Washington University School of Medicine
Digital Commons@Becker

Independent Studies and Capstones | Program in Audiology and Communication Sciences

2011

Evaluation of auditory-visual speech perception in individuals diagnosed with dementia of the Alzheimer's type

Alyse Paige Firtel

EVALUATION OF AUDITORY-VISUAL SPEECH PERCEPTION IN INDIVIDUALS DIAGNOSED WITH DEMENTIA OF THE ALZHEIMER'S TYPE

by

Alyse Paige Firtel

A Capstone Project
submitted in partial fulfillment of the
requirements for the degree of:

Doctor of Audiology

Washington University School of Medicine
Program in Audiology and Communication Sciences

May 18, 2012

Approved by:
Mitchell Sommers, Ph.D., Capstone Project Advisor
Brent Spehar, Ph.D., Capstone Project Advisor

Abstract: Auditory-visual speech perception testing was completed using word- and consonant-level stimuli in individuals with known degrees of dementia of the Alzheimer's type. The correlations with the cognitive measures and the speech perception measures (A-only, V-only, AV, VE or AE) did not reveal significant relationships.

ACKNOWLEDGEMENTS

The author acknowledges Mitchell Sommers, Ph.D. for his guidance and mentorship, extending from the beginning thoughts of this project through its completion. The author extends immense gratitude to Brent Spehar, Ph.D., co-advisor for this project, who devoted a great deal of time assisting with programming, data analysis, and manuscript revision. The author also thanks John C. Morris, MD, Virginia Buckles, Ph.D., Becky Fierberg, Betsy Grant, Martha Storandt, Ph.D., and all of the other professionals at the Charles F. and Joanne Knight Alzheimer's Disease Research Center for helping to enable this project by way of support for continued research, participant recruitment, and scheduling. Most importantly, the author thanks all of the individuals who contributed to this research as participants. The individuals who participated with dementia of the Alzheimer's type increased the author's knowledge and understanding of the progression of the disease and the need for further behavioral research within this population. This Capstone research project was supported by NIH/NIA.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

ABBREVIATIONS

AV	Auditory-visual
A-only	Auditory only
V-only	Visual only
CAVET	Children's Auditory-Visual Enhancement Test
IAC	Iowa Consonant Test
AE	Auditory Enhancement
VE	Visual Enhancement
SRT	Speech Reception Threshold
SDT	Signal Detection Threshold
SNR	Signal-to-noise ratio
ADRC	Charles F. and Joanne Knight Alzheimer's Disease Research Center
EM	Episodic memory
SM	Semantic memory
WM	Working memory
VS	Visuospatial composite

Introduction

Most individuals can recognize speech better when they are able to hear and see the speaker than when listening alone. When listening environments are noisy or acoustic information is degraded, visual speech information (i.e. information provided by lipreading) can be used to supplement the deficit. The current project focuses on individuals with dementia of the Alzheimer's type and the potential influence of the associated cognitive deficits on auditory-visual speech perception.

Hearing and Speech Perception in the Elderly

Auditory sensitivity decreases as a function of age, ultimately leading to sustained hearing loss (CHABA, 1988; Frisina, 1997; Gates & Mills, 2005). Age-related hearing loss, also known as presbycusic hearing loss, is a comprehensive diagnosis that occurs due to the natural deterioration of the auditory system over the course of one's life. This deterioration can be a result of many potential factors such as environmental noise, physiologic processes, and medical influences (CHABA, 1988). These changes typically begin with hearing loss in the high frequencies.

An individual's normal range of hearing is from 20-20,000 Hz. The sounds of speech range from 250 to 8,000 Hz, with the majority of consonant sounds falling between 2,000 to 4,000 Hz. Consonant sounds are critical to the clarity of speech. Because presbycusic hearing loss typically begins in the highest frequencies an individual may not report hearing difficulty until it begins to affect speech frequencies in this range. While presbycusic hearing loss is one of the most prevalent signs of the aging process, it is not the only auditory process that declines with age.

Auditory speech perception performance, which is the ability to understand what is being said, also declines as a function of age (Frisina, 1997). This decline is partially a result of decreased hearing sensitivity in the critical speech frequencies (2000-4000 Hz), as well as a result of increased difficulty in adverse listening situations (CHABA, 1988). An individual with declining auditory speech perception will experience a decreased ability to communicate with individuals due to their inability to perceive appropriate acoustic cues. The compromised reception of acoustic cues highlights the importance of a supplemental process to aid understanding.

Auditory-visual Speech Perception

Auditory-visual (AV) speech perception, the combination of visual cues of speech with the auditory properties of a speech signal has been shown to improve speech perception (Sumby & Pollack, 1954). This is especially true when the reception of acoustic cues is compromised. The cues that a listener extracts from a speech signal can be different for each modality. With auditory speech perception, manner of articulation and voicing cues are more easily obtained, whereas visual speech perception often provides important information for place of articulation cues (Grant, Walden, & Seitz, 1998; Grant & Seitz, 1998). In auditory-visual speech perception, the additive effect of presentation in both modalities is greater when cues are complementary compared to those that are redundant (Grant et al., 1998; Grant & Seitz, 1998); this is because auditory cues provide acoustic information that is enhanced by the articulatory features provided by speech in the visual modality.

The complementary nature of the auditory and visual speech signals contributes to the additive outcome. Cues for manner of articulation and voicing are mostly obtained through the auditory signal; perception of these speech features are difficult to extract through the visual modality. For example, cues about voicing and nasality (like in the phoneme /n/ or /m/) can be obtained through the auditory modality, even in distorted listening environments, whereas the visual signal does not transmit these features in any listening environment (Grant et al., 1998; Grant & Seitz, 1998). The visual signal conveys place of articulation cues that are important for determining the difference between phonemes in situations when the auditory cues are degraded, as when listening in background noise, reverberant conditions, or in the presence of a hearing loss. Although place information is often available through the auditory channel, if a listener's acoustic cues for place of articulation (/t/ vs. /p/) are diminished, then with speechreading[1] the listener may be able to supplement place information from the visual modality to correctly identify the phoneme.

A central question in assessing auditory-visual speech perception is to identify how much information each modality contributes to overall speech perception. For example, to establish how much information is conveyed via the visual modality one could measure visual enhancement – the improvement in speech perception that results from adding visual speech information to an ongoing auditory signal. Similarly, to determine the amount of information conveyed via the auditory modality, one could measure auditory enhancement – the amount of benefit obtained from auditory speech information as a supplement to visual speech information. Both visual enhancement and auditory enhancement are determined relative to an individual's performance, in the auditory-only and visual-only conditions respectively. Differential benefits in enhancement can result from changes in reception of information through each modality. For example, improving audition through the use of hearing aids would likely improve auditory enhancement. Similarly, all other things being equal, those with better lipreading abilities would

[1] Speechreading refers to processing visual speech information in the presence of congruent auditory signals. (Feld and Sommers, 2009)

show better visual enhancement (MacLeod & Summerfield, 1990, Sommers, Tye-Murray, & Spehar, 2005, Sumby & Pollack, 1954; Grant et al., 1998).

Factors effecting lipreading ability in the elderly

There are differences across populations for lipreading ability and AV speech perception. Lipreading ability can affect how much auditory-visual benefit an individual obtains. Varying levels of lipreading performance have been demonstrated in normal and hearing-impaired children and adults (Lyxell & Holmberg, 2000; Cienkowski & Carney, 2002; Sommers et al., 2005). One consistent finding throughout the studies that have investigated auditory-visual speech perception is that older adults typically display poorer lipreading ability than young adults (Sommers et al., 2005; Spehar, Tye-Murray, & Sommers, 2004; Cienkowski & Carney, 2002; Dancer, Krain, Thompson, Davis, & Glen, 1994). The age-related deterioration in lipreading ability may be surprising because one would expect individuals with hearing loss to supplement the lack of auditory information with visual cues. However, factors that affect lipreading ability, like poor working memory and reduced processing speed, seem to counteract this potential compensatory process. Determining additional factors that affect lipreading may help establish the source of differences among individuals.

Numerous studies have tried to elucidate the cause of differences in lipreading performance among individuals. This is because lipreading ability can affect the amount of auditory-visual benefit an individual achieves. For example, studies that have examined intelligence (Elphick, 1996, Simmons, 1959), vocabulary (Lyxell & Rönnberg, 1992), and years of education (Dancer et al., 1994) show poor correlations with lipreading ability. Further research has demonstrated some significant correlations between lipreading and verbal-inference making (Lyxell & Rönnberg, 1989), and lipreading and the ability to decipher fragmented sentences (Simmons, 1959). Recent investigations have examined the relationship between lipreading ability and working memory (Lidestam, Lyxell, & Anderson, 1999; Lyxell & Holmberg, 2000), and lipreading ability and processing speed (Lyxell & Holmberg, 2000). Due to the inconsistencies in methodology (e.g. sample, inclusion of predictor variables) it is difficult to make comparisons across studies. Feld and Sommers (2009) examined all of these measures within groups of young and older adults and found that both processing speed and spatial working memory correlate with lipreading ability.

Advanced age is associated with declines in both processing speed and spatial working memory (for a short review see Salthouse, 2010), as well as with declines in lipreading ability. Research has shown older adults consistently demonstrate poorer working memory ability and slower processing speed compared to younger adults (Salthouse, 1991,1994, 2009; Jenkins, Myerson, Hale, & Fry, 1999; Myerson, Hale, Rhee, & Jenkins, 1999; Jenkins, Myerson, Joerding, & Hale., 2000; for further review see Cerella & Hale, 1994). Additionally, this deficit is more pronounced on spatial working memory tasks than on verbal working memory tasks. Thus, the high variance in lipreading ability can be partially explained by the differences in processing speed and spatial working memory for each individual. Taken together, these findings suggest that older adults obtain less benefit from the addition of visual speech information, partially as a result of decreased cognitive performance. The current project focuses on a population that allows for the analysis of specific cognitive processes and their influence on lipreading ability, and thus auditory-visual benefit.

Dementia of the Alzheimer's Type

Dementia of the Alzheimer's type (DAT) is the most common type of dementia suffered by individuals as they age and is characterized by progressive declines in memory, thinking, and reasoning. Those suffering from DAT also exhibit difficulty with spatial relationships, visual images, word retrieval, and speech production. These difficulties occur in part due to the development of beta-amyloid plaques and twisted strands of the protein tau, resulting in impaired neural synapses and cell death. This cell death ultimately leads to the atrophy of brain tissue. The degeneration of brain structure and synapses translates into functional declines as the disease progresses. (Alzheimer's Association, 2010)

Functional declines across cognitive abilities are observed in individuals with DAT. Multiple studies have established that individuals with DAT have reduced processing speed and degraded working, semantic, and episodic memory as a result of their cognitive impairment (Baddeley, Loggie, Bressi, Della Sala, & Spinnler, 1986; Baddeley, Bressi, Loggie, Della Sala, & Spinnler, 1991; Metzler-Baddeley, 2007; Heyanka, Mackelprang, Golden, & Marke, 2010; Chertkow & Bub, 1990). Of particular interest to the current study, due to their known relationship with lipreading performance are working memory and processing speed. Working memory has been delineated into two subsystems, the visuospatial component which is responsible for storage and manipulation of visual and spatial inputs and the verbal component which is responsible for speech-based information (Baddeley et al., 1991; Jenkins et al., 1999). Research has demonstrated that older adults demonstrate declines in visuospatial working memory more so than in verbal working memory (Myerson et al., 1999). In individuals with DAT, both of these components are affected. Declines in visuospatial working memory are exhibited in individuals with DAT. This has been demonstrated in performance on the Corsi block tapping measurement which examines visual encoding, short term storage for spatial location, and sequence ordering (Baddeley et al., 1991, Spinnler, Della Sala, Bandera, & Baddeley, 1988; Fischer, 2001; Huntley & Howard, 2010). Verbal working memory becomes more impaired as the disease progresses (Baddeley et al, 1991; Huntley & Howard, 2010). This is evidenced in declining performance on digit and word span measures. Deficits in working memory are also present in dual-task measures and increase with the progression of the disease. Individuals with DAT suffer from an inability to efficiently utilize cognitive processes in performing simultaneous tasks partially as a result of deficits in central executive function (Baddeley et al., 1991). Deficits in central executive function are concerned with attention and coordinating information from a number of different sources; reduction in central executive function has the potential to further exacerbate memory and processing deficits (Baddeley et al., 1991). Individuals with DAT demonstrate processing deficits. They perform worse on tasks that demand more in-depth processing, like digit ordering and making appropriate judgments about grammatical errors, compared to healthy, older adults (MacDonald, Almor, Henderson, Kempler, & Andersen, 2001).

Auditory-visual speech perception and DAT

Auditory-visual speech perception is a task of dual modalities in which both peripheral and central systems contribute to a person's ability to benefit from combining auditory and visual speech information. Delbeuck, Collette, and Van der Linden, (2007) utilized the McGurk effect[2] to examine the brain connectivity in individuals with DAT; a cross-modal task, like the McGurk model, requires that the brain process information from the auditory and visual modalities and utilize the information in order to produce the illusory effect. The researchers analyzed performance of each group (individuals with DAT and controls) on congruent and incongruent auditory-visual stimuli with regards to type of illusory effect (fusion (di) – occurs when an auditory bilabial (bi) is dubbed with a velar (gi), or combination (bgi) – emerges when bilabial

[2] McGurk effect is an illusory phenomenon via auditory-visual speech perception in which the visual stimulus alters the auditory perception of the stimulus. (McGurk and MacDonald, 1976)

(bi) is presented visually and non-bilabial (gi) is presented auditorily) reported by the participant. The results of this study demonstrated that individuals within the DAT sample exhibited auditory-visual integration deficits. This was evidenced in fewer numbers of illusions in the McGurk condition when compared to their control subjects. Furthermore, the authors proposed that a potential contributing factor is that individuals with DAT may have increased difficulty processing visual speech information. This was noted based on a pattern of poorer performance in the visual-only conditions of the study, which included a visual discrimination task and a lipreading task. The Delbeuck et al. study supports the electrophysiologic (Golob, Miranda, Johnson, & Starr, 2001; Fingelkurts, Fingelkurts, Krause, Mottonen, & Sams, 2003) and neuroimaging (Calvert, Bullmore, Brammer, Campbell, Williams, & McGuire, 1997; Calvert, Campbell, & Brammer, 2000; Drzezga, Grimmer, Peller, Wermke, Siebner, Rauschecker, Schwaiger, & Kurz, 2005; Macaluso, George, Dolan, Spence, & Driver, 2004; Sekiyama, Kanno, Mirua, & Sugita, 2003) evidence demonstrating cross-modal deficits based on the disconnection hypothesis. The disconnection hypothesis suggests that neural dysfunction in individuals with DAT is better explained by a disturbance of the communications or connectivity between different areas of the brain rather than by disturbances in a specific location in the brain. More neuropsychological research is needed to determine the deficits in auditory-visual speech perception.

Identifying the deficits in auditory-visual speech perception may aid in further understanding the deficits in multi-modal integration that individual's with DAT exhibit. For example, MacDonald et al. (2001) demonstrated that patients with DAT perform poorly on tasks that require cross-modal integration (identifying a visual image in conjunction with information presented auditorily). In this study, participants were placed in front of a computer. They were presented with a sentence or sentence fragment auditorily and asked to name the appropriate completion of the sentence (this was assessed using different parts of speech) based on a visual presentation of potential responses (e.g. subject-verb agreement or pronouns) on the computer screen. One reason for this poor performance may be that the associative areas of the brain responsible for performing tasks of this type, and other multi-modal tasks, need to effectively connect or communicate in order to accomplish integration of the information obtained from different sensory modalities (Delbeuck et al., 2007). Without the ability to efficiently utilize information presented simultaneously through two modalities, patients may experience difficulties with auditory-visual speech perception.

There is limited behavioral research regarding auditory-visual speech perception in the DAT population. Much of the research has focused on evoked potentials (Golob et al., 2001; Fingelkurts et al., 2003) and imaging (Calvert et al., 1997; Calvert et al., 2000; Drzezga, et al., 2005; Macaluso et al., 2004; Sekiyama et al., 2003). For example, using auditory and visual evoked potentials, Golob et al. (2001) provided evidence for reduced interactions between cortical regions based on refractory effects. The refractory effect consists of reduced amplitude and latency of the evoked response observed for one stimulus when another stimulus has been presented immediately before. The absence of a refractory effect was observed in individuals with DAT when using auditory-visual stimuli; this result is potentially due to the disconnection between sensory areas in individuals with DAT. Studies using electrophysiologic and imaging techniques illustrate structural involvement and disturbance in interactions between sensory modalities. Despite the considerable physiologic evidence that dual-task deficits exist for individuals with DAT, to our knowledge, studies examining the functional implications of this deficit with regards to auditory-visual speech perception are limited (Delbeuck et al., 2007).

The current study was designed to further investigate the influences of cognitive processing on auditory-visual speech perception. The current study examines individuals with DAT because they provide an opportunity to study visual enhancement and lipreading ability within the elderly population across varying degrees of known memory loss and cognitive processing ability. One potential explanation for variance in performance is that individuals with DAT are impaired on both abilities that contribute to lipreading and consequently, will have reduced benefit from visual enhancement. A secondary purpose of this study was to compare these different measures for consonants (nonsense syllables) versus meaningful words. Previous research (Sommers et al., 2005; Grant et al., 1998) has shown that in both young adults and healthy, older adults these measures are not very well correlated and we wanted to extend this research to individuals with DAT.

Methods

Participants

Thirteen elderly adults (mean age = 78.2, SD = 4.7, 6 males, 7 females) were recruited from the participant pool maintained by the Charles F. and Joanne Knight Alzheimer's Disease Research Center (ADRC) at Washington University School of Medicine in St. Louis. All participants received compensation of $15 per hour for their participation in this study. A typical study session consisted of one, two-hour session.

Participants were placed into three groups according to their rating on the Clinical Dementia Rating assessment (CDR) (Morris, 1993). The CDR scale is a 5-point scale that identifies six domains of cognitive and functional performance with regards to dementia. The CDR scale examines: memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care. Each category is rated based on a semi-structured interview conducted with the patient and an accompanying, reliable source (e.g. spouse, family member) (Morris, 1993, see Table 1). Appendix A contains a detailed description provided by Martha Storandt, Ph.D.[3] describing the psychometric test battery of standardized tests used to establish the control for the CDR baseline. The three groups were defined as CDR 0, CDR 0.5 and CDR 1. Individuals in the CDR 0 group are healthy, older adults without DAT. Those participants in the CDR 0.5 and CDR 1 groups have received the more specific diagnosis of DAT. The diagnosis of DAT was based on extensive medical and neurological examinations, in addition to their CDR rating. Individuals with concomitant illnesses contributing to DAT and individuals with other potentially dementing disorders were excluded from this study.

---

[3] Psychometric test description provided by Martha Storandt, PhD, Professor of Psychology and Neurology at Washington University in St. Louis.

CLINICAL DEMENTIA RATING (CDR)

CLINICAL DEMENTIA RATING (CDR):	0	0.5	1	2	3

	Impairment				
	None 0	Questionable 0.5	Mild 1	Moderate 2	Severe 3
Memory	No memory loss or slight inconsistent forgetfulness	Consistent slight forgetfulness; partial recollection of events; "benign" forgetfulness	Moderate memory loss; more marked for recent events; defect interferes with everyday activities	Severe memory loss; only highly learned material retained; new material rapidly lost	Severe memory loss; only fragments remain
Orientation	Fully oriented	Fully oriented except for slight difficulty with time relationships	Moderate difficulty with time relationships; oriented for place at examination; may have geographic disorientation elsewhere	Severe difficulty with time relationships; usually disoriented to time, often to place	Oriented to person only
Judgment & Problem Solving	Solves everyday problems & handles business & financial affairs well; judgment good in relation to past performance	Slight impairment in solving problems, similarities, and differences	Moderate difficulty in handling problems, similarities, and differences; social judgment usually maintained	Severely impaired in handling problems, similarities, and differences; social judgment usually impaired	Unable to make judgments or solve problems
Community Affairs	Independent function at usual level in job, shopping, volunteer and social groups	Slight impairment in these activities	Unable to function independently at these activities although may still be engaged in some; appears normal to casual inspection	No pretense of independent function outside home Appears well enough to be taken to functions outside a family home	Appears too ill to be taken to functions outside a family home
Home and Hobbies	Life at home, hobbies, and intellectual interests well maintained	Life at home, hobbies, and intellectual interests slightly impaired	Mild but definite impairment of function at home; more difficult chores abandoned; more complicated hobbies and interests abandoned	Only simple chores preserved; very restricted interests, poorly maintained	No significant function in home
Personal Care	Fully capable of self-care		Needs prompting	Requires assistance in dressing, hygiene, keeping of personal effects	Requires much help with personal care; frequent incontinence

Score only as decline from previous usual level due to cognitive loss, not impairment due to other factors.

Table 1 Clinical Dementia Rating scale (Morris, 1993) addresses the six categories used to determine person's CDR rating.

The three groups were equal with respect to average participant age (CDR 0 Mean = 77.4 yrs (SD = 4.8), CDR 0.5 Mean = 77.2 yrs (SD = 2.7); CDR 1 Mean = 81.0 yrs (SD = 7.2) and average level of education (CDR 0 Mean = 14.3 yrs (SD = 2.3), CDR 0.5 Mean = 15.0 yrs (SD = 4.4); CDR 1 Mean = 14.0 yrs (SD = 3.4). Group equality was determined by one-way ANOVAs, one for age ($F(2,10)=0.7$; $p>.05$) and one for years of education ($F(2,10)=0.1$; $p>.05$). Each participant's age and the amount of education in years is shown in Table 2.

Participant	Age	Years of Education
1	81	16
2	70	13.5
3	80	12
4	89	16
5	78	18
6	83	13
7	76	18
8	75	16
9	76	15
10	76	12
11	74	20
12	79	10
13	79	9

Table 2 Participant's age and years of education

All participants were screened for vision and hearing. Potential participants with corrected visual acuity poorer than 20/40, as determined by a Snellen eye chart, would have been excluded from the study. The hearing screening consisted of word-level discrimination testing and pure-tone threshold testing for each ear. Audiometric tests were completed using a calibrated Madsen Auricle audiometer under headphones (TDH 49). The discrimination testing was completed using recorded CID W-22 word lists. Word lists were presented at 40 dB above a participant's pure-tone average (PTA = average of threshold obtained at 500, 1000, and 2000 Hz) to ensure adequate audibility for each word. Table 3 shows the results for the pure-tone testing threshold values, pure-tone averages, and the W-22 scores for the better ear of each participant in the study separated by group. Better ear measurements are reported here and throughout the paper because testing was conducted in the soundfield without hearing aids. Under these conditions hearing performance is expected to be associated with the ability of a participant's better ear. To determine if there were any differences in hearing ability among the three groups one-way ANOVA was conducted for the better ear PTA.

CDR	Participant	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz	8000 Hz	PTA (dB HL)	W-22 Score
0	2	20	20	20	40	40	60	26.6	88%
	3	20	15	5	15	40	80	11.6	92%
	5	20	10	5	15	50	85	10	94%
	6	15	10	5	10	45	65	8.3	88%
	9	15	10	20	40	40	75	15	82%
	Mean (sd)	18.0 (2.7)	13.0 (4.5)	11.0 (8.2)	24.0 (14.7)	43.0 (4.5)	73.0 (10.4)	12.2 (9.5)	88.8% (4.6)
0.5	1	35	45	45	35	55	70	41.6	94%
	7	15	10	15	45	45	55	23.3	92%
	10	15	15	25	25	35	65	16.6	92%
	11	5	5	5	5	60	75	5	94%
	13	25	25	30	30	50	65	28.3	90%
	Mean (sd)	19.0 (11.0)	20.0 (15.8)	24.0 (15.1)	28.0 (14.8)	49.0 (9.6)	66.0 (7.4)	22.9 (13.5)	90.8% (1.1)
1	4	40	10	5	5	40	85	6.6	96%
	8	20	15	10	5	15	50	10	90%
	12	15	15	35	55	60	85	35	84%
	Mean (sd)	25.0 (13.2)	13.3 (2.8)	16.6 (16.0)	21.6 (28.8)	38.3 (22.5)	73.3 (20.2)	17.2 (15.5)	88.7% (8.1)

Table 3 CDR indicates Clinical Dementia Rating; Pure-tone testing threshold values, Pure-tone average (PTA) calculated using 500, 1000, 2000, and 4000 Hz, and W-22 scores for the better ear of each participant in the study, separated by group.

Stimuli

Stimuli consisted of consonant and word level speech presented in three conditions: auditory only (A-only), lip-reading (V-only), and audiovisual (AV).

Consonants

Participants received five repetitions of 13 consonants in an /i-C-i/ format. The consonants tested were: /b/, /d/, /f/, /g/, /k/, /m/, /n/, /p/, /s/, /t/, /v/, /ʃ/, and /z/. Participants were presented with /ibi/, /iki/, etc. One male talker produced all of the stimuli for testing consonants.

The stimuli were digitized from existing audiovisual Laserdisc recordings of the Iowa Consonant Test (Tyler, Reece, Tye-Murray, 1986). This was achieved by connecting the output of the Laserdisc player (Laservision LD-V8000) into a commercially available PCI interface card for digitization (Matrox RT2000). Acquisition was controlled by software (Adobe Premier). Video capture was 24-bit, 720 x 480 in nTSC-standard 4:3 aspect ratio and 29.97 frames per second to best match the original analog signal. Audio was captured at 16 bits and a sampling rate of 48 kHz.

Words

Stimulus materials for word-level testing was the Children's Auditory Visual Enhancement Test (CAVET) (Tye-Murray & Geers, 2001). The CAVET consists of three lists (List A, List B, and List C) that are balanced in difficulty for speechreading. The CAVET test consists of three lists of 20 words spoken by a female talker. Recordings are of her head and shoulders speaking directly into the camera. Each one to three-syllable word is preceded by the carrier phrase "say the word". The words contained on CAVET are fairly familiar to adults, since the measure was developed for use on four to six year old hearing impaired children. It was used for the current study because it was expected to avoid floor effects in the V-only condition and ceiling level performance in the AV condition. Stimuli were digitized from VHS recordings using the same processes described above for the consonant stimuli.

Procedures

After completing the hearing and vision screenings, participants were seated in a sound treated room approximately .5 meters from a 17" ELO touch-systems monitor. They were asked to respond to each presentation by repeating the word (word-level testing) or pressing the corresponding response button on the touchscreen monitor (closed-set consonant testing) after each presentation. Within each stimulus task, testing in the A-only, V-only, and AV conditions were counterbalanced among the whole sample. Testing order was counterbalanced across the sample, such that each modality was presented first, second, and third equally often. This was to help ensure approximately the same number of participants received each condition order. Participants were informed of the upcoming presentation modality by the researcher. The experimenter verbally informed each participant that they would 'only see the talker', 'only hear the talker,' or 'both see and hear the talker'. Stimuli were presented via a PC (Dell, Precision) using software specifically designed in LabView for this study.

Audio presentations were in the soundfield using two loudspeakers positioned +/-45 degrees to the left and right of the participant when looking at the monitor. All stimuli levels were presented at 62 dB SPL A (approximately 50 dB HL) in four-talker background babble. The level of the babble noise was adjusted for each individual (see the Setting Background Babble section below). Audio was routed from the computer's Sound Blaster Live audio card to a calibrated audiometer (Madsen Auricle) then to a Sampson amplifier before being sent to the loudspeakers.

For CAVET testing, participants were informed that they would see, hear, or see and hear the female talker. They were informed that she would articulate the carrier phrase "say the word" and a target word. Participants were told the carrier phrase and were instructed to identify the word that follows the phrase. Participants were instructed to say their response aloud; if the participant was unsure of the word, he/she was encouraged to guess. Testing was conducted in four-talker background babble set at the adjusted level for each participant. The presentation modality (A-only, V-only, AV) of the lists was counterbalanced across the sample such that equal number of participants received the same order (order 1-6 described in Table 4).

Order	Modality 1	Modality 2	Modality 3
1	A-only	V-only	AV
2	V-only	A-only	AV
3	AV	V-only	A-only
4	A-only	AV	V-only
5	V-only	AV	A-only
6	AV	A-only	V-only

Table 4 Counterbalance order (1-6) of presentation modality order used for both CAVET and IAC testing

Identification of the consonant-level stimuli was measured using a closed-set test structure. During the consonant identification testing, each response was made by pressing one of 13 virtual buttons that appear on the screen after each presentation. Of particular concern with the current population, testing may have been more difficult for those with further progression of DAT. For example, after a reasonable number of consonant trials, healthy older adults likely were able to have some memory of the mapping between specific keys on the touch screen and particular sounds, whereas the DAT patients may not have been able to this as well. Before testing, the experimenter familiarized each participant with the stimuli by reviewing the consonant and appropriate touch-screen response. During practice, participants were first asked to repeat each consonant stimulus. Participants were then asked to practice identifying all 13 consonants presented in the AV condition in a background babble ratio of +10 prior to beginning testing. Some participants were slightly confused during this task (e.g. looking for “C” for the sound /iSi/). In this case, participants were reinstructed on phoneme-grapheme association. If participants were unsure of the target consonant, they were encouraged to guess in order to continue the assessment. The presentation modality (A-only, V-only, AV) of the lists was counterbalanced across the sample such that equal number of participants received the same order (recall Table 4).

Setting Background Babble

All testing was conducted in four-talker background babble and signal-to-babble levels were set individually for each participant and stimulus condition (word- and consonant- level). The main goal in setting the background babble level was to avoid floor level performance in the A-only condition and to simultaneously avoid ceiling level performance in the AV condition.

Background babble level was set independently for each participant. Given the population for the current study it was important to avoid the potential fatigue-related issues that might have resulted from presenting excessive trials. The modified ASHA SRT (ASHA, 1988) procedure was tailored to achieve the SDT levels as quickly as possible. Participants were presented with spondee words in broadband noise. The starting level was 0 SNR; the level of the noise increased by 5 dB (so as to produce, -5 dB SNR, -10 dB SNR, and so forth). Participants were presented with five spondee words at each level, until they reached a level where they stopped responding because they could not discern the stimulus. A correction factor of 10 was utilized to determine the appropriate starting level of background babble for testing. As an important note, clinical discretion was occasionally needed to adjust the signal to babble ratio levels to avoid ceiling performance in the AV condition while maintaining above floor performance in the A-only condition. In all cases where adjustments were needed testing stopped and was restarted at an easier level.

Analyses

Visual Enhancement and Auditory Enhancement

Visual enhancement describes the amount of benefit one can obtain from adding the visual channel to the auditory channel. It is calculated by the formula VE = (p(AV)-p(A)/1-p(A) where p(AV) is the probability of correctly identifying the stimuli in AV condition and p(A) is the probability of correctly identifying the stimuli in the A-only condition. Specifically, this describes the proportion of possible improvement that was achieved by adding the visual information to the auditory information. For example, if a person can achieve 40 % performance in the A-only condition there is a range of possible improvement of 60 percentage points. If a person achieves 70% in the audiovisual condition a VE of 50% has occurred because half of the possible range of improvement was accomplished. Note that using the current calculation eliminates the inherent bias of using a simple change score from A-only scores to AV performance. Dividing by the amount of possible improvement from A-only (1-p(A)), allows for the normalization of A-only performance across participants. This calculation has been used in previous studies of AV speech perception (Grant & Seitz, 1998, Grant et al., 1998, Sommers et al., 2005).

Auditory enhancement is similar to VE in that it describes the amount of benefit achieved by adding a second channel. In AE, however, the calculation describes the amount of normalized benefit from adding the auditory channel to V-only performance: AE = (p(AV)-p(V))/1-p(V) where p(AV) is the probability of correctly identifying the stimuli in AV condition and p(V) is the probability of correctly identifying the stimuli in the V-only condition. Notably, in the current study A-only performance was forced to the same level across participants, while the AE calculation normalizes for V-only scores. Because of this, the AE calculation is less likely to be influenced by differences in unimodal performance than the VE calculation.

Results

Results for the two types of stimuli (word-level and consonant-level) are described separately. For each type of stimulus results for the obtained scores in the three conditions (A-only, V-only, and AV) are described first then results for the derived measures of benefit (AE and VE) will be presented. Correlations between the two types of stimuli are then described. Finally, cognitive measures provided by the ADRC and auditory-visual performance are correlated to further explore their potential influences on measures of auditory-visual speech perception.

Children's Audiovisual Enhancement Test

Performance in A, V, and AV Conditions

Figure 1 displays mean percent correct scores, split by CDR group, in the A-only, V-only, and AV conditions for the word-level stimuli. Initial visual inspection of Figure 1 indicated that A-only scores were fairly similar across the three groups (CDR 0 mean = 35.0 % (SD = 12.7); CDR 0.5 mean = 34.0 % (SD = 12.4); CDR 1 mean = 40.0 % (SD = 10.0). The attempt to keep A-only performance at a similar level across participants by adjusting signal-to-babble ratio individually appeared successful. This was verified by one-way analysis of variance (ANOVA) that compared A-only performance across the three CDR groups that found no difference between groups ($F(2, 10) < 1$; $p > .05$; $\eta^2_p = .078$).

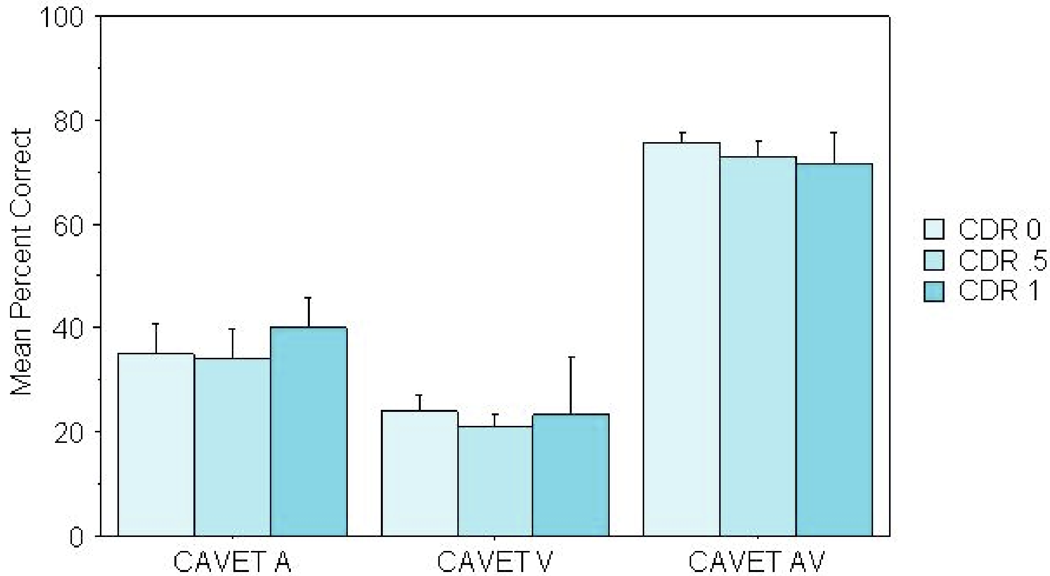

Figure 1 displays mean percent correct scores, split by CDR group, in the A-only, V-only, and AV conditions for the word-level stimuli (CAVET). Error bars indicate standard error.

Examination of V-only performance indicated that as expected all groups showed relatively low performance when compared to A-only and AV conditions. Mean percent correct scores, split by CDR group are as follows: CDR 0 mean = 24.0 % (SD = 6.5), CDR 0.5 mean = 21.0 % (SD =5.4), CDR 1 mean = 23.3 % (SD =18.9). Potential differences across the three CDR groups in V-only performance were investigated using one-way ANOVA with V-only performance entered as a between subjects variable. No difference between groups was found for the V-only scores ($F(2,10) < 1$; $p > .05$; $\eta^2_p = .064$).

Similar to the results for A-only and V-only word-level performance, the results for the AV condition suggested similar performance across CDR groups (CDR 0 mean = 75.8 % (SD = 4.2), CDR 0.5 mean = 73.0 % (SD = 6.7), CDR 1 mean = 71.6 % (SD = 10.4). Potential differences across the three CDR groups for AV performance were investigated using one-way ANOVA. No difference between groups was found for AV scores ($F(2,10) < 1$; $p > .05$; $\eta^2_p =$ .09).

Comparison of Visual and Auditory Enhancement

Figure 2 displays means for VE, split by CDR group for the word-level stimuli. None of the differences in VE scores between CDR reached statistical significance ($F(2,10) = 1.2$; $p > .05$; $\eta^2_p = .21$), however, the data may suggest a trend as a function of CDR rating.

Analyses paralleling those for VE were also completed to evaluate AE differences relative to CDR group. Figure 3 displays means for AE as a function of CDR group, for word-level stimuli. In contrast to the results for VE, scores for AE appear consistent across the three CDR groups. Similar to results for VE, however, the ANOVA did not indicate a significant main effect for AE ($F(2,10) < 1$; $p > .05$; $\eta^2_p = .11$).

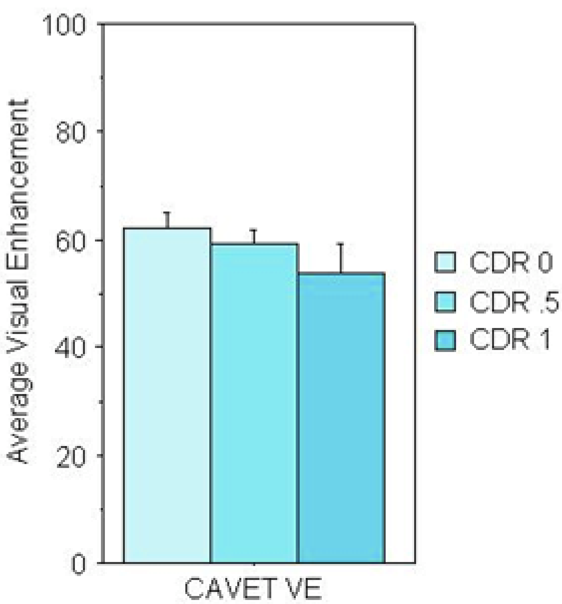

Figure 2 displays adjusted means for VE for word-level stimuli. Error bars indicate standard error.

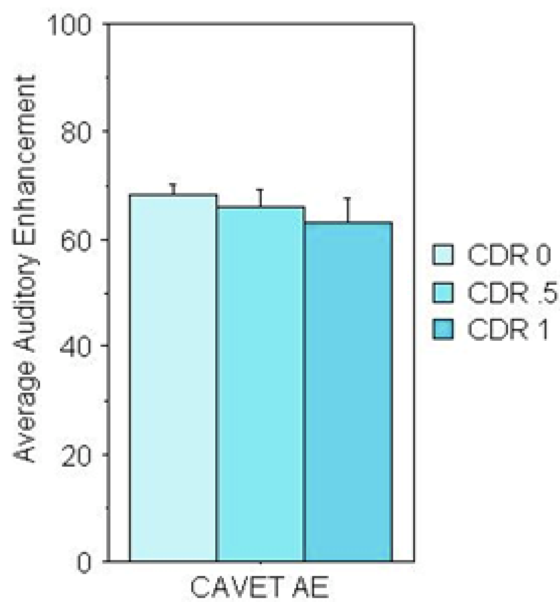

Figure 3 displays adjusted means for AE for word-level stimuli. Error bars indicate standard error.

Iowa Consonant Test

Performance in A, V, and AV Conditions

Figure 4 displays mean percent correct scores, split by CDR group, in the A-only, V-only, and AV conditions for the consonant-level stimuli. Initial visual inspection of Figure 4 indicated that A-only scores were fairly similar across the three groups: CDR 0 mean = 38.4 % (SD = 9.8), CDR 0.5 mean = 35.1 % (SD = 12.8), CDR 1 mean = 38.1 % (SD = 2.1). Similar to word-level stimuli, setting individual signal-to-babble ratio was successful in establishing similar performance across groups and values were above floor performance. The similarity across groups was verified by one-way analysis of variance (ANOVA) that compared A-only performance across the three CDR groups and found no difference between groups ($F(2, 10) < 1$; $p > .05$; $\eta^2_p = .07$).

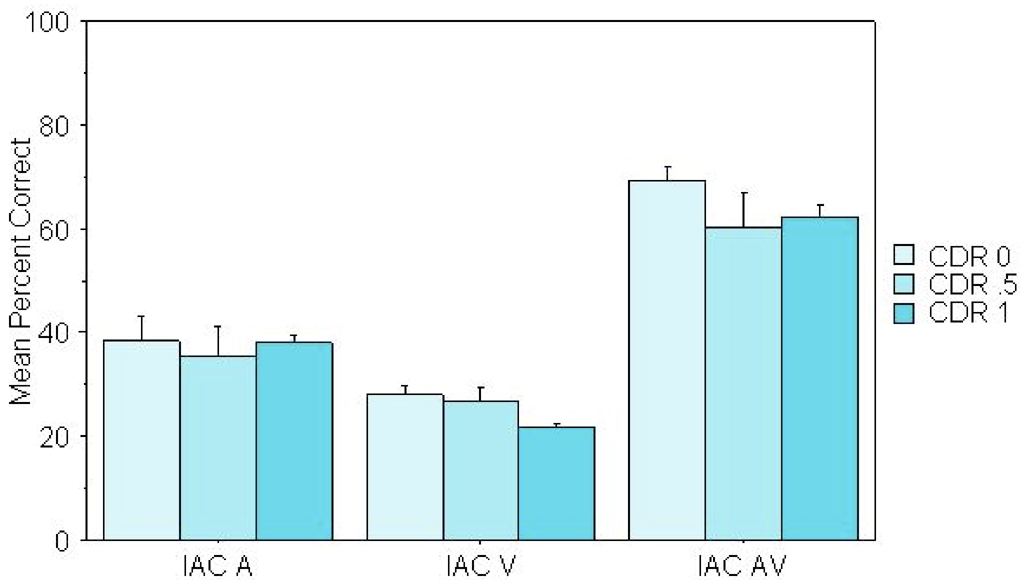

Figure 4 displays mean percent correct scores, split by CDR group, in the A-only, V-only, and AV conditions for the consonant-level stimuli (IAC).

Examination of V-only performance for consonant-level stimuli indicated relatively low performance when compared to A-only and AV conditions. Mean percent correct scores, split by CDR group are as follows: CDR 0 mean = 27.8 % (SD = 4.1), CDR 0.5 mean = 26.5 % (SD = 5.8), CDR 1 mean = 21.6 % (SD = 1.4). Potential differences across the three CDR groups in V-only performance were examined using one-way ANOVA. V-only scores across the three CDR groups were entered as a between subjects variable that found no difference between groups ($F(2,10) = 1.8$; $p > .05$; $\eta^2_p = .28$).

The results for the AV condition for consonant-level stimuli suggested a slight decline in AV speech perception ability for individuals in the impaired CDR groups: CDR 0 mean = 69.3% (SD = 6.0), CDR 0.5 mean = 60.3 % (SD = 14.6), CDR 1 mean = 62.3 % (SD = 3.9). Potential differences across the three CDR groups for AV performance were also investigated using one-way ANOVA. AV scores for the three CDR groups were entered as a between subjects variable. Results indicated no difference between groups for AV consonant performance ($F(2,10) = 1.0$; $p > .05$; $\eta^2_p = .18$).

Comparison of Visual and Auditory Enhancement

Figure 5 displays means for VE, split by CDR group, for the consonant- level stimuli. Figure 5 suggests that, for consonant-level stimuli, VE is poorer in those individuals with higher CDR ratings. Although none of the differences between the CDR groups reached statistical significance ($F(2,10) = 2.25$; $p > .05$; $\eta^2_p = .34$) it is interesting to note this trend as a function of CDR rating.

Similar to the analysis of VE, AE differences among the CDR groups were compared. Figure 6 displays means for AE as a function of CDR group, for consonant- level stimuli. Potential AE differences across the three CDR groups were investigated using one-way ANOVA. AE differences across the three CDR groups were entered as a between subjects variable and no difference was found between groups ($F(2,10) < 1$; $p > .05$; $\eta^2_p = .15$).

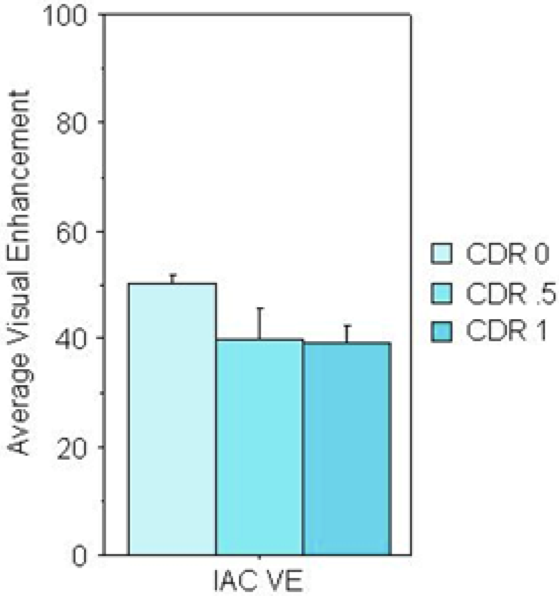

Figure 5 displays VE for consonant-level stimuli. Error bars indicate standard error.

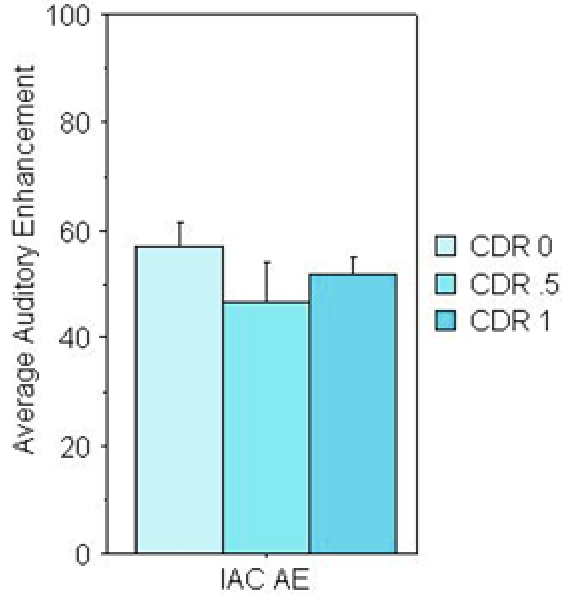

Figure 6 displays AE for consonant-level stimuli. Error bars indicate standard error.

Correlation of Words vs. Consonants

Results of the correlation analysis are shown in Table 5. Correlations between the scores for the measured variables (A-only, V-only and AV) are presented along with correlations between the derived measures of benefit (VE and AE). Of particular interest to the current study are the correlations between the two measures of enhancement among the three groups. These correlations are shown in Table 6. For the CDR 0 group there was no correlation between consonant- and word-level stimuli for the VE and AE measures ($r = .171$, $p = .806$ and $r = .529$, $p = .405$ respectively). The results for the CDR 0.5 group also did not show correlations between the two types of stimuli (VE, $r = .188$, $p = .788$ and AE, $r = .374$, $p = .579$). Due to the very small sample sizes (CDR 0, $n = 5$; CDR .5, $n = 5$; CDR 1, $n = 3$) the significance testing software was not able to calculate the degree of difference between the two types of stimuli for the CDR 1 group. Further, for the same reason, all correlations calculated within the CDR groups are difficult to interpret.

For further analysis lipreading scores for the two types of stimuli were also compared. Results indicated that, as a whole, lipreading (V-only) scores for the two types of stimuli were not correlated (r = .099, p = .755). Nor were they correlated within each of the CDR groups (CDR 0, r = .128, p = .855; CDR 0.5, r = .404, p = .544; CDR 1, could not compute significance).

Correlation Matrix

	CAVET A	CAVET V	CAVET AV	IAC A	IAC V	IAC AV	CAVET VE	CAVET AE	IAC VE	IAC AE
CAVET A	1.000	.573	.631	.245	.077	.340	-.076	.457	.400	.345
CAVET V	.573	1.000	.648	.249	.099	.283	.329	.186	.262	.281
CAVET AV	.631	.648	1.000	.322	.338	.532	.724	.867	.605	.475
IAC A	.245	.249	.322	1.000	-.039	.819	.235	.274	.398	.873
IAC V	.077	.099	.338	-.039	1.000	.316	.348	.395	.523	.079
IAC AV	.340	.283	.532	.819	.316	1.000	.388	.524	.843	.970
CAVET VE	-.076	.329	.724	.235	.348	.388	1.000	.705	.400	.316
CAVET AE	.457	.186	.867	.274	.395	.524	.705	1.000	.630	.448
IAC VE	.400	.262	.605	.398	.523	.843	.400	.630	1.000	.755
IAC AE	.345	.281	.475	.873	.079	.970	.316	.448	.755	1.000

Table 5 Correlations between the scores for the measured variables (A-only, V-only and AV) presented with correlations between the derived measures of benefit (VE and AE).

Correlation Matrix
Split By: CDR Rating
Cell: .CDR 0

	CAVET A	CAVET V	CAVET AV	IAC A	IAC V	IAC AV	CAVET VE	CAVET AE	IAC VE	IAC AE
CAVET A	1.000	.602	.552	.653	-.281	.441	-.540	.308	-.141	.449
CAVET V	.602	1.000	.530	.685	.128	.643	-.095	.069	.358	.514
CAVET AV	.552	.530	1.000	.801	-.651	.607	.402	.882	-.050	.685
IAC A	.653	.685	.801	1.000	-.613	.943	.126	.567	.452	.955
IAC V	-.281	.128	-.651	-.613	1.000	-.515	-.350	-.842	-.045	-.691
IAC AV	.441	.643	.607	.943	-.515	1.000	.180	.367	.721	.975
CAVET VE	-.540	-.095	.402	.126	-.350	.180	1.000	.534	.171	.242
CAVET AE	.308	.069	.882	.567	-.842	.367	.534	1.000	-.243	.529
IAC VE	-.141	.358	-.050	.452	-.045	.721	.171	-.243	1.000	.615
IAC AE	.449	.514	.685	.955	-.691	.975	.242	.529	.615	1.000

5 observations w ere used in this computation.

Correlation Matrix
Split By: CDR Rating
Cell: CDR .5

	CAVET A	CAVET V	CAVET AV	IAC A	IAC V	IAC AV	CAVET VE	CAVET AE	IAC VE	IAC AE
CAVET A	1.000	.568	.793	-.115	.666	.381	.061	.749	.819	.283
CAVET V	.568	1.000	.578	-.071	.404	.210	.254	.375	.410	.160
CAVET AV	.793	.578	1.000	.172	.972	.516	.656	.973	.733	.357
IAC A	-.115	-.071	.172	1.000	.282	.873	.427	.222	.462	.895
IAC V	.666	.404	.972	.282	1.000	.550	.766	.988	.677	.382
IAC AV	.381	.210	.516	.873	.550	1.000	.378	.539	.835	.982
CAVET VE	.061	.254	.656	.427	.766	.378	1.000	.664	.188	.239
CAVET AE	.749	.375	.973	.222	.988	.539	.664	1.000	.731	.374
IAC VE	.819	.410	.733	.462	.677	.835	.188	.731	1.000	.777
IAC AE	.283	.160	.357	.895	.382	.982	.239	.374	.777	1.000

5 observations w ere used in this computation.

Correlation Matrix
Split By: CDR Rating
Cell: CDR 1

	CAVET A	CAVET V	CAVET AV	IAC A	IAC V	IAC AV	CAVET VE	CAVET AE	IAC VE	IAC AE
CAVET A	1.000	.924	.961	.866	-.569	.866	.907	.537	.662	.836
CAVET V	.924	1.000	.782	.991	-.213	.610	.677	.175	.327	.563
CAVET AV	.961	.782	1.000	.693	-.775	.971	.988	.750	.844	.955
IAC A	.866	.991	.693	1.000	-.082	.500	.574	.043	.199	.449
IAC V	-.569	-.213	-.775	-.082	1.000	-.904	-.863	-.999	-.993	-.927
IAC AV	.866	.610	.971	.500	-.904	1.000	.996	.887	.948	.998
CAVET VE	.907	.677	.988	.574	-.863	.996	1.000	.843	.916	.989
CAVET AE	.537	.175	.750	.043	-.999	.887	.843	1.000	.988	.912
IAC VE	.662	.327	.844	.199	-.993	.948	.916	.988	1.000	.965
IAC AE	.836	.563	.955	.449	-.927	.998	.989	.912	.965	1.000

3 observations w ere used in this computation.

Table 6 Correlation matrices for each CDR group demonstrating correlations between auditory enhancement (AE) and visual enhancement (VE) with consonants and words.

Correlations between Cognitive Measures and AV Benefit

Results of the correlational analysis between the composite scores from the cognitive measures provided by the ADRC (episodic memory (EM), semantic memory (SM), working memory (WM), and visuospatial composite (VS)) and the derived enhancement variables (VE and AE) are shown in Table 7. For further analysis, the correlations between the cognitive measures and lipreading (V-only) scores are also shown in Table 7. Because statistical differences did not exist for VE, AE, and V-only performance between the CDR groups, comparisons between performance on the speech perception tasks and the four composite scores were made without dividing the groups. This allowed for a continuous range of cognitive function to be used in the correlation analyses. The results the correlational analysis between the speech variables and the composite variables must be interpreted with caution not only because of the small sample sizes, but also because conducting multiple correlations on the same sample can cause sporadic significant correlations to occur. For this reason a conservative *p*-value was adopted to indicate significance. The new adjusted *p*-value was determined by dividing the criteria by the number of comparisons made between the speech variables and the composite variables (.05/24 = .002). Using this strict *p*-value, none of the comparisons reached significant levels.

Correlation Matrix

	episodic	semantic	working	visuospatial	CAVET VE	CAVET AE	CAVET V	IAC VE	IAC AE	IAC V
episodic	1.000	.796	.752	.835	.322	.431	.154	.537	.191	.432
semantic	.796	1.000	.802	.785	.279	.165	.226	.408	-.042	.544
working	.752	.802	1.000	.723	.220	.201	.043	.395	-.116	.436
visuospatial	.835	.785	.723	1.000	.288	.302	.246	.432	.203	.221
CAVET VE	.322	.279	.220	.288	1.000	.705	.329	.400	.316	.348
CAVET AE	.431	.165	.201	.302	.705	1.000	.186	.630	.448	.395
CAVET V	.154	.226	.043	.246	.329	.186	1.000	.262	.281	.099
IAC VE	.537	.408	.395	.432	.400	.630	.262	1.000	.755	.523
IAC AE	.191	-.042	-.116	.203	.316	.448	.281	.755	1.000	.079
IAC V	.432	.544	.436	.221	.348	.395	.099	.523	.079	1.000

Table 7 Correlation matrix demonstrating results between the composite scores for episodic memory (EM), semantic memory (SM), working memory (WM), and visuospatial composite (VS) and auditory enhancement (AE) and visual enhancement (VE) with the correlations between the cognitive measures and lipreading scores (V-only).

A few of the comparisons, although not significant, did meet notable levels of correlation. For example, a pattern emerges between the consonant V-only scores semantic memory composite (r = .544, p = .05). Figure 7 shows the scatterplot for consonant V-only performance and the semantic memory composite. To a lesser extent, the relationship between the consonant VE scores and the composite scores for episodic memory was also noteworthy (r = .537, p = .06). Figure 8 shows the scatterplot of consonant VE scores with the composite scores for episodic memory.

Although the potential for inter-correlation between the variables is high, results of the correlational analysis also produced a few other noteworthy comparisons between scores taken from the consonant testing and the ADRC's composite scores. For example the correlation coefficient between consonant V-only scores and both the episodic (r = .432) and working memory (r = .432) composite scores were mentionable. Scatterplots for these two comparisons are shown in Figures 9 and 10 respectively.

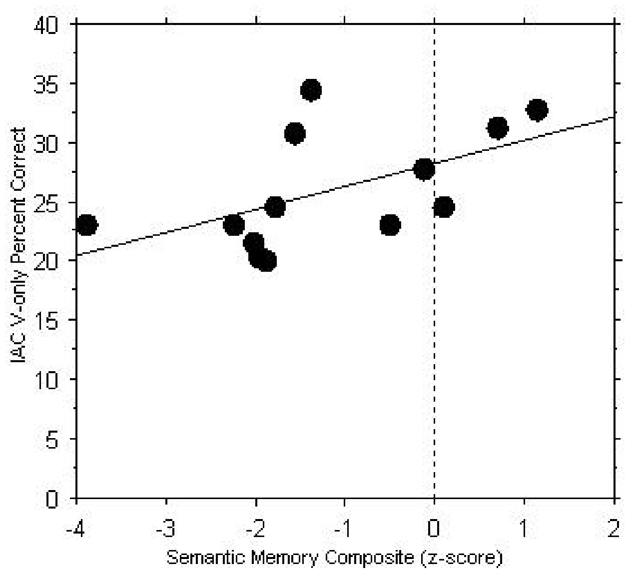

Figure 7 Scatterplot for consonant V-only performance correlated with the semantic memory composite. Solid line shows the best fitting regression line for all participants.

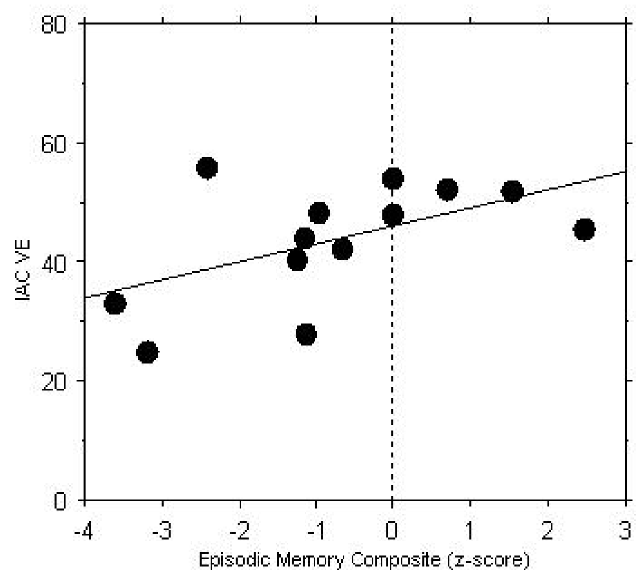

Figure 8 Scatterplot for consonant visual enhancement (VE) correlated with the episodic memory composite. Solid line shows the best fitting regression line for all participants.

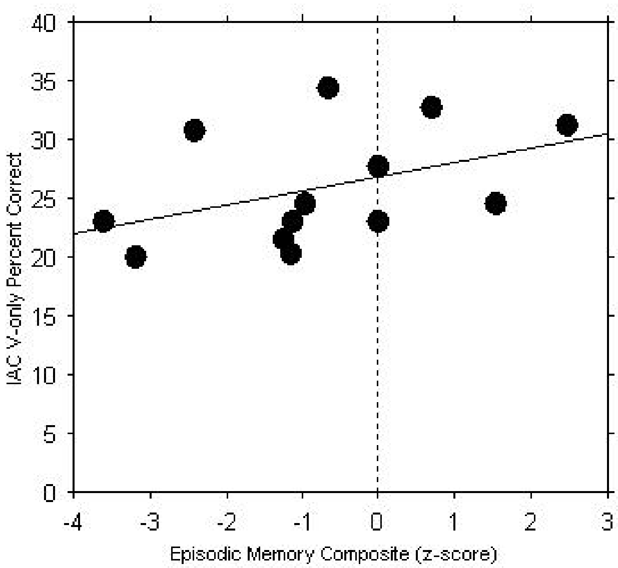

Figure 9 Scatterplot for consonant V-only performance correlated with the episodic memory composite. Solid line shows the best fitting regression line for all participants.

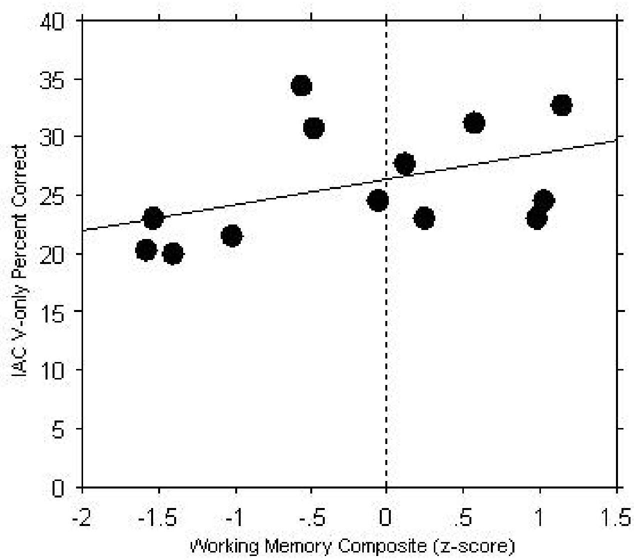

Figure 10 Scatterplot for consonant V-only performance correlated with the working memory composite. Solid line shows the best fitting regression line for all participants.

Discussion

The present study was designed to investigate the effects of DAT on the ability to benefit from combining auditory and visual speech information. Moreover, two degrees of DAT severity based on the CDR scale (0.5 and 1) allowed for comparison in terms of degrees of DAT. The statistical findings indicated that healthy, older adults and individuals with DAT exhibit comparable benefits, as indexed by measures of both VE and AE, for both consonant- and word-level stimuli. A possible trend, however, in the enhancement data was noteworthy. Specifically, VE among individuals with CDR ratings of 0.5 and 1 showed slightly less benefit from the addition of the visual channel to the ongoing auditory signal. Correlations between scores on the two different stimulus types indicated that performance on the word-level task was not correlated with performance on the consonant-level task. Finally, correlations between cognitive variables and the speech perception variables did not show correlations with the various memory composite scores and the visuospatial processing composite score provided by the ADRC. Noteworthy relationships between lipreading ability and semantic working memory along with episodic memory were found, however these relationships failed to reach significance after the multiple comparisons correction was made to the statistical analyses.

Peripheral Hearing Sensitivity

Pure-tone testing results parallel those found in previous research examining peripheral hearing loss in individuals with DAT (Gates, Karzon, Garcia, Peterein, Storandt, Morris, & Miller, 1995). Individuals with DAT continue to show a range of thresholds comparable to healthy, older adults. Pure-tone averages (500, 1000, 2000, 4000 Hz) revealed a range of normal to mild hearing loss for the better ear.

CDR status and Auditory-visual Performance

The present study investigated behavioral research in auditory-visual speech perception in individuals with DAT. Previous neuropsychological research in this population has focused on integration using the McGurk effect (Delbeuck et al., 2007). As mentioned above the McGurk effect involves simultaneously presenting participants with discrepant auditory and visual information in order to produce an illusory response. Because with the McGurk effect, successful integration is measured by the number of errors a listener makes it is difficult to determine if the response is a result of a unimodal error or, as the McGurk effect would suggest, a product of integration. A review of the confusion matrices produced by Miller and Nicely (1955) shows that the common phonemic error responses typically made in difficult listening situations are also the correct response in the McGurk effect. The present study utilized a different approach for examining different components of AV speech perception. The findings of the current study provide insight into the potential patterns in individuals with DAT for lipreading (V-only) and visual enhancement (VE) ability, as well as potential contributions of cognitive factors to auditory-visual speech perception.

The V-only mean percent correct score obtained for the DAT participants in current study were comparable to the results found in previous studies using healthy, older adults (Sommers et al., 2005; Cienkowski & Carney, 2002; Dancer et al., 1994). For example, Sommers et al. reported a mean percent correct score of approximately 20% (SD = 10) with consonant-level and approximately 30% (SD = 5) for word-level stimuli (values are interpolated from Figure 1 of Sommers et al., 2005). This value is similar to those found in our sample of individuals with DAT for either stimulus type. In the current study the CDR .5 and 1 scores for consonant-level stimuli were 26.5% (SD = 5.8) and 21.6% (SD = 1.4) respectively. Word-level scores were

21.0% (SD = 5.4) for the CDR .5 group and 23.3% (SD = 18.9) for the CDR 1 group. Thus, the findings for V-only performance from the present study are comparable with those of V-only performance in previous investigations that examined healthy older adults. Results, suggest that participants with DAT have a similar ability to encode visual speech information as healthy, older adults.

Although the findings for VE did not reveal differences among the CDR groups a pattern did emerge. Close inspection of Figures 2 and 5 show a possible avenue for future study. This possible decline in VE may be a reflection of the effects of increased cognitive impairment in capacities that contribute to an individual's ability to benefit from the addition of the visual signal to a congruent auditory signal. Viewed in conjunction with our data and previous findings citing the lack of correlation for enhancement between word- and consonant- levels task (Sommers et al., 2005, Grant & Seitz, 1998), it appears that different mechanisms mediate speech perception for words and consonants. Specifically, Grant et al. (1998) proposed an integration model that explains that AV perception of words is determined by both bottom-up extraction of auditory and visual cues coupled with top-down usage of lexical, syntactic, and semantic processing whereas for consonants AV perception is determined primarily by bottom-up processing. Thus, findings for VE with word-level stimuli would reflect the increased importance in top-down cognitive contributions for word-level speech perception. Furthermore, the areas noted to contribute to top-down processing are areas of known decline in individuals with DAT.

The correlations with the cognitive measures provided by the ADRC and the speech perception measures (A-only, V-only, AV, VE or AE) did not reveal significant relationships. On the other hand, there were some comparisons that showed potential for being correlated and the lack of significance may be due to the limited sample size. Some of these include the potential relationship between lipreading and semantic memory composite for consonant identification. Semantic memory deficits are believed to impact an individual's ability to appropriately label, name, or describe things (Warrington & Shallice, 1984). If, in future studies this trend continues, this may suggest that an individual with poor semantic memory for visual presentations may not be able to label the visual stimuli appropriately.

Limitations

To our knowledge, the current study represents one of the first demonstrations of behavioral research in auditory-visual speech perception in individuals with DAT. It is evident that the greatest limitation of the current study was the insufficient number of participants needed to detect significant differences. Additionally, when interpreting the data it is important to consider potential contributions to the absence of correlations. With regards to correlations across types of stimuli, consonant stimuli were tested using a closed-set test format whereas word-level stimuli were assessed in an open set format. Additionally the talkers for each type of stimuli differed; a male talker produced the consonant-level stimuli and a female-talker produced the word-level stimuli. Although the individual tests differ, similar measures of AV speech perception have been utilized in a number of studies and therefore, comparisons across studies are possible (Sommers et al., 2005, Tye-Murray, Sommers, and Spehar, 2007, Grant et al., 1998).

Clinical Implications

The findings from the present study reveal potential opportunities to improve patient care among the DAT population. There are limited behavioral measures that clearly delineate differences between healthy, older adults and individuals with DAT. Additional testing is needed to determine if the findings from this study regarding enhancement could provide potential avenues for developing clinically applicable behavioral measures of DAT categorization. Previous literature (Grant et al., 1998; Sumby & Pollack, 1954) demonstrates that individuals have better speech intelligibility when they can hear and see the talker, this same phenomena was observed in our population of individuals with DAT. It is important instruct caretakers to communicate with individuals with DAT using face-to-face interaction as it provides the most speech information for the individual to utilize.

Conclusions and Future Research

The current research provided novel information about auditory-visual speech perception in individuals with dementia of the Alzheimer's type. Future research should include a larger sample size in order to more thoroughly investigate the affects of dementia of the Alzheimer's type on auditory-visual speech perception. Specifically research should focus on measures of enhancement and cognitive influences. An important goal for future research may also include identifying the detailed demands imposed by integration and to specify why and how those demands correlate with enhancement measures. Finally, the present study used consonant- and word-level stimuli; future research may extend this investigation using sentence-level or discourse stimuli to investigate the role of semantic, syntactic, and comprehension resources needed for auditory-visual speech perception. Continued research on dementia of the Alzheimer's type is needed to expand our understanding of auditory-visual speech perception abilities in this population and to further understand the progression of the disease.

References

Alzheimer’s Association. (2010). Alzheimer’s Association Report: 2010 Alzheimer’s disease facts and figures. *Alzheimer’s & Dementia, 6,* 158-194.

American Speech-Language-Hearing Association. (1988). Guidelines for determining threshold level for speech. *ASHA, 30*, 85-89.

Armitage S.G. (1946). An analysis of certain psychological tests used in the evaluation of brain injury. *Psych Mono, 60*, 1-48.

Baddeley, A., Loggie, R., Bressi, S., Della Sala, S., & Spinnler, H. (1986). Dementia and working memory. *Quarterly J Experimental Psycho, Section A, 38(4)*, 603-618.

Baddeley, A., Bressi, S., Della Sala, S., Logie, R., & Spinnler, H. (1991). The decline of working memory in Alzheimer’s disease: A longitudinal study. *Brain, 114,* 2521-2542.

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C. R., & McGuire, P.K. (1997). Activation of auditory cortex during silent lipreading. *Science*, *276*, 593–596.

Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current Biology*, *10*, 649–657.

Cerella, J. & Hale, S. (1994). The rise and fall of information processing rates over the life span. *Acta Psychologica, 86,* 109-197.

Chertkow, H & Bub, D. (1990). Semantic memory loss in dementia of Alzheimer's type. What do various measures measure?. *Brain, 113(2), 397-417.*

Cienkowski K.M., & Carney A.E. (2002). Auditory–visual speech perception and aging. *Ear Hear*, *23*, 439-449.

Committee on Hearing, Bioacoustics, and Biomechanics (CHABA). (1988). Speech Understanding and Aging. *J Acoust Soc Am, 83,* 859 – 893.

Dancer, J., Krain, M., Thompson, C., Davis, P., & Glen, J. (1994). A cross-sectional investigation of speechreading in adults: Effects of age, gender, practice, and education. *Volta Review, 96*, 31–40.

Delbeuck, X., Collette, F., Van der Linden, M. (2007). Is Alzheimer's disease a disconnection syndrome? Evidence from a crossmodal audio-visual illusory experiment. *Neuropsychologia, 45(14),* 3315-3323.

Drzezga, A., Grimmer, T., Peller, M., Wermke, M., Siebner, H., Rauschecker, J.P., Schwaiger, M., & Kurz, A. (2005). Impaired cross-modal inhibition in Alzheimer's disease. *PLoS Medicine, 2(10),* 0986-0995.

Elphick, R. (1996). Issues in comparing the speechreading abilities of hearing-impaired and hearing 15 to 16 year-old pupils. *Bri J Educational Psychology, 66*, 357–365.

Feld, J. & Sommers, M. (2009). Lipreading, processing speed, and working memory in younger and older adults. *J Speech Lang Hear Res, 52,* 1555-1565.

Fingelkurts, A.A., Fingelkurts, A.A., Krause, C.M., Mottonen, R., & Sams, M. (2003) Cortical operational synchrony during audio-visual speech integration. *Brain Lang, 85(2)*, 297-312.

Fischer, M.H. (2001). Probing spatial working memory with the Corsi Blocks task. *Brain Cognition, 45,* 143-154.

Frisina, D.R., & Frisina, R.D. (1997). Speech recognition in noise and presbycusis: relations to possible neural mechanisms. *Hear Res, 106,* 95-104.

Gates, G.A., & Mills, J.H. (2005). Presbycusis. *Lancet, 366,* 1111-1120.

Gates, G.A., Karzon, R.K., Garcia, P., Peterein, J., Storandt, M., Morris, J.C., & Miller, J.P. (1995). Auditory dysfunction in aging and senile dementia of the Alzheimer's type. *Arch Neuro, 52(6),* 626-634.

Golob, E.J., Miranda, G.G., Johnson, J.K., & Starr, A. (2001). Sensory cortical interactions in aging, mild cognitive impairment, and Alzheimer's disease. *Neurobio Aging, 22(5),* 755-763.

Goodglass, H., & Kaplan, E. (1983). The Assessment of Aphasia and Related Disorders. 2 ed. Philadelphia: Lea & Febiger.

Grant, K.W., & Seitz, P.F. (1998). Measures of auditory-visual integration in nonsense syllables and sentences. *J Acoust Soc Am, 104(4),* 2438-2450.

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. *J Acoust Soc Am, 103(5)*, 2677-2690.

Grober, E., Buschke, H., Crystal, H., Bang, S., & Dresner, R. (1988). Screening for dementia by memory testing. *Neurology, 3,* 900-903.

Heyanka, D.J., Mackelprang, J.L., Golden, C.J., & Marke, C.D. (2010). Distinguishing Alzheimer's disease from vascular dementia: An exploration of five cognitive domains. *Int J Neuroscience, 120(6),* 409-414.

Huntley, J.D. & Howard, R.J. (2010). Working memory in early Alzheimer's disease: a neuropsychological review. *Int J Geriatr Psychiatry, 25*, 121–132.

Jenkins, L., Myerson, J., Hale, S., & Fry, A. (1999). Individual and developmental differences in working memory across the life span. *Psychonomic Bulletin & Review, 6*, 28-40.

Jenkins, L., Myerson, J., Joerding, J. A., & Hale, S. (2000). Converging evidence that visuospatial cognition is more age-sensitive than verbal cognition. *Psychology and Aging, 15,* 157–175.

Johnson, D.K., Storandt, M., Morris, J.C., & Galvin, J.E. (2009). Longitudinal study of the transition from healthy aging to Alzheimer's disease. *Arch Neurol, 66,* 1254-1259.

Lidestam, B., Lyxell, B., & Andersson, G. (1999). Speech-reading: Cognitive predictors and displayed emotion. *Scandinavian Audiology, 28,* 211–217.

Lyxell, B., & Holmberg, I. (2000). Visual speechreading and cognitive performance in hearing impaired and normal hearing children (11–14 years). *Bri J Educational Psychology*, *70,* 505–518.

Lyxell, B., & Rönnberg, J. (1989). Information-processing skill and speech-reading. *Bri J Audiology*, *23,* 339–347.

Lyxell, B., & Rönnberg, J. (1992). The relationship between verbal ability and sentence-based speechreading. *Scandinavian Audiology, 21,* 67–72.

Macaluso, E., George, N., Dolan, R., Spence, C., & Driver, J. (2004). Spatial and temporal factors during processing of audiovisual speech: a PET study. *NeuroImage, 21,* 725-732.

MacDonald, M.C., Almor, A., Henderson, V.W., Kempler, D., & Andersen, E.S. (2001). Assessing working memory and language comprehension in Alzheimer's disease. *Brain Lang, 78(1),* 17-42.

MacLeod, A., & Summerfield, Q. (1990). A procedure for measuring auditory and audio–visual speech-reception thresholds for sentences in noise: Rationale, evaluation, and recommendations for use. *Bri J Audiology, 24*, 29–43.

McGurk, H. & MacDonald, J. (1976). Hearing lips seeing voices. *Nature, 264(5588), 746- 748.*

Metzler-Baddeley, C. (2007). A review of cognitive impairments in dementia with Lewy bodies relative to Alzheimer's disease and Parkinson's disease with dementia. *Cortex, 43(5),* 583-600.

Miller, G.A. & Nicely, P.E. (1955). An analysis of perceptual confusions among English consonants. *J Acoust Soc Am, 27(2),* 338-352.

Morris, J. (1993). The Clinical Dementia Rating: Current version and scoring rules. *Neurology*, *43(11)*, 2412-2414.

Myerson, J., Hale, S., Rhee, S. H., & Jenkins, L. (1999). Selective interference with verbal and spatial working memory in young and older adults. *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 54,* 161–164.

Salthouse, T.A. (1991). Mediation of adult age differences in cognition by reductions in working memory. *Psych Sci, 2,* 179-183.

Salthouse, T.A. (1994). The aging of working memory. *Neuropsych, 8,* 535-543.

Salthouse, T.A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging, 30(4),* 507-514.

Salthouse, T.A. (2010). Selective review of cognitive aging. *J Int Neuropsych Soc, short review,* 1-7.

Sekiyama, K., Kanno, I., Mirua, S., & Sugita, Y. (2003). Auditory-visual speech perception examined by fMRI and PET. *Neuroscience Res, 47(3)*, 277-287.

Simmons, A. A. (1959). Factors related to lipreading. *J Speech Hear Res*, *2,* 340–352.

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

Spehar B, Tye-Murray N, Sommers M. (2004). Time-compressed visual speech and age: a first report. *Ear Hear, 25,* 565-572.

Spinnler, H., Della Sala, S., Bandera, R., & Baddeley, A. (1988). Dementia, ageing, and the structure of human memory. *Cog Neuropsych, 5(2),* 193-211.

Sumby, W.H. & Pollack, I. (1954). Visual contributions to speech intelligibility in noise. *J Acoust Soc America, 26,* 212-215.

Thurstone, L.L., Thurstone, L.G. (1949). Examiner Manual for the SRA Primary Mental Abilities Test. Chicago: Science Research Associates.

Tye-Murray, N. & Geers, A. (2001). *Children's Audiovisual Enhancement Test*. St. Louis, MO: Central Institute for the Deaf.

Tye-Murray, N., Sommers, M.S., & Spehar, B. (2007). Audiovisual integration and lipreading abilities of older adults with normal and impaired hearing. *Ear Hear, 28(5),* 656-668.

Tyler, R.D., Preece, J., & Tye-Murray, N. (1986). *The Iowa Laser Videodisk Tests.* Iowa City, Iowa: University of Iowa Hospitals.

Warrington, E.K. & Shallice, T. (1984). Category specific semantic impairments. *Brain, 107(3), 829-853.*

Waters, G.S. & Caplan, D. (1997). Working memory and on-line sentence comprehension in patients with Alzheimer's disease. *J Psycholinguistic Res, 26(4),* 377-400.

Wechsler D. (1955). Manual: Wechsler Adult Intelligence Scale. New York: Psychological Corporation.

Wechsler D. (1981). Manual: Wechsler Adult Intelligence Scale-Revised. New York: Psychological Corporation.

Wechsler D. (1987). Manual: Wechsler Memory Scale-Revised. San Antonio, TX: Psychological Corporation.

Wechsler, D., Stone, C.P. (1973). Manual: Wechsler Memory Scale. New York: Psychological Corporation.

Appendix A

Psychometric Test Battery

(Provided by Martha Storandt, Ph.D., Professor of Psychology and Neurology at Washington University in St. Louis.)

A psychometric battery assessing a broad spectrum of abilities was administered to all participants, usually a week or two after the annual clinical assessment at the ADRC. Standardized test scores were averaged to form four composites. The episodic memory composite included the sum of the three free recall trials from the Selective Reminding Test (Grober, 1988), Associate Learning from the Wechsler Memory Scale (WMS) (Wechsler & Stone, 1973), and immediate recall of the WMS Logical Memory. The semantic memory composite included the Information subtest from the Wechsler Adult Intelligence Scale (WAIS) (Wechsler, 1955), the Boston Naming Test (Goodglass, 1983), and Animal Naming (Goodglass, 1983). The working memory composite included WMS Mental Control, Digit Span Forward and Digit Span Backward, and Letter Fluency for S and P (Thurstone, 1949). The visual spatial composite included the WAIS Block Design and Digit Symbol subtests and Trailmaking Test A and B (Armitage 1946).

The reference (normative) group used to standardized most of the tests prior to forming the composites was a sample (Johnson, 2009) of 310 people (*M* age = 74.5 years, *SD* = 8.6; *M* education = 14.8 years, *SD* = 3.2) who were enrolled as CDR 0, had at least one annual follow-up, but never progressed to CDR > 0. The means and standard deviations of three measures (Selective Reminding Test, Animal Naming, Trail Making B) not included in that report were based on the same robust sample but with slightly smaller sample sizes because these three tests were added to the battery after its initiation. Beginning September 1, 2005, the funding agency required changes in three tests in the battery: WMS Logical Memory was replaced by WMS-R Logical Memory Story (Wechsler, 1987); WAIS Digit Symbol was replaced by WAIS-R Digit Symbol; WMS Digit Span Forward and Backward was replaced by the WMS-R version. All but one of the changes were trivial and therefore are ignored here. The Logical Memory change, however, was substantial (i.e., gist scoring, only one story compared with verbatim scoring of two stories). Although smaller and with less follow-up than the group used for the other measures, the reference group used for *z*-score conversion for the WMS-R Story A was the first assessment of 78 people (*M* age = 73.4 years) enrolled with CDR 0 since September, 2005, who remained CDR 0 throughout follow-up: $M = 12.49$, $SD = 3.39$.