2022-12-01

Sources of Top-Down Processing in Recognition of Repeated Speech in Noise

Liam Gleason
University of Texas at El Paso

Recommended Citation

Gleason, Liam, "Sources of Top-Down Processing in Recognition of Repeated Speech in Noise" (2022). *Open Access Theses & Dissertations*. 3680.
https://scholarworks.utep.edu/open_etd/3680

SOURCES OF TOP-DOWN PROCESSING IN RECOGNITION OF REPEATED SPEECH IN NOISE

LIAM J. GLEASON

Masters' Program in Psychology

APPROVED:

Wendy Francis, Ph.D., Chair

Jamie Desjardins, Ph.D.

Christine Potter, Ph.D.

Ana Schwartz, Ph.D.

Stephen L. Crites, Jr., Ph.D.
Dean of the Graduate School

SOURCES OF TOP-DOWN PROCESSING IN RECOGNITION OF REPEATED SPEECH IN NOISE

by

LIAM J. GLEASON

THESIS

Presented to the Faculty of the Graduate School of
The University of Texas at El Paso
in Partial Fulfillment
of the Requirements
for the Degree of

MASTER OF ARTS

Department of Psychology
THE UNIVERSITY OF TEXAS AT EL PASO
December 2022

Abstract

Prior research has shown that repetition facilitates speech recognition in difficult listening environments, as indicated by more accurate reporting and lower subjective noise ratings. Both effects were found even when sentences were read at encoding, suggesting that the top-down processing involved comes from a modality-general level of representation (Gleason & Francis, 2021). We investigated whether this top-down processing comes from the semantic or lemma level of language representation. Bilingual participants listened to sentences in English and Spanish. At test, these sentences and new sentences were presented auditorily in English with background noise. After listening to each sentence, participants reported the final word and rated the noise level. In the English encoding condition, results replicated previous priming effects in repetition accuracy and noise ratings. However, these repetition effects did not transfer across language conditions, which suggests that the top-down processing originates at the lemma level and not the semantic level.

Table of Contents

List of Tables

List of Figures

Sources of Top-Down Processing in Recognition of Repeated Speech in Noise

Perceiving and recognizing speech in difficult listening environments, such as a busy restaurant or a noisy party, is a very cognitively demanding task (Rudner et al., 2012; Wong et al., 2008). Previous research has shown that repeating information makes speech recognition easier (Jacoby et al., 1988), but the reasons are not fully understood. Some have suggested that top-down processing based on previous exposure to the information may provide an explanation (Pichora-Fuller, 2008). However, it is yet to be established which levels of language representation are involved in this process. The current study addresses this question by testing whether the speech recognition advantage for repeated information transfers across languages in bilingual listeners. We also examine possible associations of language proficiency and age of acquisition with repetition-related effects within and across languages.

Effects of Repetition on Speech Recognition

Repetition priming refers to a form of implicit memory that improves accuracy and/or the speed of responses when stimuli have been presented more than once (Gabrieli, 1998). Implicit memory for content presented in speech has been observed by presenting spoken sentences mixed with background noise. When sentences have been previously presented listeners are better able to reproduce those sentences (Jacoby et al., 1988; Sheldon et al., 2008). This effect is thought to occur because previous experience allows participants to better utilize top-down processing to decipher speech. Participants misattribute their ease in understanding to lower background noise and will rate background noise lower for sentences that have been repeated (Jacoby et al., 1988). To better understand why repetition makes speech recognition in noise easier we can turn to previous research regarding bilingual speech recognition.

Bilingual Speech Recognition

While listening to speech in noisy environments is a difficult task, it is especially difficult for bilingual listeners. Bilinguals recognize speech in noise less accurately than their monolingual counterparts, especially when they use their second language, L2 (Bradlow & Alexander, 2007; Bradlow & Bent, 2002; Lecumberri & Cooke, 2006). One explanation for this difficulty is that bilinguals are not able to use linguistic cues to the same extent that monolinguals are able to use those same cues (Bradlow & Alexander, 2007; Crandell & Smaldino, 1996; Cutler et al., 2004; Mayo et al., 1997).

One of the foundational studies on speech recognition compared the performance of simultaneous, early, and late bilinguals on the Speech Perception In Noise (SPIN) task to the performance of monolinguals (Mayo et al., 1997). Researchers found that early bilinguals were better able to maintain accurate recognition of speech with more background noise when compared to late bilinguals. This comparison led to the explanation that even when non-native listeners developed a high level of fluency, their ability to identify L2 speech in noise is greatly influenced by their age of acquisition (Mayo et al., 1997). One possibility is this difference in speech recognition may be due to bilingual listeners being less tolerant of noise levels. However, a study comparing bilinguals and monolinguals showed no group differences in acceptable noise levels, even when the acceptable noise levels were obtained in bilinguals' L2 (von Hapsburg & Bahng, 2006). Speech perception measures did not correlate with acceptable noise levels in either language for bilingual listeners. These results indicate that acceptable noise levels are independent of language in bilinguals, and therefore is not a plausible explanation for speech perception differences (von Hapsburg & Bahng, 2006).

Shi (2009) compared the performance of monolinguals and simultaneous, early, and late bilinguals on word recognition tasks in varying levels of background noise. All three groups of bilinguals performed less accurately at all noise levels when compared to monolinguals. Differences in performance between late bilingual and monolingual listeners were found even in the weakest condition of background noise, music. This result implies that late bilingual listeners are more susceptible to errors when any background noise is present (Shi, 2009). Age of acquisition and length of learning were both shown to be reliable predictors of bilingual performance. In a follow up study, Shi (2010) also examined the effects of context and acoustic degradation on speech perception of listeners with different ages of English acquisition. When comparing monolingual, early bilingual, and late bilingual performance on a SPIN task, it was found that late bilinguals received less benefit from contextual cues in sentences, while early bilinguals and monolinguals were able to utilize context to assist speech recognition in noise. Thus, the extent to which noise influenced the use of contextual cues was dependent on the listener’s age of acquisition (Shi, 2010).

Processing and Levels of Representation in Speech Recognition

The beneficial effects of repetition and context on speech recognition performance may arise because of top-down processing. To understand where difficulty arises in a bilingual’s language system it is important to consider speech recognition models that incorporate both bottom-up and top-down processing. In the context of speech recognition, bottom-up processing can be described as using incoming acoustic stimuli to work upwards until a representation of the word is accessed. Top-down processing is using previous knowledge and contextual information to assist in accessing the word representation.

To understand how bottom-up and top-down processing may work in speech recognition we consider the Bilingual Language Interaction Network for Comprehension of Speech (BLINCS; Shook & Marian, 2013). The BLINCS model proposes that auditory information moves bi-directionally between levels of language representation, see Figure 1. This model incorporates both bottom-up and top-down processing by allowing information from each level of language to move freely from one level to another. When a speaker relies on bottom-up processing in this model they would take the auditory input, and work their way up through each language level, from the phonological representation all the way up to the semantic representation. For example, if the auditory input was *frog*, the listener would start with the phonological representation, how the word *frog* sounds, then work up to the phono-lexical representation, the word *frog* as a lexical unit, and finally up to the semantic representation, what a *frog* is as a concept. Top-down processing in speech perception would work in the opposite direction. The listener would start with the semantic representation, then work their way down to the lexical and phonetic representations.

In quiet or easy listening conditions, bilinguals may rely on bottom-up processing to understand speech. However, if interference is introduced, they may rely on top-down processing to understand speech and fill in the missing information caused by difficult listening conditions. Bilinguals’ less accurate speech recognition in noise may be due to difficulties in top-down processing. Difficulty may arise while processing contextual cues due to limited proficiency, which would indicate that difficulties in top-down processing play a role (Pichora-Fuller, 2008; Shi, 2010). Previous research has shown evidence that speech recognition in noise can be facilitated when there is contextual constraint (Florentine, 1985; Gleason & Francis, unpublished data; Mayo et al., 1997; Moberly & Reed, 2019; Shi, 2010). The constraint of a sentence can be identified by the predictability of its final word. For example, "*We saw a flock of wild geese.*" is a high constraint sentence because the last word '*geese'* is easily predicted due to the nature of the first part of the sentence. At the opposite end of the spectrum is a low constraint sentence, "*Mary knows about the rug*". The final word '*rug'* is not easily predictable and therefore the sentence is considered low constraint. The level of constraint of a sentence may help listeners identify words in the sentence when speech is presented in difficult context. This use of context and constraint to assist in identifying words is an example of using top-down processing. Previous research has found higher repetition accuracy (Cooke et al., 2008; Gleason & Francis, 2021; Mayo et al., 1997) and lower subjective noise ratings (Gleason & Francis, 2021) in high constraint sentences presented in background noise.

Another possibility may be that both bottom-up and top-down processing contribute to the speech perception difficulties found in bilingual listeners (Shi, 2010). It is possible that the use of bottom-up and top-down processing depends on the quality of input (Kalikow et al., 1977; Luce & Pisoni, 1998; Moberly & Reed, 2019). Therefore, speech perception may be an interaction between incoming acoustic data and long-term language knowledge (Tuennerhoff & Noppeney, 2016). Thus, speech perception difficulties in bilinguals may occur in both bottom-up and top-down processes and explain their difficulty recognizing speech in noisy environments.

It appears that repetition and contextual effects in speech perception arise from top-down processing. However, it is not clear at which level of language these effects occur. Previously, we examined the contribution of top-down processing to speech perception and memory for spoken sentences in background noise in a single-language experiment (Gleason & Francis, unpublished data). Participants were exposed to sets of high and low constraint sentences in the auditory or visual modality in the encoding phase. Participants then completed a test phase in which sentences from the encoding phase were mixed with new sentences. Sentences at test were presented auditorily with varying levels of background noise. Participants listened to the sentences and attempted to repeat the last word of the sentence then rated the level of background noise. Having an auditory presentation of sentences at encoding lowered the subjective ratings of background noise and increased repetition accuracy, this benefit was greatest at the highest noise level. Visual presentation of sentences at encoding also increased repetition accuracy relative to new sentences and lowered noise ratings. The finding that there was any transfer at all from visual encoding to an auditory test is remarkable and suggests that the repetition effects in accuracy and noise ratings is driven in part by top-down processing. It is not clear, however, whether the top-down processing that underlies this transfer across modalities arises from the semantic level or from another level of representation that should be included in models of speech perception. This additional level of representation may be the lemma level.

The lemma level of language, as described by (Levelt, 1983), is a representational level of language that connects phonological information to semantic and syntactic information. Although this construct was initially used for speech production, it can be extended to speech comprehension. The lemma connects a word’s semantic, syntactic, phonological, and orthographic representations but does not itself contain any of these representations. Thus, it is modality general, in that it would not matter if the input were visual or auditory, the same lemma would be accessed. However, lemmas are language specific, meaning the lemma accessed will depend on the input language.

If the effects occur at the lemma level, repetition priming could facilitate identification of phonology even without the involvement of semantic processing. This would imply that the effects would transfer across modalities but would not transfer across languages. If the top-down processing arises from the semantic level, then repetition effects would be expected to transfer across both modality and language.

Current Study

The primary aim of the current study was to determine whether previously found repetition benefits in accuracy (Cooke et al., 2008; Gleason & Francis, 2021; Mayo et al., 1997) and noise rating effects (Gleason & Francis, 2021; Jacoby et al., 1988) occur at the semantic level. If repetition priming effects transfer across languages, this would indicate that top-down processing occurs at the semantic level. If repetition priming effects do not transfer across languages, this would indicate that the top-down processing occurs at a language-specific level. We therefore examined whether listening to Spanish sentences at encoding would increase accuracy and lower noise ratings when listening to the English translations in noise.

The second aim of the study was to determine what role proficiency and age of acquisition play in the ability to benefit from repetition. We anticipate results to follow the same general pattern found in previous research (Mayo et al., 1997), that higher proficiency will lead to more accurate recognition of speech in noise. It is expected that higher-proficiency speakers with an earlier age of second language acquisition will have better accuracy and lower noise ratings overall regardless of repetition, as in previous research (Mayo et al., 1997). Based on the hypothesis that lower-proficiency speakers cannot use top-down processing as efficiently as higher-proficiency speakers, we hypothesized that the effects of repetition would be weaker, to the extent that these effects depend on top-down processing. Thus, higher proficiency speakers were expected to benefit more from repetition in both repetition accuracy and subjective noise ratings.

Method

Participants

Participants (N = 60) were normal hearing bilingual students enrolled at the University of Texas at El Paso (UTEP) participating for course credit. Participants were all young adults (*Mdn age* = 20), most identified as female (68.3%) and of Latinx ethnicity (87.9%). Ninety-one percent of participants reported Spanish as their first language (*Mdn AoA* = 1), and English as their second language acquired (*Mdn AoA* = 6). All participants reported that they considered themselves to have normal hearing; potential participants who were experiencing any hearing issues did not complete the experimental protocol. Bilingual status was determined by administering the English and Spanish versions of the Woodcock-Muñoz Language Survey-Revised (WMLS-R; Woodcock, et al., 2005). In each version of the assessment, participants completed picture vocabulary and verbal analogies tasks. During the picture vocabulary task participants were shown drawings of items and were asked to name each item. The verbal analogies task required participants to complete an analogy with the most logical response. For example, "*A fish swims; a bird...*" *"Flies"*. Scores from these tasks were tallied and submitted to the IIP-WMLS-R computer program, which calculated a composite Oral Language proficiency score and an age-equivalency score (Schrank & Woodcock, 2005). To qualify for the study participants must have received an age equivalency score of 8 years old or higher on the Oral Language composite in both English (*M* = 13.43, *SD* = 5.03) and Spanish (*M* = 12.00, *SD* = 3.73) versions of the assessment. The minimum requirement of an age equivalency score of 8 years was chosen to be consistent with previous published studies in our laboratory that used the same two subtests of the WMLS-R.

Design

The experiment had a 3 (encoding language) x 5 (noise level) within-subjects design. In the encoding phase, stimuli sentences were presented in Spanish, English, or not at all. At test, English sentences were presented with five different levels of background noise (56 dB, 58 dB, 60 dB, 62 dB, and 64 dB). The dependent variables were repetition accuracy and noise ratings in the test phase.

Materials

The primary task that participants completed was a modified Speech Perception In Noise (SPIN) task (Bilger et al., 1979). Only low-constraint sentences, those in which the final word was not predictable, were selected. Low-constraint sentences were used, because our previous study showed stronger priming effects in the low-constraint condition (Gleason & Francis, 2021). To be sure that the sentences were in fact low constraint, a different group of participants (N = 35) was asked to read sentence frames and then type in the word that best completed each sentence frame. Results showed that for low constraint sentences zero percent of participants produced the target word. This confirmed that the target words in low constraint sentences were unpredictable.

From the normed sentences, 150 sentences were selected and translated to Spanish. The sentences were randomly divided into 3 sets of 50 sentences. The assignment of each set to encoding conditions and noise levels at test was counterbalanced across participants.

For use in the encoding phase, the English and Spanish sentences were recorded in a soundproof room using the audio software Audacity Version 2.3.3. The speaker for both English and Spanish sentences was a native Spanish/English bilingual female with a neutral accent. In the test phase, we used English sentence recordings (Desjardins & Doherty, 2013) that were spoken by a different female speaker with a neutral accent. These audio files were presented using the standard speakers on the experiment computer, an iMac 2013, at an average volume of 56 dB throughout the experiment.

The background noise used during practice and test in the SPIN task was speech shaped noise, which is a white noise that shares a similar frequency to normal speech. Five distinct levels of background noise were used, all the same volume or louder than the sentences. Levels from quietest to loudest were 56 dB, 58 dB, 60 dB, 62 dB, and 64 dB.

Procedure

Participants first read and signed an informed consent form. Participants then completed both the English and Spanish versions of the Woodcock-Muñoz language survey. Once the language assessments were completed researchers confirmed that the participant qualified for the experiment. After participants completed the language assessments, they completed experimental tasks. The experiment consisted of three phases, encoding, practice, and test. Verbal instructions were given at each of the three phases. During the encoding phase participants listened to 50 English sentences and 50 Spanish sentences separated into separate blocks of trials with language order counterbalanced. Participants were instructed to listen to each set of sentences silently. Within each block, sentences were presented in random order with no background noise.

The practice and test phases involved a modified SPIN task. Participants were informed that they would be listening to sentences mixed with background noise Participants listened to a sentence, attempted to report the last word of the sentence, and then rated the level of background noise on a scale of one to five (one being the quietest and five the loudest). The experimenter scored reporting accuracy in real time, using a list of expected responses. To familiarize participants with the task, they completed a practice set of ten trials, two trial for each level of background noise, before they moved onto the test phase. At test, the 100 sentences from encoding were mixed with 50 new sentences. All sentences were presented auditorily in English mixed with background noise. Sentences and background noise levels were randomly intermixed. Participants were never informed that sentences from the encoding task were repeated at test.

After completing the modified SPIN task participants completed a demographic survey and the ESPADA language background questionnaire. Both questionnaires were completed virtually using Qualtrics survey platform. Once finished, participants were debriefed, compensated and thanked for their participation.

Results

Accuracy

Accuracy results were analyzed using a repeated-measures analysis of variance, including the three encoding levels and the five noise levels at test as within-subjects factors. As expected, as background noise increased accuracy repeating the target word decreased, as indicated by a significant negative linear trend, $F(1,59) = 897.94$, $MSE = .007$, $p < .001$, $\eta_p^2 = 0.94$, see Figure 2. Participants were more accurate for sentences previously presented in English than for newly presented sentences $F(1, 59) = 11.496$, $MSE = .034$, $p = .001$, $\eta_p^2 = 0.16$, or sentences previously presented in Spanish, $F(1, 59) = 10.217$, $MSE = .029$, $p = .002$, $\eta_p^2 = 0.15$. Importantly, there was no significant difference in repetition accuracy between sentences that were previously presented in Spanish and newly presented sentences, $F(1, 59) = 0.194$, $MSE = .036$, $p = .661$, $\eta_p^2 = 0.003$, see Figure 3. There were no significant interactions between the encoding language comparisons and the linear trend on noise level (*p*s > .1).

Subjective Noise Ratings

To analyze subjective noise ratings, a separate repeated-measures analysis of variance was used. Again, the three encoding levels and the five noise levels at test were included as within-subject factors. Overall, rating results followed a similar pattern to accuracy results, As expected, participants gave higher subjective noise ratings to sentences with higher noise levels, as indicated by a significant linear trend, $F(1, 59) = 230.813$, $MSE = .278$, $p < .001$, $\eta_p^2 = 0.80$. Participants rated sentences previously presented in English as significantly quieter than newly presented sentences, $F(1, 59) = 15.44$, $MSE = .147$, $p < .001$, $\eta_p^2 = 0.21$, or sentences previously presented in Spanish, $F(1, 59) = 11.717$, $MSE = .142$, $p = .001$, $\eta_p^2 = 0.17$. There was no significant difference in noise ratings between sentences previously presented in Spanish and newly presented sentences, $F(1, 59) = 0.403$, $MSE = .118$, $p = .528$, $\eta_p^2 = 0.007$. There were no significant interactions between encoding language comparisons and the linear trend on noise level ($ps > .1$).

Proficiency and AoA

To understand the associations of participant proficiency and AoA with accuracy, correlations among these variables were computed. Participants' composite oral language proficiency scores in English were positively correlated with overall accuracy, $r(58) = .614$, $p < .001$. Spanish proficiency scores were not significantly correlated with overall accuracy, $r(58) = -.159$, $p = .225$. There was no significant correlation between English proficiency and accuracy priming for the English encoding condition, $r(58) = .189$, $p = .149$, or the Spanish encoding condition, $r(58) = .214$, $p = .100$. Similarly, no significant correlation was found between Spanish proficiency and accuracy priming for English, $r(58) = .058$, $p = .662$, or Spanish, $r(58) = .148$, $p = .258$, encoding conditions. There was a significant negative correlation between English age of acquisition and overall accuracy, $r(58) = -.345$, $p = .007$. There was also a significant negative correlation between English age of acquisition and accuracy priming in English, $r(58) = -.269$, $p = .038$. There was no significant correlation between English age of acquisition and Spanish priming, $r(58) = -.140$, $p = .285$. Spanish age of acquisition was not included in our analyses due to the highly-skewed distribution in our demographic. See Table 1 for complete repetition accuracy correlation results.

In relation to the associations of participant proficiency and AoA with subjective noise ratings, correlations among these variables were computed. English proficiency scores were not significantly correlated with overall subjective noise ratings, $r(58) = -.080$, $p = .546$. Similarly, there was no significant correlation found for Spanish proficiency and overall subjective noise ratings, *r*(58) = .122, *p* = .352. No significant correlations were found between English proficiency and noise rating priming for English, *r*(58) = .094, *p* = .474, or Spanish, *r*(58) = -.098, *p* = .457, encoding conditions. There were also no significant correlations found between Spanish proficiency and noise rating priming for English, *r*(58) = .140, *p* = .288, or Spanish, *r*(58) = .133, *p* = .312, encoding conditions. English age of acquisition was not correlated with overall subjective noise ratings, *r*(58) = .210, *p* = .107. English age of acquisition was not correlated with noise rating priming in English, *r*(58) = -.086, *p* = .515, or Spanish, *r*(58)= -.141, *p* = .283, encoding conditions. Again, Spanish age of acquisition was not included in our analyses due to a highly-skewed distribution. See Table 2 for complete subjective noise ratings correlation results.

Discussion

The current study aimed to further understand the mechanisms underlying repetition effects on bilingual speech recognition in difficult listening environments. The first aim of the current study was to determine whether previously found repetition effects on speech recognition accuracy (Cooke et al., 2008; Gleason & Francis, 2021; Mayo et al., 1997) and noise ratings (Jacoby et al., 1988) occur at the semantic or lemma level. As expected, the present results replicated previous findings that speech recognition accuracy was significantly higher for sentences presented in English during encoding than new sentences. Participants also rated background noise significantly quieter for sentences encoded in English than for new sentences.

Both the present study and the current study had English auditory encoding conditions, and in both studies, the effect sizes in priming of reporting accuracy were large ($\eta_p^2 = .235$ in the previous study; $\eta_p^2 = 0.16$ in the present study). However, while effect size for accuracy of sentences previously presented visually were medium ($\eta_p^2 = .072$), the effect size for accuracy of sentences previously presented in Spanish in the current study were small ($\eta_p^2 = 0.003$). A similar pattern is present when comparing the noise rating effect sizes across the two studies. Effect sizes for sentences presented auditorily in English at encoding were large in both studies ($\eta_p^2 = .173$ in the previous study; $\eta_p^2 = 0.21$ in the current study). Again, the effect size for priming in noise ratings for sentences previously presented visually were medium to large ($\eta_p^2 = 0.102$), and the effect size for noise ratings for sentences presented in Spanish at encoding were small ($\eta_p^2 = 0.007$). (Note that a direct statistical comparison across studies is not warranted, because the previous study included monolingual participants and high-constraint sentences.)

An important finding in the current study was the absence of a difference in accuracy or noise ratings for sentences that were previously presented in Spanish and sentences that were new to the participants. The absence of a difference between Spanish and new sentences demonstrate that these repetition-priming effects do not transfer across languages. This indicates that the top-down processing responsible for these effects occurs at a language-specific level of representation.

The BLINCS model (Shook & Marian, 2013) incorporates both bottom-up and top-down processing by allowing information to move bi-directionally across levels of language, see Figure 1. Many models of speech comprehension, such as the BLINCS model, do not include a level for representations that are modality-general but language-specific. For example, the highest level of language representation included in the BLINCS model, semantic representation, is modality-general and language-general. Our results suggest that top-down processing does not arise from this level because the repetition effects do not transfer across languages. The next level included, the phono-lexical representation, implies that representations at this level are language-specific and modality-specific, and therefore repetition effects would not be expected to transfer across languages or modalities. Previous research (Gleason & Francis, 2021) demonstrated that repetition effects in recognition accuracy and noise ratings transferred across modalities. An important aspect that should be considered when contemplating these results is the finding that reading automatically activates phonology (Berent & Perfetti, 1995; Frost & Kampf, 1993). Therefore, we cannot rule out the possibility that this phonological activation underlies the transfer across modalities. These findings paired with the current study indicate the need to include an intermediate level of representation that is modality-general but language-specific. We therefore suggest that models of speech recognition should include an intermediate level between semantic and lexical representations, a lemma level. As an example of how this level may be included, we propose a modified BLINCS model, see Figure 6.

The secondary aim of the study was to determine what role proficiency and age of acquisition play in the benefits of repetition. As seen in previous research (Mayo et al., 1997; Shi, 2009, 2010), both English proficiency and English AoA were significantly correlated with overall accuracy, indicating that higher English proficiency and earlier English AoA were associated with higher recognition accuracy. Higher English proficiency was also associated with larger priming effects in the English encoding condition but not in the Spanish encoding condition. Proficiency scores in Spanish did not correlate with the size of the repetition priming effects in accuracy in either language condition. Similarly, English age of acquisition was not correlated with accuracy priming in either language condition. Correlations involving Spanish age of acquisition were not analyzed, because the majority of participants acquired Spanish first or simultaneously with English, and therefore the form of the distribution was not appropriate for a correlational analysis.

There were no significant correlations between language proficiency and overall subjective noise ratings or noise rating priming for either language. These results were expected due to the subjective aspect of the rating scale. When rating the level of background noise at test participants were instructed to use the full extent of the scale. Because of this, noise ratings were similar across participants regardless of language proficiency.

Recognizing speech in noise is a cognitively demanding task (Rudner et al., 2012; Wong et al., 2008). Repetition based on top-down processing facilitates speech recognition (Jacoby et al., 1988; Pichora-Fuller, 2008). However, it had yet to be determined which levels of language representation are involved in this process. The current study demonstrated that repetition effects do not transfer across languages, indicating that they occur at a language-specific but modality-general level. We suggest that models of speech recognition and comprehension should include a level of representation that fits this description, such as the lemma. Our findings also replicate previous research showing that proficiency and AoA play a role in bilinguals’ ability to recognize speech in noise and indicate that listeners with higher proficiency benefit more from repetition.

Table 1: Correlations of Repetition Accuracy and Language Background Measures

	English Proficiency [a]	Spanish Proficiency [a]	Age of English Acquisition	English Priming	Spanish Priming	Overall Accuracy
English Proficiency [a]	-					
Spanish Proficiency [a]	-0.109	-				
Age of English Acquisition	-.293*	0.213	-			
English Priming	0.189	0.058	-.269*	-		
Spanish Priming	0.214	0.148	-0.14	.589**	-	
Overall Accuracy	.614**	-0.159	-.345**	0.182	0.164	-

Note. The correlation between English and Spanish priming is because the scores were created by subtracting from the same baseline.

[a] Proficiency scores used were W scores from the WMLS-R

*$p < .05$; **$p < .01$

Table 2: Correlations of Subjective Noise Ratings and Language Background Measures

	English Proficiency [a]	Spanish Proficiency [a]	Age of English Acquisition	English Priming	Spanish Priming	Overall Ratings
English Proficiency [a]	-					
Spanish Proficiency [a]	-0.109	-				
Age of English Acquisition	-.293*	0.213	-			
English Priming	0.094	0.14	-0.086	-		
Spanish Priming	-0.098	0.133	-0.141	.467**	-	
Overall Ratings	-0.08	0.122	0.21	-0.08	-0.014	-

Note. The correlation between English and Spanish priming is because the scores were created by subtracting from the same baseline.

[a] Proficiency scores used were W scores from the WMLS-R

**p* < .05; ***p* < .01

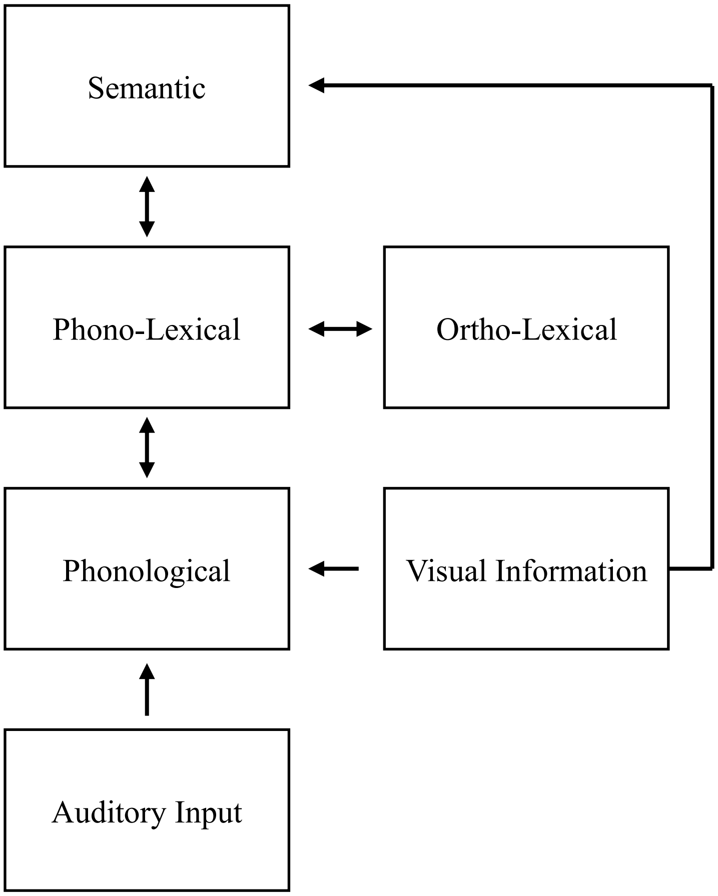

Figure 1: The Bilingual Language Interaction Network for Comprehension of Speech

Note. The Bilingual Language Interaction Network for Comprehension of Speech (Shook & Marian, 2012)

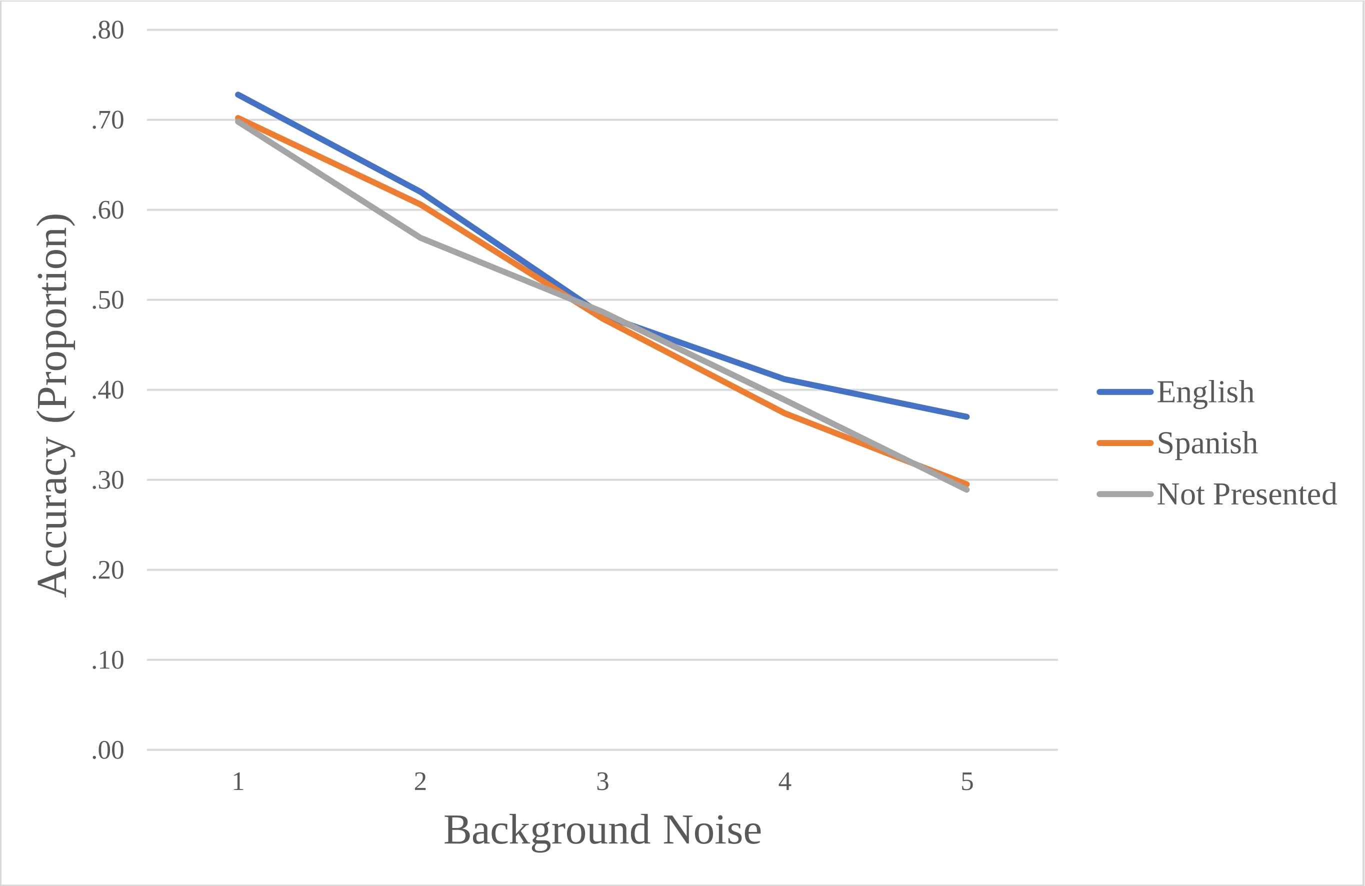

Figure 2: Repetition Accuracy Across Noise Levels

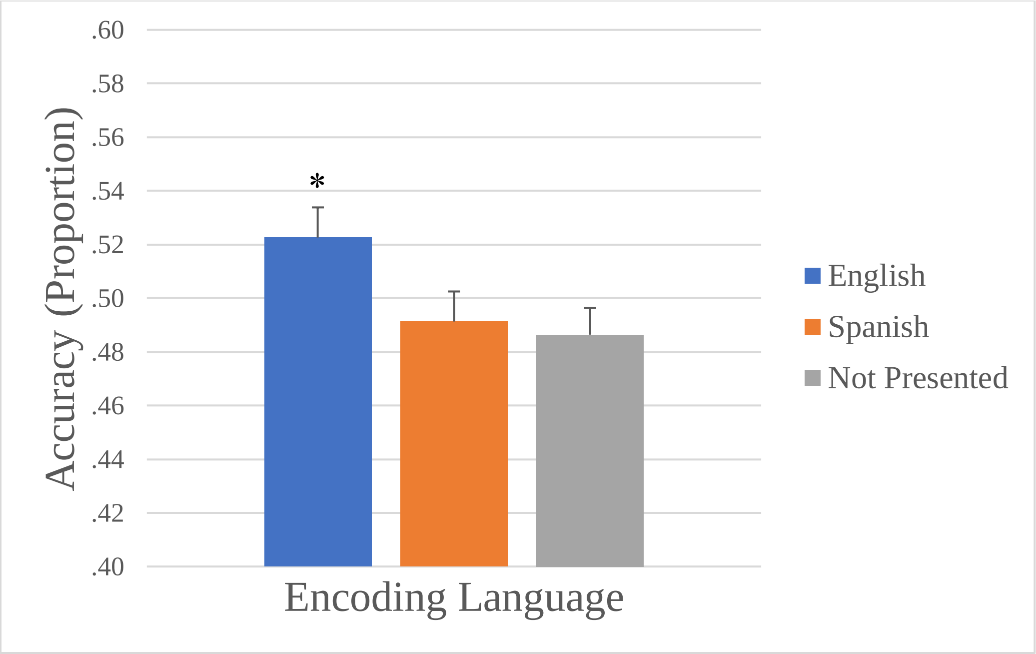

Figure 3: Repetition Accuracy as a Function of Encoding Language

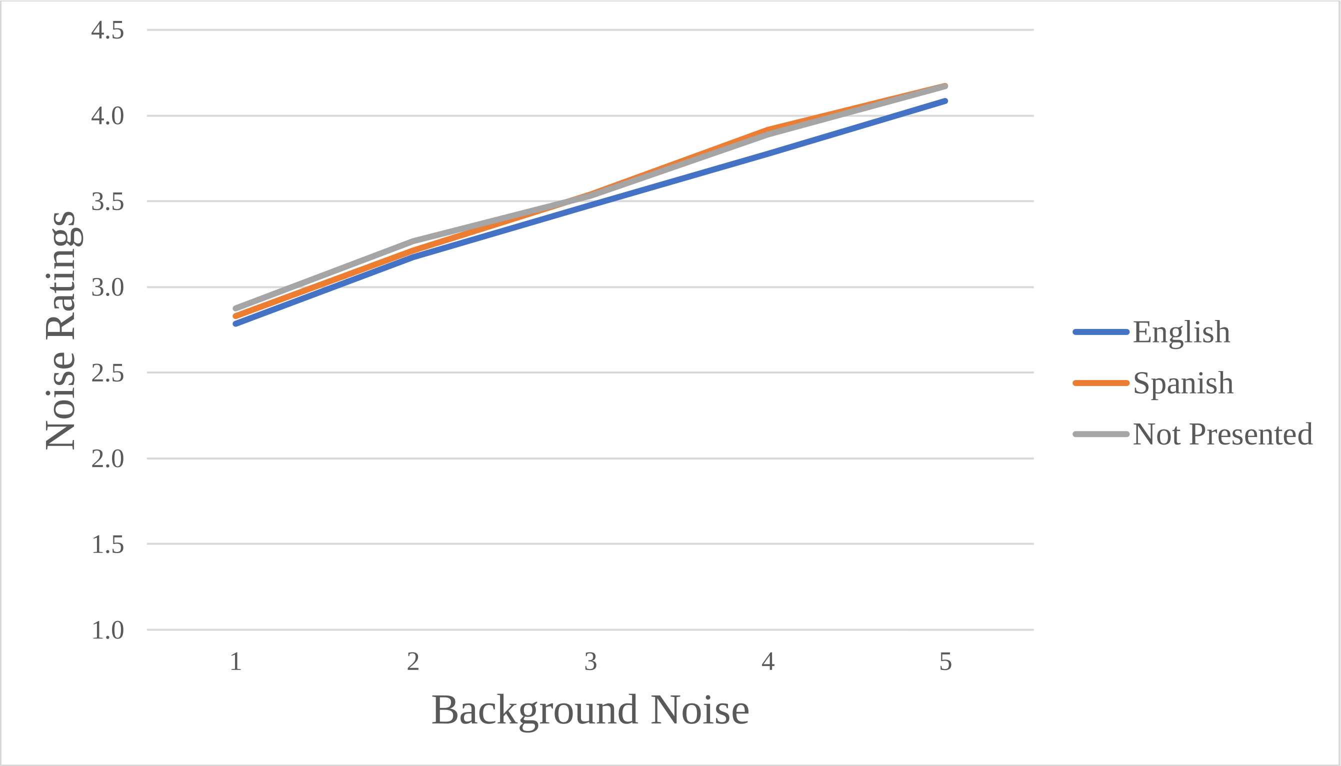

Figure 4: Subjective Noise Ratings Across Noise Levels

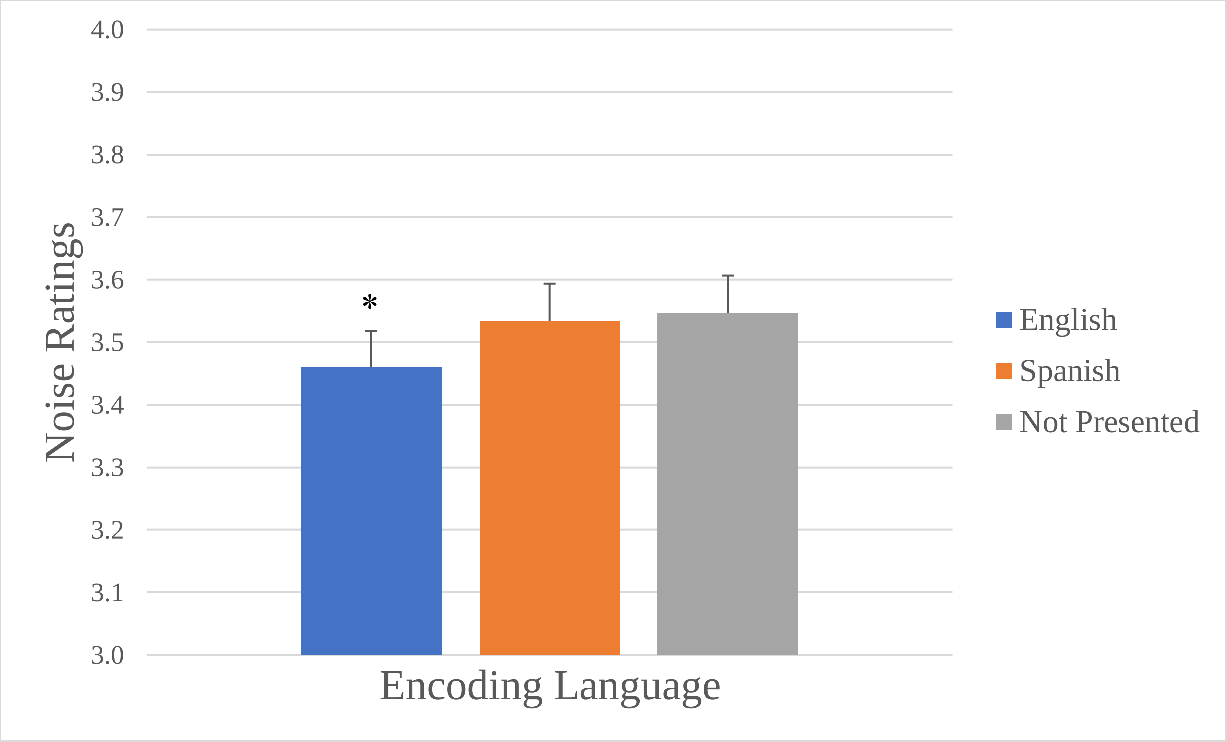

Figure 5: Subjective Noise Ratings as a Function of Encoding Language

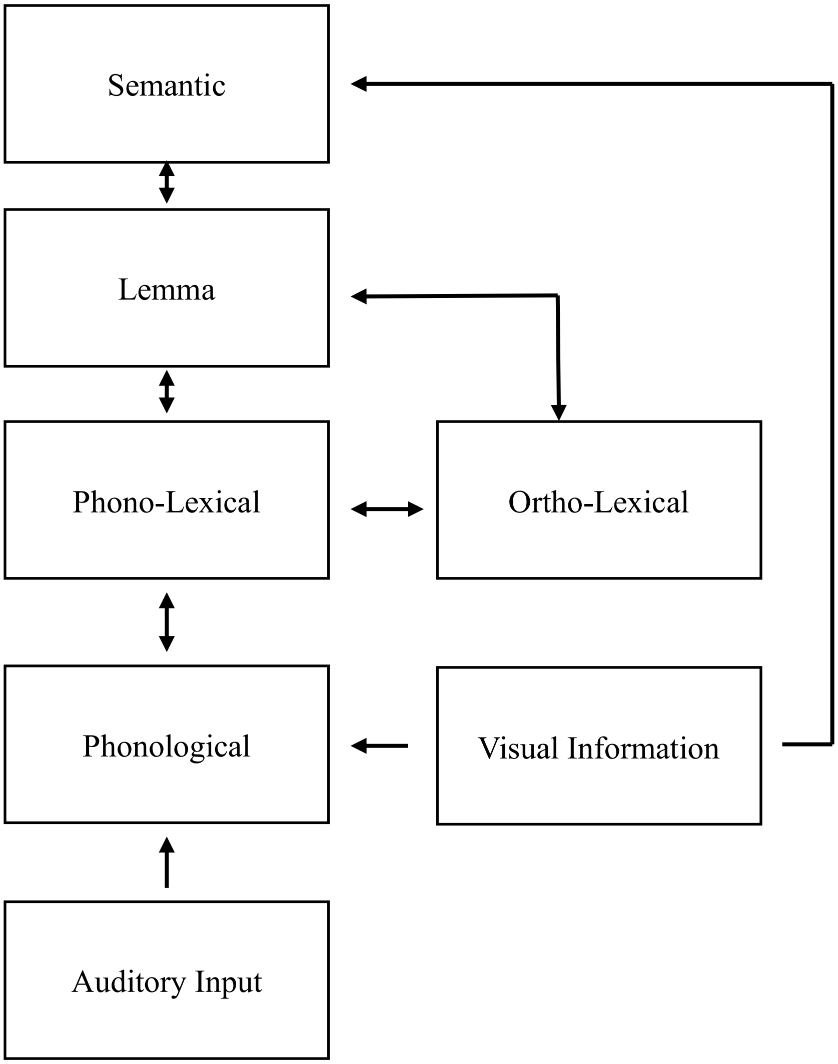

Figure 6: BLINCS Model: Revised

Note. A revised model of BLINCS (Shook & Marian, 2012) that incorporates a lemma level of language representation.

References

Berent, I., & Perfetti, C. A. (1995). A rose is a REEZ: The two-cycles model of phonology assembly in reading English. Psychological Review, 102(1), 146–184. https://doi.org/10.1037/0033-295X.102.1. 146

Bilger, R. C., Nuetzel, J. M., Rabinowitz, W. M., & Rzeczkowski, C. (1984). Standardization of a test of speech perception in noise. *Journal of Speech, Language, and Hearing Research, 27*(1), 32-48.

Bradlow, A. R., & Alexander, J. A. (2007). Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *The Journal of the Acoustical Society of America*, *121*(4), 2339–2349. https://doi.org/10.1121/1.2642103

Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *The Journal of the Acoustical Society of America*, *112*(1), 272–284. https://doi.org/10.1121/1.1487837

Cooke, M., Garcia Lecumberri, M. L., & Barker, J. (2008). The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *The Journal of the Acoustical Society of America*, *123*(1), 414–427. https://doi.org/10.1121/1.2804952

Crandell, C. C., & Smaldino, J. J. (1996). Speech Perception in Noise by Children for Whom English Is a Second Language. *American Journal of Audiology*, *5*(3), 47–51. https://doi.org/10.1044/1059-0889.0503.47

Cutler, A., Weber, A., Smits, R., & Cooper, N. (2004). Patterns of English phoneme confusions by native and non-native listeners. *The Journal of the Acoustical Society of America*, *116*(6), 3668–3678. https://doi.org/10.1121/1.1810292

Desjardins, J. L., & Doherty, K. A. (2013). Age-Related Changes in Listening Effort for Various Types of Masker Noises. *Ear and Hearing*, *34*(3), 261-272.

Florentine, M. (1985). Speech perception in noise by fluent, non-native listeners. *The Journal of the Acoustical Society of America*, *77*(S1), S106–S106. https://doi.org/10.1121/1.2022152

Frost, R., & Kampf, M. (1993). Phonetic recoding of phonologically ambiguous printed words. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19(1), 23–33. https://doi.org/10. 1037/0278-7393.19.1.23

Gabrieli, J. D. E. (1998). Cognitive neuroscience of human memory. *Annual Review of Psychology*, *49*(1), 87–115. https://doi.org/10.1146/annurev.psych.49.1.87

Gleason & Francis (2021). *Mechanisms of Predictability and Repetition-Priming Effects on Speech Recognition in Noise* [Conference presentation]. Psychonomic Society, Virtual Meeting.

Jacoby, L. L., Allan, L. G., Collins, J. C., & Larwill, L. K. (1988). *Memory Influences Subjective Experience: Noise Judgments*. 8.

Kalikow, D. N., Stevens, K. N., & Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *The Journal of the Acoustical Society of America*, *61*(5), 1337–1351. https://doi.org/10.1121/1.381436

Kathleen Pichora-Fuller, M. (2008). Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing. *International Journal of Audiology*, *47*(sup2), S72–S82. https://doi.org/10.1080/14992020802307404

Lecumberri, M. L. G., & Cooke, M. (2006). Effect of masker type on native and non-native consonant perception in noise. *The Journal of the Acoustical Society of America*, *119*(4), 2445–2454. https://doi.org/10.1121/1.2180210

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and brain sciences*, *22*(1), 1-38.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing Spoken Words: The Neighborhood Activation Model: *Ear and Hearing*, *19*(1), 1–36. https://doi.org/10.1097/00003446-199802000-00001

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. https://doi.org/10.1044/jslhr.4003.686

Moberly, A. C., & Reed, J. (2019). Making Sense of Sentences: Top-Down Processing of Speech by Adult Cochlear Implant Users. *Journal of Speech, Language, and Hearing Research*, *62*(8), 2895–2905. https://doi.org/10.1044/2019_JSLHR-H-18-0472

Rudner, M., Lunner, T., Behrens, T., Thorén, E. S., & Rönnberg, J. (2012). Working Memory Capacity May Influence Perceived Effort during Aided Speech Recognition in Noise. *Journal of the American Academy of Audiology*, *23*(8), 577–589. https://doi.org/10.3766/jaaa.23.7.7

Sheldon, S., Pichora-Fuller, M. K., & Schneider, B. A. (2008). Priming and sentence context support listening to noise-vocoded speech by younger and older adults. *The Journal of the Acoustical Society of America*, *123*(1), 489–499. https://doi.org/10.1121/1.2783762

Shi, L.-F. (2009). Normal-hearing English-as-a-second-language listeners’ recognition of English words in competing signals. *International Journal of Audiology*, *48*(5), 260–270. https://doi.org/10.1080/14992020802607431

Shi, L.-F. (2010). Perception of Acoustically Degraded Sentences in Bilingual Listeners Who Differ in Age of English Acquisition. *Journal of Speech, Language, and Hearing Research*, *53*(4), 821–835. https://doi.org/10.1044/1092-4388(2010/09-0081)

Shook, A., & Marian, V. (2013). The Bilingual Language Interaction Network for Comprehension of Speech. *Bilingualism: Language and Cognition*, *16*(2), 304–324. https://doi.org/10.1017/S1366728912000466

Tuennerhoff, J., & Noppeney, U. (2016). When sentences live up to your expectations. *NeuroImage*, *124*, 641–653. https://doi.org/10.1016/j.neuroimage.2015.09.004

von Hapsburg, D., & Bahng, J. (2006). Acceptance of Background Noise Levels in Bilingual (Korean-English) Listeners. *Journal of the American Academy of Audiology*, *17*(09), 649–658. https://doi.org/10.3766/jaaa.17.9.4

Wong, P. C. M., Uppunda, A. K., Parrish, T. B., & Dhar, S. (2008). Cortical Mechanisms of Speech Perception in Noise. *Journal of Speech, Language, and Hearing Research*, *51*(4), 1026–1041. https://doi.org/10.1044/1092-4388(2008/075)

Vita

Liam Gleason received his Bachelor of Science in Psychology with a minor in Spanish from Tennessee Technological University in 2018. His anticipated graduation date with a Ph.D. in Psychology from the University of Texas at El Paso is 2024. Liam is a member of Dr. Wendy Francis' Bilingual Cognition Laboratory and Dr. Chris Potter's laboratory, Children's Learning Project. His research interests involve bilingualism, memory, and cognition. Specifically, he is interested in how the languages we speak can impact our memory and other cognitive functions, and factors that impact speech recognition.