Illusionary Effect of Face Mask on Audio-Visual Speech Intelligibility: A Pilot Study with L2 Arabic Learners

Yahya Aldholmi[1] and Hanadi Almuhawwis[1]

[1]Department of Linguistics & Translation Studies, King Saud University

yaldholmi@ksu.edu.sa, hanadisumu@gmail.com

Abstract

In this study, we recruited a native Arabic speaker to audio-visually record 20 Modern Standard Arabic (MSA) sentences in a clear style at a normal speech rate, once while wearing a face mask and once without. Then, we presented 10 L2 Arabic learners with the nonmasked *auditory* stimuli and asked them to rate, on a 1-7 scale, the intelligibility of each sentence. The stimuli were presented in two conditions: *match*, where 10 were presented with the original *nonmasked visual stimuli*, and *mismatch*, where the other 10 were presented with the *masked visual stimuli*. Listeners tended to rate the sentences in the *mismatch* condition as *less* intelligible, demonstrating that speech intelligibility is sensitive not only to the acoustic-phonetic changes caused by masks, as previous studies suggest, but also to the visibility of face masks.

Keywords

Face masks, speech intelligibility, Arabic, L2

Introduction

The impact of face masks on speech production and perception in a variety of contexts and environments has long been explored in the literature (e.g., Bond et al., 1989; Fecher and Watt, 2011). Bond et al. (1989) found changes in vowel space articulation, F1 and F2 in particular, for speakers wearing oxygen masks, and recommended that automatic speech recognition (ASR) should consider operational environments, such as in aircrafts. Fecher and Watt (2011) also reported effects of different face-concealing garments on spectral properties, which have important implications for speech production and perception in a forensic context. Recently, due to the heavy worldwide use of masks necessitated by COVID-19, studies have revisited the topic and investigated the impact of different face mask types on speech properties (e.g., Chmelík et al., 2021; Georgiou, 2022) and intelligibility in different styles and environments (e.g., Knowles & Badh, 2022). For example, Cohn et al. (2021) showed that face-masked *clear* speech presented auditorily to listeners is more intelligible than nonmasked, perhaps because talkers consciously produced speech with utmost clarity. It is unclear, however, whether face masks that are visible to listeners, especially L2 learners, can have an impact on speech intelligibility. Specifically, L2 listeners may develop an assumption that they will face difficulty understanding a face-masked speaker, regardless of actual clarity. The current study examines the potential effect of the visibility of face masks on audio-visual speech intelligibility by L2 Arabic learners.

1 Methods

1.1 Stimuli

Twenty MSA sentences adopted from Set 2 of the stimuli used in Aldholmi (2020) were audio-visually recorded by a male talker twice, once while wearing a surgical face mask and once without. Twenty *auditory stimuli* (extracted from the nonmasked items) and *visual stimuli* (masked and nonmasked) were used for the experiment.

1.2 Talkers and listeners

The talker is a native Arabic speaker (*age* = 23) pursuing an undergraduate degree in the Arabic language. He was instructed to produce the sentences in a clear style at a normal speech rate. Every time he recorded a new sentence, he listened to the first sentence and then attempted to emulate the levels of loudness, speech rate, and clarity. The listeners were L2 Arabic intermediate learners (age, M = 27.8, SD = 4.1) with varying linguistic backgrounds (e.g., Tagalog, Yoruba, Indonesian, and English). None of them reported a history of impaired speech, hearing, or vision.

1.3 Procedure

The participants were presented with the 20 items in random order and asked to judge the intelligibility of each item on a 1-7 scale (1 = *totally unintelligible*, 7 = *totally intelligible*). All items were presented audio-visually (Figure 1), but half were accompanied by the masked visual stimuli (*mismatch* condition), while the second half were associated with the nonmasked stimuli (*match* condition). A 100-ms lag was inserted at the beginning of each item in the *match* condition to inhibit the benefits of visual cues.

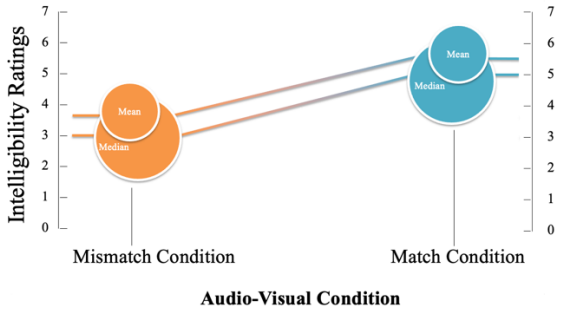
Figure 1
Illustration of audio-visual conditions

Mismatch Condition Match Condition

2 Results

As shown in Figure 2, although all 20 items were produced without a face mask, the auditory stimuli accompanied by the masked visual stimuli received lower intelligibility ratings (M= 3.8, Mdn = 3) than the ones accompanied by the nonmasked visual stimuli (M= 5.5, Mdn= 5). Over 70% of the mismatch-condition stimuli were judged as being (slightly) unintelligible, and approximately 60% of the match-condition stimuli were judged as being (slightly) intelligible. A Kruskal-Wallis test shows a statistically significant difference between the two conditions, $H(1) = 82$, $p = 0.02$.

Figure 2
Intelligibility ratings by L2 Arabic learners



3 Discussion and conclusion

The results from the current study show that wearing a face mask, regardless of speech clarity, affects L2 learners' judgements of speech intelligibility. Our results do not conflict with those from Cohn et al. (2021), who found that nonmasked *auditory* speech is less intelligible than masked, because their study examined the acoustic/auditory but not the visual domain. One may infer that, as some studies have indicated (e.g., Thibodeau et al., 2021), the absence of visual cues yielded lower ratings of intelligibility in the mismatch condition. However, we interpret our results as an illusionary effect of masks on L2 learners' impression of speech intelligibility, because the impact of the visual cues was deliberately inhibited in the match condition. Future research may reduce the intelligibility of nonmasked speech or/and recruit advanced L2 learners to examine whether the illusionary effect will persist in the masked condition.

4 References

Aldholmi, Y. (2020, October). Syllable rate vs. segment rate in perceived speech rate. In *Proc. 11th ExLing* (pp. 21-24).

Bond, Z. S., Moore, T. J., & Gable, B. (1989). Acoustic–phonetic characteristics of speech produced in noise and while wearing an oxygen mask. *The Journal of the Acoustical Society of America*, *85*(2), 907–912.

Chmelík, V., Urbán, D., Zelem, L., & Rychtáriková, M. (2021). Effect of mouth mask and face shield on speech spectrum in Slovak Language. *Applied Sciences*, *11*(11), 4829.

Cohn, M., Pycha, A., & Zellou, G. (2021). Intelligibility of face-masked speech depends on speaking style: Comparing casual, clear, and emotional speech. *Cognition*, *210*, 104570.

Fecher, N., & Watt, D. (2011, August). Speaking under Cover: The Effect of Face-concealing Garments on Spectral Properties of Fricatives. In *ICPhs* (pp. 663-666).

Georgiou, G. P. (2022). Acoustic markers of vowels produced with different types of face masks. *Applied Acoustics*, *191*, 108691.

Knowles, T., & Badh, G. (2022). The impact of face masks on spectral acoustics of speech: Effect of clear and loud speech styles. *The Journal of the Acoustical Society of America*, *151*(5), 3359-3368.

Thibodeau, L. M., Thibodeau-Nielsen, R. B., Tran, C. M. Q., & de Souza Jacob, R. T. (2021). Communicating during COVID-19: the effect of transparent masks for speech recognition in noise. *Ear and Hearing*, *42*(4), 772-781.