Undergraduate Honors Thesis Collection

2020

The Role of Gender in the Perception of Lipreading

Emma Richards
Butler University

BUTLER UNIVERSITY HONORS PROGRAM

Honors Thesis Certification

Please type all information in this section:

Applicant Emma Grace Richards
(Name as it is to appear on diploma)

Thesis title The Role of Gender in the Perception of Lipreading

Intended date of commencement May 9th, 2020

Read, approved, and signed by:

Thesis adviser(s) ______________________ 4-13-2020
Date

______________________ ________
Date

Reader(s) Mary D. Gospel, Ph.D., CCC April 13, 2020
Date

______________________ ________
Date

Certified by ______________________ ________
Director, Honors Program Date

The Role of Gender in the Perception of Lipreading

A Thesis

Presented to the Department of Communication Sciences and Disorders

College of Communication

and

The Honors Program

of

Butler University

In Partial Fulfillment

of the Requirements for Graduation Honors

Emma Grace Richards

May 8, 2020

Abstract

The experiment examines the gender differences in lipreading ability. We tested 25 fluent English-speaking adults with normal hearing and vision. Each participant was asked to watch four videos testing their lipreading ability with matched and nonmatched gendered speakers. The participants were asked to transcribe 40 sentences, which were scored on the number of target words correct. It was found that participants who watch the videos with matched genders perceived more target words than in the mismatched gender videos. The results could explain previous inconsistencies in earlier research of whether ability to lipread is improved by the gender of the speaker.

The Role of Gender in the Perception of Lipreading

Every day, we use lipreading to help better perceive the auditory signals of human speech but most of the time we do not realize how important this visual aspect of speech is. Lipreading is often used by people with hearing-impairment to compensate for their hearing loss (Tye-Murray, Spehar, Myerson, Hale, & Sommers, 2014). Individuals with normal hearing also use lipreading to assist the perception of speech, especially in noisy surroundings (Summerfield,1992). A majority of the time, the contribution of vision in speech perception goes unnoticed unless the visual image and auditory signal do not match. A famous example of this phenomenon is the McGurk effect, where an auditory illusion originates from lip movements, but does not match the auditory signal (McGurk & MacDonald, 1976). People with normal hearing also rely on visual cues as well, such as body language and facial expression in addition to lipreading. Previous studies have shown variability in lipreading ability among age groups, but results have been inconsistent with respect to gender (Tye-Murray, Spehar, Myerson, Hale, & Sommers, 2016). The current experiment examines if listeners perceive speech more accurately if the speaker is the same gender as the listener.

Various studies have investigated connections between lipreading and age, or the effect lip reading has on areas of the brain in men and women. However, the effect of matched gender between speaker and listener on lipreading abilities has not yet been researched. Given lipreading often aids in the perception of speech for individuals who have normal hearing or impaired hearing, it is sensible to determine any gender-related differences in lipreading ability. Because lipreading is used to assist the perception of speech in noisy environments, such as classrooms or crowded places, young adults are the main focus to examine how the gender of a speaker may impact the ability to lipread. The study provides more information on the link between lipreading ability and gender.

In typical experiments, the auditory and visual speech signals are generated by the same talker. Green, Kuhl, Meltzoff, and Stevens (1991) examined whether a discrepancy in the gender of the talker between the auditory and visual signals would influence the magnitude of the McGurk effect. Forty-four undergraduate students were given a visual (V), auditory (A), or Audio-Visual (AV) presentation of the male or female “talker” producing the syllables /ba/ and /ga/. For the AV stimuli, there were four types of stimuli: two where the A and V signals originated from talkers of the same gender (the control). The other two types of stimuli had talkers, both visually and auditorily, with different genders. A male talker's voice was dubbed onto a videotape containing a female talker's face, and vice versa. The results indicated that participants’ mechanism for integrating speech information from auditory and visual modalities is not disrupted when there is a gender incompatibility between the two modalities. The findings are compatible with the theoretical notion that information about voice characteristics of the talker are extracted and used to normalize the speech signal at an early stage of phonetic processing, prior to the integration of the auditory and the visual information.

Lipreading skills were first examined in older adults due to a possible connection between age-related hearing loss and lipreading skills. The rate of age-related decrease in visual (V-only) speech recognition performance is often greater than that of audiovisual (AV) performance (Tye-Murray, Spehar, Myerson, Hale, & Sommers, 2016). To investigate the impact age has on AV and V-only speech recognition performance, participants completed the Build-a-Sentence test (BAS), a sentence-level test that uses a set of target words that are inserted into the same sentence structure. The list of possible target words was shown on a computer monitor following each sentence, and participants were asked to select their responses by repeating them aloud. The stimuli were based on audiovisual recordings of the same female talker reading the lists of sentences into the camera. Four levels of blurriness were used to distort the Visual Only stimulus. The results revealed the rate of age-related decrease in Visual Only performance was more than twice the decrease in AV condition. This suggests the benefit that comes from being able to see and hear a talker remains constant throughout adulthood, and variations in AV advantage are driven by age-related changes in visual-only and auditory-only speech recognition.

Another study revealed that age-related deficits in many sensory and cognitive abilities, including lip reading abilities, are greater for men than for women (Tye-Murray, Sommers, and Spehar, 2007). Researchers tested whether age-related changes in lipreading abilities are similar for men and women. The stimuli consisted of a single female talker with a general American dialect saying consonants, words, and sentences that contained familiar vocabulary for the participants. Across consonants, words, and sentences, the younger group did significantly better than the older group, but there was no difference between gender. The findings suggested that whereas both older adult males and females performed less well than their younger counterparts, the differences between younger and older adults did not vary due to gender.

A recent study investigating whether the primary auditory cortex (PAC) is involved in silent lipreading showed men and women had different networks for processing lip reading. (Ruytjens, Albers, van Dijk, & Willemsen, 2007). The experiment tested which areas of the temporal lobe process visible speech, with a focus on the PAC. Eighteen healthy volunteers participated in this study, half male and half female. Each subject was scanned while either watching a movie with a speaker silently articulating numbers (lipreading condition) or watching a static image of the same speaker (baseline condition). Subjects were instructed to repeat internally the number seen or the number '1'. Compared to the baseline condition, silent lipreading activated temporal areas in both hemispheres with the largest activation clusters in the left hemisphere. The data showed that only women significantly activated the left PAC during visual speech processing. These results suggest that a different neural network is involved in lipreading between the two genders.

It has been shown that understanding lipreading is easier in general when the speaker is female (Daly, Bench, & Chappell, 1996). The study examined differences between men and women in eight ways speech production varies from individual to individual. These variables include: eyebrow movement, lip rounding, mouth area, mouth height, mouth width, speed of speech, teeth visibility, and visible tongue movement. The 48 participants, aged between 18 and 50 years, were English speakers with normal hearing and eyesight. The participants viewed videos of male and female talkers over two 45-minute sessions. There were pauses after each sentence which allowed the participants to repeat what they perceived. Responses were recorded by a researcher sitting behind them. It was found female talkers used more teeth visibility, more lip rounding, larger mouth height, larger mouth area, and slower speed of speech than males. All of these variables have shown to improve the visual intelligibility of talkers by both male and female listeners. These outcomes provide evidence of a visual gender dialect.

Various studies have investigated connections between lipreading and age or the effect lipreading has on the areas of the brain in males and females. However, the effect speakers with matched gender have on lipreading abilities has not be researched. This led to the formation of the research question: Will gender of a speaker affect the lipreading capabilities of the person receiving the visual stimulus if the receiver is the same gender as the speaker? Considering lipreading often aids in the perception of speech for individuals who have normal hearing or impaired hearing, it is sensible to determine the differences in lipreading ability among the genders. Since lipreading is used to assist the perception of speech in noisy environments, such as classrooms or crowded places; young adults are the main focus to examine how the gender of a speaker may impact the ability to lipread. Based on previous studies on the effect of lipreading on the brain and genders, it is expected that young adults who receive the stimulus from a matched-gendered speaker will exhibit better lipreading abilities than young adults who receive the stimulus from a mismatched gendered speaker.

To test for potential confounding variable of memory and language skills, two control tests will be run to test participants' processing speed and working memory. The test of processing speed comes from Kail and Miller (2006), who examined whether processing speed in the language domain develops in children at the same rate as global processing speed. One hundred and sixteen children aged either 9 or 14 years old were tested. Of the 10 different speeded tasks, one of the language tasks pertained to differentiating between correct and incorrect grammar. Results showed processing speed was faster for language tasks at age 9, but not at age 14. The results revealed evidence of modest continuity in processing speed across the 5-year span. Since the participants of the current study are college-aged students, it is possible to assume these participants' processing speed will be faster.

The test of memory was used because large individual differences in spoken word recognition performance have been found in deaf children after cochlear implantation based on verbal working memory abilities. Recently, it was reported that simple forward digit span measures of verbal working memory were significantly correlated with spoken word recognition scores, even after potentially confounding variables were statistically controlled for (Pisoni and

Cleary, 2003). Using the test lists and procedures from the WISC III (Wechsler, 1991), forward and backward auditory digit spans were obtained from four groups of 8- and 9-yr-old children with cochlear implants, testing a total of 176 children with cochlear implants. The WISC-III digit span memory task requires the child to repeat back a list of digits that are spoken live-voice by an experimenter at a rate of approximately one digit per second. For the “digits-forward” section of the task, the child was required to simply repeat the list as heard. For the “digits-backward” section of the task, the child was told to “say the list backward.” The results suggest that perhaps as much as 20% of the currently unexplained variance in spoken word recognition scores may be independently accounted for by individual differences in cognitive factors related to the speed and efficiency with which phonological and lexical representations of spoken words are maintained in and retrieved from working memory. A smaller percentage, perhaps about 7% of the currently unexplained variance in spoken word recognition scores, may be accounted for in terms of working memory capacity.

Methods

Participants:

To answer the question of whether the gender of a speaker affects the lipreading capabilities of the person receiving the visual stimulus, the researchers performed the following experiment. Participants consisted of 20 women and 5 men – between 18 and 25 years old –all of whom were students currently attending Butler University. An email advertisement was sent to students on the Honors Listserv and the Communication Sciences and Disorders (CSD) Listserv to ensure a varied group of the population signed up to participate in the study. The incentive of a $5 Starbucks gift card was offered to any interested participants in a bid to increase the number of students signing up.

After a student showed interest in participating in the experiment, they were sent the link to sign up for a testing time. On the scheduled day, the participants filled out an informed consent contract and a background questionnaire. The background questionnaire collected data about the student's vision, hearing, and language history. The answers from these background surveys show every participant was a native English-speaker and all had 20/20 vision or used corrective lenses. All but two participants had normal hearing, but even the participants with hearing loss were able to hear the stimuli with no issue. Once these forms were completed, the participant began testing.

Materials:

Before testing began, videos were created using iMovie software. A male and a female were recruited to be model speakers in the videos. Both models were compensated with a $5 Starbucks gift card for their involvement. Each model was given two different lists of 10 randomized utterances, one for each video they were in. Across all three videos, there were 40 randomized utterances, all containing three different nouns that served as the target words for each utterance. In total, there were 120 target words, placed in one of two sentence structures. There was a total of four video conditions: Videos One and Two had a male speaker reading the utterances, with his mouth visible. In Video One, the audio was intact, and the male model spoke in a low-pitched voice (AV condition), while Video Two had no audio (V-only condition). Videos Three and Four had a female speaker reading the utterances, with her mouth visible. In Video Three, the audio was intact, and the female model spoke in a low-pitched voice (AV condition), while Video Four had no audio (V-only condition). The AV conditions were used as a controlled variable. Once the videos were recorded, the order of the four segments was randomly edited together into 24 videos. Another laptop with Microsoft Word was used to record participants' responses for the 40 utterances they heard.

Both processing speed and memory span were tested as cognitive controls for the study. The processing speed task tested how quickly participants were able to judge grammatical sense (Kail & Miller, 2006). The researcher used the software Psyscope to create the stimuli that would present 29 sentences auditorily. Fourteen of the sentences were grammatically correct, while 15 sentences had grammatical errors: one set of five had the subject and the verb disagree, the next set of five had the incorrect word order, and the last set of five omitted a preposition. For the working memory task, a digit span test was used. This task required participants to recall a list of digits in correct serial order, or forward digit span, given to them by the researcher (Wechsler, 1991). When a list was repeated correctly, a new list was presented with an additional digit added to each new list. The researcher recorded the number of lists the participant remembered correctly. Forward digit span had a possible total of 16 points, and backward digit span had a possible total of 14 points. The results for these cognitive control tests helped determine if processing speed or working memory were related to any observed differences found in lip reading ability.

Procedure:

On the day they were scheduled for testing, the participants filled out the informed consent contract and a background questionnaire. After the necessary forms were filled out, participants sat at a table in the silent testing room, Fairbanks 013. The room was kept completely silent to prevent interference from background noise. On the table, the two laptops were set up, the first was programed with the Psyscope software and prepped to play the video stimuli. The second laptop displayed a Microsoft Word document divided into four sections, one for each video the participants would watch.

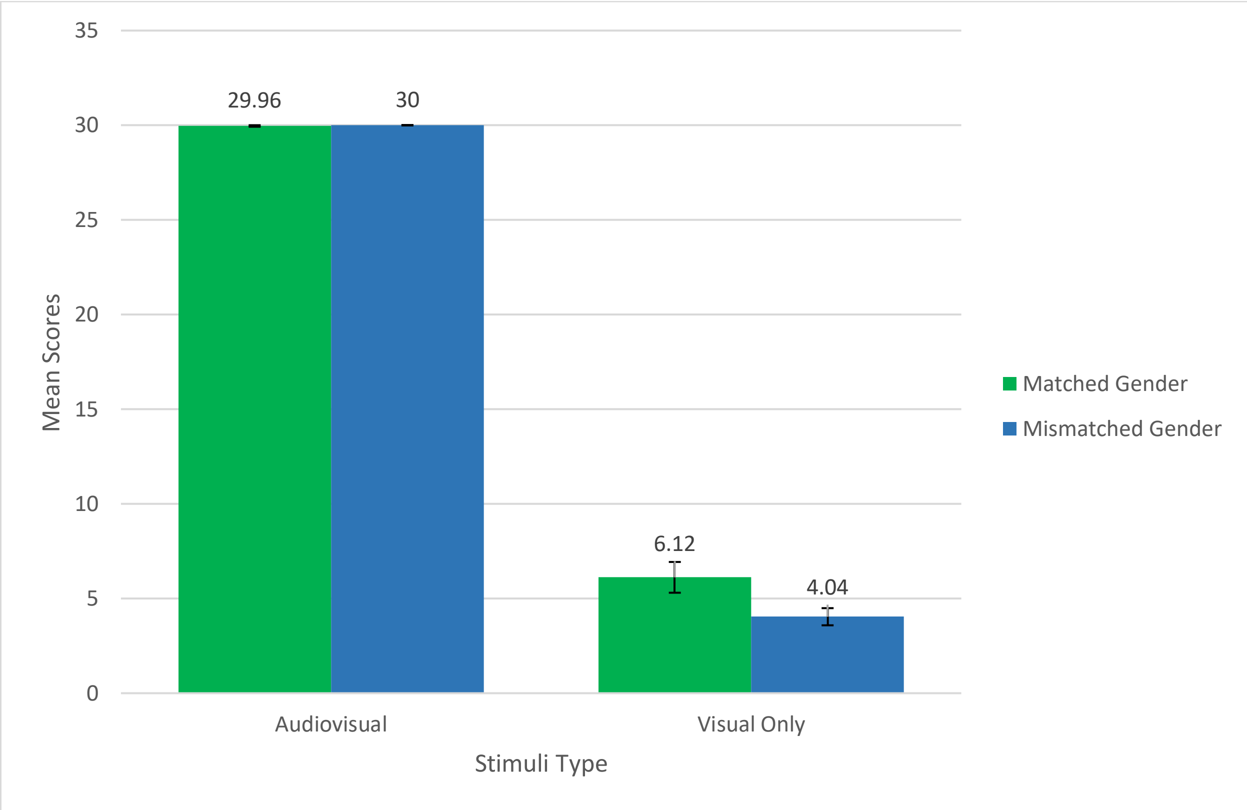
To begin the study, the participants completed the processing speed task and working memory task to determine any possible individual difference in these two skills. The participants would click “run” to begin the Psyscope software. Participants were instructed to listen to the sentences and determine if the sentence was grammatically “correct” or “incorrect”. Participants’ accuracy and reaction time were recorded. To record their answers, participants were told to click the left arrow button if they deemed the sentence “incorrect”, and the right arrow button if they deemed the sentence “correct.” For the digit span testing, participants were given a list of numbers and asked to repeat them in correct serial order, also known as forward digit span. The researchers recorded the number of lists the participant remembered correctly. The participant was then told to repeat the lists of digits using backward digit span, which required the participant to repeat the numbers in the reverse order of the list presented by the examiner. Like the forward digit span, for each correctly recalled list, the participant was tested on a new list with an additional digit added. Participants were again scored on the number of lists they remembered.

Next, the participants were instructed to watch the 11-minute video and told to write down each utterance they perceived. The participants were told to write down their best guess of what they heard or saw the model speakers say. Participants used headphones to listen to the stimuli to ensure no background noise was present. Using the second laptop equipped with Microsoft Word, the participants recorded what they perceived for the 40 utterances. All 40 utterances were presented consecutively to the participant. Once the video began and the first model, with or without sound, said one utterance, participants were given a 15-second pause to write what they believed they observed. The same process occurred for the remaining 39 utterances. Once the last utterance concluded, the participant was asked post-test questions of whether or not they knew of the male and female model used in the video. If a participant knew one of the models, they were asked what their relationship was to the model, using this scale: Met Once, Acquaintance, Friend, Best Friend, or Family. Once the survey was completed, the participant was told the experiment had finished and they were compensated with their gift card. Each testing period took an average of 25 minutes to complete.

The participants' data were divided into two groups, one consisting of all females and one consisting of all males. The participants were scored on the three tests: processing speed, working memory, and lipreading ability. To assess the processing speed, the participants were scored on the number of sentences they correctly labeled as grammatically "correct" and "incorrect, as well as their average reaction time for grammatically correct sentences and grammatically incorrect sentences. The working memory test added the scores from the 14 forward and 15 backward digit span trials together, creating the participant's digit span score. To evaluate lipreading ability, the participants were scored on the number of target words they perceive correctly out of the 120 possible target words. These scores were then averaged together with the other scores from their gender group to collect the group's data.

For each female participant, the total number of correct target words from the two female-talker conditions were averaged together with each corresponding set of the men's total number of correct target words from the two male-talker conditions, creating the average number of correct target words for the Matched Gender Group. Each of the women's total of number of correct target words from the two male-talker videos was averaged with the men's number of the total number of correct target words from the two female-talker videos, creating the data from the Mismatched Gender Group. The data analysis for the average number of correct target words for the Matched and Mismatched Gender Groups was carried out by performing an ANOVA test to determine the significance of the data. The results of the Audiovisual and Visual Only condition were also analyzed. It was predicted the Matched Gender Group would have a better average number of correct target words than the Mismatched Gender Group.

Results



At the completion of the study, it was hypothesized that the matched gender group would perform better than the mismatched gender group in the Visual Only condition. A paired-sample t-test revealed a statistically significant difference between the V-Only matched gender and mismatched gender group, $t_{(25)} = 2.972$, $p = .007$. As shown in the figure, there was not a significant difference in the data collected in the Audiovisual condition $t_{(25)} = -1.000$, $p = .327$), which is expected of the control variable. There was a significant correlation between the participants' scores for the Visual Only stimuli: those who recorded a higher number of correct target words for the matched gender video also had a higher number of correct target words for the mismatched gender group. The same correlation appeared for participants who recorded a lower number of correct target words for both the matched and mismatched gender video.

In regard to any confounding variables from the cognitive controls, there were no correlations between grammar skills or working memory skills and V-Only accuracy, with the exception of a moderately strong, negative correlation found between the matched gender groups and having a slower reaction time for the incorrect grammar phrases during the processing speed task. This relationship did not impact the data collected during the rest of the study. It is also worth noting the participants' relationship with the male and female models for the videos was not significantly correlated with participants' results.

Discussion

Since lipreading is used to assist the perception of speech in noisy environments, such as classrooms or crowded places, young adults are the main focus to examine how the gender of a speaker may impact the ability to lipread. The researcher set out to examine whether the same gendered speaker and recipient would improve lipreading ability and if the subjects' lipreading ability would decline when the speaker is a different gender. It was practical to look at the impact gender had on the subjects' lipreading ability since the visual aspects of speech are necessary for comprehension. This study provided more information on the link between lipreading ability and gender. As the results showed, the participants' lipreading accuracy was influenced by the gender of the speaker. The improvement seen in the matched gender groups could be due to how Visual Only signals are perceived by neural networks within the brain. Males and females use different networks to perceive lipreading signals, seen in the female brain activating the left PCA during with silent lipreading. Since female listeners lipread more accurately with a female speaker, and male listeners lipread more accurately with a male speaker, it might pertain to how to neural pathways work differently. For each gender, when it comes to the visual aspects of speech, it could be easier to understand what the speaker is saying when the speaker and listener share a gender. This could be caused by the different tendencies men and woman physically use while speaking, making it easier for matched genders to perceive what is being said. The results of this study could explain the previous inconsistencies in past research whether ability to lipread is improved by the matched gender of the speaker and listener. As stated previously, the further analyses on the background survey and cognitive control tasks determined grammar, memory, and familiarity did not play a role in the results of this study.

In future studies, researchers should strive to obtain a sample size of no less than 20 participants for both the male and female groups. Since there were not 20 male participants in this study, the matched and mismatched data of this group could possibly be skewed. Researchers should also try to recreate the video stimuli in a studio where background noise could be controlled, as well as lighting and sound quality. It would also be helpful to include several female and several male speakers to ensure the results were not simply the effect of the two particular speakers included in this study. Another possible improvement would be to include a video of an individual who identifies and presents as nonbinary. This would aid in determining if the gender of the speaker truly plays a role in an individual's ability to lipread.

As for further data collection, this study focused on individuals with normal hearing in quiet environments, but lipreading is often vital in noisy situations, so a condition with background noise should be added in follow-up studies. In addition, the relationship between the gender of the speaker and listener and lipreading ability should be assessed in participants with hearing impairments, not just normal hearing. Future researchers could also test different age groups to see if the relationship between gender and lipreading ability fluctuates with age. In the future, additional studies should be conducted to look into the specific brain activity differences in males and females when they are stimulated by silent lipreading to determine why listeners with matched gender to the speaker performed better than when the genders did not match. Studies pertaining to brain activity in both males and females during silent speech perception could close the gap in the current research.

References

Daly, N., Bench, J., & Chappell, H. (1996). Gender differences in visual intelligibility. *Journal of the Academy of Rehabilitative Audiology, 29*, 27-40.

Feld, J. E., & Sommers, M. S. (2009). Lipreading, processing speed, and working memory in younger and older adults. *Journal of Speech, Language, and Hearing Research, 52,* 1555-1565.

Green, K., Kuhl, P., Meltzoff, a., & Stevens, E. (1991). Integrating speech information across talkers, gender, and sensory modality: Female faces and male voices in the McGurk effect. *Perception & Psychophysics, 50,* 524-536.

Kail, R. V., & Miller, C. A. (2006). Developmental change in processing speed: Domain specificity and stability during childhood and adolescence. *Journal of Cognition and Development, 7*(1), 119-137.

Nicholls, G. H., & Ling, D. (1982). Cued speech and the reception of spoken language. *Journal of Speech & Hearing Research, 25*, 262-269.

Pisoni, D. and Cleary, M. (2003). Measures of working memory span and verbal rehearsal speed in deaf children after cochlear implantation. *Ear Hear*, 24, 106S–120S.

Ruytjens, L., Albers, Frans., van Dijk, P., & Willemsen, A. (2007). Activation in primary auditory cortex during silent lipreading is determined by sex. *Audiology & Neurotology*, *12*, 371-377.

Summerfield, Q. (1992). Lipreading and audio-visual speech perception. *Philosophical Transactions: Biological Sciences,335*(1273), 71-78.

Tye-Murray, N., Sommers, M. S., & Spehar, B. (2007). The effects of age and gender on lipreading abilities. *Journal of the American Academy of Audiology, 18,* 883-892.

Tye-Murray, N., Sommers, M., Spehar, B., Myerson, J., Hale, S., & Rose, N. S. (2008). Auditory-visual discourse comprehension by older and young adults in favorable and unfavorable conditions. *International Journal of Audiology, 47*, S31-S37.

Tye-Murray, N., Spehar, B., Myerson, J., Hale, S., & Sommers, M. (2016). Lipreading and audiovisual speech recognition across the adult lifespan: Implications for audiovisual integration. *Psychology and Aging, 31*(4), 380-389.

Tye-Murray, N., Spehar, B., Myerson, J., Hale, S., & Sommers, M. (2014) Lipreading in school-age children: the roles of age, hearing status, and cognitive ability." *Journal of Speech, Language, and Hearing Research*, 57(2), 556–565.

Acknowledgement

I would like to thank the Butler University Honors Thesis Grant for funding my research. I would also like to thank Dr. Bergeson for advising me during the research process.

Appendix:

Video 1 (Male, Audiovisual)

1. The **friends** in the **car** saw a **mouse**.
2. A **boy** ran with the **cat** and **book**.
3. **Mom** and **duck** saw the **people**.
4. The **mouse** and **whale** saw the **cats**.
5. The **girl** and **ball** ate the **food**.
6. The **bees** and **girls** played the **game**.
7. The **dog** watched the **snail** and **boy.**
8. The **book** was **wide** and **green**.
9. A **woman** and **deer** read the **book**.
10. A **baby** was **chubby** and **sleepy**

Video 2 (Male, Visual-Only)

1. **Man** watches the **dog** chase the **bow**.
2. A **lion** and **giraffe** chased the **bear**.
3. The **whale** and **mouse** swam with a **cow**.
4. A **book** and **chair** sat on the **table**.
5. **She** threw the **apple** and **cake**.
6. **They** watch the **cat** chase the **horse**.
7. A **truck** was **black** and **tiny**.
8. **He** drove to the **city** and the **park**.
9. The **pencil** and **garbage** fell on the **ground**.
10. A **bird** and **phone** flew in the **sky**.

Video 3 (Female, Audiovisual)

1. **Mom** chased the **boy** and **goat**.
2. The **dog** was **sleepy** and **cold**.
3. The **tree** and **beaver** are **green**.
4. **Bees** and **owls** flew in the **woods**.
5. A **baby** ate the **peas** and **carrots**.
6. The **computer** is **sleek** and **slow**.
7. A **bird** played the **guitar** and **beetle**.
8. **Mom** and **dog** eat the **food**.
9. **Bear** and **sister** watch a **movie**.
10. A **wolf** and **cricket** run **fast**.

Video 4 (Female, Visual-Only)

1. The **cricket** followed the **cat** and **pig**.
2. A **child** ran from the **truck** and **farm**.
3. The **mouse** and **baby** ran to the **hotel**.
4. A **clown** and **toy** scared the **kids**.
5. The **movie** was **scary** and **funny**.
6. The **woman** and **pickle** chased the **ear**.
7. A **cow** was chased by the **clown** and **badger**.
8. A **wolf** was **grey** and **pink**.
9. The **shrimp** was **small** and **crunchy**.
10. The **deer** and **fox** were **angry**.

*These were the sentences used in each of the four videos. The bolded words are the three target words used to score participants on their lipreading ability.