Received: 9 May 2023 | Revised: 12 January 2024 | Accepted: 29 February 2024

DOI: 10.1111/ijal.12553

ORIGINAL ARTICLE

WILEY

Wearing face masks in different speech styles during the COVID-19 pandemic: A study of Thai L2 English learners

Patchanok Kitikanan[1] | **Alex Ho-Cheong Leung**[2]

[1]English Department, Faculty of Humanities, Naresuan University, Tha Pho, Thailand

[2]Department of Humanities, Faculty of Arts, Design and Social Sciences, Northumbria University, Newcastle upon Tyne, UK

Correspondence
Alex Ho-Cheong Leung, Department of Humanities, Faculty of Arts, Design and Social Sciences, Northumbria University, Newcastle upon Tyne, UK.
Email:
alex.ho-cheong.leung@northumbria.ac.uk

การสวมใส่หน้ากากอนามัยในบริบทหลากหลายในยุคโรคระบาดโควิด-19: กรณีศึกษาของกลุ่มผู้เรียนภาษาอังกฤษชาวไทย

Abstract

During the COVID-19 epidemic, the effect of mask wearing on communication has been questioned and explored. However, the study on the impact of face mask wearing on L2 speech is still limited. The main goal of the present study was to explore the extent to which face masks affect interlocutors' L2 speech perception. Factors that varied were face mask (no mask, transparent mask, and disposable face mask), presentation mode (audio only vs. audiovisual), and speaking style (conversational vs. clear). The relationship between these three variables on the L2 processing was also investigated. Fifty-three Thai undergraduates who were L2 learners of English participated in an Internet-based perceptual task. They listened to 60 English sentences and typed the sentences they heard over an online form. The results showed that the participants did well when the L2 communication was in clear speech regardless of type of face mask and presentation mode. The improvement of the L2 perception of clear speech occurred even when the speaker produced sentences with disposable face mask as opposed to conversational speech suggesting that clear speech could enhance intelligibility in communication. As expected, the perceptual score was the lowest when the speech was in audio-only mode with conversational style and disposable

face mask. These results suggest that in L2 classroom contexts, speaking clearly could lessen the negative effects of the face mask and unavailability of visual information of the speaker.

KEYWORDS
COVID-19, face mask, presentation mode, speaking style, speech intelligibility

บทคัดย่อ
งานวิจัยจำนวนมากได้ศึกษาถึงผลกระทบของการสวมหน้ากากอนามัยที่มีต่อการสื่อสารในยุคโรคระบาดโควิด-19
อย่างไรก็ตาม
งานวิจัยที่ศึกษาผลกระทบของการสวมหน้ากากอนามัยที่มีต่อการสื่อสารภาษาที่สองยังคงมีจำนวนไม่มาก
วัตถุประสงค์ของงานวิจัยนี้ เพื่อสำรวจผลของการสวมหน้ากากอนามัยที่มีต่อการรับรู้ภาษาที่สอง
โดยปัจจัยที่ศึกษามีจำนวน 3 ปัจจัย ได้แก่ ชนิดของหน้ากากอนามัย (ไม่มีหน้ากาก
หน้ากากแบบโปร่งใสและหน้ากากแบบทึบ) บริบทการนำเสนอ
(ได้ยินเสียงอย่างเดียวและได้ยินเสียงพร้อมเห็นภาพ) และวิธีการพูด (พูดแบบสนทนาและพูดให้ชัดเจน)
ซึ่งงานวิจัยนี้ได้ศึกษาความสัมพันธ์ของปัจจัยทั้งสามนี้ด้วย
กลุ่มอาสาสมัครประกอบด้วยกลุ่มผู้เรียนชาวไทยระดับปริญญาตรีจำนวน 53 คน
ซึ่งเป็นคนที่เรียนภาษาอังกฤษเป็นภาษาที่สอง
โดยกลุ่มอาสาสมัครได้เข้าร่วมงานวิจัยที่ให้ฟังประโยคภาษาอังกฤษจำนวน 60 ประโยค
ผ่านช่องทางออนไลน์และอาสาสมัครพิมพ์ประโยคใส่แบบฟอร์มออนไลน์ ผลการวิจัยพบว่า
กลุ่มอาสาสมัครมีคะแนนการรับรู้ภาษาอังกฤษดีเมื่อผู้พูดพูดให้ชัดเจน
ในทุกชนิดของหน้ากากอนามัยและทุกบริบทการนำเสนอ
กลุ่มอาสาสมัครมีคะแนนดีกว่าเมื่อฟังประโยคภาษาอังกฤษที่ผู้พูดสื่อสารชัดเจน แม้ว่าจะใส่หน้ากากแบบทึบ
เทียบกับเมื่อผู้พูดสื่อสารแบบสนทนา ซึ่งแสดงให้เห็นว่า
การพูดให้ชัดเจนช่วยเพิ่มระดับความเข้าใจภาษาของผู้ฟังได้ นอกจากนี้ งานวิจัยยังพบว่า
คะแนนการรับรู้ประโยคภาษาอังกฤษต่ำที่สุดเมื่ออาสาสมัครฟัง-
การพูดที่ไม่เห็นหน้าผู้พูดและผู้พูดสื่อสารด้วยการพูดแบบสนทนาและใส่หน้ากากแบบทึบ
ซึ่งเป็นไปตามการคาดการณ์ของผู้วิจัย ผลการวิจัยเหล่านี้แสดงให้เห็นว่า ในบริบทการสื่อสารด้วยภาษาที่สองนั้น
ผู้พูดควรพูดด้วยความชัดเจน
ซึ่งจะช่วยลดอุปสรรคการสื่อสารของหน้ากากอนามัยและการสื่อสารโดยไม่เห็นหน้าผู้พูดได้มากขึ้น

คำสำคัญ:
หน้ากากอนามัย, บริบทการนำเสนอ, วิธีการพูด, ความเข้าใจภาษา, โควิด-19

1 | INTRODUCTION

One million people have died worldwide as a result of the COVID-19 epidemic to date (Wood et al., 2020). As this pandemic seems to have no end, healthcare authorities and governments suggested prophylactic practices that can decrease the chance of virus transmission. Besides hand washing and social distancing, the other effective practices that could reduce transmission of the viruses is wearing face masks. Although wearing face masks is not common in Western populations, such as Canada, France, Germany, and the United Kingdom, wearing face masks is still popular in many Asian countries, such as China, Japan, and India (Bricker, 2020). However, masks decrease the quality of speech signal by altering the direction of sound radiation and by attenuating frequencies above 1 kHz, which carry important information for perception although variation exists depending on mask types and the microphone placement (Corey

et al., 2020). Masks also cover up visual cues that improve speech recognition by adding details about speech sounds that the auditory signal is missing (Hardison, 2021; Hardison & Pennington, 2021).

Several studies showed that different types of face masks affect speech intelligibility to a different degree (e.g., Brown et al., 2021; Smiljanic et al., 2021; Yi et al., 2021). For example, in the study of Yi et al. (2021), American English speakers were recruited to listen to English speech varying in four factors: type of mask, presentation mode, speaking style, and type of background noise. The results showed that with background noise, subjects scored less when the speaker wore a disposable face mask or a transparent mask compared to wearing no mask. It is also found that transparent masks increased participants' ability to listen to target sentences as visual information is available in such condition. They also supported that clear speech, which is a communication that is adjusted to enhance speech intelligibility (see below), facilitated participants' ability to understand the sentences. As the participants in that study were American English speakers whose mother tongue is identical to the target language (i.e., English), clear speech allows them to get access to the message with ease. Another example is the study of Brown et al. (2021) investigating the effects of face mask types and noise level on speech intelligibility in young and older adults. The findings showed that in quiet condition, mask type had a minimal effect on speech intelligibility in both groups. But in context with background noise, all types of face masks negatively affect the intelligibility in both groups. The surgical face masks had the smallest impact on performance whereas transparent and cloth face masks had great impact on the intelligibility. The results of these studies support Chandrasekaran et al. (2009) that visual speech information such as mouth, lip, and tongue movements could provide important (temporal) cues about speech articulation. Although face masks have been found to be a barrier for native speakers of English in terms of speech intelligibility, speaking clearly, reducing noise, and having a good signal amplification could help improve intelligibility (Smiljanic et al., 2021).

According to Guskaroska (2021), both native and non-native speakers view speech with face masks as an impediment to speech intelligibility. However, research on face mask wearing and L2 speech learning is still limited. At the time of this research, we are only aware of one such "published study"—there could be a chance that there are other studies (e.g., ones that are non-published or published in small/local outlets) that we could not find on this topic (Lee & Hart, 2022). Lee and Hart (2022) examined the effect of face masks in L2 listening comprehension. Japanese participants of L2 English were divided into three groups according to the stimulus they received: speech through a mask, speech with no mask, and audio only. Participants received audiovisual message through listening to stimuli produced with a mask and with no mask. While the audio-only group heard the stimuli of the speaker without a face mask on. The results showed that the listeners scored worse when hearing stimulus from speakers with masks than when speakers wore no mask and when they heard audio-only stimuli. This suggests that face mask wearing affects the acoustic signal negatively and it also obscures nonverbal cues such as facial expressions and lip movements. Although the intelligibility of speech among L2 learners has been explored, some questions still remain, for example, "What type of mask is the most effective to wear for the speaker to convey a message to L2 learners?," "Would it be easier for L2 learners to receive visual information from the speaker?" or "Would hearing the speech only (without visual cues) be easier for them?," "Will clear speech help L2 learners understand even when a face mask is worn?," "Is there a relationship among the type of mask worn, presentation mode and speaking style?." These questions will be addressed in this study.

The first objective of this study is to investigate the effects of wearing face masks on speech intelligibility in different speaking styles and presentation modes. There are three variables: (1) type of mask (no mask, transparent mask, and disposable face mask), (2) presentation mode (auditory only and audiovisual), and (3) speaking style (clear speech and conversational speech). We also would like to explore how these three variables interact, and if so, how. The total number of conditions is 12. We aim to identify the most effective condition for L2 learners on speech intelligibility.

2 | LITERATURE REVIEW

2.1 | Face mask and speech intelligibility

The effect of face mask wearing on speech communication has been investigated in some linguistic contexts (e.g., English and Korean), most studies found that face masks decrease participants' ability to perceive speech signals (e.g., Bottalicoet al., 2020; Brown et al., 2021; Cohn et al., 2021; Guayasamin, 2021; Moon et al., 2022; Schwarz et al., 2022). For instance, Moon et al. (2022) explored the effect of face mask on speech perception by participants with normal hearing and hearing disorder. In a listen-and-repeat task, participants watched four types of video recordings: a female speaker with and without a face mask and a male speaker with and without a face mask. The findings showed that participants had better perception when seeing speakers without a mask, suggesting that visual cues facilitated speech perception for both groups. Another study by Schwarz et al. (2022) investigated the acoustic and visual effects of face masks on speech intelligibility and processing speed with variations in semantic predictability, that is, contextual information, of children and adults. The findings showed that both groups performed poorly and had slower response when listening to face-masked speech as opposed to speech produced without a face mask. In agreement with these studies, Bottalico et al. (2020) and Brown et al. (2021) found that all mask types are positively related to the rating of listening effort. However, few studies showed that face masks affect speech perception in more nuanced ways (e.g., Keerstock et al., 2020; Magee et al., 2020; Smiljanic et al., 2021). For example, in the study of Magee et al. (2020), the effects of three face mask types (N95, surgical, and cloth) on acoustic qualities of speech and intelligibility were investigated. Although word accuracy and sentence translation were affected by all masks, no differences were found across conditions for word or sentence intelligibility measures. Smiljanic et al. (2021) which explored the effects of background noise, protective masks, and non-native accents on speech comprehension and general communication is another of such studies. They found that for both native speakers and non-native speakers, conversational discourse produced while donning a disposable face mask did not adversely influence speech understandability in quiet listening contexts. Similarly, Cohn et al. (2021) found no difference between the face-masked speech and non-face-masked speech in the casual condition among American L1 English listeners.

2.2 | Presentation mode and language proficiency

In general, speech is commonly perceived auditorily. However, it could also be perceived with articulatory movements. The visual cues of speech have been shown to be important in perception. When a listener sees other visual cues that are incongruent with the speech they are hearing, it may lead to misperception (Tiippana, 2014). This is called the McGurk effect (McGurk & MacDonald, 1976). In this effect, the visual and auditory components of two sounds are combined, creating the impression of a third sound. In other words, a person's perception of sound is altered by the visual information they received from witnessing someone talk. There could also be an intermediate outcome where the sound perceived is neither the one that corresponds to the visual nor the speech, but something in between that contains features from both. This suggests a multisensory illusion due to the conflicting auditory and visual information. The study on the effects of visual information on intelligibility has extended to people with speech disorders. The results of the impact of visual cues on speech intelligibility in individuals with speech problems showed that intelligibility of speech was higher when listeners received audiovisual cues as opposed to the auditory speech in isolation (Erber, 1975; Hubbard & Kushner, 1980; Monsen, 1983). In L2 speech learning, audiovisual modality in perceptual training was found to increase L2 perception. For example, in the study by Hazan et al. (2005), the perception of /v/-/b/-/p/ labiodental/labial contrast of Japanese learners of English was greater among those who were trained in audiovisual training relative to those trained in the auditory training. Even though the effect of presentation mode in L2 speech learning has been explored in a number of research studies (e.g., Abraham, 2007; Becker & Sturm, 2017; Hazan et al.,

2005; Jones & Plass, 2002; Li, 2006), to our knowledge, our study is the first to investigate the effect of this factor in relation to the face mask type. The visual aids provide learners with additional modes of cognitive representation of speech which in turn enhances the perception of the target speech.

2.3 | Speaking style and language proficiency

The aim of talkers in ordinary communication is to convey their ideas in a way that is understandable to listeners. Talkers will naturally and impulsively alter their speech to accommodate the listener if they become aware that the listener is having trouble understanding them owing to hearing loss, or a distinct native tongue (Smiljanić & Bradlow, 2009). They tend to adjust their talking into "clear speech." In clear speech, speakers are likely to speak more slowly, louder, and with more exaggerated sound articulation (Smiljanić & Bradlow, 2009). Additionally, variable phonetic patterns brought about by prosodic strengthening, which refers to the "temporal and/or spatial expansion of articulation due to accent and/or prosodic boundary" (Cho, 2005), and changes in tempo from fast to slow can be found in clear speech (Cole et al., 2007). For native listeners, research studies showed that clear speech improved accuracy in communication (Smiljanic et al., 2021). However, unlike native speakers whose speech difficulty often arises from losing speech signal, the communication problem of non-native speakers commonly occurs due to a lack of experience in sound structures (e.g., sound contrasts and phonotactic rules) and higher-level linguistic structures (e.g., syntax and semantics) (Smiljanić & Bradlow, 2009). A number of studies were carried out on the effect of clear speech as perceived by non-native listeners (e.g., Smiljanic & Bradlow, 2007; Smiljanić & Bradlow, 2011). One such study is by Smiljanić and Bradlow (2011) who investigated the effect of clear speech among native and high proficiency non-native listeners. The findings revealed that non-native listeners, rather than native listeners, gained a significant benefit of clear speech from native talkers. Similarly, Smiljanic and Bradlow (2007) explored the effect of English clear speech produced by Croatian talkers and perceived by Croatian listeners. They found that for non-native listeners, clear speech was significantly more intelligible than plain speech. It has been reported that clear speech enhances speech intelligibility, even when face masks were worn. For example, Yi et al. (2021) demonstrated that speaking clearly can compensate difficult communication conditions, such as making up for the lack of visual clues and diminished audio signals. Related to that, Cohn et al. (2021) found that face-masked speech was actually more understandable than speech that was not produced with a face mask which could be due to presenters particularly altering their speech clarity when a face mask was worn.

2.4 | Predictions

For the type of mask, we anticipate that the ease of listening will decline from no mask (NO), to transparent mask (TM), and to disposable face mask (DM), respectively. The no-mask speech should be easiest since audio and visual information are fully/mostly available. The transparent-mask speech should be easier than the disposable-face-mask speech as some visual cues are still available to the listeners. For the presentation mode, the participants should score lower in the audio-only mode (AO) as opposed to the audiovisual mode (AV). For the speaking style, the participants should perform better when listening to clear speech (CLEAR) than conversational speech (CON). We are also interested in exploring the interactions among these three factors (i.e., mask condition, presentation mode, and speaking style), which result in 12 conditions. However, we offer no a priori predictions for them. Table 1 shows the 12 contrasts for the three factors.

TABLE 1 The 12 conditions for the three variables.

Audio only (30)		Audiovisual (30)	
Conversational (15)	Clear (15)	Conversational (15)	Clear (15)
No mask (5)	No mask (5)	No mask (5)	No mask (5)
Transparent (5)	Transparent (5)	Transparent (5)	Transparent (5)
Disposable (5)	Disposable (5)	Disposable (5)	Disposable (5)

Note: The numbers in parentheses are the number of the sentences for each context.

3 | METHODOLOGY

3.1 | Participants

Fifty-three Thai undergraduates aged between 19 and 23 years (Mean = 20.07, SD = 0.75) took part in the study. They were English-major students in their second year of studies. Their Cambridge English Placement Test (CEPT) scores were between 10 and 47 (Mean = 28.21, SD = 6.13). The mean CEPT scores suggested that the level of English proficiency of these participants was A2 which means they were basic users of English. They had studied English as a foreign language in Thailand since kindergarten. English had been used mainly in the classroom rather than in their natural daily contexts. Pearson's correlation coefficient showed that there was no correlation between CEPT score and accuracy in the L2 perception ($p > 0.05$). The L2 perception score was the overall mean accuracy of word identification accuracy scores per participant.

3.2 | Materials

The stimuli were 240 words in 60 sentences taken from Yi et al. (2021). They were produced by a female native American English speaker. Each sentence contains four target words, which are capitalized here for clarity, such as "My DOCTOR WORKS in that BUSY HOSPITAL." and "The CHICKEN SOUP was a TASTY MEAL." Unlike the study of Yi et al. (2021), the effect of background noise is not the focus of our study—the stimuli hence contained no competing noise. The 60 sentences were divided into two main groups: 30 sentences with audio only and 30 sentences with both audio and visual cues. In each condition (i.e., audio only and audiovisual), 15 sentences were produced in clear speech whereas the other 15 sentences were in conversational speech. The clear speech sentences and conversational speech sentences were then sub-divided into three groups (five sentences for each group): no mask, transparent mask, and disposable face mask (see Table 1). The mean duration of the clear speech was 3154.37 ms whereas the mean duration of the conversational speech was 1610.59 ms (Yi et al., 2021). The same speaker read each sentence across all conditions, that is, while wearing a transparent mask, a disposable face mask, and while wearing no face mask. For the contrast in speaking patterns, that is, clear speech and conversational speech, she started out with clear speech and then switched to a conversational style. In conversational speech, the speaker was instructed to use conversational language by speaking as though she was conversing with a friend. In clear speech, the speaker was instructed to speak as if she was conversing with a person who had hearing loss and found it difficult to comprehend her. Besides these 60 sentences, there were another 12 sentences used for practice, with 6 being audio only and the other 6 being audiovisual sentences. Of those practice sentences with audio only, two were produced with a disposable face mask, two were produced with a transparent mask, and the other two were produced with no mask. The same procedure was followed for the audiovisual practice sentences. Table A in the appendix shows the list of sentences with their contexts.

3.3 | Data collection

The tested sentences and practice sentences were randomized in Microsoft Excel for presentation. Microsoft PowerPoint was used to present 72 sentences, including practice and target sentences. To ensure ecological validity and due to the restrictions imposed by the COVID pandemic, we did not request participants to take part in a lab. They listened to the stimuli over headphones via MS Teams at a convenient place to them. They typed (transcribed) their responses in an online questionnaire on Google Form. The questionnaire comprised three sections: the consent form, the practice session, and the test session. Each stimulus was presented for 40 s; each slide was projected for that duration. The listeners were instructed to report any testing-related issues, such as a delay between the accompanying visual and aural stimuli. But none of them had reported any issue. The whole experiment took approximately 1 h to complete. Participants had given written informed consent and received extra credits in a course as compensation for their time. All experimental procedures were approved by Naresuan University Institutional Review Board (COA No. 010/2023, IRB No. P2-0319/2565).

3.4 | Data analysis

The data from Google Form were transferred to Microsoft Excel. The target words of each sentence were divided into four rows, resulting in four rows per sentence. Each participant had 240 items (60 sentences × 4 items). The total number of observations was 12,720 (53 participants × 240 items). The right answer was coded as "correct" and the wrong answer was coded as "wrong." American English and British English spelling were both accepted, for example, "favorite" and "favourite" were both coded as "correct." The sentences with misspellings were counted as incorrect. Binomial generalized linear mixed models (GLMMs) were used via the *lme4* package in R statistical software (Bates et al., 2015) to analyze the data. The dependent variable was binomial—whether the answer was correct or wrong. The independent variables were *mask type*, *presentation mode*, and *speech style*. For *mask type*, three types were classified: no mask, transparent mask, and disposable face mask. For *presentation mode*, there were two modes: audio only and audiovisual. For *speech style*, there were also two styles: clear and conversational. These combinations resulted in 12 conditions of communication (2 presentation modes × 2 speech styles × 3 mask type). The listener and item were random intercepts. The model comparison was done between a one-level model, a two-level model, and a three-level model with interaction of all factors. The *anova* function was used to show the Akaike Information Criterion (AIC), according to Zheng et al. (2014) that the model with smaller AIC was a better one than the one with a greater AIC. For our study, the optimal model was the model with three-way interaction. To see all pairs of contrasts, Tukey's HSD (Honest Significant Difference) post hoc test was calculated using the *emmeans* package in R statistical software (Lenth et al., 2018).

4 | RESULTS

Of the 12 conditions, it was found that the number of the correct answers in the audio-only mode in conversation speech with disposable face mask yielded the lowest scores whereas that in the audiovisual in clear speech with disposable face mask yielded the highest scores. The number of correct answers yielded for stimuli read in clear speech was always greater than those in the conversation speech, regardless of the presentation modes (i.e., audio vs. audiovisual) and the type of face masks (i.e., disposable masks, transparent masks, and no mask). The numbers of the total correct answers in each condition, out of 1060, are shown through the descriptive results in Figure 1.

In addition, after comparing GLMMs, significant differences among the three variables were found in the three-way-interaction GLMM ($p < 0.01$) (Table 2), suggesting that all variables were responsible in explaining the perception

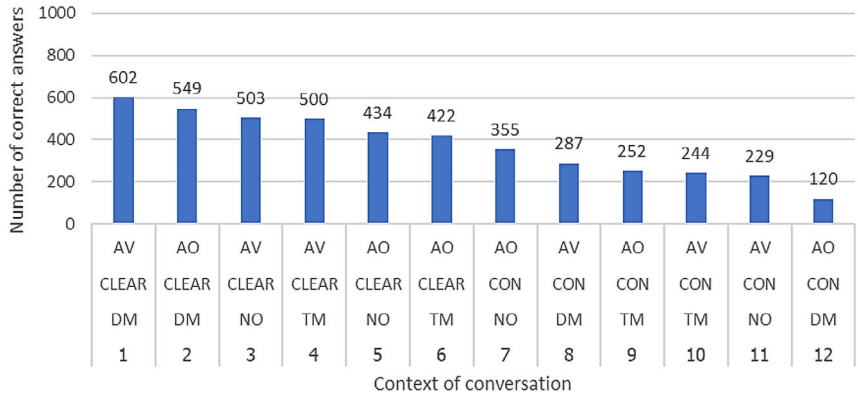

FIGURE 1 Number of correct answers in sentences for comparisons between presentation modes (audio only: AO, audiovisual: AV), speech styles (clear: CLEAR and conversation: CON), and types of face masks (no mask: NO, transparent mask: TM, and disposable face mask: DM). [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 2 Estimates for intercept and factors from generalized linear mixed model for the exploration of L2 English speech.

	Estimate	Std. error	*z*-value	Pr(> |z|)
(Intercept)	0.1162	0.6029	0.193	0.847129
Audiovisual	−0.397	0.5904	−0.672	0.501316
Conversation	4.7369	0.6038	7.845	4.33e-15
No mask	1.4202	0.5883	2.414	0.015778
Transparent mask	1.4553	0.5885	2.473	0.013407
Audiovisual: Conversation	−1.5798	0.8395	−1.882	0.059852
Audiovisual: No mask	−0.3189	0.8316	−0.383	0.701380
Audiovisual: Transparent mask	−0.1970	0.8317	−0.237	0.812781
Conversation: No mask	−4.0677	0.8404	−4.840	1.30e-06
Conversation: Transparent mask	−3.1015	0.8417	−3.685	0.000229
Audiovisual: Conversation: No mask	3.4325	1.1792	2.911	0.003603
Audiovisual: Conversation: Transparent mask	2.2832	1.1798	1.935	0.052967

of L2 English speech by Thai learners; hence, only the interactions will be reported and discussed below. Each condition is numbered as follows:

1. Audiovisual mode with clear speech and disposable face mask;
2. Audio-only mode with clear speech and disposable face mask;
3. Audiovisual mode with clear speech and no mask;
4. Audiovisual mode with clear speech and transparent mask;
5. Audio-only mode with clear speech and no mask;
6. Audio-only mode with clear speech and transparent mask;
7. Audio-only mode with conversational speech and no mask;
8. Audiovisual mode with conversational speech and disposable face mask;
9. Audio-only mode with conversational speech and transparent mask;

10. Audiovisual mode with conversational speech and transparent mask;
11. Audiovisual mode with conversational speech and no mask;
12. Audio-only mode with conversational speech and disposable face mask.

Overall, the participants scored well when the speaker produced sentences in clear speech. This is particularly true when the speaker talked in clear speech with disposable face mask as the score of correct answers in these conditions were significantly greater than the conditions when the speaker talked in conversational speech across face mask types and presentation modes.

To be specific, for the face mask effect, in both audio-only and audiovisual modes, there were no significant differences between the L2 perceptual scores among the mask conditions in clear speech ($p > 0.05$ for all contrasts). In audiovisual mode in conversational speech, no significant differences between the L2 perceptual scores among the mask conditions were found ($p > 0.05$ for all contrasts). In audio-only mode, the number of correct answers when the speaker wore no mask was significantly higher than that when the speaker wore disposable face mask ($b = -2.65$, $SE = 0.60$, $p < 0.01$). There were no significant differences between the number of correct answers when the speaker wore no face mask versus when the speaker wore transparent mask, and when the speaker wore disposable face mask versus when she wore transparent mask ($p > 0.05$ for all contrasts).

For the clear-speech effect, participants performed better when listening to clear speech as opposed to conversational speech, except in two conditions. In audio-only mode when the speaker was wearing disposable face mask, the number of correct answers in the clear speech was significantly greater than when the speaker produced sentences in conversational style ($b = 4.74$, $SE = 0.61$, $p < 0.01$). In audiovisual mode the number of correct answers in the clear speech was significantly higher than in the conversational style across face mask conditions (Transparent mask: $b = 2.34$, $SE = 0.59$, $p < 0.01$; Disposable face mask: $b = 3.16$, $SE = 0.60$, $p < 0.01$; and No mask: $b = 2.52$, $SE = 0.60$, $p < 0.01$). In audio-only mode, no significant differences were found between the number of correct answers in the clear and conversational speech when the speaker wore no face mask and when wearing transparent face mask ($p > 0.05$ for all contrasts). For audiovisual effect, no significant differences were found across all the conditions ($p > 0.05$ for all contrasts).

Apart from that, the results from GLMM with three-way interaction showed that in the audio-only mode when the speaker wore a disposable face mask and spoke with clear speech, the number of correct answers was significantly greater than when the speaker produced sentences in conversational style across presentation modes and type of face masks ($b = 4.74$, $SE = 0.61$, $p < 0.01$ for AO, CON, and DM; $b = 2.76$, $SE = 0.59$, $p < 0.01$ for AV, CON, and DM; $b = 2.09$, $SE = 0.59$, $p < 0.05$ for AO, CON, and NO; $b = 3.23$, $SE = 0.60$, $p < 0.01$ for AV, CON, and NO; $b = 3.09$, $SE = 0.60$, $p < 0.01$ for AO, CON, and TM; and $b = 3.20$, $SE = 0.59$, $p < 0.01$ for AV, CON, and TM). Similarly, in the audiovisual mode when the speaker wore a disposable face mask and spoke with clear speech, the number of correct answers was significantly greater than when the speaker produced sentences in conversational style across presentation modes and type of face masks ($b = 5.13$, $SE = 0.61$, $p < 0.01$ for AO, CON, and DM; $b = 3.16$, $SE = 0.60$, $p < 0.01$ for AV, CON, and DM; $b = 2.49$, $SE = 0.60$, $p < 0.01$ for AO, CON, and NO; $b = 3.62$, $SE = 0.60$, $p < 0.01$ for AV, CON, and NO; $b = 3.49$, $SE = 0.60$, $p < 0.01$ for AO, CON, and TM; and $b = 3.60$, $SE = 0.60$, $p < 0.01$ for AV, CON, and TM). In the audiovisual mode when the speaker wore no face mask and spoke with clear speech, the number of correct answers was significantly greater than in the audio-only mode in conversational speech with transparent mask ($b = 2.39$, $SE = 0.60$, $p < 0.01$), in the audiovisual mode in conversational speech with transparent mask ($b = 2.50$, $SE = 0.60$, $p < 0.01$), and in the audiovisual mode in conversational speech with no mask ($b = 2.52$, $SE = 0.60$, $p < 0.01$). In the audiovisual mode in clear speech with transparent mask, the number of correct answers was significantly greater than in the audio-only mode in conversational speech with transparent mask ($b = 2.23$, $SE = 0.59$, $p < 0.01$), and in the audiovisual mode in conversational speech with transparent mask ($b = 2.34$, $SE = 0.59$, $p < 0.01$).

The GLMM results also showed that in audio-only mode in conversational speech with disposable face mask, the number of correct answers was significantly lower than in the audiovisual mode in clear speech with transparent mask and no mask ($b = -3.88$, $SE = 0.61$, $p < 0.01$ for with mask; $b = -4.03$, $SE = 0.61$, $p < 0.01$ for with no mask), in the

audio-only mode in clear speech with transparent mask and no mask ($b = -3.28$, $SE = 0.60$, $p < 0.01$ for with mask; $b = -3.32$, $SE = 0.61$, $p < 0.01$ for with no mask), and in the audio-only mode in conversation speech with no mask ($b = -2.65$, $SE = 0.60$, $p < 0.01$). They showed that in the audiovisual mode in conversation speech with disposable face mask, the number of correct answers was significantly lower than that in the audiovisual mode in clear speech with no mask ($b = -2.06$, $SE = 0.59$, $p < 0.05$). They revealed that in the audiovisual mode in conversation speech with no mask, the number of correct answers was significantly lower than that in the audiovisual mode in clear speech with transparent mask ($b = -2.37$, $SE = 0.60$, $p < 0.01$).

The statistical details of the contrasts of type of face masks (no mask: NO, transparent mask: TM, and disposable face mask: DM) in different speaking styles (clear: CLEAR and conversation: CON) and presentation modes (audio only: AO and audiovisual: AV) can be found in Table B.

5 | DISCUSSION

In the present study, we set out to investigate the potential effect of three variables, *type of face mask*, *presentation mode*, and *speaking style*, on L2 speech perception among Thai learners of English. As the findings reported above suggest, it is the interactions among these factors that turn out to be worthy of exploration. In general, we have found a great impact of clear speech on enhancing speech intelligibility for the participants. The finding that there was no difference in the scores of correct answers among different face mask conditions in clear speech in both audio-only and audiovisual modes contradicts previous studies which showed a negative effect of face masks in the perception of speech signals (e.g., Bottalico et al., 2020; Brown et al., 2021; Cohn et al., 2021; Guayasamin, 2021; Moon et al., 2022; Schwarz et al., 2022). The impact of clear speech has an even greater impact than face mask wearing and presentation mode. In other words, irrespective of the type of face mask and presentation mode, speaking clearly improved the listening ability of L2 learners. This is shown by the first six highest scores of the speech intelligibility that occurred when the L2 learners listened to the stimuli with clear speech across type of face masks and presentation modes even though the aggregate for each was not high (see Figure 1). The effect of clear speech could also be seen across face mask settings when the number of correct answers in the clear speech in the audiovisual mode was significantly higher than that in the conversational style. The result of the benefit of clear speech is consistent with the previous studies reviewed (e.g., Smiljanic & Bradlow, 2007; Smiljanić & Bradlow, 2011). The results also demonstrated that the conditions with clear speech, disposable face mask with both audiovisual and audio-only modes were more intelligible than all other conditions with conversational speech. This might be because speakers specifically modify their speech clarity when wearing face masks, that is, the speech tended to be more effort-intensive and specially aimed for clarity (Cohn et al., 2021). This concurs with Guskaroska (2021) which found that when wearing face masks, speakers tend to employ various strategies to adjust their speech to aid communication. Such strategies include pronouncing sounds or words more clearly to enhance understanding in the speech, speaking more slowly, loudly, and with a wider intonation range. When these adjustments are present in clear speech, the speech intelligibility increases. This leads to the significant differences observed parallel to findings in our study.

In addition, no significant difference in the L2 perception was found between these six types of clear-speech stimuli, regardless of face mask types and presentation modes. In conjunction with the observed difficulties participants had for conversational speech, our findings corroborate with Yi et al. (2021) who found that in most situations, conversational speech are less intelligible than clear speech. This suggests that speaking clearly has a greater effect on speech intelligibility than presentation mode and face mask type. In other words, if the talker speaks clearly, the intelligibility of the speech is high regardless of whether the listener receives visual information and/or wearing mask or not. It also suggests that to increase understanding of speech as perceived by L2 learners, it is important to speak clearly. To be specific, teachers might use strategies to make their speech clearer to the L2 learners such as accurately and precisely pronouncing sounds, taking pauses between phrases and sentences, and speaking at a slower pace.

Our findings also showed that the conversational speech produced with disposable mask in audio-only mode received the lowest scores of speech intelligibility. This is not surprising as this condition contains many variables that were expected to hinder communication. The speech understanding of this type of stimuli was significantly lower than the clear speech with disposable mask across presentation modes, and the clear speech in audiovisual mode regardless of face mask types. These findings support the hypothesis that when the listeners lacked visual cues while listening to conversational-style speech which was spoken with disposable mask on, they would find it the most difficult.

This study found no audiovisual effect across all conditions. This might be because the compensatory effect of "clear speech" was so salient that any "visual" information parallel to what existing literature reported was masked/are not able to emerge in the statistical model. The second explanation is that simultaneous processing of visual and aural information might raise cognitive burden, leading to mental fatigue and lowering speech comprehension.

When comparing the pairs of conversational speech, we found no significant difference in all pairs, except for the contrast between audio-only mode in conversational speech with no mask and audio-only mode in conversational speech with disposable face mask on. Conversational speech with no mask spoken in the audio-only mode was found to significantly improve the speech intelligibility among L2 learners in comparison with the one with disposable face mask. This supports the claim that wearing face mask might hinder speech communication. This finding finds support in various previous studies (e.g., Bottalico et al., 2020; Brown et al., 2021; Guayasamin, 2021). For example, Keerstock et al. (2020) which found negative impact of face mask when native speakers were exposed to foreign-accented texts articulated by a non-native English speaker while wearing a face mask in a noisy environment. Similarly, Truong et al. (2021) observed that face mask had a negative impact on German listeners' recall of spoken German sentences. However, Smiljanic et al. (2021) did not find any negative impact of wearing face mask in speech intelligibility between the conversational speech with mask and with no mask by both non-native talker and native talker in quiet settings. This difference could be attributed to the participants and presentation mode. In the study of Smiljanic et al. (2021), the participants were native speakers of English and they listened to stimuli in audiovisual mode. In our study, the participants were non-native speakers at beginner level. These possibly suggest that wearing face mask has more negative effect on L2 speech understanding but this might not be the case for native listeners. They also suggest that although not being provided with the whole facial expressions, some visual cues, such as eye and eyebrow movements might have been useful in giving listeners some useful speech information.

6 | LIMITATIONS AND DIRECTIONS FOR FUTURE STUDIES

The COVID-19 pandemic has posed ongoing challenges to teaching and research alike, our project was unfortunately not immune from them. This study employed stimuli recorded in a quiet laboratory setting with no background noise, which might be rather different from listening to speech in a more natural setting. Future studies might be carried out with stimuli that contain background noise, such as speech-shaped noise, wide-band noise, and talker babble. However, it could be argued that the use of stimuli without background noise is reasonably representative of the teaching of L2 via emergency online method where the teachers instruct students in a quiet room over the Internet. An additional limitation of this study could potentially stem from the one-time participation of each participant, coupled with the uniformity of presentation order for all individuals. In future investigations, it may be beneficial to implement a counter-balanced presentation order scheme. This methodological adjustment could facilitate a more nuanced examination, enabling us to disentangle the potential confounding influences of "during task improvements" or the impact of fatigue. Also, since this study was carried out during the peak of the pandemic it was not possible to request participation beyond individuals' homes. This also meant that we had less control over the research procedure, since participants listened to stimuli over their own headphones. The clarity and loudness of the test items were not controlled for by us. It is possible that some participants might hear the stimuli better than other participants because of their headphone setting. However, allowing the participants to use their headphones and partaking in the research at their own home is perhaps the best option out of an unprecedented situation when on site teaching in Thailand

was limited. At that time, only online instruction was allowed, and therefore online research seemed to be the most appropriate method to prevent the spread of the virus.

As the world is slowly recovering from the pandemic, future studies could rectify some of these limitations, for example, by having participants listen to the stimuli over the headphone with standardized settings in a sound-proof booth to ensure the sound quality. Follow-up work can also collect reaction time data. This information could be useful in estimating the amount of effort involved in processing speech across the testing conditions.

7 | CONCLUSION

The impacts of type of mask, presentation mode, and speaking style have been explored in this study. Even though previous studies have investigated each of these factors in isolation to an extent particularly in an L1 context, at the time of writing, our research is the first to consider how the three factors interact in an emergency online L2 teaching and learning setting. Our findings demonstrated that speaking with clear speech is useful in aiding L2 speech perception. The results showed that when disposable mask was involved, L2 learners did significantly better with clear speech, as opposed to speech in conversational style. None of the pairings that involved L2 clear speech was found to be significantly different from one another in the perception task. Speaking clearly could help with difficult communication circumstances, such as diminished audio signals and the (partial) absence of visual clues caused by the wearing of a face mask. While the results showed that perceiving L2 speech in audio-only mode with conversational style produced with a disposable face mask was the most challenging out of all tested conditions for L2 listeners, the audiovisual presentation with clear speech was also found to be beneficial to the processing of L2 communication. We hope that being aware of these findings will help facilitate communications among interlocutors including in teaching and learning settings where the availability of audiovisual cues is (partially) impeded (e.g., emergency online teaching and/ or when face masks are involved).

ACKNOWLEDGMENTS

We would like to thank the participants for taking part in our study. We would also like to thank the editors, the anonymous reviewers, and the editorial team for their input on our initial submission. Their feedback has helped enhance our work. We would like to thank Hoyoung Yi for sharing their stimuli.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Patchanok Kitikanan https://orcid.org/0000-0003-1802-3182
Alex Ho-Cheong Leung https://orcid.org/0000-0003-2175-9860

TRANSPARENT PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/ijal.12553

ENDNOTE

[1] We would like to thank one of the reviewers in helping us draw out the important areas to note in our statistical results which are directly relevant to our research questions. These are now highlighted in Table B.

REFERENCES

Abraham, L. B. (2007). Second-language reading comprehension and vocabulary learning with multimedia. *Hispania*, *90*, 98–108.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Becker, S. R., & Sturm, J. L. (2017). Effects of audiovisual media on L2 listening comprehension: A preliminary study in French. *Calico Journal*, *34*(2), 147–177. https://doi.org/10.1558/cj.26754

Bottalico, P., Murgia, S., Puglisi, G. E., Astolfi, A., & Kirk, K. I. (2020). Effect of masks on speech intelligibility in auralized classrooms. *The Journal of the Acoustical Society of America*, *148*(5), 2878–2884. https://doi.org/10.1121/10.0002450

Bricker, D. (2020). *More people say they're wearing masks to protect themselves from COVID-19 since March. There's been an increase in the number of people wearing masks. Game Changers (Ipsos)*. https://www.ipsos.com/en/more-people-say-theyre-wearing-masks-protect-themselves-covid-19-march

Brown, V. A., Van Engen, K. J., & Peelle, J. E. (2021). Face mask type affects audiovisual speech intelligibility and subjective listening effort in young and older adults. *Cognitive Research: Principles and Implications*, *6*(1), 49. https://doi.org/10.1186/s41235-021-00314-0

Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A., & Ghazanfar, A. A. (2009). The natural statistics of audiovisual speech. *PLoS Computational Biology*, *5*(7), e1000436. https://doi.org/10.1371/journal.pcbi.1000436

Cho, T. (2005). Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of/ɑ, i/in English. *The Journal of the Acoustical Society of America*, *117*(6), 3867–3878.

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. https://par.nsf.gov/servlets/purl/10275347

Cole, J., Kim, H., Choi, H., & Hasegawa-Johnson, M. (2007). Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from Radio News speech. *Journal of Phonetics*, *35*(2), 180–209. https://doi.org/10.1016/j.wocn.2006.03.004

Corey, R. M., Jones, U., & Singer, A. C. (2020). Acoustic effects of medical, cloth, and transparent face masks on speech signals. *The Journal of the Acoustical Society of America*, *148*(4), 2371–2375. https://doi.org/10.1121/10.0002279

Erber, N. P. (1975). Auditory-visual perception of speech. *Journal of Speech and Hearing Disorders*, *40*(4), 481–492. https://doi.org/10.1044/jshd.4004.481

Guayasamin, M. A. C. (2021). *Face masks and speech perception: Emotions and intelligibility perceived by monolingual and bilingual speakers*. University of Kentucky.

Guskaroska, A. (2021). The role of face masks on speech intelligibility for native and nonnative speakers of English: Graduate students' experiences. *Journal of Academic Language and Learning*, *15*(1), 117–133.

Hardison, D. M. (2021). Multimodal input in second-language speech processing. *Language Teaching*, *54*(2), 206–220. https://doi.org/10.1017/S0261444820000592

Hardison, D. M., & Pennington, M. C. (2021). Multimodal second-language communication: Research findings and pedagogical implications. *RELC Journal*, *52*(1), 62–76. https://doi.org/10.1177/0033688220966635

Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication*, *47*(3), 360–378. https://doi.org/10.1016/j.specom.2005.04.007

Hubbard, D. J., & Kushner, D. (1980). A comparison of speech intelligibility between esophageal and normal speakers via three modes of presentation. *Journal of Speech, Language, and Hearing Research*, *23*(4), 909–916. https://doi.org/10.1044/jshr.2304.909

Jones, L. C., & Plass, J. L. (2002). Supporting listening comprehension and vocabulary acquisition in French with multimedia annotations. *The Modern Language Journal*, *86*(4), 546–561.

Keerstock, S., Meemann, K., Ransom, S. M., & Smiljanic, R. (2020). Effects of face masks and speaking style on audio-visual speech perception and memory. *The Journal of the Acoustical Society of America*, *148*(4), 2747. https://doi.org/10.1121/1.5147635

Lee, B. J., & Hart, E. T. (2022). Facemask occlusion's impact on L2 listening comprehension. *Speech Communication*, *139*, 45–50. https://doi.org/10.1016/j.specom.2022.03.005

Lenth, R., Love, J., & Herve, M. (2018). *emmeans: Estimated marginal means, aka least-squares means*. https://CRAN.R-project.org/package=emmeans

Li, C.-H. (2006). *What you know and see can help you: An examination of the effects of advance organizers, visual images, and gender differences on the video-based listening comprehension of EFL college students*. Unpublished doctoral dissertation, Indiana University of Pennsylvania, Indiana.

Magee, M., Lewis, C., Noffs, G., Reece, H., Chan, J. C., Zaga, C. J., Paynter, C., Birchall, O., Azocar, S. R., Ediriweera, A., Kenyon, K., Caverlé, M. W., Schultz, B. G., & Vogel, A. P. (2020). Effects of face masks on acoustic analysis and speech perception:

Implications for peri-pandemic protocols. *The Journal of the Acoustical Society of America*, *148*(6), 3562–3568. https://doi.org/10.1121/10.0002873

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746–748.

Monsen, R. B. (1983). The oral speech intelligibility of hearing-impaired talkers. *Journal of Speech and Hearing Disorders*, *48*(3), 286–296. https://doi.org/10.1044/jshd.4803.286

Moon, I.-J., Jo, M., Kim, G.-Y., Kim, N., Cho, Y.-S., Hong, S.-H., & Seol, H.-Y. (2022). How does a face mask impact speech perception? *Paper presented at the Healthcare*, *10*(9), 1709. https://doi.org/10.3390/healthcare10091709

Schwarz, J., Li, K. K., Sim, J. H., Zhang, Y., Buchanan-Worster, E., Post, B., Gibson, J. L., & McDougall, K. (2022). Semantic cues modulate children's and adults' processing of audio-visual face mask speech. *Frontiers in Psychology*, *13*, 879156. https://doi.org/10.3389/fpsyg.2022.879156

Smiljanic, R., & Bradlow, A. (2007). *Proceedings of the XVIth International Congress of Phonetic Sciences*, 661–664.

Smiljanić, R., & Bradlow, A. R. (2009). Speaking and hearing clearly: Talker and listener factors in speaking style changes. *Language and Linguistics Compass*, *3*(1), 236–264. https://doi.org/10.1111/j.1749-818X.2008.00112.x

Smiljanić, R., & Bradlow, A. R. (2011). Bidirectional clear speech perception benefit for native and high-proficiency non-native talkers and listeners: Intelligibility and accentedness. *The Journal of the Acoustical Society of America*, *130*(6), 4020–4031. https://doi.org/10.1121/1.3652882

Smiljanic, R., Keerstock, S., Meemann, K., & Ransom, S. M. (2021). Face masks and speaking style affect audio-visual word recognition and memory of native and non-native speech. *The Journal of the Acoustical Society of America*, *149*(6), 4013–4023. https://doi.org/10.1121/10.0005191

Tiippana, K. (2014). What is the McGurk effect? *Frontiers in Psychology*, *5*, 725. https://doi.org/10.3389/fpsyg.2014.00725

Truong, T. L., Beck, S. D., & Weber, A. (2021). The impact of face masks on the recall of spoken sentences. *The Journal of the Acoustical Society of America*, *149*(1), 142–144. https://doi.org/10.1121/10.0002951

Wood, D., Adeline, S., Talbot, R., & Wilburn, T. (2020). Coronavirus world map: Tracking the spread of the outbreak. *National Public Radio*. Retrieved from [Coronavirus world map: Tracking the spread of the outbreak].

Yi, H., Pingsterhaus, A., & Song, W. (2021). Effects of wearing face masks while using different speaking styles in noise on speech intelligibility during the COVID-19 pandemic. *Frontiers in Psychology*, *12*, 682677. https://doi.org/10.3389/fpsyg.2021.682677

Zheng, Y., McPherson, K., & Smith, P. (2014). Effects of early and late treatment with L-baclofen on the development and maintenance of tinnitus caused by acoustic trauma in rats. *Neuroscience*, *258*, 410–421. https://doi.org/10.1016/j.neuroscience.2013.11.032

How to cite this article: Kitikanan, P., & Leung, A. H.-C. (2024). Wearing face masks in different speech styles during the COVID-19 pandemic: A study of Thai L2 English learners. *International Journal of Applied Linguistics*, *34*, 1123–1140. https://doi.org/10.1111/ijal.12553

APPENDIX 1

TABLE A The target sentences and their contexts (type of face masks [no mask: NO, transparent mask: TM, disposable face mask: DM]; different speaking styles [conversation: CON and clear: CLEAR]; and presentation modes [audio only: AO and audiovisual: AV]).

No.	Sentence	Context
1	The CHEF COOKS PASTA EVERY day	AV-CON-NO
2	The STRESSFUL WEEK ENDED at the BAR	AV-CON-DM
3	These BROWN MUSHROOMS TASTE AMAZING	AO-CL-DM
4	The MAP SHOWS the CITY ROADS	AO-CL-NO
5	Our COUSIN STARTS SCHOOL TOMORROW	AV-CL-DM
6	His FIRST GIRLFRIEND was ATTRACTIVE and SMART	AV-CON-NO
7	The STRONG WIND COOLED the AIR	AO-CL-NO
8	The GIRL LOVES SWEET CANDY	AV-CON-DM
9	The MARKET was CROWDED and TOO FAR	AV-CL-DM
10	The CROWD WATCHED the TALENTED PERFORMER	AO-CON-DM
11	The FIVE STUDENTS were LATE for CLASS	AV-CL-NO
12	The THIRSTY KID DRINKS JUICE	AV-CON-TM
13	The PRIVATE UNIVERSITY is NOT CHEAP	AV-CL-NO
14	The GROUP HEARD SLOW MUSIC	AV-CON-TM
15	My GRANDPARENTS TOOK PICTURES on VACATION	AO-CON-DM
16	The CHICKEN SANDWICH CAME with SALAD	AV-CON-TM
17	The CLEAN BEACHES have CLEAR WATER	AV-CL-TM
18	The FATHER and DAUGHTER SAW the MOVIE	AO-CON-TM
19	The KING and QUEEN PLANNED a PARTY	AV-CON-NO
20	That PRETTY GIRL WON a PRIZE.	AO-CON-NO
21	The CHEF MADE FRESH NOODLES.	AV-CL-TM
22	The YOUNGEST SISTER WATCHES TV	AV-CL-TM
23	She WASHED and DRIED her CURLY HAIR	AO-CON-TM
24	The RAIN LASTED for MANY WEEKS	AV-CL-NO
25	The HAPPY CHILDREN LAUGH at the STORY	AO-CON-NO
26	The TALENTED ARTIST DREW a PICTURE	AV-CON-NO
27	The MACHINE MADE a TERRIBLE NOISE	AO-CL-DM
28	The RABBIT and MOUSE EXPLORED the FIELD	AO-CL-TM
29	The STUDENT STUDIES in the QUIET ROOM	AV-CON-DM
30	The GOAT EATS DRY LEAVES	AV-CON-NO
31	The FOREIGN LADY DREAMED of her HOME	AO-CON-TM
32	The STRONG ARMY WON the BATTLE	AV-CL-TM

(Continues)

TABLE A (Continued)

No.	Sentence	Context
33	The POPULAR CLUB is OFTEN FULL	AO-CL-NO
34	The HUNGRY TEENAGERS EAT SNACKS	AO-CL-TM
35	My YELLOW SHOES CAME in a BAG	AO-CL-DM
36	The WHITE HORSE LIVES on a FARM	AO-CON-NO
37	The HOT SUN WARMED the POOL	AV-CL-DM
38	The SHY GUEST SPEAKS QUIETLY	AV-CON-TM
39	The SAD PETS NEED LOVE	AO-CON-DM
40	That BOOK COSTS TEN DOLLARS	AV-CL-TM
41	The MEAN TEACHER is NEVER NICE	AV-CL-DM
42	The FLAGS FLY HIGH and PROUD	AO-CL-DM
43	The PAINTER USED SOFT BRUSHES	AO-CL-TM
44	His SPEECH was BORING and TOO LONG	AV-CL-NO
45	The VEGETABLES GREW in the GREEN GARDEN	AV-CON-TM
46	My BROTHER SLEEPS until LATE MORNING	AO-CL-TM
47	Her LOUD COUGH SOUNDED HORRIBLE	AV-CL-NO
48	The ARTIST VISITED MANY MUSEUMS	AO-CL-DM
49	The BLUEBERRY PIE BAKED in the OVEN	AO-CON-NO
50	The GARDENER GREW COLORFUL PEPPERS	AO-CON-NO
51	Our BAND PRACTICES in my SMALL GARAGE	AO-CL-NO
52	Her FAVORITE PANTS WERE RUINED	AO-CON-TM
53	The THREE COUSINS did their MATH HOMEWORK	AO-CON-DM
54	Her UNCLE WAITS QUIETLY for the ANSWER	AO-CL-NO
55	The ADULTS LEARNED to DANCE in SCHOOL	AV-CON-DM
56	The ANGRY BEAR SCARED the CAMPERS	AO-CL-TM
57	The MAN LOST his HOUSE KEY	AV-CL-DM
58	He LOST his WHITE HAT TODAY	AO-CON-DM
59	The CHEAP DRINKS ATTRACT CUSTOMERS	AV-CON-DM
60	A GOOD FRIEND TELLS the TRUTH	AO-CON-TM

TABLE B [1] The contrasts of type of face masks (no mask: NO, transparent mask: TM, and disposable face mask: DM) in different speaking styles (conversation: CON and clear: CLEAR), and presentation modes (audio only: AO and audiovisual: AV), based on Tukey's HSD post hoc test.

Contrast	Estimate	Standard error	*z* value	*p*-value
AO,CLEAR,DM—AO,CON,DM	4.737	0.606	7.812	<0.0001
AO,CLEAR,DM—AV,CON,DM	2.760	0.594	4.646	0.0002
AO,CLEAR,DM—AO,CON,NO	2.089	0.592	3.530	0.0213
AO,CLEAR,DM—AV,CON,NO	3.226	0.596	5.414	<0.0001
AO,CLEAR,DM—AO,CON,TM	3.091	0.595	5.196	<0.0001
AO,CLEAR,DM—AV,CON,TM	3.200	0.594	5.384	<0.0001
AV,CLEAR,DM—AO,CON,DM	5.134	0.613	8.380	<0.0001
AV,CLEAR,DM—AV,CON,DM	3.157	0.597	5.288	<0.0001
AV,CLEAR,DM—AO,CON,NO	2.486	0.596	4.174	0.0018
AV,CLEAR,DM—AV,CON,NO	3.623	0.600	6.036	<0.0001
AV,CLEAR,DM—AO,CON,TM	3.488	0.598	5.832	<0.0001
AV,CLEAR,DM—AV,CON,TM	3.597	0.599	6.001	<0.0001
AV,CLEAR,NO—AO,CON,TM	2.386	0.594	4.015	0.0035
AV,CLEAR,NO—AV,CON,TM	2.496	0.595	4.194	0.0016
AV,CLEAR,NO—AV,CON,NO	2.522	0.596	4.229	0.0014
AV,CLEAR,TM—AO,CON,TM	2.229	0.593	3.760	0.0093
AV,CLEAR,TM—AV,CON,TM	2.339	0.593	3.946	0.0045
AO,CON,DM—AV,CLEAR,TM	−3.876	0.605	−6.401	<0.0001
AO,CON,DM—AV,CLEAR,NO	−4.033	0.608	−6.634	<0.0001
AO,CON,DM—AO,CLEAR,TM	−3.282	0.606	−5.411	<0.0001
AO,CON,DM—AO,CLEAR,NO	−3.317	0.606	−5.477	<0.0001
AO,CON,DM—AO,CON,NO	−2.647	0.606	−4.370	0.0008
AV,CON,DM—AV,CLEAR,NO	−2.056	0.594	−3.461	0.0269
AV,CON,NO—AV,CLEAR,TM	−2.365	0.595	−3.977	0.0040
AO,CLEAR,DM—AV,CLEAR,DM	−0.397	0.593	−0.670	1.0000
AO,CLEAR,DM—AO,CLEAR,NO	1.420	0.589	2.411	0.3981
AO,CLEAR,DM—AV,CLEAR,NO	0.704	0.592	1.190	0.9898
AO,CLEAR,DM—AO,CLEAR,TM	1.455	0.590	2.468	0.3602
AO,CLEAR,DM—AV,CLEAR,TM	0.861	0.591	1.458	0.9516
AV,CLEAR,DM—AO,CLEAR,NO	1.817	0.595	3.056	0.0930
AV,CLEAR,DM—AV,CLEAR,NO	1.101	0.595	1.850	0.7899
AV,CLEAR,DM—AO,CLEAR,TM	1.852	0.595	3.112	0.0795
AV,CLEAR,DM—AV,CLEAR,TM	1.258	0.594	2.119	0.6087
AO,CON,DM—AV,CON,DM	−1.977	0.607	−3.259	0.0514
AO,CON,DM—AV,CON,NO	−1.511	0.610	−2.477	0.3546
AO,CON,DM—AO,CON,TM	−1.646	0.607	−2.712	0.2203
AO,CON,DM—AV,CON,TM	−1.537	0.608	−2.526	0.3237

(Continues)

TABLE B (Continued)

Contrast	Estimate	Standard error	*z* value	*p*-value
AV,CON,DM−AO,CLEAR,NO	−1.340	0.591	−2.267	0.5003
AV,CON,DM−AO,CON,NO	−0.671	0.593	−1.132	0.9933
AV,CON,DM−AV,CON,NO	0.466	0.596	0.781	0.9998
AV,CON,DM−AO,CLEAR,TM	−1.305	0.591	−2.206	0.5452
AV,CON,DM−AV,CLEAR,TM	−1.899	0.592	−3.206	0.0604
AV,CON,DM−AO,CON,TM	0.331	0.595	0.556	1.0000
AV,CON,DM−AV,CON,TM	0.440	0.595	0.740	0.9999
AO,CLEAR,NO−AV,CLEAR,NO	−0.716	0.590	−1.213	0.9881
AO,CLEAR,NO−AO,CON,NO	0.669	0.589	1.136	0.9931
AO,CLEAR,NO−AV,CON,NO	1.806	0.594	3.039	0.0975
AO,CLEAR,NO−AO,CLEAR,TM	0.035	0.589	0.060	1.0000
AO,CLEAR,NO−AV,CLEAR,TM	−0.559	0.589	−0.949	0.9986
AO,CLEAR,NO−AO,CON,TM	1.671	0.592	2.822	0.1704
AO,CLEAR,NO−AV,CON,TM	1.780	0.592	3.004	0.1071
AV,CLEAR,NO−AO,CON,NO	1.385	0.592	2.339	0.4486
AV,CLEAR,NO−AO,CLEAR,TM	0.751	0.591	1.270	0.9827
AV,CLEAR,NO−AV,CLEAR,TM	0.157	0.591	0.266	1.0000
AO,CON,NO−AV,CON,NO	1.137	0.595	1.910	0.7531
AO,CON,NO−AO,CLEAR,TM	−0.634	0.590	−1.074	0.9957
AO,CON,NO−AV,CLEAR,TM	−1.228	0.590	−2.081	0.6365
AO,CON,NO−AO,CON,TM	1.001	0.593	1.688	0.8739
AO,CON,NO−AV,CON,TM	1.111	0.594	1.871	0.7773
AV,CON,NO−AO,CLEAR,TM	−1.771	0.595	−2.978	0.1151
AV,CON,NO−AO,CON,TM	−0.135	0.597	−0.227	1.0000
AV,CON,NO−AV,CON,TM	−0.026	0.598	−0.044	1.0000
AO,CLEAR,TM−AV,CLEAR,TM	−0.594	0.589	−1.009	0.9975
AO,CLEAR,TM−AO,CON,TM	1.635	0.592	2.763	0.1958
AO,CLEAR,TM−AV,CON,TM	1.745	0.593	2.940	0.1270
AO,CON,TM−AV,CON,TM	0.109	0.595	0.184	1.0000