Multilingual Speech Processing in Simulated Online Learning Conditions

Galia Freiwirth

LING 449-001: Honours Essay

University of British Columbia

Dr. Molly Babel

April 29, 2021

Abstract

The COVID-19 pandemic required that students and instructors make a shift in their routines, from attending lectures in person to attending online. A vast number of studies show that when listening to speech in quiet conditions, bilinguals are just as successful at perceiving the speech signal as monolinguals, but when listening to speech in noise, the perceptual abilities of bilinguals decrease considerably. Additionally, research on vocabulary size indicates that a larger vocabulary size can be beneficial when listening to speech in adverse conditions. Nonetheless, only few studies have analyzed the abilities of listeners with knowledge of more than two languages on speech perception tasks, and to our knowledge, the relationship between vocabulary size and speech-perception-in-noise (SPIN) tasks has yet to be studied. This study aimed to address the gap by assessing the performance of multilinguals in adverse listening conditions, quantifying multilingualism with a language entropy measure while also considering English vocabulary size. Students at the University of British Columbia were recruited to participate in a SPIN task, auditory vocabulary assessment, and multilingualism questionnaire. Speech processing challenges were present for all participants in the adverse conditions, but even more so for participants with lower vocabulary assessment scores. There was no effect of entropy on SPIN performance. The results replicate previous SPIN task findings, extend the findings to multilinguals, and suggest that there may be an added layer of difficulty for individuals with smaller vocabulary sizes in online environments.

Keywords: speech-perception-in-noise; vocabulary; multilingualism; entropy; online class; virtual learning

Introduction

The COVID-19 pandemic required that students and instructors across the globe make drastic shifts in their academic routines: from learning and teaching in person, to attending classes and meetings on various platforms that support video-calling online. Inevitably, the transition to online learning came with a myriad of challenges, including a loss in quality and clarity of the speech signal from the compromised audio. In addition to the presence of technical challenges, a potential, additional obstacle for online speech perception surfaced for those who speak more than one language. A number of previous studies have shown that when listening to speech in quiet conditions, bilinguals are just as successful at perceiving the speech signal as monolinguals, but when listening to speech in noise, the perceptual abilities of bilinguals decrease considerably (Lucks Mendel, & Widner, 2016; Mayo et al., 1997; Rosenhouse et al., 2006; Tabri et al., 2015; Von Hapsburg & Peña, 2002; Schmidtke, 2016; Weiss & Dempsey, 2008). Although multilinguals vary a great deal in terms of their language background, usage, self-perceived accents, abilities, and proficiency, as well as their vocabulary size, the question of how multilinguals perceive speech in online learning environments is still worthy of being researched; this is particularly so due to the large multilingual population in Vancouver, British Columbia, where the study is being conducted.

Lucks Mendel and Widner (2016) measured the abilities of adult monolingual English listeners and bilingual Spanish listeners (proficient in English) on three speech-perception-in-noise (SPIN) tasks: the quick speech-in-noise (QuickSIN), the Bamford-Kowal-Bench speech-in-noise (BKB-SIN), and the words-in-noise (WIN). For all SPIN tasks, the sentences in noisy conditions were played in the presence of multi-talker babble. They found that scores for the QuickSIN and WIN tests were relatively similar compared to that of the BKB-SIN test. The

BKB-SIN test consists of simpler sentence structures and added semantic contexts, resulting in the participants attaining the highest scores on said test. Across all three tests, the results of bilinguals were lower than monolinguals, and comparable to listeners with hearing impairments (even though the bilinguals' hearing was unimpaired), exhibiting the challenges bilinguals face to perceive speech in noisy conditions.

Similarly, Rosenhouse et al. (2006) conducted a study to examine the effects of amalgamated adverse listening conditions on bilinguals' (Arabic L1 Hebrew L2) speech perception. In their experiment, they combined two different speech rates, regular and fast, with two different background conditions, quiet and noisy, yielding four different testing conditions. In each of the four conditions, participants listened to eight sentences in Arabic and eight sentences in Hebrew. For trials in conditions with noise, sentences were played in the presence of multi-talker babble of eight female speakers' simultaneous speech. The authors found that the bilinguals experienced greater challenges with processing the Hebrew sentences in the task, were affected more by background noise than speech rate, and were more affected by the combination of speech rate and background noise in Hebrew than Arabic. They explain further that because of Arabic diglossia, it is possible that participants struggled more with Hebrew as it represented their L3, not their L2, and therefore made processing speech in adverse conditions particularly difficult.

In a study by Tabri et al. (2015) that looked at trilinguals' performance on SPIN tasks, the authors found that the performance of trilingual listeners (speakers of Arabic, English, and French) was comparable with the performance of bilingual listeners (speakers of Arabic and English). They analyzed the English speech perception of participants using a SPIN task with various levels of noise intensity in the background. Half of the sentences were of high predictability (offering contextual cues to listeners) and half of the sentences were of low predictability (requiring listeners to use acoustic-phonetic information). When presented in noisy conditions, the sentences were played in the presence of 12-talker babble noise. Trilingual listeners performed slightly worse than bilingual listeners, exhibiting the steepest decline in speech-in-noise processing abilities, but not by a significant margin. All the participants had acquired their languages at an early age (before six), and even with high proficiencies, their perceptual abilities were still affected by the background noise.

Previous research has largely focused on the perceptual abilities of bilinguals in SPIN tasks but has not given much consideration (besides Tabri and colleagues' (2015) study) to how multilinguals – specifically those with knowledge of more than two languages – would perform on such a task. Additionally, research on vocabulary sizes has revealed that greater vocabulary knowledge and size enhances adults' speech recognition in adverse listening conditions (Banks et al., 2015; Janse & Adank, 2012), and that bilinguals have smaller vocabularies in each of their languages (Bialystok et al., 2010). However, there yet to be research conducted on how vocabulary size may be correlated with individuals' scores on SPIN experiments. Finally, language entropy has been used to characterize the language use of bilinguals in different contexts – namely work, home, and social environments (Gullifer & Titone, 2020) – but a connection between language entropy scores and SPIN task performance remains unstudied. The goals of the current study are to assess speech in noise performance in adverse listening conditions for multilingual individuals, quantifying multilingualism with a language entropy measure while also considering English vocabulary size. The current work is inspired by the existing virtual learning environment and the multilingual student body of the University of British Columbia (UBC). Seeing as bilinguals and trilinguals in previous SPIN studies have consistently performed poorly, relative to monolinguals' performance, it is hypothesized that multilinguals will share similar struggles in this experiment. The current study includes four levels of difficulty in the SPIN task, and it is hypothesized that as the level of difficulty increases through increased speech distortion, the speech recognition abilities of multilinguals will decrease as well. It is hypothesized that participants with higher scores in the vocabulary assessment will attain higher scores on the SPIN task. Finally, it is hypothesized that participants with lower entropy scores will attain higher scores on the SPIN task.

Methodology

Participants

The participants included in this study were students at the University of British Columbia, who were recruited online through the Department of Linguistics: UBC Linguistics Sign-Up Sona System. In total, there were 128 adult-aged participants, of which, 98 self-identified as female and ranged in age from 18 to 42 years (M = 22 years), 27 self-identified as male and ranged in age from 18 to 27 years (M = 21 years), and three self-identified as 'other' and ranged in age from 20 to 24 years (M = 23 years). With regards to language backgrounds, seven participants had experience with one language (monolingual), 30 participants had experience with two languages (bilingual), and the remaining 91 participants were multilinguals: 47 had experience with three languages, 29 had experience with four languages, 11 had experience with five languages, one had experience with six languages, two had experience with seven languages, and one had experience with nine languages. English was the most reported dominant language, followed by Mandarin, Cantonese, Japanese, Korean, Russian, Tagalog, Chinese (not otherwise specified), Hindi, Hungarian, Spanish, and Thai. Secondary languages reported by participants included French, Punjabi, Vietnamese, Albanian, Chinuk Wawa, German, Hebrew, Indonesian, Norwegian, Oriya, Polish, Portuguese, Serbian, Tamil and Telugu. The study was completed by participants mostly located in the Lower Mainland area of British Columbia, Canada, but also in other Canadian provinces, as well as in China, Hong Kong, Japan, the Philippines, Singapore, Spain, Taiwan, the United States, and the United Arab Emirates.

Materials and Stimuli

This study was administered online to the participants in two parts. The first part consisted of a SPIN task. The second part consisted of an auditory vocabulary assessment and multilingualism questionnaire.

Speech-Perception-in-Noise (SPIN) Task

For the SPIN task, we used a list of 144 unique English declarative sentences (Baese-Berk, personal communication). The list contains 48 sentences of high predictability (e.g., "Red and green are colours."), 48 sentences of low predictability (e.g., "Mom talked about the pie."), and 48 sentences that are semantically anomalous (e.g., "The black top ran the spring."). As mentioned by Tabri et al. (2015), having high predictability sentences allows for listeners to use the contextual cues to transcribe their responses, while the other sentence types require listeners to solely rely on the acoustic-phonetic information to transcribe what they hear. The sentences were read aloud by a female speaker with a Pacific Northwest English accent, recorded using a FIFINE (2019) T669 USB microphone in a carpeted room with a blanket hung up behind the microphone. The recording was digitized at 44.1 kHz and 32-bit float on Audacity (2021) and saved as an individual .wav file. Using Praat (Boersma & Weenink, 2020), the onset and offset boundaries of each sentence in the recording were demarcated, labeled with the sentence content in a TextGrid file, and extracted to their own sound files. Finally, all sentences were RMSamplitude normalized to 70 dB and 500 milliseconds of silence were added at the onset and the offset of each sound file.

The sound files were then manipulated to create different adverse listening conditions. To create the bandpass-limited items, files were passed through a Hahn pass band with lower and upper thresholds of 300 and 3400 Hz, respectively, with a 50 Hz bandwidth. To create the speech-in-noise items, the filtered and full spectrum .wav files were mixed with multi-talker babble at 0 dB SNR. All together, this resulted in four stimuli types: full spectrum + clear, full spectrum + multi-talker babble, bandwidth-limited + clear, bandwidth-limited + multi-talker babble.

Auditory Vocabulary Assessment

The vocabulary assessment used in this study was an adaptation of Lemhöfer & Broersma's (2012) Lexical Test for Advanced Learners of English (LexTALE). It was modified by Dr. Molly Babel to use in the Speech in Context Lab at UBC as an auditory vocabulary assessment (Babel, 2020). The task was first developed for Babel and Mellesmoen's (2019) project, where the scores of their participants produced a normal distribution of percent correct on the task. The modified Auditory LexTALE is composed of 60 trials in which participants heard a single speech utterance and had to decide whether the utterance was an existing English word (e.g., "remuda," "celestial," "muddy") or not an existing English word (e.g., "purrage," "pulsh," "rebondicate").

A was used survey to administer the LexTALE task. The back button was restricted to prevent participants from hearing the stimuli more than once per trial. Utterances in this task were presented in the same order for all participants.

Multilingualism Questionnaire

Participants completed the Language Experience and Proficiency Questionnaire (LEAP-Q) (Marian et al., 2007). The LEAP-Q was created by the Northwestern Bilingualism and Psycholinguistics Research Laboratory to assess language profiles in bilinguals and multilinguals. The questions relate to participants' language history, use, attitudes, and self-rated proficiency. This part of the survey consisted of 13 questions and was administered on the same survey, immediately succeeding the auditory LexTALE task.

The auditory vocabulary assessment and multilingualism questionnaire were included together in Part Two of the study, administered using a survey on Qualtrics online software (Qualtrics, 2020).

Procedure

Part One: SPIN Task

The first part of this study was conducted online and took approximately 30 minutes to complete. Before the task began, participants were given written instructions and a consent form to read. Once they consented to participate, they were instructed to complete the study in a quiet space and began the headphone check (Woods et al., 2017). Participants listened to three tones per round of the headphone check and had to determine which tone was the quietest. They were required to select the correct tone in four out of six rounds in order to proceed to the SPIN task.

In the SPIN task, listeners were presented with 36 trials of each noise type, comprised of 12 sentences from each sentence type. These items were presented in a randomized order for each participant, and within each participant, no sentence was repeated across noise or sentence types. Participants were instructed to type exactly what they heard. Two breaks were inserted for participants to take after the first third of the SPIN task, and again after the second third of the task. At the end of the task participants were thanked for their time. Participants received partial course credit for completing the task.

Part Two: Auditory Vocabulary Assessment and Multilingualism Questionnaire

The second part of this study was also conducted online and took approximately 30 minutes to complete. Participants were presented with the consent form again prior to beginning Part Two. To commence the modified LexTALE vocabulary assessment, participants were given instructions to categorize each item as “Yes, a word” or “No, not a word” when they heard each utterance. In cases where they may have been unsure of the exact meaning of the utterance but believed it existed, they were instructed to click on "Yes, a word." In cases where they were unsure about the existence of the word in English, they were instructed to click on "No, not a word." Participants had an unlimited amount of time per trial. After the vocabulary assessment, participants were instructed to fill out the LEAP-Q thoroughly and sincerely. At the end of the task, participants were once again thanked for their participation in this study and granted partial course credit for completing the task.

The LexTALE vocabulary assessment was purposely administered after the SPIN task, to avoid participants potentially being influenced by their performance on said assessment.

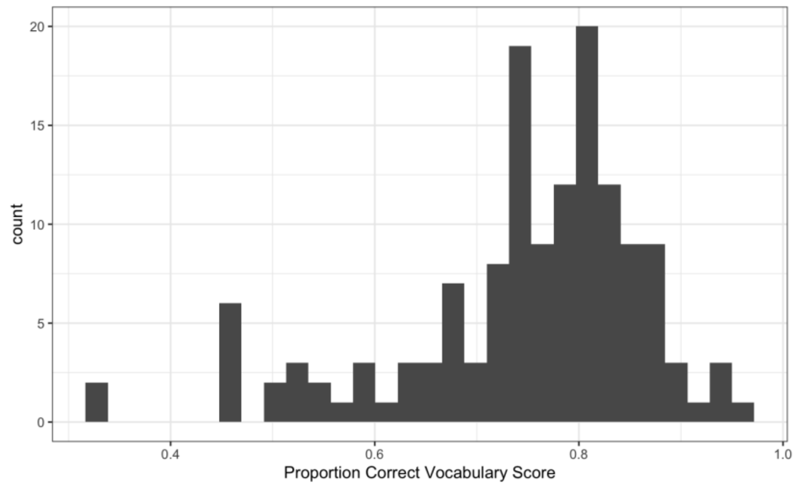
Results

Vocabulary and Language Entropy Scores

Null responses on words were marked as incorrect, and individual scores were computed as the mean proportion correct across the 60 trials of the task. Figure 1 presents a histogram of the distribution of these scores.

Figure 1

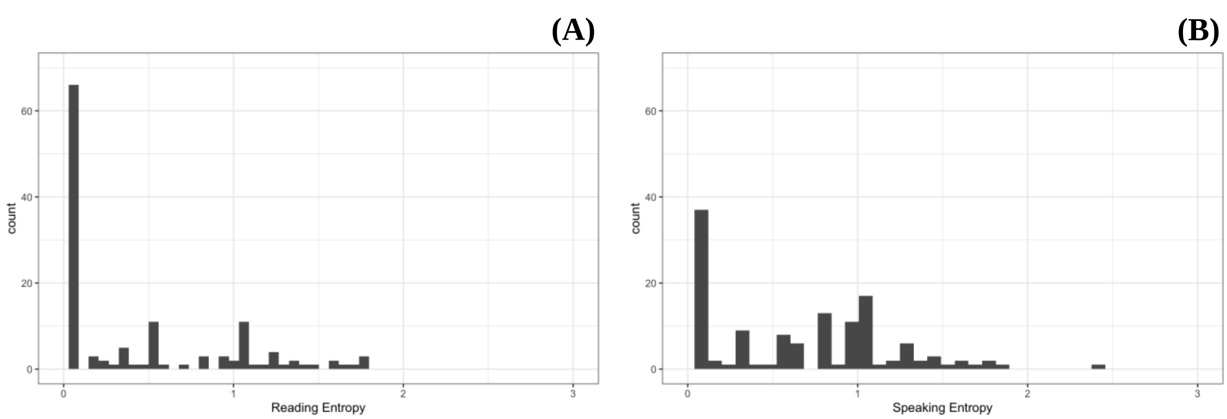
Proportion correct vocabulary scores on LexTALE vocabulary assessment.

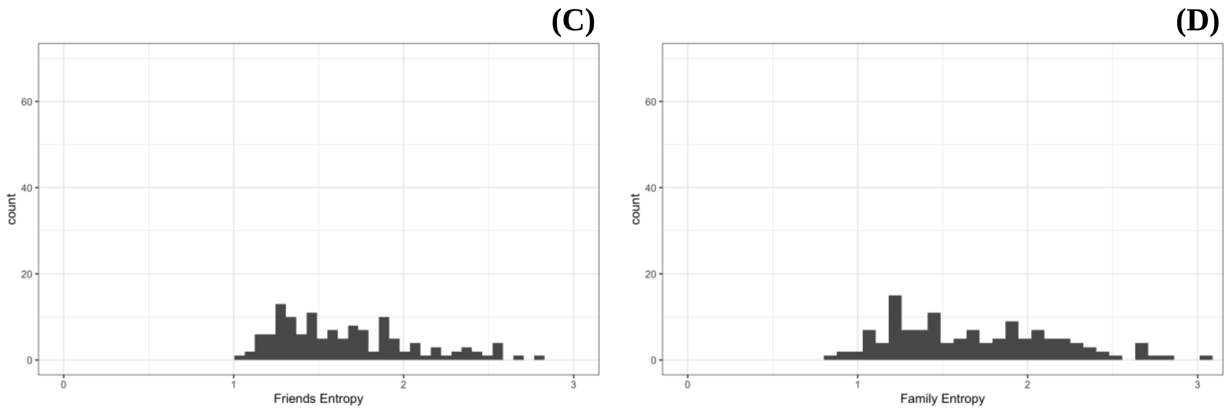


Language entropy scores were calculated following the methods described in Gullifer and Titone (2020) for reading, speaking, friends, family, watching, self-instruction, and formal instruction. Distributions of the entropy scores for reading, speaking, friends, and family are shown in Figure 2.

Figure 2

Distribution of entropy scores for (A) reading, (B) speaking, (C) friends, and (D) family.





Because of the large number of NAs for the two instruction question modules, only entropy values for reading, speaking, friends, family, and watching were included in a principal components analysis (PCA) intended to reduce the dimensionality of these language use contexts. Subsequently removing subjects with NAs in these columns reduced the number of participants from 134 to 129. A PCA of these dimensions revealed the loadings reported in Table 1; the importance of these components is summarized in Table 2. The first principal component was related to reading, speaking, friends, and watching, and accounted for approximately 43% of the variance. The loading of the second principal component was dominated by family entropy with some positive loading from friends, as well. This component accounted for 22% of the variance in the model. A scree plot of the PCA proportion variance is in Figure 3.

Table 1

Loadings of the PCA for entropy dimensions: reading, speaking, friends, family, and watching.

Entropy Dimension	PC1	PC2	PC3	PC4	PC5
Reading	0.5368957	-0.25816548	0.4348802	0.08871998	-0.66940386
Speaking	0.5738076	0.02578004	0.398203	0.03897548	0.71413967
Friends	0.3978873	0.57514637	-0.2625681	-0.64554271	-0.15882368

Family	-0.1893857	0.7700944	0.4074633	0.43479177	-0.12655992
Watching	0.4339431	0.09405982	-0.6460225	0.62035506	-0.02570252

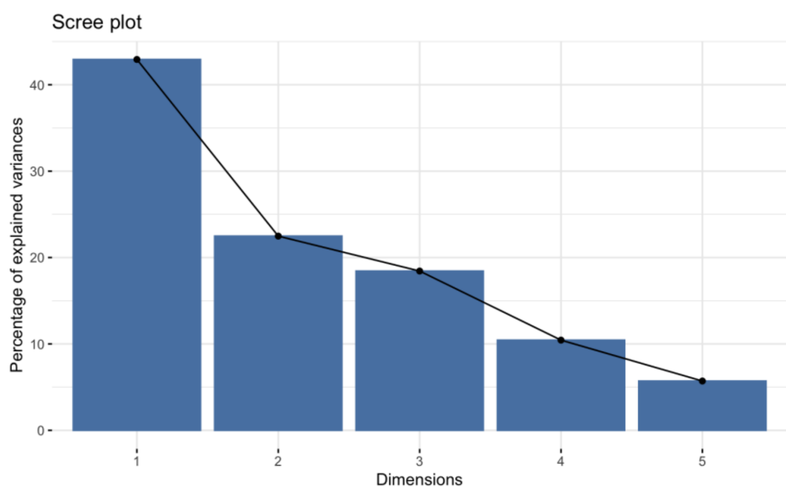
Table 2

Importance of components from the PCA.

	PC1	PC2	PC3	PC4	PC5
Standard deviation	1.4651	1.0603	0.9602	0.7226	0.53424
Proportion of Variance	0.4293	0.2248	0.1844	0.1044	0.05708
Cumulative Proportion	0.4293	0.6541	0.8385	0.9429	1.00000

Figure 3

Scree plot of the PCA proportion variance.



The vocab and entropy scores were combined with responses on the SPIN task, with 128 participants completing all tasks. Amongst these participants, vocabulary scores were negatively correlated with entropy PC1 [$t(126) = -4.03$, $r = -0.34$, $p < 0.001$], suggesting that those with lower vocabulary scores had higher levels of language entropy.

Sentence Transcription Performance

Performance on the sentence transcription task was scored using a fuzzy logic token sort ratio (TSR) in python (Bosker, 2020, 2021), which scores sentence transcription accuracy from 0 to 100. These scores were used as the dependent measure in a series of linear mixed effects regression models. The possible independent variables in the analysis were sentence type, noise masking, bandwidth, vocabulary, and the first principal component for entropy. Sentence type was Helmert coded to compare High (⅓) and Low (-⅓) Predictability sentences to Semantically Anomalous sentences (⅔) and then the High (½) and Low (-½) Predictability sentences with each other (Anomalous = 0). Noise masking was sum coded (clear = 1, noise = -1), as was bandwidth (full spectrum = 1, bandpass limited = -1). Vocabulary scores were centered and scaled. A series of mixed effects linear regression model with random intercepts for participant and sentence was built, starting with sentence type, and progressively adding variables as main effects and interactions, as warranted by model comparison. The optimal model output is in Table 3. This is the model with vocabulary scores, but without entropy. The entropy scores did not improve the model.

Table 3

Linear mixed effects regression model outputs.

Effect	Estimate	Std. Error	t-value
(Intercept)	89.60656	0.58122	154.17
set_type.f1	-7.24595	0.89183	-8.125
set_type.f2	-3.18825	1.02985	-3.096
filter1	6.293	0.24818	25.357
bandwidth1	2.96008	0.09381	31.554
lextale.score.centered	3.59554	0.38778	9.272

set_type.f1:filter1	4.02006	0.19855	20.247
set_type.f2:filter1	1.47721	0.23034	6.413
set_type.f1:bandwidth1	0.37123	0.19869	1.868
set_type.f2:bandwidth1	0.08428	0.23015	0.366
filter1:bandwidth1	-1.81432	0.09305	-19.498
set_type.f1:lextale.score.centered	1.49671	0.19664	7.611
set_type.f2:lextale.score.centered	2.17427	0.22723	9.569
filter1:lextale.score.centered	-1.28471	0.2377	-5.405
bandwidth1:lextale.score.centered	-0.53762	0.09307	-5.777
set_type.f1:filter1:bandwidth1	0.43832	0.19731	2.222
set_type.f2:filter1:bandwidth1	-0.05484	0.22804	-0.24
set_type.f1:filter1:lextale.score.center	0.18334	0.19763	0.928
set_type.f2:filter1:lextale.score.center	-0.87752	0.22803	-3.848
set_type.f1:bandwidth1:lextale.score.cen	0.2406	0.19732	1.219
set_type.f2:bandwidth1:lextale.score.cen	-0.23045	0.22811	-1.01
filter1:bandwidth1:lextale.score.centere	0.02069	0.0931	0.222
set_type.f1:filter1:bandwidth1:lextale.s	-0.28185	0.19743	-1.428
set_type.f2:filter1:bandwidth1:lextale.s	0.24283	0.22812	1.065

A t-value great than |2| indicates that the effect is significant. Based on this, the effects indicate that high predictability (mean TSR score = 90.41107) and low predictability sentences (mean TSR score = 93.45569) were different from semantically anomalous sentences (mean TSR score = 84.71673), as well as different from each other. There was an effect of filter (which is what masking noise is called in the model), indicating that clear conditions (mean TSR score = 95.62749) and noise conditions (mean TSR score = 83.05012) were different from each other. There was an effect of bandwidth, indicating that the full bandwidth (mean TSR score = 92.53729) differed from the limited bandwidth (mean TSR score = 86.69400). There was an effect of LexTALE score, indicating that a difference exists between high scores on the

LexTALE assessment, and low scores on the LexTALE assessment. Additionally, there was a Pearson's product-moment correlation of 0.6702021 (95 percent confidence interval: 0.5620368 to 0.7558293) between the TSR and LexTALE scores ($t = 10.136$, $df = 126$, p-value $< 2.2e-16$).

The interaction of sentence type and filter is shown in Figure 4. Although semantically anomalous sentences reduced intelligibility in clear and noisy trials, the reduction in intelligibility for semantically anomalous sentences in noisy trials was greater. The interaction of filter and bandwidth is shown in Figure 5. While limiting the bandwidth reduced the intelligibility in clear and noisy trials, the reduction in intelligibility for the noise masking on the bandwidth-limited trials was greater. The effect of vocabulary score and sentence type on intelligibility is shown in Figure 6. While a difference in intelligibility scores is present for participants in the entire range of vocabulary scores, the difference in intelligibility scores across different sentence types is even greater for participants with lower vocabulary scores. The effect of vocabulary score and filter on intelligibility is shown in Figure 7. While a difference in intelligibility scores is present for participants in the entire range of vocabulary scores, the difference in intelligibility scores across different filter conditions is even greater for participants with lower vocabulary scores. The effect of vocabulary score and bandwidth on intelligibility is shown in Figure 8. While a difference in intelligibility scores is present for participants in the entire range of vocabulary scores, the difference in intelligibility scores across different bandwidths (full or limited) is even greater for participants with lower vocabulary scores. The interaction of filter, bandwidth, and sentence type is shown in Figure 9. While limiting the bandwidth reduced the intelligibility in clear and noisy trials for both semantically anomalous and semantically coherent sentences, the reduction in intelligibility for the noise masking on the bandwidth-limited trials was greater, and even greater for the semantically anomalous trials. The effect of vocabulary score, filter, and sentence type on intelligibility is shown in Figure 10. While a difference in intelligibility scores is present for participants in the entire range of vocabulary scores, the difference in intelligibility scores across different filter conditions, for both high predictability and low predictability sentences, is even greater for participants with lower vocabulary scores.

Figure 4

Box plot displaying scores of sentence types (high predictability, low predictability, and anomalous) in clear condition and noise (filter) condition.

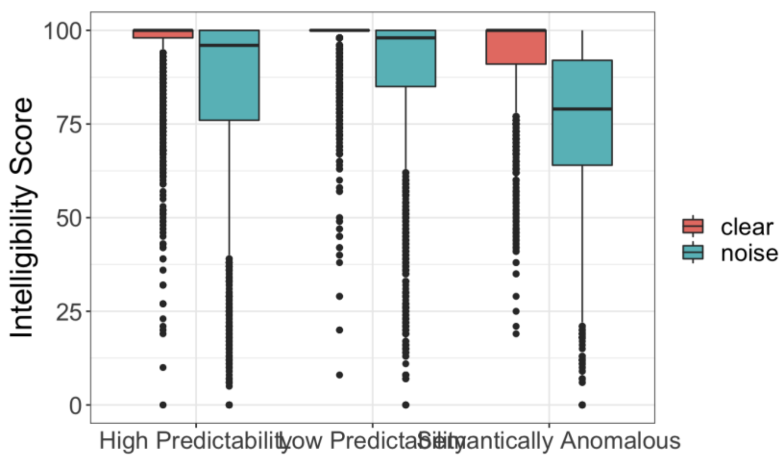


Figure 5

Box plot displaying scores of sentences in clear condition and noise (filter) condition with full bandwidth and limited bandwidth.

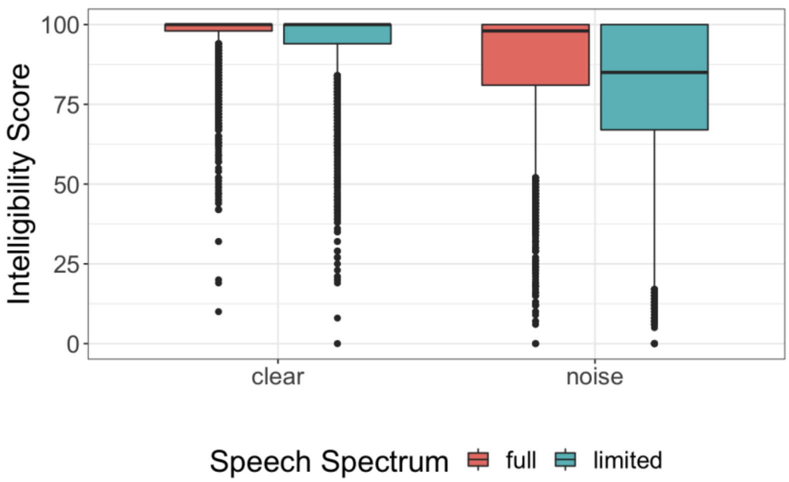


Figure 6

The effect of LexTALE vocabulary assessment scores and sentence types (high predictability, low predictability, and anomalous) on intelligibility.

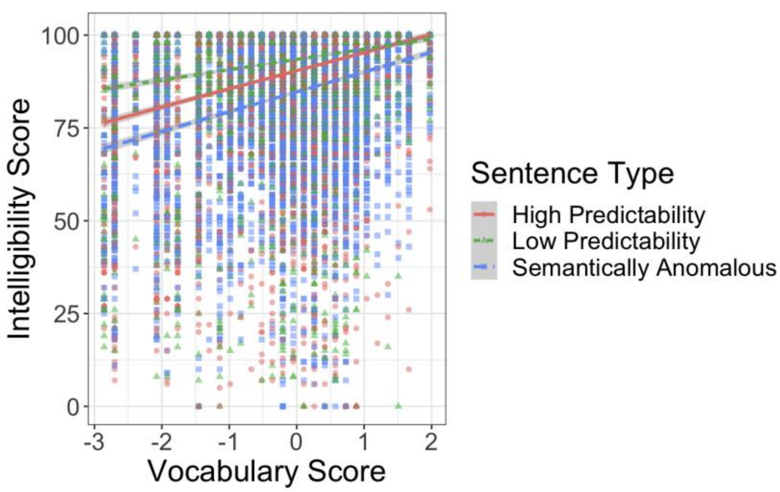


Figure 7

The effect of LexTALE vocabulary assessment scores and filter (noise condition and clear condition) on intelligibility.

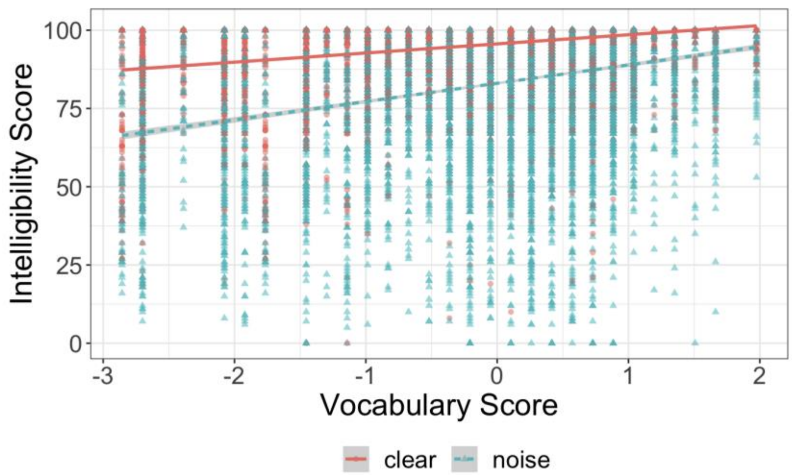


Figure 8

The effect of LexTALE vocabulary assessment scores and bandwidth (full and limited) on intelligibility.

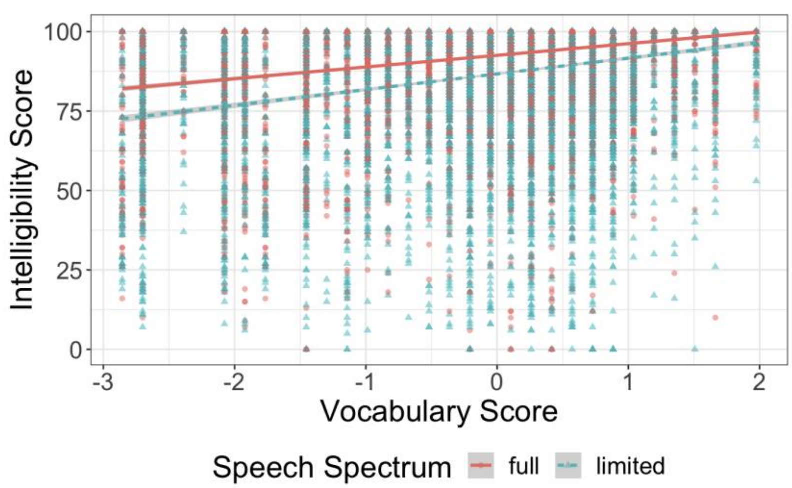


Figure 9

Box plot displaying scores of sentences in clear condition and noise (filter) condition with full bandwidth and limited bandwidth, separated by sentence type (high and low predictability, known as coherent, from semantically anomalous).

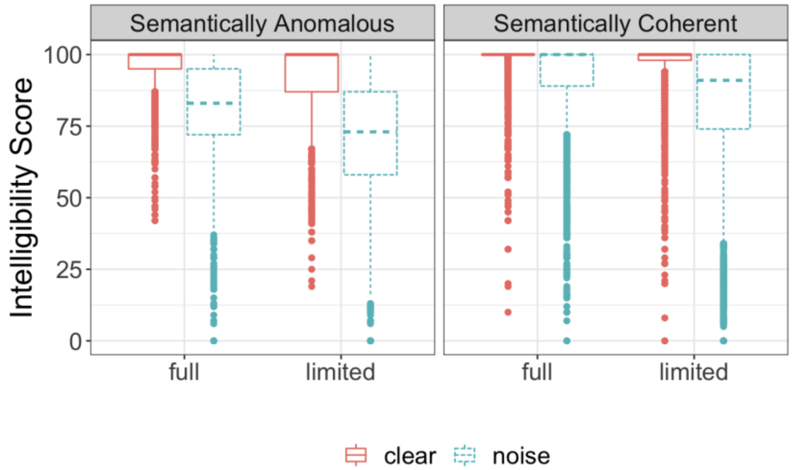
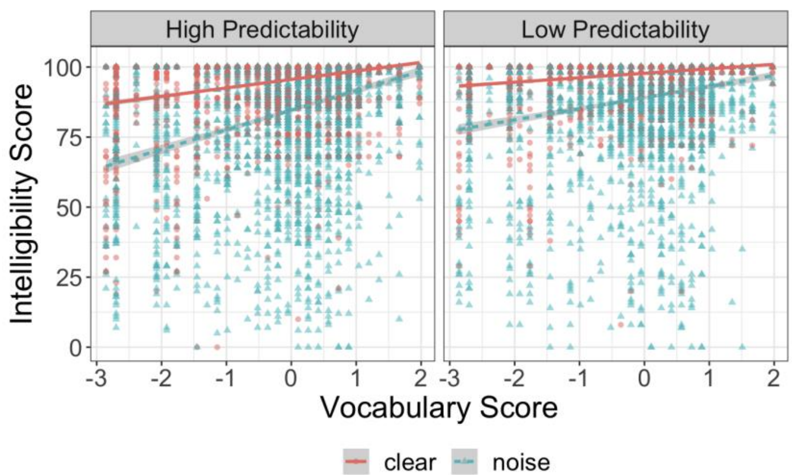


Figure 10

The effect of LexTALE vocabulary assessment scores, filter conditions (clear and noise), and sentence type (high and low predictability) on intelligibility.



Discussion and Conclusions

Previous studies conducted on speech perception show that in comparison to monolingual listeners, bilinguals are just as good at listening to speech in quiet conditions but tend to experience challenges with processing speech in adverse listening conditions, such as in SPIN tasks (Lucks Mendel, & Widner, 2016; Mayo et al., 1997; Rosenhouse et al., 2006; Tabri et al., 2015; Von Hapsburg & Peña, 2002; Schmidtke, 2016; Weiss & Dempsey, 2008). As expected, in the current study participants struggled to listen to speech that was filtered (noise condition), bandwidth-limited, and containing sentences that were semantically anomalous. In addition, participants who scored lower on the auditory vocabulary assessment were seen to experience greater challenges with listening to speech in adverse conditions, more so than the participants who achieved higher scores on the auditory vocabulary assessment.

In line with Gullifer and Titone's (2020) study, for the purposes of measuring language entropy as an estimate of participants' social diversity of language use, information was extracted from the LEAP-Q (Marian et al., 2007) portion of the multilingualism questionnaire, regarding the participants' age of acquisition, language exposure and use in different environments, self-reported accent perceptions, and abilities in their additional languages. Based on the findings from Gullifer and Titone (2020), and considering the fact that the current study included monolingual English-speaking participants, it was expected that participants who attained lower entropy scores (also known as language used in a "compartmentalized manner") would obtain higher scores on the SPIN task. Instead, in the current study, language entropy did not influence SPIN task scores and in turn, cannot be used to predict SPIN task performance.

Evidence from bilingual research shows that bilinguals have smaller vocabulary sizes in each of their languages but approximately the same sized vocabulary as a monolingual, when the words from their different languages are combined overall (Bialystok et al., 2010). As mentioned above, vocabulary size did have a significant effect on SPIN task performance, supporting findings from previous studies on speech processing in adverse listening conditions (Banks et al., 2015; Janse & Adank, 2012). In the challenging trials of the SPIN task, such as in the ones with semantically anomalous sentences or in filtered speech conditions (with noise), participants with lower vocabulary scores experienced greater difficulty with transcribing what they heard than participants who scored higher on the vocabulary assessment. This finding is particularly relevant to the context of online learning during the COVID-19 pandemic, as it could suggest that multilingual students experience a potential, greater struggle to hear their virtual meetings in comparison to their monolingual peers.

The current study was inspired by the virtual learning environment and the multilingual student body at UBC. A demographic survey at UBC from 2012 revealed that nearly 50% of domestic students spoke a language other than English as their first language, and over 60% of international students spoke a language other than English as their first language (Coutts, 2012). As indicated by the 2016 Canadian census (Statistics Canada, 2017), the amount of language diversity and multilingualism in the British Columbia region is continuously rising, suggesting the percentage of people from the UBC community who speak a language other than English may be greater than what was reported in 2012. It is necessary to take these reports into consideration so that UBC creates an inclusive environment for their students and staff.

This study is important as it sheds light on how multilinguals perform on experimental tasks, as well as how they would potentially process speech in real-world, online environments. In addition to academic environments, the current study invites the reader to take into consideration the severity of the adverse listening conditions that multilinguals face in other environments. Online university lectures are just one example of an environment where degraded speech may negatively impact the listener; other fundamental environments to consider include telemedicine, teletherapy, and job interviews. Additionally, the study serves as a reminder to all to create good listening environments in online meetings, using coherent speech, in quiet spaces, and with good quality microphones.

References

Audacity Team. (2021). *Audacity: Free audio editor and recorder*. (Version 3.0.0) [Computer application]. https://audacityteam.org/

Babel, M. (2020, December 8). *Modified Auditory LexTale for Native English Listeners*. Retrieved from https://osf.io/a7ftx/

Babel, M., & Mellesmoen, G. (2019). Perceptual adaptation to stereotyped accents in audio-visual speech. In *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia* (pp. 1044-1048).

Banks, B., Gowen, E., Munro, K. J., & Adank, P. (2015). Cognitive predictors of perceptual adaptation to accented speech. *The Journal of the Acoustical Society of America, 137*(4), 2015-2024.

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism (Cambridge, England), 13*(4), 525.

Boersma, P., & Weenink, D. (2020). *Praat: Doing phonetics by computer* (Version 6.1.14) [Computer program]. http://www.praat.org/

Bosker, H. R. (2020, October 28). *Using fuzzy string matching for automated assessment of listener transcripts in speech intelligibility studies*. Retrieved from osf.io/73dnj

Bosker, H. R. (2021). Using fuzzy string matching for automated assessment of listener transcripts in speech intelligibility studies. *Behavior Research Methods.* doi: 10.3758/s13428-021-01542-4.

Coutts, L. (2012, August 15). *Who are our Students? Implications for Teaching and Learning |*

UBC Centre for Teaching, Learning and Technology. UBC Centre for Teaching, Learning and Technology. https://ctlt.ubc.ca/2012/08/15/who-are-our-students-implications-for-teaching-and-learning/

FIFINE. (2019). *USB Microphone* (Model T669) [Apparatus]. https://fifinemicrophone.com/collections/microphones/products/fifine-usb-microphone-t669

Gullifer, J. W., & Titone, D. (2020). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition, 94*(1), 1-18.

Janse, E., & Adank, P. (2012). Predicting foreign-accent adaptation in older adults. *Quarterly Journal of Experimental Psychology, 65*(8), 1563-1585.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods, 44,* 325-343.

Lucks Mendel, L., & Widner, H. (2016). Speech perception in noise for bilingual listeners with normal hearing. *International Journal of Audiology, 55*(2), 126-134.

Marian, Blumenfeld, & Kaushanskaya (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech Language and Hearing Research, 50*(4), 940-967.

Mayo, L. H., Florentine, M., & Buus, S. (1997). Age of second-language acquisition and perception of speech in noise. *Journal of Speech, Language, and Hearing Research, 40*(3), 686-693.

Qualtrics. (2020). *Qualtrics* [Computer survey software]. Provo, Utah, USA. https://www.qualtrics.com/

Rosenhouse, J., Haik, L., & Kishon-Rabin, L. (2006). Speech perception in adverse listening conditions in Arabic-Hebrew bilinguals. *International Journal of Bilingualism, 10*(2), 119-135.

Schmidtke, J. (2016). The bilingual disadvantage in speech understanding in noise is likely a frequency effect related to reduced language exposure. *Frontiers in Psychology, 7*, 678.

Statistics Canada. (2017). *Focus on Geography Series, 2016 Census*. Statistics Canada Catalogue no. 98-404-X2016001. Ottawa, Ontario. Data products, 2016 Census.

Tabri, D., Chacra, K. M. S. A., & Pring, T. (2015). Speech perception in noise by monolingual, bilingual and trilingual listeners. *International Journal of Language & Communication Disorders, 1*-12.

Von Hapsburg, D., & Peña, E. D. (2002). Understanding bilingualism and its impact on speech audiometry. *Journal of Speech, Language, and Hearing Research*.

Weiss, D., & Dempsey, J. J. (2008). Performance of bilingual speakers on the English and Spanish versions of the Hearing in Noise Test (HINT). *Journal of the American Academy of Audiology, 19*(1), 5-17.

Woods, K. J., Siegel, M. H., Traer, J., & McDermott, J. H. (2017). Headphone screening to facilitate web-based auditory experiments. *Attention, Perception, & Psychophysics, 79*(7), 2064-2072.