# The Impact of Visual Input on the Ability of Bilateral and Bimodal Cochlear Implant Users to Accurately Perceive Words and Phonemes in Experimental Phrases

by

Cimarron Ludwig

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved April 2015 by the
Graduate Supervisor Committee:

Julie Liss, Chair
Michael Dorman
Tamiko Azuma

ARIZONA STATE UNIVERSITY

MAY 2015

ABSTRACT

A multitude of individuals across the globe suffer from hearing loss and that number continues to grow. Cochlear implants, while having limitations, provide electrical input for users enabling them to “hear” and more fully interact socially with their environment. There has been a clinical shift to the bilateral placement of implants in both ears and to bimodal placement of a hearing aid in the contralateral ear if residual hearing is present. However, there is potentially more to subsequent speech perception for bilateral and bimodal cochlear implant users than the electric and acoustic input being received via these modalities. For normal listeners vision plays a role and Rosenblum (2005) points out it is a key feature of an integrated perceptual process. Logically, cochlear implant users should also benefit from integrated visual input. The question is how exactly does vision provide benefit to bilateral and bimodal users. Eight (8) bilateral and 5 bimodal participants received randomized experimental phrases previously generated by Liss et al. (1998) in auditory and audiovisual conditions. The participants recorded their perception of the input. Data were consequently analyzed for percent words correct, consonant errors, and lexical boundary error types. Overall, vision was found to improve speech perception for bilateral and bimodal cochlear implant participants. Each group experienced a significant increase in percent words correct when visual input was added. With vision bilateral participants reduced consonant place errors and demonstrated increased use of syllabic stress cues used in lexical segmentation. Therefore, results suggest vision might provide perceptual benefits for bilateral cochlear implant users by granting access to place information and by augmenting cues for syllabic stress in the absence of acoustic input. On the other hand vision did not provide the bimodal participants significantly increased access to place and

stress cues. Therefore the exact mechanism by which bimodal implant users improved speech perception with the addition of vision is unknown. These results point to the complexities of audiovisual integration during speech perception and the need for continued research regarding the benefit vision provides to bilateral and bimodal cochlear implant users.

# DEDICATION

This master's thesis is dedicated to my wife, Jennifer, and my three daughters, Annabella, Amalie, and Brielle. Thank you for being so wonderful, especially during the past two years.

ACKNOWLEDGEMENTS

Thank you to the subjects that participated in this study, Shuai Wang for initially helping me start my project, Denae Morris for reliability coding, Sara Cook Rene Utianski, and Saul Frankford for helping with data analysis, and my committee, Julie Liss, Michael Dorman, and Tamiko Azuma, for all of your support and guidance.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

# OPERATIONAL DEFINITIONS

## Lexical Boundary Terms Defined:

### Words Correct

The total number of words counted as correct. "A" and "the" are used interchangeably. Adding an –s or –ed to the end of the target word is counted as correct unless the total number of syllables is altered.

### Percent Words Correct

The percent of the total number of words presented within experimental phrases that were accurately perceived.

### Insert Strong Error Type

A lexical boundary was inaccurately inserted before a strong syllable.

### Insert Weak Error Type

A lexical boundary was inaccurately inserted before a weak syllable.

### Delete Strong Error Type

A lexical boundary was inaccurately deleted before a strong syllable.

### Delete Weak Error Type

A lexical boundary was inaccurately deleted before a weak syllable.

### Add Syllable Error

Syllable(s) inaccurately added to the experimental phrase by the listener.

### Delete Syllable

Syllable(s) inaccurately deleted from the experimental phrase by the listener.

### Phonemic Resemblance Error Types:

### Error Types:

- Vowel
- Consonantal place
- Consonantal manner error
- Consonantal voice error
- Phoneme inserted in subject response
- Phoneme deletion in subject response

# I. INTRODUCTION

It may be posited that, to fully interact with our environment, all of our senses must be intact. O'Donoghue (2013) specifically addresses hearing loss. He writes "deafness impairs quality of life by relentlessly dismantling the machinery of human communication" (p. 1190). But the environment is rife with multisensory converging and complimentary cues, and the human brain leverages usable senses to compensate for those that are lost. The utility of cochlear implants in addressing deafness is a testament to these facts. The relatively crude electrical impulses delivered by cochlear implants meld with all other sources of information in the brain and the environment to allow users to "hear" and communicate. Because of this interaction among the sensory input in the environment, the way this input is transmitted, and the resulting speech perception, there is much to be learned about the way cochlear implants restore communication abilities.

A particularly interesting question has to do with the recent clinical shift to providing cochlear implants to both ears, or bilaterally, rather than to just one. This shift represents a response to research that has shown bilateral benefits (Dorman & Gifford 2010, van Hoesel 2012). van Hoesel (2012) states the benefit of bilateral cochlear implants relates to gains made in spatial hearing ability. Specifically, the large effect of the head shadow can be used advantageously at either ear and offer improvements in localization. In day-to-day listening, especially in environments with multi-talkers and non-speech noise this can be of significant benefit in affording the listener the ability to direct his attention to the auditory signal of choice.

In persons with one deaf ear and the other with some residual hearing, the trend now is to implant the deaf ear and optimally amplify hearing in the better ear. These patients are known as "bimodal" as they have electrical stimulation from the cochlear implant and acoustic stimulation to the better ear (Dorman & Gifford 2010, van Hoesel 2012). Subjects fitted with a hearing aid in the contralateral ear benefit from the addition of low-frequency information that is generally poorly coded by the implant. Such low-frequency information can offer significant advantages in terms of pitch discrimination and speech perceptibility (van Hoesel 2012). Low-frequency input contributes power to speech and the pitch information provides a benefit in terms of rhythm and prosodic contour.

The question becomes what other inputs are available to the listener. For normal listeners vision plays a role in the perception of speech. Rosenblum (2008) points out "virtually any time we are speaking with someone in person, we use information from seeing movement of their lips, teeth, tongue, and non-mouth facial features, and likely have been doing so all our lives" (p.405). Rosenblum (2005) also notes that visual information is not just an additional piece of information in speech perception, but rather a key feature of a multisensory integrated process. Grant, Tufts, & Greenberg (2007) state during face-to-face communicative events, the listener extracts and integrates information from both the acoustic and optic speech signals. Integration takes place within the auditory modality, across the acoustic frequency spectrum, and across sensory modalities, including the optic and acoustic signals. It would follow that cochlear implant users would benefit from visual input as well.

This study aimed to determine what kind of benefit to speech perception is provided to bilateral and bimodal cochlear implant users with vision. To make this determination, specific measures were examined. These measures included: a) percent words correct b) consonant errors and c) lexical boundary errors. These were chosen because of their relevance to overall speech perceptibility. Importantly, the analysis of errors allows us to identify potential sources underlying the overall intelligibility scores.

Word recognition is paramount in speech recognition. By recognizing individual words, individuals are able to determine the parts that make up the whole. The measure of percent words correct provides a straightforward metric indicating how easy or difficult speech is to perceive. In this study, it was hypothesized that bilateral and bimodal participants will benefit from the addition of visual information and this will result in an increased percent words correct for both groups. But finding that vision helps does not inform the question of how vision helps. Figuring out clues to the "how" requires a look at both what listeners were able to get right and what they got wrong.

One error analysis measure that was used to address the question of how vision helps is the examination of consonant errors. Specifically, what is the phonemic relationship between a misperceived word and the correct target with regard to place, manner, and voice?

This is important when one considers the potential impact of hearing loss on the perception of consonants. In a study conducted by Grant, Tufts, & Greenberg (2007) hearing-impaired listeners were unable to extract as much information on a consonant recognition task in auditory-only conditions when compared to normal-hearing listeners. This comes as no surprise considering the degradation of the auditory signal hearing-impaired listeners experience.

Given the fact consonants are generally perceived as discrete linguistic segments that can be defined as combinations of distinctive binary features it is possible to examine consonants and take a closer look for error patterns. The behavior of the articulators in terms of whether airflow is continuous or periodic, the manner in which the articulators produce restriction, and the place within the oral cavity where the constriction is produced (Svisrky et al. 2011) can all be accessed.

Of the three features, manner, voice, and place, the most visible by far is place of articulation. Manner and voicing mostly involve unseen structures including the velopharyngeal port and the larynx and generally are not transmitted well visually or require a trained eye. However, Grant, Tufts, & Greenberg (2007) note that place cues are transmitted well visually for hearing-impaired subjects, and these subjects performed at the same level as normal-hearing subjects during visual-only trials. They concluded that hearing-impaired listeners were able to overcome deficits with regard to auditory place information by using visual cues.

It stands to reason that bilateral and bimodal cochlear implant users might also exhibit improvement in parsing out consonantal place cues during speech perception tasks with the addition of vision as well. It is therefore hypothesized that if bilateral and bimodal cochlear implant participants are using the visual information available, their perceptual errors will share the same place of articulation as the target consonants. And if this is the case, we can conclude that part of the increase in percent words correct with the addition of vision can be attributed to the visibility of place of articulation in bilateral and bimodal cochlear implant users.

The second error analysis measure for this study focuses on the fact speech perception is not just about getting the phonemes correct. Rather, it means taking the continuous acoustic signal and figuring out where one word ends and another begins. When speech is clear and easy to understand this is a simple task because the string of words is actually heard. But when speech input is degraded, as is the case for cochlear implant users, listeners often have to work to find the boundaries between words.

Therefore, lexical segmentation measures were examined based on a model for speech perception called Metrical Segmentation Strategy (Cutler and Butterfield 1992; Cutler and Norris 1988; Liss et al. 1998) that indicates the ability to lexically segment connected speech is an important component in the perceptibility of speech. According to this model and subsequent support, listeners use syllabic strength to parse the acoustic stream they are receiving into recognizable words. In particular, listeners pay attention to the juxtaposition of strong and weak syllables.

According to Liss et al. (1998) Cutler and Carter (1987) have found support for this model in the statistical probabilities of syllabic strength within the English language. They found word-initial syllables and single-syllable words are associated with strong syllables. Furthermore, according to Liss et al. (1998) Cutler and Carter (1987) reported that these words are most commonly open-class words such as nouns, adjectives, and verbs whereas weak syllables are most commonly associated with second syllable placement or closed-class words such as pronouns and articles.

Cutler and Carter (1987) approximated 90% of English content words found in conversational speech have a strongly stressed syllable at the beginning of the word. Metrical Segmentation Strategy thereby postulates that communicative members should treat the strongly stressed syllable as though it was the beginning of the word (Cutler & Norris 1988). Participants paying attention to lexical cues and implementing lexical segmentation strategies will primarily have more insert before strong than insert before weak errors.

In testing, it was found that the pattern of errors committed by listeners was of a certain character if they recognized strong syllables at word onset. Specifically, listeners were predisposed to insert lexical boundaries before strong syllables. It is also predicted that there would be a greater likelihood of deleting lexical boundaries before weak syllables (Cutler and Butterfield 1992; Liss et al 1998). According to Smith et al. (1989) this pattern of results suggests that listeners do attend to strong syllables and their robust acoustic cues. For this study, the impact of vision on participant's ability to attend to lexical cues was examined. It is hypothesized that vision would augment both bilateral and bimodal cochlear implant users' ability to access these lexical cues and promote lexical segmentation. If this is happening there would be an increased likelihood that participants would insert lexical boundaries before strong syllables as opposed to inserting lexical boundaries before weak syllables and have more deletions before weak syllables than strong.

In summary it is hypothesized that for both bilateral and bimodal cochlear implant users vision will benefit speech perception as evidenced by an increased percent words correct. If this increase is because of better place of articulation information consonant errors will include greater consonant place accuracy. If the increase is because of better cues to syllabic stress lexical boundary errors will be of the predicted variety.

## II. METHODS

### A. Participants

Thirteen (13) participants with cochlear implants participated in the experiment. Participants were recruited from within Arizona and from other states including North Carolina, Minnesota, Texas, Florida, and Ohio. Eight (8) participants were bilaterally implanted. Within this group there were 3 females and 5 males ranging in age from 21 to 87 years of age. The 8 bilateral participants had an average age of 61. All participants but 1 have had implants for at least 2 years, with implant dates ranging from 2002 to 2011. One (1) participant had the first implant in 2008, but the second cochlear implantation wasn't implanted until 2013. Five (5) participants had bimodal input from both a cochlear implant and hearing aid. All bimodal participants were female, range in age from 57 to 83 years of age, and have an average age of 68. Bimodal participants' implantation dates ranged from 1998 to 2007.

### B. Speech Stimuli

Experimental phrases previously generated by Liss et al. (1998) were presented to each subject. Each experimental phrase contained three to five words and consisted of six syllables with alternating stress. One hundred and sixty (160) experimental phrases were used in the study (see Appendix A). Half the phrases contained a strong-weak syllabic stress pattern (e.g., *balance clamp and body*) and the other half of the phrases contained a weak-strong syllabic stress pattern (e.g., *create her spot of art*). The phrases were semantically anomalous but syntactically plausible. All experimental phrases were presented to the participants via an adult male talker. For the audiovisual condition, video of the adult male talker along with the audio signal was presented to the participants. Video presentation of stimuli occurred via a computer monitor placed immediately in front of participants at a standardized viewing distance.

**C. Procedure**

Testing was completed in a double-walled sound booth at the Speech and Hearing Sciences Department, Arizona State University, between January 2014 and September 2014. All subjects, bilateral and bimodal, completed a short practice session. During the practice session, participants were presented with experimental phrases similar in composition to those used during the experiment. Participants were presented with practice stimuli in both audio only and audiovisual presentation. Further, the signal-to-noise ratio for each participant was identical to the signal-to-noise ratio used during actual experimentation. Participants sat at a distance of 3 feet from the computer monitor used for concurrent visual presentation. Auditory input in both auditory and audiovisual conditions was presented at 60 decibels of sound pressure level.

During the experiment each subject was presented with all 160 experimental phrases. Phrases were divided into 8 blocks of 20 phrases. Each block consisted of 10 strong-weak syllabic stress pattern phrases and 10 weak-strong syllabic stress pattern phrases. Four (4) blocks of 20 phrases were presented with auditory-input only and 4 blocks of 20 phrases were presented with auditory and visual input. The order of presentation of auditory-input only phrase blocks and auditory-visual input phrase blocks was randomized. All 160 experimental phrases were presented to each subject in one setting.

In order to identify the sources of intelligibility benefit, it was necessary to generate error responses for analysis. To accomplish this and avoid a potential ceiling effect, multi-talker babble noise was used to mask the experimental phrases. Speech intelligibility scores were first obtained using AzBio sentences. According to Dorman et al. (2013) these are sentences that were developed as an alternative to conventional and well-known sentence test materials. Sentences are of an equivalent difficulty level and for use in the evaluation of performance over time and across conditions (Spahr et al., 2012). Since the experimental phrases used are semantically anomalous but syntactically plausible and the AzBio sentences are syntactically and semantically plausible the signal-to-noise ratio (SNR) was consequently adjusted. The SNR for experimental phrases targeted approximately 30% intelligibility as a baseline. For bimodal subjects, the SNR was between +3decibel(dB) and +15dB. Bilateral listeners had an SNR between +3dB and +10dB.

Individual SNR values can be found in following tables:

**Table 1: Bilateral Subject SNR**

| | |
|---|---|
| BL1 | +5dB |
| BL2 | +7dB |
| BL3 | +3dB |
| BL4 | +5dB |
| BL5 | +6dB |
| BL6 | +10dB |
| BL7 | +4dB |
| BL8 | +4dB |

**Table 2: Bimodal Subject SNR**

| | |
|---|---|
| BM1 | +7dB |
| BM2 | +6dB |
| BM3 | +6dB |
| BM4 | +15dB |
| BM5 | +3dB |

### D. Data Analysis

The total data set consists of 13 participant files (8 bilateral and 5 bimodal). Each participant file has 160 experimental phrases (80 strong-weak stress; 80 weak-strong stress). All experimental phrases (2080 total phrases) were analyzed and then re-analyzed for multiple measures. These measures included three subsets of analysis, total words correct, lexical segmentation and phonemic resemblance.

Lexical segmentation measures included: total words correct, percent words correct, insert before strong, insert before weak, delete before strong, delete before weak, delete syllable, and add syllable. Lexical segmentation examples are included below.

Table 3: Lexical Segmentation Examples

| LEXICAL BOUNDARY TYPE | TARGET PHRASE | PARTICIPANT RESPONSE |
|---|---|---|
| **Words Correct** | *balance clamp and bottle* | *balance clamp and body (3 words correct)* |
| **Inserting Strong** | ***account** for who could knock* | ***a count** for who could knock* |
| **Inserting Weak** | *balance clamp and **bottle*** | *balance clamp and **bought her*** |
| **Delete Strong** | frame ***her seed** to answer* | frame ***herself** to answer* |
| **Delete Weak** | ***bush is** chosen after* | ***gorgeous** chosen after* |
| **Add Syllable** | *pooling pill or **cattle*** | *worry wool a **chemical*** |
| **Delete Syllable** | *constant willing **walker*** | *constant willing **walk*** |

Response words that were analyzed for phonemic resemblance and analyzed for consonant errors included those without lexical boundary errors, syllable deletions, or syllable additions. Not meeting these indices would have precluded the response words from meeting strict predetermined qualifiers necessary for a response word to be considered phonemically resemblant to the target word.

In order for a response word to qualify as being phonemic resemblant to a target word specific predetermined criteria had to be met. For 1-syllable target words with more than 3 phonemes, the subject response word was considered phonemically resemblant to the target word if at least 50% of target phonemes were correctly identified. When a 1-syllable word only had 3 phonemes, the subject response word needed to only have 1 correct phoneme. Participant responses to 2-syllable words had to contain the correct number of syllables and at least 50% of the phonemes in the target word had to be matched in order for the word to be considered phonemically resemblant.

Participant response words that were judged to be phonemically resemblant per the above stated criteria were then analyzed with regard to the specific phonemes that comprised the target word and consequent response. Phonemic resemblance analysis included various measures including the total number of phonemes and the amount of correct phonemes for applicable words within experimental phrases. Next, the type of error was recorded. Error types included: vowel, consonantal place, consonantal manner, consonantal voice, extra phoneme insertion, and phoneme deletion. While data were acquired for all error types, only consonant errors were examined in this study.

Table 4: Phonemic Resemblance Error Examples

| PHONEMIC RESEMBLANCE EXAMPLES | TARGET WORD | PARTICIPANT RESPONSE |
|---|---|---|
| **Error Vowel (EV)** | shape | sheep |
| **Error Place (EP)** | lead | league |
| **Error Manner (EM)** | log | dog |
| **Error Voice (EV)** | hat | bat |
| **Phonemic Insertion (PI)** | age | page |
| **Phonemic Deletion (PD)** | treats | trees |

It was possible for a participant response to have more than one word phonemically resemblant within a target phrase. In the following example two of the three words in the target phrase are inaccurately perceived by the participant but are also phonemically resemblant.

Target: ***darker*** *painted* ***baskets***

Response: ***farther*** *pizza* ***caskets***

It was also possible for a single response word to have more than one error and still be considered phonemically resemblant. In the following example the subject response word did just that. In this particular response the subject response contained a phonemic insertion error as well as a vowel error.

Target: ***oyster***

Response: ***poster***

Finally, it was also possible for a single incorrect phoneme provided in the participant response to fall within multiple phonemic error categories. In the following example, the participant substituted [k] for [n] and consequently had a single error that spanned three categories, consonantal place, consonantal manner, and consonantal voicing.

Target: ***stain***

Response: ***stake***

After all participant data were analyzed on word correct, lexical error, and phonemic error measures, analysis focused on group data for both the bilateral and bimodal subject sets separately. For each group, total words correct, percent words correct, lexical-boundary error ratios, and consonantal feature errors were evaluated and will be discussed later in the results section.

**E. Reliability**

All subject files were analyzed a second time by the original coder (intra-judge). Twenty five percent of subject files were randomly selected and analyzed by a second trained judge (inter-judge) to obtain reliability estimates for the coding aspect of analysis. Inter-judge agreement between the first judge using the reanalyzed data set and the set analyzed by the second trained judge was high and is noted below.

**Table 5: Bilateral Data Coding Analysis Reliability – Inter-judge Percent Agreement**

| Total Words Correct | Lexical Boundary Errors | Consonantal Phonemic Resemblance |
|---|---|---|
| 98.13% | 99.30% | 97.29% |

**Table 6: Bimodal Data Coding Analysis Reliability – Inter-judge Percent Agreement**

| Total Words Correct | Lexical Boundary Errors | Consonantal Phonemic Resemblance |
|---|---|---|
| 99.50% | 99.38% | 97.00% |

## III. RESULTS

**Figure 1** reflects the percent of the total number of words in experimental phrases each bilateral cochlear implant participant was able to accurately perceive. This value is further delineated for each subject in terms of the percent of the total number of words correct for both the audio-only condition and the audiovisual condition. As a group the bilateral users were able to accurately identify 37.3% of the words in target phrases without vision and 54.8% of those words with vision. A one-tailed t-test was performed comparing the percent of the total number of words in experimental phrases that were correctly perceived in the auditory-only and audiovisual conditions for each subject. For this analysis a one-tailed t-test was utilized because it was believed vision would not have the effect of reducing performance and all subjects had more words correct when vision was added to the audio signal. There was a significant effect for vision, $t(7) = -9.75$, $p<.001$. Thus, bilateral cochlear implant listeners were significantly more apt to correctly perceive words in experimental phrases when a visual signal was provided compared to when an auditory signal alone was given.

**Figure 1: Bilateral Subjects – Percent Words Correct**

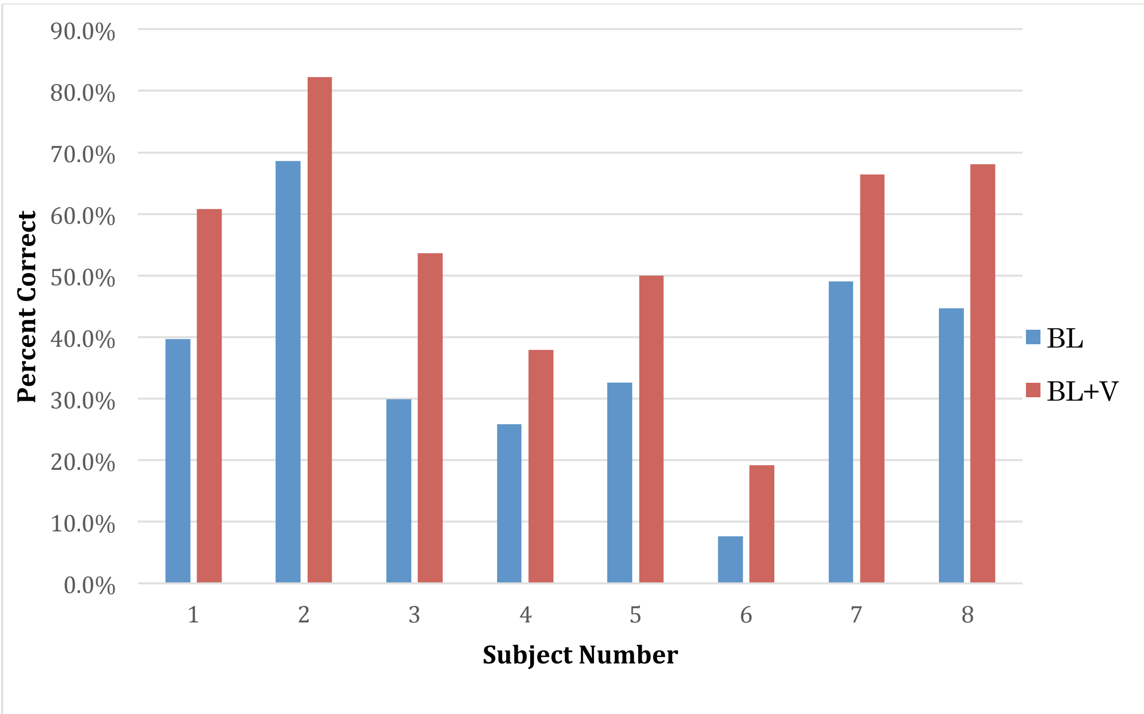



**Figure 2** reflects the percent of the total number of words in experimental phrases each bimodal cochlear implant participant was able to correctly perceive. Further, this value is delineated for each subject in terms of the percent of the total number of words correct for both the audio-only condition and the audiovisual condition. As a group the bimodal participants were able to accurately identify 32.2% of the words in target phrases without vision and 56.0% of those words with vision. In line with the analysis that occurred concerning bilateral cochlear implant users, a one-tailed t-test was performed comparing the percent of the total number of words in experimental phrases that were correctly perceived in the auditory-only and

audiovisual conditions for each subject. Again, all bimodal participants had shown in increase in the number of words correct when vision was added so interest was only in whether this effect was significant in this direction. There was a significant effect for vision, $t(4) = -4.95$, $p<.01$. Thus, bimodal cochlear implant listeners also were significantly more apt to correctly perceive words in experimental phrases when a visual input was added to the auditory input.

**Figure 2: Bimodal Subjects – Percent Words Correct**

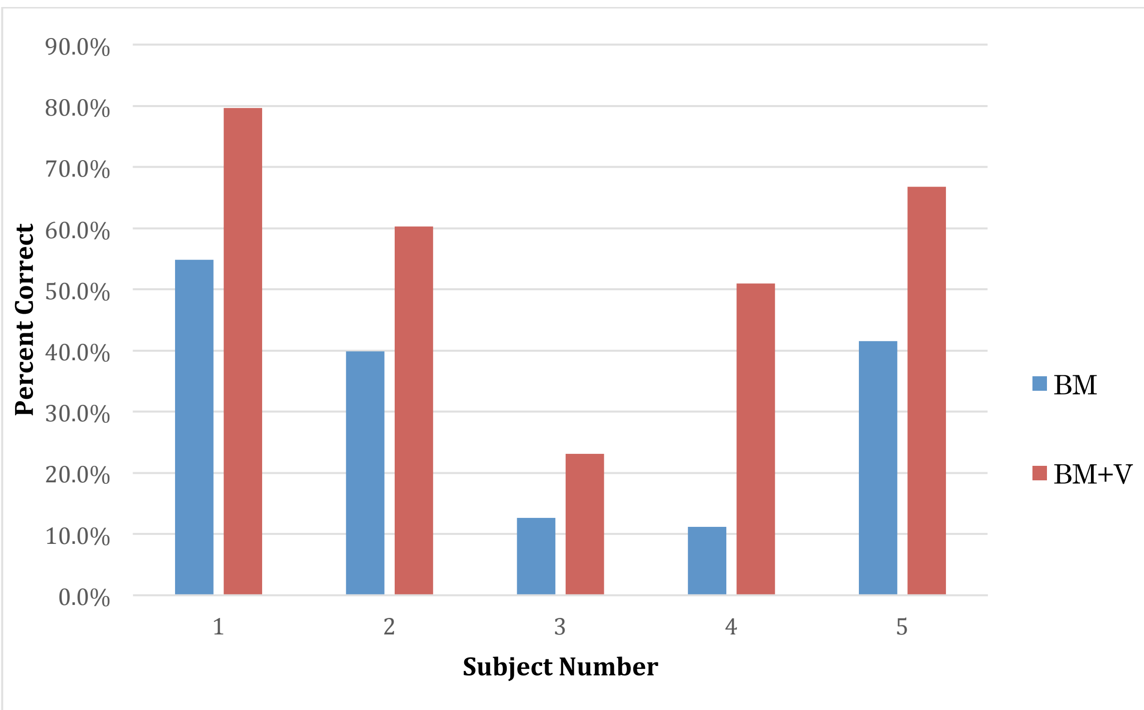



**Figure 3** reflects the percent of each type of consonant error (place, manner, voice) relative to the total number of consonant errors committed by the bilateral cochlear implant participants. The percentage of consonant errors with incorrect place of articulation decreased significantly with the addition of vision for the bilateral participants ($t(7) = 1.05$, $p<.01$). However, there was no benefit of vision for manner

($t(7) = 1.005$, $p > .01$ ); and a significantly greater proportion of voicing errors with vision ($t(7) = -3.2765$, $p < .01$) Thus, for bilateral participants, the addition of vision only benefitted place of articulation for consonants.

**Figure 3: Bilateral Consonantal Feature Analysis**

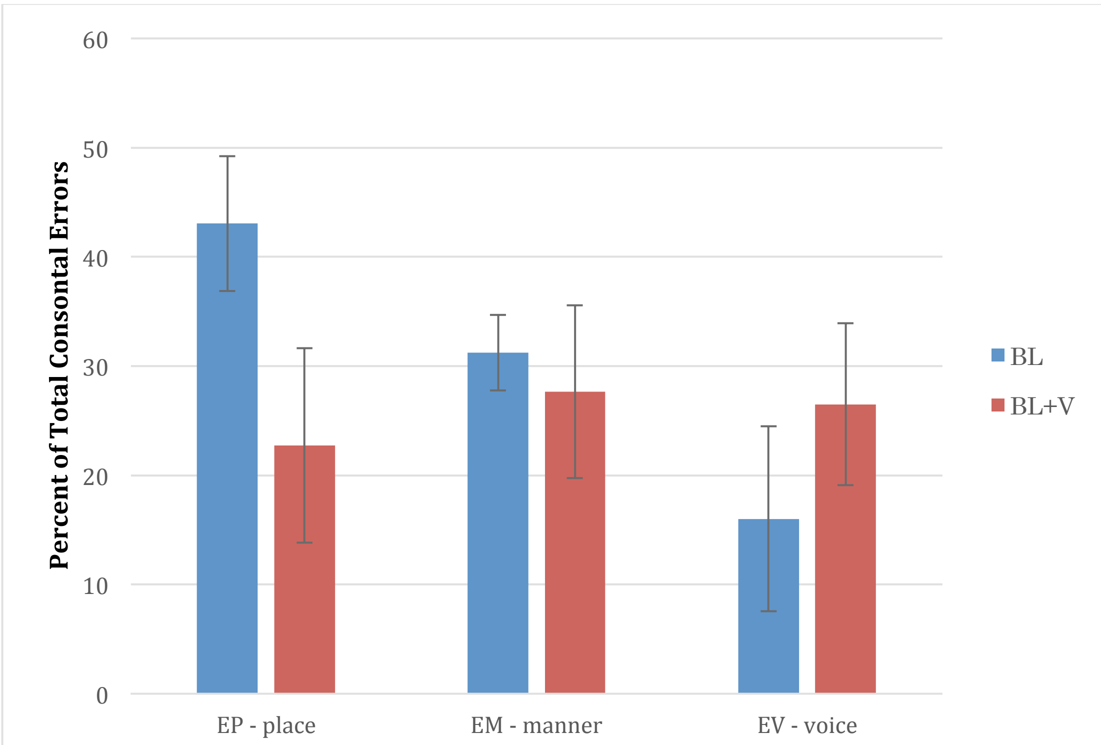



**Figure 4** reflects the percent of each type of consonant error (place, manner, voice) relative to the total number of consonantal errors committed by the bimodal cochlear implant participants. Results for place were not significant, ($t(4) = 0.4$, $p > .05$). Similarly, manner ($t(4) = 0.87$, $p > .05$) and voicing ($t(4) = -0.35$, $p > .05$) features did not benefit from the addition of vision for these participants. Thus, there is not a significant difference in consonant errors for each subtype between bimodal cochlear implant users using only an auditory signal versus having audiovisual input.

**Figure 4 – Bimodal Consonantal Feature Analysis**

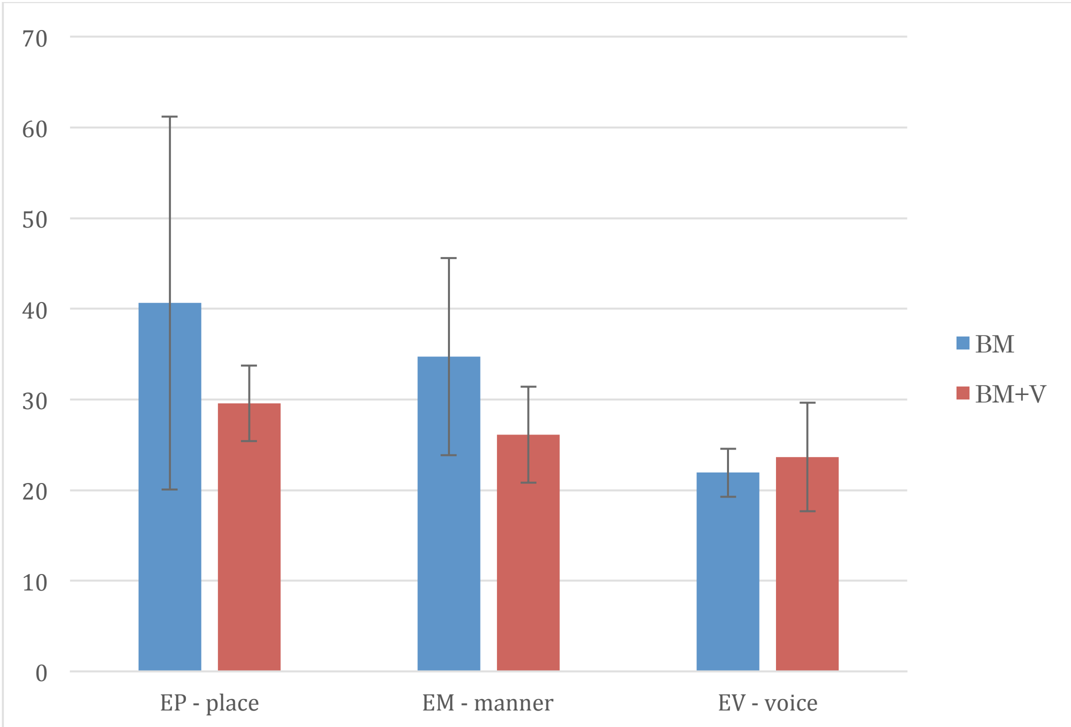



**Figure 5** reflects ratios indicative of specific types of lexical boundary errors committed by the bilateral cochlear implant participants. Two ratios were examined to determine if support for utilization of Metrical Segmentation Strategy hypothesis existed. The primary support for this comes from the ratio of insert strong errors to insert weak errors. To a lesser extent the ratio of delete weak errors to delete strong errors is of interest as it is predicted to support the hypothesis. The strength of agreement and support for participants utilizing Metrical Segmentation Strategy are commensurate with the ratio values, with a value of "1" indicating insertion and deletion errors occur with equal frequency before strong and weak syllables. The greater the strength of adherence to the predicted pattern in support of Metrical Segmentation Strategy, the greater the positive distance from "1". Bilateral

cochlear implant users not integrating a visual signal had insertion errors 1.66 times more often before a strong syllable than a weak one and insertion errors 2.1 times more often before a strong syllable than a weak one with the addition of visual input. Without vision deletion errors occurred 0.23 times more often before a weak syllable than a strong one and with vision they occurred 0.42 times more often before a weak than strong syllable.

A nonparametric goodness-of-fit Chi-square analysis comparing the four error type rates (insert before strong, insert before weak, delete before strong, and delete before weak) used in error ratios was performed, comparing error type rates for participant responses without vision and with vision. There was a significant effect, $X^2$ (3, N=285) = 8.83, p= 0.032 indicating error-rate values were drawn from two different populations. While bilateral cochlear implant users did attend to lexical segmentation cues with and without vision, the addition of visual input enhanced participant attendance to lexical cues in support of Metrical Segmentation Strategy.

**Table 7: Bilateral Subjects – Total Lexical Boundary Errors by Type**

| | Insert Strong | Insert Weak | Delete Strong | Delete Weak |
|---|---|---|---|---|
| Audio Only | 118 | 71 | 78 | 18 |
| With Vision | 105 | 50 | 50 | 21 |

**Figure 5: Bilateral Subjects – Lexical Boundary Error Ratios**

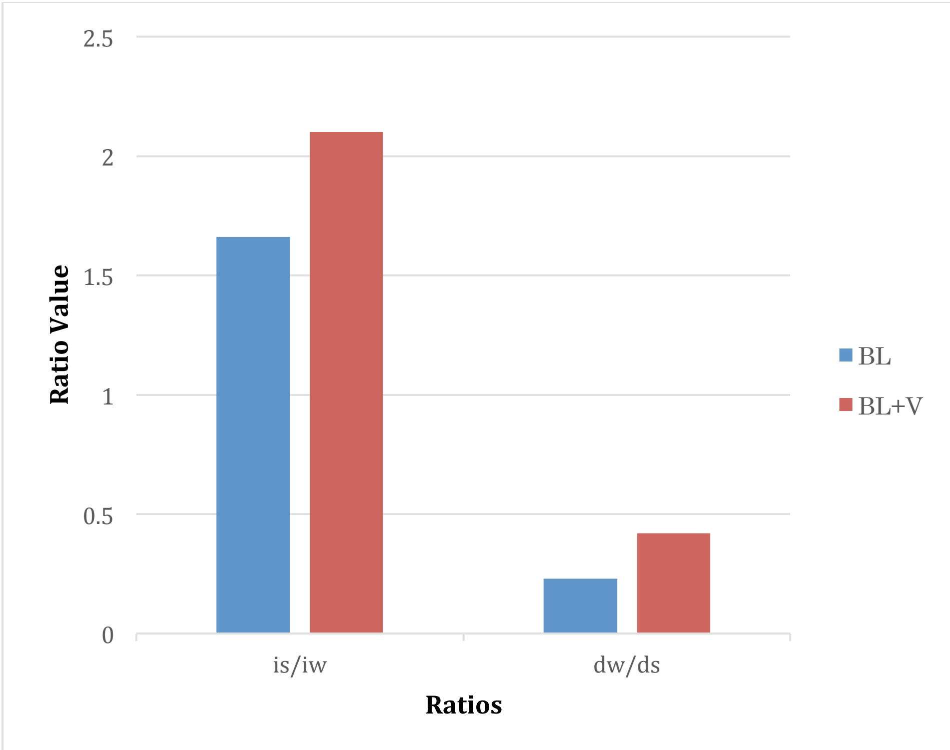


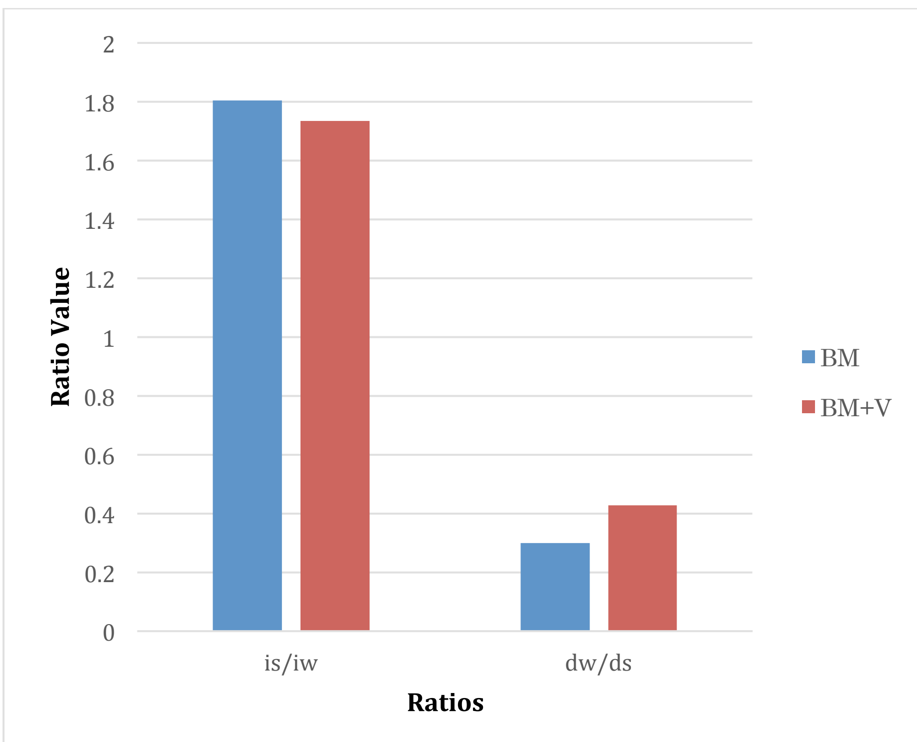

**Figure 6** reflects the ratios of specific types of lexical boundary errors committed by the bimodal cochlear implant users. Bimodal cochlear implant users not integrating a visual signal had insertion errors 1.80 times more often before a strong syllable than a weak one. When vision was added the same participants committed insertion errors 1.74 times more often before a strong than weak syllable. Deletion errors occurred 0.30 times more often before a weak syllable than a strong one without vision and 0.43 times more often with vision. A nonparametric goodness-of-fit analysis Chi-square analysis comparing the four values describing error rates among bimodal cochlear implant users without vision to the four values describing rates among users with the addition of vision was not significant (p=0.31). Thus, while bimodal cochlear implant users relied on syllable stress to identify word

boundaries, the addition of vision did not result in a different pattern of lexical boundary errors.

**Table 8: Bimodal Subjects – Total Lexical Boundary Errors by Type**

| | Insert Strong | Insert Weak | Delete Strong | Delete Weak |
|---|---|---|---|---|
| Audio Only | 101 | 56 | 30 | 9 |
| With Vision | 59 | 34 | 21 | 9 |

**Figure 6: Bimodal Subjects – Lexical Boundary Error Ratios**



## IV. DISCUSSION AND CONCLUSION

The present study investigated how the addition of vision augments information received via auditory and electrical inputs in bilateral and bimodal cochlear implant participants. Specifically, this impact was measured across three markers of perceptibility: percent words correct, frequency of consonant errors, and lexical boundary errors.

It was hypothesized that both bilateral and bimodal participants would have benefit from the addition of vision and the overall percent words correct would increase. Further, it was hypothesized if the reason was because vision added consonant place information then errors in this regard would be reduced. And finally, it was hypothesized if vision granted access to prosodic information, lexical errors would follow certain patterns.

In addressing the first hypothesis, we confirmed that vision significantly benefits speech perception for both bilateral and bimodal cochlear implant users. When vision was added both bilateral and bimodal cochlear implant users were able to more accurately transcribe whole words in phrases. All participants demonstrated this gain in percent words correct and it was significant at the group level for both. The magnitude of benefit between bilateral and bimodal cochlear implant participants was not significantly different ($F(1,11)=000$, $p=.999$)

MacLeod and Summerfield (1987) point out that most individuals find it easier to hear and understand speech when they can see whomever is talking. Rouger, J. et al. (2009) point out that single cochlear implant users maintain a robust level of speech, or lip reading. This finding is true even after multiple years of cochlear implant use and electric signal integration. But the question remains was there more going on in terms of vision granting access to place and lexical cues.

Further, the question remains did percent words correct improve for the same reasons regarding each group of cochlear implant users? Recall that the bilateral participants have implants in both ears and are only receiving electrical input. They appear to show benefit from vision. With visual input the frequency of consonant place errors was significantly reduced.

Bilateral cochlear implant users also demonstrated a reliance on syllabic stress cues for lexical segmentation in support of Metrical Segmentation Strategy. It postulates that the highest number of errors will be insertions before strong syllables and that was the case both with and without vision. Although deletion errors did not follow the predicted pattern both with and without vision there were relatively few deletion errors overall and these errors are not the primary marker of Metrical Segmentation Strategy. Most importantly, and robustly, listeners tended to treat strong syllables as word onsets when they made errors. And of significance is the fact that when vision was added this tendency increased. This suggests there is something in the visual signal that improves identification of strong syllables and these listeners used this information.

In summary, vision improved speech perception for bilateral cochlear implant users. This gain occurred because of two reasons as evidenced by the error analysis data. For one, visual cues provided increased access to place of articulation. And finally visual cues provided information regarding syllable strength.

Recall, bimodal cochlear implant users have a more diverse input than bilateral users. Bimodal users have electric input in one ear and acoustic input via a hearing aid in the other. In this regard they are already potentially integrating an electric and auditory signal. Regarding place cues, bimodal cochlear implant users did not show any place of articulation benefit with vision. Additionally, there was not a significant reduction in manner or voice errors as well.

However, it is worth noting that there was high variability between bimodal participants' performance. This is not surprising considering a greater skill level regarding the integration of auditory and visual cues, or higher integration efficiency, will almost lead to improved performance in auditory-visual tasks (Grant, Tufts, & Greenberg 2007; Grant, Walden, and Seitz, 1998). And it is assumed that efficient integrators are better at using cues from multiple sources in speech recognition tasks (Grant, Walden, and Seitz, 1998). It might just be in this instance that some bimodal cochlear implant users are more efficient at integrating all input signals, including vision, electric, and acoustic inputs.

Bimodal cochlear implant users did use syllabic stress to identify word boundaries with and without vision. But, bimodal cochlear implant users did not demonstrate an improved access to these prosodic cues in support of Metrical Segmentation Strategy with the addition of visual input. This is evidenced by the fact lexical boundary error distributions did not significantly differ between the audio and audiovisual data sets. In fact, the ratio of insert strong errors to insert weak errors was almost identical with and without the addition of visual input for bimodal users.

In summary and in reference to the original hypotheses outlined in this study bimodal cochlear implant users did have improved speech perception as demonstrated by increased words correct with the addition of vision. However this improvement is not attributable to place of articulation information. Nor is it the case the improvement in percent words correct is attributable to syllable strength information as outlined by Metrical Segmentation Strategy. This raises questions as to what might be the source of improvement.

It might be the case that their access to acoustic information renders the addition of visual information less important for low-level processing such as phoneme identification and lexical segmentation. It is reasonable to hypothesize this might in fact be due to the bimodal cochlear implant users already getting additional acoustic input from the hearing aid that washes out any effect vision might have on perceptibility. Studies have shown that the additional low-frequency speech signals provided by the hearing aid in bimodal cochlear implant users provide prosody cues and help with the perceptual organization of the signal generated by the cochlear implant (Cullingham & Zeng 2011; Nittrouer & Chapman 2009). That is, bimodal cochlear implant users do not need to access the same lexical segmentation strategies that bilateral cochlear implant users do because the auditory input is more robust. Jang et al. (2014) also point out that pitch and fine temporal structure low-frequency information can be ascertained from residual hearing via acoustic amplification.

Perhaps visual information simply reinforces the acoustic input they are already receiving via the hearing aid to result in improved performance. Or it might be the case, as eluded to earlier, that there is just too much variability in performance, especially given the small sample size of the bimodal cochlear implant users. This is an important point in this study. Due to high subject-response variability and the overall small bimodal participant sample size relatively low power was achieved during statistical analysis of bimodal participant data.

However, if the differences regarding how bilateral and bimodal cochlear implant users are able to improve speech perceptibility as demonstrated through measures such as percent words correct are replicated in future studies and deemed real, they suggest a more complex relationship among acoustic and visual cues in speech perception than previously believed.

In future studies a larger sample size should be acquired. Additionally, to vest out the potentially complicated interaction of vision and acoustic input other possible sources of improved intelligibility granted by vision should be explored. Such measures might include lip-reading ability and general problem-solving ability with acoustically degraded signals such as those experienced by bilateral and bimodal cochlear implant users.

REFERENCES


Cullington, H. E., & Zeng, F. (2011). Comparison of bimodal and bilateral cochlear implant users on speech recognition with competing talker, music perception, affective prosody discrimination, and talker identification. *Ear and Hearing, 32(1),* 16-30.

Cutler, A. and Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language,* 31, 218-236.

Cutler, A. and Carter, D.M. (1987). The predominance of strong syllables in the English vocabulary. *Comput. Speech Language, 2*, 133-142.

Cutler, A. and Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology, 14*, 113-121.

Dorman, M and Gifford, R. (2010). Combining acoustic and electric stimulation in the service of speech recognition. *Int J Audiol.*, *49(12),* 912–919.

Dorman, M. F., Spahr, A., Gifford, R. H., Cook, S., Zhang, T., Loiselle, L., Schramm, D. (2012). Current research with cochlear implants at Arizona State University. *Journal of the American Academy of Audiology*, *23(6),* 385–395.

Grant, K. W., Tufts, J. B., & Greenberg, S. (2007). Integration efficiency for speech perception within and across sensory modalities by normal-hearing and hearing-impaired individuals. *The Journal of the Acoustical Society of America, 121(2),* 1164-1176.

Grant, K.W., Walden, B.E., and Seitz, P.F. (1998). Auditory-visual speech recognition by hearing-impaired subjects: Consonant recognition, sentence recognition, and auditory-visual recognition. *J. Acoustical Society of America, 100*, 2415-2424.

Jang, J. H., Lee, J. H., Chang, S. O., & Oh, S. H. (2014). Effect of aided hearing in the nonimplanted ear on bimodal hearing. *Otology & Neurotology, 35(10),* e270-e276.

Lansford, K. L., Liss, J. M., Caviness, J. N., & Utianski, R. L. (2011). A cognitive-perceptual approach to conceptualizing speech intelligibility deficits and remediation practice in hypokinetic dysarthria. *Parkinson's Disease, 2011*, 1-9.

Liss, J.M., Spitzer, S., Caviness, J.N., Adler, C.A., Edwards, B. (1998). Syllabic strength and lexical boundary decisions in the perception of hypokinetic dysarthric speech. *J Acoust Soc Am., 104*, 2457-2466.

MacLeod, A., Summerfield, Q. (1987). Quantifying the contribution of vision to speech perception in noise. *British journal of audiology* ***21(2),*** 131–141.

Mattys, S. L., White, L., & Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology: General, 134(4),* 477-500.

Nittrouer, S., & Chapman, C. (2009). The effects of bilateral electric and bimodal Electric—Acoustic stimulation on language development. *Trends in Amplification, 13(3),* 190-205.

O'Donoghue, G. (2013). Cochlear implants-science, serendipity, and success. *The New England Journal of Medicine, 369(13),* 1190.

Rouger, J., Fraysse, B., Deguine, O., & Barone, P. (2008). McGurk effects in cochlear-implanted deaf subjects. *Brain Research, 1188*, 87-99.

Rosenblum L.D. (2005). The primacy of multimodal speech perception. InPisoni D., Remez R.(Eds.), *Handbook of speech perception* (pp. 51–78). Malden, MA: Blackwell.

Rosenblum, L. D. (2008). Speech perception as a multimodal phenomenon. *Current Directions in Psychological Science, 17(6),* 405-409.

Spahr, A. J., Dorman, M. F., Litvak, L. M., Van Wie, S., Gifford, R. H., Loizou, P. C., et al. (2012). Development and validation of the AzBio sentence lists. *Ear and Hearing, 33(1),* 112-117.

Svirsky, M. A., Sagi, E., Meyer, T. A., Kaiser, A. R., & Teoh, S. W. (2011). A mathematical model of medial consonant identification by cochlear implant users. *The Journal of the Acoustical Society of America, 129(4),* 2191-2200.

van Hoesel, R. J. M. (2012). Contrasting benefits from contralateral implants and hearing aids in cochlear implant users. *Hearing Research, 288(1-2)*, 100.

# APPENDIX A

## EXPERIMENTAL PHRASES

Phrases

| Strong-Weak | Weak-Strong |
|---|---|
| absent fields did morning | a reason guests contained |
| age of centered wagons | account for who was knocked |
| ancient leading students | address her meeting time |
| balance clamp and bottle | admit the gear beyond |
| birth the notice symbol | advance but sat appeal |
| bolder ground from justice | afraid beneath demand |
| both enjoyed was processed | allow assured remains |
| bush is chosen after | alone become restrained |
| butcher shook the middle | amend estate approach |
| cheap control in paper | and spoke behind her sin |
| closer showing metal | appear to wait or turn |
| constant willing walker | arrived had land can scale |
| cool the jar in private | assume to catch control |
| darker painted baskets | attack became concerned |
| deep conceived the feature | attend the trend success |
| defend years to something | avoid or beat command |
| distant leaking basement | award his drain away |
| done with finest handle | aware reviewed such trees |
| drove arrive the other | before his wish was strong |
| during pattern programs | begins excess is near |
| each informed from flowers | begun his crown belief |
| figure proves from normal | beside a sunken bat |
| force of focus moment | career despite research |
| form object with knowing | commit such used advice |
| fort believed such orders | compare events of bank |
| frame her seed to answer | complete it thought or troops |
| friendly moon was sectioned | conduct ideal had songs |
| functions aim his acid | confused but roared again |
| headed wheels with stories | connect the beer device |
| higher patient concept | create her spot of art |
| hold a page of fortune | debate reply was mean |
| housing drawn in samples | define respect instead |
| hundred printed license | degree prevents from games |
| kick a tad above this | depend is longer sound |
| lake is pressure sofa | derived extent with streets |
| listen final station | desire can bar accept |
| mark a single ladder | detail require such risk |

| | |
|---|---|
| mate denotes a judgement | direct can sweet extreme |
| may the same pursued it | divide across retreat |
| measure fame with legal | effect his wage but stood |
| mister types is fashion | embark or take her sheet |
| mode campaign for budget | extend but please his stones |
| model sad and local | for coke a great defeat |
| motion double garden | forget the joke below |
| music useful rising | had eaten junk and train |
| narrow seated member | had value plants to mind |
| now return can money | her owners arm the phone |
| orders fairly level | improve in driving cloud |
| pain can follow agents | increase a grade sedate |
| passing plus a factor | indeed a tax ascent |
| people bought such minor | intense perfects such coasts |
| pick a chain for action | its harmful note abounds |
| plan reduced its setting | itself a band provides |
| plenty please or causes | mistake delight for heat |
| pooling pill or cattle | obtain contracts from tasks |
| push her equal culture | or spent sincere aside |
| question major nature | perceive sustained supplies |
| rampant boasting captain | permits achieved but lied |
| resting older earring | present relief among |
| rhythm under artist | recall had future through |
| rocking modern poster | regards because had class |
| rode the lamp for teasing | remove and name for stake |
| round and bad for carpet | required attempt maintain |
| rowing farther matters | secure but lease apart |
| seat for locking runners | submit his cash report |
| sight about this deeper | suggest its price reserve |
| signal breakfast pilot | support with dock and cheer |
| sinking rather tundra | technique but sent result |
| spackle enter broken | this daughter presence rules |
| speaking clear is power | to sort but fear inside |
| stable wrist and load it | transcend almost betrayed |
| target keeping season | unique exchange in holes |
| tension known from pleasure | unless escape can learn |
| thinking charged the hearing | unseen machines agree |
| thrown had special office | around without such roads |
| truth improved in shelters | assign exists perhaps |
| vital seats with wonder | exam of joy began |

| | |
|---|---|
| wanted find the finger | refer to good from league |
| weather pure was surfaced | refused percent to goal |
| world repeats with feelings | surprise was might between |