Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Acoustic markers of vowels produced with different types of face masks

Georgios P. Georgiou a,b,*

a Department of Languages and Literature, University of Nicosia, Nicosia, Cyprus
b Department of Foreign Languages, RUDN University, Moscow, Russia

Abstract

The wide spread of SARS-CoV-2 led to the extensive use of face masks in public places. Although masks offer significant protection from infectious droplets, they also impact verbal communication by altering speech signal. The present study examines how two types of face masks affect the speech properties of vowels. Twenty speakers were recorded producing their native vowels in a /pVs/ context, maintaining a normal speaking rate. Speakers were asked to produce the vowels in three conditions: (a) with a surgical mask, (b) with a cotton mask, and (c) without a mask. The speakers’ output was analyzed through Praat speech acoustics software. We fitted three linear mixed-effects models to investigate the mask-wearing effects on the first formant ($F_1$), second formant ($F_2$), and duration of vowels. The results demonstrated that $F_1$ and duration of vowels remained intact in the masked conditions compared to the unmasked condition, while $F_2$ was altered for three out of five vowels (/e a u/) in the surgical mask and two out of five vowels (/e a/) in the cotton mask. So, both types of masks altered to some extent speech signal and they mostly affected the same vowel qualities. It is concluded that some acoustic properties are more sensitive than other to speech signal modification when speech is filtered through masks, while various sounds are affected in a different way. The findings may have significant implications for second/foreign language instructors who teach pronunciation and for speech therapists who teach sounds to individuals with language disorders.

1. Introduction

During the COVID-19 outbreak, face masks have become must-use equipment for billions of people in an attempt for the world nations to effectively restrain the spread of the disease [30]. The role of face masks, and especially certain types of masks such as surgical and N95, has been proven important for the reduction of virus transmission [8,26,31] since they prevent the emission of infected particles and droplets from expiratory activities such as breathing, coughing, sneezing, and speaking [21,15].

Although face masks offer protection from virus infection, they negatively affect verbal communication and consequently intelligibility [3,10,22,27]. Some studies have shown that speech signal is modified through masks [6,10,13,34,35], decreasing the transmission of speech by an estimated 3–4% [32]. Also, when a speaker wears a face mask, the listener cannot follow the articulatory gestures of the speaker (e.g., lips, jaw movements), which are equally important as auditory cues for the understanding of speech [18]; among other, this might negatively impact the communication of individuals with hearing loss or communication disorders [7].

Face masks can be categorized into three main types: respirators, medical masks, and cloth masks (Nguyen et al., 2021). Various mask types offer different levels of protection and they affect speech signal to a different extent depending on their fabric, shape, and fit. Corey et al. [10] investigated how different types of face masks attenuate speech signal. Specifically, they tested surgical masks, N95 and KN95 respirators, six cloth masks made out of different fabrics, two cloth masks with transparent windows, and a plastic shield. They found that frequencies of more than 1 kHz were attenuated through all mask types. With respect to the effect of mask type on speech signal, surgical masks offered the best acoustic performance followed by loosely woven cotton masks. Masks with transparent windows and plastic shields offered worse acoustic performance than medical and cloth masks. In general, there was a variation in the quality of speech signal among cloth masks made of different materials. Maryn et al. [29] assessed how the sound properties of speech and noise are affected by three different types of respiratory protective masks (disposable surgical mask, FFP2 mask, and transparent plastic mask). They investigated sound properties that corresponded to voice production, voice quality, articulation, and resonance. The largest acoustic
demonstrated that there were no significant differences for vowel compared measures of voice markers in 16 adults who produced for females wearing surgical masks. Finally, Nguyen et al. (2021) found for the enhancement (CPPs) for /a/. The results indicated that there were no significant differences for the phonological inventory of the language. Georgiou[17] tested whether face masks exercised any effect on the spectral and quantitative characteristics of the Greek vowels /i e a o u/. The findings revealed that specific vowels /a/ and /i:/, instead of the full range of vowels found in the phonological inventory of the language. Georgiou[17] tested whether face-covering had a significant effect on the formants of /i e a o u/. The author concluded that there was a significant effect of facecover on the sounds' intensity as well as the centre of gravity, standard deviation, skewness, and kurtosis. This would make /f0/ and /f1/ indistinguishable, having significant implications for forensic studies. The effect of face-covering (helmet, mask, niqāb) on the acoustic properties of vowels (F1, F2, duration) was examined by Abbasi et al. [1]. The results indicated that all types of face-covering had a significant effect on the formants of /a/, while mask did not affect any speech feature of /a:/ and /i:. This study analyzed the speech properties of two Pahari central vowels only, that is /a/ and /a:/, instead of the full range of vowels found in the phonological inventory of the language. Georgiou[17] tested whether face masks exercised any effect on the speech material that was analyzed included the voiceless sibilants /s f/ and non-sibilants /l r/. The results revealed that specific mask type on the acoustic properties of sounds are more vulnerable to alteration. The effect of face-covering (helmet, mask, niqāb) on the acoustic features of sounds are more vulnerable to alteration. The effect of mask type was not considered. Joshi et al.[23] investigated the effect of different types of face masks on voice measures such as sound intensity level, fundamental frequency, jitter, shimmer and harmonics-to-noise ratio to assess how voice and verbal communication are affected. All subjects voiced a vocal sample with and without surgical masks. No significant differences were found among the masked–unmasked conditions for any of the acoustic markers, while a non-significant reduction of vocal intensity was observed in the majority of the subjects.

Few studies have studied the effect of mask-wearing on individual sounds or specific types of sounds. Fecher & Watt[13] tested how various types of face-covering impact the properties of fricatives. Ten speakers were asked to produce 18 English consonants. The speech material that was analyzed included the voiceless sibilants /s f/ and non-sibilants /l r/. The results revealed that specific masks attenuated some significant speech properties of sibilants rendering the distinction between sibilants and non-sibilants challenging. Saigusa[35] studied the effect of face-covering (motorcycle helmet, balaclava, and plastic mask) on the acoustic features of fricatives /f0 v/. The author concluded that there was a significant effect of facewear on the sounds' intensity as well as the centre of gravity, standard deviation, skewness, and kurtosis. This would make /f0/ and /f1/ indistinguishable, having significant implications for forensic studies. The effect of face-covering (helmet, mask, niqāb b) on the acoustic properties of vowels (F1, F2, duration) was examined by Abbasi et al. [1]. The results indicated that all types of face-covering had a significant effect on the formants of /a/, while mask did not affect any speech feature of /a:/ and /i:. This study analyzed the speech properties of two Pahari central vowels only, that is /a/ and /a:/, instead of the full range of vowels found in the phonological inventory of the language. Georgiou[17] tested whether face masks exercised any effect on the spectral and qualitative characteristics of the Greek vowels /i e a o u/. The findings showed a few differences between the masked and the unmasked conditions which mostly affected the F2 of some vowels. The effect of mask type was not considered. Joshi et al.[23] investigated the effect of different types of face masks on voice measures such as sound intensity level, fundamental frequency, jitter, shimmer and harmonics-to-noise ratio to assess how voice and verbal communication are affected. All subjects voiced a vocal sample with and without surgical masks. No significant differences were found among the masked–unmasked conditions for any of the acoustic markers, while a non-significant reduction of vocal intensity was observed in the majority of the subjects.

2. Methodology

2.1. Participants

Twenty speakers of Cypriot Greek (nmales = 12, nfemales = 8) participated in the study. Their age range was 20 – 55 years (Mage = 33.1, SD = 9.36), and they were born and raised in Cyprus. All of them originated from moderate-income families. The participants did not report any visual, auditory, or language deficits. Prior to their participation in the speech experiments, they were informed about the study goals and signed a consent form.

2.2. Stimuli

The stimuli consisted of the 5 Cypriot Greek vowels /i e a o u/. The vowels were embedded in monosyllabic nonsense words with the structure of /pVs/. The participants produced the words in the carrier phrase “to < keyword> íne kaló” (“the < keyword > is good”) as speaking to a friend.

2.3. Procedure

The test was completed in a sound-attenuated room and participants were tested individually. The researcher advised participants to appropriately sit in front of a PC monitor, maintaining a consistent distance from it. The carrier phrases were presented through a Microsoft PowerPoint sheet and participants were instructed to pronounce them as speaking to a friend, maintaining a normal pace. The phrases were presented in a random order in 4 repetitions with an optional two-minute break at the midpoint. In total, the participants produced 20 randomized utterances each (5 vowels × 4 repetitions). The output was recorded through Zoom H4n audio recorder at a 44.1 kHz sampling rate and saved as...
.wav files with a resolution of 24 bits. The distance between the participants’ mouth and the recorder was approximately 30 cm and it was uniform amongst speakers. The target vowels, which were part of /pV/s words, were isolated and segmented via Praat [5]. The participants had completed the test three times in three different conditions: without a mask, with a surgical mask, and with a cotton mask. The masks were completely covering the mouth and the nose of the participants and exactly fitted their face. The characteristics of the masks are shown in Table 1.

2.4. Acoustic analysis

The output was sent to Praat for speech analysis. The determination of the target vowels’ offsets and onsets have been decided on the basis of spectrograms and the waveforms. We provided measures on F1, F2, and duration. Formants comprise the acoustic resonances of the human vocal tract. F1 corresponds to vowel height, that is, how high is the tongue in the mouth (high vowels – mid vowels – low vowels), and F2 corresponds to vowel position, that is, how back is the tongue in the mouth (front vowels – central vowels – back vowels). Fig. 1 shows the spectrogram and waveform of vowel /i/.

For the generation of all tracks, the length of windows was 0.025 m.s., the pre-emphasis 50 Hz, and the spectrogram view range 5500 Hz. For the measurement of spectral features, the initial point of the vowel’s acoustic analysis was specified as the end of the noise of the first consonant /p/ and the onset of the vowel (V). The last point of the vowel’s acoustic analysis was specified as the end of the vowel (V) and the beginning of the noise of the second consonant /s/. The target vowels’ boundaries have been manually set by the researcher. Vowel formants were measured at their midpoint. Vowel duration was calculated as the distance between vowel onset and vowel offset. A Praat script was used to automatically extract the participants’ measurements (F1, F2, duration), which were ultimately saved on a Microsoft Excel file.

2.5. Statistical analysis

For our analysis, we used mixed-effects models from the lmerTest package [24] in R [33]. The dependent variables were F1, F2, and vowel duration. The productions were normalized for peak intensity in Praat. To mitigate any gender effects, formants were normalized with the NORM software package [37]. The final model included vowel (5 Greek vowels), condition (no mask/surgical mask/cotton mask), and vowel × condition as fixed slopes, while speaker was modeled as a random slope. The pairwise comparisons have been conducted through the emmeans package [25] from R. The total number of the analyzed productions was 1200 (i.e., 20 speakers × 3 conditions × 5 vowels × 4 repetitions).

3. Results

We fitted three linear mixed-effects models to investigate the effect of face mask type on F1, F2, and duration of vowels as produced by native speakers. Table 2 shows the results of the analysis. As it is illustrated in Fig. 2, vowels /i/ or /o/ seem to spectrally overlap in all conditions, while /a/ and /u/ seem to spectrally overlap in the surgical-cotton mask condition and the no mask-cotton mask condition respectively. By contrast, vowel /e/ does not occupy the same position in the vowel space in the three conditions (particularly regarding F2), vowel /a/ in the no mask condition is more peripheral compared to the other two conditions, and vowel /u/ in the surgical mask condition is more peripheral than the other two conditions.

The findings of the statistical analysis indicated that there was a significant effect (at α = 0.05) of vowel with respect to F1 of all vowels ([i]: β = −387.05, SE = 14.07, t = −27.50, p < 0.001; [e]: β = −160.00, SE = 14.07, t = −11.37, p < 0.001; [o]: β = −178.90, SE = 14.07, t = −12.71, p < 0.001; [a]: β = −332.60, SE = 14.07, t = −23.635, p < 0.001). There was also a significant effect of no mask condition ([i]: β = 33.45, SE = 14.07, t = −2.38, p < 0.05) but a non-significant effects of surgical mask condition ([i]: β = 26.80, SE = 14.07, t = 1.90, p = 0.06). In addition, it was observed an interaction of /e/ and /u/ × no mask condition ([β = −39.95, SE = 19.90, t = −2.00, p < 0.05; β = −43.85, SE = 19.90, t = −2.20, p < 0.05), and /o/ × surgical mask condition ([β = −44.10, SE = 19.90, t = −2.22, p < 0.05). To further check these effects, we employed the emmeans package from R to perform posthoc Tukey tests. The tests showed no significant differences between all corresponding vowels in the no mask-cotton mask, no mask-surgical mask, and cotton mask -surgical mask conditions. This shows that the F1 of vowels was similar in the three conditions and thus the use of mask did not alter this acoustic feature.

With respect to F2, there was a significant effect for all vowels ([i]: β = 1439.80, SE = 28.62, t = 50.30, p < 0.001; [e]: β = −521.45, SE = 18.218, t = 18.22, p < 0.001; [o]: β = −179.30, SE = 28.62, t = −6.26, p < 0.001; [a]: β = −119.45, SE = 28.62, t = −4.17, p < 0.001). Also, there was a significant effect of no mask condition ([β = 210.45, SE = 28.62, t = 7.35, p < 0.001) but a non-significant effect of surgical mask condition ([β = 38.45, SE = 28.62, t = 1.34, p = 0.18]); therefore, speech signal differed when speakers produced the vowels with a mask vs without a mask. Finally, there was a significant interaction of all vowels but /e/ × no mask condition ([i]: β = −246.25, SE = 40.48,
Table 2
Results of the linear mixed-effects models for the effects of vowel, condition and vowel × condition on F1, F2, and Duration. The table shows the estimate for the intercept and the coefficients of the model, the standard errors (SE), the degrees of freedom (df), the t-value, and the p-value (Pr>|t|). The intercept is vowel /a/ in cotton mask condition.

|   | Estimate | SE  | df   | t value | Pr>|t| |   |
|---|----------|-----|------|---------|-------|---|
| F1 | (Intercept) | 674.95 | 10.60 | 238.40 | 63.696 | <2e-16 *** |
|   | e        | −160.00 | 14.07 | 266.00 | −11.370 | <2e-16 *** |
|   | i        | −387.05 | 14.07 | 266.00 | −27.504 | <2e-16 *** |
|   | o        | −178.90 | 14.07 | 266.00 | −12.713 | <2e-16 *** |
|   | u        | −322.60 | 14.07 | 266.00 | −23.635 | <2e-16 *** |
|   | no mask  | 33.45   | 14.07 | 266.00 | 2.377  | 0.0182 *  |
|   | surgical | 26.80   | 14.07 | 266.00 | 1.904  | 0.0579  |
|   | e:no mask| −39.95  | 19.90 | 266.00 | −2.007 | 0.0457  |
|   | i:no mask| −24.65  | 19.90 | 266.00 | −1.239 | 0.2166  |
|   | o:no mask| −36.60  | 19.90 | 266.00 | −1.839 | 0.0670  |
|   | u:no mask| −43.85  | 19.90 | 266.00 | −2.203 | 0.0284  |
|   | e:surgical| −38.70  | 19.90 | 266.00 | −1.945 | 0.0529  |
|   | i:surgical| −23.45  | 19.90 | 266.00 | −1.178 | 0.2397  |
|   | o:surgical| −44.10  | 19.90 | 266.00 | −2.216 | 0.0275  |
|   | u:surgical| −9.25   | 19.90 | 266.00 | −0.465 | 0.6425  |
| F2 | (Intercept) | 1001.35 | 20.24 | 285.00 | 49.476 | <2e-16 *** |
|   | e        | 521.45  | 28.62 | 285.00 | 18.218 | <2e-16 *** |
|   | i        | 1439.80 | 28.62 | 285.00 | 50.303 | <2e-16 *** |
|   | o        | −179.30 | 28.62 | 285.00 | −6.264 | 1.38e-09 *** |
|   | u        | −119.45 | 28.62 | 285.00 | −4.173 | 3.96e-05 *** |
|   | no mask  | 210.45  | 28.62 | 285.00 | 7.353  | 2.07e-12 *** |
|   | surgical | 38.45   | 28.62 | 285.00 | 1.343  | 0.18023 |
|   | e:no mask| 42.60   | 40.48 | 285.00 | 1.052  | 0.29350 |
|   | i:no mask| −246.25 | 40.48 | 285.00 | −6.084 | 3.78e-09 *** |
|   | o:no mask| −197.60 | 40.48 | 285.00 | −4.882 | 1.75e-06 *** |
|   | u:no mask| −232.35 | 40.48 | 285.00 | −5.740 | 2.42e-08 *** |
|   | e:surgical| 389.60  | 40.48 | 285.00 | 9.625  | <2e-16 *** |
|   | i:surgical| 59.30   | 40.48 | 285.00 | 1.465  | 0.14403 |
|   | o:surgical| 4.85    | 40.48 | 285.00 | 0.120  | 0.90471 |
|   | u:surgical| 107.45  | 40.48 | 285.00 | 2.555  | 0.00939 |
| Duration | (Intercept) | 190.950 | 4.323 | 284.548 | 44.171 | <2e-16 *** |
|   | e        | −15.600 | 6.081 | 266.000 | −2.565 | 0.0109 |
|   | i        | 42.080  | 6.081 | 266.000 | −7.063 | 1.42e-11 *** |
|   | o        | −0.750  | 6.081 | 266.000 | 0.378  | 0.70619 |
|   | u        | −39.550 | 6.081 | 266.000 | −6.504 | 3.86e-10 *** |
|   | no mask  | 6.500   | 6.081 | 266.000 | 0.107  | 0.9150  |
|   | surgical | 15.100  | 6.081 | 266.000 | 2.483  | 0.0136  |
|   | e:no mask| −6.950  | 8.600 | 266.000 | −0.808 | 0.4197  |
|   | i:no mask| −0.900  | 8.600 | 266.000 | 0.134  | 0.8919  |
|   | o:no mask| −12.800 | 8.600 | 266.000 | −1.488 | 0.1378  |
|   | u:no mask| −10.900 | 8.600 | 266.000 | −1.267 | 0.2061  |
|   | e:surgical| 4.400   | 8.600 | 266.000 | 0.512  | 0.6093  |
|   | i:surgical| −13.250 | 8.600 | 266.000 | −1.541 | 0.1246  |
|   | o:surgical| −10.150 | 8.600 | 266.000 | −1.230 | 0.2390  |
|   | u:surgical| −6.700  | 8.600 | 266.000 | −0.779 | 0.4365  |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

F2 of each corresponding vowel differs in all conditions, we conducted pairwise comparisons with the Tukey postdoc test using the `emmeans` package from R. It was found that the F2 of vowel /a/ differed significantly in the cotton – surgical masks (β = −145.90, SE = 28.6, t = −5.10, p = 0.0001, d = −1.61) and the no mask – surgical mask conditions (β = −167.80, SE = 28.6, t = −5.86, p < 0.001, d = −1.21). Also, the F2 of vowel /a/ differed significantly in the no mask – surgical mask (β = 172.00, SE = 28.6, t = 6.01, p < 0.001, d = 1.90) and the no mask – cotton mask conditions (β = −210.45, SE = 28.6, t = −7.35, p < 0.001, d = −2.33). Finally, there was a significant difference for the F2 of /i/ in the no mask – surgical mask (β = −175.00, SE = 28.6, t = −6.11, p < 0.001, d = −1.93), no mask – cotton mask (β = −253.05, SE = 28.6, t = −8.84, p < 0.001, d = −2.80), and cotton – surgical mask conditions (β = 428.05, SE = 28.6, t = −14.96, p < 0.001, d = −4.73). So, only F2 was affected by mask wearing. Specifically, the F2 of two vowels was altered when speakers produced them with a cotton mask and the F2 of three vowels was altered when speakers produced them with a surgical mask.

The duration of vowels was also examined in the three conditions. Fig. 3 illustrates the duration of each vowel as produced by...
the native speakers in all conditions. Some initial differences can be spotted. First, most of the vowels produced without a mask had shorter durations than those produced in the masked conditions. The duration of the vowels produced with a surgical mask was longer than those produced with a cotton mask.

The analysis demonstrated that there was a significant effect of vowels on duration for /i/ (\(b = -42.950, SE = 6.081, t = -7.06, p < 0.001\)), /e/ (\(b = -15.600, SE = 6.081, t = -2.57, p < 0.05\)), and /u/ (\(b = -39.550, SE = 6.081, t = -6.50, p < 0.001\)). Also, there was a significant effect of surgical mask on duration (\(b = 15.100, SE = 6.081, t = 2.48, p < 0.05\)). The pairwise comparisons with the Tukey post-hoc test revealed that there were not any significant differences between corresponding vowels in the three conditions apart from the duration of /e/ in the no mask – surgical mask conditions (\(b = -25.80, SE = 6.08, t = -4.24, p = 0.028, d = -1.34\)). So, duration was not importantly altered in the masked vs non-masked conditions.

To examine whether vowels can be distinguished using face-masks, we compared the vowels’ acoustic features (F1, F2, and Duration) produced with surgical and cotton masks with those produced without a mask. Vowel signal alteration occurs when at least one of each vowel’s acoustic features in the two masked conditions significantly differs from the corresponding feature in the no mask condition. As seen in Table 3, F2 and duration of vowel /e/ in surgical mask and F2 of the same vowel in cotton mask differ from those in the no mask condition. F2 of vowel /a/ also differs between surgical/cotton and no mask conditions, while F2 of vowel /u/ in surgical mask differs from that in the no mask condition. Thus, the signal of vowels /e a u/ is altered when produced with a surgical mask, and the signal of vowels /e a/ is altered when produced with a cotton mask (see Table 3).

Table 3

| Vowel | Surgical mask F1 | F2 | Duration |
|-------|------------------|----|----------|
| /i/   | 0.20             | *  | *        |
| /e/   | *                | *  | *        |
| /a/   | *                | *  | *        |
| /o/   |                  |    |          |
| /u/   |                  |    |          |

* = significant differences from the no mask condition

4. Discussion

The present study examined the speech properties of vowels as produced by their native speakers with and without face masks. Specifically, 20 Cypriot Greek speakers were asked to produce the 5 Cypriot Greek vowels in carrier phrases with cotton and surgical masks and without a face mask. Their productions were analyzed through speech analysis software. F1, F2, and duration of the vowels in the three conditions were extracted automatically from a Praat script.

The findings of this study demonstrated that face masks modify to some extent speech signal since some speech properties of vowels produced with face masks differed significantly from the properties of vowels produced without a face mask. This corroborates earlier findings that observed signal modification in speech produced through masks (e.g., [2,13,22,35]). Fecher [12] argued that there are two main factors that determine the alteration of speech signal produced with face mask-wearing. First, two important phonetic theories, the source-filter theory of speech production developed by Fant [11] and the quantal theory of speech [36] support that even a minor modification of articulatory gestures can result in the production of sounds that are acoustically different from those that are typically produced. Speakers may try to reposition their articulators when speaking through face masks in order to produce more clear speech. This may have a significant effect on the speech properties of the produced sounds. Second, depending on the material of the face mask, an amount of energy may be absorbed by the mask which usually completely fits the face. Finally, when air particles hit the material of the mask, additional turbulences may be produced, affecting in that way the quality of speech signal. Thus, indeed face masks alter speech signal at least to some extent and what it really comes out is not what was actually produced.

The findings portrayed that not every vowel under investigation was affected in the same way by the use of face masks. Similarly, Cohn et al. [9] pointed out that face masks impact intelligibility to a larger degree depending on the type of articulation. For instance, face masks can decrease the aspiration of voiceless plosives or hinder the full lip projection of labials. In our study, vowels /i o/ remained unaffected with the use of any type of mask, whereas the properties of vowels /e a u/ were altered in the masked conditions. An outstanding finding that emerged from the present study is that F1 and duration (except in the case of /e/ in the surgical mask) aspects of vowels did not differ between the masked and no masked conditions, while F2 differed as some vowels were more or less peripheral in the masked conditions. In other words, face masks mostly impact the front/back movement of the tongue rather than its low/high position in the mouth. This agrees to some extent with the findings of Joshi et al. [23] who observed a transmission loss for F2 but not for F1 (and F0) with the use of face masks. So, it is suggested that speech signal alteration might be sound-specific and some acoustic properties might be more sensitive to changes compared to other when face masks are used.

One of the main aims of this study was to investigate the effect of mask type on vowel properties. For this purpose, two face mask types were used: a surgical and a cotton mask. Both mask types affected the same vowels (except /u/) and the same properties of vowels. For example, vowel /a/ did not differ in the surgical – cotton mask conditions but differed in the surgical – no mask and the...
cotton – no mask conditions. However, this was not the case for vowel /e/ which differed both in the surgical – no mask and the cotton – no mask conditions and the surgical – cotton mask conditions. This practically means that although speech signal is altered through masks affecting the same vowels, the speech signal of some vowels may be affected in a different way. This is interpreted as a result of the material of the mask which acts as a filter of the sound properties. The literature suggests that surgical masks alter the quality of speech signal to a lesser extent than other types of masks (e.g., [10,29,38]), and that cotton jersey (generic) masks offer a good acoustic performance that is comparable to that of surgical masks [10]. The results of this study are partially consistent with previous findings as both the surgical and the cotton masks affected only one acoustic feature of vowels and for some vowels only, yielding that other important acoustic features such as F1 and duration remain intact. Although masks affected only one vowel property, we believe that this can create challenges in pronunciation teaching where listeners have to receive explicit qualitative speech input.

The results can inform second/foreign language instructors who teach pronunciation during the pandemic. It was found that the acoustic properties of certain vowels were affected by face mask-wearing. In particular, vowels /e a/ in the surgical mask and vowels /e a/ in the cotton mask were produced with different articulatory properties compared to the no mask condition; therefore, the speech signal of these vowels was altered. So, considering that Arabic learners of Greek often cannot accurately perceive and produce the Greek vowel /e/ due to the effect of their first language (see [16,19,20]), instructors should be careful when they teach this vowel as face mask-wearing alters its speech properties. They should seek alternative ways to teach sounds. For instance, the use of masks may be avoided when much distance is kept from the students, or computer-based speech software can be used for the teaching of sounds. The use of lapel microphones is another good practice as most masks have little effect on them [10]. The results can also offer important implications for health care practitioners. For example, the teaching of native sounds to individuals with language disorders can be challenging for speech therapists. They should consider that speech signal cannot be accurate with the use of masks and thus they should find the best possible way to teach problematic speech sounds.

5. Conclusions

Effects of face mask-wearing on the acoustic properties of speech sounds were revealed. One limitation of this study is that gender was not considered in the acoustic analysis. In addition, considering that some vowels were affected in a different way in the three conditions, it remains to more precisely investigate why some vowels are affected more by face mask-wearing in comparison to other vowels. Future studies should also take into consideration the effect of face mask-wearing on other acoustic properties of vowels such as F3 and fundamental frequency.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to thank all individuals for their participation in the experiments. This paper has been supported by the RUDN University Strategic Academic Leadership Program.

References

[1] Abbasi KK, Khan AQ, Abbasis BA. Effects of forensically relevant face coverings on acoustic properties of Pahari central vowels. Int Rev Soc Sci 2021;9(1):231–25.
[2] Atcherson SR, Finley ET, McDowell BR, Watson C. More speech degradations and considerations in the search for transparent face coverings during the COVID-19 pandemic. Audiol Today 2020.
[3] Atcherson SR, Mender Veit W, Moore WJ, Patro C, Lee S, Pousson M, et al. The effect of conventional and transparent surgical masks on speech understanding in individuals with and without hearing loss. J Am Acad Audiol 2017;28(01):58–67.
[4] Bent T, Bradlow AR, Smith B. Segmental errors in different word positions and their effects on intelligibility of non-native speech: All’s well that begins well. In: Bohn O-S, Murray M, editors. Language Experience in Second Language Speech Learning: In honor of James Emil Flege. Amsterdam/Philadelphia: John Benjamins; 2007. p. 207–29.
[5] Boersma P, Weenink D. Praat: doing phonetics by computer [Computer program]. (2021). Retrieved from http://www.fon.hum.uva.nl/praat/.
[6] Bond ZS, Moore TJ, Cable G. Acoustic–phonetic characteristics of speech produced in noise and while wearing an oxygen mask. J Acoust Soc Am 1989;85(2):907–12.
[7] Chodosh J, Weinstein BE, Blustein J. Face masks can be devastating for people with hearing loss. BMJ Open 2020;7(10):e026683.
[8] Coclite D, Napoletano A, Gianola S, Del Monaco A, D’Angelo D, Fauzi A, et al. Face mask use in the community for reducing the spread of COVID-19: a systematic review. Front Med 2020;7.
[9] Cohn M, Picha A, Zellou G. Intelligibility of face-masked speech depends on speaking style: comparing casual, clear, and emotional speech. Cognition 2021;210:104570.
[10] Corey RM, Jones U, Singer AC. Acoustic effects of medical, cloth, and transparent face masks on speech signals. J Acoust Soc Am 2020;148(4):2371–5.
[11] Fant G. Acoustic theory of speech production. The Hague: Mouton Publishers; 1960.
[12] Fecher N. Effects of forensically-relevant facial concealment on acoustic and perceptual properties of consonants Doctoral dissertation. University of York; 2014.
[13] Fecher N, Watt D. Speaking under cover: the effect of face-concealing garments on spectral properties of fricatives. ICPhS 2011:663–6.
[14] Fiorella ML, Cavalaro G, Di Nicola V, Quaranta N. Voice differences when wearing and not wearing a surgical mask. J Voice 2021.
[15] Georgiou GP, Kiliani A. The use of aspirated consonants during speech may increase the transmission of COVID-19. Med Hypotheses 2020;144:109937.
[16] Georgiou GP. Discrimination of L2 Greek vowel contrasts: evidence from learners with Arabic L1 background. Speech Commun 2018;102:68–77.
[17] Georgiou GP. The effect of face mask on the acoustic properties of vowels. In: Brunn SD, Gilbreath D, editors. COVID-19 and an Emerging World of Ad Hoc Geographies. Springer; 2022. in press.
[18] Georgiou GP. Speech perception in visually impaired individuals might be diminished as a consequence of monomodal cue acquisition. Med Hypotheses 2020;143:110088.
[19] Georgiou GP. Effects of phonetic training on the discrimination of second language sounds by learners with naturalistic access to the second language. J Psycholinguist Res 2021;50(3):707–21.
[20] Georgiou GP. ThemiustochioC Vowel learning in diglossic settings: evidence from Arabic-Greek learners. Int J Bilingualism 2021;25(1):135–50.
[21] Georgiou GP, Georgiou GP, Georgiou GP, Kiliani A. How the language we speak determines the transmission of COVID-19. Irish J Med Sci (1971-) 2021;190(4):1433–48.
[22] Golden A, Weinstein B, Shimam N. How do medical masks degrade speech perception. Hearing Rev 2020;27(5):8–9.
[23] Joshi A, Procter T, Kulesz PA. COVID-19: acoustic measures of voice in individuals wearing different face masks. J Voice 2021, in press.
[24] Kuznetsova A, Brockhoff PB, Christensen RHB, Jensen DS. Tests in Linear Mixed Effects Models. R package version 3.1–2. (2020).
[25] Lenth R, Singmann H, Jonathon Love J, Buerkner P, Herve M. Marginal Means, aka Least-Squares Means. R package version 2020;3:2.
[26] Leung NHL, Chu DKW, Shiu EYC, Chan K-H, McDevitt JJ, Hau BJP, et al. Respiratory virus shedding in exhaled breath and efficacy of face masks. Nat Med 2020;26(5):676–80.
[27] Magee M, Lewis C, Noffs G, Reece H, Chang JCS, Zaga C, et al. Effects of face masks on acoustic analysis and speech perception: implications for peri-pandemic protocols. J Acoust Soc Am 2020;148(6):3562–8.
[28] Maryn Y. Initial Exploration at Home: Acoustic Voice Markers With and Without Disposable Surgical Face Mask. Lokeren, Belgium. (2020). https://www.phonanium.com/videos-and-tutorials/#toggle-id-14.
[29] Maryn Y, Wuyts FL, Zarowski A. Are acoustic markers of voice and speech signals affected by nose-and-mouth-covering respiratory protective masks? J Voice 2021.
[30] Matusiak Ł, Szepietowska M, Krajewski PK, Białynicki-Birula R, Szepietowski K. The use of masks and thus they should find the best possible way to teach problematic speech sounds. PLOS ONE 2021;16(1):e0245688.
Palmiero AJ, Symons D, Morgan JW, Shaffer RE. Speech intelligibility assessment of protective facemasks and air-purifying respirators. J Occup Environ Hygiene 2016;13(12):960–8.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2021). https://www.R-project.org/.

Saeidi R, Huhtakallio I, Alku P. Analysis of face mask effect on speaker recognition. In: Interspeech. p. 1800–4.

Stevens KN, Keyser SJ. Quantal theory, enhancement and overlap. J Phonetics 2010;38(1):10–9.

Thomas ER, Kendall T. NORM: The Vowel Normalization and Plotting Suite. Online. Resource. (2007). URL: http://lingtools.uoregon.edu/norm/norm1.php.

Toscano JC, Toscano CM, Fu Q-J. Effects of face masks on speech recognition in multi-talker babble noise. PLoS ONE 2021;16(2):e0246842.