# *Effects of Interfering Speech in a Foreign Language on Speech Understanding*

**Cheyanne K. Waller, McNair Scholar**
**The Pennsylvania State University**

**McNair Faculty Research Adviser:**
**Navin Viswanathan, Ph.D.**
**Associate Professor & Professor in Charge of the Undergraduate Program**
**Department of Communication Sciences and Disorders**
**College of Health and Human Development**
**The Pennsylvania State University**

**Abstract:** Measuring language proficiency is a challenging process that typically includes a combination of standardized test and self-reporting assessments. The disadvantage of these assessments is that the standardized test can be time consuming, and the self-reported measures can be biased. To address these disadvantages, current research has studied the effects of interfering speech in a foreign language and language proficiency levels on speech understanding. To evaluate whether there is a correlation between language proficiency levels and performance scores on speech-in-speech recognition task we must first test the validity of speech-in-speech recognition task. To do so, we had monolingual English listeners participate in a speech-in-speech recognition task to see if they demonstrate a linguistic release from masking. We use monolingual English listeners because we hypothesize that bilingual listeners with high proficiency in their second language will have similar scores to the monolingual listeners. So, monolingual listeners scores must be included to compare with the bilingual listeners. We then plan on having bilingual participants take part in two tasks. In the first task will be the speech-in-speech recognition task, where participants will hear a target sentence in the presence of foreign and native background noise, with the job of correctly identifying the color and number heard from the target. The next task will be a standard language proficiency task. We expect to see a positive correlation with higher language proficiency levels correlating to higher performance scores on speech-in-speech recognition task. Making speech-in-speech recognition task an alternative way of measuring language proficiency,

## I. Introduction

Schools and businesses need to be able to measure an individual's language proficiency. Public Schools also rely on language proficiency tests for the academic placement for their bilingual students. For example, international students planning to attend a college or university in the United States, must take an exam prior to starting classes to assess their level of language proficiency. During the hiring process, companies may also consider language proficiency. According to the Corporate Finance Institute, CFI, potential employees are assessed on their language proficiency on a 0-5 language proficiency scale. With 0 being no/limited proficiency and 5 being native/bilingual proficiency (CFI Education Inc, 2015). With the reliance on language proficiency exams in the public and private sector for schools and business, there needs to be an efficient approach to assessing language proficiency.

The current approach to testing language proficiency is through standardized assessments. The most common method used is the Test of English as a Foreign Language, also known as TOEFL. Other common Language Proficiency exams include the International English Language Testing System (IELTS), Certificate of Advanced English (CAE), and the Certificate of Proficiency in English (CPE). These proctored exams are administered to assess one's proficiency level in English (Academic Positions, 2020).

The TOEFL consist of 4 sections: reading, listening, speaking, and writing. In the reading section, participants read passages and respond to questions. In the listening section, participants listen to a brief lecture or classroom discussion and then answer questions about what they heard. In the speaking section, participants have a conversation about things already discussed in the exam and other familiar topics. In the writing section, participants read a passage, listen to a recording, and then type out a response. This test takes approximately four hours to complete (TOEFL iBT® Test Content, 2021). Another method of assessing language proficiency is through self-reported measures such as language history questionnaires (LHQ) or language experience and proficiency questionnaires (LEAP-Q). Both the TOEFL and self-reported measures, LHQ and LEAP-Q, have disadvantages. The TOEFL takes an extended amount of time to complete, and the self-reported measures could be unreliable if the person underreported or overestimated their language proficiency level. Therefore, goal of this research is to determine if a correlation exist between language proficiency levels and performance scores on speech-in-speech recognition task to find an alternative way of measuring language proficiency that does not rely on self-reported measures.

Current research has assessed the effect second language proficiency has on speech-in-speech recognition tasks (Francis et al., 2018). Speech-in-speech recognition task are when participants are assessed on their ability to comprehend target speech in the presence of background noise, also known as a masker. The widely studied phenomenon, Linguistic Release from Masking (LRM), also looks at which situations make completing speech-in-speech recognition tasks easier. LRM is when participants have an easier time recognizing target speech when the competing speech is in a different language (Calandruccio et al., 2010; Brouwer et al., 2012). The research also shows that language familiarity plays a role in the performance scores of speech-in-speech recognition task. If a participant is proficient with both languages, they will demonstrate a greater linguistic release from masking than individuals who are less proficient with one or both languages (Francis et al., 2018).

This study looked at the performance scores of monolingual English listeners on a speech-in-speech recognition task. Before identifying a correlation between language proficiency levels and speech-in-speech recognition scores, validity of the speech-in-speech recognition task had to be established. Past data had to be replicated by having monolinguals listeners complete a speech-in-speech recognition task. The participants showed a higher performance score when the target sentence was in a different language than the masker compared to when the target and masker were in the same language. The purpose of using monolingual listeners was to analyze their scores and compare them to the scores of the bilingual listeners. Past research has showed that monolingual listeners tend to have higher performance scores on speech-in-speech recognition task when the target and masker differ in language. By comparing the data, we hypothesize that bilingual participants who are highly proficient in their second language will have similar performance scores to that of the monolingual listeners.

To expand upon this study an assessment of the correlation of language proficiency and speech-in-speech recognition scores will be completed by Dutch-English bilinguals. To assess

the correlation, Dutch-English bilingual participants will complete a language proficiency battery followed by an LRM quick test. Past research suggests that when bilingual individuals have high proficiency in their second language they perform similarly to monolingual listeners on speech-in-speech recognition tasks (Miller, 2019). We hypothesize that Dutch-English bilinguals with high English proficiency will have a greater amount of linguistic release from masking that is similar to the amount of linguistic release from masking monolingual listeners experience. We also hypothesize that Dutch-English bilinguals with low English proficiency will have a low amount of linguistic release from masking. In other words, when the target and masker are in different languages, versus when they are in the same language, there will be similar performance scores and it will show a low amount of linguistic release from masking.

## I. Experiment I: Pilot Study of Coordinate Response Measures

Experiment 1 used Coordinate Response Measures (CRM) as the speech-in-speech recognition task on English monolinguals to see if they would demonstrate a linguistic release from masking when the target and background speech differ. The purpose behind using monolingual English listeners was to have comparable results from known individuals with high proficiency levels in English.

### A. Methods

#### *Participants*

*a. Listeners.* Participants included three English Monolingual women from the Speech, Language and Cognition Lab at The Pennsylvania State University.

### B. Stimuli

a. *Talker Stimuli* – The Coordinate Response Measure (CRM) wavefiles are in the "broadband" format with 44.1 kHz sampling rates. The CRM wavefiles consist of eight talkers and have the sentence structure of: "Ready [Call sign] go to [Color] [Number] now. The talker stimuli were presented at 65dB to constitute normal conversation conditions.

b. *Masker Stimuli* – All stimuli were recorded in a sound booth at 44.1 kHz sampling rate and a 16-bit resolution. Both maskers were presented at 71dB creating a signal to noise ratio, SNR = -6dB. Both masker stimuli were composed of semantically anomalous sentences to focus the linguistic properties of the masking language. In this study, a syntactically normal sentence test (SNST) was used for the maskers. Dutch translations of these SNST sentences came from Brouwer et al. (2012).

### C. Procedures

Participants completed the experiment on Labvanced®, a browser-based experimental software. This task consisted of a headphone evaluation, practice trials, following by the actual experiment.

In the headphone evaluation, participants had to listen to three pure tone sounds and determine which sound was the softest. If they were able to identify the softest sound, they were allowed to move on to the practice trial and the actual task.

Written instructions were provided, stating that participants would hear a female talker who would say a call sign, color, and number. The format was, "Ready [call sign], go to [color] [number] now." The task was for the participants to identify the color and the number stated in each trial by selecting the button of the correct color and number. The possible colors included: Red, Blue, Green, and White. The possible numbers included in this study were numbers one through eight.

## II. Experiment II: Dutch-English Bilinguals Language Proficiency Battery and CRM task

In a future experiment, we will examine the English language proficiency level of Dutch-English bilinguals through a series of five tasks. We will also examine the Dutch-English bilinguals' ability to recognize target speech in the presence of background speech.

### A. Methods

*Participants*

a. *Listeners.* Participants will include Dutch-English Bilinguals. Participants will be selected through Prolific. Prolific is a Research Participant Repository that allows researchers to quickly find participants by launching studies to thousands of participants in minutes. The participants in the study will fill out a questionnaire with a series of questions about their language history.

### B. Stimuli

a. *Boston Naming Task* – This task consists of 30 black and white drawings that depict everyday objects and items arranged in an order where they begin to increase in difficulty. (Kaplan et al., 1983)

b. *Verbal Fluency Task* - In the letter task we will use the standard letters: "B, M, D, and T" These letters are used due to the large number of words that begin with these letters. In the category task, there were four categories selected. Those four categories were: "*Animals, Clothing, Musical Instruments, and Vegetables*."

c. *Nonword Repetition Task* - Here participants will hear 1-syllable words ranging in length from three, four, five, up to six words. All participants will hear the words in the same order, starting with the shortest list of words and going up to the longest list of words. Participants will listen to 16 sets of utterances.

d. *Lexical Decision Task* - Words for this task are from www.lextale.com. The task will include 60 words. The words within the lexTALE will range from 4 to 12 letters with a mean of 7.3 letters. According to the CELEX database, the words will also have a mean frequency that was between 1-26 occurrences per million with a mean of 6.4 (Baayan, Piepenbrock, & Guilkens, 1995).

e. *Talker Stimuli* – The Coordinate Response Measure (CRM) wavefiles are in the "broadband" format with 44.1 kHz sampling rates. The CRM wavefiles consist of eight talkers and have the sentence structure of: "Ready [Call sign] go to [Color] [Number] now. The talker stimuli will be presented at 65dB to constitute normal conversation conditions.

f. *Masker Stimuli* – All stimuli will be recorded in a sound booth at 44.1 kHz sampling rate and a 16-bit resolution. Both maskers will be presented at 71dB creating a signal to noise ratio, SNR = -6dB. Both masker stimuli will be composed of semantically anomalous sentences to focus the linguistic properties of the masking language. Syntactically normal sentence test (SNST) will be used for the maskers. Dutch translations of these SNST sentences will come from Brouwer et al. (2012).

### C. Procedures

Participants will access the experiment using a browser-based experimental software, Labvanced®. Participants will receive written instructions on Labvanced®. They will be informed that they are taking part in the first task which will include five sections. The task will begin with the first section, the Boston Naming Test (BNT). This test is an instrument used to assess naming ability and the ability to retrieve words. (Kaplan et al., 1983). Participants will be exposed to 30 black and white pictures representing everyday objects that are organized in a way

that will increases in difficulty as time passes. The participants will be asked to name each item shown in the picture.

The next section will be the Verbal Fluency Task. The Verbal Fluency Task will assess verbal functioning (Lezak et al., 2012). This task will provide insight into the participant's verbal ability such as lexical retrieval ability and lexical knowledge and assess their executive control. (Lezak et al., 2012). Through this assessment, we will be able to predict deficits in executive control or verbal ability. The Verbal Fluency Task will be comprised of two parts, the letter task, and the category task. In the letter task, participants will see four letters: B, M, D, and T. For each letter they will have 60 seconds to list as many words that begin with the letter. In the category task, participants will see four categories: Animals, Clothing, Musical Instruments, and Vegetables. Then they will be given 30 seconds to list items within each category.

Following the Verbal Fluency task is the Nonword Repetition Task. For this task, the participants will hear a string of one syllable nonwords and repeat each nonword to the best of their ability. There will be 16 trials of nonwords sentences that will get longer as the experiment continues. The task will begin with a string of three words and will increase up to a string of six words.

Finally, the first part of the experiment will conclude with the Lexical Decision Task. The purpose of this task will be to measure the participants' vocabulary knowledge in English. The task will consist of 60 words including nouns, adjectives, verbs, verb participles, and adverbs. Some of the words in the list will be English words and the rest will be nonwords that do not exist in English. In this task, the word will appear on the screen and participants will determine whether the world is a real English word or not. They will be instructed to press the 'A' key if they believe the word is a real English word and the 'L' key if they believe it is not a real English word.

The second part of the experiment will be a speech-in-speech recognition task. This task will consist of a headphone evaluation, practice trials, and the speech-in-speech recognition task. In the headphone evaluation, participants will listen to three pure tone sounds and determine which sound is the softest. If they can identify the softest sound, they will be allowed to move on to the practice trial and the actual task.

Written instructions will be provided, stating that participants will hear a female talker who will say a call sign, color, and number. The format will be, "Ready [call sign], go to [color] [number] now." The task will require the participants to identify the color and the number stated in each trial by selecting the button of the correct color and number. The possible colors included are: Red, Blue, Green, and White. The possible numbers included in this study are numbers one through eight.

**Table 1**

*Percentage of Words Correct with English Targets*

| Participants | Dutch Maskers | English Maskers | Expected |
|---|---|---|---|
| Participant #1 | 86% | 88% | No |
| Participant #2 | 86% | 78% | Yes |
| Participant #3 | 81% | 81% | No |
| Participant #4 | 84% | 82% | Yes |
| **Average** | 85% | 81% | Yes |

*Note.* Speech-in-speech recognition performance scores (percent correct) for 4 listeners in the presence of single-talker Dutch maskers and single-talker English maskers with an SNR of -6dB.

The column shows the percentage correct with Dutch maskers, the percentage correct with English maskers, and if those results were expected by stating either “Yes” or “No”.

**IV. Expected Results**

Before gathering data from all subjects, four pilot trials were conducted and analyzed. Table ` depicts performance scores (percent of color and number correctly identified) in each condition, English target with English Masker and English target with Dutch masker. The intensity level of the target sentence was 65dB and the masker sentence was 71dB, contributing to a Signal to Noise Ratio, SNR = -6dB. The results, however, were not as expected. Some participants received a linguistic release from masking, which can be shown by higher percentages in Dutch versus English masker and a “Yes” in the expected column. Yet, some participants either had higher percentages with the English masker or the same percentage regardless of the language the masker was in. Leading to the conclusion that the single-talker masker made the task too simple, causing there to not be a linguistic release from masking. However, if the difficulty of the experiment is increased by making the maskers two-talkers instead of a single talker, we expect there to be a greater linguistic release from masking. After increasing the difficulty level of the experiment, we expect a greater linguistic release from masking where all four monolingual participants will have a higher percentage correct score with the Dutch masker compared to the English masker. Leading to our hypothesizes that Dutch-English bilinguals with a higher English proficiency will score similarly to the monolingual participants. The similarities being that participants will have a higher percentage correct when maskers are in Dutch rather than English. We also hypothesize that Dutch-English bilinguals with a low English proficiency will score significantly lower with respects to the bilinguals with higher English proficiency, depicted in figure 2. If a correlation between proficiency level in second language and performance scores is found, the CRM task can be used as a speech-in-speech recognition task such as the CRM task as an alternative method to measuring language proficiency that does not rely on self-reported measures.

**Figure 2**

*Expected Percentage Correct for Dutch-English Bilinguals with English Targets*

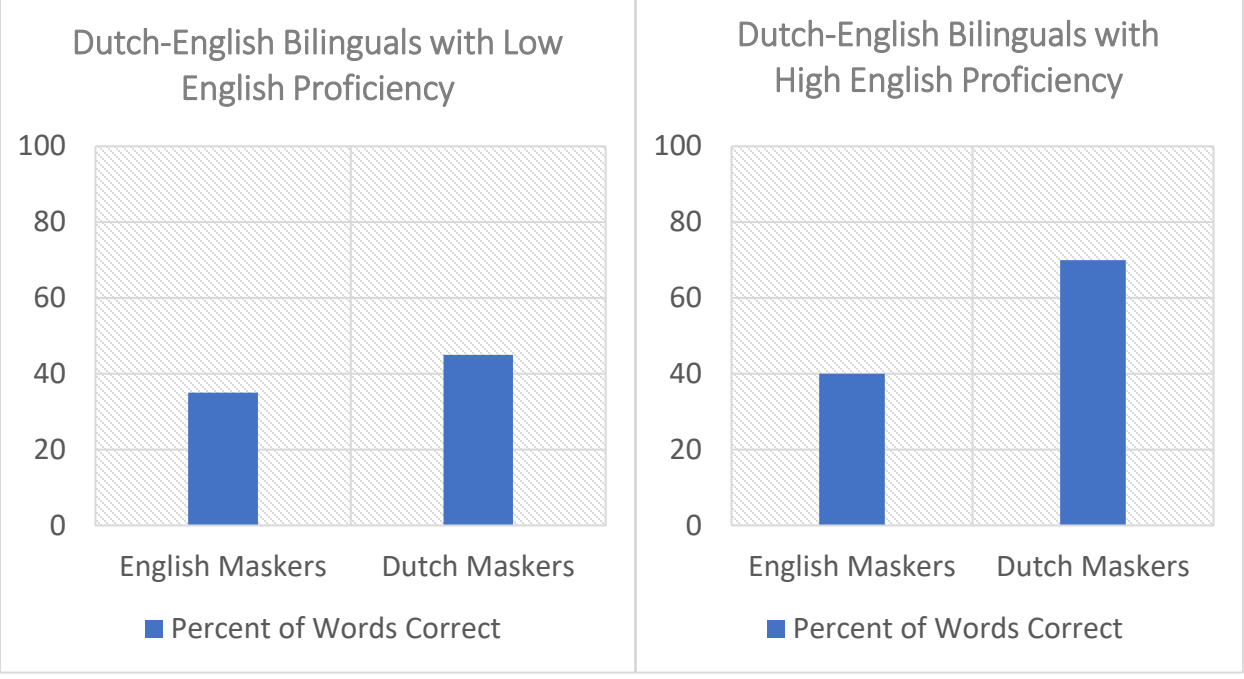

Fig. 2. Due to Linguistic Release from Masking we expect to see a greater performance score percentage when the masker is in Dutch with an English target than a masker in English with an English target. These boxes depict the expected results when Dutch-English bilinguals with varying levels of English proficiency complete a speech-in-speech recognition task with two maskers and an SNR = -6 dB.

## V. Discussion

Previous studies have demonstrated that listeners receive a benefit, or a linguistic release from masking, in speech-in-speech recognition task when the target speech and masker differ in language (Calandruccio et al., 2010; Brouwer et al., 2012). Other studies have looked at the effects second language proficiency has on recognition of speech in native and nonnative competing speech (Francis et al., 2018)? Through this study we will expand on past research to investigate if there is a correlation between the amount of LRM a participant receives and their level of proficiency in a second language. This question will be addressed by analyzing participants scores on a two-part experiment, utilizing a language proficiency task and a speech-in-speech recognition task.

From the piloted data, we found that monolingual English speakers had a slightly higher performance score when target and masker differed, showing that LRM exist. In future studies, we will increase the amount of LRM a participant experiences, by having the participants listen to a two-talker masker instead of a single-masker during the experiment. Where there will be two sentences of background noise along with the target sentence. With the goal of increasing the difficulty level of the experiment and showing a greater LRM. To conclude, we hope to expand our pool of participants to Dutch-English bilinguals to see if there is a positive correlation between higher English proficiency level and higher performance scores on the speech-in-speech recognition task leading to a greater LRM. Proving that the amount of language proficiency is a factor that contributes to the performance score of the speech-in-speech recognition task while also allowing the speech-in-speech recognition task to be an alternative method to measuring language proficiency.

## Acknowledgments

This research was supported by The Pennsylvania State University Ronald E. McNair Post-Baccalaureate Achievement Program and The Pennsylvania State University faculty.

**References**


Bench, J., Kowal, A., and Bamford, J. (1979). "The BKB (Bamford-KowalBench) sentence lists for partially-hearing children," Br. J. Audiol. 13, 108–112

Brouwer, S., Van Engen, K. J., Calandruccio, L., & Bradlow, A. R. (2012). Linguistic Contributions to speech-on-speech masking for native and non-native listeners: Language familiarity and semantic content. *Acoustical Society of America.*

Brungart, D. S. (2001). "Informational and energetic masking effects in the perception of two simultaneous talkers," J. Acoust. Soc. Am. 109(3), 1101–1109.

Calandruccio, L., Dhar, S., & Bradlow, A. R. (2010). Speech-on-speech masking with variable access to the linguistic content of the masker speech. *Acoustical Society of America,* 860-869.

Calandruccio, L., Brouwer, S., Van Engen, K. J., Dhar, S., and Bradlow, A. R. (2009). "Non-native speech perception in the presence of competing speech noise," Talk presented at the American Speech-Language-Hearing Association (ASHA), November 20, New Orleans, LA.

Danhauer, J, L., Crawford, S., Edgerton, B, J. (1984). English, Spanish, and Bilingual Speakers' Performance on a Nonsense Syllable Test (NST) Speech Sound Discrimination. *Journal of Speech and Hearing Disorders.*

Eddins, D, A., Liu, C. (2011). Psychometric properties of the coordinate response measure corpus with various types of background interference. *Acoustical Society of America,* 177-183.

Francis, A. L., Tigchelaar, L. J., Zhang, R., & Zekveld, A. A. (2018). Effects of Second Language Proficiency and Linguistic Uncertainty on Recognition of Speech. *Journal of Speech, Language, and Hearing Research.*

Friesen, D, C., Luo, L., Luk, G., Bialystok, E. (2015). Proficiency and control in verbal fluency performance across the lifespan for monolinguals and bilinguals. *Language, Cognition, and Neuroscience,* 238-250.

Harrington, M. (2006). The Lexical Decision Task as a Measure of L2 Lexical Proficiency. *EUROSLA.*

Medvedev, O, N., Sheppard C., Monetta, L., Taler, V., (2019). The BNT-38: Applying Rasch Analysis to Adapt the Boston Naming Test for Use with English and French Monolinguals and Bilinguals. *Journal of Speech, Language, and Hearing Research.*

Miller, M. K., Calandruccio, L., Buss, E., et al. (2019). Masked English Speech Recognition Performance in Younger and Older Spanish-English Bilingual and English Monolingual Children. *Journal of Speech, Language, and Hearing Research.*

Uchihara, T., Saito, K., (2016). Exploring the relationship between productive vocabulary knowledge and second language oral ability. *The Language Learning Journal,* 64-75

Wauters, L., Marquardt, T. P. (2017). Category, Letter, and Emotional Verbal Fluency in Spanish-English Bilingual Speakers: A Preliminary Report. *Archives of Clinical Neuropsychology*, 444-457.