Spring 4-19-2021

The Acoustic Effect of Masks on Speech and Singing

Anna Birkemeier
acbirke@bgsu.edu

The Acoustic Effect of Masks on Speech and Singing

Anna Birkemeier

HONORS PROJECT

Submitted to the University Honors Program
at Bowling Green State University in partial
fulfillment of the requirements for graduation with
UNIVERSITY HONORS

19 April, 2021

Dr. Ronald Scherer, Ph.D., Department of Communication Sciences and Disorders, Advisor
Dr. Emily Pence Brown, Ph.D., College of Musical Arts, Advisor

ABSTRACT

Objectives / Introduction:

Mask-wearing has become commonplace for the public during the SARS-CoV-2 (COVID-19) pandemic. This has caused communication difficulties such as muffled speech, causing a lowered speech intelligibility and worse audibility. The purpose of this study was to determine the acoustic filtering effects of various masks.

Methods / Study Design:

The acoustic effects of a surgical mask, a Singer's Mask, and an N95 mask were obtained. A single subject spoke (and repeated) the vowels /a,i,u/ and the consonants /s, ʃ/, and sang the vowels on G4 and C5 with and without the masks, respectively. Spectra of the vowels and long-term average spectra of the consonants were used to obtain comparative results.

Results:

All of the masks created high frequency energy attenuation from zero to 15 dB in various locations of the acoustic spectrum. The masks typically act as a low pass filter, meaning that the mask attenuates higher frequencies. For speech, the N95 mask had the highest level of attenuation for the vowels and consonants around the formants and salient locations. The surgical mask had the least amount of attenuation around the formants and salient locations.

Conclusions:

Communication difficulties with using masks tend to reduce energy in frequencies above the first formant. This may lead to difficult listening conditions by reducing the ability to understand what is said or what is sung.

INTRODUCTION

Effective communication and vocal performance rely on both high intelligibility (clearly produced speech sounds) and audibility (the ability to hear the speaker). Speech therapy and music voice lessons also serve the same purpose- to aid in the improvement of the communication of a message either by speaking in the case of speech language pathology, or by a performance in the case of a musician. In the context of the SARS-CoV-2 (COVID-19) pandemic, both professions have been drastically changed to account for safety while speaking and singing. According to Mittal et al. (2020), the "transmission of respiratory infections such as COVID-19 is primarily via virus-laden fluid particles (i.e. droplets and aerosols) that are formed in the respiratory tract of an infected person and expelled from the mouth and nose during breathing, talking, coughing and sneezing" (Mittal et al., 2020, page 894).

Health professionals, essential workers, and now the public are commonly wearing masks as a form of Personal Protective Equipment (PPE) that acts to "provide 'inward' protection by filtering virus-laden aerosolized particles that would otherwise be inhaled by an uninfected person, and 'outward' protection by trapping virus-laden droplets expelled by an infected person" (Mittal et al., 2020, page 894). The American Speech-Language and Hearing Association (ASHA) has recommended that speech language pathologists follow the Center for Disease Control's (CDC) guidelines regarding when and how to use PPE, specifically with using an N95 mask or a surgical mask for maximal protection during therapy services (ASHA, 2020).

Vocalists on the other hand, specifically when they participate in group singing, have been under scrutiny ever since news broke of a choir rehearsal in Washington that infected 32 individuals and 20 from probable secondary transmission, including two deaths. Singing has since been referred to as a superemmiter, and choral associations have been working toward methods and guidelines for safe singing. The Journal of Voice has recommended singing outdoors, limiting choir sizes, shortening rehearsal times, and wearing masks to combat the spread of the virus (Naunheim et al., 2020).

The usage of masks has come with difficulties in the scope of speech language pathology and vocal performance. Professionals in the field have anecdotally noted vocal fatigue, a hoarse voice quality, and a reduction in vocal range after using a mask (Gupta & Pack, 2020). Masks seem to muffle sound, leading to worse speech recognition and prompting the wearer to raise the amplitude of the voice and in general to speak louder, which in turn can cause these negative effects.

There are recent studies that have shown that masks act as low pass filters for speech, meaning that it attenuates higher frequencies. The reduction of energy in the higher frequencies may make speech less intelligible (Goldin et al., 2020). The acoustic filtering is due to the weave and material of the masks, where the more breathable the mask is, the better the speech signal will be. Goldin et al. (2020) conducted a study on the acoustic effects of a surgical mask and two

types of N95 masks. The study used a GRAS (GRAS Sound & Vibration, Beaverton, OR.) head and torso simulator and measured white noise. The results indicated that masks attenuate frequencies around 2000 Hz to 7000 Hz. The level of attenuation ranged from 3 dB with a surgical mask to 12 dB using the two N95 masks.

Corey et al. (2020) researched the acoustic attenuation of several masks, including a surgical mask, various cloth masks, and a N95 respirator. This study had a single participant who recorded 30 second speech samples, where the participant attempted to use equal effort for each recording. The results suggested that there is little attenuation below 1000 Hz, and the strongest attenuation for all the masks was above 4000 Hz.

Bottalico et al. (2020) conducted research on the effect of masks on speech intelligibility. This study used a fabric mask, a surgical mask, and a N95 mask. Recordings available to the public were played through a Head and Torso Simulator with a Mouth Simulator to ensure equal effort level throughout the study. The recordings were played to 40 participants from their personal computers. Each participant measured the speech intelligibility on a visual analog scale (0-100). The greatest attenuation for frequencies was above 2000 Hz, which are important frequencies for speech understanding (Bottalico et al., 2020). The fabric mask had the most attenuation while the surgical and N95 mask had similar levels of attenuation.

While there are several studies on mask attenuation for speech, there are no current studies that use singing to test mask attenuation. The current studies also do not look at attenuation in reference to the location of formants, which are the most important frequencies for distinguishing between vowels. The purpose of this study was to determine the acoustic filtering effects of various masks. The research questions for this study were 1.) What are the filtering aspects of each mask? 2.) What mask will perform the best acoustically (that is, match the non-mask production the best)?

METHODOLOGY

Participant

The Primary Investigator (PI) was the single participant for this study. The PI was an undergraduate senior at a mid-sized public university in the Midwestern United States majoring in communication disorders who also has taken classical vocal lessons for 10 years.

Materials/Masks

The masks used for this study were an N95 mask, the Singer's Mask, and a surgical mask. The N95 and surgical mask have been utilized in medical settings as they appear to offer good protection against the virus. The Singer's Mask is marketed heavily to musicians because of its bill-like shape that apparently allows for better articulation. **Figure 1** displays the front and side views of the chosen masks. The surgical mask (pictured left) is loose fitting and disposable. The Singer's Mask (pictured center) protrudes to create a space between the wearer's lips and the end of the mask. The N95 mask (pictured right) has a tight fit to create a seal around the wearer's nose and mouth (FDA, 2020).

Figure 1. The front and side views of the masks chosen for this study. Pictured from left to right are the surgical mask, the Singer's Mask, and the N95 mask.

Data Collection

Recordings were made on the PI's Macbook through Praat. All recording sessions were midday, and the PI warmed up her voice before recording to ensure consistent voice quality.

For the vowels /a,i,u/ (as in Bob, bee, and boo), the PI produced the phonemes without the mask on the face for 3 seconds then with the mask on the face for 3 seconds. This process was repeated 3 times for each phoneme resulting in 3 tokens for each vowel. The phonemes were spoken as well as sung with vibrato on G4 (392 Hz) and C5 (523 Hz) for each mask. The intended fundamental frequency (f_o) for the spoken vowels was approximately 380 Hz because to the participant, this frequency seemed to be highly repeatable. For each repetition of token productions, the same gesture was used (level of effort, loudness, mouth position for each vowel, head position). Also, the same distance from the mouth to the computer microphone was used (58.42 cm)

For the sustained /s, ʃ/ productions (as in see and she), the PI produced the phonemes without the mask for 8 seconds and with the mask for 8 seconds. This process was repeated to

create 2 tokens without the mask and 2 tokens with the mask for each mask. Again, the same effort, loudness, articulator positioning, and head position was used, with a distance of 58.42 cm between the mouth and computer microphone.

Analysis

The /a,i,u/ recordings were analyzed with Praat software using the Spectral Slice feature, which displays the frequency and intensity spectrum at a chosen moment during the signal. The frequency and intensity of the first 13 harmonics (12 harmonics for the recordings sung at C5) were visually determined using the cursor and recorded into Excel. This yielded 3 spectra without a mask and 3 spectra with a mask for each mask type, which were then averaged to create an average spectrum without a mask and with a mask for each mask type.

The /s, ʃ/ recordings were analyzed using RTSect to get a Long-Term Average Spectrum (LTAS). For each mask, an LTAS was created for a trial without the mask, then a separate LTAS was created for a trial with the mask. The no mask LTAS and mask LTAS were placed on top of each other in Microsoft Word and the transparency feature was used to display the LTAS on the same graph. This process was repeated to give two LTAS graphs per phoneme. The spectra were outlined in Microsoft Paint by visual averaging the fluctuating top of the LTAS spectra to provide a reasonable smoothing of the LTAS. The PI manually measured the intensity difference between the no mask smoothed LTAS and mask smoothed LTAS using a ruler at every 1000 Hz.

Reliability

The reliability of the phonemes /a,i,u/ will be discussed relative to the similarity in production of the tokens in the results section. The participant recorded 3 tokens for each condition, and the average was taken to create a single spectrum for comparison. To account for human variability, we have decided to label 5 dB or more significant attenuation difference.

RESULTS

No Mask Spectra

Spoken Vowels /a,i,u/

Figure 2 shows the results of the spectra for the spoken sustained /a/ when the mask was not worn on the face. *Figure A* depicts the spectra for the tokens during the N95 mask recording session. The blue, orange, and gray lines depict separate tokens. The average f_o of the 3 tokens was 359.45 Hz with a frequency difference of 3.67 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.1 dB (near 360 Hz or H1) to 11.3 dB (near 2150 Hz or H6). *Figure B* shows the results of the spectra for the spoken sustained /a/ tokens during the Singer's Mask recording session when the mask was not worn. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 387.20 Hz with frequency difference of 1.57 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.3 dB (near 4250 Hz or H11) to 13.7dB (near 5030 Hz or H13). *Figure C* shows the results of the spectra for the spoken sustained /a/ tokens during the surgical mask recording session when the mask was not worn. The blue, orange, and gray lines

depict separate tokens. The average f_o of the 3 tokens was 382.83 Hz with a frequency difference of 5.4 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 3 dB (near 380 Hz or H1) to 13.2 dB (near 2680 Hz or H7). *Figure D* displays all 9 tokens of the spoken sustained /a/ without the mask on the face. The average f_o of the tokens was 376.49 Hz with a frequency difference of 31.03 Hz. The difference range of the corresponding harmonics for the entire spectrum was 9.1 dB (near 360 Hz or H1) to 29.3 dB (near 4200 Hz or H11). *Figure E* displays the average spectra of all spoken /a/ tokens. The average f_o of the tokens was 376.49 Hz. The primary peaks indicate the region of the second and third formant frequencies for the spoken /a/ vowel by the single subject.

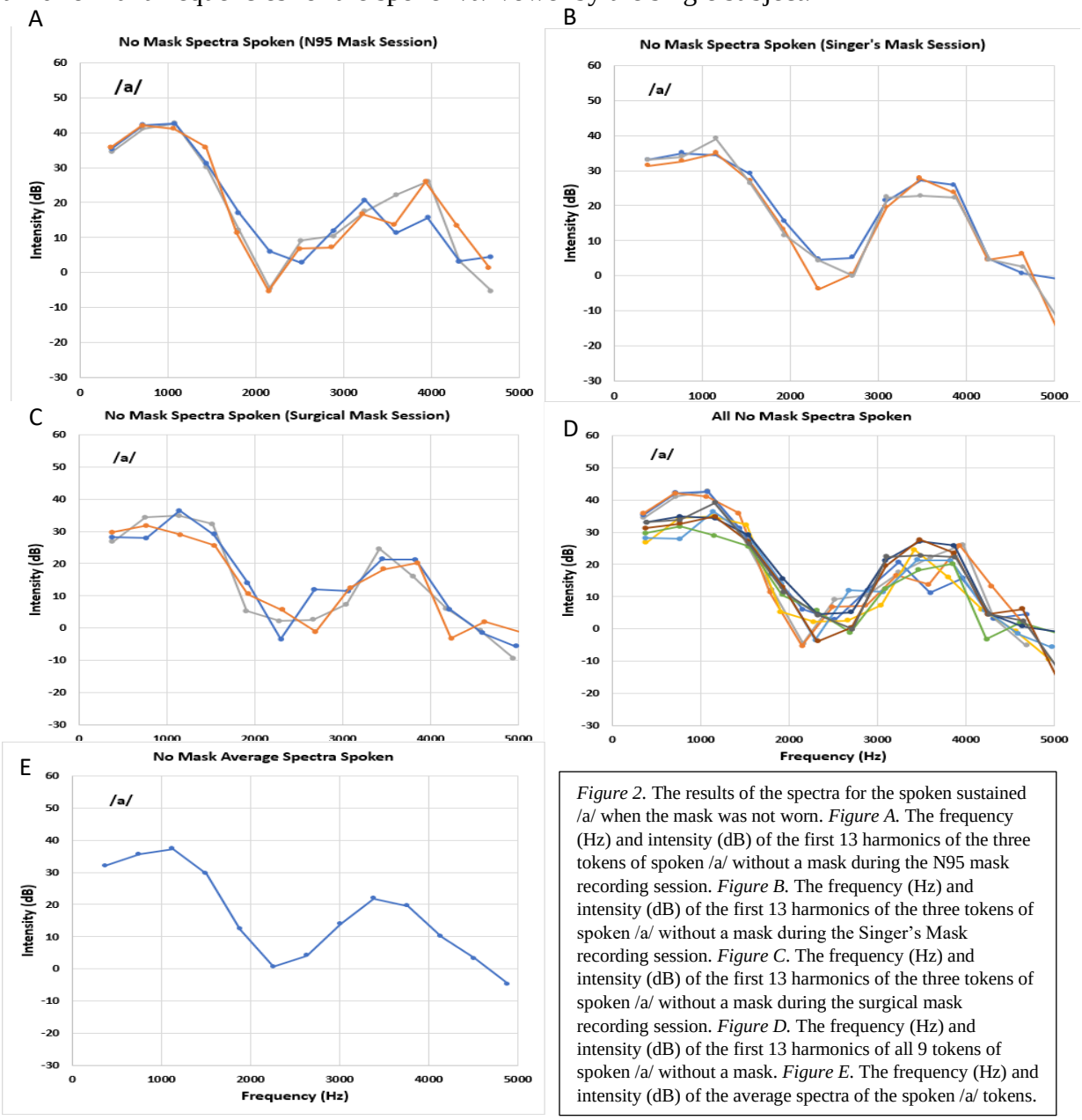

Figure 2. The results of the spectra for the spoken sustained /a/ when the mask was not worn. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of spoken /a/ without a mask during the N95 mask recording session. *Figure B.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of spoken /a/ without a mask during the Singer's Mask recording session. *Figure C.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of spoken /a/ without a mask during the surgical mask recording session. *Figure D.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of all 9 tokens of spoken /a/ without a mask. *Figure E.* The frequency (Hz) and intensity (dB) of the average spectra of the spoken /a/ tokens.

Figure 3 shows the results of the spectra for the spoken sustained /i/ when the mask was not worn on the face. *Figure A* shows the results of the spectra for the spoken sustained /i/ tokens during the N95 mask recording session when the mask was not on the face. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 360.44 Hz with a difference of 1.59 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.8 dB (near 360 Hz or H1) to 9.1 dB (near 2160 Hz or H6). *Figure B* displays the results of the spectra for the spoken sustained /i/ tokens during the Singer's Mask recording session when the mask was not worn. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 389.19 Hz with a frequency difference of 1.47 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.6 dB (near 1150 Hz or H3) to 13.6 dB (near 5050 Hz or H13). *Figure C* shows the results of the spectra for the spoken sustained /i/ during the surgical mask recording session when the mask was not worn. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 387.90 Hz with a frequency difference of 3.4 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.7dB (near 5050 Hz or H13) to 13.5 dB (near 3850 Hz or H10). *Figure D* shows the spectra of all spoken sustained /i/ tokens without a mask. The colored lines depict separate tokens. The average f_o of the tokens was 379.18 Hz with a frequency difference among the f_o was 30.59 Hz. The difference range in intensity for the corresponding harmonics for the entire spectrum was 6.2 dB (near 4100 Hz or H11) to 23.7 dB (near 5000 Hz or H13). *Figure E* displays the average spectra of all spoken sustained /i/ tokens without a mask. The average f_o of the tokens was 379.18 Hz and had an intensity of 43.08 dB. The primary peaks indicate the region of the second and third formant frequencies for the spoken /i/ vowel by the single subject.

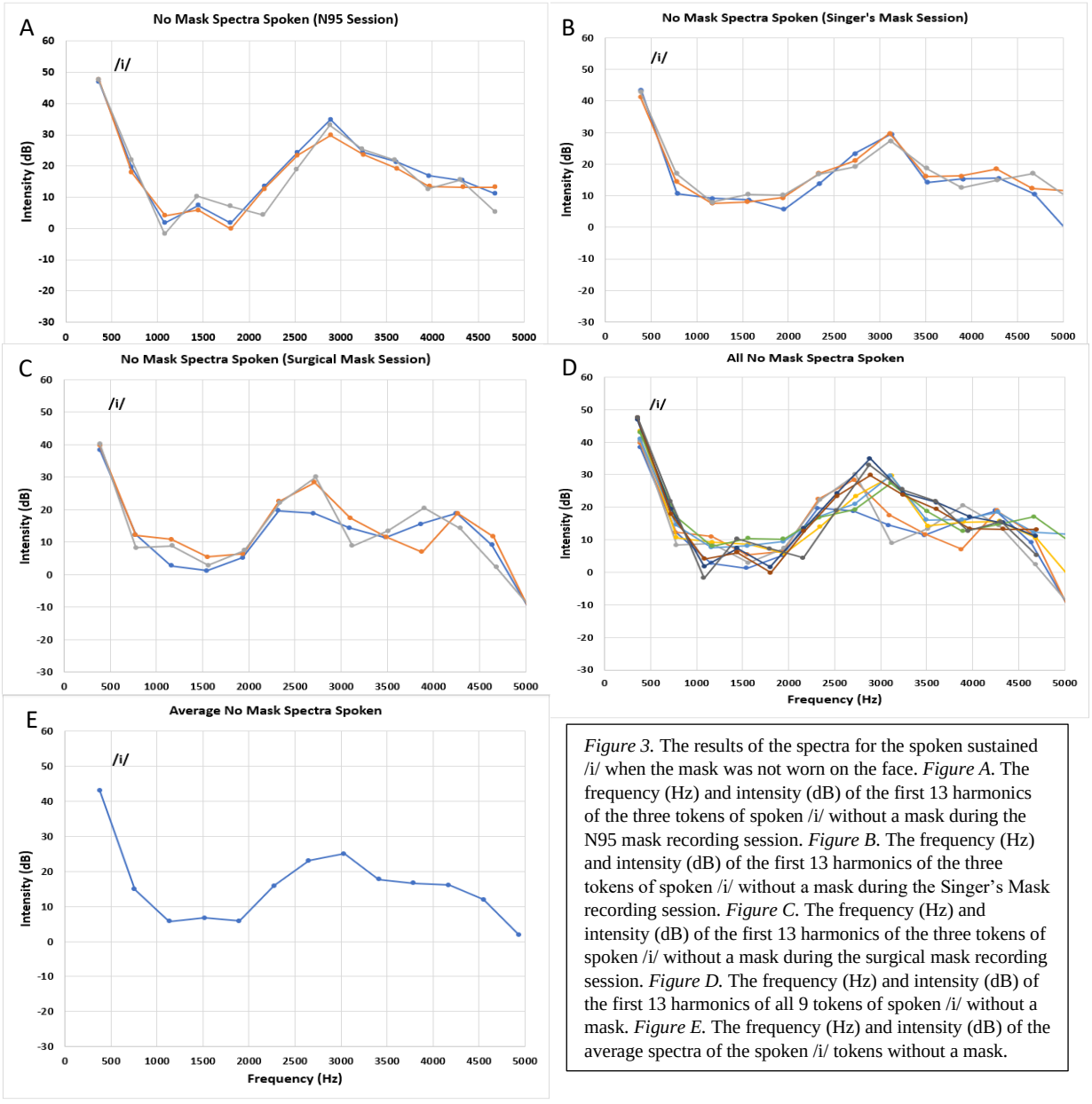

Figure 3. The results of the spectra for the spoken sustained /i/ when the mask was not worn on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of spoken /i/ without a mask during the N95 mask recording session. *Figure B.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of spoken /i/ without a mask during the Singer's Mask recording session. *Figure C.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of spoken /i/ without a mask during the surgical mask recording session. *Figure D.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of all 9 tokens of spoken /i/ without a mask. *Figure E.* The frequency (Hz) and intensity (dB) of the average spectra of the spoken /i/ tokens without a mask.

Figure 4 shows the results of the spectra for the spoken sustained /i/ when the mask was not worn on the face. *Figure A* shows the results of the spectra for the spoken sustained /u/ tokens during the N95 mask recording session when the mask was not worn. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 361.03 Hz with a frequency difference of 0.29 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.5 dB (near 360 Hz or H1) to 15.1 dB (near 4000 Hz or H11). *Figure B* shows the results of the spectra for the spoken sustained /u/ tokens during the Singer's Mask recording session when the mask was not worn. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o was 2.06 Hz with an average of 387.50 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.7 dB (near 2700 Hz or H7) to 25.7 dB (near 5050 Hz or H13). *Figure*

C shows the results of the spectra for the spoken sustained /u/ during the surgical mask recording session without a mask. The blue, orange, and gray lines depict separate tokens. The intended f_o was approximately 380 Hz and the maximum frequency difference among the f_o was 0.75 Hz. The difference range for the intensity of the corresponding harmonics was 2.5 dB (near 380 Hz or H1) to 16.3 dB (near 5000 Hz or H13). *Figure D* shows the spectra of all spoken sustained /u/ tokens without a mask. The colored lines depict separate tokens. The maximum frequency difference among the f_o was 28.69 Hz. The difference range in intensity for the corresponding harmonics for the entire spectrum was 6.9 dB (near 3700 Hz or H10) to 32.4 dB (near 4000 Hz or H12). *Figure E* displays the average spectra of all spoken sustained /u/ tokens. The average f_o of the tokens was 379.24 Hz. The primary peaks indicate the region of the second and third formant frequencies for the spoken /u/ vowel by the single subject.

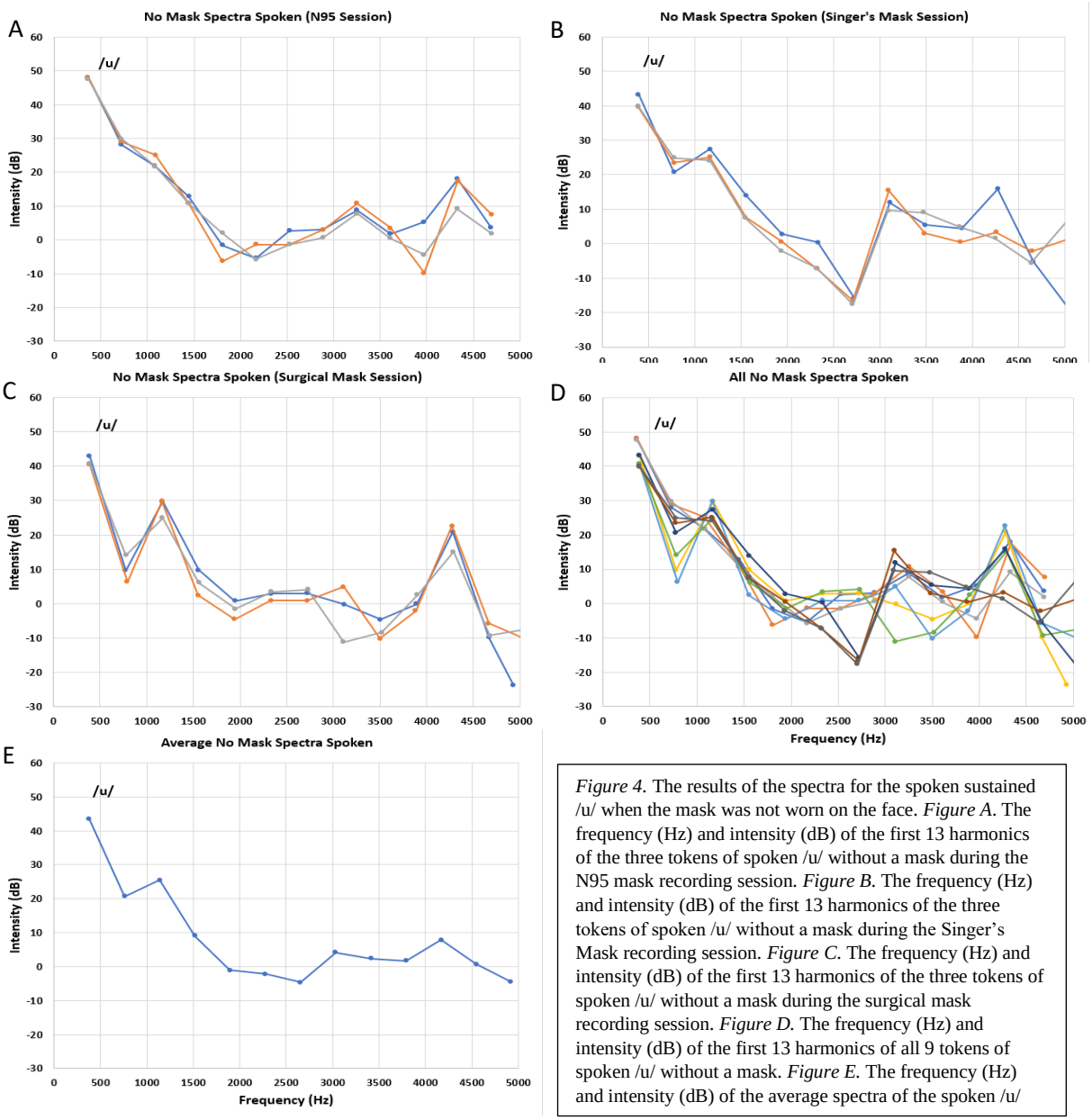

Figure 4. The results of the spectra for the spoken sustained /u/ when the mask was not worn on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of spoken /u/ without a mask during the N95 mask recording session. *Figure B.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of spoken /u/ without a mask during the Singer's Mask recording session. *Figure C.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of spoken /u/ without a mask during the surgical mask recording session. *Figure D.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of all 9 tokens of spoken /u/ without a mask. *Figure E.* The frequency (Hz) and intensity (dB) of the average spectra of the spoken /u/

Vowels on G4 /a,i,u/

Figure 5 shows the results of the spectra for the sustained /a/ sung on G4 when the mask was not worn on the face. *Figure A* displays the results of the spectra for the /a/ tokens sung sustained at G4 during the N95 mask recording session when the mask was not worn. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o was 3.04 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.4 dB (near 400 Hz or H1) to 12.4 dB (near 4200 Hz or H10). *Figure B* shows the results of the spectra for the /a/ tokens sung sustained at G4 during the Singer's Mask recording

session when the mask was not worn on the face. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o was 19.05 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.7 dB (near 1700 Hz or H4) to 19.8 dB (near 5400 Hz or H13). *Figure C* shows the results of the spectra for the /a/ tokens sung sustained at G4 during the surgical mask recording session without the mask on the face. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o 3.5 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.1 dB (near 2100 Hz or H5) to 6.5 dB (near 4550 Hz or H11). *Figure D* shows the spectra of all tokens of /a/ sung sustained on G4 without a mask. The colored lines depict separate tokens. The maximum frequency difference among the f_o 19.05 Hz. The difference range in intensity for the corresponding harmonics for the entire spectrum was 6.9 dB (near 400 Hz or H1) to 26.1 dB (near 4600 Hz or H11). *Figure E* displays the average spectra of all /a/ tokens sung on G4. The average f_o of the tokens was 416.64 Hz and had an intensity of 37.73 dB. The primary peaks indicate the region of the second and third formant frequencies for the /a/ vowel by the single subject.

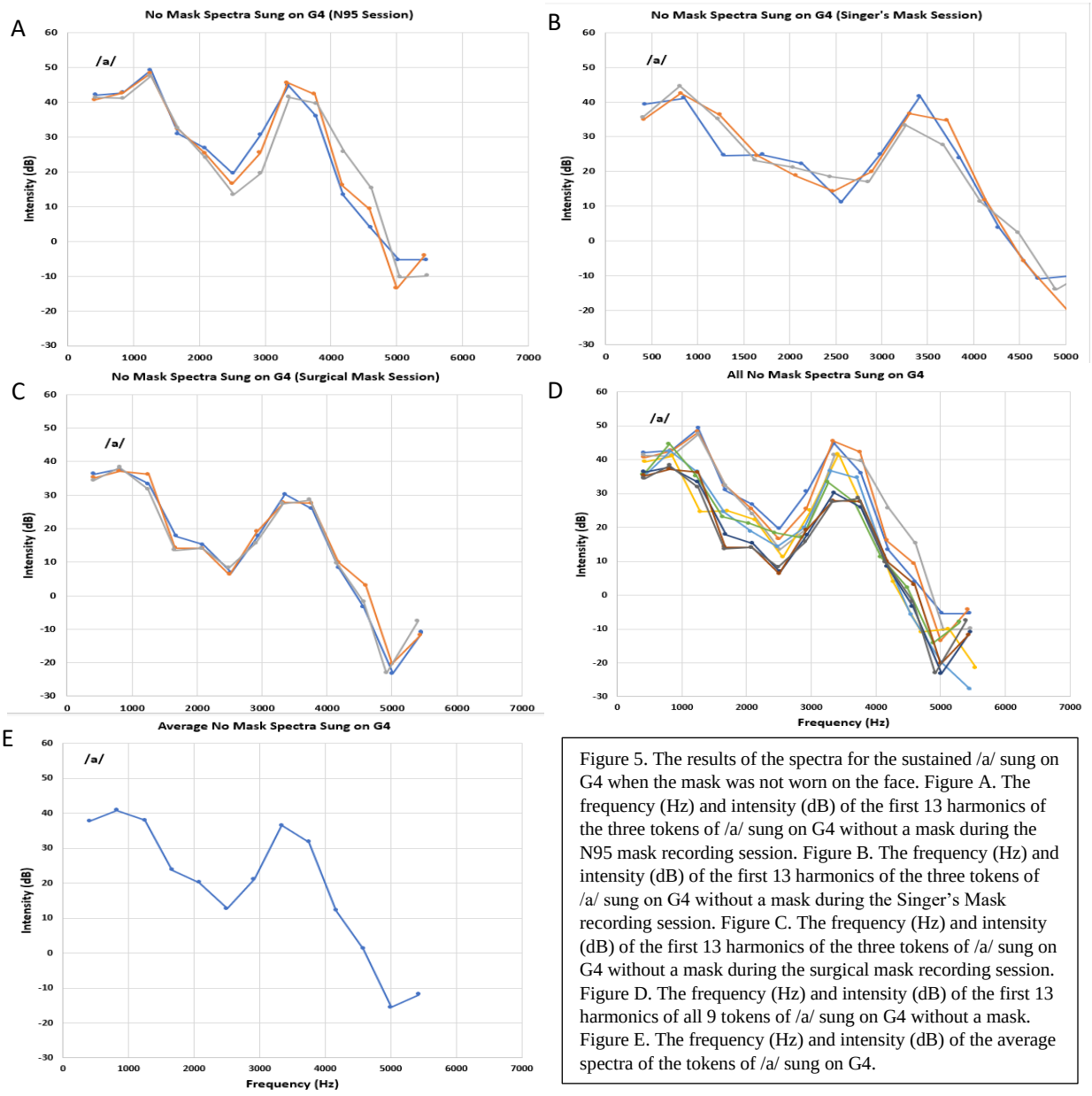

Figure 5. The results of the spectra for the sustained /a/ sung on G4 when the mask was not worn on the face. Figure A. The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of /a/ sung on G4 without a mask during the N95 mask recording session. Figure B. The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of /a/ sung on G4 without a mask during the Singer's Mask recording session. Figure C. The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of /a/ sung on G4 without a mask during the surgical mask recording session. Figure D. The frequency (Hz) and intensity (dB) of the first 13 harmonics of all 9 tokens of /a/ sung on G4 without a mask. Figure E. The frequency (Hz) and intensity (dB) of the average spectra of the tokens of /a/ sung on G4.

Figure 6 shows the results of the spectra for the sustained /i/ sung on G4 when the mask was not worn on the face. *Figure A* displays the results of the spectra for the /i/ tokens sung sustained at G4 during the N95 mask recording session without the mask on the face. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o 2.02 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.6 dB (near 4500 Hz or H11) to 13.9 dB (near 5300 Hz or H13). *Figure B* displays the results of the spectra for the /i/ tokens sung sustained at G4 during the Singer's Mask recording session when the mask was not worn on the face. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o 8.51 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.3

dB (near 1650 Hz or H4) to 22.4 dB (near 5300 Hz or H13). *Figure C* shows the results of the spectra for the /i/ tokens sung sustained at G4 during the surgical mask recording session when the mask was not worn. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o 4.7 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.2 dB (near 2500 Hz or H6) to 17.4 dB (near 5400 Hz or H13). *Figure D* shows the spectra of all tokens of /i/ sung sustained on G4 without a mask. The colored lines depict separate tokens. The maximum frequency difference among the f_o was 12.12 Hz. The difference range in intensity for the corresponding harmonics for the entire spectrum was 6.1 dB (near 400 Hz or H1) to 23.6 dB (near 5300 Hz or H13). *Figure E* displays the average spectra of all /i/ tokens sung on G4. The average f_o of the tokens was 410.16 Hz and had an intensity of 43.67 dB. The primary peaks indicate the region of the second and third formant frequencies for the /i/ vowel by the single subject.

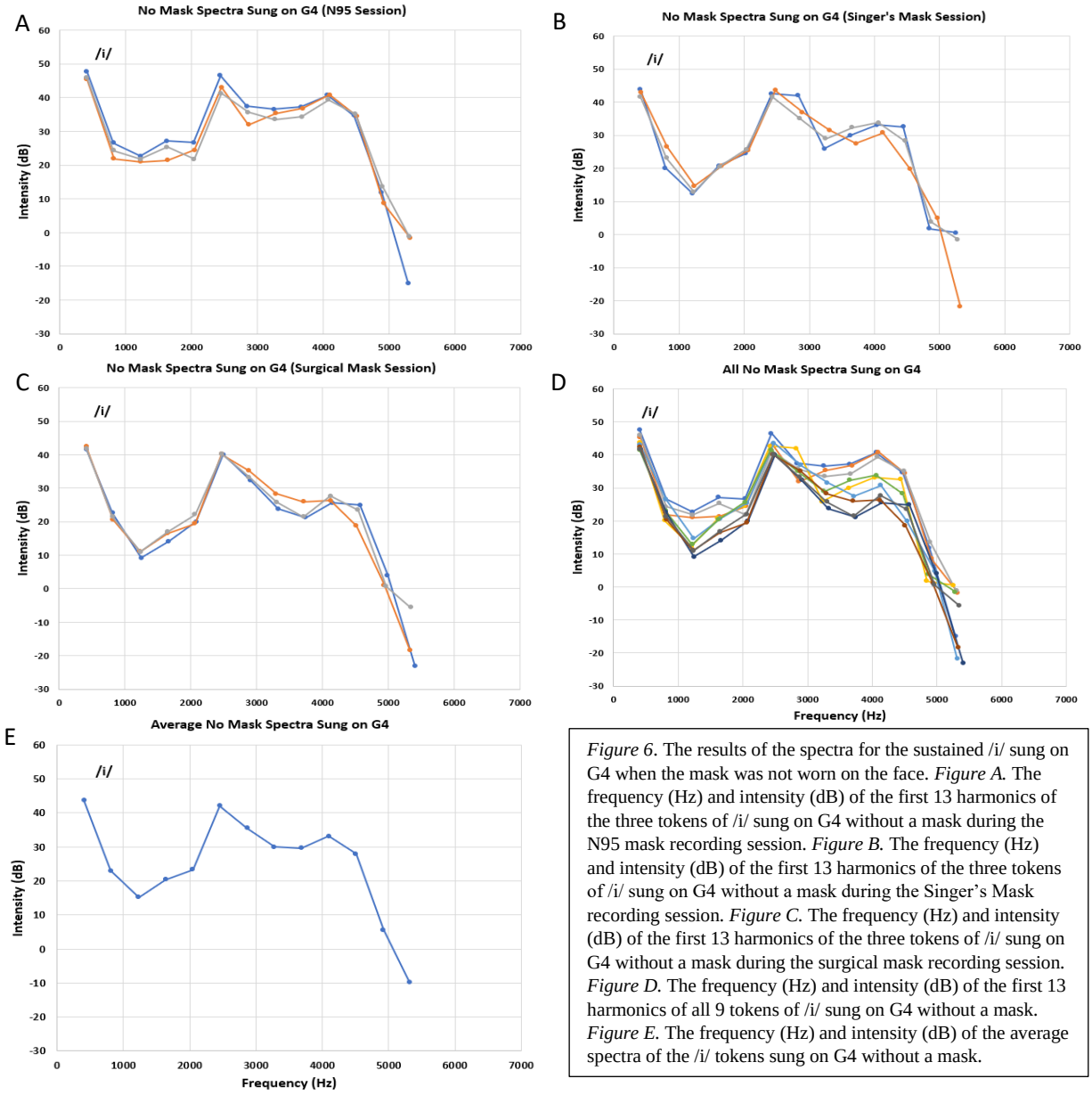

Figure 6. The results of the spectra for the sustained /i/ sung on G4 when the mask was not worn on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of /i/ sung on G4 without a mask during the N95 mask recording session. *Figure B.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of /i/ sung on G4 without a mask during the Singer's Mask recording session. *Figure C.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of /i/ sung on G4 without a mask during the surgical mask recording session. *Figure D.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of all 9 tokens of /i/ sung on G4 without a mask. *Figure E.* The frequency (Hz) and intensity (dB) of the average spectra of the /i/ tokens sung on G4 without a mask.

Figure 7 shows the results of the spectra for the sustained /u/ sung on G4 when the mask was not worn on the face. *Figure A* shows the results of the spectra for the /u/ tokens sung sustained at G4 during the N95 mask recording session when the mask was not worn on the face.

The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o was 4.3 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.3 dB (near 400 Hz or H1) to 10.8 dB (near 5300 Hz or H13). *Figure B* shows the results of the spectra for the /u/ tokens sung sustained at G4 during the Singer's Mask recording session when the mask was not worn on the face. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o was 4.7 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.7 dB (near 2000 Hz or H5) to 11.7 dB (near 5350 Hz or H13). *Figure C* shows the results for the spectra for the /u/ tokens sung sustained at G4 during the surgical mask recording session when the mask was not worn. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o was 5.21 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.6 dB (near 400 Hz or H1) to 12.3 dB (near 5000 Hz or H12). *Figure D* shows the spectra of all tokens of /u/ sung sustained on G4 without a mask. The colored lines depict separate tokens. The maximum frequency difference among the f_o was 10.35 Hz. The difference range in intensity for the corresponding harmonics for the entire spectrum was 5.2 dB (near 2450 Hz or H6) to 20.8 dB (near 5350 Hz or H13). *Figure E* displays the average spectra of all /u/ tokens sung on G4. The average f_o of the tokens was 409.66 Hz and had an intensity of 44.6 dB. The primary peaks indicate the region of the second and third formant frequencies for the /u/ vowel by the single subject.

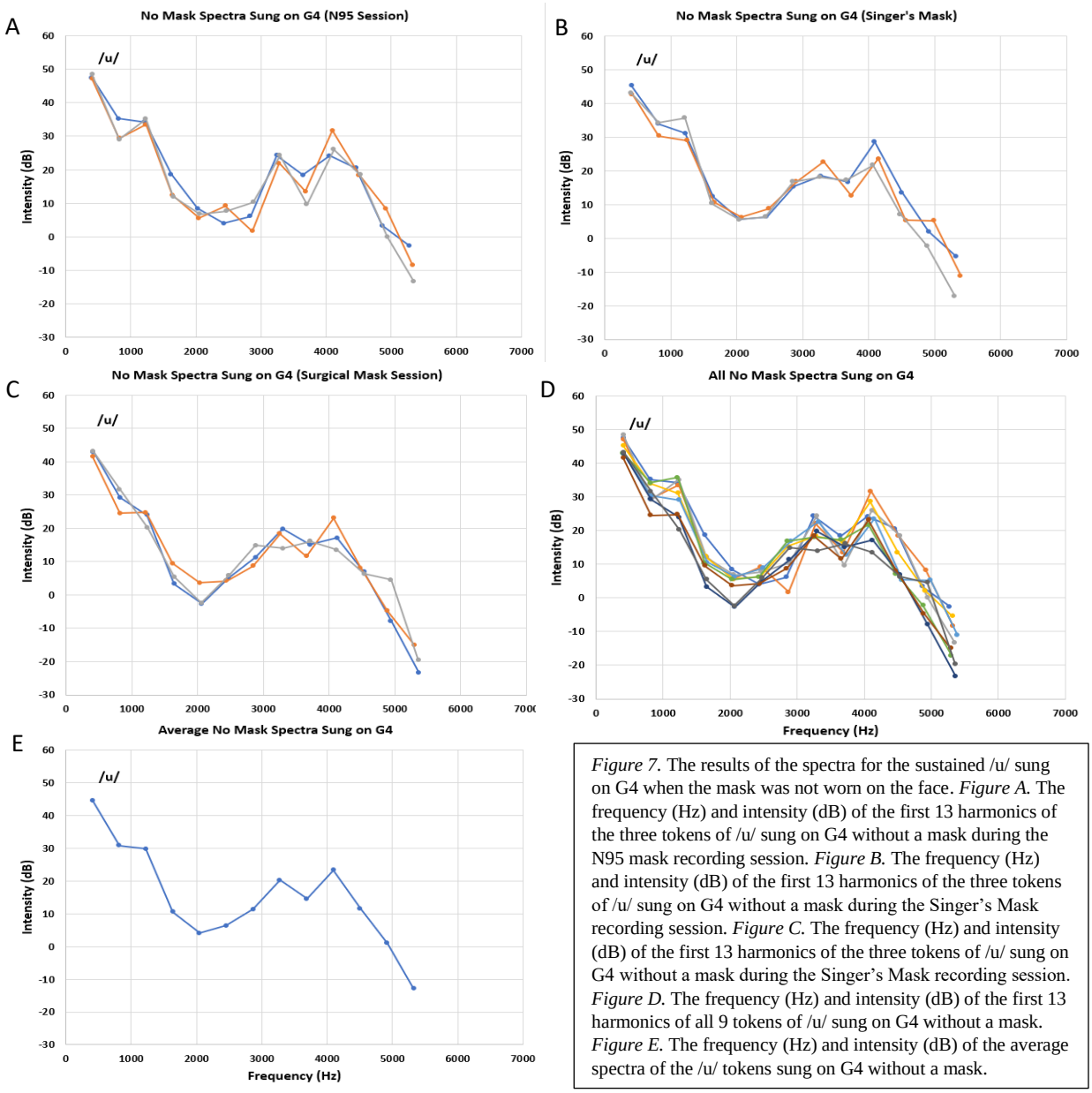

Figure 7. The results of the spectra for the sustained /u/ sung on G4 when the mask was not worn on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of /u/ sung on G4 without a mask during the N95 mask recording session. *Figure B.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of /u/ sung on G4 without a mask during the Singer's Mask recording session. *Figure C.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of /u/ sung on G4 without a mask during the Singer's Mask recording session. *Figure D.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of all 9 tokens of /u/ sung on G4 without a mask. *Figure E.* The frequency (Hz) and intensity (dB) of the average spectra of the /u/ tokens sung on G4 without a mask.

Figure 8 shows the results of the spectra for the sustained /a/ sung on C5 when the mask was not worn on the face**.** *Figure A* shows the results for the spectra for the /a/ tokens sung sustained at C5 during the N95 mask recording session when the mask was not worn on the face. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o 12.56 Hz. The difference range in intensity for the corresponding harmonics for the

entire spectrum was 1.5 dB (near 550 Hz or H1) to 15.8 dB (near 6000 Hz or H11). *Figure B* shows the results of the spectra for the /a/ tokens sung sustained at C5 during the Singer's Mask recording session when the mask was not worn. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o 17.65 Hz. The difference range in intensity for the corresponding harmonics for the entire spectrum was 2.1 dB (near 550 Hz or H1) to 17.8 dB (near 4300 Hz or H8). *Figure C* displays the results of the spectra for the /a/ tokens sung sustained at C5 during the surgical mask recording session when the mask was not worn. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o 11.61 Hz. The difference range in intensity for the corresponding harmonics for the entire spectrum was 1.7 dB (near 550 Hz or H1) to 22.1 dB (near 6600 Hz or H12). *Figure D* displays the spectra of all tokens of /a/ sung sustained on C5 without a mask. The colored lines depict separate tokens. The maximum frequency difference among the f_o was 17.65 Hz. The difference range in intensity for the corresponding harmonics for the entire spectrum was 5.9 dB (near 500 Hz or H1) to 24.1 dB (near 6000 Hz or H11). *Figure E* displays the average spectra of all /a/ tokens sung on C5. The average f_o of the tokens was 541.64 Hz and had an intensity of 42.49 dB. The primary peaks indicate the region of the second and third formant frequencies for the /a/ vowel by the single subject.

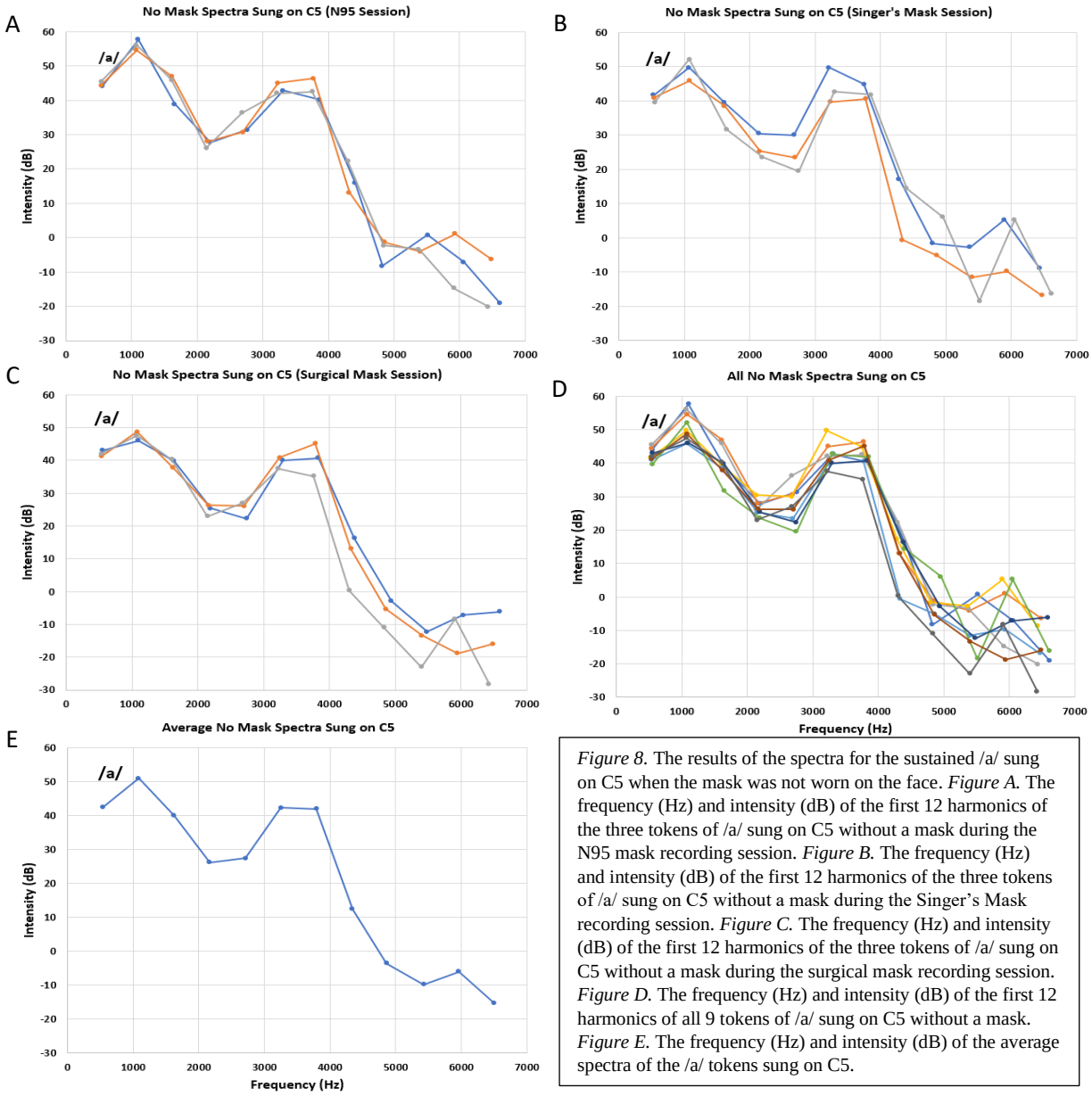

Figure 8. The results of the spectra for the sustained /a/ sung on C5 when the mask was not worn on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 12 harmonics of the three tokens of /a/ sung on C5 without a mask during the N95 mask recording session. *Figure B.* The frequency (Hz) and intensity (dB) of the first 12 harmonics of the three tokens of /a/ sung on C5 without a mask during the Singer's Mask recording session. *Figure C.* The frequency (Hz) and intensity (dB) of the first 12 harmonics of the three tokens of /a/ sung on C5 without a mask during the surgical mask recording session. *Figure D.* The frequency (Hz) and intensity (dB) of the first 12 harmonics of all 9 tokens of /a/ sung on C5 without a mask. *Figure E.* The frequency (Hz) and intensity (dB) of the average spectra of the /a/ tokens sung on C5.

Figure 9 shows the results of the spectra for the sustained /i/ sung on C5 when the mask was not worn on the face**.** *Figure A* shows the results of the spectra for the tokens /i/ sung sustained at C5 during the N95 mask recording session when the mask was not worn on the face. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o was 5.28 Hz. The difference range for the intensity of the corresponding harmonics

for the entire spectrum was 0.9 dB (near 1600 Hz or H3) to 10.7 dB (near 5300 Hz or H10). *Figure B* shows the results of the spectra for the /i/ tokens / sung sustained at C5 during the Singer's Mask recording session when the mask was not worn on the face. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o was 17.32 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 2.9 dB (near 1600 Hz or H3) to 19.5 dB (near 5400 Hz or H10). *Figure C* displays the results for the spectra for the /i/ tokens / sung sustained at C5 during the surgical mask recording session when the mask was not worn on the face. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o was 23.23 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.8 dB (near 1600 Hz or H3) to 25.1 dB (near 6500 Hz or H12). *Figure D* shows the spectra of all tokens of /i/ sung sustained on C5 without a mask. The colored lines depict separate tokens. The maximum frequency difference among the f_o was 23.23 Hz. The difference range in intensity for the corresponding harmonics for the entire spectrum was 8.2 dB (near 550 Hz or H1) to 40.8 dB (near 2700 Hz or H5). *Figure E* displays the average spectra of all /i/ tokens sung on C5. The average f_o of the tokens was 538.31 Hz with an intensity of 50.41 dB. The primary peaks indicate the region of the second and third formant frequencies for the /i/ vowel by the single subject.

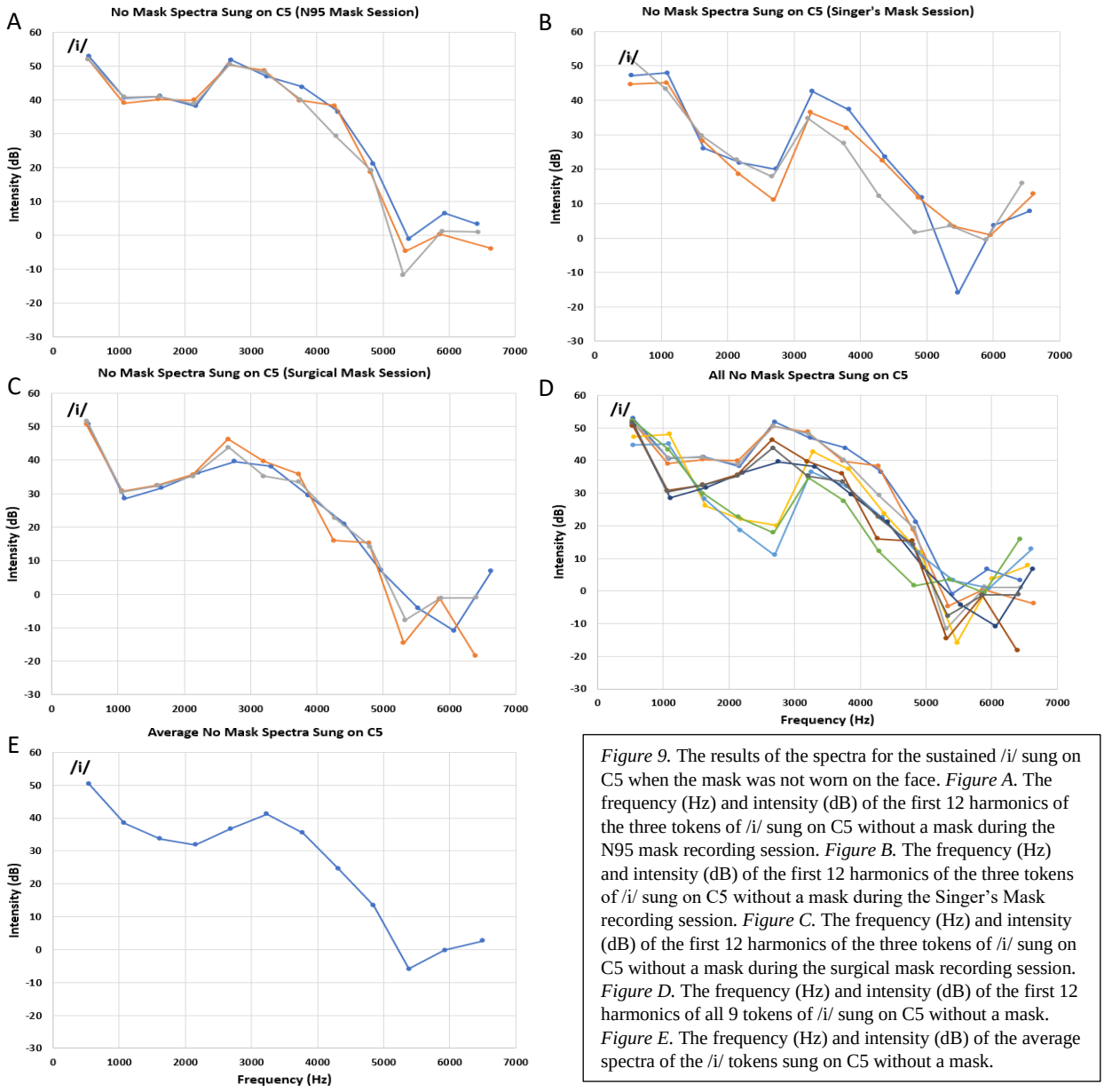

Figure 9. The results of the spectra for the sustained /i/ sung on C5 when the mask was not worn on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 12 harmonics of the three tokens of /i/ sung on C5 without a mask during the N95 mask recording session. *Figure B.* The frequency (Hz) and intensity (dB) of the first 12 harmonics of the three tokens of /i/ sung on C5 without a mask during the Singer's Mask recording session. *Figure C.* The frequency (Hz) and intensity (dB) of the first 12 harmonics of the three tokens of /i/ sung on C5 without a mask during the surgical mask recording session. *Figure D.* The frequency (Hz) and intensity (dB) of the first 12 harmonics of all 9 tokens of /i/ sung on C5 without a mask. *Figure E.* The frequency (Hz) and intensity (dB) of the average spectra of the /i/ tokens sung on C5 without a mask.

Figure 10 shows the results of the spectra for the sustained /u/ sung on C5 when the mask was not worn on the face**.** *Figure A* shows the results of the spectra for the /u/ tokens sung sustained at C5 during the N95 mask recording session when the mask was not worn on the face. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o was 32.2 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.2 dB (near 5400 Hz or H10) to 14.2 dB (near 550 Hz or H1). *Figure B* shows the results for the spectra for the /u/ tokens sung sustained at C5 during the Singer's Mask recording session when the mask was not worn on the face. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o was 14.49

Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.1 dB (near 1100 Hz or H2) to 22.4 dB (near 6500 Hz or H12). *Figure C* shows the results of the spectra for the /u/ tokens sung sustained at C5 during the surgical mask recording session when the mask was not worn on the face. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o was 14.82 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.5 dB (near 550 Hz or H1) to 17.8 dB (near 5900 Hz or H11). *Figure D* shows the spectra of all tokens of /u/ sung sustained on C5 without a mask. The colored lines depict separate tokens. The maximum frequency difference among the f_o was 45.04 Hz. The difference range in intensity for the corresponding harmonics for the entire spectrum was 10.8 dB (near 1600 Hz or H3) to 29.8 dB (near 4400 Hz or H10). *Figure E* displays the average spectra of all /u/ tokens sung on C5. The average f_o of the tokens was 538.64 Hz and had an intensity of 49.06 dB. The primary peaks indicate the region of the second and third formant frequencies for the sung /u/ vowel by the single subject.

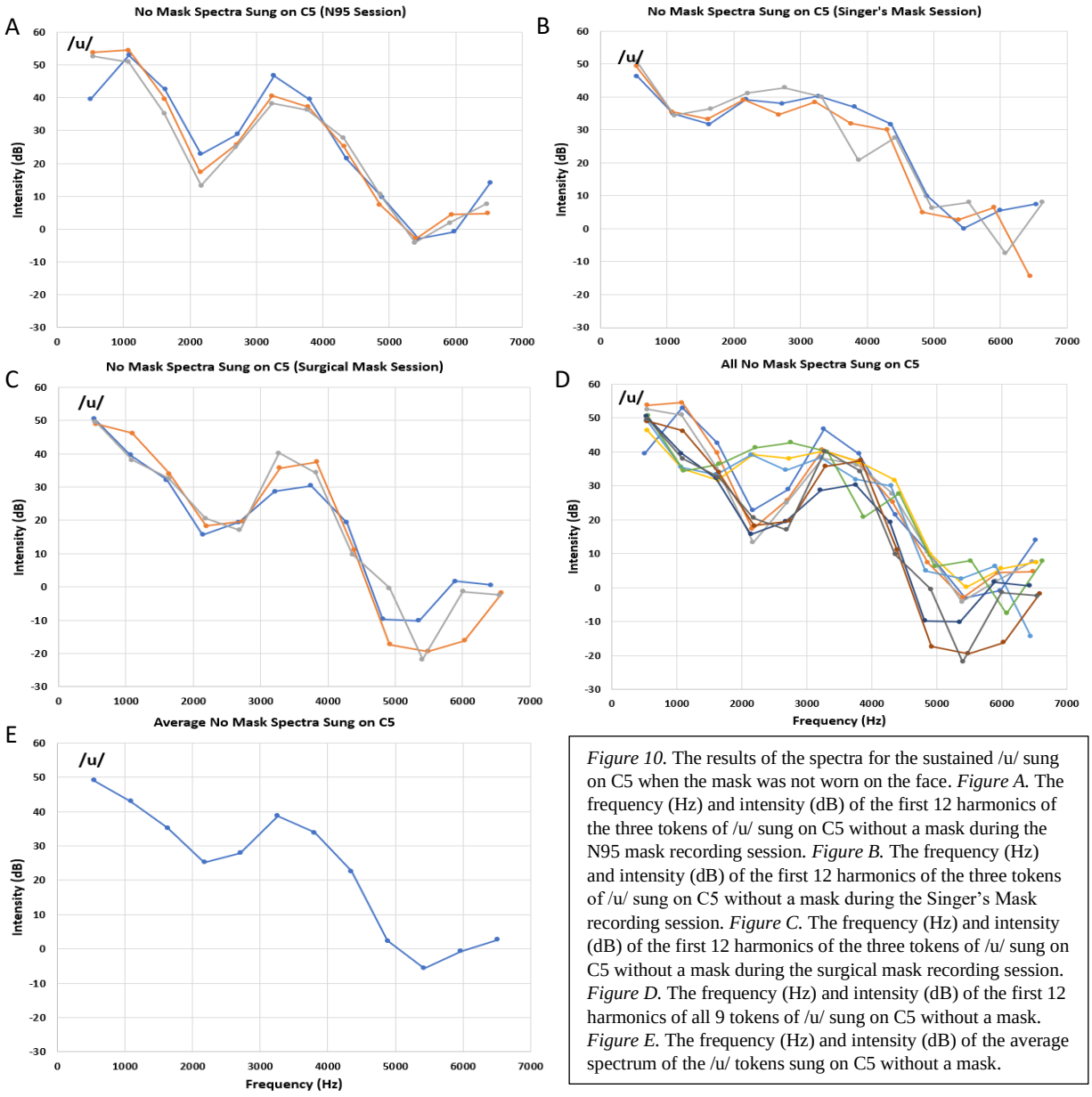

Figure 10. The results of the spectra for the sustained /u/ sung on C5 when the mask was not worn on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 12 harmonics of the three tokens of /u/ sung on C5 without a mask during the N95 mask recording session. *Figure B.* The frequency (Hz) and intensity (dB) of the first 12 harmonics of the three tokens of /u/ sung on C5 without a mask during the Singer's Mask recording session. *Figure C.* The frequency (Hz) and intensity (dB) of the first 12 harmonics of the three tokens of /u/ sung on C5 without a mask during the surgical mask recording session. *Figure D.* The frequency (Hz) and intensity (dB) of the first 12 harmonics of all 9 tokens of /u/ sung on C5 without a mask. *Figure E.* The frequency (Hz) and intensity (dB) of the average spectrum of the /u/ tokens sung on C5 without a mask.

N95 Mask

Vowels Spoken /a,i,u/

Figure 11 shows the results and comparisons of spoken /a/ with the N95 mask on the face. *Figure A* displays the results of the spectra for the spoken sustained /a/ tokens with the N95 mask on the face. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o was 2.75 Hz. The difference range in intensity of the corresponding harmonics for the entire spectrum was 0.5 dB (near 700 Hz or near the second

harmonic, H2) to 18.8 dB (near 2150 Hz or near the sixth harmonic, H6). *Figure B* displays the average spectra of spoken sustained /a/ with the N95 mask and without the N95 mask on the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 1.1 Hz, with the no mask f_o at 359.49 Hz and the mask f_o at 358.39 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.8 dB (near 350 Hz or H1) to 14.07 dB (near 3950 Hz or H11). *Figure C* displays the difference between the no mask average spectrum and the mask average spectrum in intensity (dB) for a spoken sustained /a/. The intensity difference of the f_o was 1.06 dB, with the mask spectrum higher than the no mask spectrum. The figure suggests that there is a gradual increase in the intensity difference among the harmonics since the regression line has a positive slope. However, the R2 term is very small (R2 = 0.1204), and thus this general relationship does not hold. However, there are three locations where the difference between the spectra indicates 5 or more dB, at H4, H8, and H11. H4 is near the second formant F2 and H11 is near the third or fourth formant (F3 or F4), and thus these two differences may create a perceptual quality change between the no mask and masked conditions.

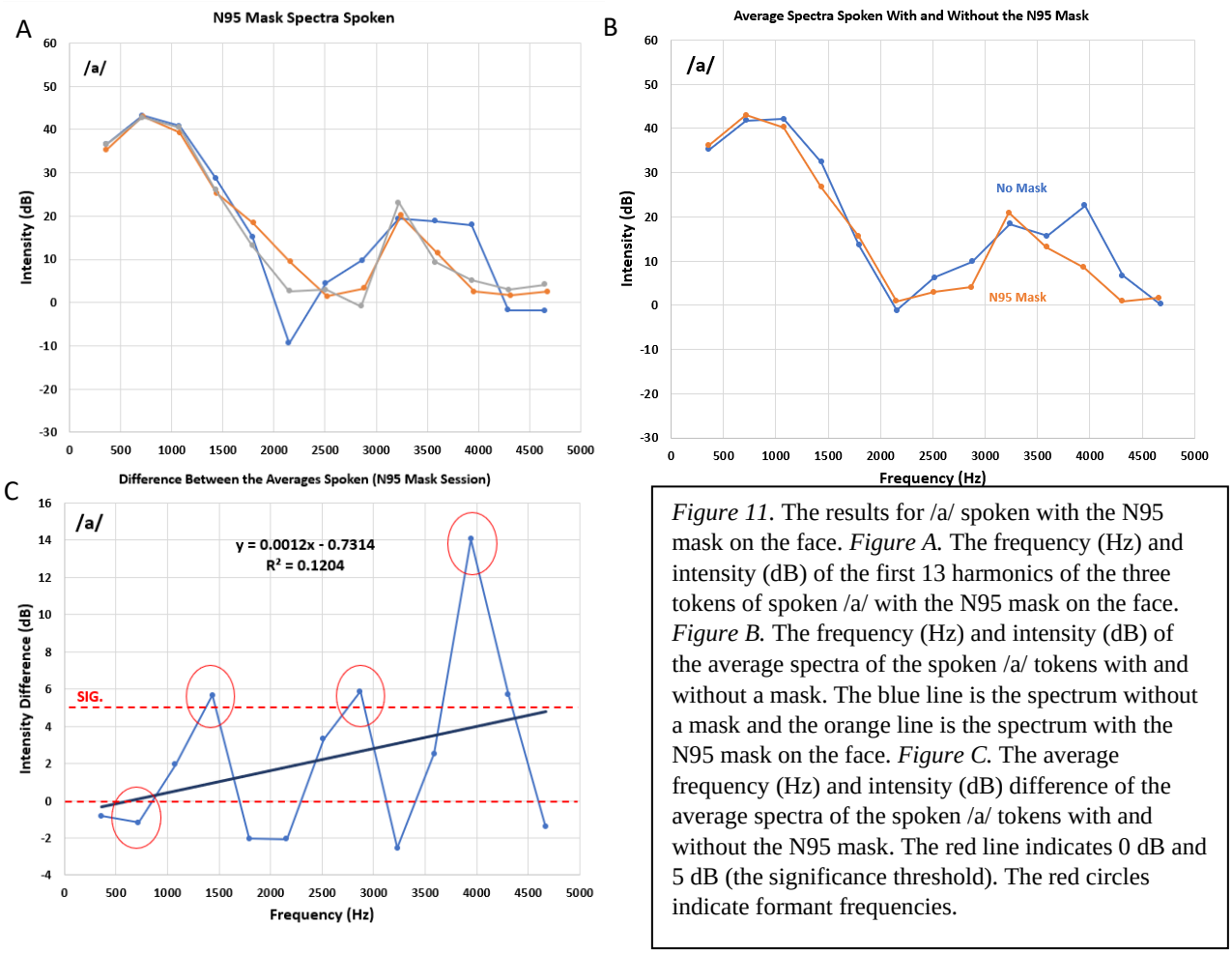

Figure 11. The results for /a/ spoken with the N95 mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of spoken /a/ with the N95 mask on the face. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the spoken /a/ tokens with and without a mask. The blue line is the spectrum without a mask and the orange line is the spectrum with the N95 mask on the face. *Figure C.* The average frequency (Hz) and intensity (dB) difference of the average spectra of the spoken /a/ tokens with and without the N95 mask. The red line indicates 0 dB and 5 dB (the significance threshold). The red circles indicate formant frequencies.

Figure 12 shows the results and comparisons of spoken /i/ with the N95 mask on the face. *Figure A* shows the results of the spectra for the spoken sustained /i/ tokens with the N95 mask on the face. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o was 1.8 Hz with the average f_o at 360.67 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.6 dB (near 350 Hz or H1) to 12.4 dB (near 1400 Hz or H4). *Figure B* displays the average spectra of spoken sustained /i/ with the mask and without the N95 mask on the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 0.07 Hz, with the no mask f_o at 360.44 Hz and the mask f_o 360.37 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.13 dB (near 350 Hz or H1) to 11.67 dB (near 4300 Hz or H13). *Figure C* displays the difference between the no mask average spectrum and the mask average spectrum in intensity (dB) for a spoken sustained /i/. The difference of the f_o was 0.13 dB, with the mask spectrum higher than the no mask spectrum. The figure suggests that there is a gradual increase in the intensity difference among the harmonics since the regression line has a positive slope of 0.0025. However, the R2 term is small (R2 = 0.403), and thus this general relationship does not hold. However, there are several locations where the difference between the spectra indicates a difference of 5 or more dB, specifically at H7, H8, H9, H10, and H11. These harmonics are near the third and fourth formant (F3 and F4), thus these two differences may create a perceptual quality change between the no mask and masked conditions.

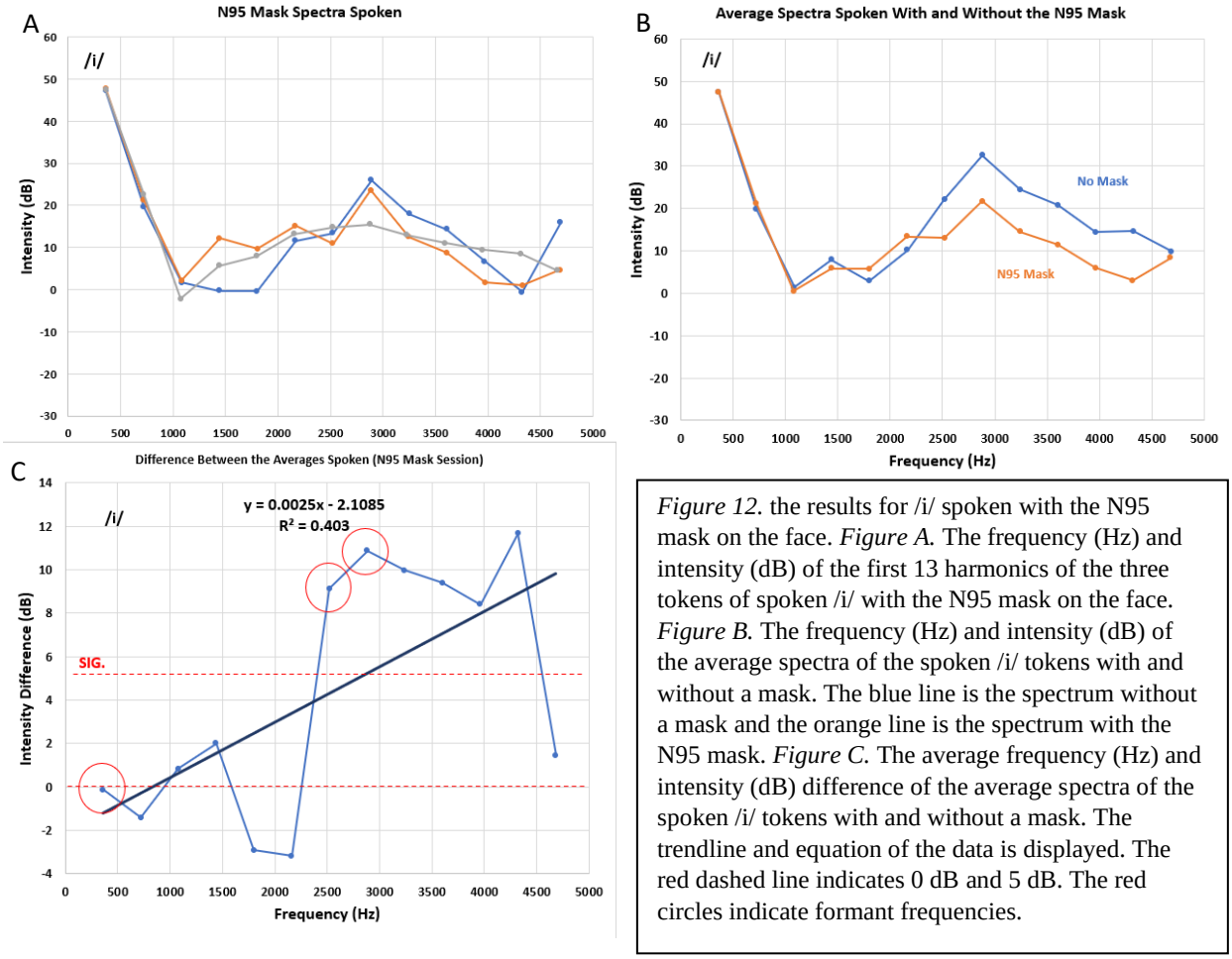

Figure 12. the results for /i/ spoken with the N95 mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of spoken /i/ with the N95 mask on the face. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the spoken /i/ tokens with and without a mask. The blue line is the spectrum without a mask and the orange line is the spectrum with the N95 mask. *Figure C.* The average frequency (Hz) and intensity (dB) difference of the average spectra of the spoken /i/ tokens with and without a mask. The trendline and equation of the data is displayed. The red dashed line indicates 0 dB and 5 dB. The red circles indicate formant frequencies.

Figure 13 shows the results and comparisons of spoken /u/ with the N95 mask on the face. *Figure A* shows the results of the spectra for the spoken sustained /u/ tokens with the N95 mask on the face. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 361.1 Hz and the maximum frequency difference among the f_o was 4.24 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.4 dB (near 350 Hz or H1) to 10.5 dB (near 2600 Hz or H7). *Figure B* displays the average spectra of spoken sustained /u/ with the N95 mask and without a mask on the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 0.07 Hz, with the no mask f_o at 361.03 Hz and the mask f_o at 361.1 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.3 dB (near 700 Hz or H2) to 7.77 dB (near 4350 Hz or H11). *Figure C* displays the difference in between the no mask average spectrum and the mask average spectrum in intensity (dB) for a spoken sustained /u/. The difference of the f_o was 0.07 dB, with the no mask spectrum higher than the mask spectrum. The figure suggests that there is a gradual increase in the intensity difference among the harmonics since the regression line has a positive slope. However, the R2 term is small (R2 = 0.1321), and

thus this general relationship does not hold. However, there are locations where the difference between the spectra indicates a difference of 5 or more dB, specifically at H3 and H12. These harmonics are near F2 and F4, thus the difference in intensity may create a perceptual quality change between the no mask and masked conditions.

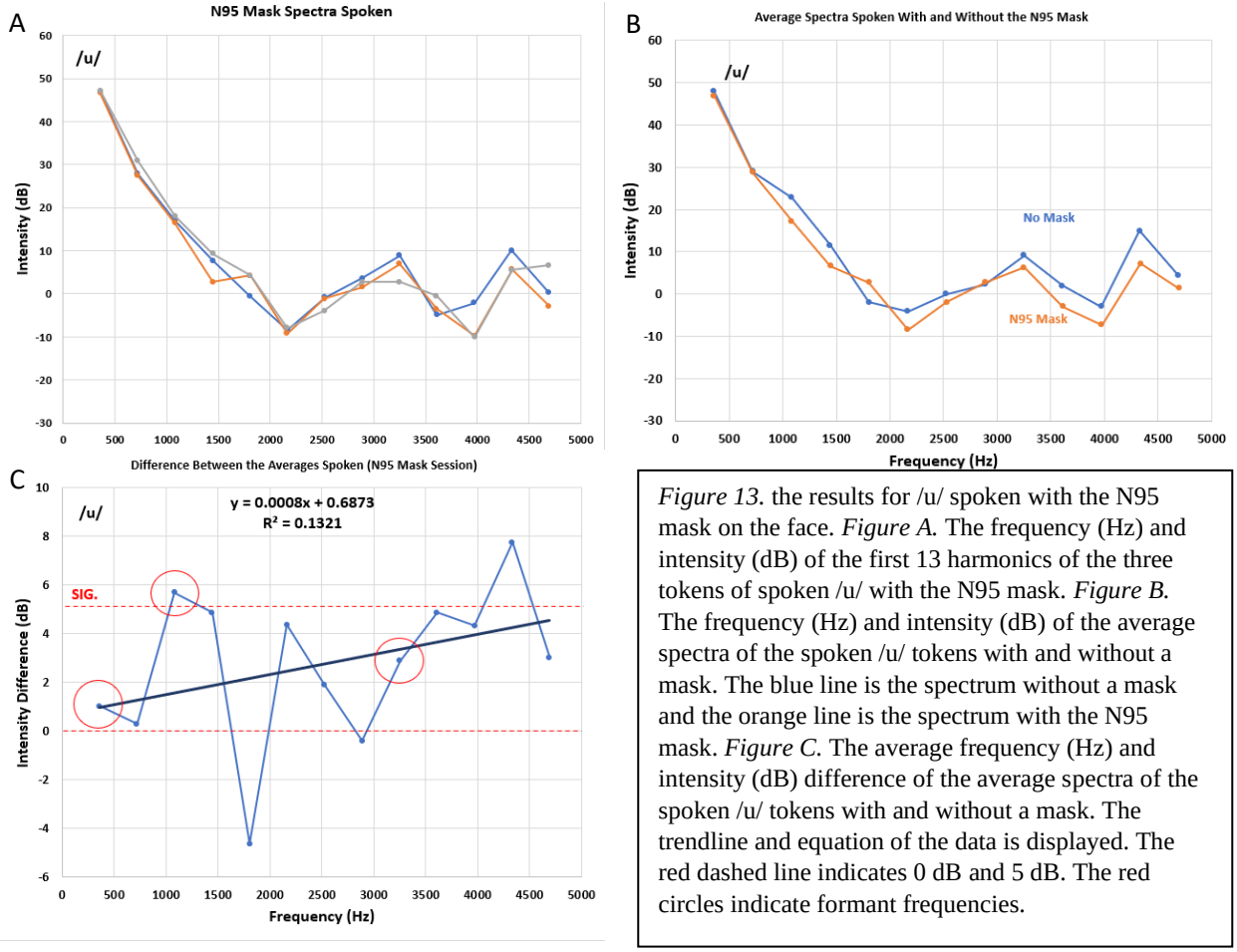

Figure 13. the results for /u/ spoken with the N95 mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of spoken /u/ with the N95 mask. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the spoken /u/ tokens with and without a mask. The blue line is the spectrum without a mask and the orange line is the spectrum with the N95 mask. *Figure C.* The average frequency (Hz) and intensity (dB) difference of the average spectra of the spoken /u/ tokens with and without a mask. The trendline and equation of the data is displayed. The red dashed line indicates 0 dB and 5 dB. The red circles indicate formant frequencies.

Vowels on G4 /a,i,u/

Figure 14 shows the results and comparisons of /a/ sung on G4 with the N95 mask on the face. *Figure A* shows the results for the spectra for the /a/ tokens sung sustained on G4 with the N95 mask on the face. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o was 2.75 Hz with an average f_o of 409.84 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.7 dB (near 400 Hz or H1 and 800 Hz or H2) to 12.9 dB (near 3700 Hz or H9). *Figure B* displays the average spectra of /a/ sung on G4 with the N95 mask and without the mask. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 7.12 Hz, with the no mask f_o at 416.96 Hz and the mask f_o at 409.84 Hz. The difference range for the intensity of the corresponding harmonics

for the entire spectrum was 0.43 dB (near 2500 Hz or H6) to 10.33 dB (near 4550 Hz or H11). *Figure C* displays the difference between the no mask average spectrum and the N95 mask average spectrum in intensity (dB) for sustained /a/ sung on G4. The difference of the f_o was dB, with the no mask spectrum higher than the N95 mask spectrum. The figure suggests that there is a gradual increase in the intensity difference among the harmonics since the regression line has a positive slope. However, the R2 term is small (R2 = 0.1406), and thus this general relationship does not hold. However, there are several locations where the difference between the spectra indicates a difference of 5 or more dB, specifically at H3, H11, and H13. These harmonics are close to F2 and F3, thus the difference in intensity may create a perceptual quality change between the no mask and masked conditions.

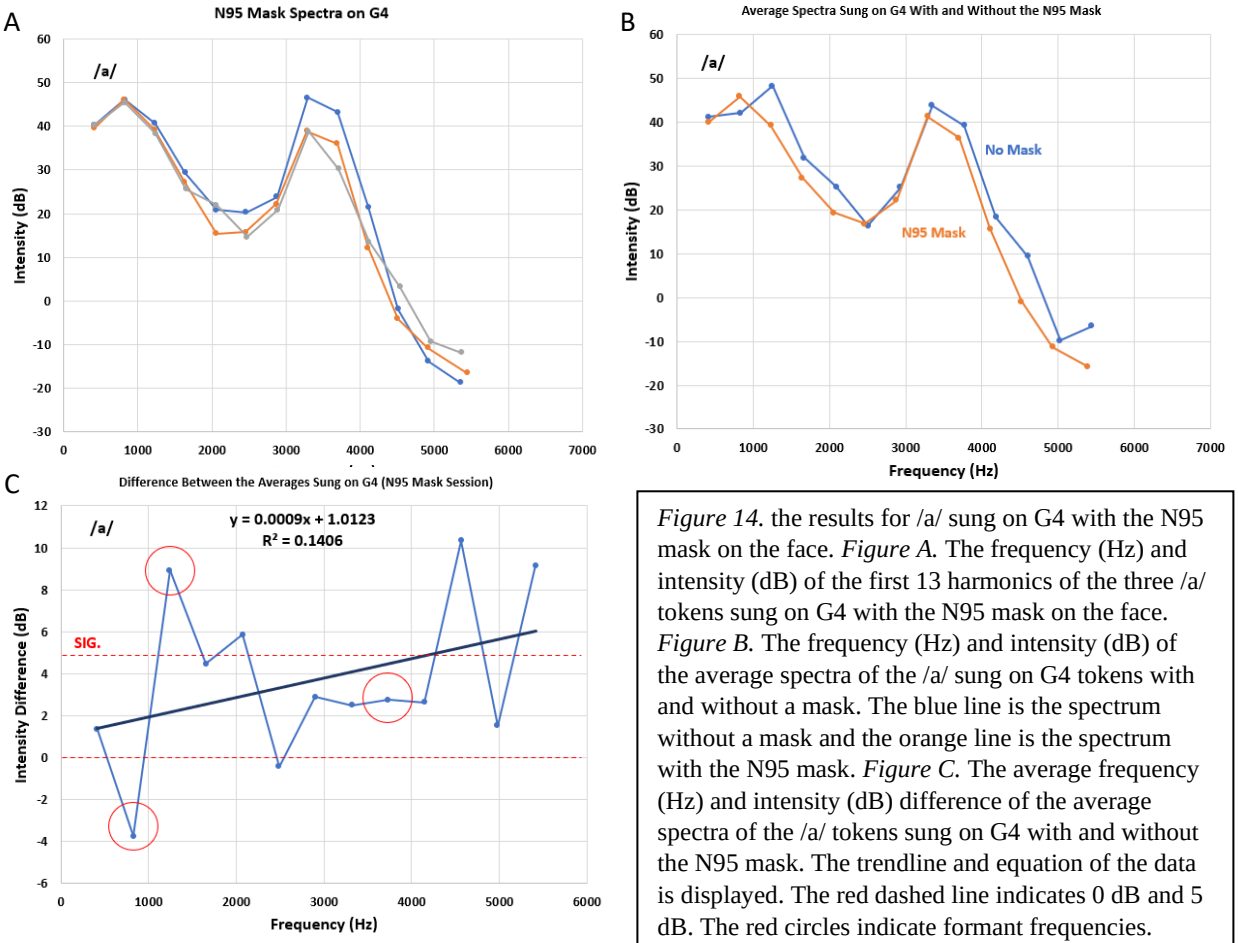

Figure 14. the results for /a/ sung on G4 with the N95 mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three /a/ tokens sung on G4 with the N95 mask on the face. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the /a/ sung on G4 tokens with and without a mask. The blue line is the spectrum without a mask and the orange line is the spectrum with the N95 mask. *Figure C.* The average frequency (Hz) and intensity (dB) difference of the average spectra of the /a/ tokens sung on G4 with and without the N95 mask. The trendline and equation of the data is displayed. The red dashed line indicates 0 dB and 5 dB. The red circles indicate formant frequencies.

Figure 15 shows the results and comparisons of /a/ sung on G4 with the N95 mask on the face. *Figure A* shows the results of the spectra of the tokens for /i/ sung on G4 with the N95 mask on the face. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o was 2.46 Hz, with the average f_o at 409.15 Hz. The difference

range for the intensity of the corresponding harmonics for the entire spectrum was 1.1 dB (near 1650 Hz or H4) to 10.3 dB (near 2850 Hz or H7). *Figure B* displays the average spectra of sustained /i/ sung on G4 with the N95 mask and without the mask on the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 3.81 Hz, with the no mask f_o at 408.69 Hz and the mask f_o at 409.15 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.87 dB (near 400 Hz or H1) to 5.4 dB (near 4500 Hz or H11). *Figure C* displays the difference between the no mask average spectrum and the mask average spectrum in intensity (dB) for /i/ sung on G4. The difference of the f_o was 1.87 dB, with the mask spectrum higher than the no mask spectrum. The figure suggests that there is a gradual increase in the intensity difference among the harmonics since the regression line has a positive slope. However, the R2 term is small (R2 = 0.1753), and thus this general relationship does not hold. However, there is a location where the difference between the spectra indicates a difference of 5.4 dB at H11. This harmonic is near F4, thus the difference in intensity may create a perceptual quality change between the no mask and masked conditions.

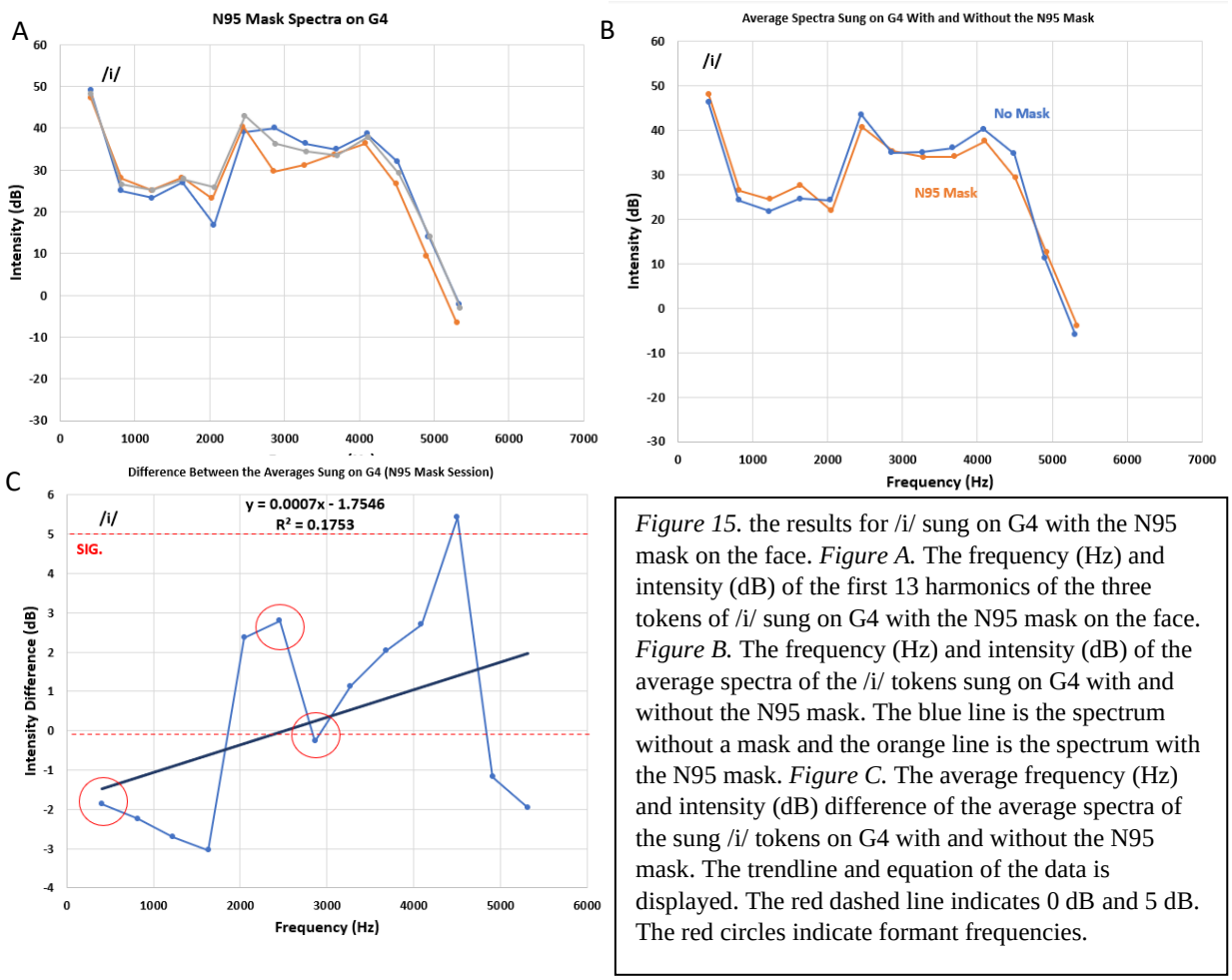

Figure 15. the results for /i/ sung on G4 with the N95 mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of /i/ sung on G4 with the N95 mask on the face. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the /i/ tokens sung on G4 with and without the N95 mask. The blue line is the spectrum without a mask and the orange line is the spectrum with the N95 mask. *Figure C.* The average frequency (Hz) and intensity (dB) difference of the average spectra of the sung /i/ tokens on G4 with and without the N95 mask. The trendline and equation of the data is displayed. The red dashed line indicates 0 dB and 5 dB. The red circles indicate formant frequencies.

Figure 16 shows the results and comparisons of /a/ sung on G4 with the N95 mask on the face. *Figure A* shows the results of the spectra of the /u/ tokens sustained on G4 with the N95 mask on the face. The blue, orange, and gray lines depict separate tokens. The maximum frequency difference among the f_o was 2.95 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.6 dB (near 4500 Hz or H11) to 15.5 dB (near 2850 Hz or H7). *Figure B* displays the average spectra of sustained /u/ sung on G4 with the N95 mask and without the mask. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 1.61 Hz with the no mask f_o at 409.18 Hz and the mask f_o at 407.57 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.73 dB (near 400 Hz or H1) to 6.0 dB (near 5000 Hz or H12). *Figure C* displays the difference between the no mask average spectrum and the N95 mask average spectrum in intensity (dB) for sustained /u/ sung on G4. The difference of the f_o was 0.73 dB, with the no mask spectrum higher than the N95 mask spectrum. The figure suggests that there is a gradual increase in the intensity difference among the harmonics since the regression line has a positive slope. While the R^2 term is large ($R^2 = 0.5719$), the locations of significant spectra differences are not close to the formants, so the intensity differences displayed may not attribute to perceptual changes in the vowel. F2 has a negative difference, meaning the mask spectrum was higher than the no mask spectrum. This may indicate that the N95 mask had created a resonance that increased the intensity of this frequency.

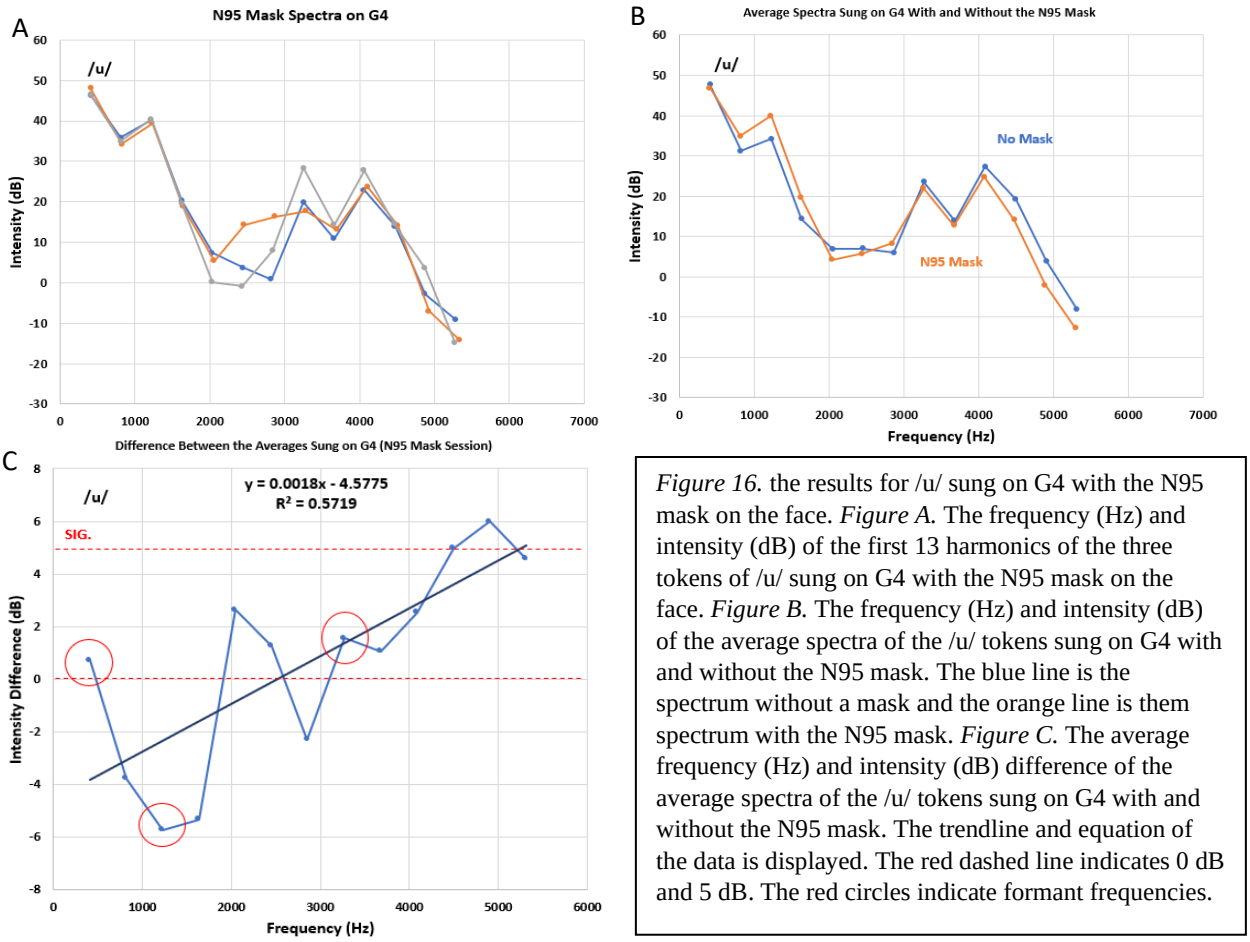

Figure 16. the results for /u/ sung on G4 with the N95 mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of /u/ sung on G4 with the N95 mask on the face. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the /u/ tokens sung on G4 with and without the N95 mask. The blue line is the spectrum without a mask and the orange line is them spectrum with the N95 mask. *Figure C.* The average frequency (Hz) and intensity (dB) difference of the average spectra of the /u/ tokens sung on G4 with and without the N95 mask. The trendline and equation of the data is displayed. The red dashed line indicates 0 dB and 5 dB. The red circles indicate formant frequencies.

Vowels on C5 /a,i,u/

Figure 17 shows the results and comparisons of /a/ sung on C5 with the N95 mask on the face. *Figure A* displays the results of the spectra of the /a/ tokens sung sustained on C5 with the N95 mask on the face. The blue, orange, and gray lines depict separate tokens. The average f_o was 545.70 Hz with a difference of 1.66 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.2 dB (near 1100 Hz or H2) to 17.8 dB (near 4900 Hz or H9). *Figure B* displays the average spectra of sustained /a/ sung on C5 with the N95 mask and without the mask on the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 1.06 Hz, with the no mask f_o at 541.62 Hz and the mask f_o at 545.70 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.43 dB (near 4900 Hz or H9) to 8.17 dB (near 5500 Hz or H10). *Figure C* displays the difference between the no mask average spectrum and the N95 mask average spectrum in intensity (dB) for sustained /a/ sung on C5. The difference of the f_o was 0.9 dB, with the N95 mask spectrum higher than the no mask spectrum. The figure suggests that there is a gradual increase in the intensity difference among the harmonics since the regression line has a positive slope. However, the R2 term is small (R2 = 0.1352), and thus this general relationship does not

hold. However, there are locations where the difference between the spectra indicates a difference of 5 or more dB, specifically at H4 and H10. H4 is near F2, thus it may cause perceptual changes heard while wearing the mask on the face. While H10 has a high attenuation, it is negligible because the intensity is about 50 dB lower than the formant frequencies.

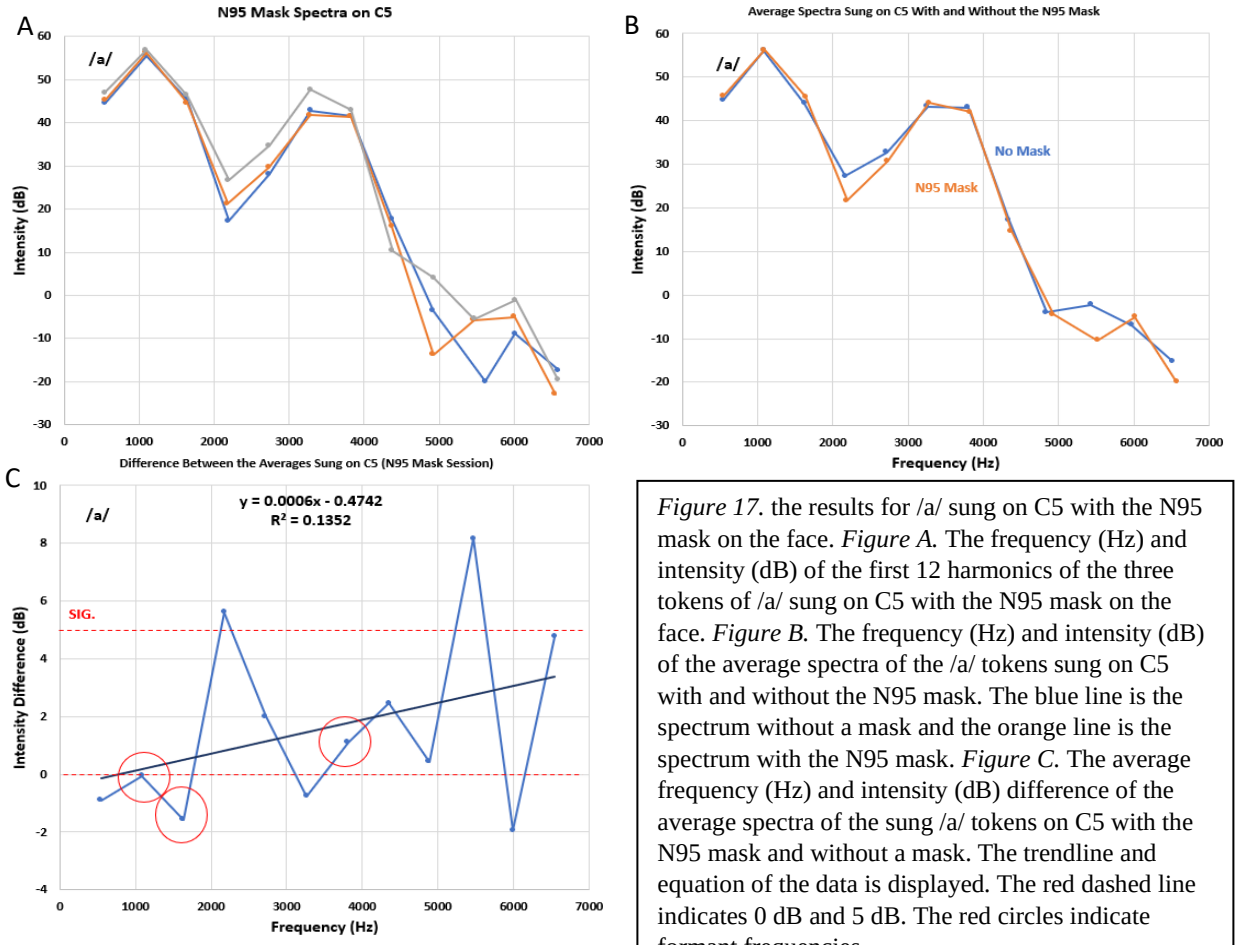

Figure 17. the results for /a/ sung on C5 with the N95 mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 12 harmonics of the three tokens of /a/ sung on C5 with the N95 mask on the face. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the /a/ tokens sung on C5 with and without the N95 mask. The blue line is the spectrum without a mask and the orange line is the spectrum with the N95 mask. *Figure C.* The average frequency (Hz) and intensity (dB) difference of the average spectra of the sung /a/ tokens on C5 with the N95 mask and without a mask. The trendline and equation of the data is displayed. The red dashed line indicates 0 dB and 5 dB. The red circles indicate formant frequencies.

Figure 18 shows the results and comparisons of /i/ sung on C5 with the N95 mask on the face. *Figure A* shows the results for the spectra of the tokens for sustained /i/ sung on C5 with the N95 mask on the face. The blue, orange, and gray lines depict separate tokens. The average f_o was 538.88 Hz with a difference of 10.38 Hz among the tokens. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.0 dB (near 1050 Hz or H2) to 14.7 dB (near 5900 Hz or H12). *Figure B* displays the average spectra of sustained /i/ sung on C5 with the N95 mask and without the mask on the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 2.96 Hz, with the no mask f_o at 535.92 Hz and the mask f_o at 538.88 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.06 dB (near 540 Hz or H1) to 11.77 dB (near 4850 Hz or H9). *Figure C* displays the difference between the no mask average spectrum and the N95 mask

average spectrum in intensity (dB) for sustained /i/ sung on C5. The difference of the f_o was 0.067 dB, with the N95 mask spectrum higher than the no mask spectrum. The figure suggests that there is a gradual increase in the intensity difference among the harmonics since the regression line has a positive slope. However, the R^2 term is small ($R^2 = 0.1059$), and thus this general relationship does not hold. While there are locations where the difference between the spectra indicates a difference of 5 or more dB, specifically at H4, H9, and H11, the differences are negligible because they are about 50 dB lower than the formant frequencies, thus not relevant to the perceptual sound heard.

Figure 18. the results for /i/ sung on C5 with the N95 mask on the face. Figure A. The frequency (Hz) and intensity (dB) of the first 12 harmonics of the three tokens of sung /i/ on C5 with the N95 mask. Figure B. The frequency (Hz) and intensity (dB) of the average spectra of the /i/ tokens sung on C5 with the N95 mask and without the mask. The blue line is the spectrum without a mask and the orange line is the spectrum with the N95 mask. Figure C. The average frequency (Hz) and intensity (dB) difference of the average spectra of the /i/ tokens sung on C5 with the N95 mask and without a mask. The trendline and equation of the data is displayed. The red dashed line indicates 0 dB and 5 dB. The red circles indicate formant frequencies.

Figure 19 shows the results and comparisons of /u/ sung on C5 with the N95 mask on the face. *Figure A* shows the results of the spectra of the sustained /u/ tokens sung on C5 with the N95 mask on the face. The blue, orange, and gray lines depict separate tokens. The average f_o was 542.63 Hz with a difference of 7.43 Hz among the tokens. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.4 dB (near 550 Hz or H1) to 23.4 dB (near 6000 Hz or H11). *Figure B* displays the average spectra of sustained /u/ sung on C5 with the N95 mask and without the mask. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency

difference among the f_o was 1.06 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.1 dB (near 1600 Hz or H3) to 9.73 dB (near 5400 Hz or H10). *Figure C* displays the difference between the no mask average spectrum and the N95 mask average spectrum in intensity (dB) for sustained /u/ sung on C5. The difference of the f_o was 3.43 dB, with the N95 mask spectrum higher than the no mask spectrum. The figure suggests that there is a gradual increase in the intensity difference among the harmonics since the regression line has a positive slope. The R^2 term is relatively larger ($R^2 = 0.6249$), so the positive relationship holds. While there are locations where the difference between the spectra indicates a difference of 5 or more dB, specifically at H8 and H10, these areas are not close to formants and negligible due to being around 30 dB to 50 dB less than the formant frequencies. However, the negative differences at the beginning of the spectrum are near F1 and F2, so the mask may be enhancing frequencies, thus creating a perceptual difference.

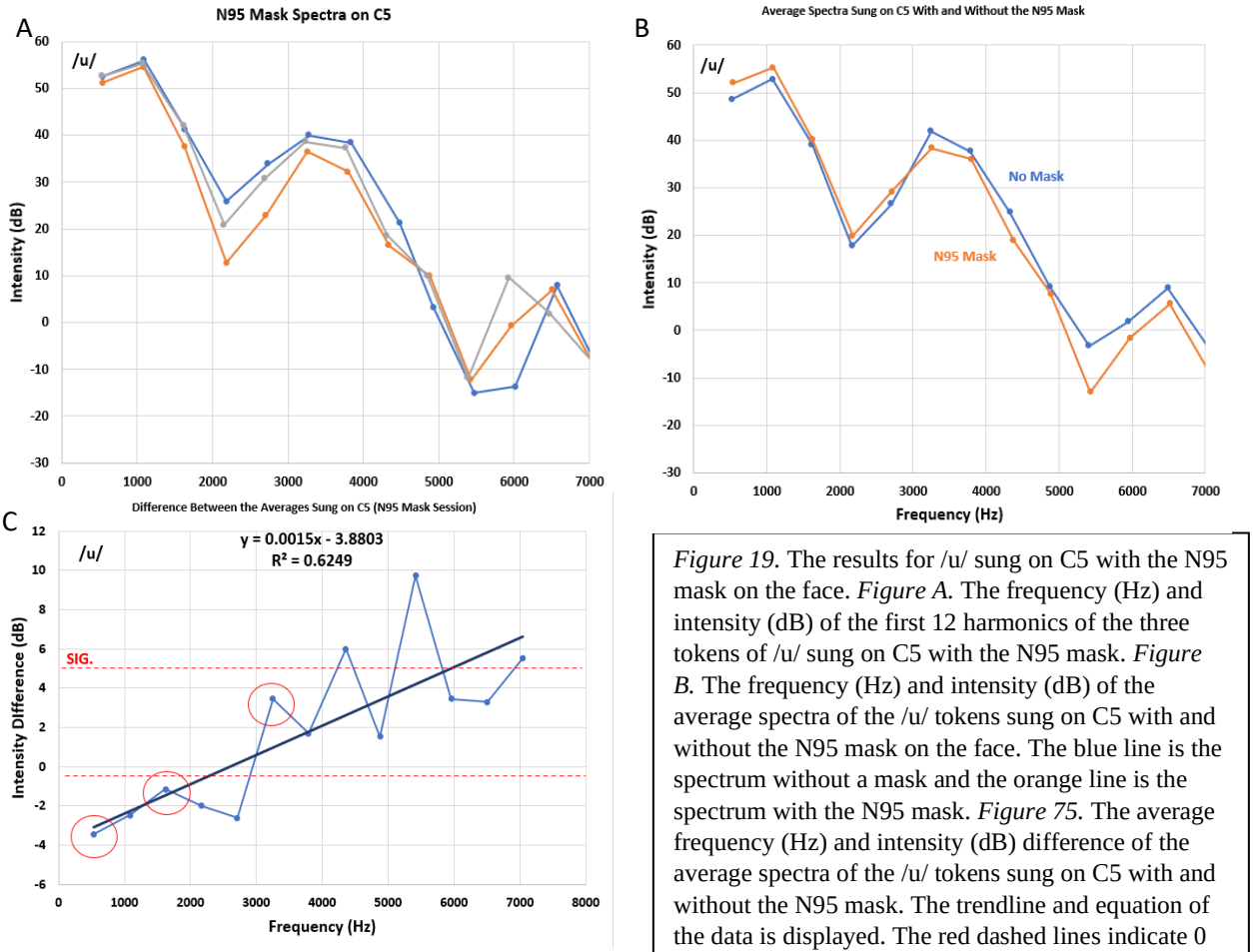

Figure 19. The results for /u/ sung on C5 with the N95 mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 12 harmonics of the three tokens of /u/ sung on C5 with the N95 mask. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the /u/ tokens sung on C5 with and without the N95 mask on the face. The blue line is the spectrum without a mask and the orange line is the spectrum with the N95 mask. *Figure 75.* The average frequency (Hz) and intensity (dB) difference of the average spectra of the /u/ tokens sung on C5 with and without the N95 mask. The trendline and equation of the data is displayed. The red dashed lines indicate 0 dB and 5 dB. The red circles indicate formant frequencies.

Long Term Average Spectra (LTAS) /s, ʃ/

Figure 20 shows the results and LTAS comparisons of /s/ with the N95 mask worn on the face. *Figure A* displays the trial 1 LTAS for /s/ without with the N95 mask and without the

mask on the face. The blue line indicates the LTAS without a mask and the red line indicates the LTAS with the N95 mask on the face. The main difference in spectra lies in the salient frequency location for an /s/ production, between 5000 Hz and 7000 Hz. *Figure B* displays the trial 2 LTAS for /s/ with the N95 mask and without the mask on the face. The blue line indicates the LTAS without a mask and the red line indicates the LTAS with the N95 mask on the face. Again, the main difference in spectra lies in the salient frequency location for an /s/ production, between 5000 Hz and 7000 Hz *Figure C* displays the intensity difference between the no mask LTAS and N95 mask LTAS for trial 1 at every 1000 Hz (blue line) and for trial 2 at every 1000 Hz (orange line). The maximum difference for the entire spectrum was 16.2 dB at 6000 Hz for trial 1 and 13.4 dB at 6000 Hz for trial 2. The salient location for /s/ is located around 6000 Hz, thus the highest attenuation at this area may create a perceptual quality change between the no mask and masked conditions. That is, the perception of /s/ may be greatly reduced during speech due to being about 15 dB reduced when the mask was worn.

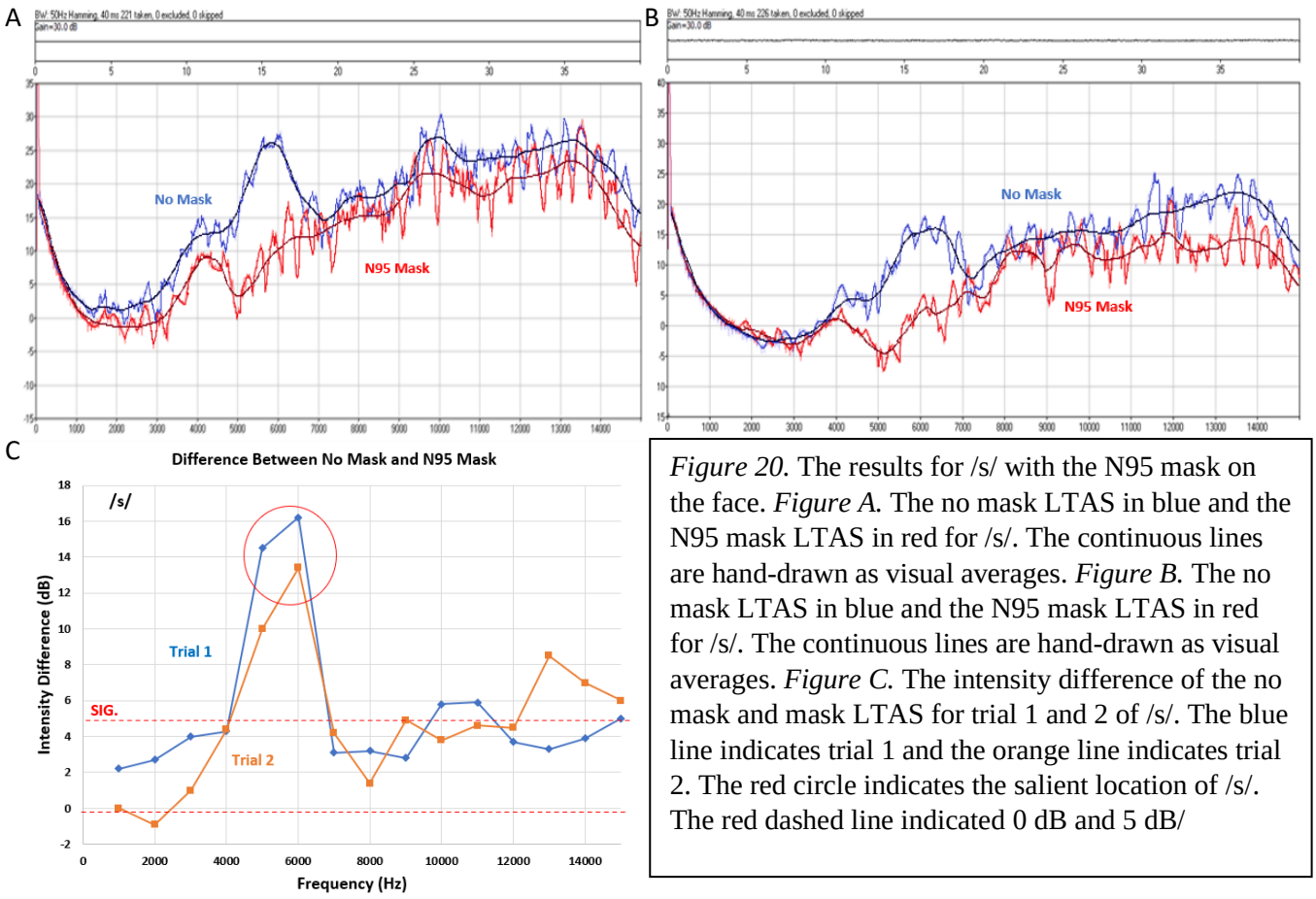

Figure 20. The results for /s/ with the N95 mask on the face. *Figure A.* The no mask LTAS in blue and the N95 mask LTAS in red for /s/. The continuous lines are hand-drawn as visual averages. *Figure B.* The no mask LTAS in blue and the N95 mask LTAS in red for /s/. The continuous lines are hand-drawn as visual averages. *Figure C.* The intensity difference of the no mask and mask LTAS for trial 1 and 2 of /s/. The blue line indicates trial 1 and the orange line indicates trial 2. The red circle indicates the salient location of /s/. The red dashed line indicated 0 dB and 5 dB/

Figure 21 shows the results and LTAS comparisons of /ʃ/ with the N95 mask worn on the face. *Figure A* displays the trial 1 LTAS for /ʃ/ without a mask and with the N95 mask on the face. The blue line indicates the LTAS without a mask and the red line indicates the LTAS with the N95 mask on the face. The main difference in spectra lies in the salient frequency location for an /ʃ/ production, between 3000 Hz and 5000 Hz. *Figure B* displays the trial 2 LTAS for /ʃ/

without a mask and with the N95 mask on the face. The blue line indicates the LTAS without a mask and the red line indicates the LTAS with the N95 mask on the face. The main difference in spectra lies in the salient frequency location for an /ʃ/ production, between 3000 Hz and 5000 Hz. *Figure C* displays the intensity difference between the no mask LTAS and N95 mask LTAS for trial 1 at every 1000 Hz (blue line) and trial 2 at every 1000 Hz (orange line). The maximum difference for the entire spectrum was 12.7 dB at 14000 Hz for trial 1 and 11 dB at 14000 Hz for trial 2. The salient location for / ʃ / is located around 3000 Hz, which had attenuation of 7.9 dB for trial 1 and 4 dB for trial 2. This attenuation may create a perceptual quality change between the no mask and masked conditions.

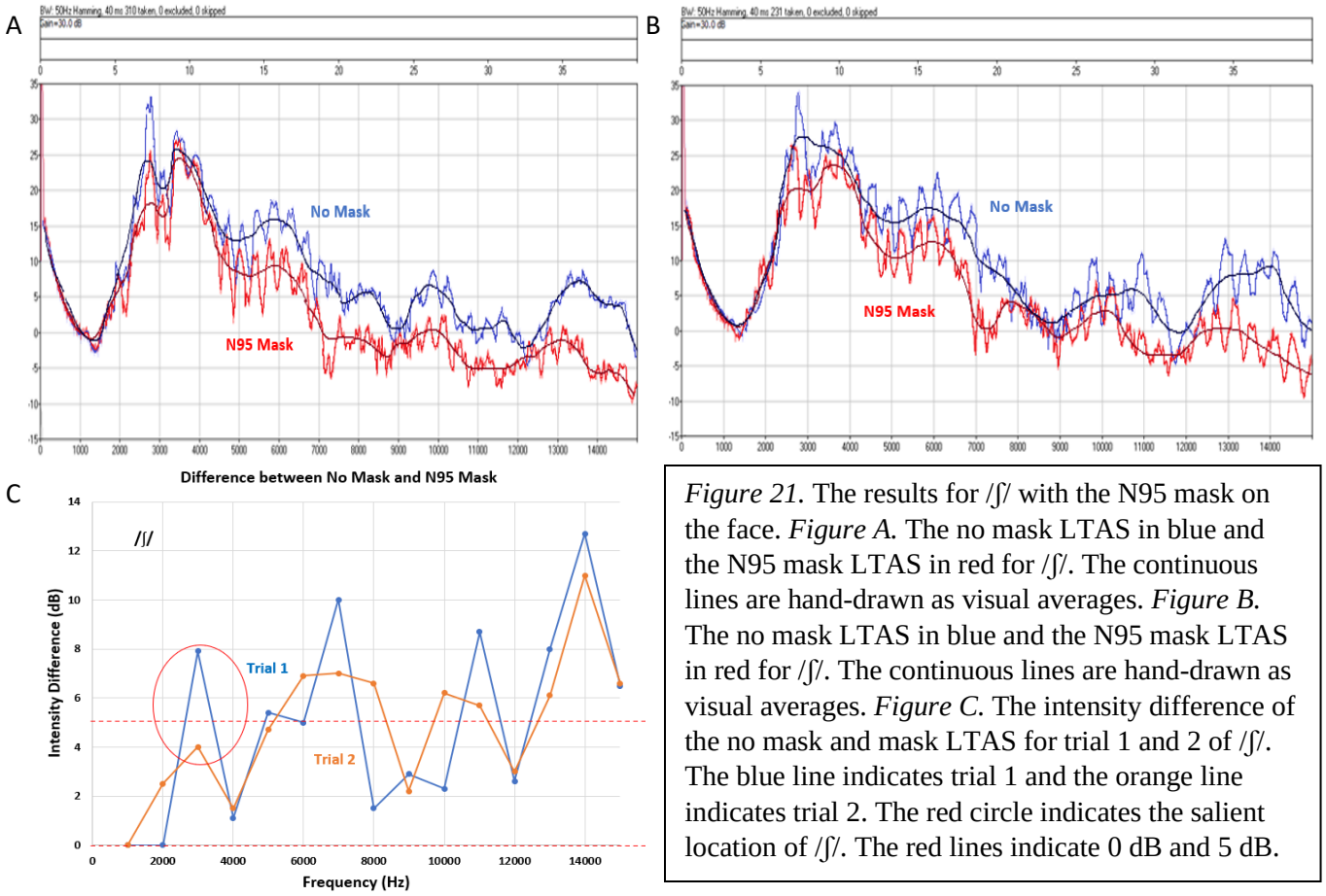

Figure 21. The results for /ʃ/ with the N95 mask on the face. *Figure A.* The no mask LTAS in blue and the N95 mask LTAS in red for /ʃ/. The continuous lines are hand-drawn as visual averages. *Figure B.* The no mask LTAS in blue and the N95 mask LTAS in red for /ʃ/. The continuous lines are hand-drawn as visual averages. *Figure C.* The intensity difference of the no mask and mask LTAS for trial 1 and 2 of /ʃ/. The blue line indicates trial 1 and the orange line indicates trial 2. The red circle indicates the salient location of /ʃ/. The red lines indicate 0 dB and 5 dB.

Singer's Mask

Vowels Spoken /a,i,u/

Figure 22 shows the results and comparisons of spoken /a/ with The Singer's mask on the face. *Figure A* shows the results of the spectra of the spoken sustained /a/ tokens with the Singer's Mask on the face. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 386.39 Hz and the maximum frequency difference among the f_o was 1.56 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum

was 1.1 dB (near 1500 Hz or H4) to 8.7 dB (near 4200 Hz or H11). *Figure B* displays the average spectra of spoken sustained /a/ with the mask and without the Singer's Mask on the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 0.81 Hz and the difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.33 dB (near 750 Hz or H2) to 10 dB (near 3500 Hz or H9). *Figure C* displays the difference between the no mask average spectrum and the mask average spectrum in intensity (dB) for a spoken sustained /a/. The intensity difference of the f_o was 1.17 dB, with the no mask spectrum higher than the mask spectrum. The figure suggests there is a slight increase in the intensity difference among the harmonics since the regression line has a positive slope. The R^2 term is small ($R^2 = 0.0172$), so the general relationship does not hold. However, there are several locations where the attenuation is 5 dB or more, specifically H6, H7, H8, and H9. These harmonics are close to F3, thus the difference in intensity may create a perceptual quality change between the no mask and masked conditions.

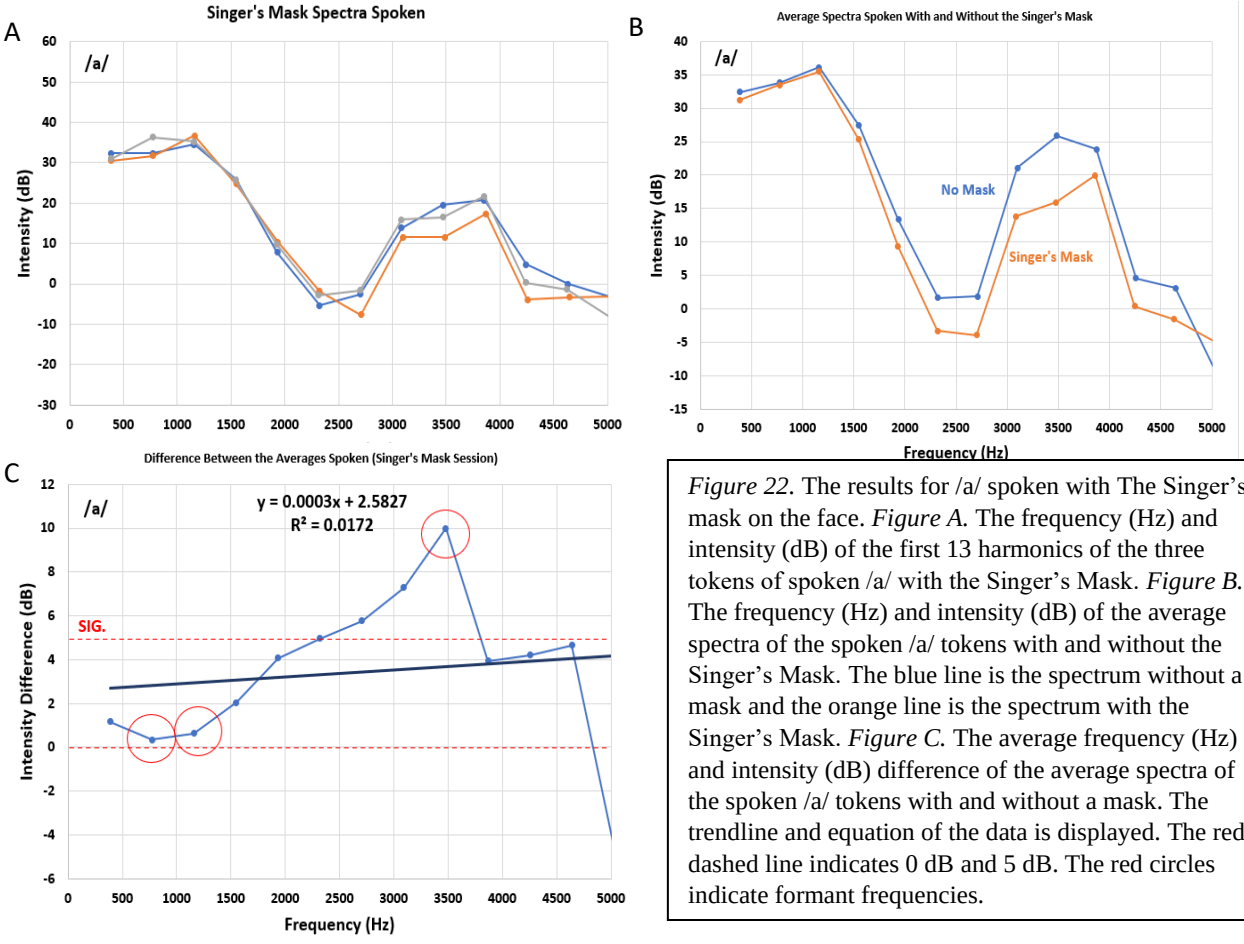

Figure 22. The results for /a/ spoken with The Singer's mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of spoken /a/ with the Singer's Mask. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the spoken /a/ tokens with and without the Singer's Mask. The blue line is the spectrum without a mask and the orange line is the spectrum with the Singer's Mask. *Figure C.* The average frequency (Hz) and intensity (dB) difference of the average spectra of the spoken /a/ tokens with and without a mask. The trendline and equation of the data is displayed. The red dashed line indicates 0 dB and 5 dB. The red circles indicate formant frequencies.

Figure 23 shows the results and comparisons of spoken /i/ with the Singer's Mask on the face. *Figure A* shows the results of the spectra of the spoken sustained /i/ tokens with the Singer's Mask on the face. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 388.05 Hz and the maximum frequency difference among the f_o was 0.6 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.7 dB (near 380 Hz or H1) to 8.5 dB (near H6, H7, or H11). *Figure B* displays the average spectra of spoken sustained /i/ with the mask and without the Singer's Mask on the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 1.13 Hz and the difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.43 dB (near 380 Hz or H1) to 8.56 dB (near 3500 Hz or H9). *Figure C* displays the difference between the no mask average spectrum and the Singer's Mask average spectrum in intensity (dB) for a spoken sustained /i/. The intensity difference of the f_o was 0.43 dB, with the no mask spectrum higher than the mask spectrum. The figure suggests there is a slight increase in the intensity difference among the harmonics since the regression line has a positive slope. The R^2 term is small (R^2 = 0.3898), so the general relationship does not hold. However, there are several locations where the attenuation is 5 dB or more, specifically H6, H9, and H11. These harmonics are close to F2 and F3, thus the difference in intensity may create a perceptual quality change between the no mask and masked conditions.

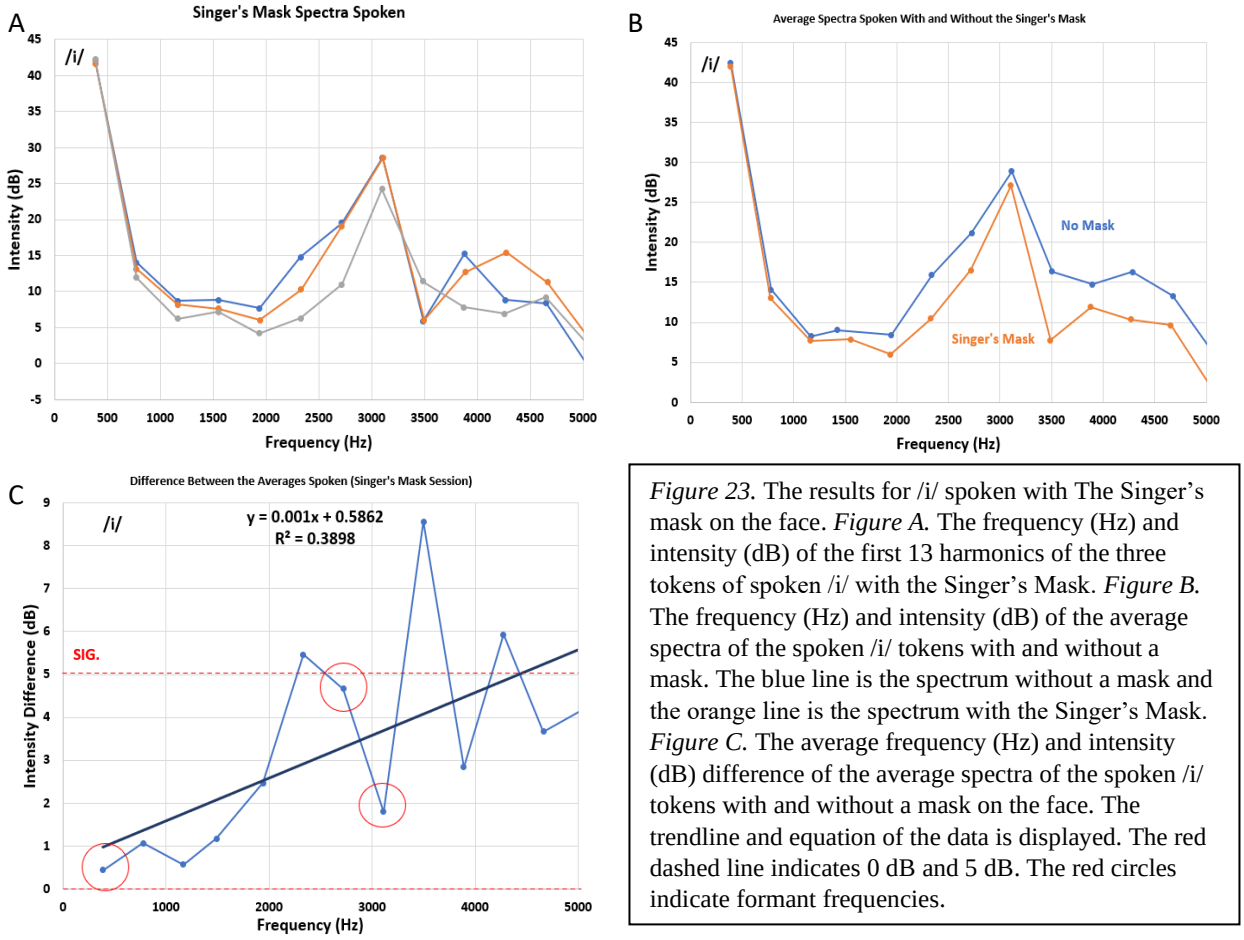

Figure 23. The results for /i/ spoken with The Singer's mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of spoken /i/ with the Singer's Mask. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the spoken /i/ tokens with and without a mask. The blue line is the spectrum without a mask and the orange line is the spectrum with the Singer's Mask. *Figure C.* The average frequency (Hz) and intensity (dB) difference of the average spectra of the spoken /i/ tokens with and without a mask on the face. The trendline and equation of the data is displayed. The red dashed line indicates 0 dB and 5 dB. The red circles indicate formant frequencies.

Figure 24 shows the results and comparisons of spoken /u/ with the Singer's Mask on the face. *Figure A* shows the results of the spectra of the spoken sustained /u/ tokens with the Singer's Mask on the face. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 386.09 Hz and the maximum frequency difference among the f_o was 4.24 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.4 dB (near 380 Hz or H1) to 10.5 dB (2650 Hz or H7). *Figure B* displays the average spectra of spoken sustained /u/ with the Singer's Mask and without the mask on the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 1.41 Hz and the difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.3 dB (near 2700 Hz or H7) to 9.76 dB (near 2300 Hz or H6). *Figure C* displays the difference in between the no mask average spectrum and the mask average spectrum in intensity (dB) for a spoken sustained /u/. The difference of the f_o was 2.87 dB, with the no mask spectrum higher than the mask spectrum. The figure suggests there is a slight increase in the intensity difference among the harmonics since the regression line has a positive slope. The R^2 term is small ($R^2 = 0.0892$), so the general relationship does not hold. There are several locations where the attenuation is above 5 dB or more, specifically H4, H5, and H10. These harmonics are not close to formants thus the attenuation is not perceptually relevant.

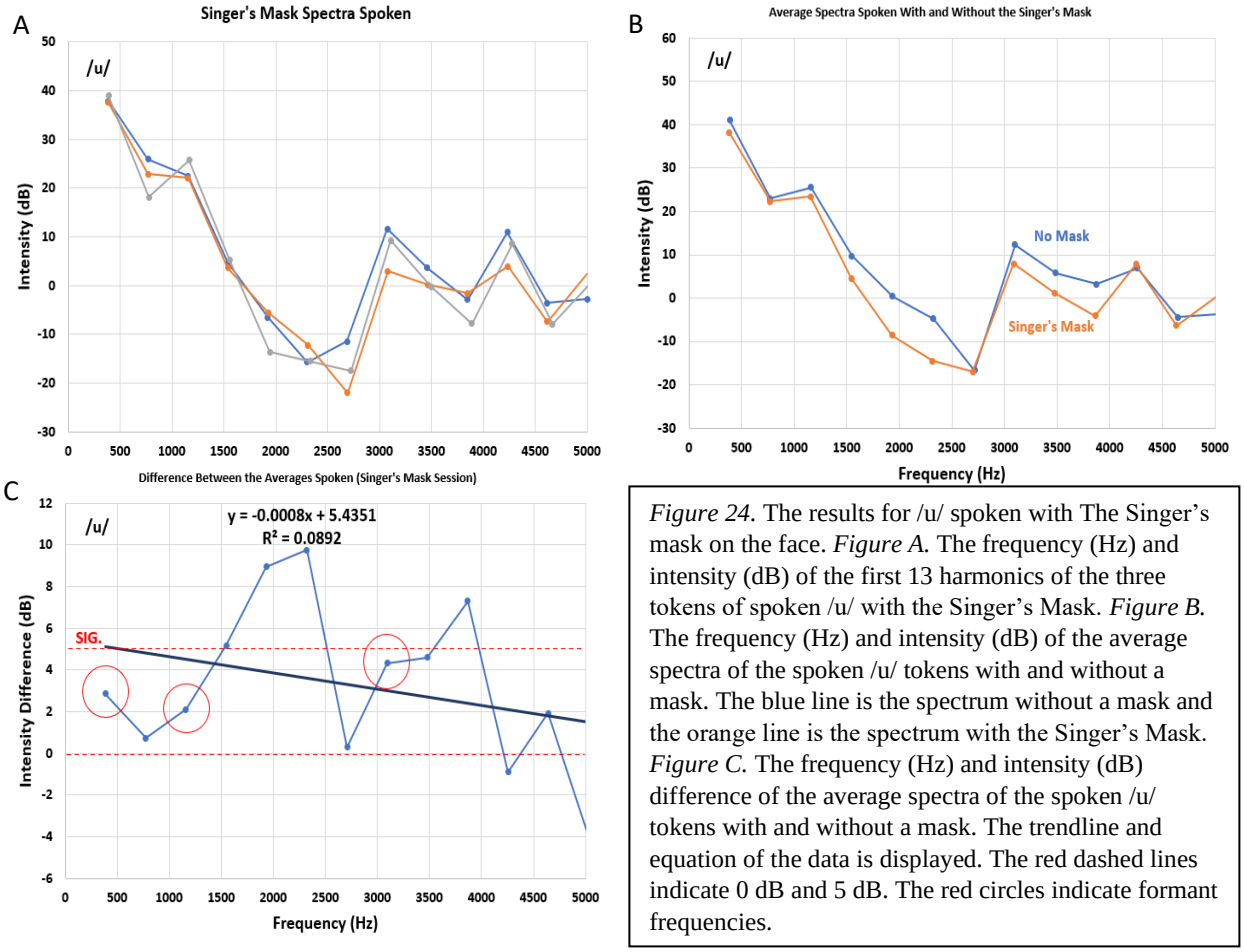

Figure 24. The results for /u/ spoken with The Singer's mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of spoken /u/ with the Singer's Mask. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the spoken /u/ tokens with and without a mask. The blue line is the spectrum without a mask and the orange line is the spectrum with the Singer's Mask. *Figure C.* The frequency (Hz) and intensity (dB) difference of the average spectra of the spoken /u/ tokens with and without a mask. The trendline and equation of the data is displayed. The red dashed lines indicate 0 dB and 5 dB. The red circles indicate formant frequencies.

Vowels on G4 /a,i,u/

Figure 25 shows the results and comparisons of /a/ sung on G4 with the Singer's Mask on the face. *Figure A* shows the results for the spectra of the /a/ tokens sung sustained on G4 with the Singer's Mask on the face. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 414.47 Hz and the maximum frequency difference among the f_o was 12.33 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.8 dB (near 1250 Hz or H3) to 15.2 dB (4550 Hz or H11). *Figure B* displays the average spectra of sustained /a/ on G4 with the Singer's mask and without the mask on the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 1.12 Hz and the difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.6 dB (near 1250 Hz or H3) to 7.83 dB (near 2300 Hz or H6). *Figure C* displays the difference in between the no mask average spectrum and the mask average spectrum in intensity (dB) for /a/ sung on G4. The intensity difference of the f_o was 0.43 dB, with the no mask spectrum higher than the mask spectrum. The figure suggests there is a slight decrease in the intensity difference among the harmonics since the regression line has a negative slope. The R^2 term is small (R^2 = 0.1044), so the general relationship does not hold. There are several locations where the attenuation is above 5 dB or more, specifically H5 and H6. These harmonics are not close to a

formant, so the intensity differences displayed may not attribute to perceptual changes in the vowel.

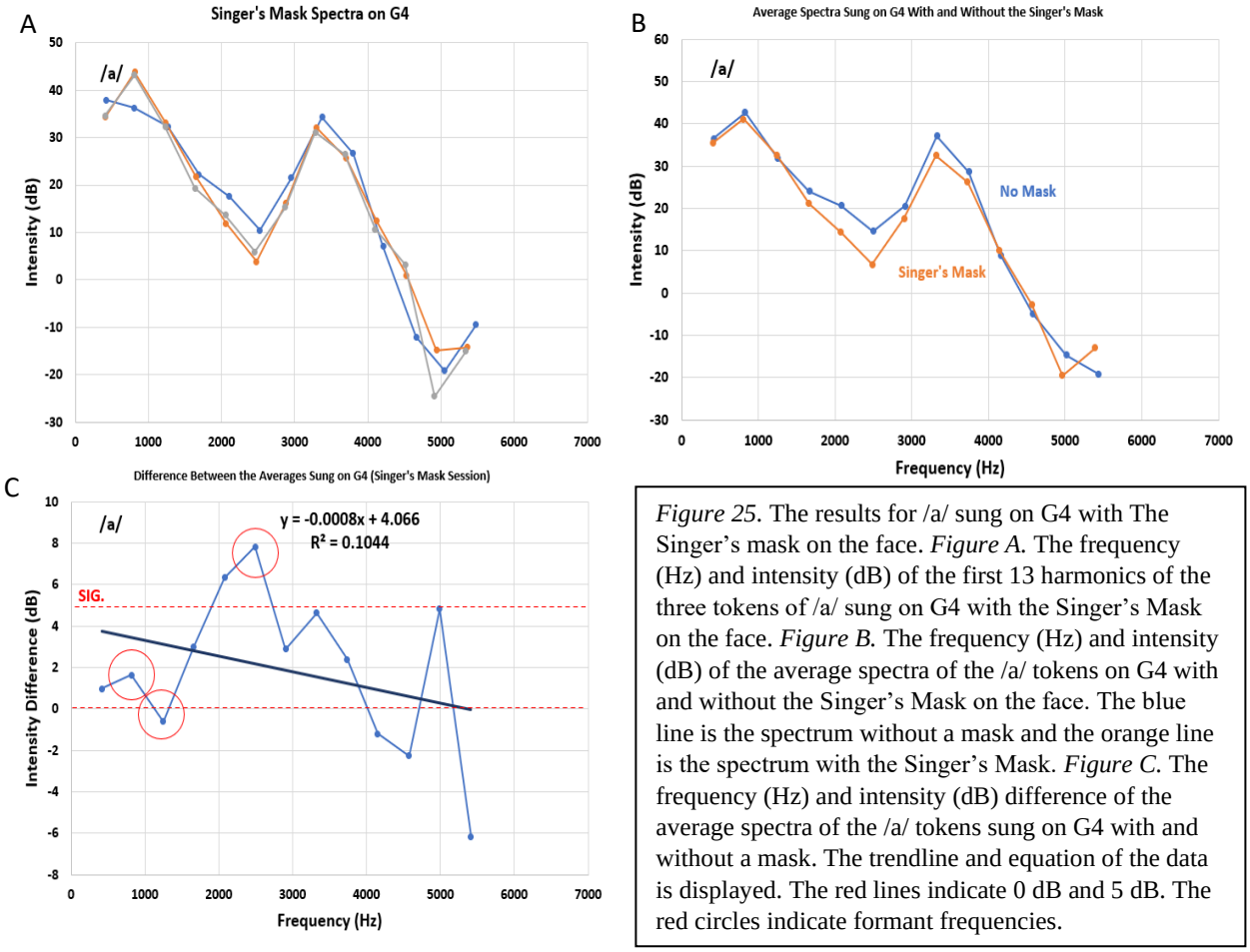

Figure 25. The results for /a/ sung on G4 with The Singer's mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of /a/ sung on G4 with the Singer's Mask on the face. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the /a/ tokens on G4 with and without the Singer's Mask on the face. The blue line is the spectrum without a mask and the orange line is the spectrum with the Singer's Mask. *Figure C.* The frequency (Hz) and intensity (dB) difference of the average spectra of the /a/ tokens sung on G4 with and without a mask. The trendline and equation of the data is displayed. The red lines indicate 0 dB and 5 dB. The red circles indicate formant frequencies.

Figure 26 shows the results and comparisons of /i/ sung on G4 with the Singer's Mask on the face. *Figure A* shows the results of the spectra of the /i/ tokens sung sustained on G4 with the Singer's Mask on the face. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 405.08 Hz and the maximum frequency difference among the f_o was 4.19 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 2.2 dB (near 1600 Hz or H4) to 14.1 dB (5200 Hz or H13). *Figure B* displays the average spectra of sustained /i/ sung on G4 with the Singer's mask and without the mask on the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 3.05 Hz and the difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.2 dB (near 400 Hz or H1) to 3.47 dB (near 800 Hz or H2). *Figure C* displays the difference in between the no mask average spectrum and the Singer's Mask average spectrum in intensity (dB) for a sustained /i/ on G4. The intensity difference of the f_o was 0.2 dB, with the no mask spectrum

higher than the mask spectrum. The figure suggests there is a slight decrease in the intensity difference among the harmonics since the regression line has a negative slope. The R^2 term is small ($R^2 = 0.3474$), so the general relationship does not hold. There are no areas where the attenuation is more than 5 dB, so the intensity differences displayed may not attribute to perceptual changes in the vowel.

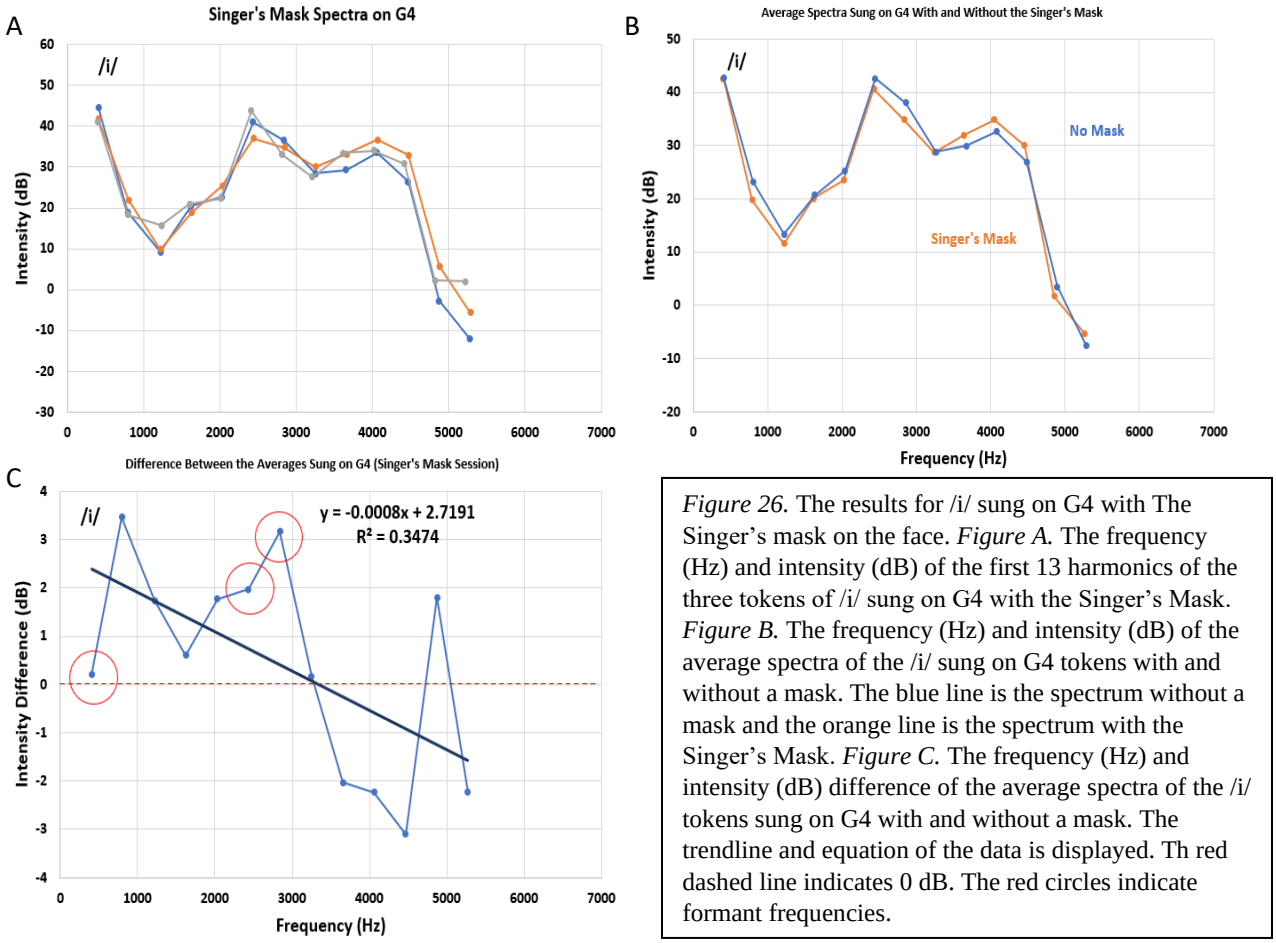

Figure 26. The results for /i/ sung on G4 with The Singer's mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of /i/ sung on G4 with the Singer's Mask. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the /i/ sung on G4 tokens with and without a mask. The blue line is the spectrum without a mask and the orange line is the spectrum with the Singer's Mask. *Figure C.* The frequency (Hz) and intensity (dB) difference of the average spectra of the /i/ tokens sung on G4 with and without a mask. The trendline and equation of the data is displayed. Th red dashed line indicates 0 dB. The red circles indicate formant frequencies.

Figure 27 shows the results and comparisons of /u/ sung on G4 with the Singer's Mask on the face. *Figure A* shows the results of the spectra of the /u/ tokens sung sustained on G4 with the Singer's Mask on the face. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 408.51 Hz and the maximum frequency difference among the f_o was 2.35 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 2.0 dB (near 400 Hz or H1) to 13.6 dB (5300 Hz or H13). *Figure B* displays the average spectra of sustained /u/ sung on G4 with the Singer's Mask and without the mask on the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 0.69 Hz and the difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.1 dB (near 1600 Hz or H4) to 17.6 dB (near 4900 Hz or H12). *Figure C* displays the difference

between the no mask average spectrum and the mask average spectrum in intensity (dB) for a sustained /u/ on G4. The difference of the f_o was 1.9 dB, with the no mask spectrum higher than the mask spectrum. The figure suggests there is a slight increase in the intensity difference among the harmonics since the regression line has a positive slope. The R^2 term is small (R^2 = 0.4003), so the general relationship does not hold. There are two locations where the attenuation is 5 dB or more, specifically at H7 and H12. These harmonics are not close to a formant, so the intensity differences displayed may not attribute to perceptual changes in the vowel.

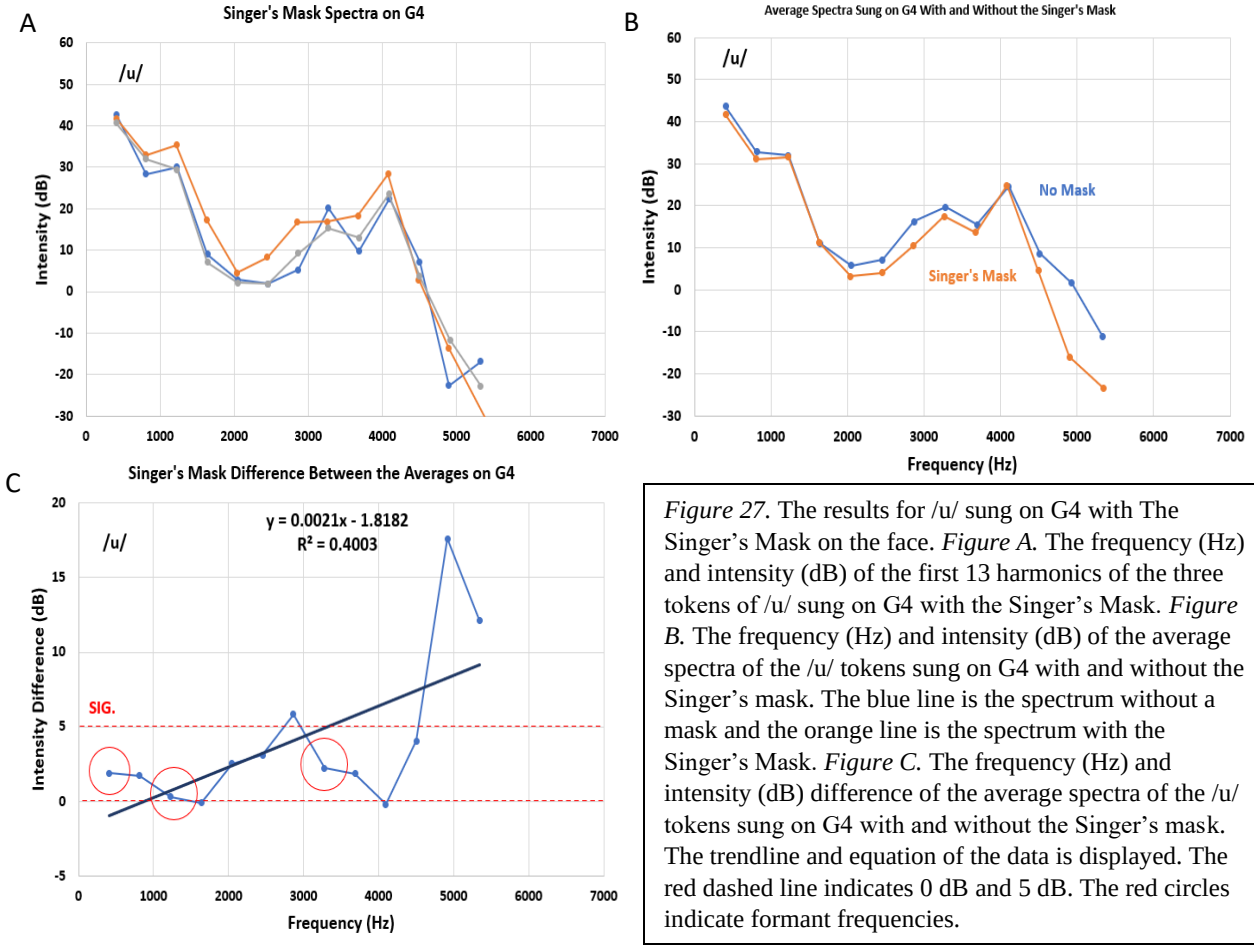

Figure 27. The results for /u/ sung on G4 with The Singer's Mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of /u/ sung on G4 with the Singer's Mask. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the /u/ tokens sung on G4 with and without the Singer's mask. The blue line is the spectrum without a mask and the orange line is the spectrum with the Singer's Mask. *Figure C.* The frequency (Hz) and intensity (dB) difference of the average spectra of the /u/ tokens sung on G4 with and without the Singer's mask. The trendline and equation of the data is displayed. The red dashed line indicates 0 dB and 5 dB. The red circles indicate formant frequencies.

Vowels on C5 /a,i,u/

Figure 28 shows the results and comparisons of /a/ sung on C5 with the Singer's Mask on the face. *Figure A* shows the results of the spectra for the /a/ tokens sung sustained on C5 with the Singer's Mask on the face. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 540.76 Hz and the maximum frequency difference among the f_o was 8.0 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.0 dB (near 540 Hz or H1) to 17.6 dB (5900 Hz or H11). *Figure B* displays the average spectra of sustained /a/ on C5 with the Singer's mask and without the mask on the face. The blue line represents the spectrum without the mask and the orange line represents the

spectrum with the mask. The maximum frequency difference among the f_o was 0.68 Hz and the difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.27 dB (near 1600 Hz or H3) to 11.83 dB (near 9500 Hz or H11). *Figure C* displays the difference between the no mask average spectrum and the Singer's Mask average spectrum in intensity (dB) for a sustained /a/ sung on C5. The difference of the f_o was 1.7 dB, with the no mask spectrum higher than the mask spectrum. The figure suggests there is a slight increase in the intensity difference among the harmonics since the regression line has a negative slope. The R^2 term is small ($R^2 = 0.2021$), so the general relationship does not hold. There are several locations where the attenuation is above 5 dB or more, specifically at H4, H9, H10, and H11. H11 is near F4, thus the difference in intensity may create a perceptual quality change between the no mask and masked conditions.

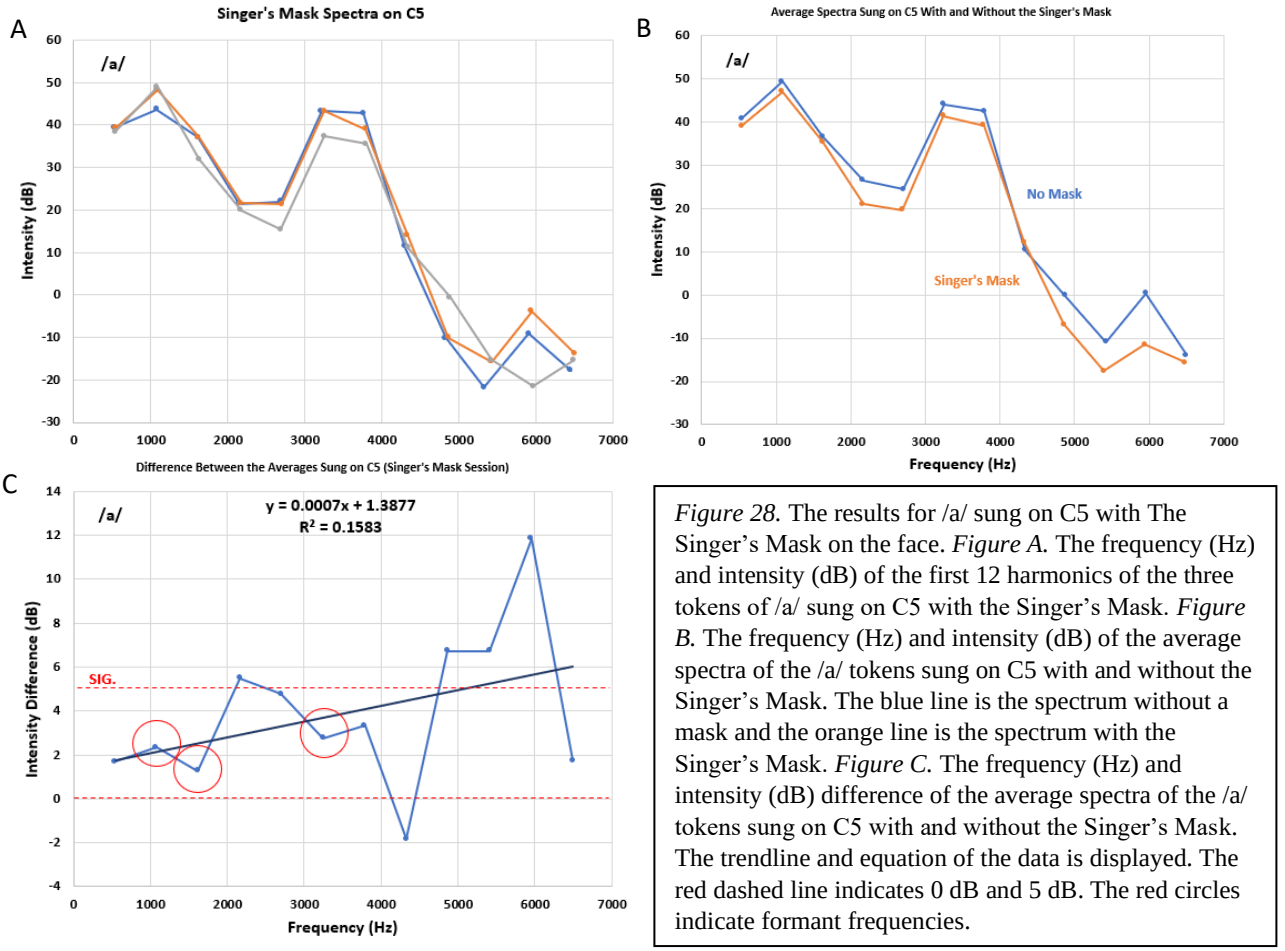

Figure 28. The results for /a/ sung on C5 with The Singer's Mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 12 harmonics of the three tokens of /a/ sung on C5 with the Singer's Mask. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the /a/ tokens sung on C5 with and without the Singer's Mask. The blue line is the spectrum without a mask and the orange line is the spectrum with the Singer's Mask. *Figure C.* The frequency (Hz) and intensity (dB) difference of the average spectra of the /a/ tokens sung on C5 with and without the Singer's Mask. The trendline and equation of the data is displayed. The red dashed line indicates 0 dB and 5 dB. The red circles indicate formant frequencies.

Figure 29 shows the results and comparisons of /i/ sung on C5 with the Singer's Mask on the face. *Figure A* shows the results of the spectra for the /i/ tokens sung on C5 with the Singer's Mask on the face. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 543.71 Hz and the maximum frequency difference among the f_o was 4.85 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.1 dB (near 540 Hz or H1) to 17.4 dB (near 6000 Hz or H11). *Figure B* displays the average spectra of sustained /i/ sung on C5 with the Singer's Mask and without the mask on the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 2.01 Hz and the difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.83 dB (near 1000 Hz or H2) to 8.33 dB (near 4900 Hz or H9). *Figure C* displays the difference between the no mask average spectrum and the mask average spectrum in intensity (dB) for a sustained /i/ sung on C5. The difference of the f_o was 1.83 dB, with the no mask spectrum higher than the mask spectrum. The figure suggests there is a slight increase in the intensity difference among the harmonics since the regression line has a positive slope. The R^2 term is small ($R^2 = 0.0204$), so the general relationship does not hold. There are several locations where the attenuation is above 5 dB or more, specifically H4, H5, H6, H7, and H9. These harmonics are near F2 and F3 thus the difference in intensity may create a perceptual quality change between the no mask and masked conditions.

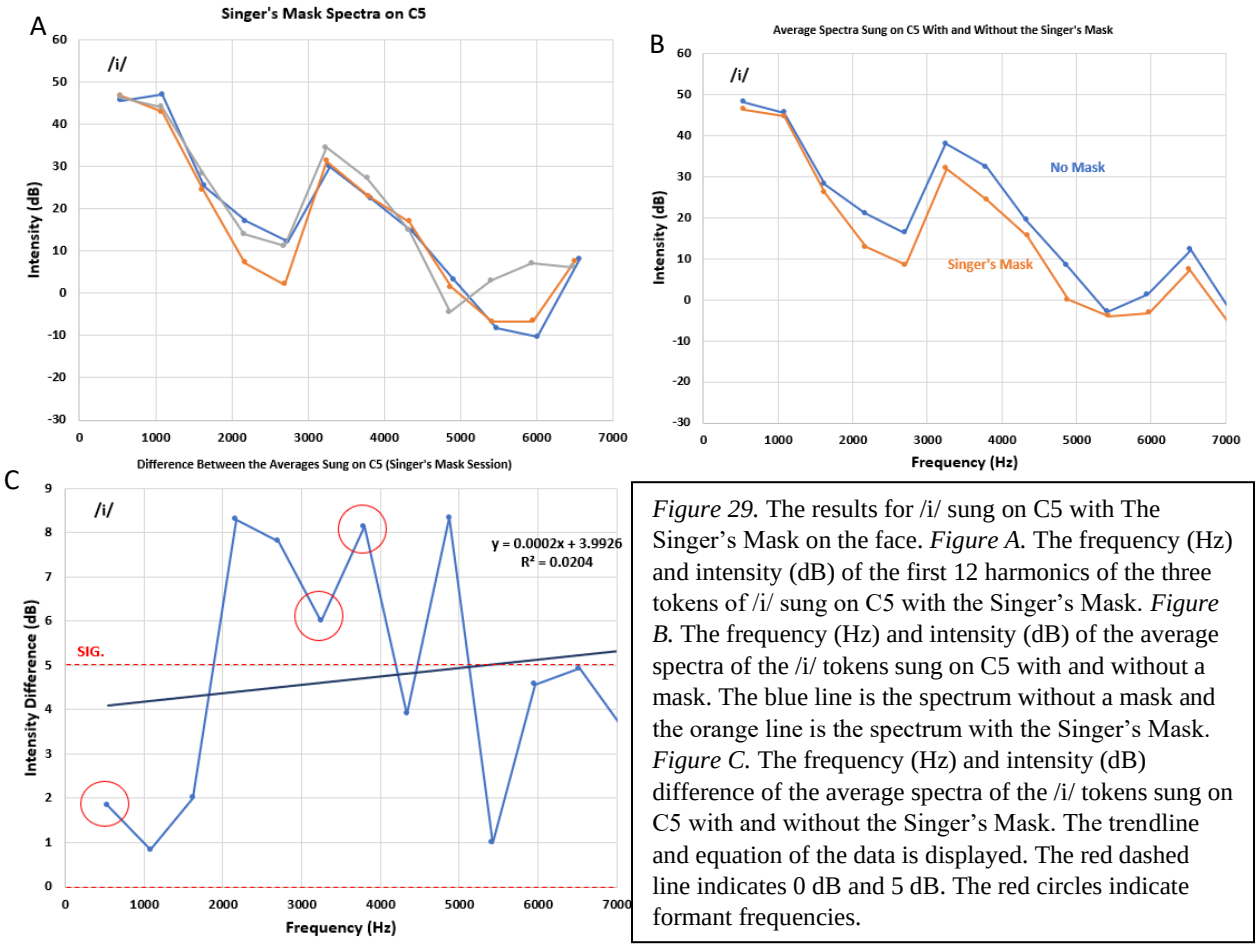

Figure 29. The results for /i/ sung on C5 with The Singer's Mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 12 harmonics of the three tokens of /i/ sung on C5 with the Singer's Mask. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the /i/ tokens sung on C5 with and without a mask. The blue line is the spectrum without a mask and the orange line is the spectrum with the Singer's Mask. *Figure C.* The frequency (Hz) and intensity (dB) difference of the average spectra of the /i/ tokens sung on C5 with and without the Singer's Mask. The trendline and equation of the data is displayed. The red dashed line indicates 0 dB and 5 dB. The red circles indicate formant frequencies.

Figure 30 shows the results and comparisons of /u/ sung on C5 with the Singer's Mask on the face. *Figure A* shows the results of the spectra of the /u/ tokens sung sustained on C5 with the Singer's Mask on the face. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 538.38 Hz and the maximum frequency difference among the f_o was 1.64 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 3.0 dB (near 1600 Hz or H3) to 16.4 dB (6400 Hz or H12). *Figure B* displays the average spectra of sustained /u/ sung on C5 with the Singer's Mask and without the mask on the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 6.81 Hz and the difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.73 dB (near 6000 Hz or H11) to 7.77 dB (near 2300 Hz or H10). *Figure C* displays the difference in between the no mask average spectrum and the mask average spectrum in intensity (dB) for a sustained /u/ sung on C5. The intensity difference of the f_o was 0.87 dB, with the no mask spectrum higher than the mask spectrum. The figure suggests there is a slight increase in the intensity difference among the harmonics since the regression line has a positive slope. The R^2 term is small ($R^2 = 0.1202$), so the general relationship does not hold. There are several locations where the attenuation is above 5 dB or more, specifically, H8, H10, and H12. These harmonics are not close to a formant, so the intensity differences displayed may not attribute to perceptual changes in the vowel.

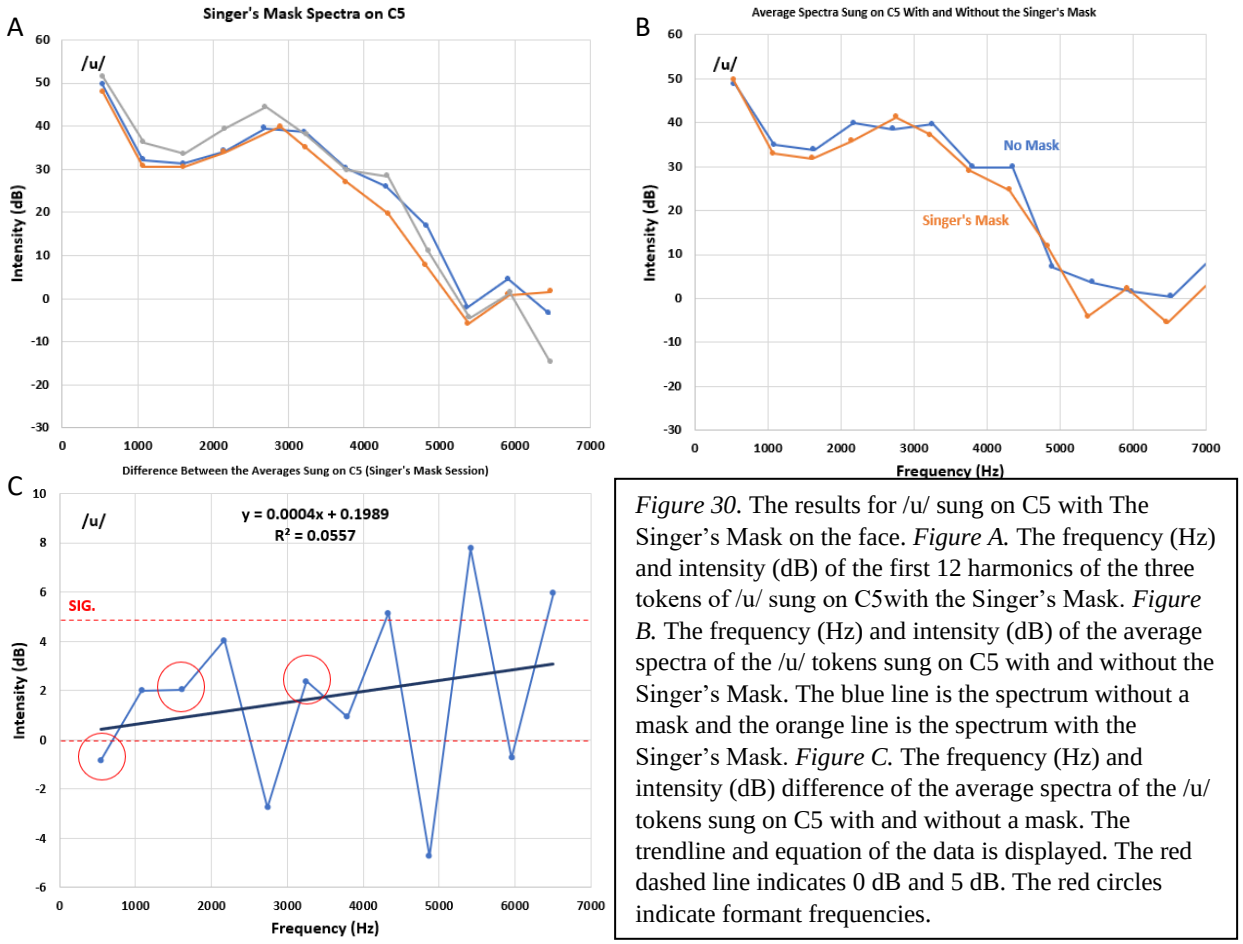

Figure 30. The results for /u/ sung on C5 with The Singer's Mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 12 harmonics of the three tokens of /u/ sung on C5with the Singer's Mask. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the /u/ tokens sung on C5 with and without the Singer's Mask. The blue line is the spectrum without a mask and the orange line is the spectrum with the Singer's Mask. *Figure C.* The frequency (Hz) and intensity (dB) difference of the average spectra of the /u/ tokens sung on C5 with and without a mask. The trendline and equation of the data is displayed. The red dashed line indicates 0 dB and 5 dB. The red circles indicate formant frequencies.

LTAS Comparisons /s,ʃ/

Figure 31 shows the results and LTAS comparisons of /s/ with the Singer's Mask worn on the face. *Figure A* displays the trial 1 LTAS for /s/ without a mask and with the Singer's Mask. The blue line indicates the LTAS without a mask and the red line indicates the LTAS with the Singer's Mask on the face. The main difference in spectra lies in the salient frequency location for an /s/ production, between 5000 Hz and 7000 Hz. *Figure B* displays the trial 2 LTAS for /s/ without a mask and with the Singer's Mask. The blue line indicates the LTAS without a mask and the red line indicates the LTAS with the Singer's Mask on the face. Again, the main difference in spectra lies in the salient frequency location for an /s/ production, between 5000 Hz and 7000 Hz. *Figure C* displays the intensity difference between the no mask LTAS and the Singer's Mask LTAS for trial 1 at every 1000 Hz (blue line) and for trial 2 at every 1000 Hz (orange line). The maximum difference for the entire spectrum was 10.2 dB at 7000 Hz for trial 1 and 12 dB at 6000 Hz for trial 2. The salient location for /s/ is located around 6000 Hz, thus the highest attenuation at this area may create a perceptual quality change between the no mask and masked conditions. That is, the perception of /s/ may be greatly reduced during speech due to being about 10 dB to12 dB reduced when the mask was worn.

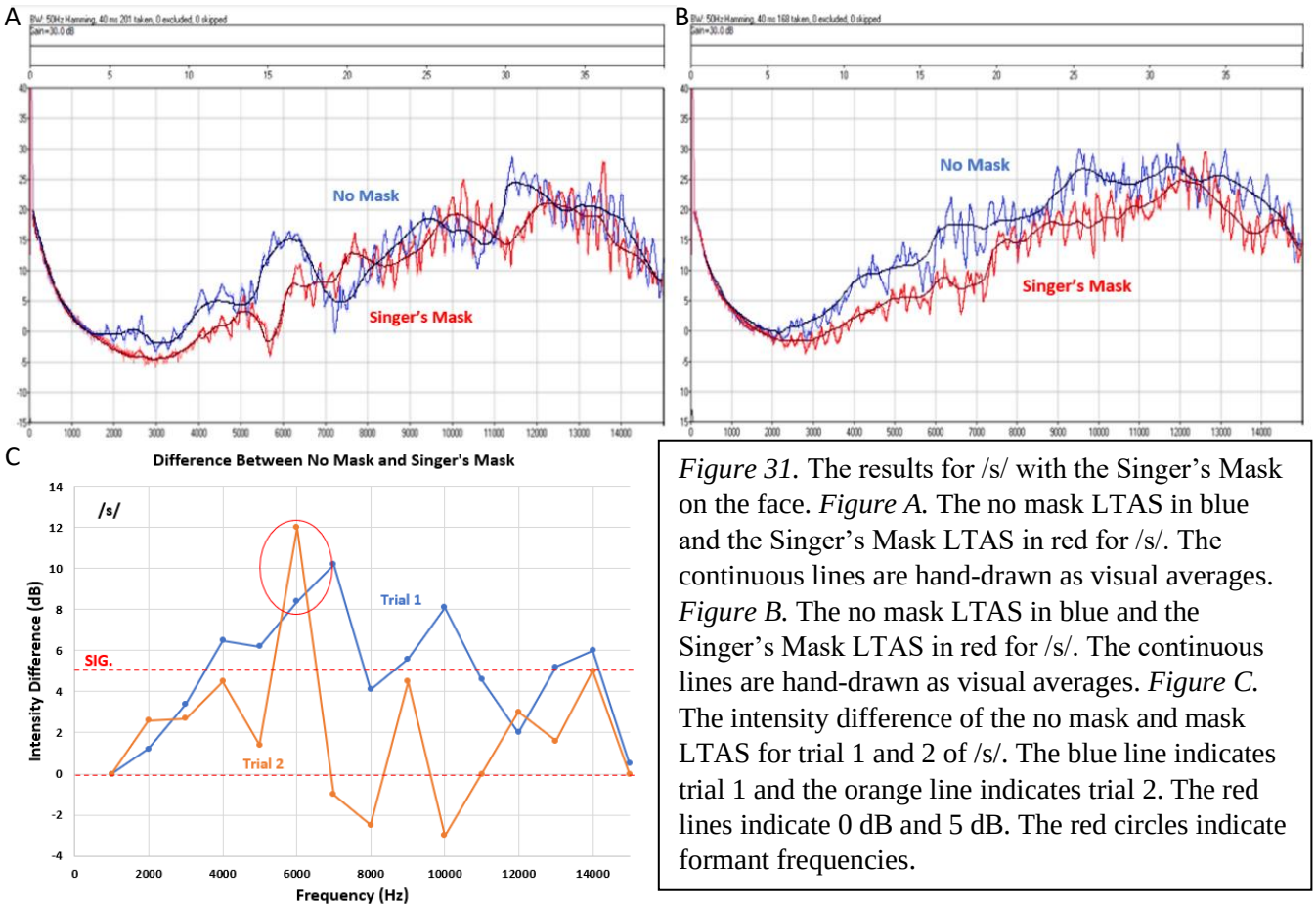

Figure 31. The results for /s/ with the Singer's Mask on the face. *Figure A.* The no mask LTAS in blue and the Singer's Mask LTAS in red for /s/. The continuous lines are hand-drawn as visual averages. *Figure B.* The no mask LTAS in blue and the Singer's Mask LTAS in red for /s/. The continuous lines are hand-drawn as visual averages. *Figure C.* The intensity difference of the no mask and mask LTAS for trial 1 and 2 of /s/. The blue line indicates trial 1 and the orange line indicates trial 2. The red lines indicate 0 dB and 5 dB. The red circles indicate formant frequencies.

Figure 32 shows the results and LTAS comparisons of /ʃ/ with the Singer's Mask worn on the face. *Figure A* displays the trial 1 LTAS for /ʃ/ without a mask and with the Singer's Mask on the face. The blue line indicates the LTAS without a mask and the red line indicates the LTAS with the Singer's Mask on the face. The main difference in spectra lies in the salient frequency location for an /ʃ/ production, between 3000 Hz and 5000 Hz. *Figure B* displays the trial 2 LTAS for /ʃ/ without a mask and with the Singer's Mask. The blue line indicates the LTAS without a mask and the red line indicates the LTAS with the Singer's Mask on the face. Again, the main difference in spectra lies in the salient frequency location for an /ʃ/ production, between 3000 Hz and 5000 Hz. *Figure C* displays the intensity difference between the no mask LTAS and the Singer's Mask LTAS for trial 1 at every 1000 Hz (blue line) and for trial 2 at every 1000 Hz (orange line). The maximum difference for the entire spectrum was 6.8 dB at 12000 Hz for trial 1 and 9 dB at 8000 Hz for trial 2. The salient location for /ʃ/ is located around 3000 Hz to 5000 Hz, thus the highest attenuation at this area may create a perceptual quality change between the no mask and masked conditions. The peak attenuation for at the salient location for trial 1 was 4.9 dB at 4000 Hz and for trial 2 the peak attenuation was 7.6 dB at 3000 Hz.

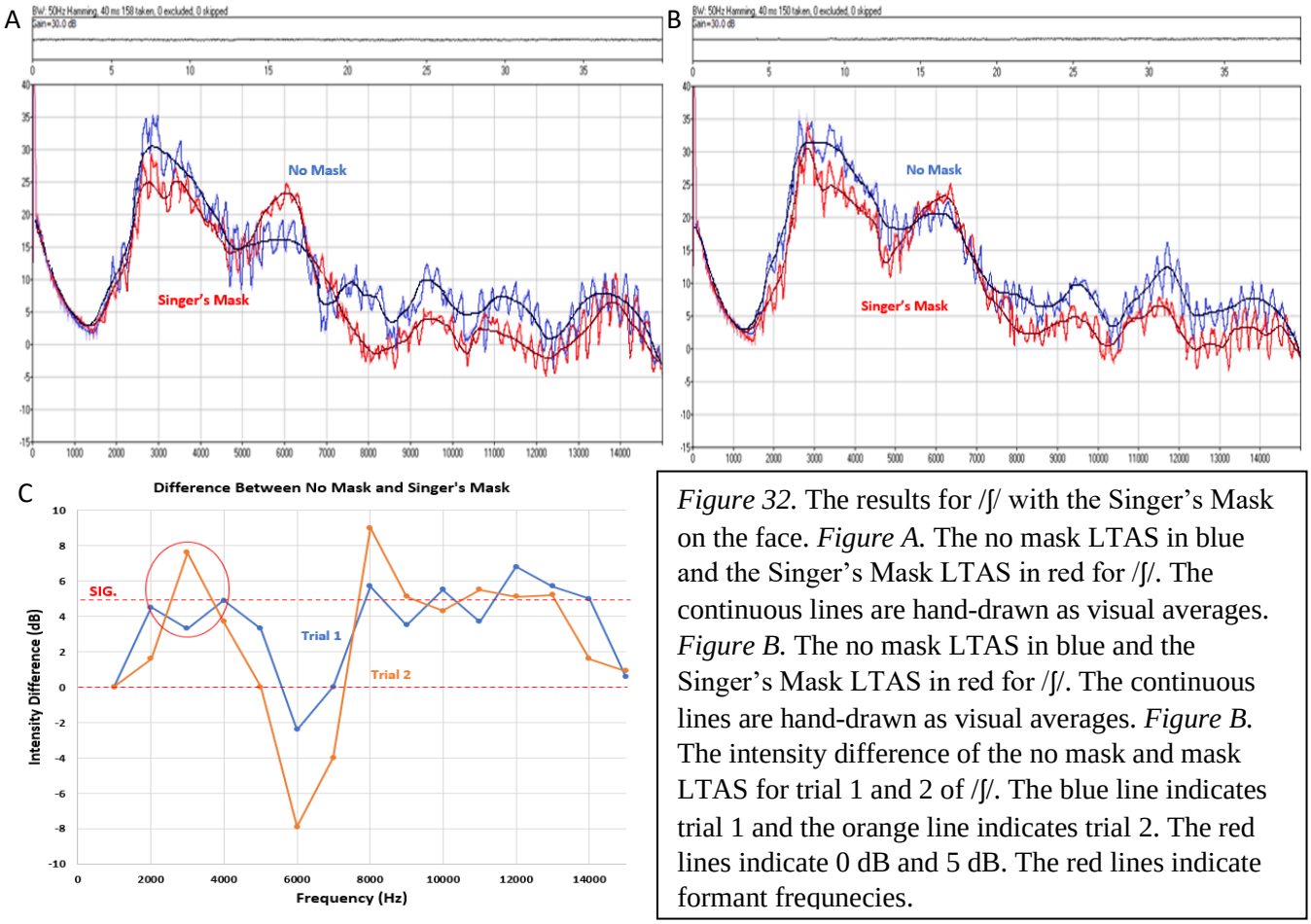

Figure 32. The results for /ʃ/ with the Singer's Mask on the face. *Figure A.* The no mask LTAS in blue and the Singer's Mask LTAS in red for /ʃ/. The continuous lines are hand-drawn as visual averages. *Figure B.* The no mask LTAS in blue and the Singer's Mask LTAS in red for /ʃ/. The continuous lines are hand-drawn as visual averages. *Figure B.* The intensity difference of the no mask and mask LTAS for trial 1 and 2 of /ʃ/. The blue line indicates trial 1 and the orange line indicates trial 2. The red lines indicate 0 dB and 5 dB. The red lines indicate formant frequnecies.

Surgical Mask

Vowels Spoken /a,i,u/

Figure 33 shows the results and comparisons of /a/ spoken with the surgical mask on the face. *Figure A* shows the results of the spectra for the spoken sustained /a/ tokens with the surgical mask on the face. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 382.63 Hz and the maximum frequency difference among the f_o was 4.95 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 2.4 dB (near 3400 Hz or H9) to 10.4 dB (near 2300 Hz or H3). *Figure B* displays the average spectra of spoken sustained /a/ with the surgical mask and without the mask on the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 0.20 Hz and the difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.17 dB (near 2300 Hz or H6) to 7.63 dB (near 3800 Hz or H10). *Figure C* displays the difference between the no mask average spectrum and the mask average spectrum in intensity (dB) for a spoken sustained /a/. The difference of the f_o was 0.67 dB, with the no mask spectrum higher than the mask spectrum. The figure suggests there is a slight increase in the intensity difference among the

harmonics since the regression line has a positive slope. The R^2 term is very small (R^2 = 0.IE-04), so the general relationship does not hold. There is a location where the attenuation is above 5 dB or more, specifically at H10. H10 is close to F3, so the attenuation may attribute to perceptual changes heard while wearing the mask on the face.

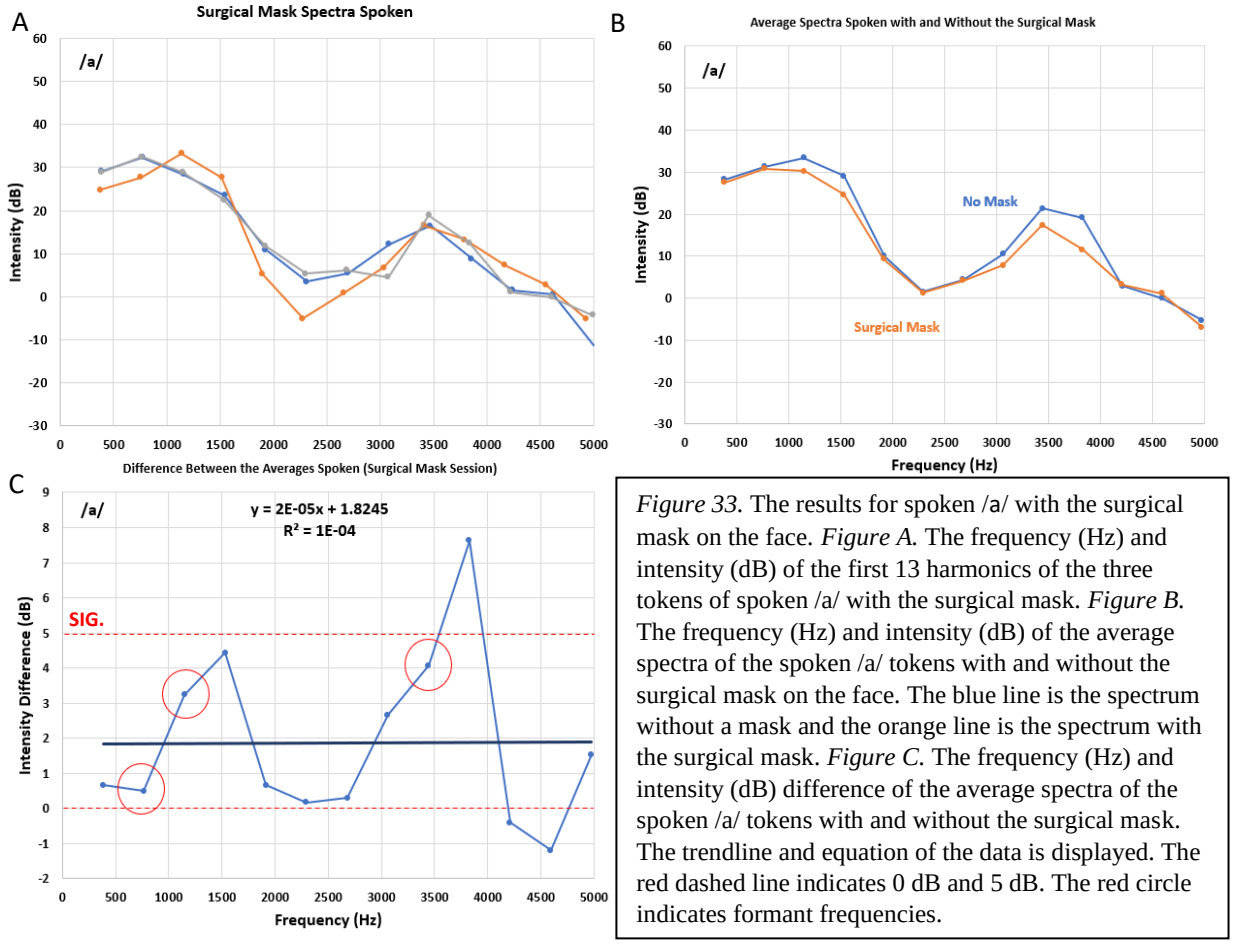

Figure 33. The results for spoken /a/ with the surgical mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of spoken /a/ with the surgical mask. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the spoken /a/ tokens with and without the surgical mask on the face. The blue line is the spectrum without a mask and the orange line is the spectrum with the surgical mask. *Figure C.* The frequency (Hz) and intensity (dB) difference of the average spectra of the spoken /a/ tokens with and without the surgical mask. The trendline and equation of the data is displayed. The red dashed line indicates 0 dB and 5 dB. The red circle indicates formant frequencies.

Figure 34 shows the results and comparisons of /i/ spoken with the surgical mask on the face. *Figure A* shows the results of the spectra for the spoken sustained /i/ with the surgical mask on the face. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 388.34 Hz and the maximum frequency difference among the f_o was 2.53 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.3 dB (near 1500 Hz or H4) to 20.3 dB (near 700 Hz or H2). *Figure B* displays the average spectra of spoken sustained /i/ with the surgical mask and without the mask on the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 0.43 Hz and the difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.27 dB (near H1 and H6) to 6.2 dB (near 4600 Hz or H12). *Figure C* displays the difference in between the no mask average spectrum and the mask average spectrum in intensity (dB) for a spoken sustained /i/. The

difference of the f_o was 0.27 dB, with the no mask spectrum higher than the mask spectrum. The figure suggests there is a slight decrease in the intensity difference among the harmonics since the regression line has a negative slope. The R^2 term is higher ($R^2 = 0.5952$), so the general relationship does hold. There are several locations where the attenuation is above 5 dB or more, specifically H8, H10, H11, and H12. These harmonics are close to F3 and F4, so this attenuation may attribute to perceptual changes heard while wearing the mask on the face.

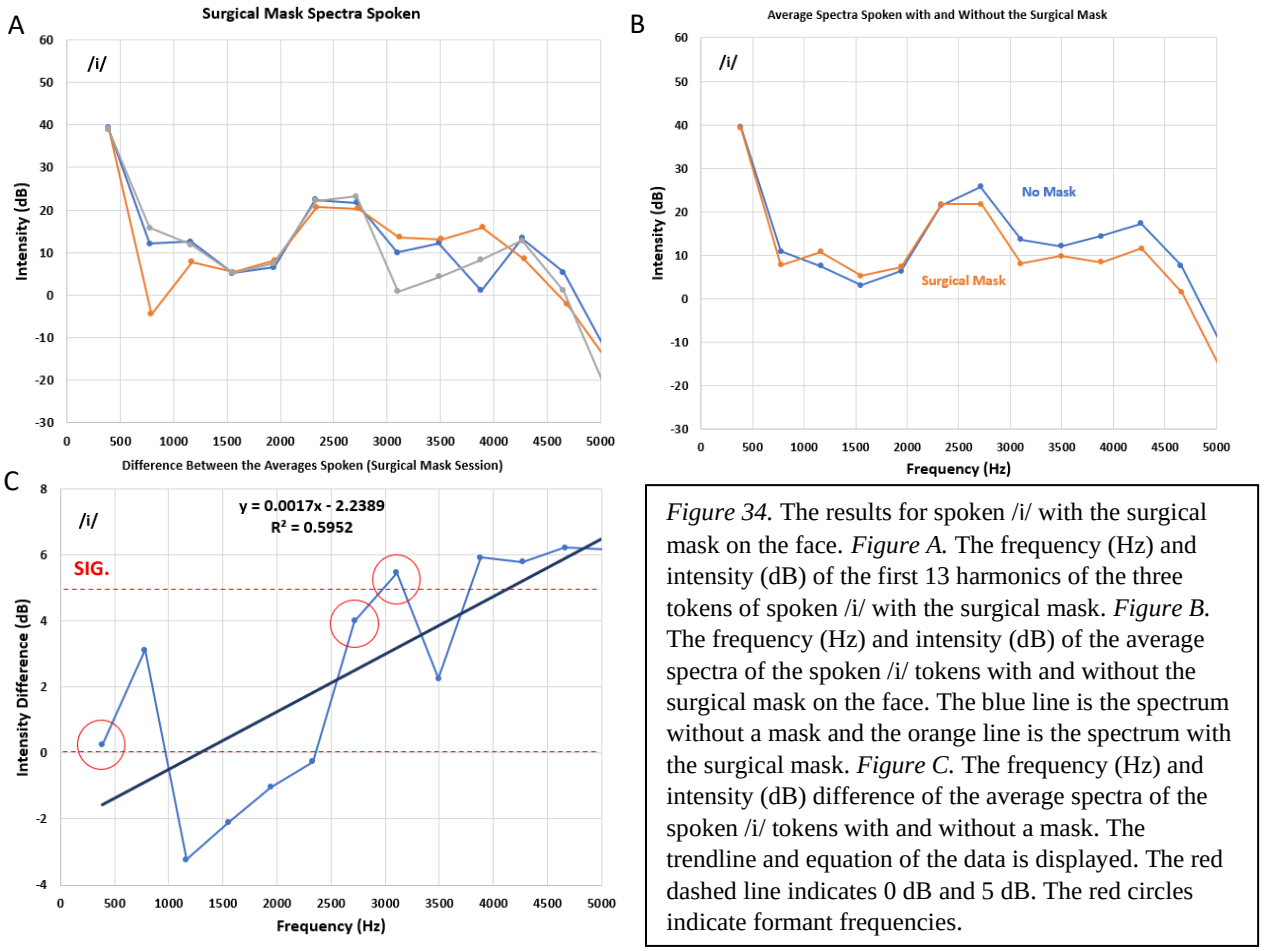

Figure 34. The results for spoken /i/ with the surgical mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of spoken /i/ with the surgical mask. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the spoken /i/ tokens with and without the surgical mask on the face. The blue line is the spectrum without a mask and the orange line is the spectrum with the surgical mask. *Figure C.* The frequency (Hz) and intensity (dB) difference of the average spectra of the spoken /i/ tokens with and without a mask. The trendline and equation of the data is displayed. The red dashed line indicates 0 dB and 5 dB. The red circles indicate formant frequencies.

Figure 35 shows the results and comparisons of /u/ spoken with the surgical mask on the face. *Figure A* shows the results for the spectra of the spoken sustained /u/ with the surgical mask on the face. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 388.02 Hz and the maximum frequency difference among the f_o was 2.16 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.0 dB (near 4200 Hz or H11) to 21.0 dB (near 700 Hz or H2). *Figure B* displays the average spectra of spoken sustained /u/ with the surgical mask and without the mask on the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 1.18 Hz and the difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.0 dB (near 700 Hz or H2) to 11.1 dB (near 3100 Hz or H8). *Figure C* displays the difference between the no mask

average spectrum and the mask average spectrum in intensity (dB) for a spoken sustained /u/. The difference of the f_o was 1.83 dB, with the no mask spectrum higher than the mask spectrum. The figure suggests there is a slight increase in the intensity difference among the harmonics since the regression line has a positive slope. The R^2 term is small (R^2 = 0.2322), so the general relationship does not hold. There are several locations where the attenuation is above 5 dB or more, specifically H7, H8, H9 and H12. These harmonics are close to F3, so the intensity differences displayed may attribute to perceptual changes in the vowel.

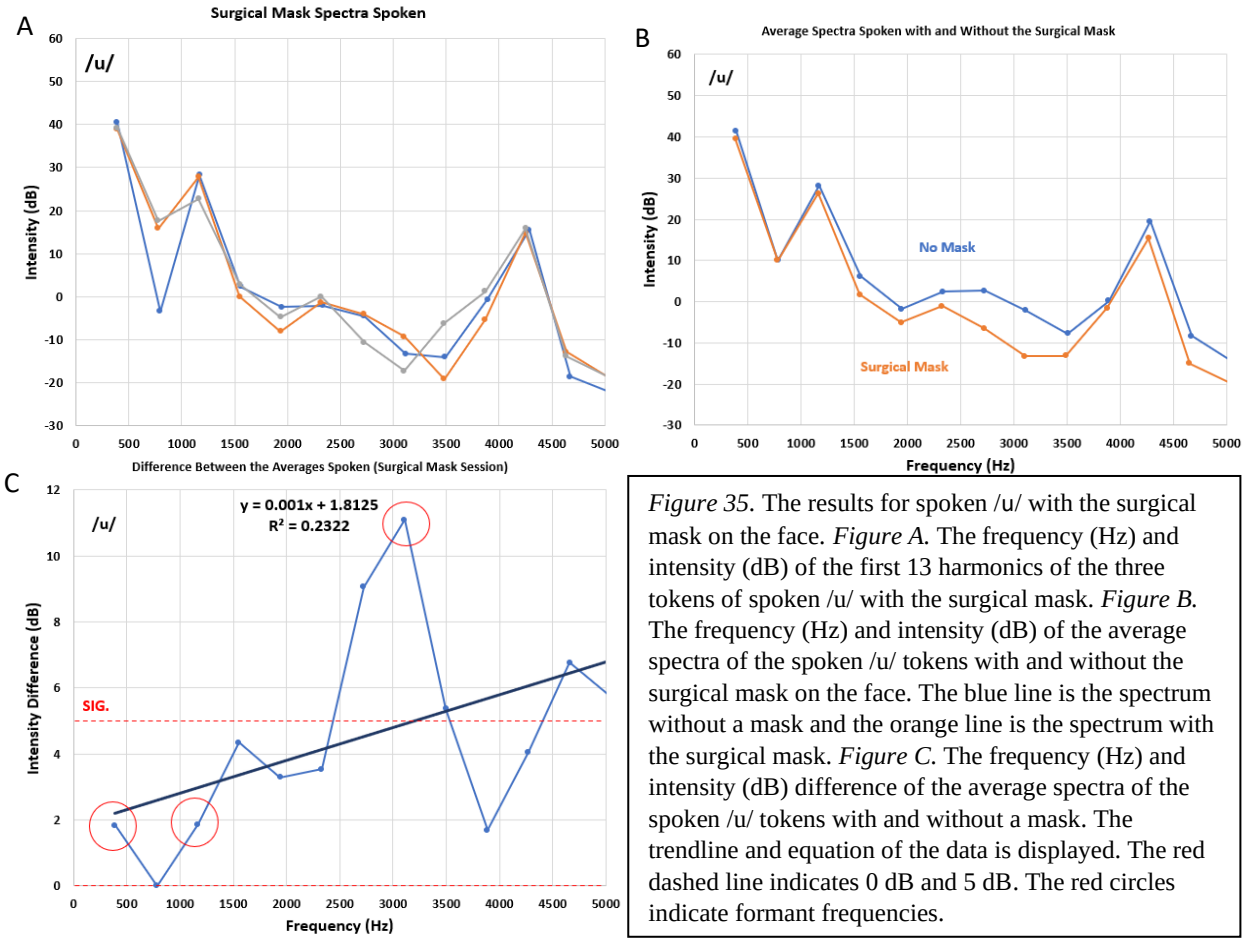

Figure 35. The results for spoken /u/ with the surgical mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of spoken /u/ with the surgical mask. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the spoken /u/ tokens with and without the surgical mask on the face. The blue line is the spectrum without a mask and the orange line is the spectrum with the surgical mask. *Figure C.* The frequency (Hz) and intensity (dB) difference of the average spectra of the spoken /u/ tokens with and without a mask. The trendline and equation of the data is displayed. The red dashed line indicates 0 dB and 5 dB. The red circles indicate formant frequencies.

Vowels on G4 /a,i,u/

Figure 36 shows the results and comparisons of /a/ sung on G4 with the surgical mask on the face. *Figure A* shows the results for the spectra of the tokens for sustained /a/ sung on G4 with the surgical mask on the face. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 413.1 Hz and the maximum frequency difference among the f_o was 14.5 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.7 dB (near 3300 Hz or H8) to 28.9 dB (near 4500 Hz or H11). *Figure B* displays the average spectra of /a/ sung on G4 with the surgical mask and without the mask on the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 4.27 Hz and the

difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.7 dB (near 400 Hz or H1) to 12.67 dB (near 5400 Hz or H13). *Figure C* displays the difference between the no mask average spectrum and the mask average spectrum in intensity (dB) for /a/ sung on G4. The difference of the f_o was 1.7 dB, with the no mask spectrum higher than the mask spectrum. The figure suggests there is a slight increase in the intensity difference among the harmonics since the regression line has a positive slope. The R^2 term is small ($R^2 = 0.0473$), so the general relationship does not hold. There are several locations where the attenuation is above 5 dB or more, specifically H3, H9, H12, and H13. These harmonics are close to F2 and F3, so the intensity differences displayed may attribute to perceptual changes in the vowel when the mask is worn.

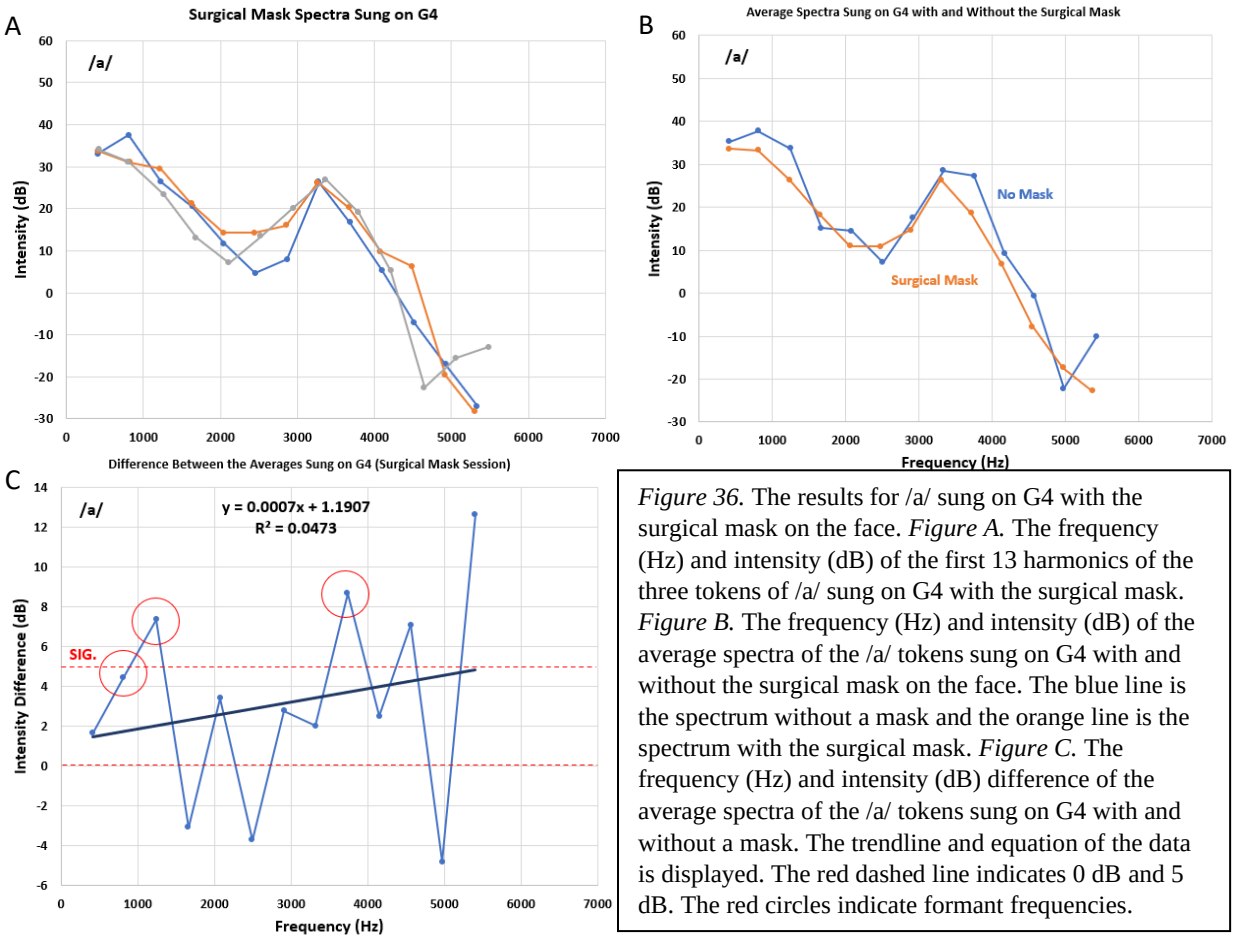

Figure 36. The results for /a/ sung on G4 with the surgical mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of /a/ sung on G4 with the surgical mask. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the /a/ tokens sung on G4 with and without the surgical mask on the face. The blue line is the spectrum without a mask and the orange line is the spectrum with the surgical mask. *Figure C.* The frequency (Hz) and intensity (dB) difference of the average spectra of the /a/ tokens sung on G4 with and without a mask. The trendline and equation of the data is displayed. The red dashed line indicates 0 dB and 5 dB. The red circles indicate formant frequencies.

Figure 37 shows the results and comparisons of /i/ sung on G4 with the surgical mask on the face. *Figure A* shows the results for the spectra of the tokens for sustained /i/ sung on G4 with the surgical mask on the face. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 408.58 Hz and the maximum frequency difference among the f_o was 7.11 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.5 (near 400 Hz or H1) to 14.5 dB (4900 Hz or H12). *Figure B* displays the average spectra of sustained /i/ sung on G4 with the surgical mask and without the mask on the face. The blue line represents the spectrum without the mask and the orange line represents the

spectrum with the mask. The maximum frequency difference among the f_o was 5.07 Hz and the difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.0 dB (near 400 Hz or H1) to 6.27 dB (near 2900 Hz or H7). *Figure C* displays the difference between the no mask average spectrum and the mask average spectrum in intensity (dB) for /i/ sung on G4. The difference of the f_o was 0.0 dB. The figure suggests there is a slight decrease in the intensity difference among the harmonics since the regression line has a negative slope. The R^2 term is small ($R^2 = 0.0038$), so the general relationship does not hold. There is a location where the attenuation is above 5 dB or more, specifically at H7. H7 is close to F3, so the intensity differences displayed may attribute to perceptual changes in the vowel when the mask is worn.

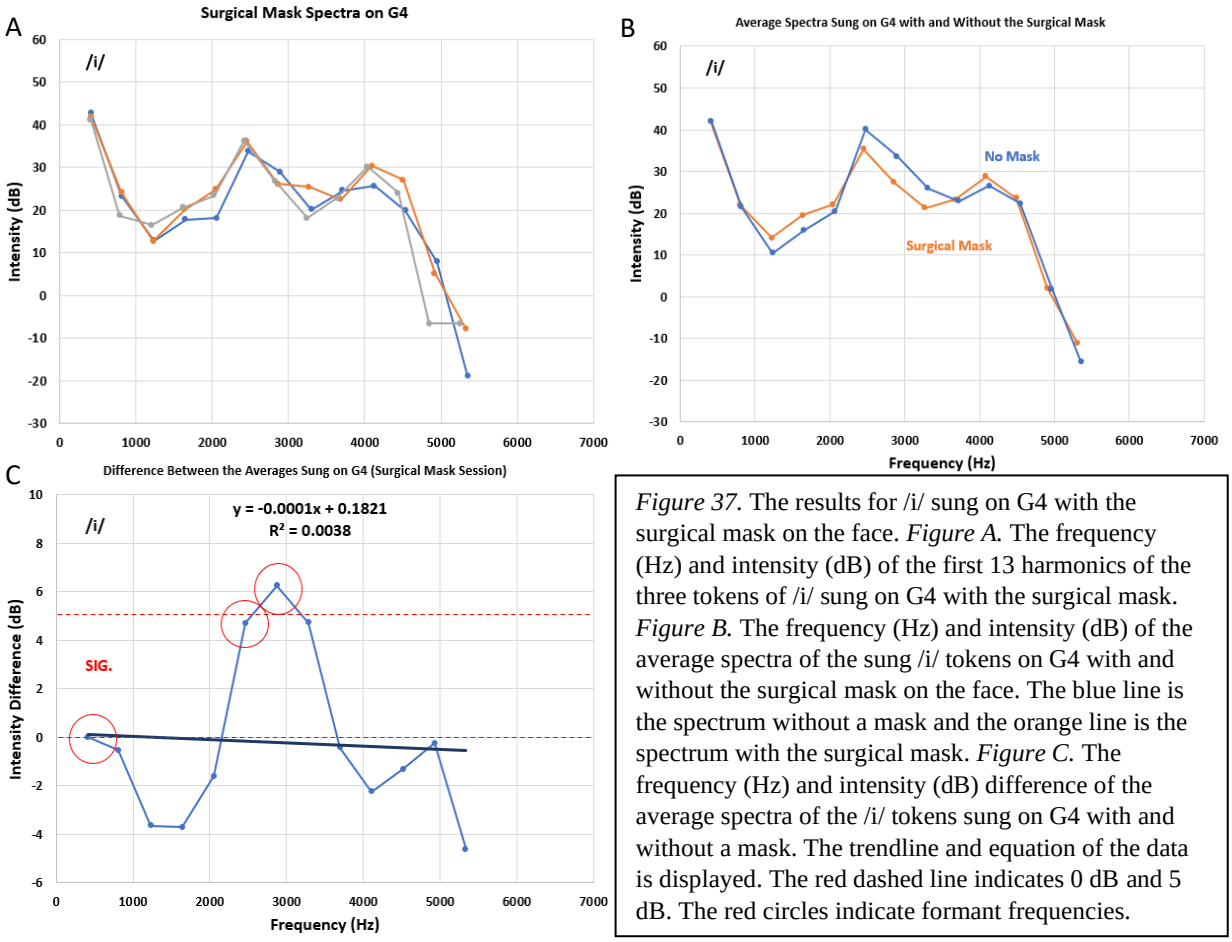

Figure 37. The results for /i/ sung on G4 with the surgical mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of /i/ sung on G4 with the surgical mask. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the sung /i/ tokens on G4 with and without the surgical mask on the face. The blue line is the spectrum without a mask and the orange line is the spectrum with the surgical mask. *Figure C.* The frequency (Hz) and intensity (dB) difference of the average spectra of the /i/ tokens sung on G4 with and without a mask. The trendline and equation of the data is displayed. The red dashed line indicates 0 dB and 5 dB. The red circles indicate formant frequencies.

Figure 38 shows the results and comparisons of /u/ sung on G4 with the surgical mask on the face. *Figure A* shows the results for the spectra of the tokens for sustained /u/ sung on G4 with the surgical mask on the face. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 411.90 Hz and the maximum frequency difference among the f_o was 7.36 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.5 dB (near H1 and H2) to 12.45 dB (near 4900 Hz or H12). *Figure B* displays the average spectra of sustained /u/ sung on G4 with the surgical mask and without the mask on

the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 1.29 Hz and the difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.1 dB (near 400 Hz or H1) to 7.6 dB (near 5300 Hz or H12). *Figure C* displays the difference in between the no mask average spectrum and the mask average spectrum in intensity (dB) for /u/ sung on G4. The difference of the f_o was 0.1 dB, with the mask spectrum higher than the no mask spectrum. The figure suggests there is a slight increase in the intensity difference among the harmonics since the regression line has a positive slope. The R^2 term is small ($R^2 = 0.3056$), so the general relationship does not hold. There are several locations where the attenuation is above 5 dB or more, specifically H7 and H13. These harmonics are not close to a formant, so the intensity differences displayed may not attribute to perceptual changes in the vowel.

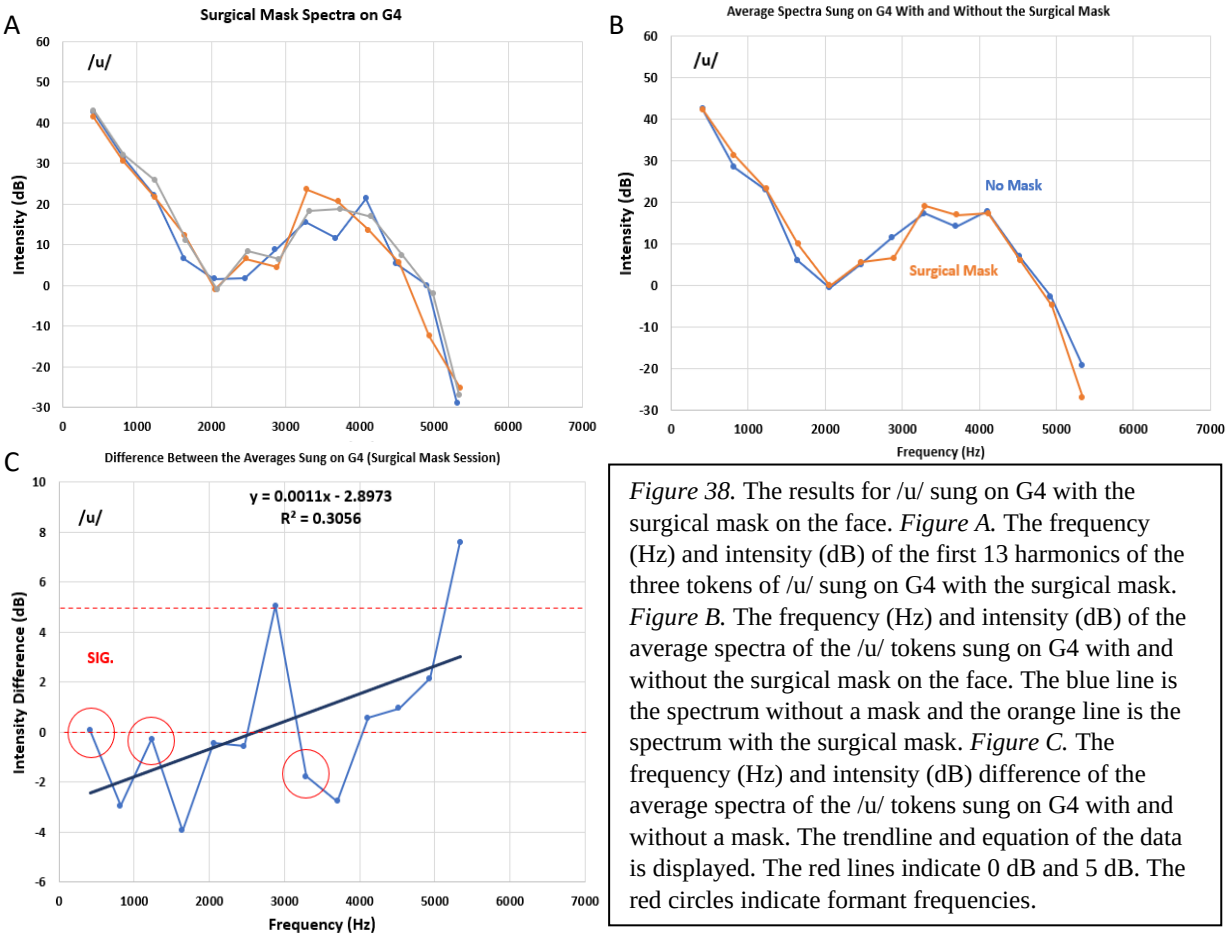

Figure 38. The results for /u/ sung on G4 with the surgical mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 13 harmonics of the three tokens of /u/ sung on G4 with the surgical mask. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the /u/ tokens sung on G4 with and without the surgical mask on the face. The blue line is the spectrum without a mask and the orange line is the spectrum with the surgical mask. *Figure C.* The frequency (Hz) and intensity (dB) difference of the average spectra of the /u/ tokens sung on G4 with and without a mask. The trendline and equation of the data is displayed. The red lines indicate 0 dB and 5 dB. The red circles indicate formant frequencies.

Vowels on C5 /a,i,u/

Figure 39 shows the results and comparisons of /a/ sung on C5 with the surgical mask on the face. *Figure A* shows the results for the spectra of the tokens for sustained /a/ sung on C5 with the surgical mask on the face. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 539.63 Hz and the maximum frequency difference among the f_o was 13.6 Hz. The difference range for the intensity of the corresponding harmonics for the entire

spectrum was 1.1 dB (near 2100 Hz or H4) to 16.2 dB (near 6000 Hz or H11). *Figure B* displays the average spectra of sustained /a/ sung on C5 with the surgical mask and without the mask on the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 2.24 Hz and the difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.43 dB (near 2700 Hz or H5) to 11.37 dB (near 5400 Hz or H10). *Figure C* displays the difference between the no mask average spectrum and the mask average spectrum in intensity (dB) for /a/ sung on C5. The difference of the f_o was 1.4 dB, with the no mask spectrum higher than the mask spectrum. The figure suggests there is a slight increase in the intensity difference among the harmonics since the regression line has a positive slope. The R^2 term is small (R^2 = 0.1537), so the general relationship does not hold. There are several locations where the attenuation is above 5 dB or more, specifically H3, H7, H9 and H10. Some of these harmonics are close to F2 and F3, so the intensity differences displayed may attribute to perceptual changes in the vowel.

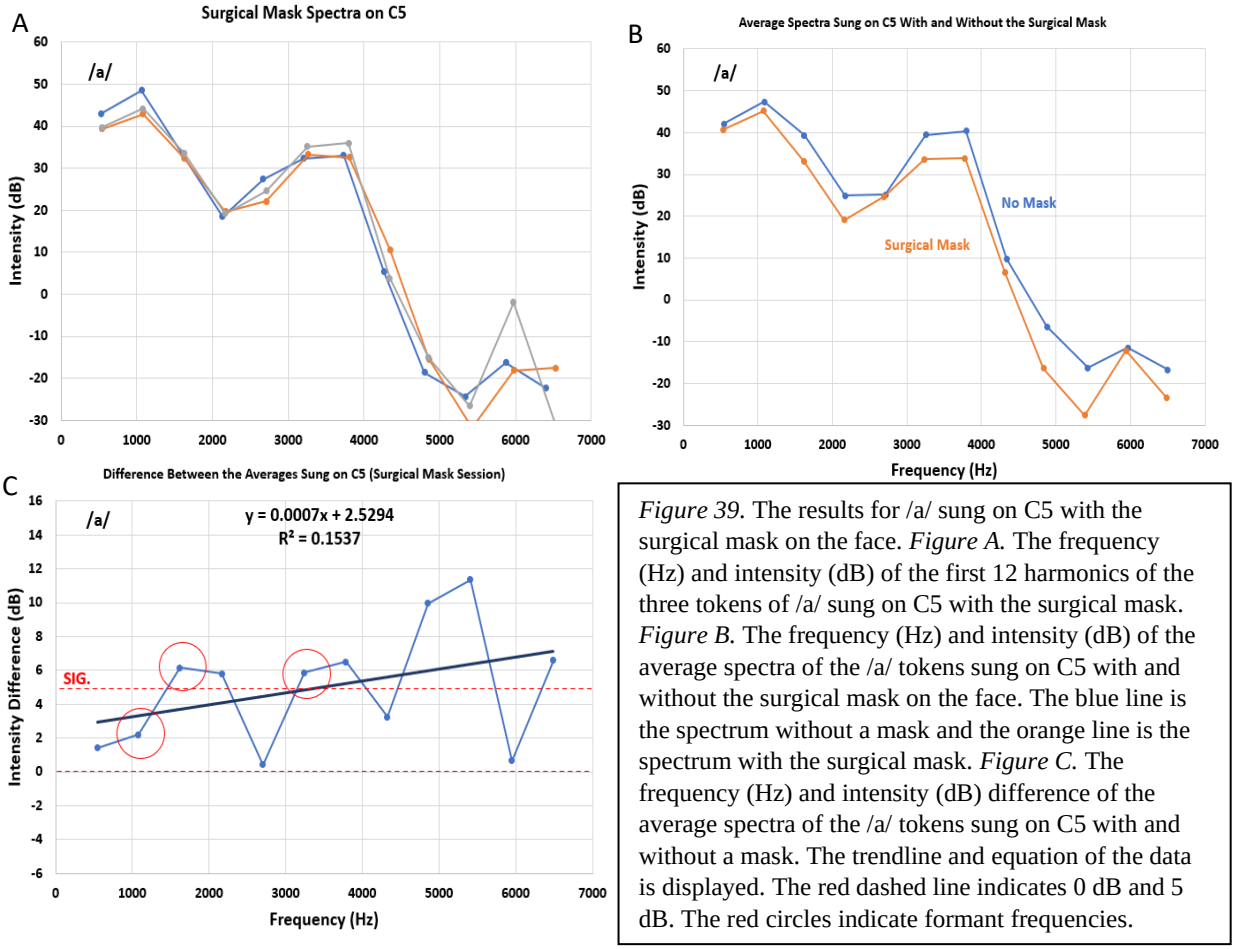

Figure 39. The results for /a/ sung on C5 with the surgical mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 12 harmonics of the three tokens of /a/ sung on C5 with the surgical mask. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the /a/ tokens sung on C5 with and without the surgical mask on the face. The blue line is the spectrum without a mask and the orange line is the spectrum with the surgical mask. *Figure C.* The frequency (Hz) and intensity (dB) difference of the average spectra of the /a/ tokens sung on C5 with and without a mask. The trendline and equation of the data is displayed. The red dashed line indicates 0 dB and 5 dB. The red circles indicate formant frequencies.

Figure 40 shows the results and comparisons of /i/ sung on C5 with the surgical mask on the face. *Figure A* shows the results for the spectra of the tokens for sustained /i/ sung on C5 with the surgical mask on the face. The blue, orange, and gray lines depict separate tokens. The

average f_o of the tokens was 542.96 Hz and the maximum frequency difference among the f_o was 6.0 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 1.1 dB (near 1600 or H3) to 15.9 dB (4300 Hz or H8). *Figure B* displays the average spectra of sustained /u/ sung on C5 with the surgical mask and without the mask on the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 5.64 Hz and the difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.3 dB (near 4300 Hz or H8) to 10.23 dB (near 2700 Hz or H5). *Figure C* displays the difference between the no mask average spectrum and the mask average spectrum in intensity (dB) for /i/ sung on C5. The intensity difference of the f_o was 2.6 dB, with the no mask spectrum higher than the mask spectrum. The figure suggests there is a slight decrease in the intensity difference among the harmonics since the regression line has a negative slope. The R^2 term is small (R^2 = 0.0484), so the general relationship does not hold. There are several locations where the attenuation is above 5 dB or more, specifically H4, H5, H6, H10, and H11. Some of these harmonics are close to F2 and F3, so the intensity differences displayed may attribute to perceptual changes in the vowel.

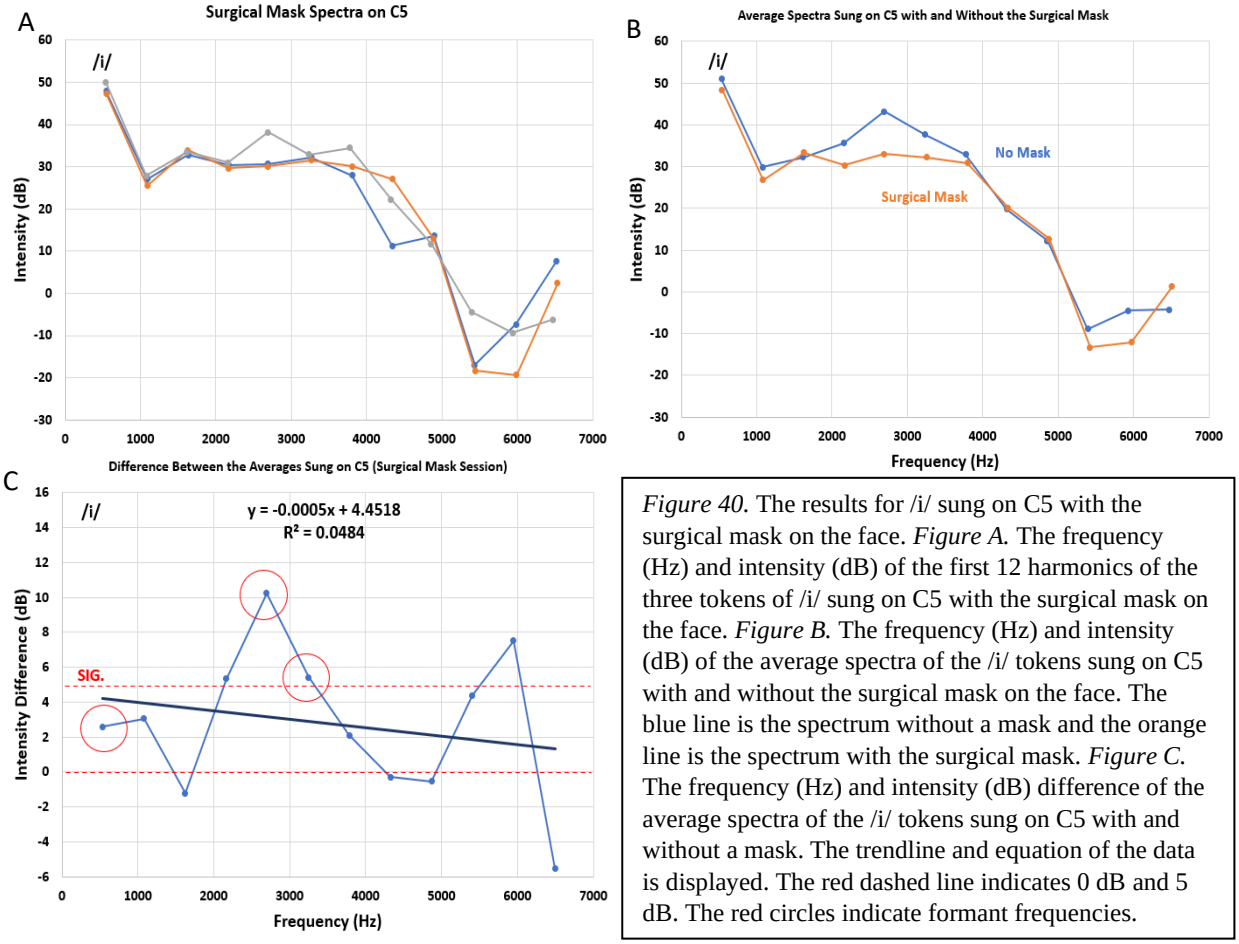

Figure 40. The results for /i/ sung on C5 with the surgical mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 12 harmonics of the three tokens of /i/ sung on C5 with the surgical mask on the face. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the /i/ tokens sung on C5 with and without the surgical mask on the face. The blue line is the spectrum without a mask and the orange line is the spectrum with the surgical mask. *Figure C.* The frequency (Hz) and intensity (dB) difference of the average spectra of the /i/ tokens sung on C5 with and without a mask. The trendline and equation of the data is displayed. The red dashed line indicates 0 dB and 5 dB. The red circles indicate formant frequencies.

Figure 41 shows the results and comparisons of /u/ sung on C5 with the surgical mask on the face. *Figure A* shows the results for the spectra of the tokens for sustained /u/ sung on C5

with the surgical mask on the face. The blue, orange, and gray lines depict separate tokens. The average f_o of the tokens was 542.69 Hz and the maximum frequency difference among the f_o was 19.29 Hz. The difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.9 dB (near 3300 Hz or H6) to 11.9 dB (6500 Hz or H12). *Figure B* displays the average spectra of /sustained /u/ sung on C5 with the surgical mask and without the mask on the face. The blue line represents the spectrum without the mask and the orange line represents the spectrum with the mask. The maximum frequency difference among the f_o was 0.10 Hz and the difference range for the intensity of the corresponding harmonics for the entire spectrum was 0.2 dB (near 2100 Hz or H4) to 13.33 dB (near 6500 Hz or H12). *Figure C* displays the difference between the no mask average spectrum and the mask average spectrum in intensity (dB) for a sustained /u/ sung on C5. The intensity difference of the f_o was 0.4 dB, with the no mask spectrum higher than the mask spectrum. The figure suggests there is a slight increase in the intensity difference among the harmonics since the regression line has a positive slope. The R^2 term is higher ($R^2 = 0.4116$), so the general relationship does hold. However, this is due to the large intensity difference of H12, which has a negative intensity and therefore is not perceptually relevant. While several locations have high attenuation (namely at H6, H11, and H12), the high variability of the mask spectra may cause this to be insignificant.

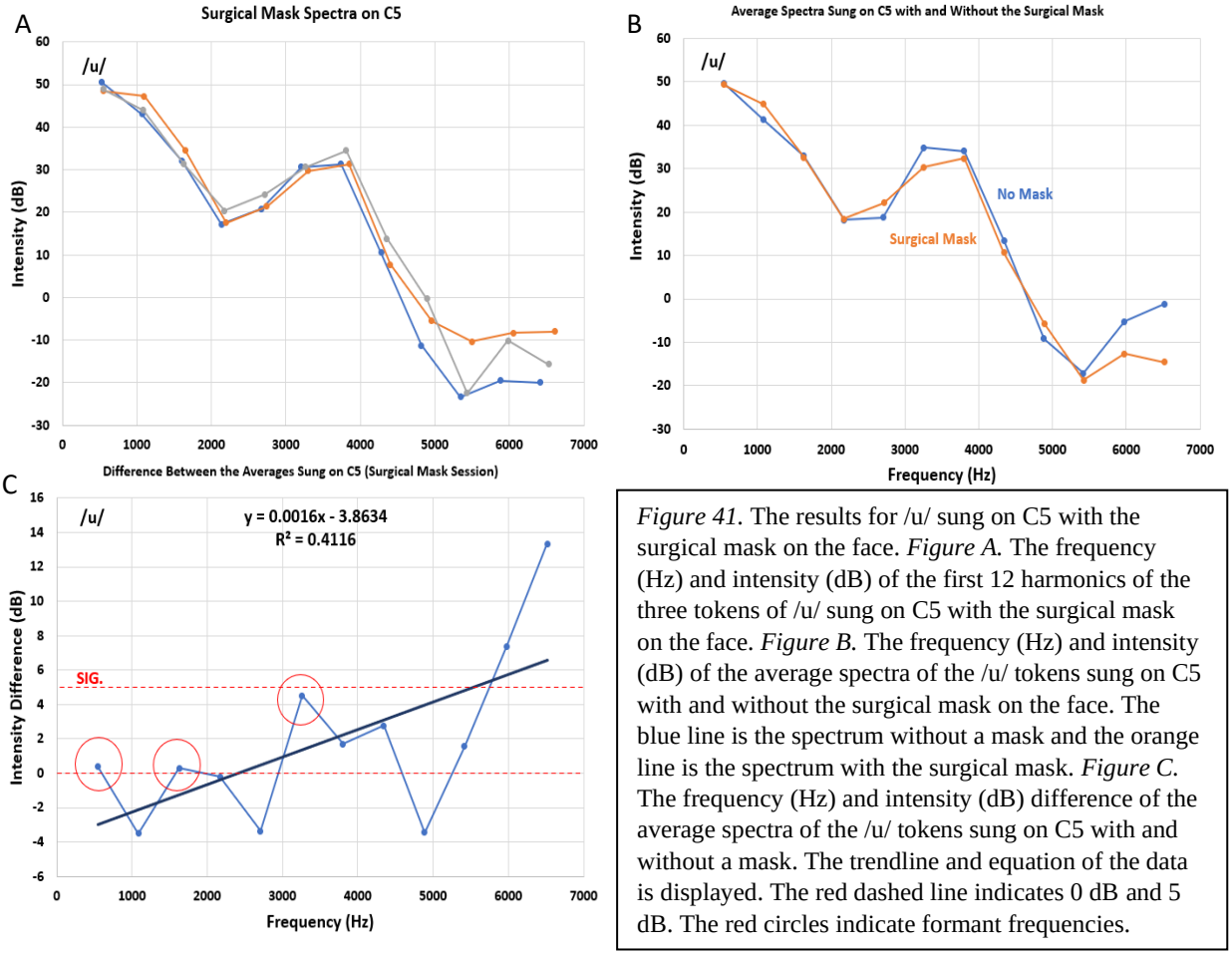

Figure 41. The results for /u/ sung on C5 with the surgical mask on the face. *Figure A.* The frequency (Hz) and intensity (dB) of the first 12 harmonics of the three tokens of /u/ sung on C5 with the surgical mask on the face. *Figure B.* The frequency (Hz) and intensity (dB) of the average spectra of the /u/ tokens sung on C5 with and without the surgical mask on the face. The blue line is the spectrum without a mask and the orange line is the spectrum with the surgical mask. *Figure C.* The frequency (Hz) and intensity (dB) difference of the average spectra of the /u/ tokens sung on C5 with and without a mask. The trendline and equation of the data is displayed. The red dashed line indicates 0 dB and 5 dB. The red circles indicate formant frequencies.

LTAS Comparisons /s, ʃ/

Figure 42 shows the results and LTAS comparisons of /s/ with the surgical mask worn on the face. *Figure A* displays the trial 1 LTAS for /s/ without a mask and with the surgical mask. The blue line indicates the LTAS without a mask and the red line indicates the LTAS with the surgical mask on the face. The main difference in spectra lies in the salient frequency location for an /s/ production, between 5000 Hz and 7000 Hz. *Figure B* displays the intensity difference between the no mask LTAS and surgical mask LTAS for trial 1 at every 1000 Hz. The recordings for trial 2 were corrupted most likely due to internal electronic problems that made the results insignificant. The maximum difference for the entire spectrum was 11.0 dB at 6000 Hz. The salient location for /s/ is located around 6000 Hz, thus the highest attenuation at this area may create a perceptual quality change between the no mask and masked conditions. That is, the perception of /s/ may be greatly reduced during speech due to being about 11 dB reduced when the mask was worn.

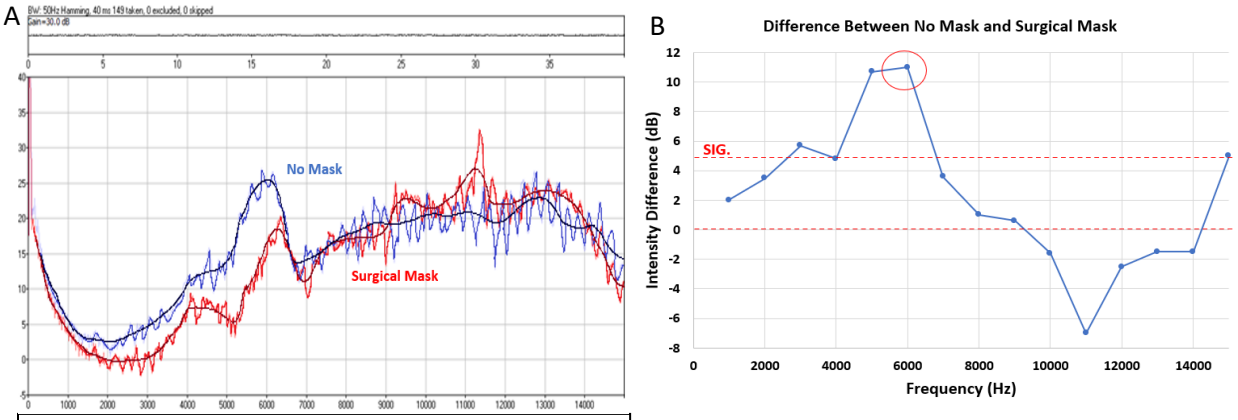

Figure 42. The results for /s/ with the surgical mask on the face. *Figure A.* The no mask LTAS in blue and the surgical mask LTAS in red for /s/. The continuous lines are hand-drawn as visual averages. *Figure C.* The intensity difference of the no mask and surgical mask LTAS for trial 1 of /s/. The red lines indicate 0 dB and 5 dB. The red circle indicates the salient location of /s/.

Figure 43 shows the results and LTAS comparisons of /ʃ/ with the surgical mask worn on the face. *Figure A* displays the trial 1 LTAS for / ʃ/ without a mask and with the surgical mask. The blue line indicates the LTAS without a mask and the red line indicates the LTAS with the surgical mask on the face. The main difference in spectra lies in the salient frequency location for an /ʃ/ production, between 3000 Hz and 5000 Hz. *Figure B* displays the trial 2 LTAS for / ʃ/ without a mask and with the surgical mask. The blue line indicates the LTAS without a mask and the red line indicates the LTAS with the surgical mask on the face. The main difference in

spectra lies in the salient frequency location for an /ʃ/ production, between 3000 Hz and 5000 Hz. *Figure C* displays the intensity difference between the no mask LTAS and surgical mask LTAS for trial 1 at every 1000 Hz (blue line) and for trial 2 at every 1000 Hz (orange line). The maximum difference for the entire spectrum was 8.5 dB at 13000 Hz for trial 1 and 8.0 dB at 6000 Hz for trial 2. The salient location for /ʃ/ is located around 3000 Hz to 5000 Hz, thus the attenuation at this area may create a perceptual quality change between the no mask and masked conditions.

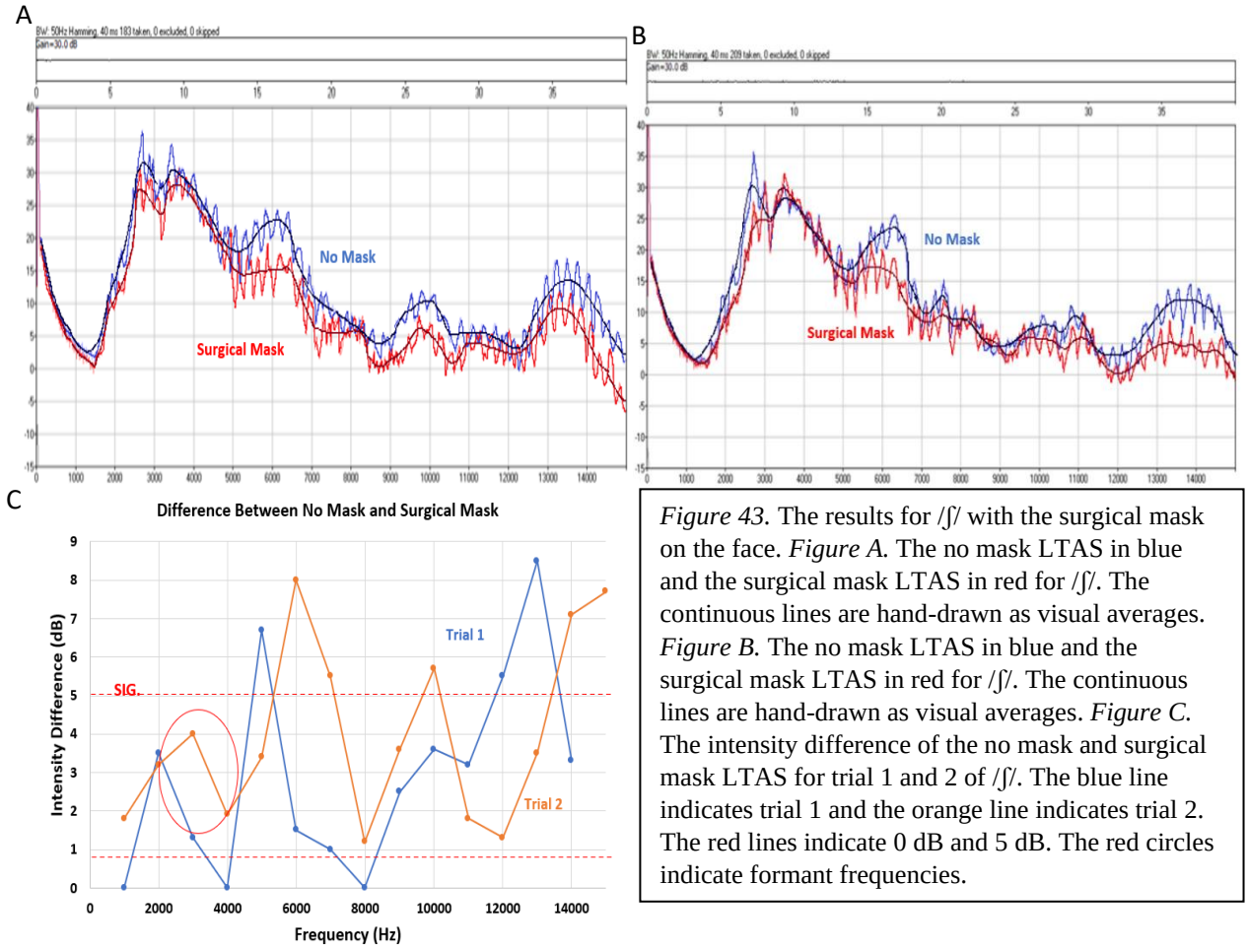

Figure 43. The results for /ʃ/ with the surgical mask on the face. *Figure A.* The no mask LTAS in blue and the surgical mask LTAS in red for /ʃ/. The continuous lines are hand-drawn as visual averages. *Figure B.* The no mask LTAS in blue and the surgical mask LTAS in red for /ʃ/. The continuous lines are hand-drawn as visual averages. *Figure C.* The intensity difference of the no mask and surgical mask LTAS for trial 1 and 2 of /ʃ/. The blue line indicates trial 1 and the orange line indicates trial 2. The red lines indicate 0 dB and 5 dB. The red circles indicate formant frequencies.

DISCUSSION

Spectral Attenuation for Vowels /a,i,u/

Due to human variability, we have decided to label 5 dB or more significant attenuation difference. **Table 1** displays the average attenuation of the spoken vowels /a,i,u/ near F1, F2, and F3 for the N95 mask, the Singer's Mask, and the surgical mask. Based on the given attenuation difference for significance, only the N95 mask had a significant difference of 6.7 dB that could account for perceptual changes hear when the mask is worn.

Table 1. The average attenuation of spoken /a,i,u/ at F1, F2, and F3 for the N95 mask, the Singer's Mask, and the surgical mask.

	/a/	/i/	/u/
N95 Mask	**3.5 dB**	**6.7 dB**	**3.2 dB**
Singer's Mask	**3.6 dB**	**4.6 dB**	**3.1 dB**
Surgical Mask	**2.6 dB**	**3.26 dB**	**4.9 dB**

Because previous studies have found masks to act as a low pass filter, **Table 2** displays the average attenuation of the spoken vowels /a,i,u/ near F2 and F3 for the N95 mask, the Singer's Mask, and the surgical mask. This was to calculate the average attenuation of formants above 1000 Hz, which is where attenuation occurred for all tokens. Based on the given attenuation difference for significance, the N95 mask displayed significant attenuation for /a,i/, the Singer's Mask had attenuation with /a,i/ and the surgical mask displayed attenuation for /u/. Based on these values, the surgical mask had the least amount of attenuation and the N95 mask had the most attenuation.

Table 2. The average attenuation of spoken /a,i,u/ at F2 and F3 for the N95 mask, the Singer's Mask, and the surgical mask.

	/a/	/i/	/u/
N95 Mask	**5.75 dB**	**10.0 dB**	**4.3 dB**
Singer's Mask	**5.3 dB**	**6.6 dB**	**3.2 dB**
Surgical Mask	**3.65 dB**	**4.7 dB**	**6.5 dB**

Table 3 displays the average attenuation of the vowels /a,i,u/ sung on G4 near F1, F2, and F3 for the N95 mask, the Singer's Mask, and the surgical mask. Based on the given attenuation difference for significance (5 dB), only /a/ with the surgical mask displayed significant attenuation. For the vowels sung on G4, the attenuation is significantly lower than the spoken vowels, which may mean that the masks attenuate singing less than speech. There are several areas where the mask spectra had a higher intensity than the no mask spectra, as shown by the negative values in the table below. Because these values are near formants, the masks may have created a resonance that increased the intensity around the formants.

Table 3. The average attenuation of /a,i,u/ sung on G4 at F1, F2, and F3 for the N95 mask, the Singer's Mask, and the surgical mask.

	/a/	/i/	/u/
N95 Mask	**2.66 dB**	**0.22 dB**	**-1.14 dB**
Singer's Mask	**2.95 dB**	**1.79 dB**	**1.48 dB**
Surgical Mask	**6.85 dB**	**3.66 dB**	**-0.65 dB**

Table 4 displays the average attenuation of the vowels /a,i,u/ sung on G4 near F2 and F3 for the N95 mask, the Singer's Mask, and the surgical mask. Based on the given attenuation difference for significance (5 dB), there was significant attenuation with the N95 mask for /a/, as well as with the surgical mask for /a,i/. While the surgical mask performed worse acoustically compared to the spoken vowels, there was overall less attenuation among the N95 mask and Singer's Mask.

Table 4. The average attenuation of /a,i,u/ sung on G4 at F2 and F3 for the N95 mask, the Singer's Mask, and the surgical mask.

	/a/	/i/	/u/
N95 Mask	**5.85 dB**	**1.27 dB**	**-2.08 dB**
Singer's Mask	**3.62 dB**	**2.59 dB**	**2.53 dB**
Surgical Mask	**8.03 dB**	**5.49 dB**	**-1.02 dB**

Table 5 displays the average attenuation of the vowels /a,i,u/ sung on C5 near F1, F2, and F3 for the N95 mask, the Singer's Mask, and the surgical mask. Based on the given attenuation difference for significance (5 dB), there was no significant attenuation for the N95 mask. The Singer's Mask had significant attenuation for /i/, and the surgical mask had significant attenuation also for /i/.

Table 5. The average attenuation of /a,i,u/ sung on C5 at F1, F2, and F3 for the N95 mask, the Singer's Mask, and the surgical mask

	/a/	/i/	/u/
N95 Mask	**-0.18 dB**	**0.73 dB**	**-0.37 dB**
Singer's Mask	**2.12 dB**	**5.32 dB**	**1.18 dB**

Surgical Mask	**4.96 dB**	**6.09 dB**	**1.73 dB**

Table 6 displays the average attenuation of the vowels /a,i,u/ sung on C5 near F2 and F3 for the N95 mask, the Singer's Mask, and the surgical mask. Based on the given attenuation difference for significance (5 dB), the N95 mask had no significant attenuation. The Singer's Mask had significant attenuation for /i/. The surgical mask had significant attenuation for /a,i/. Thus, for C5 the surgical mask had the greatest amount of significant attenuation, and the N95 mask had the least amount of significant attenuation.

Table 6. The average attenuation of /a,i,u/ sung on C5 at F2 and F3 for the N95 mask, the Singer's Mask, and the surgical mask.

	/a/	**/i/**	**/u/**
N95 Mask	**-0.24 dB**	**1.07 dB**	**1.17 dB**
Singer's Mask	**2.02 dB**	**7.07 dB**	**2.2 dB**
Surgical Mask	**6.34 dB**	**7.83 dB**	**2.4 dB**

The average acoustic spectra for the vowels /a,i,u/ with the masks worn on the face all had a pattern of high frequency attenuation, starting around 1000 Hz when compared to the average spectra with no mask. The fundamental frequency (f_o) and the second harmonic (H2) typically had little to no attenuation across all the mask spectra, indicating that the masks act as a "low pass filter" because the lowest harmonics below 1000 Hz tended to not be affected. Based on the tables above, for speech the N95 mask performed the worse acoustically, while the surgical mask performed the best. Interestingly, this did not transfer to the sung vowels. For G4, the surgical mask performed the worst acoustically with the Singer's Mask performing the best. Similarly, for C5, the surgical mask performed the worst acoustically with the N95 mask performing the best.

Figure 44 displays the intensity differences at F1, F2, and F3 for /a,i,u/ spoken with the N95 mask, the Singer's Mask, and the surgical mask. The slight positive slope and the range of intensity differences from 1 dB to 7 dB suggest that as the frequency increases, the intensity difference also increases. Thus, showing the low pass filter characteristics of the masks.

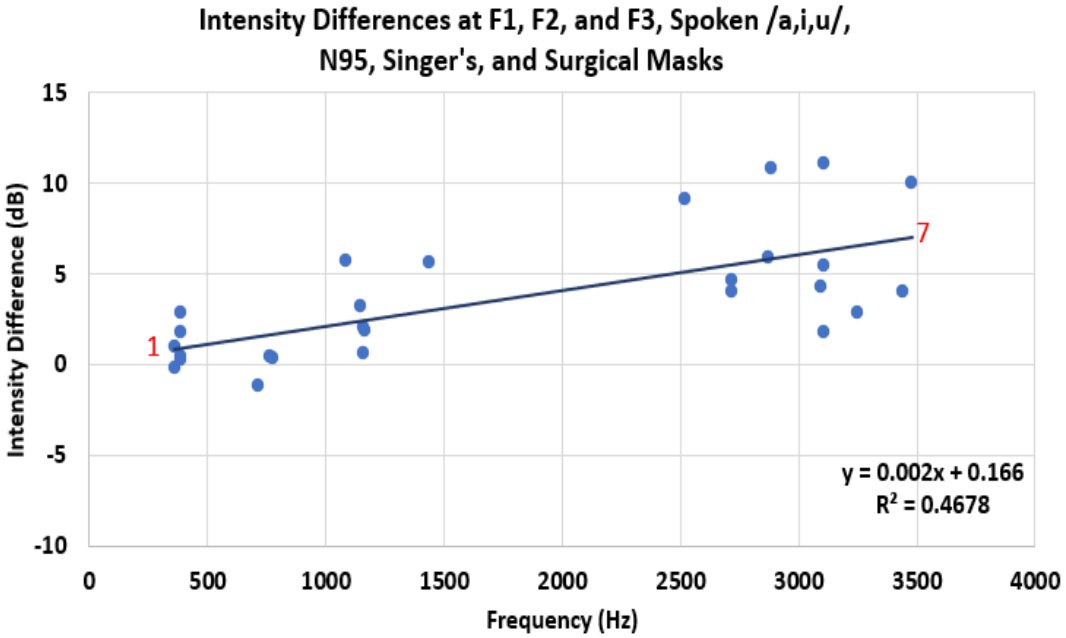

Figure 44. The intensity differences (dB) at F1, F2, and F3 for /a,i,u/ spoken with the N95 mask, the Singer's Mask, and the surgical mask.

Spectral Attenuation for Fricatives /s, ʃ/

The intensity difference from the LTAS with the mask and without the mask showed high attenuation typically around 5000 Hz to 6000 Hz for /s/. As these frequencies are around the salient location of /s/ production and perception, this attenuation is significant to the perceptual changes that the masks cause. That is, attenuations up to 16 dB that were found across the masks should create a perceptually evident change, leading to an increased inability to hear the /s/ sound. While there are areas where the mask LTAS is higher in intensity than the no mask LTAS, it is not near its salient location (5000 Hz to 7000 Hz). Thus, these differences are most likely not perceptually important. **Table 7** displays the average intensity difference of trial 1 and trial 2 for /s/ for the N95 mask, the Singer's Mask, and the surgical mask. Based on the given attenuation difference for significance (5 dB), all three masks have significant attenuation for /s/. The N95 mask had the most attenuation, while the surgical mask had the least attenuation.

Table 7. The average attenuation of trial 1 and trial 2 for /s/ for the N95 mask, the Singer's Mask, and the surgical mask.

N95 Mask	**14.25 dB**
Singer's Mask	**10.15 dB**

Surgical Mask	**8.1 dB**

The attenuation of /ʃ/ was variable, but there was overall less attenuation at the salient location of /ʃ/ (3000 Hz to 5000 Hz) compared to the attenuation of /s/ at its salient location for all three masks. **Table 8** displays the average intensity difference of trial 1 and trial 2 for /ʃ/ for the N95 mask, the Singer's Mask, and the surgical mask. Based on the given attenuation difference for significance (5 dB), all three masks have significant attenuation for /ʃ/. Similar to /s/, the N95 mask had the most attenuation, while the Singer's Mask had the least attenuation.

Table 8. The average attenuation of trial 1 and trial 2 for /ʃ/ for the N95 mask, the Singer's Mask, and the surgical mask.

N95 Mask	**6.9 dB**
Singer's Mask	**4.5 dB**
Surgical Mask	**4.85 dB**

Figure 45 displays the intensity difference at the salient locations of /s, ʃ/ with the N95 mask, the Singer's Mask, and the surgical mask. The slight positive slope and intensity difference range of 5 dB to 12 dB indicate that as the frequency increases, the intensity difference also increases. Thus, showing the low pass filter characteristics of the masks.

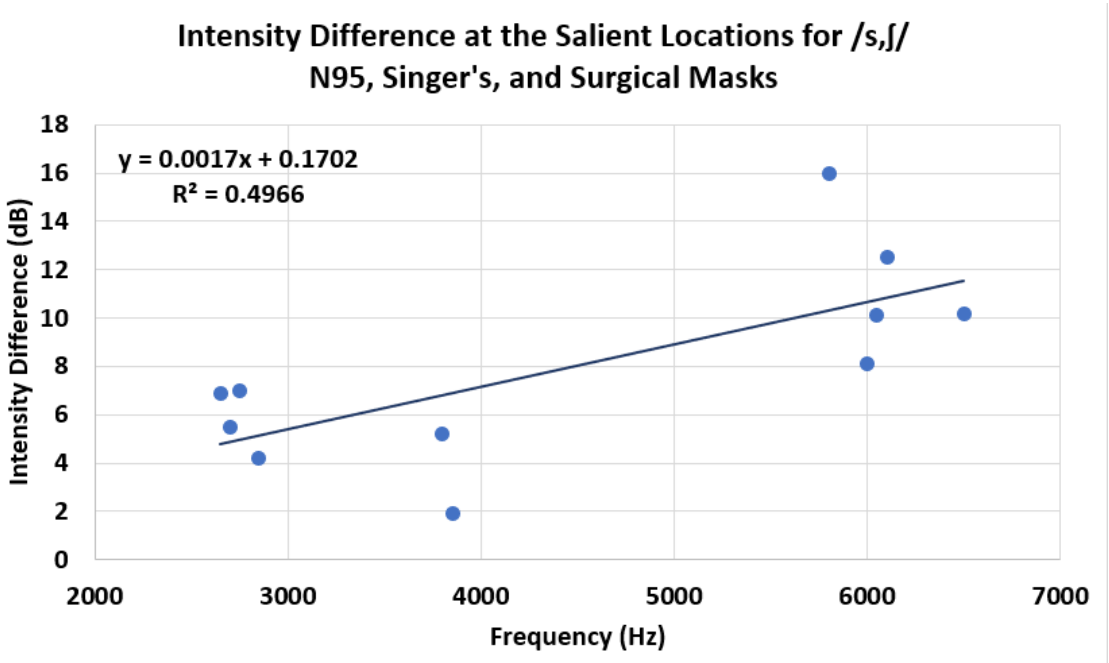

Figure 45. The intensity differences (dB) at the salient locations of /s,ʃ/ with the N95 mask, the Singer's Mask, and the surgical mask.

Figure 46 displays the intensity difference at F1, F2, and F3 for /a,i,u/ spoken and the salient locations of /s, ʃ/ with the N95 mask, the Singer's Mask, and the surgical mask. The slight positive slope and intensity difference range of 1 dB to 12 dB indicate that as the frequency increases, the intensity difference also increases. Thus, showing the low pass filter characteristics of the masks.

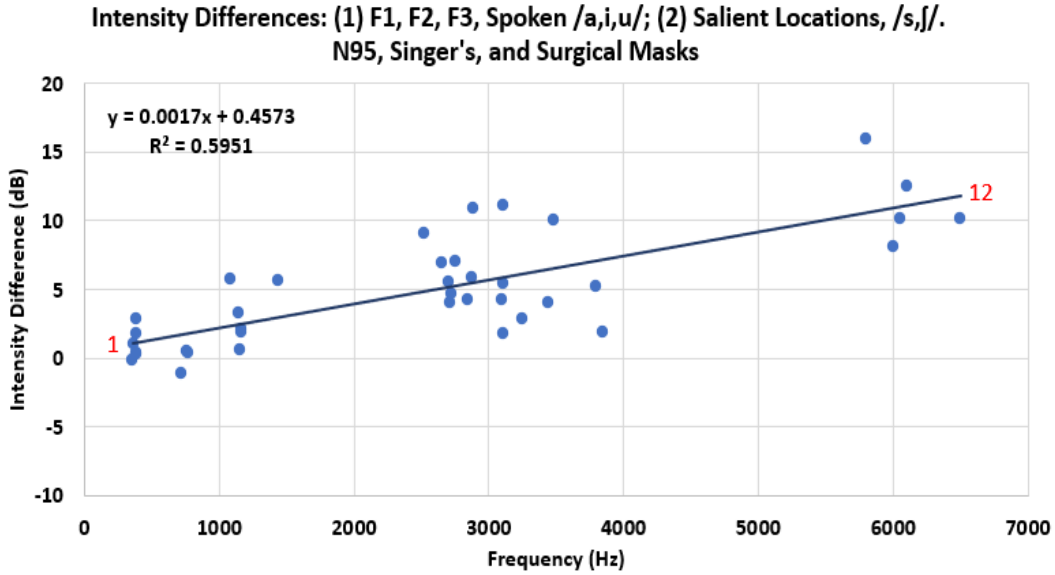

Figure 46. The intensity differences (dB) at F1, F2, and F3 for /a,i,u/ the salient locations of /s,ʃ/ with the N95 mask, the Singer's Mask, and the surgical mask.

Comparisons Across Other Studies

Past studies on the acoustic filtering of masks have noted that masks act as a low pass filter, meaning that the high frequencies are attenuated (Goldin et al., 2020). In this study, masks also attenuated higher frequencies. A study by Corey et al. (2020) found that the masks had little effect under 1000 Hz, which was also seen in this study. The study reported a peak attenuation of around 4 dB for the surgical mask and an attenuation of around 6dB for the N95 mask, all of which was above 4000 Hz. Goldin et al. (2020) also conducted a study noting an attenuation range of 3 dB with a surgical mask to 12 dB using two N95 masks. The current study gave similar results, with the N95 mask having more attenuation than the surgical mask for spoken vowels and consonants. While previous studies only looked at overall mask attenuation, our study added the element of formants to determine if attenuation was perceptually relevant.

Clinical and Performance Relevance

This research project has given insight into which mask may allow for the best intelligibility for speech and singing. This is extremely important because mask wearing is now

commonplace for the public and the professions of Speech Language Pathology and Vocal Performance have been drastically changed to account for safety while speaking and singing. For speech, our results suggest that a surgical mask might be preferred for acoustic reasons. While the N95 mask performed the worse acoustically in this study, it may be the safest for protection against COVID-19. For singers, the Singer's Mask may be preferred for acoustic reasons, as it was the best for sung vowels and the second for consonants. Masks may affect sung vowels less than speech. We hypothesize that there is less attenuation during singing because the distribution of harmonics among the formants is wider (less match of harmonics to formants).

Limitations

The limitations of this study were primarily due to human variability in producing the vowels and consonants. Because of this limitation, we offer a generous attenuation difference that appears to be significant (5 dB). There was variability in the no mask and mask spectra across recording sessions. This may have been caused by small differences in effort level, lung pressure, vocal tract shape, placement of the mask on face, and the mask edge seal on the face. It is inadequate for only one human subject to have tested these masks due to variability of vowel and consonant production.

More research should be done on the acoustic filtering of masks in order to make more generalizations. Additional research should involve more tokens (5 to 7 rather than 3) and should involve a larger sample of speakers (n = >1). This would permit a statistical design rather than a central tendency single subject design.

References

Bottalico, P., Murgia, S., Puglisi, G., Astolfi, A., & Kirk, K., (2020) "Effect of masks on speech intelligibility in auralized classrooms", *The Journal of the Acoustical Society of America, 148,* 2878-2884. https://doi.org/10.1121/10.0002450.

Center for Devices and Radiological Health. (2020). N95 respirators, SURGICAL masks, and face masks. https://www.fda.gov/medical-devices/personal-protective-equipment-infection-control/n95-respirators-surgical-masks-and-face-masks

Corey, R., Jones, U. and Singer, A., (2020). Acoustic effects of medical, cloth, and transparent face masks on speech signals. *The Journal of the Acoustical Society of America. 148*(4). http://doi.org/10.1121/10.0002279.

Goldin, A., Weinstein, B., & Shiman, N. (2020). How Do Medical Masks Degrade Speech Reception? Speech blocked by surgical masks becomes a more important issue in the Era of COVID-19. Hearing Review, 27(5), 8–9.

Gupta, R. and Pack, M. (2020). "Masked Phonation: How are Our Voices Affected", PAVA Informant, 1(3).

Mittal, R., Ni, R & Seo, J. (2020). The flow of physics and COVID-19. Journal of Fluid Mechanics, 894 (F2). https://doi.org/10.1017/jfm.2020.330.

Using Masks for In-Person Service Delivery During COVID-19: What to Consider (2020). American Speech-Language-Hearing Association. https://www.asha.org/practice/using-masks-for-in-person-service-delivery-during-covid-19-what-to-consider/