterly Journal of Experimental Psychology*, *61*(7), 961–967. <https://doi.org/10.1080/17470210801908476>
- Lenarz, M., Sönmez, H., Joseph, G., Büchner, A., & Lenarz, T. (2012). Long-Term performance of cochlear implants in postlingually deafened adults. *Otolaryngology–Head and Neck Surgery*, *147*(1), 112–118. <https://doi.org/10.1177/0194599812438041>
- Llorach, G., Kirschner, F., Grimm, G., & Hohmann, V. (2020). Video recordings for the female German Matrix Sentence Test (OLSA). Zenodo. <https://doi.org/10.5281/zenodo.3673062>
- Llorach, G., Kirschner, F., Grimm, G., Zokoll, M. A., Wagener, K. C., & Hohmann, V. (2022). Development and evaluation of video recordings for the OLSA matrix sentence test. *International Journal of Audiology*, *61*(4), 311–321. <https://doi.org/10.1080/14992027.2021.1930205>
- Mcleod, R. W. J., Gallagher, M., Hall, A., Bant, S. P., & Culling, J. F. (2022). Acoustic analysis of the effect of personal protective equipment on speech understanding: Lessons for clinical environments. *International Journal of Audiology*, 1–6. <https://doi.org/10.1080/14992027.2022.2070780>
- Muzzi, E., Chermaz, C., Castro, V., Zaninoni, M., Saksida, A., & Orzan, E. (2021). Short report on the effects of SARS-CoV-2 face protective equipment on verbal communication. *European Archives of Oto-Rhino-Laryngology*, *278*(9), 3565–3570. <https://doi.org/10.1007/s00405-020-06535-1>
- Myerson, J., Spehar, B., Tye-Murray, N., Van Engen, K., Hale, S., & Sommers, M. S. (2016). Cross-modal informational masking of lipreading by babble. *Attention, Perception, & Psychophysics*, *78*(1), 346–354. <https://doi.org/10.3758/s13414-015-0990-6>
- Palmiero, A. J., Symons, D., Morgan, J. W., & Shaffer, R. E. (2016). Speech intelligibility assessment of protective facemasks and air-purifying respirators. *Journal of Occupational and Environmental Hygiene*, *13*(12), 960–968. <https://doi.org/10.1080/15459624.2016.1200723>
- Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., & Barone, P. (2007). Evidence that cochlear-implanted deaf patients are better multisensory integrators. *Proceedings of the National Academy of Sciences*, *104*(17), 7295–7300. <https://doi.org/10.1073/pnas.0609419104>
- Schreitmüller, S., Frenken, M., Bentz, L., Ortman, M., Walger, M., & Meister, H. (2018). Validating a method to assess lipreading, audiovisual gain, and integration during speech reception with cochlear-implanted and normal-hearing subjects using a talking head. *Ear & Hearing*, *39*(3), 503–516. <https://doi.org/10.1097/AUD.0000000000000502>
- Sönnichsen, R., Llorach Tó, G., Hochmuth, S., Hohmann, V., & Radeloff, A. (2022). How face masks interfere with speech understanding of normal-hearing individuals: Vision makes the difference. *Otology & Neurotology*, *43*(3), 282–288. <https://doi.org/10.1097/MAO.0000000000003458>
- Stevenson, R. A., Sheffield, S. W., Butera, I. M., Gifford, R. H., & Wallace, M. T. (2017). Multisensory integration in cochlear implant recipients. *Ear & Hearing*, *38*(5), 521–538. <https://doi.org/10.1097/AUD.0000000000000435>
- Strelnikov, K., Rouger, J., Barone, P., & Deguine, O. (2009). Role of speechreading in audiovisual interactions during the recovery of speech comprehension in deaf adults with cochlear implants. *Scandinavian Journal of Psychology*, *50*(5), 437–444. <https://doi.org/10.1111/j.1467-9450.2009.00741.x>
- Tye-Murray, N., Spehar, B., Myerson, J., Hale, S., & Sommers, M. (2016). Lipreading and audiovisual speech recognition across the adult lifespan: Implications for audiovisual integration. *Psychology and Aging*, *31*(4), 380–389. <https://doi.org/10.1037/pag0000094>
- Van Tasell, D. J., & Hawkins, D. B. (1981). Effects of guessing strategy on speechreading test scores. *American Annals of the Deaf*, *126*(7), 840–844. <https://doi.org/10.1353/aad.2012.1284>
- Vos, T. G., Dillon, M. T., Buss, E., Rooth, M. A., Bucker, A. L., Dillon, S., Pearson, A., Quinones, K., Richter, M. E., Roth, N., Young, A., & Dedmon, M. M. (2021). Influence of protective face coverings on the speech recognition of cochlear implant patients. *The Laryngoscope*, *131*(6), E2038–E2043. <https://doi.org/10.1002/lary.29447>
- Wagner, K., Hochmuth, S., Ahrlich, M., Zokoll, M., & Kollmeier, B. (2014, March 13). Der weibliche Oldenburger Satztest. The female version of the Oldenburg sentence test. [Conference session]. 17. Jahrestagung der Deutschen Gesellschaft für Audiologie, Oldenburg, Germany. <http://www.uzh.ch/orl/dga-ev/publikationen/tagungsbaende/tagungsbaende.html>
- Yakel, D. A., Rosenblum, L. D., & Fortier, M. A. (2000). Effects of talker variability on speechreading. *Perception & Psychophysics*, *62*(7), 1405–1412. <https://doi.org/10.3758/BF03212142>
- Zaltz, Y., Buganim, Y., Zechoval, D., Kishon-Rabin, L., & Perez, R. (2020). Listening in noise remains a significant challenge for cochlear implant users: Evidence from early deafened and those with progressive hearing loss compared to peers with normal hearing. *Journal of Clinical Medicine*, *9*(5), 1381. <https://doi.org/10.3390/jcm9051381>
- Zheng, Y., Koehnke, J., & Besing, J. (2017). Combined effects of noise and reverberation on sound localization for listeners with normal hearing and bilateral cochlear implants. *American Journal of Audiology*, *26*(4), 519–530. [https://doi.org/10.1044/2017\\_AJA-16-0101](https://doi.org/10.1044/2017_AJA-16-0101)