Use Authorization

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at Idaho State University, I agree that the Library shall make it freely available for inspection. I further state that permission to download and/or print my thesis for scholarly purposes may be granted by the Dean of the Graduate School, Dean of my academic division, or by the University Librarian. It is understood that any copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Signature ________________________________

Date ____________________________________

An Exploration of the Effects of Opioid Pain Medication on Speech Perception and Audiovisual Integration Abilities in Middle Aged Adults

by

Jessica E. Simmons

A thesis submitted in partial fulfillment

of the requirements for the degree of

Master of Science in the Department of Communication Sciences and Disorders

Idaho State University

August 2015

Committee Approval

To the Graduate Faculty:

The members of the committee appointed to examine the thesis of Jessica Simmons find it satisfactory and recommend that it be accepted.

Nicholas Altieri, Ph.D.
Major Advisor

Wendy Morgan, M.S., CCC-SLP
Committee Member

Dotty Sammons, Ph.D.
Graduate Faculty Representative

Idaho State
UNIVERSITY

Office for Research Integrity
921 South 8th Avenue, Stop 8046 • Pocatello, Idaho 83209-8046

June 10, 2014

Nicholas Altieri, PhD
Mail Stop 8116
CSED
Pocatello, ID 83209

RE: Your application dated 6/10/2014 regarding study number 4090: Impact of Chronic Opioid Medication Use on Neuro-Cognitive and Language Functions (Mountain West Clinical Translational Research)

Dear Dr. Altieri:

Thank you for your response to requests from a prior review of your application for the new study listed above. Your study is eligible for expedited review under FDA and DHHS (OHRP) designation.

This is to confirm that your application is now fully approved. The protocol is approved through 6/10/2015.

You are granted permission to conduct your study as most recently described effective immediately. The study is subject to continuing review on or before 6/10/2015, unless closed before that date.

Please note that any changes to the study as approved must be promptly reported and approved. Some changes may be approved by expedited review; others require full board review. Contact Thomas Bailey (208-282-2179; fax 208-282-4723; email: humsubj@isu.edu) if you have any questions or require further information.

Sincerely,

Ralph Baergen, PhD, MPH, CIP
Human Subjects Chair

Phone: (208) 282-2592 • Fax: (208) 282-4723 • www.isu.edu/research
ISU is an Equal Opportunity Employer

TABLE OF CONTENTS

List of Figures……………………………………………………………………………vii

List of Tables…………………………………………………………………………....viii

Abstract…………………………………………………………………………………...ix

Introduction………………………………………………………………………………1

Age and Hearing Loss…………………………………………………………….3

Opioid Medication and Hearing Loss…………………………………………….5

Audiovisual Speech Integration…………………………………………………..6

Measures of Audiovisual Speech Integration…………………………………….7

Gain Scores and Accuracy Measures……………………………………..7

Age and Multisensory Integration………………………………………...8

Processing Speed…………………………………………………………...13

Hearing Loss and Multisensory Integration……………………………………..14

Reaction time and Capacity Measures…………………………………...14

Response Time……………………………………………………………...14

Hypotheses…………………………………………………………………………19

Methods……………………………………………………………………………………20

Participants……………………………………………………………………….20

Materials…………………………………………………………………………...20

Procedures…………………………………………………………………………21

Data Collection and Instrumentation……………………………………………22

Results……………………………………………………………………………………24

Sentence Recognition…………………………………………………………………24

Measures of Audiovisual Gain………………………………………………..25

Speeded-Word Recognition……………………………………………………...27

Reaction Time……………………………………………………………27

Accuracy…………………………………………………………………28

Capacity…………………………………………………………………29

Discussion………………………………………………………………………………….34

Processing Speech and Capacity Measures……………………………………..37

References………………………………………………………………………………….40

Appendix…………………………………………………………………………………...44

List of Figures

Figure 1A, Box plots showing sentence recognition accuracy scores………………….......26

Figure 1B, Box plots showing the distribution of gain scores using both formulas described above……………………………………………………………………......28

Figure 2, Participant averaged capacity measures. The right panel shows *C(t)* and the left shows *C_I(t)*………………………………………………………………………......31

Figure 3, Box plots showing six capacity measures……………………………………..34

List of Tables

Table 1, Mean reaction time on speeded word recognition task measured in milliseconds (ms)……………………………………………………………………………………29

Table 2, Mean, Median, and Maximum capacity scores………………………………...33

Abstract

Speech can be perceived through the auditory system, the visual system, and a combination of the audio and visual systems. Individuals can often compensate missing auditory information through the use of lip reading and other observable non-linguistic signals. Opioids are potentially ototoxic drugs that can affect hearing and neurocognitive functions, both of which play a role in speech perception. However, there is lack of research indicating the extent to which opioid drugs affect speech perception. This thesis will investigate the extent to which potentially ototoxic opioid drugs adversely affect multimodal speech integration and speech perception. It is important for speech language pathologists to know how much medications containing opiates can affect their clients' abilities to perceive their speech during assessment and treatment. It is hypothesized that opioid usage will positively affect speech integration and perception abilities in middle aged adults. It is expected that knowledge gained from this study can help speech language pathologists and audiologists better understand the abilities of their clients.

Introduction

Though speech perception is primarily an auditory function in normal-hearing listeners, noise can often hinder a person's ability to hear the spoken message clearly (Sommers, Tye-Murray, & Spehar, 2005). Decades of research has shown that listeners can fill in missing auditory information through the use of lip reading, or being able to see a talker's articulators when processing the speech signal (Erber, 2003; Sommers, Tye-Murray, & Spehar, 2005). In 1954, Sumby and Pollack used 256 spondaic words (i.e. baseball, cupcake) as well as supplementary tests with monosyllabic words, spondee words, and trisyllabic phrases. Sumby and Pollack (1954) delivered noise over a gas-tube that was mixed electrically with the speech signal. The speech signal and noise were delivered over various signal-to-noise ratios. Sumby and Pollack (1954) determined that there was an 80% gain by adding in the visual information of seeing the speaker's face when the signal to noise ratio was -30 dB. Harry Mcgurk and John Macdonald (1976) studied the effects of audiovisual speech integration. They determined that the visual information aids in understanding of the message. When the participants only received audio information, they had difficulty discriminating between /ba/ and /ga/. However, when they added the visual component, the participants accurately discriminated between the two syllables (Mcgurk & Macdonald, 1976). These two studies are important to the study of audiovisual speech integration in that they relay the fact that visual information is important when listening to a speaker's message.

However, a person's age and cognitive functioning, as well as hearing ability, can hinder their ability to hear and understand a spoken message. As people age, their ability to hear often declines—a condition called presbycusis (Tye-Murray, Sommers, & Spehar, 2007). In turn, this affects their ability to hear spoken language. In 2005, Sommers, Tye-Murray, and Spehar conducted a study that studied the effect of age audiovisual speech integration. The authors studied 44 older adults and 38 younger adults to identify vowel-consonant-vowel syllables, words in carrier phrases, and sentences using auditory-only, visual-only, and audiovisual trials. The signals were degraded using multi-talker babble. When controlling for auditory performance across age groups, results indicated that the older adults performed poorer in the audiovisual trials due to reduced speechreading abilities—not impaired integration abilities. Additionally, Bergeson, Pisoni, Reese, and Kirk, (2004) studied rate of hearing loss and audiovisual (AV) speech integration abilities. They determined that postligually deafened individuals with a cochlear implant (CI) with progressive and sudden hearing loss performed the best in the AV area. The individuals with progressive hearing loss performed better than the individuals with sudden hearing loss in the visual only condition, indicating the rate of hearing loss correlates with lipreading abilities (Bergeson, et al., 2004).

Pharmacological factors may also affect sensory and cognitive performance—including audiovisual speech integration. Interestingly, evidence from case studies has suggested that chronic opioid pain medication use may contribute to hearing loss (Nair, Cienkowski, & Michaelides, 2010), and deficits in certain cognitive functions such as working memory and executive functions (Rapeli et al., 2006). The impairment of these functions has the potential of impacting these individuals' ability to perceive auditory speech information and, in turn, affect their ability to communicate. Significantly, opioid drugs are becoming increasingly prescribed to manage cancer pain as well as non-cancer pain (Smith & Bruckenthal, 2010), which may lead to an increase of sensory impairments among these individuals—although the extent of the impairment is still under debate.

However, there are no controlled studies, to our knowledge, regarding the effects of potentially ototoxic opioid drugs on multimodal speech integration, and speech perception abilities in general. With the increased use of opioid drugs for pain management, it is important to systematically investigate how such drugs may affect speech recognition capabilities. The reason is that speech can be perceived in many ways: through the auditory system, through the visual system, and through a combination of the auditory and visual systems. The following sections will discuss factors already known to be associated with hearing loss. These include age-related hearing loss and hearing loss caused by chronic or short term medication use.

Age and Hearing Loss

Over 31 million non-institutionalized American adults are affected by hearing loss, one-third of who are over the age of 65 (Fausti, Wilmington, P. Helt, W. Helt, & Konard-Martin, 2005). It is well known that hearing loss often accompanies aging (Fausti et al., 2005; Spehar, 2005; Tye-Murray et al., 2007). This condition is referred to as presbycusis and typically affects listeners between 70 and 80 years of age (Jonsson, Rosenthall, Guase-Nilsson, & Steen, 1998, Tye-Murray et al., 2007). This type of high-frequency hearing loss happens to be the most common type of sensory impairment in the elderly (Dayasiri et al., 2011). Presbycusis is distinguished by its bilateral sloping audiogram (Pronk, Deeg, & Kramer, 2013) and is *sensorineural*—meaning the hair cells in the cochlea are damaged and cannot be rehabilitated (Fausti et al., 2005). Individuals with presbycusis generally have age-normal hearing in the low frequencies then gradually worsening hearing starting between 2000 and 3000 Hertz (Hz), reaching between a 60 and 80 decibels (dB) hearing loss at 8000 Hz (Jonsson et al., 1998).

Overall, approximately forty million working Americans have sensorineural hearing loss due to excessive noise exposure (Fausti et al., 2005). Similar to age related hearing loss, noise induced hearing loss also shows evidence for an age-normal audiogram for low frequency tones; however, the high-frequency range typically show a notch in which thresholds increase, and then decrease again (Fausti et al., 2005). The specific frequencies and decibel levels vary with each person and length of noise exposure (Fausti et al., 2005). For example, elderly veterans were often exposed to frequent loud noises, which makes them more susceptible to developing noise-induced hearing loss in addition to developing presbycusis (Fausti et al., 2005).

Losing the ability to hear speech sounds in the higher frequency range due to presbycusis and noise-induced hearing loss often makes speech sounds difficult to understand (Erber, 2003; Fausti et al., 2005; Humes, 2013). Fausti et al. (2005) described the configuration of the audiogram for presbycusis, noise-induced, and medication-induced hearing loss. The authors described it as a sloping, high-frequency configuration. Fausti et al. (2005) described this type of configuration as affecting the high-frequencies first, then progressing toward the lower frequencies. The most common complaint associated with this type of audiogram is that an individual can hear the other person talking but the speech sounds muffled (Humes, 2013), which is consistent with the characteristics of high frequency hearing loss. Erber (2003) explained that when a person's hearing in conversational speech becomes impaired at about 35 dB below 1000 Hz. Erber (2003) states that an individual with high frequency hearing loss is likely to compensate by watching the speaker’s mouth movements and gestures to aid in comprehension of the message. However, these individuals may have age related vision loss which would hinder their ability to compensate for their hearing loss in conversation by not being able to perceive the oral and facial cues that accompany speech (Erber, 2003). Presbycusis can impact a person socially and emotionally as it can lead to social and emotional loneliness (Dayasiri, et al., 2011; Erber 2003). By being unable to hear clearly, individuals find it more difficult to communicate and may shut down.

It is important to investigate the extent to which opioid medication usage affects hearing and whether or not there are similar affects between opioid-related hearing loss and presbycusis. The following section will therefore discuss hypothesized associations between hearing loss, and opioid pain medication use.

Opioid Medication and Hearing Loss

In addition to aging and noise exposure, certain classes of drugs can cause hearing loss (Fausti et al., 2005; Roizen, 2003). Many ototoxic drugs that treat cancer vigorously and adversely affect a person’s bilateral sensorineural hearing (Fausti et al., 2005; Roizen, 2003). Fasusti et al., (2005) described that hearing loss affects 31 million Americas, and in particular, veterans who were exposed to loud noises and ototoxic drug treatments. They determined that a significant amount of patients who receive ototoxic drugs obtain some degree of hearing loss and tinnitis. While older adults are certainly at risk for hearing loss due to medication use, it is also possible for infants to develop sensorineural hearing loss due to ototoxic drugs taken by their pregnant mothers or by medications they are placed on after birth (Roizen, 2003). Similar to age related hearing loss (i.e., presbycusis) hearing loss due to drug ototoxicity is typically progressive beginning in the high-frequencies (Roizen, 2003). Sensorineural hearing loss has been found in infants who were prenatally exposed to alcohol, trimethadione, methyl mercury, and iodine deficiency (Roizen, 2003). Roizen (2003) also mentioned that older children who are being treated on cis-platinum for cancer may develop hearing loss as well.

As previously discussed, hearing loss can greatly affect communication by interfering with a listener's ability to understand speech sounds, especially in a noisy environment. This in turn, hinders a patient's quality of life after cancer treatment (Fausti et al., 2005).

Audiovisual Speech Integration

Hearing loss is associated with the impaired ability to hear speech in noisy environments, such as restaurants and those with competing talkers (Adams, Gordon-Hickey, Moralas, & Moore, 2011). Interestingly, individuals with hearing loss typically require more time to process the content of an utterance (Adams et al., 2011), although they may still efficiently integrate auditory and visual signals when accuracy or speed are considered (Altieri & Hudock, 2014a; 2014b) efficiently integrate auditory and visual signals speech integration (Altieri & Townsend, 2011). As previously discussed, when the auditory sensory modality is compromised by hearing loss, noise, or a combination of these factors, a listener's reliance on the visual modality becomes crucial during face-to-face communication. This is reflected by benefits measured by a benefit in speed and accuracy that one gets from viewing the visual signal (Altieri & Townsend, 2011; Altieri et al., 2013; Erber, 2003; Grant et al., 1998; Sumby & Pollack, 1954). Specifically, hearing-impaired individuals may perform better (i.e. faster, better, and/or more accurately) when auditory and visual stimuli are presented together (AV), followed by when auditory only (A-only) stimuli are presented then visual only (V-only). (Altieri & Hudock, 2014; Grant et al., 1998; Sommers, Tye-Murray, & Spehar, 2005; Tye-Murray, Sommers, & Spehar, 2007; Winneke & Phillips, 2011). As previously stated, Sumby and Pollack (1954) concluded that individuals can receive up to an 80% gain with increased signal to noise ratio when adding the visual component to auditory information. This corresponds to the equivalent of a 15 dB increase in the auditory signal. Mcgurk and Macdonald (1976) determined that integrating audio and visual components is important for speech perception. When the participants only had auditory input, they reported hearing /baga/ or /gaba/. However, when they listened to the auditory stimuli and watched the visual stimuli together, they appropriately identified the syllables the speaker was saying. With opioids potentially affecting hearing, opioid users may have to depend heavily on their audiovisual speech integration abilities to understand a speaker.

Before describing the study and hypotheses, we shall briefly turn our attention to discussion on how to assess AV processing. The following sections provide more detailed information regarding measures of audiovisual speech integration in normal-hearing and hearing-impaired listeners alike and focuses on measures of accuracy, speed, and capacity. Examples from the literature are provided where appropriate.

Measures of Audiovisual Speech Integration

Gain Scores and Accuracy Measures

Audiovisual (AV) speech integration has been measured in different ways. In one of the first AV integration experiments, Sumby and Pollack (1954) used response rate to determine individuals' integration abilities. A warning light would appear one second before a spondee word was presented in order to prepare the participants. 129 participants were used in this study. They wore headphones to transmit background noise and the speech signal. Half of the participants faced the speaker while the other half faced away. They were required to write the word they perceived as quickly as possible. If they were uncertain of the word, they were to choose from a restricted word bank. The researchers determined that speech intelligibility decreased as background noise increased. However, they discovered that lip reading is an integral part of speech recognition in noise; the benefit is especially noticeable when there is a low speech-to-noise ratio (Sumby & Pollack, 1954). More current measures of integration are described in the following sections:

Age and Multisensory Integration

Grant et al. (1998) measured consonant and sentence recognition with auditory (A) only, visual (V) only, and audiovisual (AV) stimuli in hearing impaired individuals. Grant et al. (1998) were trying to determine the most important cues for AV speech recognition that could be taken from each A and V speech signals in consonant recognition. They used 18 consonants in a vowel-consonant-vowel pattern (aCa). The stimuli were presented in A-only, V-only, and AV trials, and the participants were to select the appropriate consonant from a set of consonants. Grant et al. (1998) determined that the participants benefitted from speechreading. Second, the researchers were trying to determine whether a person's ability to integrate separate A and V cues could be measured separately from their ability to recognize syllables, words, and sentences. Finally, Grant et al. (1998) wanted to identify the most important "top-down" processes that are not related to signals that could contribute to variability among individuals in AV speech recognition in sentences. They used the same process as for consonant recognition. They used phonetically balanced, low-context sentences each containing five key words. Grant et al. (1998) determined that each participant’s AV score was higher than both their A-only and V-only scores. They discovered that participants with better V sentence recognition performed better on the AV sentence recognition because these individuals may be better at extracting visual cues such as word stress and segmentation. Another possibility as to why some individuals are better speech-readers than others could be that they have higher-level cognitive skills which work in addition to the bottom-up information that is extracted from speechreading (Grant et al., 1998). These individuals likely form linguistic wholes from the fragments they perceive, which is called *perceptual closure* (Grant et al., 1998).

Sommers et al. (2005) sought to determine the extent to which age had an effect on combining audio and visual signals, and moreover, whether there was a difference between recognition of consonants, words, and sentences. They presented their participants with A-only, V-only, and AV trials of each vowel-consonant-vowel syllables, words following carrier phrases, and meaningful sentences with background babble that was configured individually to each participant so they would each receive 50% correct on the A-only trials. Similar to the Grant et al. (1998) study, Sommers et al. (2005) tested consonant recognition in a vowel-consonant-vowel shape /iCi/. For sentence recognition, the carrier phrase “Say the word ‘___.’” Audio and visual enhancement (AE and VE respectively) were measured with the following formula:

VE = (AV – A)/1 – A

which was adapted from Grant and Seitz (1998), Grant et al. (1998), Rabinowitz, Eddington, Delhorne, and Cuneo (1992). This method compares VE from a wide range of

A and V scores (Sommers et al. 2005). Here, VE measures the extent to which the provision of visual information enhances processing compared to auditory-only processing. Here, AV-A is divided by (1-A) since it is measures "visual enhancement" relative to the total benefit obtainable via the visual processing. The researchers also calculated auditory enhancement, AE, using the formula:

$$AE = (AV - A)/1 - V.$$

Although less common than measuring VE, measuring AE normalizes for the age differences that may be seen in V performance, therefore is less likely to be affected by unimodal performance differences because of the A accuracy across ages. Sommers et al. (2005) used age as a covariate when implementing an analysis of covariance (ANCOVA) to further control for differences in unimodal performances such as age differences in A or V scores.

To summarize the results, it was observed that consonants were identified better than words, and words were identified better than sentences. As expected, the authors also determined that scores for AV stimuli were significantly higher than A, and A stimuli were identified significantly better than V stimuli. Additionally, they also determined that the younger adults scored better overall than the older adults. However, Sommers et al. (2005) determined that there was a lack of significant differences for the age-groups after controlling for A-only performance across younger and older listeners. Sommers et al. (2005) determined that the AV scores for older adults in the domains of consonant and word recognition was poorer than for younger adults. However, older adults performed better in AV sentence recognition that the younger adults due to the V-only differences across the group (Sommers et al., 2005). One consequence of this finding was that the authors argued that V performance is controlled in part by common mechanisms that are independent from lexical or semantic constraints. Sommers et al. (2005) determined that the mechanisms that mediate AV perception of sentences and words may be independent from the mechanisms used for AV perception of consonants in older adults. They determined this because all correlations between V performances were significant except for the correlation between consonants and words in older adults. This could be due to changes in response times as a person ages. The results from this study suggest that age reduces speechreading abilities rather than integration abilities.

Finally, when measuring the correlation of AE and VE, Sommers et al. (2005) determined that the correlation of VE for words and sentences was not significant for younger adults, but was significant for older adults, and VE was not significant and had a negative correlation for consonants and sentences. This could be due to impaired speechreading abilities. Some consonants may look similar, especially to someone with impaired speechreading abilities. As for AE, correlation was not significant for younger adults between consonants and sentences but was negative and significant in older adults. The authors therefore concluded that the mechanisms that mediate AE and VE in consonants are different than the mechanisms used in AE and VE in words and sentences in both younger and older adults. Due to the correlations observed, the researchers determined that the participants were able to improve recognition of words and sentences when both auditory and visual information were present. However, Sommers et al. (2005) argued that it is likely that aging is responsible for declines in crucial capacities that are responsible for encoding V information which are independent of hearing.

Similarly, Tye-Murray et al. (2007) measured performance on recognizing consonants, words, and sentences in A-only, V-only, and AV trials in background babble adjusted for equivalence for each participant. The researchers tested adults with normal hearing and adults with mild-moderate hearing loss. They used a computer screen for the visual presentations and headphones for the audio presentations.

Tye-Murray et al. (2007) adopted the formula for assessing audio and visual enhancement from Sommers et al. (2005). To assess integration, Tye-Murray et al. (2007) used the Prelabeling (PRE) model which was developed by Braida (1991), which uses consonant confusion error patterns from both A and V trials to predict the AV performance of the participants. They also used the following formula adopted from Blamely et al. (1989):

$$(p)\ AV\ Predicted = 1 - ((1 - (p)A)(1 - (p)V))(3)$$

This equation depicts that A and V perception are independent from each other. Integration enhancement is then calculated using the following formula:

$$IE = (p)\ AV\ Observed - (p)\ AV\ Predicted / (1 - (p)\ AV\ Predicted)(4)$$

Tye-Murray et al. (2007) determined that there was no significant difference between the two groups in A consonant performance. However, performance in the normal hearing group for A word recognition was significantly higher, but the hearing impaired (HI) group scored higher in the sentence recognition task. The HI group scored significantly better than the normal hearing (NH) group when using V stimuli in words, but there was no significant difference in consonants or sentences (Tye-Murray et al., 2007). Tye-Murray et al. (2007) also determined that hearing loss did not affect visual only presentations of viseme categories of consonants.

Tye-Murray et al. (2007) determined that the HI group performed better than the NH group with the AV stimuli in sentences but not consonants and words. They also determined that both groups received similar benefits for AE and VE when the A and V stimuli were combined.

Tye-Murray et al. (2007) used two approaches to determine whether there was a difference in integration abilities between the HI and NH groups. For the first approach, they chose individuals who differed by less than 3.4% on both A and V tasks. They found that the individuals had similar integration abilities regardless of their group. The second approach Tye-Murray et al. (2007) used to measure integration was to compare the predicted and actual AV performance using IE (Sommers et al., 2005) and the PRE (Braida, 1991). For all three stimulus types, the predicted scores were significantly lower than the obtained scores. Between the two groups, no significant differences were found between consonants and words, but for sentences, the participants with HI exhibited significantly higher scores than the NH group. Tye-Murray et al. (2007) determined that although the HI group performed better on the AV sentence recognition tasks that does not necessarily mean they are better integrators.

Processing Speed

More recently, researchers have begun to investigate processing speed as a component of multisensory integration ability. Traditionally, processing speed has been ignored as a viable measure, which is unfortunate since speech is a dynamic process that requires the listener to recognize features in real-time.

Hearing loss and Multisensory Integration

Reaction time and Capacity measures

Response Time

Winneke and Phillips (2011) studied 20 younger adults (YAs) and 19 older adults (OAs). However, three YAs and two OAs were excluded due to poor behavioral performance which differed more than two standard deviations. All had age appropriate sensory abilities as well as appropriate cognitive functioning. Eighty spoken object names served as the stimuli, split into two equal groups of natural and manmade objects. The words were recorded of a woman speaker's head, face, and neck. The researchers used three different presentation conditions: auditory-only (A-only), visual-only (V-only), and audiovisual (AV) speech. In each condition, the participants also received a multitalker babble mask which was individually adjusted for each participant to ensure the auditory information was the same across the participants. The participants were to respond as quickly and as accurately as possible as to whether the word was a natural or manmade object. Winneke and Phillips (2011) also measured brain activity through continuous EEG. Winneke and Phillips (2011) determined that responses in the $(AV)_{speech}$ trials were more accurate than the A-only trials, which were more accurate than the V-only trials. OAs performed worse in the V-only trials than the YAs, which indicates poor lipreading abilities. Winneke and Phillips (2011) also measured response time (RT) of the participants using race model analysis, which states that trials in the multisensory domain are not faster than the fastest unisensory response (Miller, 1982). The race model is violated when the probability *(p)* of one response time is higher in the multisensory condition than in the unisensory condition, i.e., $p(AV) > p(A + V) - p(A \text{ X } V)$, which suggests interaction between two sensory channels (Miller, 1982). Winneke and Phillips (2011) placed the RT interval from 0.4-s to 2.5-s into time bins of 10-ms. They then calculated the likelihood that a response happened at a RT or faster by plotting them as cumulative distribution function (CDFs). The data was then analyzed by paired t-tests at each time bin for a race model violation. The CDF values for RT in YAs and OAs were higher in the $(AV)_{speech}$ trials than in the joint probability for unisensory responses in the 590 ms-1240 ms time bins, which violates the race model. This indicates that the faster RTs in the $(AV)_{speech}$ trials were likely because of the two unisensory information channels interacting (Winneke & Phillips, 2011). The researchers suggested that the information gathered from the RTs and accuracy measures indicate that fewer neural resources were used to achieve better performance. This suggests that $(AV)_{speech}$ is more efficient than unimodally.

Altieri and Townsend (2011) outlined basic experimental designs and how they can be used when carrying out double factorial experiments. They also used RT based capacity measures to show evidence that auditory and visual cues are processed in parallel, with interactions between channels. Basic experimental designs involve identifying targets in one or two channels (involving the direction of attention). The participants were required to simply answer yes or no to a target under either single target trials with only an auditory tone, single target trials with only a visual dot presented, redundant target trials where both were presented, and target-absent trials. Basic experimental designs include accuracy and reaction time (RT); however, these types of experimental designs require many trials to obtain measurable RT. Altieri and Townsend (2011) noted that this basic experimental design can be adapted for further research. The first factor manipulated concerned the number of targets present or channels available (A- or V-only versus AV). Altieri and Townsend (2011) stated that this is important for calculating capacity, which they define as processing efficiency as a function of channels available. They argued that accuracy-only measures are typically only used when comparing performance between populations such as children with cochlear implants or elderly adults with hearing impairment to normal hearting listeners. The second factor manipulated was the saliency of each of the channels. The auditory and visual signals would be manipulated to induce faster and slower reaction times with high and low levels of saliency, respectively (Altieri & Townsend, 2011). Altieri and Townsend (2011) note an important assumption to this methodology, *selective influence*. This refers to the notion that the salience manipulation must be effective when changing the processing speed for a sub-process. In their experiment, Altieri and Townsend (2011) assessed selective influence by checking the empirical cumulative distribution functions (CDFs or survivor functions) in each condition and specifically ordered them. The empirical survivor functions yield a finer interaction measure than measuring RTs alone. Survivor functions, however, provide the probability that a process has not been finished by time *t*.

Altieri and Townsend (2011) assessed the workload capacity $C(t)$ of the system, which measures how the number of channels affects efficiency time *t* of processing. They wanted to determine if there was cost, benefit, or no change when both auditory and visual channels were present. To calculate $C(t)$, the hazard function $H(t) = -\log(S(t))$. The equation is as follows:

$$C(t) = H_{AV}(t)/[H_A(t) + H_V(t)]$$

In their pilot study, Altieri and Townsend (2011) investigated the decision rule and processing architecture in discriminating between two words ("base" and "face") with A-only, V-only and AV signals. They also assessed the processing capacity. They studied three conditions using an auditory signal-to-noise ratio of -18dB, S/N of -12 dB, and the clear S/N ratio.

Wenger and Gibson (2004) described capacity as how much work the person who is observing can perform within a certain amount of time. Traditionally, capacity has been measured as the level of average response time (RT) or mean accuracy (Wenger & Gibson, 2004). In 2014, Altieri and Hudock described capacity, or *C(t)*, as a RT-based measure that assesses the extent to which listeners efficiently combine different sources of information. They described that it measures processing speed as the workload changes. They wanted to determine whether having visual information contribute to a more efficient use of the cognitive-linguistic resources. Altieri and Hudock predicted that *C(t)* will be better than the accuracy-only measures because processing speed is taken into account with measuring *C(t)* and because the audiovisual trials can be compared to the well-defined parallel independent race models. C_I*(t)* is a capacity measure that uses both reaction and accuracy. Altieri and Hudock (2014) showed that high and low frequency hearing loss were predicative of capacity scores. Altieri and Hudock (2014) stated that the parallel independent race models provided hypotheses where auditory and visual information do not interact and are independently processed.

Altieri and Hudock (2014) conducted a study that investigated the association between hearing ability, as measured by an audiometer, and audiovisual speech integration as measured by speed and accuracy. They used a capacity approach to attempt to answer questions that the studies only measuring accuracy missed. Alieri and Hudock (2014) also indicated that measuring gain may be problematic because there is a possibility of a ceiling effect because they are correlated with performance with one sense. Altieri and Hudock (2014) investigated the gain from AV input in sentence recognition, as well as hearing ability in relation to audio-only and visual-only sentence recognition. Altieri and Hudock (2014) had randomized A only, V only, and AV trials to measure word and sentence recognition. In the first experiment, the researchers were investigating how traditional measures of AV gain using accuracy correlated with ability of hearing. The measure the researchers used for gain was:

$$AV_{GAIN} = p(AV) - \max\{p(A), p(V)\}$$

Experiment 2 was where they were measuring capacity of integration. The speech in the A and AV trials was distorted similar to how an individual who uses a cochlear implant hears instead of babble because the babble may mask certain auditory signal cues. First, they observed a significant positive correlation between low and high frequency pure tone thresholds and AV gain which indicates—this indicates that poorer hearing predicts AV gain in terms of accuracy. They determined that low frequency ability may be a better predictor of capacity than accuracy, suggesting that capacity measures may be better than accuracy measures. Altieri and Hudock (2014) suggest that the capacity measure utilizes both speed and accuracy, if the integration abilities of normal hearing and hearing impaired listeners, and can determine if speed, accuracy, or both determines the locus of the integration potential.

This thesis will focus on opioid usage and its effects on audiovisual speech integration using comprehensive speed and accuracy based measures of multisensory processing capabilities. This study is important because of the increased use of prescription opioid drugs, and the unknown effects that these drugs may have on multimodal communication abilities. Based on the above statements, the following question is raised: To what extent do potentially ototoxic opioid drugs affect multimodal speech integration and speech perception in middle aged adults" This study will examine participants' ability to identify audio only, visual only, and audiovisual stimuli and the capacity to which they identify these stimuli.

Hypotheses

It is hypothesized that opioid usage will be associated with an increase in capacity, and hence, speech integration abilities overall due to decreased hearing abilities. A moderate correlation between C_I(t) and audiometric thresholds was recently shown by Altieri and Hudock (2014). Therefore, for the first experiment, it is hypothesized that the experimental group will demonstrate higher levels of gain due to their possible decreased hearing abilities. They will demonstrate greater accuracy in combined AV trials than the control group. In experiment two, it is hypothesized that the experimental group will demonstrate decreased RTs overall, but higher capacity.

Methods

Participants

Participants of this study will include 20-30 middle aged adults between the ages of 40 and 60 years of age. The participants will be obtained via random sampling through the Health West, Inc. database in Pocatello, ID. In order to qualify, the participants must have been using prescription opioid pain medication for six or more months, be within 40 to 60 years of age, and have normal or corrected-to-normal vision. A control group of 20 to 30 participants will also be selected via word of mouth and study postings online in Pocatello, Idaho. The control group will consist of aged matched peers with no history of opioid use and normal or corrected-to-normal vision.

Materials

The materials used for both Experiments 1 and 2 have been obtained from Altieri and Hudock (2014a; 2014b).

Sentence recognition: A flat screen Dell computer with a refresh rate of 60 Hz will be used to present the stimuli. *Beyer-Dynamic* 100 headphones will be used to present the auditory stimuli. The sentence stimuli will consist of 75 sentences obtained from the CUNY sentence database (Boothroyd, Hanin, & Hnath, 1985). Each of the sentences will be spoken by a female talker and consist of 25 auditory-only and visual-only stimuli and 25 audiovisual stimuli. The stimuli will be obtained from a laser video disk and made into a 720 x 480 pixel video at a digitized rate of 30 frames per second. The audio files will be presented at 48 Hz. The auditory stimuli will be removed using Adobe Audition for visual-only stimuli. Stimuli were presented using E-Prime 2.0 (http://www.pstnet.com/eprime.cfm) software. The sentences will be subdivided into 3, 5, 7, 9, and 11 words with five sentences per length for each stimulus set in order to simulate the varying sentence lengths in normal conversation (Altieri et al., 2011). The sentences can be found in the Appendix.

Speeded word recognition: The materials used for this study will be adopted from Altieri and Hudock (In Press). The materials will consist of videos of two female talkers obtained from the Hoosier Multi-Talker Database (Sherffert, Lacks, & Hernandex, 1997). The following monosyllabic words were obtained from the talkers: *Date, Page, Gain, Shop, Boat, Tile, Job,* and *Mouse.* Each word will be distinct in the auditory, visual, and audiovisual domains. Each video will be edited using *Adobe After Effects C4* and made into a 720 X 480 pixel clip at a rate of 30 frames per second. The audio files will be sampled at 48 kHz (16 bits) and the auditory signal will be degraded using the 8-channel sinewave CI simulator and played over *Beyer-Dynamic* 100 headphones. The range of the audio, visual, and audiovisual files will range from 800-1000 ms. Stimuli will be presented using E-Prime 2.0 (http://www.pstnet.com/eprime.cfm) software.

Procedures

Sentence recognition: Trials will be presented in separate blocks containing 25 auditory-only, 25 visual-only, and 25 audiovisual trials. A blank grey screen will be shown for each auditory-only trial. The trials will be randomized across participants to avoid order effects. Participants will be seated about 24 inches away from the computer screen. A black dot on a solid gray background was presented which will cue the participant to press the space bar to begin. The trial will begin with the talker speaking one of the sentences. The participants are to type in what they thought the talker said in the dialog box that will appear after the sentence is complete. Each sentence will be presented only once and no feedback will be given.

Speeded word recognition: The audiovisual, audio-only, and visual-only trials will be randomly presented in one block. There will be 128 audiovisual trials (8 words x 2 talkers x 2 recordings x 4 repetitions), 128 auditory-only trials, and 128 visual-only trials. For the auditory-only trials, a blank gray screen will be presented. A black dot on a solid gray background will be presented to signal the beginning of the trial. Each trial will consist of auditory-only, visual-only, or audiovisual stimuli, and the auditory stimuli will be played at about 70 dB SPL.

Data Collection and Instrumentation

Sentence recognition: Scoring will be adopted from Altieri and Townsend (2011). Whenever the participant types in a word correctly, that word will be scored as correct. Word order will not be a criterion for the word to be counted correct and misspelled words will be manually corrected. The proportion of words correct will be scored for each sentence. The independent variable in the speech recognition portion of the study will be opioid use, and the dependent variable will be the amount of AV gain.

The measure of gain, which will be adopted from Altieri and Wenger (2013), is: $AV_{Gain}=p(AV)—max\{p(A),p(V)\}$. The relationship between opioid usage and the AV gain in the CUNY sentence recognition will be demonstrated in this portion of the study. Information from this analysis will demonstrate the extent to which both the control group and the experimental group identify appropriate responses in audio-only, visual-only, and audiovisual trials.

Speeded word recognition: Responses will be recorded using the computer keyboard. The numbers one through eight will be matched with a label with each word from the stimulus set. The participants will be instructed to press the button corresponding to the word "as quickly and accurately as possible," and their responses will be timed from the onset of each trial. Responses will be based on the auditory-only, visual-only, and audiovisual stimuli. Auditory only trials will be played with a blank computer screen, and visual-only trials will be played without sound coming through the headphones. Inter-trial intervals will randomly vary from 750-1,000 ms. Forty-eight practice trials will be presented at the beginning of the experiment and feedback will be provided as "correct" or "incorrect" for the practice trials but not for the experiment.

Capacity (i.e. "change in workload") will be measured using the empirical function ($C\_I(t)$ and $C(t)$). The independent variable in the speech recognition portion of the study will be opioid use, and the dependent variable will be the capacity for multimodal speech integration. Information from this analysis will demonstrate the extent to which individuals in the control and the experimental groups integrate AV information and how much gain is received from combining modalities. Individuals who receive a capacity value larger than one will be considered faster and more accurate speech integrators. The listeners who receive a capacity value less than one will be considered poor integrators and do not benefit from visual cues.

Results

Sentence Recognition

The results include data obtained from 25 adult control participants, all without a reported history of neurological impairment. Qualitatively, the results indicate that audiovisual accuracy scores were near ceiling (90-100%); both auditory and visual-only scores showed evidence for greater individual variability compared to audiovisual scores. Additionally, auditory-only scores were lower than audiovisual accuracy. On average, visual-only accuracy scores were generally low (under 20% correct); this in particular reflects the inherent difficulty of lip-reading (see Altieri, Pisoni, & Townsend, 2011).

Figure 1A shows box-plots of the accuracy scores for the auditory-only (A-only), visual-only (V-only), and audiovisual (AV) CUNY sentence recognition experiment. The upper and lower portions of the box-plots denote the 75th and 25th percentiles, respectively. The middle line represents the median, and the whiskers represent 1.5 times the interquartile range. The + sign indicates statistical outliers that lie beyond 1.5 times the interquartile range. The mean accuracy scores are as follows: A-only, 76.00% (SD = 6.62); V-only, 11.86% (SD = 6.64); and AV, 94.64% (SD = 3.16). Overall, more variability was observed in the V-only conditions when compared to the A-only and AV conditions.

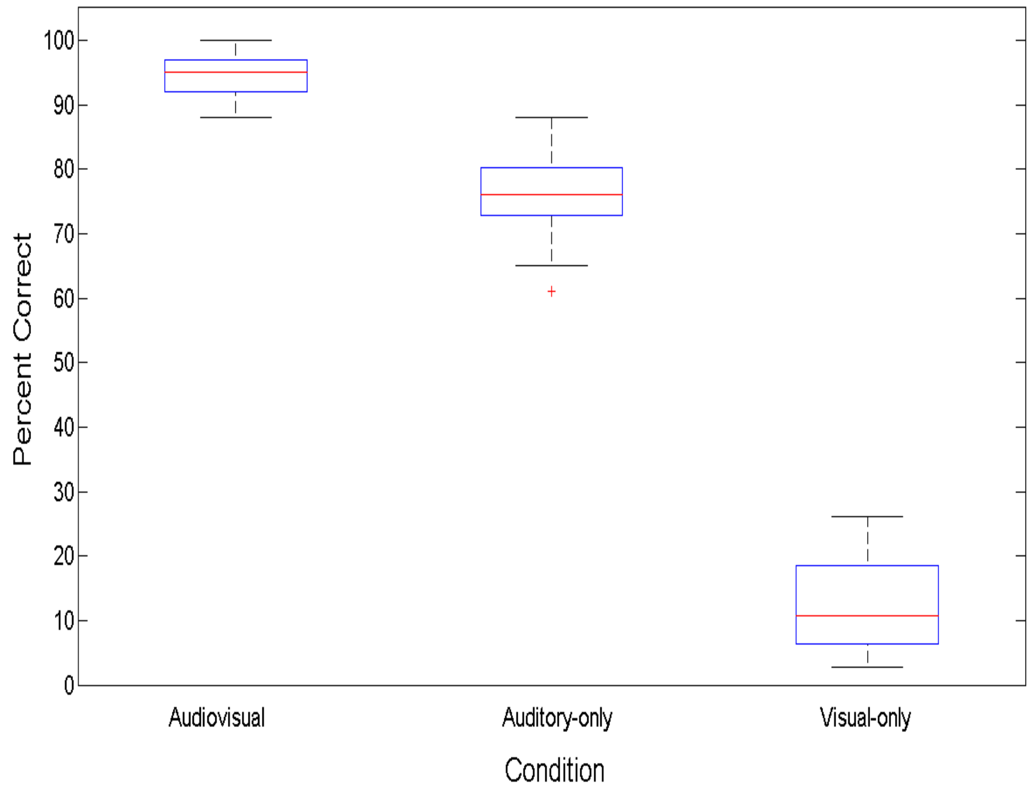

Figure 1A. Box plots showing sentence recognition accuracy scores.

Measures of Audiovisual Gain

Next, audiovisual gain scores are displayed in the box-blots in Figure 1B. Gain scores were used to demonstrate the extent to which the participants combined the auditory and visual information relative to predictions derived from unisensory conditions. Scores above 0, for both measures described in the following paragraph, would indicate that the listener “integrated” the audiovisual information since independence assumptions were violated. As the results qualitatively indicate, the gain scores for both measures were greater than 0; hence, we assume that the participants are integrating the A and V domains and combining the cues.

To summarize, the first measure of gain, AV- A, compares auditory visual processing relative to auditory alone. This “visual gain” measure computes the benefit of having visual cues present in conjunction with the auditory signal, compared to the auditory signal alone.

For this next measure, the formula AV – [A + V – A×V] is designed to compare auditory plus visual (AV) scores to independent model predictions shown in the brackets. The term in the brackets, A + V, represents the summed probability that a person is correct in either the auditory or visual-only domains respectively, whereas A × V represents the probability that a person is correct in both the auditory and visual-only domains. Hence, the quantity in the brackets “[A + V – A×V]” represents the probability of a listener being correct given the presence of an auditory or visual cue, while assuming independence between the signals. The null hypothesis, in this case independent model, predicts that AV – [A + V – A×V] = 0 because the obtained audiovisual accuracy scores would be equal to race model predictions.

To summarize, we see that Figure 1B demonstrates that since the participants scored above zero, they performed better than independent model predictions and therefore are combining the auditory and visual information to better understand the message. When using the independent model formula of gain, AV – [A + V – A×V], the mean gain across the participants is 15.72 (SD = 6.13). When using the formula AV- A representing visual gain, the mean gain across participants is 18.64 (SD = 7.22). Both measures provide converging evidence that visual information facilitates auditory recognition since independence assumptions were violated.

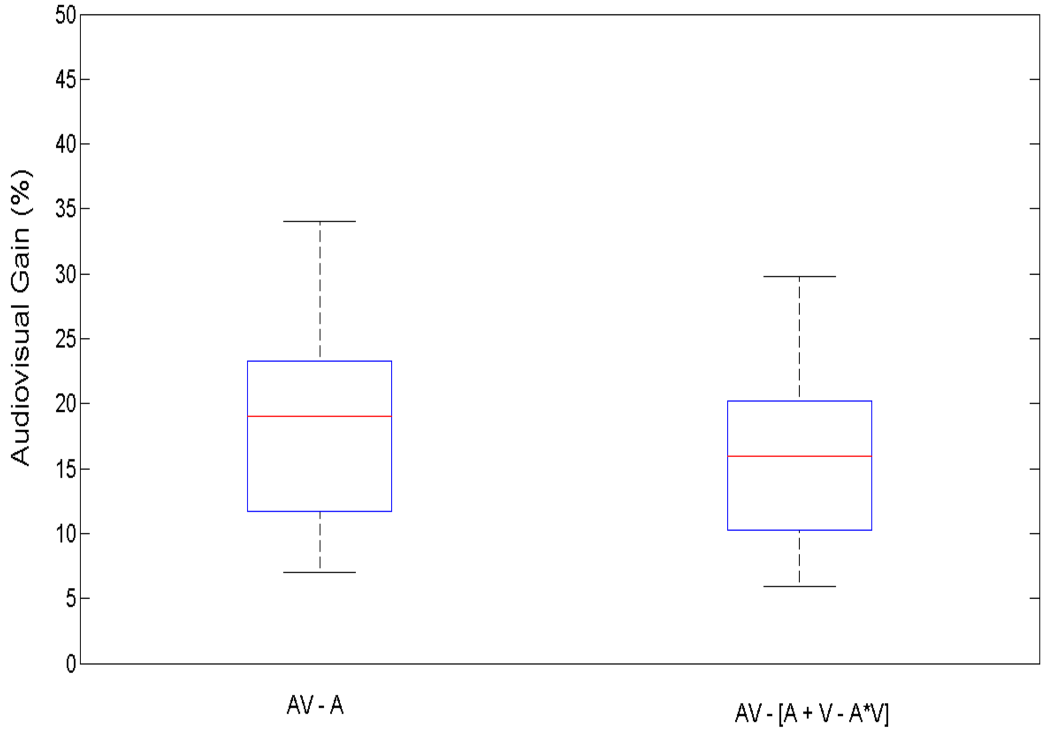

Figure 1B. Box plots showing the distribution of gain scores using both formulas described above.

Speeded-Word Recognition

Reaction Time (RT)

For this experiment, RT measures were used to demonstrate the speed at which the participants responded to the A-only, V-only, and AV stimuli by selecting from a list of eight words in a speeded response task. Table 1 depicts the mean RTs for the A-only, V-only, and AV stimuli. Similar to mean accuracy scores, the findings demonstrate that the RTs for the V-only conditions were significantly slower than the A-only and the AV trials: the RT difference was over 600 ms, indicating that the amount was substantial. Paired samples t-tests were carried out to determine the significance of combining the auditory and visual information. When comparing the RT for the A and V stimuli, the value of the t-test was 1.05E-13, indicating the participants were faster when exposed to the auditory information than when they were only exposed to the visual information. The t-test comparing the AV and V-only stimuli determined that the participants benefitted greatly from the AV stimuli, with a t-value of 1.24E-14. The t-value of the RT of the AV and A-only stimuli was 0.27, indicating that the participants did not significantly benefit from adding the visual information to the auditory stimuli. A more sensitive measure than means, capacity measures will be used to assess audiovisual integration efficiency (Altieri et al., 2014; see also Townsend, 1990).

Table 1

Mean reaction time on speeded word recognition task measured in milliseconds (ms).

	Auditory-only	**Visual-only**	**Audiovisual**
Mean RT (SD)	1871 (202)	2521 (316)	1889 (209)
Mean Accuracy (SD)	98.50(1.16)	75.00(7.14)	99.00(.93)

Accuracy

Accuracy measures, measured in percent correct, were used to determine the correctness of the responses given from the participants in the A-only, V-only, and AV trials. Table 1 above shows the mean accuracy across all three domains: A-only, V-only, and AV in the speeded word recognition task. The V-only accuracy scores were lower when compared to both the A-only and AV scores. The A-only and AV scores were near ceiling, and there was not a lot of variability; this indicates that individuals with normal hearing do not really benefit from adding the visual information when selecting from a list of eight words.

Capacity

Capacity is a non-parametric measure used to examine the rate at which listeners benefit from combing auditory and visual information using RT distributions (*C(t)*), or both RT distributions combined with accuracy (*C_I(t)*). First, the participant-averaged RT-only measure of capacity, *C(t),* was computed by averaging the capacity values across the response times for each participant. The capacity integration function, *C_I(t),* measures the rate in which the participants respond and the accuracy of their responses. *C_I(t)* was measured in a similar way as *C(t),* by averaging all the RTs and accuracy for each participant.

In Figure 2, the right panel plots the RT-only capacity, *C(t)*. The left panel of Figure 2 plots the average *C_I(t)* across the participants. In each case, the blue line represents the capacity function values across different time points, whereas the dotted lines indicate 1.5 standard deviations of the mean. Recall that the benchmark for unlimited capacity (i.e., independence) is "1". Together, the results in Figure 2 demonstrate that listeners processed audiovisual speech signals in manner indicating slightly limited capacity. The interpretation of these results is that the typical adult normal hearing listener did not benefit significantly from the visual information in this task. Interestingly, slower RTs tended to yield more limited capacity, whereas faster response times yielded unlimited or super capacity for integration indicating the faster the response time (and accuracy), the better integrators individuals are.

To briefly review, recall that unlimited capacity occurs when a listener is using the same amount of resources to complete a task with added channels and is represented by the number 1. For example, if a person recognizes a word in either the A or V domains as quickly and accurately as they do in the AV domain, they are thought to have unlimited capacity. Super capacity is when a person has more resources or information to process information and is represented by a number greater than 1. Limited capacity occurs when a listener has to divide their attentional resources in order to recognize information and is represented by numbers less than 1. When looking at *C_I(t)* versus *C(t)* in Figure 2, individuals appear to have similar integration abilities when only RT is taken into account when compared to when accuracy is added in to the measure, although there is less variability in the RT only measures. The standard error of measure (SEM) is represented by the grey lines on Figure 2. The SEM shows that individuals have a range of super capacity to limited with faster RTs and strictly limited capacity with slower RTs. This is the same when looking at both *C(t)* and *C_I(t).*

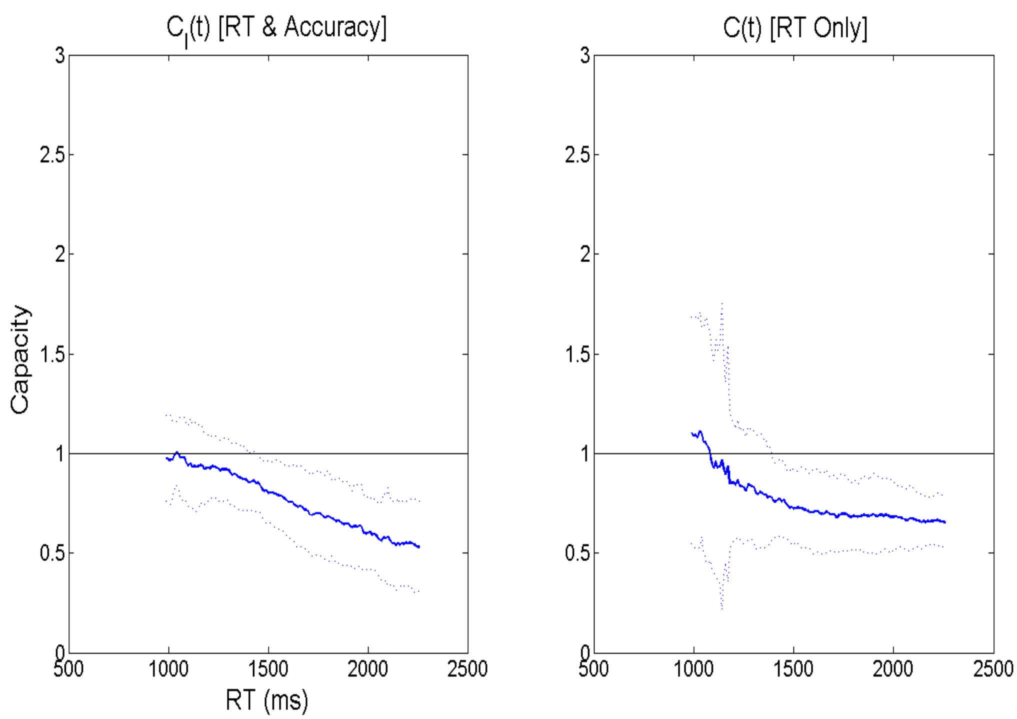

Figure 2: Participant averaged capacity measures. The right panel shows *C(t)* and the left shows *C_I(t).*

Next, summary measures of the continuous capacity coefficient were computed across participants. The continuous capacity measure was used to easier compute statistics on. In total, six summary measures were utilized: mean, median, and maximum *C(t),* and mean, median, and maximum *C_I(t)* to compare the scores of each participant.

The mean *C(t)* represents the average capacity across the participants. The median *C(t)* is the middle number in the data set when scores are arranged from smallest to largest, and the maximum *C(t)* is the highest score among the participants—this latter measure essentially indicates a listener's overall integration potential. The mean, median, and maximum *C_I(t)* measure is similar to the summary C(t) scores; however, recall that this formula also takes in accuracy.

It is important to measure the mean scores for each capacity measure to determine how the participants performed as a whole; this is especially important since capacity is a dynamic measure that can change across processing times. For example, a participant may show evidence of super capacity when responding quickly, although their slower RTs may point to unlimited or even limited capacity. It is therefore important to measure the median for each capacity measure to determine how the median relates to the mean and distribution of scores. Finally, it is important to determine the maximum capacity function to demonstrate the highest score the participant set achieved. Figure 3 shows box and whisker plots that represent the summary *C_I(t)* and *C(t).* Figure 3 shows the mean, median, and maximum *C_I(t)* and *C(t)* using box-plots. Similar to the accuracy scores in the previous section, the middle line represents the median, while the surrounding lines represent the 25^{th} and 75^{th} percentiles. The whiskers represent the 1.5 times the interquartile range, and "+" represents the outliers. Table 2 below shows the distribution of mean, median, maximum capacity of the participants.

Table 2

Mean, Median, and Maximum capacity scores

Mean C_I(t)	**Median C_I(t)**	**Max C_I(t)**	**Mean C(t)**	**Median C(t)**	**Max C(t)**
0.76	0.75	1.13	0.73	0.69	1.46
SD = 0.14	SD = 0.17	SD = 0.20	SD = 0.17	SD = 0.15	SD = 0.72

The boxplot of the mean *C_I(t)* demonstrates that the listeners are not really "integrating" both the auditory and visual signals because their average score is not above 1. Generally, these normal-hearing adult listeners process audiovisual information similarly to auditory-only. However, the box-plot of the maximum *C_I(t)* indicates that some listeners are able to integrate both signals because their integration score is greater than 1—at least for some range of RT values. When looking at RT-only capacity, the mean box-plot of *C(t)* demonstrates that the listeners are still not able to integrate the auditory and visual signals, but when looking at the maximum, the capacity score is 1.46 indicating some listeners are capable of integrating the two signals for a range of RTs (this may be observed in Figure 2 as well for faster RTs). Taken together, the continuous measures of capacity described before in Figure 2 showed results consistent with the summary statistics depicted in box-plots in Figure 3 as expected.

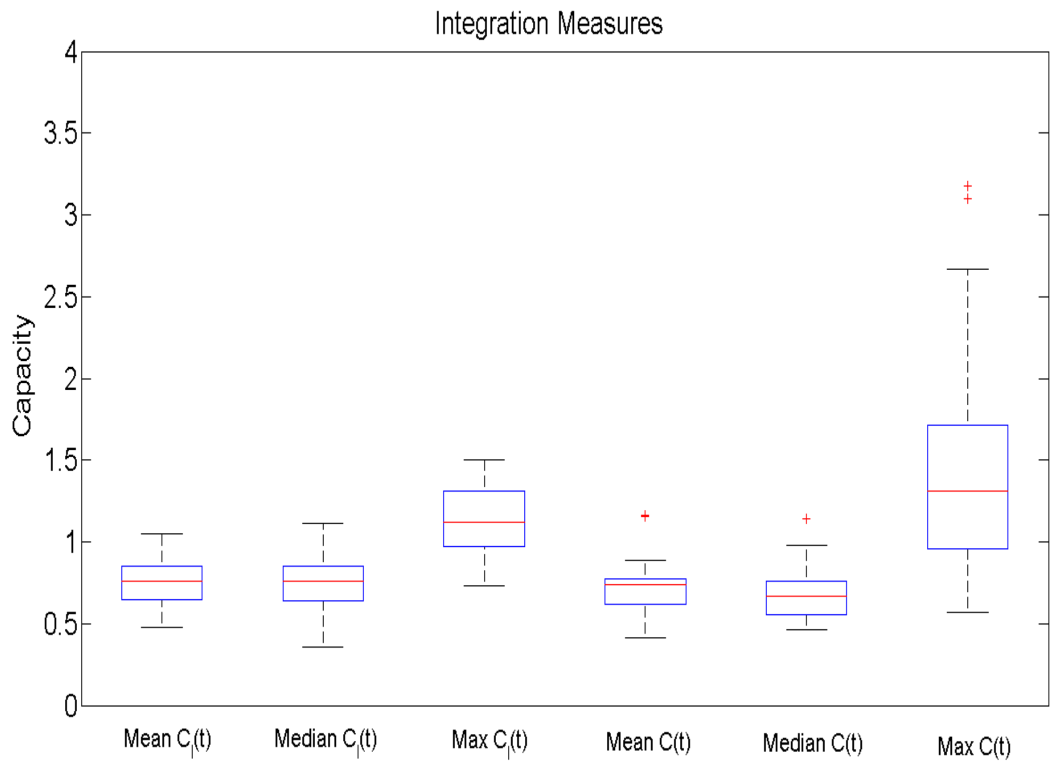

Figure 3: Box plots showing six capacity measures.

Discussion

Audiovisual speech integration abilities have the potential to be affected by many sensory and cognitive factors. For example, pure tone hearing thresholds may affect a person's ability to integrate auditory and speech information. Age related hearing loss, known as presbycusis, causes a person to lose his or her high frequency hearing in both ears, which makes certain speech sounds speech sounds difficult to understand (Erber, 2003; Fausti et al., 2005; Humes, 2013). Individuals with high-frequency hearing loss tend to watch the speaker's mouth to compensate for the lack of auditory information (Erber, 2003). One hypothesis has been that individuals with gradual presbycusis would be better integrators of auditory and visual information since they have learned how to use visual cues to better understand a spoken message (Altieri & Hudock, 2014).

In addition to aging, hearing loss can be induced due to environmental factors, such as the use of opioid medication which was explored in this thesis. Ototoxic drugs are believed to adversely affect a person's bilateral sensorineural hearing thresholds (Fasuti et al., 2005; Roizen, 2003). Drug related hearing loss, such as presbycusis, is typically progressive and begins in the high-frequencies (Roizem, 2003). With this information, it can be thought that people with drug induced hearing loss would also be good integrators of speech as indexed by increased capacity or gain (i.e., benefit in accuracy) from sentence recognition. However, integration abilities may be hindered by possible diminishing cognitive abilities due to the use of certain drugs.

Accuracy Measures

To briefly review, there are several available measures of audiovisual speech integration that use accuracy, reaction time, or a combination thereof. In 1954, Sumby and Pollack used comparisons between audiovisual accuracy and unisensory accuracy to determine an individual's integration abilities. In their experiment, Sumby and Pollack (1954) determined that lip reading represent an integral part of speech recognition in noise. Importantly, Sumby and Pollack's general results were replicated in the CUNY sentence recognition experiment using our 25 control participants.

Similarly, Grant et al. (1998) measured consonant and sentence recognition with A-only, V-only, and AV stimuli with hearing impaired individuals using accuracy measures. They determined that each participant's AV score was higher than their A-only and V-only scores. Going beyond Sumby and Pollack's previous findings, the authors also determined that the participants with better V-only sentence recognition skills performed better on the AV sentence recognition task. They determined that this was likely due to the fact that these individuals might be better at using visual cues such as segmentation and word stress. Grant et al. (1998) said that another reason why some individuals may be good at lip reading is because they have higher-level cognitive skills. These individuals likely use *perceptual closure*, which means they form linguistic wholes from the fragments they perceive.

Sommers et al. (2005) sought to determine whether or not there was a difference between recognition of consonants, words, and sentences in the A-only, V-only, and AV domains as well as audio and visual enhancement (AE and VE, respectively). VE measures the extent to which visual information enhances processing compared to A-only processing using the formula VE = AV – A)/1 – A. AE normalizes the age differences that may be seen in the V performance using the formula AE = (AV – A)/1 – V. Summers et al. (2005) determined that the scores were significantly higher for AV than

A-only, and A-only scores were significantly higher than V-only scores. They determined that VE for words and sentences was not significant for younger adults, but it was significant for older adults, and had a negative correlation for consonants and sentences. Sommers et al (2005) determined that this could be due to impaired speechreading abilities. AE was not significant for older younger in consonants and negative and significant in older adults. Sommers et al. (2005) determined that the participants were able to improve speech recognition of words and sentences when both auditory and visual information were present. However; the researchers argued aging is likely responsible for declines in the capacities that are responsible for encoding the visual information. Information from this study can be used to predict how individuals who use opioid pain medication might combine auditory and visual speech signals. They may not benefit from AE and VE due to declines in capacity similar to that of aging.

Compared to the current study, the results are similar for word recognition accuracy. Participants in the current study had higher accuracy in the AV, 99.00 % (SD = .93 %) modality than the A-only 98.50 % (SD = 1.16%), and V-only 75.00 (SD = 7.14%), modalities. However, the difference in accuracy in the A-only and AV domains was not significant, suggesting that the participants in the current study were not good integrators in regards to accuracy only measures. Individuals with high frequency hearing loss, especially gradual hearing loss, as seen in presbycusis, would likely integrate visual cues to compensate for the missing auditory signal, which would likely make them better integrators than the control group in the present study (Erber, 2003). Individuals who use opioid pain medication may have lower accuracy on word recognition tasks when compared to the control group due to possible cognitive deficits, which may cause lack of integration abilities.

Processing Speed and Capacity Measures

More recently, Altieri and Townsend (2011) used RT based capacity measures to demonstrate that auditory and visual cues are processed in parallel with interactions between the two channels. The motivation for this approach was the realization that speech recognition occurs dynamically, and unfolds in real-time; in other words, accuracy-only measures miss a lot of important information.

Their basic experimental design included measuring RT (as well as accuracy plus RT in a more recent and updated capacity measure (see Altieri et al., 2014). To measure capacity, Altieri and Townsend (2011) manipulated either the A-only or V-only versus the AV stimuli. They stated that this is important because capacity is defined as processing efficiency as a function of channels (i.e., auditory, visual, or both) available in the display. The second factor the researchers manipulated was the saliency of each channel. Altieri and Townsend (2011) measured workload capacity, *C(t),* which measures how the number of channels used affects the efficiency time *t* of processing. The researchers used the following equation to measure *C(t)*:

$$C(t) = H_{AV}(t)/[H_A(t) + H_V(t)]$$

In 2014, Altieri and Hudock (2014) used a capacity measure that measured both RT and accuracy. They also measured gain from AV input using the following formula: $AV_{GAIN} = p(AV) - \max\{p(A),p(V)\}$. First, Altieri and Hudock (2014) noted a significant positive correlation between low and high frequency pure tone thresholds and AV gain scores from CUNY sentence recognition. This particular result indicates that poorer hearing ability predicts AV gain and capacity in regards to accuracy. In the present study, the participants demonstrated a mean *C(t)*, which was measured by RT only, was 0.73. The mean *C_I(t)*, which is measured by RT and accuracy, is 0.75. These numbers indicated that, on average, the listeners are not good at integrating both the auditory and visual signals. However, the maximum *C_I(t)* is 1.13, indicating it is possible for an individual to integrate the auditory and visual signals effectively. This information may be useful in predicting whether or not opioid users will benefit from the AV signal. If these opioid users have gradual hearing loss, they may be better at extracting both the auditory and visual information; however, if these opioid users have decreased cognitive abilities, they may have slower RTs, which would hinder their capacity and ability to efficiently combine auditory and visual speech cues. Since capacity better explains integration efficiency compared to "gain," the possible deficits of opioid users may be more noticeable in the speeded word recognition task than in the sentence recognition task where only gain (accuracy benefit) is measured.

However, this study cannot give information regarding the cognitive effects that may affect a person using opioid pain medication. A variety of cognitive deficits related to attention, working memory, and inhibition may in fact hinder a person's ability to combine speech information across modalities. A thesis study parallel to this current study (Arnold, 2015), showed some emerging evidence for a weak correlation between auditory sentence recognition and picture recognition (capacity using cued memory tasks), $r(23) = 0.30$ and between maximum *C(t)* and picture recognition (capacity using cued memory tasks), $r(23) = 0.23$. This suggests is the existence of a trend regarding cognition and memory, and AV speech integration abilities. Given the information from these two studies, further research could begin to look deeper into the correlation between AV speech integration abilities and factors affecting it, such as cognitive abilities as this parallel study may suggest.

Opioid users may have decreased cognitive abilities, which may lead to decreased AV speech integration abilities. Future research will look to determine the effects of opioid pain medication on speech perception abilities. It is thought that individuals in the opioid group will have decreased audiovisual gain and capacity when compared to the control group due to possible decreased cognitive ability as a result of the opioid pain medication (Rapeli et al., 2006). Longitudinal studies should be completed to determine the effects of opioid pain medication over time, and how these drugs may adversely affect speech and hearing abilities. Longitudinal studies should also look at whether or not a period of abstinence from the medication helps increase speech integration abilities. Even further research will look to determine what other factors may influence such integration abilities, such as hearing loss, cognition, and age. The information gained from future research will allow professionals and consumers to understand how the opioid pain medication can impact functional and face-to-face aspects of daily living. This information would also help determine the best treatment options for this population, in terms of hearing aid settings or perhaps cognitive based therapy. Information gained from the subsequent research will allow professionals to understand people’s speech integration abilities relative to hearing abilities, age, and cognitive abilities.

References

Adams, E. M., Gordon-Hickey, S., Moralas, H., & Moore, R. (2011). Effect of rate-alteration on speech perception in noise in older adults with normal hearing and hearing impairment. *American Journal of Audiology, 21,* 22-32. doi:10.1044/1059-0889(2011/10-0023)

Altieri, N. & Hudock, D. (2014). Assessing variability in audiovisual speech integration skills using capacity and accuracy measures. In press.

Altieri, N., Pisoni, D. B., & Townsend, J. T. (2011). Some normative data on lip-reading skills. *Journal of the Acoustical Society of America, 130,* 1-4.

Altieri, N., Stevenson, R. A., Wallace, M. T., & Wenger, M. J. (2013). Learning to associate auditory and visual stimuli: Behavioral and neural mechanisms. *Brain Topography*, online first. doi:10.1007/s10548-013-0333-7

Altieri, N. & Townsend, J. T. (2011). An assessment of behavioral dynamic information processing measures in audiovisual speech perception. *Frontiers in Psychology, 2*, 1-15. doi: 10.3389/fpsyg.2011.00238

Arnold, Sydni (2015). Effects of chronic opioid use on short-term working memory and executive functions. In preparation.

Bergeson, T. R., Pisoni, D. B., Reese, L., & Iler Kirk, K. (2003). Audiovisual speech perception in adult cochlear implant users: Effects of sudden Vs. progressive hearing loss.

Blamey, P.J., Cowan, R.S.C., Alcantra, J.I., Whitford, L.A., & Clark, G.M. (1989). Speech perception using combinations of auditory, visual, and tactile information. *Journal of Rehabilitation Research and Development, 26*, 15-24.

Braida L.D. (1991) Crossmodal integration in the identification of consonant segments. *Quarterly Journal of Experimental Psychology*, *43*(3), 647-77.

Dayasiri, M. B. K. C., Dayasena, R. P., Jayasuriya, C., Penera, D. S. C., Kuruppu, K. A. M. P., &Peris, M. N. P. (2011). Quantitative analysis of the effect of the demographic factors on presbycusis in Sri Lanka. *Australasian Medical Journal*, *4*(3), 118-122. doi:http//dx.doi.org/10.4066/AMJ.2011.562

Erber, N. P. (2003). Use of hearing aids by older people: Influence of non-auditory factors (vision, manual dexterity). *International Journal of Audiology, 42*(2), 2 S21-2 S25. Retrieved from http://web.a.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=afc9c1b5-d60a-43e3-88f8-639b3b62c82b%40sessionmgr4003&vid=5&hid=4104

Fausti, S. A., Wilmington, D. J., Helt, P. V., Helt, W. J., & Konard-Martin, D. (2005). Hearing and health care: The need for improved hearing loss prevention and hearing conservation practices. *Journal of Rehabilitation Research & Development*, *42*(4), 45-62. Retrieved from http://web.a.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=445e3fcd-cade-4c63-b47f-58b89c1dc5db%40sessionmgr4001&vid=2&hid=4212http://web.a.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=445e3fcd-cade-4c63-b47f-58b89c1dc5db%40sessionmgr4001&vid=2&hid=4212

Grant, K. W., Walden, B. E., & Seitz, P. F. (1998). Auditory-visual speech recognition by hearing impaired subjects: Consonant recognition, sentence recognition, and auditory-visual integration. *Journal of the Acoustical Society of America*, *103*(5). 2677-2690. doi:10.1121/1.422788

Humes, L. (2013). Understanding the speech-understanding problems of older adults. *American Journal of Audiology*, *22*, 303-305. doi: 10.1044/1059-0889(2013/12-0066)

Jonsson, R., Rosenhall, U., Gause-Nilsson, I., & Steen, B. (1998). Auditory function in 70- and 75- year olds of four age cohorts: A cross-sectional and time-lag study of presbycusis. *Scandinavian Audiology, 27*(2), 81-93. Retrieved from http://web.a.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=75dcd1bb-ec4d-4aac-a2c9-eb7d0a221681%40sessionmgr4005&vid=2&hid=4201

McGurk H. & Macdonald J. (1976). Hearing lips and seeing voices. *Nature, 264*, 746-748. doi:10.1038/264746a0

Miller, J.O. (1982). Divided attention: Evidence for coactiviation with redundant with redundant signals. *Cognitive Psychology*, *14*, 247-279.

Pronk, M., Deeg, D. J. H., & Kramer, S. E. (2013). Hearing status in older persons: A significant determinant of depression and loneliness? Results from the longitudinal aging study Amsterdam. *American Journal of Audiology, 22,* 316-320. doi:10.1044/1059-0889(2013/12-0069)

Rapeli, P., Kivisaari, R., Autti, T., Kahkonen, S., Puuskari, V., Jokela, O., & Kalska, H. (2006). Cognitive function during early abstinence from opioid dependence: A comparison to age, gender, and verbal intelligence match controls. *BMC Psychiatry, 6*(9). doi:10.1186/1471-244X/6/9

Roizen, N. J. (2003). Nongenetic Causes of Hearing Loss. *Mental Retardation and Developmental Disabilities Research Reviews*, *9,* 120-127. doi: 10.1002/mrdd.10068

Sumby, W. H. & Pollack, I. (1953). Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America*, *26*(2), 212-215.

Sommers, M. S., Tye-Murray, N., &Spehar, B. (2005). Auditory-visual speech perception and auditory-visual enhancement in normal-hearing younger and older adults. *Ear & Hearing*, *26*(3), 263-275.

Tye-Murray, N., Sommers, M., S., &Spehar, B. (2007). Audiovisual integration and lipreading abilities of older adults with normal and impaired hearing. *Ear & Hearing*, *28*, 656-668. doi:10.1097/AUD.0b013e31812f7185

Wenger, M.J., Gibson, B.S. (2004) Using hazard functions to assess changes in processing capacity in an attentional cuing paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(4), 708-719.

Winneke, A. H., & Phillips, N. A. (2011). Does audiovisual speech offer a foundation of youth for old ears? An event-related brain potential study of age differences in audiovisual speech perception. *Psychology and Aging, 26*(2), 427-438. doi:10.1037/a0021683

Appendix

AV Take your medicine
AV Buy those snow boots
AV Do not scratch your rash or it will just become worse
AV Do not be late today
AV Can the plumber fix the leaking faucet
AV Put both cars in the garage before it starts to snow
AV Alcohol can damage your liver
AV We need to renovate the beach house this summer
AV Computers save time
AV Don’t fool around on the high diving board because it’s dangerous
AV Get out the snow shovel
AV The only way to diet is not to eat
AV Did anyone sing at your wedding reception
AV Do not add too much salt to the soup
AV When you're in London make sure you eat fish and chips
AV Are musicals popular
AV Where did he buy that gray suit
AV Clean the fish tank before you buy those goldfish
AV Pass the ball to him
AV where are all the employee time cards
AV How much is that black dress in the window
AV Buy a new garage door
AV Where's the nurse
AV Does labor day always fall on the first Monday of September
AV Please do not change that radio station
A Have you and your fiancée set a date for the wedding
A Is she wearing the blue dress to the theater
A Clean the cassette player before using it
A When should we have the awards dinner
A Where are the newlyweds going to spend their honeymoon
A Do the ski gloves fit
A He swims fast
A The weather forecast for tomorrow calls for sunshine and low humidity
A Did you warm up the baby’s bottle of milk
A Did the office furniture arrive
A Rents are high
A We used to collect rocks and shells when we were young
A Dancers have a good sense of rhythm
A Isn’t this coffee too sweet

A Polish your shoes
A The bee stung the little girl while she was picking flowers
A The school is closed for Labor day
A The value of the dollar fell with the deficit
A Ask her father
A Quit your job if you are not satisfied with your salary
A Put snow tires on the car today
A If she is not careful shell get a sunburn
A Well pay the telephone bill
A He should floss his teeth
A Has spring arrived
V What will we make for dinner when our neighbors come over
V Is your sister in school
V Does your boss give you a bonus every year
V Don’t spend so much on new clothes
V What’s your recipe for cheesecake
V Is your nephew having a birthday party next week
V What’s the humidity
V Let the children stay up for Halloween
V He plays the bass in a jazz band every Monday night
V How long does it take to roast a turkey
V Which team won
V Take your vitamins every morning after breakfast
V People who invest in stocks and bonds now take some risks
V Those albums are very old
V Aren’t dishwashers convenient
V Is it snowing or raining right now
V The school will be closed for Washington’s Birthday and Lincolns Birthday
V Your check arrived by mail
V Professional musicians must practice at least three hours everyday
V Are whales mammals
V Did the basketball game go into overtime
V When he went to the dentist he had his teeth cleaned
V Well plant roses this spring
V I always mail in my loan payments on time
V Sneakers are comfortable