

## Article

# Visual Cues to Speakers' Religious Affiliation and Listeners' Understanding of Second Language French Speech

Sara Kennedy <sup>1,\*</sup>, Pavel Trofimovich <sup>1</sup> , Rachael Lindberg <sup>1</sup>  and Oguzhan Tekin <sup>2</sup>

<sup>1</sup> Department of Education, Concordia University, Montréal, QC H3G 1M8, Canada; pavel.trofimovich@concordia.ca (P.T.); rachael.lindberg@concordia.ca (R.L.)

<sup>2</sup> Institute for the Study of University Pedagogy, University of Toronto Mississauga, Mississauga, ON L5L 1C6, Canada; oguzhan.tekin@utoronto.ca

\* Correspondence: sara.kennedy@concordia.ca

**Abstract:** Previous research has shown that speakers' visual appearance influences listeners' perception of second language (L2) speech. In Québec, Canada, the context of this study, pandemic mask mandates and a provincial secularism law elicited strong societal reactions. We therefore examined how images of speakers wearing religious and nonreligious coverings such as medical masks and headscarves influenced the comprehensibility (listeners' ease of understanding) and intelligibility of L2 French speech. Four L2 French women from first language (L1) Arabic backgrounds wore surgical masks while recording 40 sentences from a standardized French-language speech perception test. A total of 104 L1 French listeners transcribed and rated the comprehensibility of the sentences, paired with images of women in four visual conditions: uncovered face, medical mask, hijab (headscarf), and niqab (religious face covering). Listeners also completed a questionnaire on attitudes toward immigrants, cultural values, and secularism. Although intelligibility was high, sentences in the medical mask condition were significantly more intelligible and more comprehensible than those in the niqab condition. Several attitudinal measures showed weak correlations with intelligibility or comprehensibility in several visual conditions. The results suggest that listeners' understanding of L2 sentences was negatively affected by images showing speakers' religious affiliation, but more extensive follow-up studies are recommended.

**Keywords:** intelligibility; comprehensibility; second language; French; religious affiliation; medical mask; attitudes; stereotyping; hijab; niqab



**Citation:** Kennedy, Sara, Pavel Trofimovich, Rachael Lindberg, and Oguzhan Tekin. 2024. Visual Cues to Speakers' Religious Affiliation and Listeners' Understanding of Second Language French Speech. *Languages* 9: 154. <https://doi.org/10.3390/languages9050154>

Academic Editor: Claudine Bowyer-Crane

Received: 18 February 2024

Revised: 12 April 2024

Accepted: 15 April 2024

Published: 24 April 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Listeners rely on various sources of information, including a speaker's appearance, to understand speech (Babel 2022). In a classic study, Rubin (1992) paired a lecture recorded by an American English speaker with an image of either a Caucasian or an Asian woman. Although the speaker was identical in both cases, the students who were led to believe that the speaker was Asian downgraded the speaker in social evaluations and showed diminished understanding of the lecture compared to those who thought that the speaker was Caucasian. In essence, listeners engaged in stereotyping, where their assumptions about the speaker, such as how accented she sounds or how easy she is to understand, influenced their experience with speech. Even though appearance-driven stereotyping is a popular area of research, where the role of speaker ethnicity has been explored in relation to various listener-based measures (e.g., Kang and Rubin 2009; Kang and Yaw 2021), it is unclear how visual cues which signal a speaker's religious affiliation impact the listener. Our goal was therefore to examine whether listeners' understanding of second language (L2) speech differs for utterances presented with images showing religious (hijab and niqab) versus nonreligious (surgical mask) face coverings. Broadly speaking, we focused on the role of visual cues in L2 speech perception, examining whether listeners' assumptions and attitudes have implications for how L2 speech is understood.

### 1.1. Background Literature

Speakers' visual characteristics can affect listeners' understanding of speech as assessed via objective and subjective measures (Witkowska et al. 2023). Objective measures include those which capture listeners' accuracy in decoding speech (e.g., transcription), whereas subjective measures involve those where listeners evaluate their ease or difficulty of understanding (e.g., via scalar ratings of comprehensibility). In some studies, for example, when first language (L1) English listeners viewed images of non-Caucasian faces and evaluated speech samples from L1 English speakers, listeners transcribed the speech less accurately compared to when they viewed images of Caucasian faces or no pictures at all (Babel and Russell 2015; Kutlu et al. 2022a). Other work revealed no difference in listeners' transcriptions or comprehensibility ratings after hearing L1 Dutch or English speech while viewing Caucasian or non-Caucasian faces (Hanulíková 2018; Melguy and Johnson 2021). For L2 speech, Taylor Reid (2022) similarly found no differences when English-dominant listeners rated the comprehensibility of speech paired with faces of five ethnicities. Thus, listeners' assumptions of a speaker's ethnicity can influence their understanding of speech. However, these effects appear to be context-specific in that they are more or less likely to emerge for listeners residing in or drawn from locations that are linguistically and culturally less versus more diverse (Kutlu et al. 2022a).

To date, only one study known to us has examined the evaluation of speech paired with guises with or without visual religious symbols. Lybaert et al. (2022) presented L1 Dutch Belgian listeners with a female speaker using Standard Dutch or Colloquial Dutch, with the visual guise involving a combination of a Flemish or Maghrebi (Moroccan) name and a picture of a woman with or without a hijab (headscarf). Listeners' comprehensibility ratings did not differ between any of the guises insofar as a speaker paired with the hijab guise was rated as easy to understand as those presented in other guises. In contrast, research findings on listeners' understanding of speakers either wearing or presented as wearing face coverings are more straightforward (Badh and Knowles 2023). In conditions without background noise, L1 English listeners understand L1 English speakers physically wearing surgical masks or face coverings such as balaclavas to the same degree as they understand speakers wearing no masks (Llamas et al. 2008; Toscano and Toscano 2021). In the few studies where speakers who recorded audio samples via medical masks were paired with images of the masked speakers (including L2 speakers), listeners often understood the speakers significantly better than non-masked speakers, likely due to greater attention paid to masked speech (Pycha et al. 2022; Smiljanic et al. 2021). Thus, when listeners' assumptions are not manipulated, there is often little difference in how listeners understand speakers with or without face coverings (including religious ones). When such differences emerge, they occur across many types of face coverings, such as masks, balaclavas, neck gaiters, or a piece of tape across a speaker's mouth (Badh and Knowles 2023).

Whether or not listeners' attitudes are manipulated via speech presented in various visual guises, the relationship between a speaker's appearance and listeners' understanding likely depends on listeners' reported attitudes and experience. Lindemann (2002) showed that L1 English interlocutors with negative attitudes toward L1 Korean speakers rated their experience of communicating in English with L1 Korean speakers significantly lower than did L1 English interlocutors with positive attitudes. Sheppard et al. (2017) presented excerpts from L2 English speakers varying in intelligibility to university faculty who differed in their experience teaching L2 students. Faculty who demonstrated more negative attitudes toward international students rated the speakers less comprehensible than did faculty with more positive attitudes, and faculty with relatively less experience with L2 speech also rated comprehensibility lower compared to more experienced faculty. Jeong et al. (2021) focused on intelligibility, comprehensibility, and acceptance of a speaker's accent for English speech samples from speakers of six different L1s. Comprehensibility was positively correlated with acceptance attitudes for all speakers except for the L1 Tamil speaker, suggesting that listeners perceived the speech to be easily understood when they had favorable attitudes toward it (see also Simon et al. 2022). Generally speaking,

then, listeners who have positive attitudes toward or greater experience with L2 speech or speakers also tend to find L2 speech more intelligible and comprehensible.

### 1.2. The Current Study

Our aim in this study was to examine a previously underexplored research area, namely, the role of visual cues to a speaker's religious identity in listeners' understanding of L2 speech. In Québec (Canada), the context of our work, two recent events created circumstances well suited for addressing this objective. In 2019, the provincial law Bill 21 was passed with the goal of maintaining secularism in Québec society ([Loi sur la laïcité de l'État 2019](#)). Supported by a majority of residents ([Leger Marketing 2021](#)), Bill 21 restricts the hiring and career paths of certain governmental employees (e.g., teachers, prosecutors) who wear visible religious symbols at work and requires people accessing public services to show an open (uncovered) face. And in response to the increase in COVID-19 cases during the pandemic, the Québec government enacted an almost two-year mandate for obligatory mask-wearing in public indoor spaces, which ended in May 2022. Initially very high, the support for mask-wearing was still substantial by the mandate's end ([Institut national de santé publique du Québec 2020, 2022](#)), even though some residents associated mask-wearing with authoritarian restriction of liberties ([Stewart et al. 2023](#)). In Québec, therefore, attitudes toward the use of face coverings exist in the context of a law restricting the wearing of religious symbols, including such face coverings as niqab (a veil covering a woman's face, including mouth and nose) and hijab in certain contexts, and in the context of an extended pandemic mask mandate for public indoor spaces.

We took advantage of this unique situation to explore how L1 speakers of Québec French, who represent the linguistic and cultural majority in the province, understand L2 French speech paired with various images (guises) of covered and uncovered faces. Because Québec residents have in their majority endorsed restrictions on the wearing of religious symbols but have generally supported the use of medical face masks, these attitudinal perspectives create an expectation that a speaker's visual characteristics (i.e., face coverings that do or do not carry religious information) might influence listeners' understanding of L2 speech. We therefore specifically examined how listeners' understanding of L2 French speech is affected by a speaker's wearing of hijab and niqab (religious face coverings), compared to medical masks (nonreligious face coverings) and no face covering at all. These visual conditions differ in the degree to which they obstruct a speaker's face and signal a speaker's religious affiliation. Whereas both niqabs and surgical masks obstruct the view of a speaker's mouth, only niqabs carry information about a speaker's religious affiliation. In contrast, for speakers without face coverings, there is no obstruction to a speaker's mouth, yet hijabs are likely to be perceived as a religious symbol ([Bakht 2022](#); [Piela 2021](#)). This study is unique because few studies in this area have targeted L2 French, and no research, apart from [Lybaert et al. \(2022\)](#), has examined the role of religious symbols in listeners' understanding of L2 speech.

If visual information impacts listeners' experience with L2 speech, we could anticipate that speech presented with images of speakers wearing religious face coverings would be associated with lower listener understanding than the same speech presented alongside images of speakers wearing face masks. Essentially, the societal majority's rejection of public displays of religious symbols might harm understanding of a speaker who is wearing a religious face covering, whereas the general support for mask-wearing should not impair understanding of a speaker with a face mask on. However, if listeners' understanding of L2 speech does not reflect the broader society's attitudes, listeners might generally find it difficult to process L2 speech regardless of the face covering, in which case they would experience similar difficulty listening to speech paired with images of religious and nonreligious face coverings obstructing a speaker's face, compared to the same utterances heard in the context of no face covering. Finally, considering that listeners have person-specific attitudes toward and experience with L2 speech and individual speakers, we also anticipated that listeners' understanding would be associated with measures of their

attitudes and experience with the use of face coverings in the public domain. Our study was guided by two research questions:

1. Do L1 French listeners differ in how intelligible and comprehensible they find L2 utterances when these utterances are presented under different visual conditions?
2. What associations exist between the intelligibility and comprehensibility of L2 utterances presented under different visual conditions and variables capturing listeners' attitudinal and experiential profiles?

## 2. Methods and Materials

### 2.1. Audio Recordings

The target audios were recorded by four female L1 Arabic speakers, all adult learners of French, born and raised in Kuwait (2), Syria, and Libya (1 each). The speakers, who were on average 24 years old ( $SD = 2.92$ ) and had resided in Québec for a mean of 4.52 years ( $SD = 3.12$ ), were pursuing various bachelor's degrees at an English-medium university in Montréal, Québec. Four speakers were recorded rather than a single speaker to minimize the effect of single-speaker familiarity on listeners' understanding (Case et al. 2018). Regarding their French competence, two speakers self-reported their French level as CEFR B1 (low-intermediate), while the other two rated their French proficiency as CEFR B2 (high-intermediate). Using a 100-point scale, they additionally self-assessed their French-speaking skill at a mean of 49.75 ( $SD = 22.90$ ), where 100 meant "very competent." They had been studying French for a mean of 5.69 years ( $SD = 4.17$ ) and reported using the language for a mean of 72% daily ( $SD = 18.88$ ).

To create audio materials, the speakers initially recorded 80 French sentences (5–8 words each) drawn from the *Test de phrases dans le bruit* (Sentences in Noise Test), a standardized instrument for the evaluation of speech perception among French Canadians (Lagacé et al. 2010). The test includes a balanced list of sentences featuring highly frequent vocabulary but varying in semantic predictability of the final word in each sentence. Half of the sentences contain a final word that is predictable from the preceding sentential context (e.g., *Le fermier travaille à la ferme*, "The farmer works at the farm") while the other half contain a final element that is not easily predictable from the context (e.g., *J'ai acheté une nouvelle table*, "I bought a new table"). Because the test was designed for presentation in noise, where the greatest difference in the intelligibility of high- versus low-predictability sentences occurs at a signal-to-noise ratio of  $-1.5$  dB (Lagacé et al. 2010), we did not consider predictability as a variable. Instead, we expected a balanced set of sentences varying in semantic complexity to provide the optimal level of difficulty for L1 French listeners (i.e., avoiding ceiling or floor effects).

The speakers recorded all 80 sentences in individual Zoom sessions with an English–French bilingual researcher. They first practiced reading aloud the first 40 randomly selected sentences for about 10 min, speaking at a natural conversational pace, and then recorded those sentences at their own pace. Each sentence was presented on a separate slide, with fillers included at the beginning and the end to avoid list intonation. The speakers repeated the same procedure after a short break, practicing and then recording the remaining 40 sentences. Importantly, all speakers wore a 3-ply disposable surgical mask while practicing and recording all sentences. The recorded sentences were then checked for quality (in terms of speaking rate and presence of disfluency, incidence of background static and noise, and uniformity of intonation), and the final set of 40 sentences (20 high and 20 low in predictability) was chosen so that the selected sentences sounded most comparable across the four speakers, as judged by the researcher (see Appendix A for the full sentence list). Each sentence was saved as a separate file and then normalized for peak amplitude (90 dB) using MP3Gain.

Because the four speakers varied in their self-assessed French proficiency (from low- to high-intermediate), we recruited external raters to assess the recordings to ensure that the materials were comparable across the speakers. We presented four sentences randomly drawn from each speaker's recordings in a randomized list of 16 sentences for assessment

by 10 L1 Québec French speakers (7 females, 3 males), all young adults (19–30 years old). They used 100-point sliding scales in an online interface to evaluate each recording for comprehensibility (how easy each speaker was to understand) and accentedness (how strongly accented each speaker sounded). The speakers' comprehensibility was rated at a mean of 62.46 ( $SD = 23.87$ ), whereas their accentedness was evaluated at a mean of 44.59 ( $SD = 22.32$ ), where 100 meant “easy to understand” and “not at all accented”, respectively. Comparisons of the four speakers' ratings (using paired-samples  $t$  tests) showed that the four speakers were judged to be similarly comprehensible,  $t_s < 1.42$ ,  $p_s > 0.163$ ,  $d_s < 0.22$ , and similarly accented,  $t_s < 1.73$ ,  $p_s > 0.091$ ,  $d_s < 0.27$ .

## 2.2. Listeners

We recruited 104 L1 Québec French speakers (85 females, 17 males, 1 nonbinary, 1 queer) as listeners using various social media outlets. Listeners, who were on average 26.40 years old ( $SD = 7.18$ ), were all born and raised in Québec, having lived there most of their lives ( $M = 26.07$  years,  $SD = 7.15$ ). They reported varied educational experiences, where their highest current or obtained degree ranged from secondary or diploma credentials (6) to BA (59), MA (23), and PhD (16) degrees. Only nine listeners (8.7%) reported having taken a linguistics course as part of their prior training. Listeners reported predominantly speaking French in everyday interactions and estimated that approximately half of their daily conversations were with L2 French speakers (see Table 1 for a summary of listeners' background characteristics).

**Table 1.** Listeners' background characteristics.

Variable	<i>M</i>	<i>SD</i>	Range
Age (years)	26.40	7.18	20–68
Daily speaking in French (%)	87.00	14.77	14–100
Daily listening in French (%)	66.91	23.45	8–100
L2 French contact (%)	54.19	31.02	0–100
Immigrants should adopt Québec values <sup>a</sup>	63.61	26.53	0–100
It is okay not to speak French well <sup>a</sup>	60.15	28.71	0–100
No to religious symbols or practices in public domain <sup>a</sup>	37.80	31.18	0–100
Religious practices are due to family pressure <sup>a</sup>	50.91	23.10	0–100
Everyone should be free to choose to wear a mask	57.45	37.26	0–100

Note. <sup>a</sup> 100-point agree–disagree scale (100 = “strongly agree”).

Listeners also responded to a brief questionnaire about their attitudes toward immigrants, cultural values, and religious practices in Québec (see Appendix B) so that these reactions could be explored in relation to their responses to L2 speakers' utterances. For attitudes toward immigrants, listeners used a 100-point scale to indicate their agreement with a statement about Québec values (*If people from other countries move to Québec, they should be ready to adopt Québec values*) and a statement about the use of French (*I think it's okay for someone to spend the rest of their life in Québec even if they don't speak French well*), where 0 indicated “strongly disagree” and 100 meant “strongly agree.” To capture their perceptions of Québec secularism, listeners used a similar scale to respond to a statement about the use of religious symbols in the public domain (*Québécois have a right to expect that when they are in secular institutions like government buildings or schools, they will not encounter anyone showing any kind of religious practice or symbol*) and a statement about immigrants' religious practices (*People engage in religious practices like going to confession or wearing a headscarf because of pressure from their family or community*). Finally, to measure listeners' opinion about medical masks, they provided a rating in response to the statement about their freedom to choose to wear or not to wear a mask (*Everyone should be free to make their own choice to wear masks or not to wear them*). As a group (see Table 1), listeners were moderately supportive of the idea that immigrants should adopt Québec values and that speaking French well is not required. They also generally rejected a full ban on displays of religious symbols and



practices and did not strongly endorse the belief that immigrants' religious practices are caused by pressure from their family or community, although individual listeners varied widely in their perceptions. Finally, listeners generally felt neutral regarding their personal freedom to choose to wear medical masks, although their responses varied.

### 2.3. Materials and Procedure

To create different visual conditions, we developed four sets of images, each depicting a young woman (a) not wearing any head covering, (b) wearing a 3-ply surgical mask (which corresponded to the actual procedure followed during stimulus recording), (c) wearing a hijab (headscarf), and (d) wearing a niqab. None of the images represented the actual speakers; the images were drawn from the Bogazici Face Database (Saribay et al. 2018), the London Face Database (DeBruine and Jones 2017), and photos of volunteers from Taylor Reid (2022). The four photographed women were all young adults, portrayed against a neutral background and wearing gray-colored clothing. The original images were pre-rated by 10 external raters, all residents of Montréal, who first selected the depicted woman's ethnicity from five broad ethnic categories and then rated her fit to the chosen category using a 100-point scale (100 = "very typical"). The four women were selected as Middle Eastern at a mean rate of 78%, with no statistically significant difference in the typicality ratings, according to paired-samples *t* tests,  $t_s < 1.63$ ,  $p_s > 0.138$ ,  $d_s < 0.51$ . A graphic designer then created three additional versions of each original image, with each woman shown wearing a medical mask, a headscarf, and a niqab, as illustrated for one speaker in Figure 1 (see Appendix C for the full image set).



**Figure 1.** Illustration of the images used across the visual conditions.

We organized the 40 target sentences from the four speakers in four counterbalanced survey versions using the Latin square design (see Appendix D for item randomizations). Each survey version contained all 40 sentences, with each speaker contributing a unique set of 10 sentences (five high and five low in predictability), balanced for sentence length. In each survey version, each sentence set was then randomly paired with an image of one of four women, again using the Latin square design, so that each set of 10 sentences (representing a unique speaker) was presented in one of four visual conditions (full face, medical mask, headscarf, niqab). The 104 listeners were assigned randomly to one of the four survey versions, with 26 listeners per version. In the end, across all versions, listeners heard each of the four L1 Arabic speakers an equal number of times (each speaker producing a total of 40 sentences over all four versions), with all utterances equally presented in four conditions cued by the four depicted females, each shown in all visual guises.

We evaluated the recordings along two listener-based dimensions: intelligibility and comprehensibility. For intelligibility, considered an objective measure of understanding commonly measured via a listener's transcription accuracy (Derwing and Munro 2015), we captured the accuracy with which listeners transcribed the content of the utterances presented under the four visual conditions. For comprehensibility, which is a holistic, subjective rating of a listener's processing effort (Derwing and Munro 2015), we focused on how easy or difficult to understand were the utterances heard in those visual conditions. After presenting each utterance, therefore, we asked listeners to transcribe it and rate how difficult it was to understand via a 100-point scale, where 0 meant "very difficult" and 100 indicated "very easy".

Data collection took place over approximately 4 weeks in the summer of 2022, soon after the 21-month mandatory mask mandate for all public indoor places was lifted in Québec in early May 2022. Listeners evaluated the target recordings individually via a personal link to a LimeSurvey online interface, with a request to dedicate at least 30 min for this task and to complete it in a quiet location using a headset or good quality audio speakers. Data were collected remotely rather than in a physical lab setting to promote access to the study for listeners who may have found it challenging to travel to and from a university lab for data collection. Before signing the consent form, listeners were informed that the goal of the study was to understand how Québec French speakers evaluate short sentences spoken by L2 French speakers. To illustrate the materials, they were given a sample written sentence that was similar to the ones they would hear. They then read a brief description of the transcription task and the comprehensibility scale and familiarized themselves with the procedure by listening to an L1 English speaker saying a sample French sentence paired with an image of a Caucasian female. Listeners then evaluated the 40 target sentences, each paired with one of four images. After they initiated each new survey page, they saw the image shown at the top of the page and heard the audio played once automatically after 5 s, with 40 sentences presented in a unique random order to each listener (see Appendix E for a screenshot of the rating interface). For the transcription, they typed the content of each utterance in a text box, skipping any word(s) that they could not understand. For comprehensibility, they moved the scale toggle from its initial position at the midpoint (50) to the value that corresponded to their perceived difficulty. Finally, listeners filled out the background questionnaire and completed the attitudes questionnaire. Each received CAD 15 as compensation for participation.

### 3. Data Analysis

We operationalized L2 utterance intelligibility as the accuracy of listeners' transcription of each sentence, where an intelligibility score was defined as the ratio of all accurately transcribed words in each sentence to the total number of words per sentence. Homophonous words (e.g., *a* and *à*, *il mange* and *ils mangent*, *belle passe* and *belles passes*) and misspellings (e.g., *frapper* for *frappé*, *tout* for *tous*, *coverture* for *couverture*) were accepted if they were recognizable; however, words missing morphemes were coded as inaccurate. For example, if a listener transcribed the sentence *Le frappeur a frappé la balle* ("The batter hit the ball") as *Le frappeur a attrapé la balle* ("The batter caught the ball"), the score for this sentence was 5/6 or 0.83; however, if the same sentence was transcribed as *Il fera peur à attraper la balle* ("It will be scary to catch the ball"), the score was 3/6 or 0.50. After an initial coding of all data by one researcher, all transcriptions from 16 randomly selected listeners (15%) were coded by a bilingual research assistant, with global inter-coder reliability (two-way mixed average-measure intraclass correlations) reaching 0.99 and values for individual listeners ranging between 0.97 and 1.00. In light of high coding agreement, the original coding decisions were finalized, and mean intelligibility scores were computed across all utterances, yielding four scores per listener (one in each visual condition). The reliability of listeners' comprehensibility ratings (two-way mixed average-measure intraclass correlations) were also high across all survey versions in the medical mask condition (0.84–0.94), the full face condition (0.91–0.96), the headscarf condition (0.83–0.96), and the niqab condition (0.90–0.97). Given high internal consistency, mean comprehensibility ratings were similarly computed across all utterances per listener, separately in each visual condition.

In terms of various statistical assumptions, the intelligibility scores were negatively skewed, reflecting high utterance intelligibility for listeners, while the comprehensibility ratings showed normal distributions. There were no concerns regarding homogeneity of variances because of the fully within-participants design; however, because the data violated the assumption of sphericity (with Mauchly's tests yielding statistically significant values), the Greenhouse-Geisser correction was applied to report the averaged tests of statistical significance in repeated-measures ANOVAs. For all statistical analyses, the alpha level for statistical significance was set at 0.05 and was Bonferroni-adjusted for multiple

comparisons. Effect sizes were interpreted based on field-specific guidelines (Plonsky and Oswald 2014), using Cohen's  $d$  for repeated-measures comparisons (0.60, 1.00, 1.40) and  $r$  for correlation strength (0.25, 0.40, and 0.60), where each value designates small, medium, and large effects, respectively.

#### 4. Results

##### 4.1. Utterance Intelligibility and Comprehensibility

The first research question asked whether L1 French listeners differed in the intelligibility and comprehensibility of L2 utterances presented under different visual conditions. As summarized in Table 2, the intelligibility scores were on average very high (0.93–0.96), but individual listeners varied in their transcription accuracy (0.60–1.00). Similarly, the mean comprehensibility ratings were generally high (67–72), confirming that the speakers were relatively easy to understand; nevertheless, a fairly broad range of ratings (28–100) suggested that some listeners experienced more effort than other listeners when listening to L2 utterances. Intelligibility and comprehensibility were unrelated in the medical mask condition ( $r = 0.12$ ), but the two dimensions showed small-to-medium associations in the other conditions ( $r = 0.31$ – $0.43$ ).

**Table 2.** Descriptive statistics for sentence intelligibility and comprehensibility and the association between the two measures by listening condition.

Condition	Intelligibility			Comprehensibility			Correlation $r(102)$
	$M (SD)$	95% CI	Range	$M (SD)$	95% CI	Range	
Medical mask	0.96 (0.05)	[0.95, 0.97]	0.60–1.00	72 (15)	[70, 76]	40–100	0.12
Full face	0.94 (0.05)	[0.93, 0.95]	0.71–1.00	72 (16)	[69, 75]	36–99	0.43 ***
Headscarf	0.94 (0.04)	[0.93, 0.95]	0.81–1.00	69 (16)	[66, 72]	28–98	0.31 ***
Niqab	0.93 (0.06)	[0.92, 0.94]	0.62–1.00	67 (16)	[64, 70]	40–98	0.35 ***

Note. \*\*\*  $p < 0.001$  (two-tailed).

In terms of the relationships within each set of measures, as shown in Table 3, the intelligibility scores showed weak-to-medium associations across the four visual conditions, where the two strongest associations emerged between the medical mask and the niqab condition ( $r = 0.54$ ), where the speaker's face was covered, and between the full face and the headscarf condition ( $r = 0.44$ ), where the speaker's face was unobstructed. The comprehensibility ratings revealed medium-to-strong intercorrelations, with the strongest link ( $r = 0.74$ ) observed between the mask and the niqab condition, both of which concealed the speaker's face.

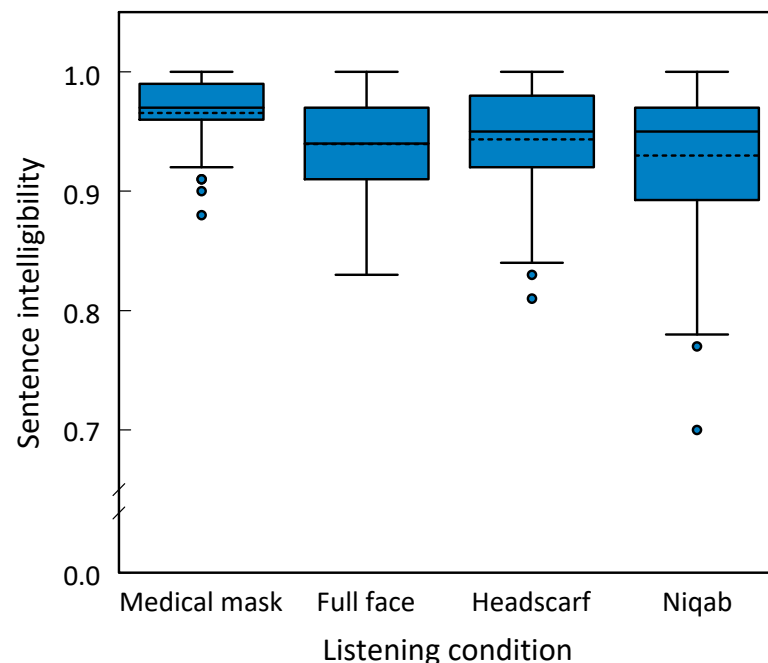
**Table 3.** Intercorrelations between sentence intelligibility scores (bottom triangle) and sentence comprehensibility ratings (top triangle).

Condition	1	2	3	4
1 Medical mask	—	0.51 ***	0.71 ***	0.74 ***
2 Full face	0.39 ***	—	0.70 ***	0.53 ***
3 Headscarf	0.26 **	0.44 ***	—	0.56 ***
4 Niqab	0.54 ***	0.13	0.17	—

Note. \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$  (two-tailed).

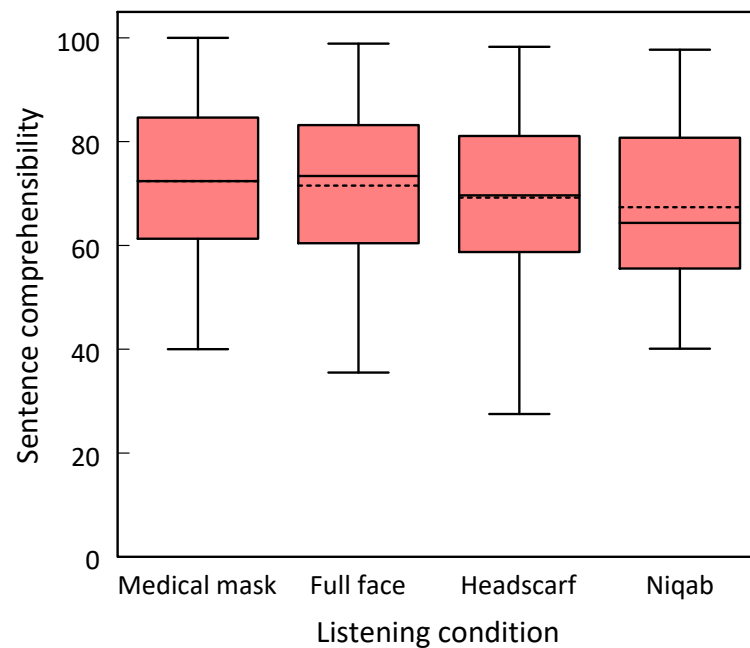


To determine if the intelligibility scores differed across the four visual conditions, we computed an ANOVA where visual condition (medical mask, full face, headscarf, niqab) served as a repeated-measures variable. This analysis revealed a statistically significant  $F$  ratio,  $F(2.28, 234.90) = 12.32, p < 0.001$ , partial  $\eta^2 = 0.11$ . Follow-up pairwise comparisons further indicated that the utterances heard in the medical mask condition were significantly more intelligible than those heard in the full face condition,  $t(103) = 4.99, p < 0.001, d = 0.49$ , the headscarf condition,  $t(103) = 3.86, p = 0.001, d = 0.38$ , and the niqab condition,  $t(103) = 6.38, p < 0.001, d = 0.63$ , all of which were in turn comparable to each other,  $t(103) < 1.95, p > 0.322, d < 0.19$ . As illustrated in Figure 2 and confirmed via nonoverlapping 95% confidence intervals (CIs) for the mean intelligibility scores (Table 2), the most pronounced difference in intelligibility scores concerned L2 utterances presented in the niqab and the medical mask conditions, in favor of the mask condition (with a small effect size).



**Figure 2.** Boxplots of intelligibility scores by listening condition, with dashed lines indicating group means.

A similar ANOVA comparing the comprehensibility ratings also revealed a statistically significant  $F$  ratio,  $F(2.50, 257.08) = 5.91, p < 0.001$ , partial  $\eta^2 = 0.05$ . Follow-up pairwise comparisons further indicated that the utterances heard in the medical mask condition were rated as significantly more comprehensible than those experienced in the headscarf condition,  $t(103) = 2.71, p = 0.048, d = 0.27$ , and the niqab condition,  $t(103) = 4.60, p < 0.001, d = 0.45$ . Additionally, the utterances heard in the full face condition were rated as more comprehensible than those in the niqab condition,  $t(103) = 2.82, p = 0.035, d = 0.28$ . No other comparisons reached statistical significance,  $t(103) < 1.92, p > 0.347, d < 0.18$ . Again, as illustrated in Figure 3 and confirmed further via nonoverlapping 95% CIs for the mean comprehensibility ratings in Table 2, the most pronounced difference for comprehensibility concerned the niqab and the medical mask conditions, in favor of the mask condition (with a small effect size).



**Figure 3.** Boxplots of comprehensibility ratings by listening condition, with dashed lines indicating group means.

#### 4.2. Listener Variables

The second research question focused on possible associations between the intelligibility and comprehensibility of L2 utterances presented under different visual conditions and several listener variables capturing their attitudinal and experiential profiles. Our intent was to provide potential explanations for why listeners may have found L2 utterances drawn from the same stimulus set to be more or less intelligible or comprehensible when those utterances were paired with different visuals. For this analysis, we examined the role of nine background variables in listeners' intelligibility and comprehensibility performance. Four variables measured listeners' demographic and language use characteristics (age, self-rated amount of daily French use in speaking, listening, and communicating with L2 speakers), whereas the remaining five captured listeners' attitudinal reactions to immigrants, their language, religious symbols, and practices, and finally listeners' opinions about mask-wearing. Despite a relatively large total sample size ( $N = 104$ ), as summarized in Table 4 (for intelligibility) and Table 5 (for comprehensibility), only a few associations approached or surpassed the benchmark for a weak relationship ( $r > |0.25|$ ), and none involved listeners' demographic or language use variables or their reactions toward medical mask-wearing.

**Table 4.** Correlations between listeners' background variables and intelligibility scores.

Variable	Mask	Face	Scarf	Niqab
Age (years)	−0.13	−0.01	−0.05	−0.18
Daily speaking in French (%)	0.08	−0.11	−0.04	0.08
Daily listening in French (%)	0.11	−0.12	−0.09	0.05
L2 French contact (%)	−0.11	0.01	0.07	−0.10
Immigrants should adopt Québec values <sup>a</sup>	0.05	−0.18	−0.02	−0.04
It is okay not to speak French well <sup>a</sup>	−0.14	−0.06	−0.25 **	0.05
No to religious symbols or practices in public domain <sup>a</sup>	−0.02	−0.12	−0.04	−0.03
Religious practices are due to family pressure <sup>a</sup>	−0.06	−0.36 ***	−0.23 *	−0.05
Everyone should be free to choose to wear a mask	0.05	−0.04	−0.05	0.01

Note. <sup>a</sup> 100-point agree–disagree scale (100 = “strongly agree”). \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$  (two-tailed).

**Table 5.** Correlations between listeners' background variables and comprehensibility ratings.

Variable	Mask	Face	Scarf	Niqab
Age (years)	−0.16	0.00	−0.05	−0.11
Daily speaking in French (%)	0.11	0.04	0.01	0.05
Daily listening in French (%)	0.09	0.03	0.06	0.09
L2 French contact (%)	0.00	0.05	0.10	0.16
Immigrants should adopt Québec values <sup>a</sup>	−0.02	−0.06	0.01	−0.09
It is okay not to speak French well <sup>a</sup>	0.18	0.06	0.07	0.18
No to religious symbols or practices in public domain <sup>a</sup>	−0.12	−0.03	−0.08	−0.23 *
Religious practices are due to family pressure <sup>a</sup>	−0.05	−0.18	−0.14	−0.07
Everyone should be free to choose to wear a mask	0.01	−0.02	−0.03	0.01

Note. <sup>a</sup> based on a 100-point agree–disagree scale (100 = “strongly agree”). \*  $p < 0.05$  (two-tailed).

In terms of intelligibility, listeners who believed more strongly that immigrants' religious practices reflect various pressures from their family or community tended to show poorer intelligibility for L1 Arabic speakers' utterances presented in the full face condition ( $r = -0.36$ , weak-to-medium effect) and in the headscarf condition ( $r = -0.23$ , weak effect). In addition, listeners who believed more strongly that it is okay for immigrants not to speak French well while living in Québec similarly showed poorer intelligibility scores for L1 Arabic speakers' utterances in the headscarf condition ( $r = -0.25$ , weak effect). With respect to comprehensibility, listeners who more strongly rejected the use of religious symbols and practices in the public domain perceived L1 Arabic speakers' utterances as more difficult to understand, but only in the niqab condition ( $r = -0.23$ , weak effect). Finally, the medical mask condition (which depicted the actual procedure followed during stimulus recording) appeared to be the most neutral as far as listeners' background and attitudinal variables were concerned, in the sense that no association emerged either for intelligibility or comprehensibility in this condition.

## 5. Discussion

### 5.1. Listener Understanding and Visual Condition

The first research question asked whether L1 French listeners differed in their intelligibility and comprehensibility of L2 utterances presented under different visual conditions. Listeners found all utterances to be generally intelligible and comprehensible, which reflected the fairly high proficiency level of the speakers and the relatively low difficulty level of the materials, composed of high-frequency vocabulary and intended for presentation in noise. Nonetheless, differences emerged for the same utterances evaluated under different visual guises. For intelligibility, the utterances were more intelligible to listeners when they were heard in the medical mask than in all other visual contexts, including religious face coverings. For comprehensibility, the utterances were more comprehensible when they were presented in the medical mask condition than with images of religious coverings. Overall, listener understanding was best in the medical mask guise but poorest in the niqab guise.

If visual condition had contributed no effect, all sentences should have been equally intelligible and comprehensible. And yet the utterances presented with an image of a masked speaker elicited greater intelligibility than those paired with a photograph of a person wearing no face covering (full face condition). Considering that all materials had been recorded by masked speakers, it is of little surprise that listeners performed the best when the acoustic information matched the visual cues. Indeed, all audio recordings likely sounded slightly low-pass filtered due to the surgical mask material, with the spectral envelope attenuated by an estimated 0–2 dB in the 1–8 kHz frequency range (Nguyen et al. 2021; Wolfe et al. 2020), which would be consistent with the image of a speaker in a surgical mask. In addition, greater intelligibility in the medical mask condition may have been due

to increased listener effort. [Pycha et al. \(2022\)](#) and [Smiljanic et al. \(2021\)](#) showed that speech from speakers who were actually wearing masks was often significantly better understood than speech from speakers without masks, and [Brown et al. \(2021\)](#) suggested that listeners increase their listening effort when they are aware that a speaker is masked. An image of a masked person may have thus prompted listeners to attend to the speech signal carefully, in the sense that they allocated more cognitive resources to the understanding of speech, which enhanced their accuracy in transcribing sentences ([Pichora-Fuller et al. 2016](#)). Finally, due to Québec's extended obligatory mask mandate (lasting 21 months and lifted about 6 weeks prior to data collection), many listeners would have had recent experience listening to many masked speakers, so an image of a masked person paired with the actual speech produced through a surgical mask may have aided listener comprehension because the input reflected listeners' frequent and recent perceptual experience ([Bent and Baese-Berk 2021](#)).

Assuming that the medical mask condition—as a context with matching audiovisual cues—provides a useful baseline for listener experience, it is striking that listeners found the utterances paired with images of women wearing religious coverings to be less intelligible and less comprehensible than essentially the same utterances presented with a photograph of a woman in a surgical mask. In fact, a niqab provides practically as much obstruction to the face as a surgical mask but carries a different connotation as a symbol of religious affiliation versus a device used in healthcare. If listeners had indiscriminately shown preference for or attention to speech produced through any type of mouth covering, the utterances heard in the niqab condition should have been as intelligible and comprehensible as those in the medical mask guise. Because this was not the case, a more plausible explanation is that the images of religious head garments such as niqabs and hijabs (headscarves) activated listeners' stereotypes about L2 speakers ([Babel and Russell 2015](#); [Kang and Rubin 2009](#); [Kang and Yaw 2021](#); [Kutlu et al. 2022a, 2022b](#)). These stereotypes may have involved various expectations, such as that the speaker belongs to a particular ethnic or religious group, that the speaker is accented or hard to understand, or that the speaker might not speak French as well as other speakers, and these expectations in turn may have influenced listeners' experience, decreasing speaker intelligibility and comprehensibility, as in [Rubin's \(1992\)](#) initial study of this phenomenon. If such stereotypes were indeed operative, they seemed to affect listener experience (at least in terms of intelligibility) broadly, irrespective of whether the image depicted a Middle Eastern woman with a religious garment (i.e., in the headscarf and niqab guises) or without one (i.e., in the full face condition). Even though the actual listeners in this study were not asked to categorize speaker origin, a similar listener sample had identified all women depicted in the full face condition as Middle Eastern with nearly 80% accuracy. Thus, an image of an ostensibly Middle Eastern woman without a religious covering may have activated listener stereotypes in a way that impaired listeners' intelligibility, but seeing a Middle Eastern woman wearing a niqab led to the numerically poorest level of utterance intelligibility and comprehensibility.

It is not entirely unexpected that the poorest level of utterance intelligibility and comprehensibility for listeners occurred under the visual guise which presented listeners with a well-known symbol of religious affiliation—a garment covering most of a woman's head and face. Listeners' stereotypical reaction to this image, along with the consequences of this reaction for their speech processing, may have arisen from listeners' prior societal experiences around religious symbols, especially niqabs. The number of women in Québec wearing niqabs is not precisely known, but informed estimates range from 50 to 100 in total out of around 150,000 Muslim women ([Ouimet 2016](#)). Like almost all Québec residents, listeners would likely have had very little to no face-to-face experience speaking and listening to a woman wearing a niqab. The overwhelming majority of Québec residents will instead encounter French speech by women in niqabs only via the media, almost exclusively when these women are speaking about how Québec's secularism policies and laws affect their lives as niqab wearers. In Québec, as in many Western societies, cultural and popular media have constructed niqabs, along with burkas, as symbols of oppression, inequality, lack of agency, and insecurity ([Inglis 2017](#); [Kearns 2017](#)), without regard for the actual

perspectives of the women wearing these garments (Bakht 2022; Jahangeer 2020). It may be that for a majority of Québec residents, their very limited, indirect experience of women wearing niqabs is further affected by a largely undifferentiated portrayal of Muslim women in the media as victimized and powerless (Ahadi 2009; Benhadjoudja 2017; Piela 2021; Selby 2014), contributing to listeners' perception of women and their speech as unconsciously linked to such negatively-valenced concepts as "obstacle", "impediment", "obscurity", or "challenge". As a result, listeners' performance on intelligibility and comprehensibility tasks may have been reduced. Because listeners' comments about their understanding were not solicited during data collection, this speculative interpretation must be revisited in future work.

## 5.2. Listener Understanding and Listener Profiles

The second research question asked whether there were associations between the intelligibility and comprehensibility of L2 utterances presented under different visual conditions and variables capturing listeners' attitudinal and experiential profiles. Nine variables were analyzed, four measuring listeners' demographic and language use characteristics, four measuring listeners' attitudinal reactions to immigrants and their religious symbols and practices, and one targeting listeners' opinions about mask-wearing. Unlike demographic and language use variables, several attitudinal variables revealed weak associations with intelligibility or comprehensibility scores. For intelligibility, listeners who believed more strongly that the religious practices of immigrants reflected pressures from their family or community were more likely to show poorer intelligibility for speech presented in the full face and the headscarf condition. Moreover, listeners who believed more strongly that it is okay for immigrants living in Québec not to speak French well tended to show poorer intelligibility for speech presented in the headscarf condition. For comprehensibility, listeners who believed more strongly that religious symbols and practices should not be seen in secular institutions rated speech harder to understand, but only in the niqab condition.

The obtained negative associations between listeners' beliefs and their judgments of L2 intelligibility and comprehensibility likely stemmed from several pre-conceived ideas held by at least some listeners about the role of language and religion in society. Listeners may have believed that some individuals from Québec's minority groups are subject to family or community pressures and therefore lack agency in whether and how they wish to express their personal and religious identity (Ahadi 2009; Benhadjoudja 2017; Jahangeer 2020; Selby 2014). Listeners may have also thought that some L2-speaking immigrants rarely learn sufficient French (Bridgman et al. 2022; Gagnon 2023). Similarly, listeners might have considered that religion is quintessentially a private matter and therefore some religious symbols (e.g., hijabs, turbans), unlike other symbols (e.g., crosses) which are considered cultural artifacts, do not belong in the public domain (Patrick et al. 2019). All these relationships are consistent with established prior findings, where negative expectations or stereotypes about a speaker's ethnoracial, linguistic, or religious belonging have been shown to lead listeners to underperform on measures of listening comprehension and intelligibility (e.g., Kang and Rubin 2009; Kang and Yaw 2021; Kutlu et al. 2022b; Rubin 1992). Put differently, if listeners expect the speaker to be powerless in decision-making, disinclined to learn or ineffective in learning the majority group's language, or resistant to the majority group's values, they may actually understand the speaker less or report more struggle with understanding the speaker.

The relationships between listeners' beliefs and their understanding of L2 utterances depended on the visual context. First and foremost, there were no associations for either intelligibility or comprehensibility with any listener variable in the medical mask guise, which again highlights that in this study, the medical mask condition seemed to act as an attitude-neutral baseline for evaluating listener understanding. For intelligibility, the three associations with listener attitudes emerged in the visual contexts in which most women from an ostensibly Middle Eastern background will appear to other Québec residents—either with no religious cover or wearing a headscarf—given that it is uncom-



mon to see women wearing niqabs in Québec's public spaces (Ouimet 2016). Thus, if listeners' attitudes shape their actual understanding of minoritized L2 speakers, these attitudinal influences occur in situations that resemble listeners' typical, everyday experiences. By contrast, for comprehensibility, which is a subjective measure of listening effort, the obtained association with listener attitudes emerged in the niqab guise, which ostensibly portrayed the speaker's religious affiliation in the most stereotypically prominent light. Here, listeners' beliefs may have amplified their perception of listening effort in the presence of visual cues (i.e., a woman wearing a niqab) that were open to highly subjective, clichéd, or idiosyncratic interpretations. While it would be premature to offer any definitive conclusions, it is nevertheless intriguing that listener attitudes patterned with objective measures of understanding (intelligibility) in the visual contexts that were typical of listeners' actual daily experience, whereas listener attitudes patterned with subjective measures of understanding (comprehensibility) in the visual context illustrating a non-prototypical, biased experiential scenario (Inglis 2017; Kearns 2017). Finally, listeners' opinions about having personal freedom to decide to wear a medical mask showed no association with intelligibility or comprehensibility of L2 utterances. Until further evidence is available, this finding implies that listeners' general attitudinal reactions toward mask-wearing were of little consequence to their actual experience with speech.

## 6. Limitations and Future Work

In terms of limitations of this work, the present materials elicited high levels of intelligibility and comprehensibility from listeners, which may have moderated potential differences across visual conditions or suppressed associations between the two measures of understanding and listeners' attitudinal and experiential variables. Even though online surveys with speech samples played on listeners' personal devices have become common (e.g., Callesano and Carter 2019; Slade et al. 2024), the decision to prioritize participants' virtual access to recording and listening sessions over researchers' control of listening conditions meant that acoustic and perceptual measurements at the physical source of recording and listening were not taken. It would be important, therefore, to determine if the current findings would hold for the same materials presented in noise or recorded by speakers whose L2 proficiency is less advanced, using procedures that allow for stricter control over recording and listening conditions. Although the target materials, drawn from a standardized test, included utterances that were matched in length and complexity, listener understanding might differ for content that is semantically anomalous (e.g., more vs. less familiar to listeners) or situationally charged (e.g., listening to a request for help vs. a socially loaded political commentary), given that utterance content determines the degree of listener understanding (Kennedy and Trofimovich 2008; Tekin and Trofimovich 2023). With respect to materials, it would also be important to compare listeners' responses to L2 utterances recorded by speakers from other ethnolinguistic backgrounds such as Latinx, Caucasian, or Southeast Asian, with or without surgical masks and religious face coverings. Finally, because most listeners in this study were recruited in Montréal, which is home to residents representing over 100 ethnic and cultural origins (Statistics Canada 2021), it is perhaps unsurprising that listeners' individual profiles of language use (e.g., self-rated amount of daily French use in speaking and in communicating with L2 speakers) showed no associations with measures of listener understanding. In future work, therefore, it might be useful to recruit listeners who differ more widely in the type and diversity of their linguistic networks (Kutlu et al. 2022a, 2022b).

We found it difficult to directly compare the present findings to those reported previously, because no prior research has examined listener understanding of L2 speech comparing visual conditions involving speakers wearing religious and nonreligious coverings. As in prior work, we found that speakers who are recorded wearing masks are understood well (Badh and Knowles 2023). The only study where a religious symbol (hijab) was one of the visual conditions did not show any differences in comprehensibility by condition (Lybaert et al. 2022), whereas here, the comprehensibility of speakers presented

in the niqab guise was significantly lower than for speakers presented in two nonreligious visual guises. Thus, we call for follow-up research to extend and clarify these findings, with the goal of providing a comprehensive view of how religious and nonreligious coverings shape listeners' understanding of L2 speech.

In terms of listeners' experience with and attitudes toward L2 speech, we used only a handful of single questionnaire items that focused on listeners' experience with L2 French speech. A longer, more extensive attitudinal test battery might have allowed us to show stronger relationships or to uncover those that we failed to capture with relevance to intelligibility and comprehensibility. The target attitudinal items were also quite specific to Québec's societal issues, including concerns about maintaining a secular society and integration of ethnic and linguistic minorities into a majority francophone province inside a largely anglophone Canada. These concerns have been publicly discussed for decades, and listeners' responses to attitudinal items may have been subject to social desirability bias (Anderson 2019), where some listeners might have responded to the questionnaire items in ways they believed were reflective of general societal attitudes rather than their personal opinions. In future, debriefing interviews with listeners would be useful to explore how they arrive at their responses. Finally, it would be worthwhile to replicate this dataset in contexts other than Québec to determine the degree of cross-contextual and cross-situational invariance in listener understanding of L2 speech with and without visual cues to a speaker's religious affiliation.

## 7. Conclusions

The current study is the first we are aware of to explore how different religious and nonreligious visual conditions might affect listeners' understanding of L2 French utterances, as well as how intelligibility and comprehensibility might be associated with listeners' attitudes and experiences under different visual conditions. Effects of visual conditions were not straightforward, largely because comprehension of L2 utterances was very high. However, the consistent and significant gap in intelligibility and comprehensibility between one or more visual conditions and the niqab condition suggests that listeners' completion of the listening tasks was negatively affected by visual cues available in the niqab condition; this condition may be associated, for these listeners at least, with the notion of difficult or problematic speech. Associations between listeners' experiential and attitudinal profiles and their performance on intelligibility and comprehensibility tasks were few and generally weak. These questionnaire items were the first, to our knowledge, to elicit listeners' attitudes about secularism in Québec's society in the context of listeners' reactions to L2 French speech under religious and nonreligious visual conditions. Social desirability bias may have underpinned some of the listeners' responses. The generally few associations between listeners' attitudes on secularism in Québec and their reactions to L2 French speech may also reflect instability in attitudes toward the legal framework for secularism (e.g., Bill 21). Surveys on Québécois' attitudes toward Bill 21 have shown a decrease over time in general support for Bill 21, with women, especially younger women, reporting less support than men (Rukavina 2022). This underlines the importance of longitudinal data targeting secularism attitudes, in relation to L2 speech processing and learning, in future research to capture societal shifts over time.

**Author Contributions:** Conceptualization, S.K. and P.T.; methodology, all authors; software, R.L. and O.T.; validation, P.T., R.L. and O.T.; formal analysis, P.T. and R.L.; investigation, R.L. and O.T.; resources, S.K. and P.T.; data curation, R.L. and O.T.; writing—original draft, S.K., P.T. and R.L.; writing—review & editing, S.K. and P.T.; visualization, P.T.; supervision, S.K. and P.T.; project administration, S.K. and P.T.; funding acquisition, S.K. and P.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Social Sciences and Humanities Research Council of Canada to S.K. and P.T., grant numbers 430-2020-1134 and 430-2022-00710.

**Institutional Review Board Statement:** The study was approved by Concordia University’s Human Research Ethics Committee (Certificate 30009422, approved 30 March 2022; Certificate 30013555, approved 26 August 2021).

**Informed Consent Statement:** Informed consent was obtained from all participants involved in the study.

**Data Availability Statement:** All data and materials for this study are publicly available via the OSF (<https://doi.org/10.17605/osf.io/cd2xp>).

**Acknowledgments:** We thank all study participants. We also gratefully acknowledge the assistance we received from Anamaria Bodea with establishing coding reliability and from Kym Taylor Reid with the design of visual stimuli.

**Conflicts of Interest:** The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

#### Appendix A. Target French Sentences (1–20 = Low Predictability, 21–40 = High Predictability)

1. Ce cheval appartient au roi.	This horse belongs to the king.
2. Ma cousine a trouvé un gros os.	My cousin found a big bone.
3. Les deux amis ont fait la paix.	The two friends made peace.
4. Il grave son nom sur du bronze.	He engraves his name on bronze.
5. Le gros chien a léché le bol.	The big dog licked the bowl.
6. Il mange ses pâtes avec du sel.	He eats his pasta with salt.
7. Il faudra mettre une deuxième couche.	You will need to apply a second coat.
8. Ils sont tous partis au camp.	They all went to camp.
9. Elle mange beaucoup trop de viande.	She eats too much meat.
10. Son ami lui a prêté ses notes.	His friend lent him his notes.
11. J’ai acheté une nouvelle table.	I bought a new table.
12. Cette couverture est faite en laine.	This blanket is made of wool.
13. Le fermier va nourrir ses vaches.	The farmer goes to feed his cows.
14. Elle va apporter de l’eau.	She’s going to bring water.
15. Mes enfants jouent avec une toile.	My children play with a canvas.
16. Claudie a découvert une mine.	Claudie discovered a mine.
17. Ce livre est rempli de belles phrases.	This book is full of beautiful sentences.
18. Ce chandail n’a pas de prix.	This sweater is priceless.
19. Ma poubelle est pleine de sable.	My trash can is full of sand.
20. Jules n’aime pas jouer dans le bain.	Jules doesn’t like playing in the bath.
21. Le fermier travaille à la ferme.	The farmer works on the farm.
22. Elle lui fait signe de la main.	She waves at him.
23. Je vais chez ma tante et mon oncle.	I’m going to my aunt and uncle.
24. Tous les trains roulent sur des rails.	All trains run on rails.
25. Le chanteur a une très belle voix.	The singer has a very beautiful voice.
26. Elle se rongait les ongles.	She bit her nails.
27. On voit mieux la lune pendant la nuit.	We see the moon better at night.
28. Elles font des bonhommes de neige.	They make snowmen.
29. J’aime mâcher de la gomme.	I like to chew gum.
30. Le frappeur a frappé la balle.	The batter hit the ball.
31. Le bateau passe sous le pont.	The boat passes under the bridge.
32. Les éléphants ont une trompe.	Elephants have a trunk.
33. Anne a vu des jongleurs au cirque.	Anne saw jugglers at the circus.
34. Le chasseur va à la chasse.	The hunter goes hunting.
35. Le jardinier arrose ses plantes.	The gardener waters his plants.
36. Les girafes ont un grand cou.	Giraffes have long necks.
37. Je ne connais même pas son nom.	I don’t even know his name.
38. Jacques met tout dans son sac à dos.	Jacques puts everything in his backpack.
39. L’avion vole haut dans le ciel.	The plane is flying high in the sky.
40. Ce joueur d’hockey fait de belles passes.	This hockey player makes great passes.

### Appendix B. Brief Attitudinal Questionnaire

Si des gens d'autres pays s'installent au Québec, ils devraient être prêts à adopter les valeurs québécoises. (*If people from other countries move to Quebec, they should be ready to adopt Quebec values.*)

fortement en désaccord ————— tout à fait d'accord

Je pense que c'est acceptable pour quelqu'un de passer le reste de sa vie au Québec même s'il ne parle pas bien le français. (*I think it's okay for someone to spend the rest of their life in Quebec even if they don't speak French well.*)

fortement en désaccord ————— tout à fait d'accord

Les Québécois ont le droit de s'attendre à ce que lorsqu'ils se trouvent dans des institutions laïques comme des bâtiments gouvernementaux ou des écoles, ils ne rencontrent personne qui manifeste une pratique ou un symbole religieux. (*Quebeckers have a right to expect that when they are in secular institutions like government buildings or schools, they will not encounter anyone showing any kind of religious practice or symbol.*)

fortement en désaccord ————— tout à fait d'accord

Les gens s'engagent dans des pratiques religieuses comme aller à la confession ou porter un hijab à cause de la pression de leur famille ou de leur communauté. (*People engage in religious practices like going to confession or wearing a headscarf because of pressure from their family or community.*)

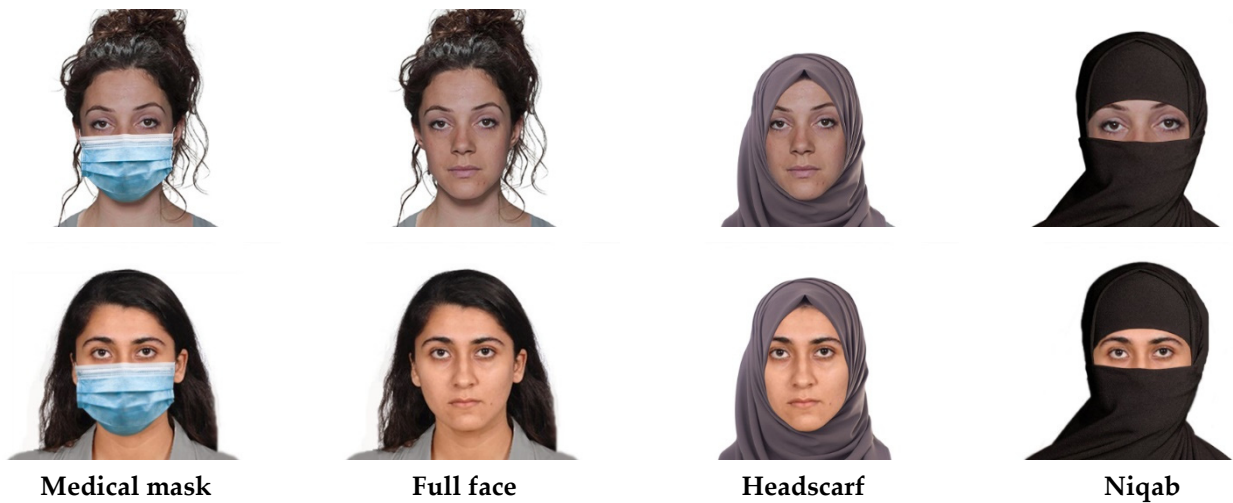
fortement en désaccord ————— tout à fait d'accord

Je pense que chacun devrait être libre de faire son propre choix de porter des masques ou de ne pas en porter. (*I believe everyone should be free to make their own choice to wear masks or not to wear them.*)

fortement en désaccord ————— tout à fait d'accord

### Appendix C. Full Image Set Used to Create Speaker Identity





Appendix D. Item Randomizations Across Survey Versions

Speaker	Survey Version 1				Survey Version 2	
	Sentences	Images		Sentences	Images	
Speaker 1	1–5 & 36–40	Full face	Image 1	11–15 & 26–30	Headscarf	Image 2
Speaker 2	16–20 & 26–30	Niqab	Image 2	6–10 & 36–40	Full face	Image 1
Speaker 3	11–15 & 21–25	Mask	Image 3	1–5 & 31–35	Niqab	Image 4
Speaker 4	6–10 & 31–35	Headscarf	Image 4	16–20 & 21–25	Mask	Image 3

Speaker	Survey Version 3				Survey Version 4	
	Sentences	Images		Sentences	Images	
Speaker 1	16–20 & 21–25	Mask	Image 3	6–10 & 31–35	Niqab	Image 4
Speaker 2	11–15 & 31–35	Headscarf	Image 4	1–5 & 21–25	Mask	Image 3
Speaker 3	6–10 & 36–40	Full face	Image 2	16–20 & 26–30	Headscarf	Image 1
Speaker 4	1–5 & 26–30	Niqab	Image 1	11–15 & 36–40	Full face	Image 2

Appendix E. Screenshot of the Rating Interface

J'ai entendu cette locutrice dire:



## References

- Ahadi, Daniel. 2009. L’Affaire Hérouxville in context: Conflicting narratives on Islam, Muslim women, and identity. *Journal of Arab and Muslim Media Research* 2: 241–60. [CrossRef] [PubMed]
- Anderson, Joel R. 2019. The moderating role of socially desirable responding in implicit–explicit attitudes toward asylum seekers. *International Journal of Psychology* 54: 1–7. [CrossRef]
- Babel, Molly. 2022. Adaptation to social-linguistic associations in audio-visual speech. *Brain Sciences* 12: 845. [CrossRef]
- Babel, Molly, and Jamie Russell. 2015. Expectations and speech intelligibility. *The Journal of the Acoustical Society of America* 137: 2823–33. [CrossRef]
- Badh, Gursharan, and Thea Knowles. 2023. Acoustic and perceptual impact of face masks on speech: A scoping review. *PLoS ONE* 18: e0285009. [CrossRef] [PubMed]
- Bakht, Natasha. 2022. Getting to know the other: Niqab-wearing women in liberal democracies. *Religions* 13: 361. [CrossRef]
- Benhadjoudja, Leila. 2017. Laïcité narrative et sécularonationalisme au Québec à l’épreuve de la race, du genre et de la sexualité. *Studies in Religion/Sciences Religieuses* 46: 272–91. [CrossRef]
- Bent, Tessa, and Melissa Baese-Berk. 2021. Perceptual learning of accented speech. In *The Handbook of Speech Perception*, 2nd ed. Edited by Jennifer S. Pardo, Lynne C. Nygaard, Robert E. Remez and David B. Pisoni. New York: Wiley, pp. 428–64. [CrossRef]
- Bridgman, Aengus, Rosalie Nadeau, and Dietlind Stolle. 2022. A distinct society? Understanding social distrust in Quebec. *Canadian Journal of Political Science* 55: 107–27. [CrossRef]
- Brown, Violet A., Kristin J. Van Engen, and Jonathan E. Peelle. 2021. Face mask type affects audiovisual speech intelligibility and subjective listening effort in young and older adults. *Cognitive Research: Principles and Implications* 6: 49. [CrossRef]
- Callesano, Salvatore, and Phillip M. Carter. 2019. Latinx perceptions of Spanish in Miami: Dialect variation, personality attributes and language use. *Language & Communication* 67: 84–98. [CrossRef]
- Case, Julie, Scott Seyfarth, and Susannah V. Levi. 2018. Short-term implicit voice-learning leads to a Familiar Talker Advantage: The role of encoding specificity. *The Journal of the Acoustical Society of America* 144: EL497–EL502. [CrossRef] [PubMed]
- DeBruine, Lisa, and Benedict Jones. 2017. Face Research Lab London Set. Available online: [https://figshare.com/articles/dataset/Face\\_Research\\_Lab\\_London\\_Set/5047666/3](https://figshare.com/articles/dataset/Face_Research_Lab_London_Set/5047666/3) (accessed on 16 February 2021).
- Derwing, Tracey M., and Murray J. Munro. 2015. *Pronunciation Fundamentals: Evidence-Based Perspectives for L2 Teaching and Research*. Amsterdam: John Benjamins.
- Gagnon, Audrey. 2023. Not just civic or ethnic, but mostly cultural: Conceptions of national identity and opinions about immigration in Quebec. *Nations and Nationalism* 29: 1076–92. [CrossRef]
- Hanulíková, Adriana. 2018. The effect of perceived ethnicity on spoken text comprehension under clear and adverse listening conditions. *Linguistics Vanguard* 4: 20170029. [CrossRef]
- Inglis, David. 2017. Cover their face: Masks, masking, and masquerades in historical-anthropological context. In *The Routledge International Handbook to Veils and Veiling*. Edited by Anna-Mari Almila and David Inglis. London: Routledge, pp. 278–92.
- Institut national de santé publique du Québec. 2020. *Faits saillants du 5 novembre 2020*. Available online: <https://www.inspq.qc.ca/covid-19/sondages-attitudes-comportements-quebecois/29-octobre-2020> (accessed on 16 October 2023).
- Institut national de santé publique du Québec. 2022. *Faits saillants du 16 décembre 2022*. Available online: <https://www.inspq.qc.ca/covid-19/sondages-attitudes-comportements-quebecois/16-dec-2022> (accessed on 16 October 2023).
- Jahangeer, Roshan Arah. 2020. Anti-veiling and the charter of Québec values: “Native testimonials,” erasure, and violence against Montreal’s Muslim women. *Canadian Journal of Women and the Law* 32: 114–39. [CrossRef]
- Jeong, Hyeseung, Anna Elgemark, and Bosse Thorén. 2021. Swedish youths as listeners of global Englishes speakers with diverse accents: Listener intelligibility, listener comprehensibility, accentedness perception, and accentedness acceptance. *Frontiers in Education* 6: 651908. [CrossRef]
- Kang, Okim, and Donald L. Rubin. 2009. Reverse linguistic stereotyping: Measuring the effect of listener expectations on speech evaluation. *Journal of Language and Social Psychology* 28: 441–56. [CrossRef]
- Kang, Okim, and Katherine Yaw. 2021. Social judgement of L2 accented speech stereotyping and its influential factors. *Journal of Multilingual and Multicultural Development*, advance online publication. [CrossRef]
- Kearns, Matthew. 2017. Gender, visibility and violence: Visual securitization and the 2001 war in Afghanistan. *International Feminist Journal of Politics* 19: 491–505. [CrossRef]
- Kennedy, Sara, and Pavel Trofimovich. 2008. Intelligibility, comprehensibility, and accentedness of L2 speech: The role of listener experience and semantic context. *Canadian Modern Language Review* 64: 459–90. [CrossRef]
- Kutlu, Ethan, Mehrgol Tiv, Stefanie Wulff, and Debra Titone. 2022a. Does race impact speech perception? An account of accented speech in two different multilingual locales. *Cognitive Research: Principles and Implications* 7: 1–16. [CrossRef]
- Kutlu, Ethan, Mehrgol Tiv, Stefanie Wulff, and Debra Titone. 2022b. The impact of race on speech perception and accentedness judgements in racially diverse and non-diverse groups. *Applied Linguistics* 43: 867–90. [CrossRef]
- Lagacé, Josée, Benoît Jutras, Christian Giguère, and Jean-Pierre Gagné. 2010. Development of the Test de Phrases dans le Bruit (TPB). *Canadian Journal of Speech-Language Pathology and Audiology* 34: 261–70. Available online: [https://cjslpa.ca/files/2010\\_CJSLPA\\_Vol\\_34/No\\_04\\_226-303/Lagace\\_Jutras\\_Giguere\\_Gagne\\_CJSLPA\\_2010.pdf](https://cjslpa.ca/files/2010_CJSLPA_Vol_34/No_04_226-303/Lagace_Jutras_Giguere_Gagne_CJSLPA_2010.pdf) (accessed on 14 June 2021).

- Leger Marketing. 2021. *Federal Leaders Debate: French Language and Bills 21 and 96 in Quebec*. Available online: <https://2g2ckk18vixp3neolz4b6605-wpengine.netdna-ssl.com/wp-content/uploads/2021/09/Leaders-debate-report-11679-251-VF.pdf> (accessed on 16 October 2023).
- Lindemann, Stephanie. 2002. Listening with an attitude: A model of native-speaker comprehension of non-native speakers in the United States. *Language in Society* 31: 419–41. [CrossRef]
- Llomas, Carmen, Philip Harrison, Damien Donnelly, and Dominic Watt. 2008. Effects of different types of face coverings on speech acoustics and intelligibility. *York Papers in Linguistics Series* 9: 80–104.
- Loi sur la laïcité de l'État. 2019. Available online: <https://www.legisquebec.gouv.qc.ca/fr/document/lc/L-0.3> (accessed on 1 February 2024).
- Lybaert, Chloé, Sarah Van Hoof, and Bart Deygers. 2022. The influence of ethnicity and language variation on undergraduates' evaluations of Dutch-speaking instructors in Belgium: A contextualized speaker evaluation experiment. *Language & Communication* 84: 1–19. [CrossRef]
- Melguy, Yevgeniy V., and Keith Johnson. 2021. General adaptation to accented English: Speech intelligibility unaffected by perceived source of non-native accent. *The Journal of the Acoustical Society of America* 149: 2602–14. [CrossRef]
- Nguyen, Duy Duong, Patricia McCabe, Donna Thomas, Alison Purcell, Maree Doble, Daniel Novakovic, Antonia Chacon, and Catherine Madill. 2021. Acoustic voice characteristics with and without wearing a facemask. *Scientific Reports* 11: 5651. [CrossRef]
- Ouimet, Louis-Philippe. 2016. *Comprendre ces Québécoises qui portent le niqab*. Radio-Canada. Available online: <https://ici.radio-canada.ca/nouvelle/1003891/niqab-femme-religion-temoignage-quebec-islam> (accessed on 25 January 2024).
- Patrick, Margaretta, W. Y. Alice Chan, Hicham Tifati, and Erin M. Reid. 2019. *Religion and Secularism: Four Myths and Bill 21*. Toronto: Canadian Race Relations Foundation.
- Pichora-Fuller, M. Kathleen, Sophia E. Kramer, Mark A. Eckert, Brent Edwards, Benjamin W. Y. Hornsby, Larry E. Humes, Ulrike Lemke, Thomas Lunner, Mohan Matthen, Carol L. Mackersie, and et al. 2016. Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear and Hearing* 37: 5S–27S. [CrossRef]
- Piela, Anna. 2021. *Wearing the Niqab: Muslim Women in the UK and the US*. London: Bloomsbury Publishing.
- Plonsky, Luke, and Frederick L. Oswald. 2014. How big is “big”? Interpreting effect sizes in L2 research. *Language Learning* 64: 878–912. [CrossRef]
- Pycha, Anne, Michelle Cohn, and Georgia Zellou. 2022. Face-masked speech intelligibility: The influence of speaking style, visual information, and background noise. *Frontiers in Communication* 7: 874215. [CrossRef]
- Rubin, Donald L. 1992. Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education* 33: 511–31. [CrossRef]
- Rukavina, Steve. 2022. New Research Shows Bill 21 Having ‘Devastating’ Impact on Religious Minorities in Quebec. *CBC News*. Available online: <https://www.cbc.ca/news/canada/montreal/bill-21-impact-religious-minorities-survey-1.6541241> (accessed on 25 January 2024).
- Saribay, S. Adil, Ali F. Biten, Erdem O. Meral, Pinar Aldan, Vít Třebický, and Karel Kleisner. 2018. The Bogazici face database: Standardized photographs of Turkish faces with supporting materials. *PLoS ONE* 13: e0192018. [CrossRef]
- Selby, Jennifer A. 2014. Un/veiling women's bodies: Secularism and sexuality in full-face veil prohibitions in France and Québec. *Studies in Religion/Sciences Religieuses* 43: 439–66. [CrossRef]
- Sheppard, Beth E., Nancy C. Elliott, and Melissa M. Baese-Berk. 2017. Comprehensibility and intelligibility of international student speech: Comparing perceptions of university EAP instructors and content faculty. *Journal of English for Academic Purposes* 26: 42–51. [CrossRef]
- Simon, Ellen, Chloé Lybaert, and Koen Plevoets. 2022. Social attitudes, intelligibility and comprehensibility: The role of the listener in the perception of non-native speech. *Vigo International Journal of Applied Linguistics* 19: 177–222. [CrossRef]
- Slade, Kate, Alanna Beat, Jennifer Taylor, Christopher J. Plack, and Helen E. Nuttall. 2024. The effect of motor resource suppression on speech perception in noise in younger and older listeners: An online study. *Psychonomic Bulletin & Review* 31: 389–400. [CrossRef]
- Smiljanic, Rajka, Sandie Keerstock, Kirsten Meemann, and Sarah M. Ransom. 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: 4013–23. [CrossRef]
- Statistics Canada. 2021. *Census Profile, 2021 Census of Population: Profile Table, Montréal, QC*. Available online: <https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/prof/details/page.cfm?Lang=E&SearchText=montreal&DGUIDlist=2021S0503462&GENDERlist=1,2,3&STATISTIClist=1&HEADERlist=0> (accessed on 15 November 2023).
- Stewart, Michelle, Maxime Bérubé, Samuel Laperle, Sklaerenn Le Gallo, and Stéphanie Panneton. 2023. “Victims of the system”: Anti-government discourse and political influencers online. In *Virtual Identities and Digital Culture*. Edited by Victoria Kannen and Aaron Langille. New York: Routledge, pp. 57–68.
- Taylor Reid, Kym. 2022. *Exploring Social Attitudes Toward Second Language Speakers of English Across Canada*. Doctoral thesis, Concordia University, Montreal, QC, Canada. Available online: <https://spectrum.library.concordia.ca/id/eprint/991004> (accessed on 13 July 2022).
- Tekin, Oguzhan, and Pavel Trofimovich. 2023. En français or in English? Examining perceived social roles of international students in response to their French and English speech. *Canadian Modern Language Review* 79: 204–27. [CrossRef]

- Toscano, Joseph C., and Cheyenne M. Toscano. 2021. Effects of face masks on speech recognition in multi-talker babble noise. *PLoS ONE* 16: e0246842. [\[CrossRef\]](#)
- Witkowska, Marta, Silvia Filippi, Magdalena Formanowicz, and Caterina Saitner. 2023. Sociophonetics and language prejudice. Accent matters—A socio-psychological perspective on sociophonetics. In *The Routledge Handbook of Sociophonetics*. Edited by Christopher Strelluf. London: Routledge, pp. 342–64. [\[CrossRef\]](#)
- Wolfe, Jace, Joanna Smith, Sara Neumann, Sharon Miller, Erin C. Schafer, Amy L. Birath, Tina Childress, Catharine McNally, Caleb McNiece, Jane Madell, and et al. 2020. Optimizing communication in schools and other settings during COVID-19. *The Hearing Journal* 73: 40–45. [\[CrossRef\]](#)

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.