

ABSTRACT

HARDEE, JEFFERSON B. Reexamining Synthetic Speech: Intelligibility and the Effects of Age, Task, and Speech Type on Recall. (Under the direction of Christopher B. Mayhorn.)

Synthetic speech is a technology that can be utilized to convey information and aid people in their tasks. Older adults in particular are a population that may be able to benefit from synthetic speech, and they are a population that has been investigated in a limited capacity. The current researchers intended to elucidate lingering conflicts in previous research on the intelligibility and recall of word and stories in synthetic speech for older and younger adults and how that compared to similar conditions in natural speech. Twenty-four older and 24 younger adults completed intelligibility and recall tasks with word lists and stories. Results indicated that older adults had a more difficult time with all speech, natural speech was easier to understand and remember than synthetic speech, and stories were easier to recall than words. Results also indicated that older adults had a more difficult time understanding synthetic words as compared to natural words than younger adults. In addition, older adults improved differentially with the recall of stories as opposed to words when compared to the younger adult group. Potential directions for synthetic speech software design and future research are discussed.

**REEXAMINING SYNTHETIC SPEECH: INTELLIGIBILITY AND THE
EFFECTS OF AGE, TASK, AND SPEECH TYPE ON RECALL**

By
JEFFERSON B. HARDEE

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Master of Science

PSYCHOLOGY

Raleigh, North Carolina

2007

APPROVED BY:

Dr. Jason Allaire

Dr. Eric Wiebe

Chair of Advisory Committee
Dr. Christopher B. Mayhorn

BIOGRAPHY

Born in Wilmington, North Carolina, I have lived all my life between the coast and Raleigh. I graduated from South Brunswick High School near Southport, North Carolina in 1999, and I received my Bachelor's Degree in Computer Science from North Carolina State University in the spring of 2003. While I had thought computers would be my life, I came to realize that my interests were broader than I had initially thought. While pursuing that degree I became interested in Psychology, received a minor in Psychology, and applied to graduate school to continue my education. These past years have been a wonderful experience, and I hope that my growth as a student and professional can continue in the future with the same vigor.

TABLE OF CONTENTS

	Page
List of Tables	iv
Introduction	1
Perceptual Declines Associated with Age	1
Cognitive Declines Associated with Age	2
Delivering Cognitive Support via Technological Interventions	4
Older Adults and Computers	6
Potential Solution: Synthetic Speech	7
Comprehension and Intelligibility of Synthetic Speech	9
Recall of Synthetic Speech	10
Research Hypotheses	11
Method	12
Participants	12
Design	13
Materials / Apparatus	14
Procedure	15
Results	17
Intelligibility Analysis	17
Recall Analysis	18
Correlational Analysis	19
Regression Analysis	21
Discussion	24
Limitations	27
Future Research	28
References	30
Appendices	36
Appendix A: Word Lists for Intelligibility Task	37
Appendix B: Words Lists for Word Recall Task	39
Appendix C: Logical Memory Stories from Wechsler Memory Scale	42
Appendix D: Logical Memory Stories Created for the Present Experiment	43
Appendix E: Informed Consent Form for Younger and Older adults	44
Appendix F: Protocol and Instructions	49
Appendix G: Debriefing Form	51

List of Tables

		Page
Table 1	Mean scores and standard errors for word Intelligibility	18
Table 2	Mean scores and standard errors for word and story recall	19
Table 3	Correlations between measures of speech and cognitive and perceptual variables	20
Table 4	Intelligibility regressions	22
Table 5	Recall regressions	23

Introduction

Recent demographic trends indicate that the older population is increasing and that the average American is getting older. Consider the projection that the population of Americans over the age of 65 will grow from 34.9 million in 2000 to 54.5 million in 2020 (U.S. Census Bureau, 2004). The percentage of Americans over the age of 65 is expected to grow from 12.4 percent in 2000 to 16.3 percent in 2020 (U.S. Census Bureau, 2004). The U.S. Census Bureau projects that the population and percentage of older adults will continue to rise through the year 2050. Accompanying increased chronological age, perceptual and cognitive changes are well documented (Craik & Salthouse, 2000; Park & Schwartz, 2000).

Perceptual Declines Associated with Age

Generally, perception refers to the ability to sense and recognize information in the environment. All of the sensory modalities demonstrate age-related declines in sensitivity (Scialfa, Ho, & Laberge, 2004). However, the sense of sight is commonly regarded as the most important sensory system for day to day life. Older adults experience significant difficulties with respect to visual acuity (Kline & Scialfa, 1997; Scialfa et al., 2004). Estimates indicate that 68 percent of those with severe visual impairments are 65 or older (Kline & Scialfa, 1997).

Visual acuity is the ability to perceive and discriminate fine detail and is extremely important when attempting to read text. The most common visual problem in older adults, presbyopia, can increase blurring in vision. Presbyopia occurs when the lens of the eye becomes less malleable and cannot focus light properly on the receptors in the eye.

There are a host of other problems that are associated with age that can contribute to declines in vision including cataracts, glaucoma, and macular degeneration. A cataract is an

area of the lens of the eye that is breaking down, becoming opaque, and causing light to be scattered or blocked before reaching the receptors in the eye. Glaucoma is a condition of high pressure inside the eye which effectively causes a portion of the eye to die by cutting off the blood supply, resulting in gaps in vision, especially in the periphery. Macular degeneration is a condition that is most often associated with the thinning of the receptor tissue in the central areas of vision, which can eventually lead to blindness in the center of vision. All of these ocular diseases can cause a significant loss of visual acuity (Kline & Scialfa, 1997; Scialfa et al., 2004), which in turn can impair function on a number of everyday tasks such as reading or following traffic signals.

Clearly the problems that are associated with vision in old age are numerous and can be potentially severe. If visual information is small in size, such as with text fonts, then older adults will have difficulty perceiving that information properly (Hartley, 1994). Imagine sitting down to read the newspaper and finding that you are having trouble discriminating between letters in the headlines. Perhaps you cannot see the letters at all. In these situations, there is a distinct need for some type of compensatory mechanism to improve functionality. Because cognitive declines may also contribute to loss of functionality, the cognitive characteristics that can develop with age must first be examined before any means of compensation can be considered.

Cognitive Declines Associated with Age

Reading is a common task that requires the use of cognitive abilities. Text comprehension, which is generally slower for older than younger adults, involves the recognition of words, combining words into clauses or ideas, drawing conclusions and inferences, establishing temporal guides, and processing references to previous text (Kemper

& Mitzner, 2001). While a number of cognitive processes are involved with text comprehension, it has been theorized that working memory limitations in particular affect the post-interpretative processes involved in retention and information manipulation (Caplan & Waters, 1999). This can affect such things as the number of correct responses to questions about the content or inferences drawn from a piece of text. Thus, working memory is a key component in the process of receiving, retaining, and manipulating information.

One of the most well documented changes in cognition is the degradation of working memory that is associated with advancing age. Working memory is the short-term storage and manipulation of information in memory (Baddeley & Hitch, 1974). Working memory is used in nearly all daily activities ranging from reading the newspaper to making a meal. Aging research has shown decrements between the working memory of young and old adults (Craik, 2000; Salthouse, 1994). There is also evidence that working memory declines are related to the amount of material that is being remembered or manipulated, such that increasing the presentation of material results in increased difficulty for older adults (Craik, 1994). Slower processing speed has been a recurring explanation for the declines in cognitive tasks that are experienced by older adults (Salthouse, 1991). When pieces of information are presented rapidly, such as speech that has many words per minute, slower processing speed can greatly increase the demands placed on working memory. Each piece of data takes an increased amount of time to incorporate into memory. Capacity shrinkage (Craik, 1986) and a failure to inhibit distracting information (Hasher & Zacks, 1988) have also been used as explanations for this decline in working memory. Any of these explanations can contribute to an explanation of the effects of cognitive workload.

Cognitive workload is best described as the degree to which mental resources are used in the execution of a task. When working memory is taxed by either a large volume of incoming data or data that is semantically complex, workload is considered to be high. High workload can cause lapses in both perception and cognition (Sweller, 1993). The problem that older adults face with workload is their already diminished capacity with working memory. If working memory is diminished, then that leaves fewer resources to allocate. Thus, workload that may be bearable for younger adults with normal working memory capacity can be overwhelming for older adults. It is clear that older adults have some specific difficulties with memory and cognitive abilities. The next question that must be answered is how this situation can be remedied.

Delivering Cognitive Support via Technological Interventions

If cognitive tasks can be designed to utilize intact capabilities and minimize the impact of cognitive processes that undergo age-related declines, then powerful cognitive support can be delivered (Mayhorn, Rogers, & Fisk, 2004). Here cognitive support refers to the assistive aspects of technology that enhance the mental capabilities (and avoid the limitations) of users. Two key methods of providing cognitive support in this domain which utilize similar principles are context and redundancy gain.

While certain types of memory decline, others remain intact well into old age (Smith, 2002). Previous research has shown that context can be a powerful mediator of memory. The cognitive strength that is being emphasized is the stable functioning of semantic memory over the lifespan. Semantic memory is thought to store general facts or knowledge. Research has found little to no difference in the retention of semantic memories for young and old adults (Bowles, 1993; Light, 1992). In other words, older adults can have more trouble with

situations that are unfamiliar to them because they cannot draw on their maintained semantic knowledge to create situational context which helps guide their response (Hancock, Fisk, & Rogers, 2001). Older adults may benefit more from sentence context when trying to understand the spoken word, especially under conditions that are not conducive to speech recognition such as high noise or divided attention (Perry & Wingfield, 1994; Wingfield, Aberdeen, & Stine, 1991).

In addition to the benefits of context, redundant cues can be an important means of providing cognitive support to older adults. Redundancy gain refers to the idea that stimuli that represent the same information will be more easily recognized and understood if they are presented simultaneously in different sensory modalities (Wickens, Lee, Liu, & Becker, 2004). The use of multiple codes during memory encoding, specifically verbal and spatial codes, may improve the memorability of a word or object (Paivio, 1986). The vision problems discussed earlier that can occur in advanced age are poignant examples that any of the individual senses is not always enough. For instance, an older adult with severe cataracts will greatly benefit from having a verbal speech signal presented in concert with a reading task. The perceptual weakness of the eyes can be corrected through the combined strength of multiple modalities and cues. Emerging technologies are one means of delivering cognitive support mechanisms such as redundancy gain and context effects.

To illustrate how technology can play this role, imagine that an older adult is preparing to take her medications. She has a problem with visual acuity that is caused by her advanced presbyopia. She often makes errors while reading the product labels, and sometimes mistakes one medication for another. Some of her medications must be ingested with food, and some must be taken more than once per day. Potentially life threatening errors

may occur if the heart medication is taken once every two days instead of twice per day as intended. Consider a technological system that reads the medication label aloud to her as she follows along with her eyes, an excellent example of redundancy gain. The envisioned speech signal would help the older adult fill in the perceptual gaps in her reading. Visually indistinct words and letters could be differentiated by listening to the speech signal. Therefore, the redundant information captures a more complete message.

However, redundant cues and context are not just for older adults. The extent to which younger adults can benefit may not be as large, but they do still benefit from context and redundant cues. Designing products to serve individuals with functional limitations or disabilities can also improve the performance of those without limitations and can often reduce the long-term costs of developing that product (Nayak, 1995; Vanderheiden, 1997). These are some of the basic benefits of Universal Design, the idea that systems can be designed to be utilized for nearly all potential users. However, if a technology that could provide computer-based compensatory aids to older adults is to be incorporated into everyday life, then there must be some assurance that it will be used by that population.

Older Adults and Computers

A common myth concerning older adults is that they are somehow afraid of technology or unable to learn it effectively. In fact, older adults do often utilize modern technologies, and they are willing to learn about new technologies such as the personal computer. Researchers have found that there is little evidence to support the claim that older adults are reluctant to learn about computers (Morrell & Echt, 1997; Rogers, Mayhorn, & Fisk, 2004).

Contrary to this myth, there is evidence that computers can fit into the lives of older adults, and that computers have been established successfully in the homes of older adults (Czaja, Guerrier, Nair, & Landauer, 1993; Hahm & Bikson, 1989). There is also evidence that older adults are interested in learning more about computers and the internet (Morrell, Mayhorn, & Bennett, 2000).

There are actually many ways that computers can enrich the lives of older adults, and include electronic communication (e.g., e-mail) and information services (Czaja et al., 1993). Insurance and healthcare information can be accessed via the computer and internet as well. While a small segment of older adults uses the internet, the degree to which older adults use the internet is largely influenced by their demographics and access to computers and not chronological age. Chronological age seems to play a role, but its impact is not nearly as substantial (Morrell et al., 2000).

Considering that older adults demonstrated interest in computers and the information that can be obtained via the internet, it is not surprising that the percentage of older adults that are using computers has grown tremendously in recent years. The usage of computers in older populations has grown from 5% to 40% between 1993 and 2002 (Morrell et al., 2000; Rogers et al., 2004). Clearly the older adult population is one that can benefit from a technology based cognitive support for visual stimuli, especially reading.

Potential Solution: Synthetic Speech

One technology that might be employed to assist the older adult to both read and listen to the content of a medication label in the previous hypothetical example is already in the marketplace in the form of synthetic speech, specifically text-to-speech devices.

Synthetic speech is a technology that has developed since the advent of computers, and refers

to the ability to produce an approximation to the human speech signal, usually with a computerized device (Greene, Manous, & Pisoni, 1986). Researchers have attempted to study synthetic speech to understand what sets it apart from natural speech and how it can be utilized to assist the public.

Synthesized speech has been used in aircraft cockpits, computer-aided instruction, and as an aid to the visually impaired (Paris, Gilson, & Thomas, 1995). Synthetic speech is also being used in voice mail systems, public transportation systems, and as an aid to those without the ability to speak. While some synthetic speech systems use digitized recordings of the human voice, text-to-speech (TTS) devices use a string of printed text and convert it to an output signal with no human intervention (Greene et al., 1986). These devices are useful because they have no inherent limits to the vocabulary that they can pronounce. Thus, they can be used for reading text to the blind or the visually impaired.

TTS devices operate by taking a string of text and converting that text into a phonemic representation, similar to the phonetics one might find in a dictionary. The device then uses that representation to create digitized sounds that mimic human speech, compressing sounds and utilizing time lapses to achieve the production of words (Allen, Hunnicut, & Klatt, 1987).

One of the more successful TTS devices on the market is DECTalk. DECTalk has proven to have a relatively high degree of intelligibility which was documented at 95 percent with individual words (Greene et al., 1986). Intelligibility refers to the percentage of language units that can be understood and identified. These TTS devices can capitalize on redundancy gain by providing a speech signal in unison with the written text.

Older adults can be a prime audience for speech synthesizers due to various age-related declines such as vision. Producing redundant signals across the auditory channel can mitigate these deficits in vision. Speech synthesizers can make a vision-based task, such as reading the mail or newspaper, plausible or much easier for visually impaired users. However, older adults do face some difficulties while using these devices.

Comprehension and Intelligibility of Synthetic Speech

Comprehension of synthetic speech is consistently more difficult than with natural speech (Greene et al., 1986; Kangas & Allen, 1990; Mirenda & Beukelman, 1987). Smither (1993) and other researchers have linked synthetic speech to an increased cognitive workload and demand (Ralston, Pisoni, Lively, Greene, & Mullennix, 1991) and postulate that this increased workload is the cause of comprehension problems in synthetic speech.

Therefore, comprehension of synthetic speech is theorized to be more difficult for listeners with degraded cognitive abilities such as memory (Koul, 2003; Smither, 1993). Because age-related decreases in working memory have been frequently observed (Salthouse, 1994), older adults are likely to experience difficulty understanding synthetic speech. Several studies have shown advanced age to have no significant impact on the intelligibility or comprehension of synthetic speech (Drager & Reichle, 2001a; Humes, Nelson, & Pisoni, 1991; Humes, Nelson, Pisoni, & Lively, 1993). However, these data are inconclusive because others have found significant main effects of advanced age (Drager & Reichle, 2001b; Kangas & Allen, 1990; Roring, Hines, & Charness). Clearly, the question of whether age affects synthetic speech intelligibility or comprehension is one that needs to be reexamined.

Finally, context can play a powerful role in helping to correct some of the deficits that can occur with synthetic speech intelligibility (Drager & Reichle, 2001a). Participants in Drager and Reichle's (2001a) work were asked to repeat synthetic sentences that were presented alone or preceded by a context-setting story. They found that participants repeated more words correctly when sentences were preceded by a story. Roring et al. (2007) also found an effect of context when identifying monosyllabic words. Accompanying the issue of basic intelligibility, the recall of synthetic speech is an area that has received some attention as well. Important questions about the effects of age still remain in that domain.

Recall of Synthetic Speech

To date most investigations have focused on word recall to determine how well individuals can remember the contents of synthetic speech (Humes et al., 1993; Smither, 1993). Simple word by word understanding of synthetic speech is not enough to conclude that older adults can use synthetic speech as effectively as younger adults. One must also consider how much can be recalled by a listener. If a message or piece of information can not be recalled when it is important, then it does not produce the intended response in the appropriate situation. It would be analogous to the listener not hearing the message at all.

Recall of synthetic speech has been investigated to some extent, but results in this area of inquiry are inconsistent. Some researchers have found no effect of advanced age with the recall of synthetic speech (Humes et al., 1993), whereas others have found a main effect of age (Smither, 1993). However, a poignant limitation of Smither's work was that the recall of words was limited to the digits from zero to nine. In addition, Smither did not conduct a hearing test to assess the limitations of older adults. Additional research is needed to provide insight into this conflicting pattern of results.

Furthermore, the effects of context have not been studied with respect to synthetic recall. Theory and other research would suggest that context would have a beneficial effect for older and younger adults, and that older adults may benefit more than younger adults (Hancock et al., 2001; Perry & Wingfield, 1994; Wingfield et al., 1991). Therefore it is necessary to investigate the recall of discourse, as well as words, in an effort to directly measure the effects of context in synthetic speech recall.

Thus, the purpose of this research was to investigate the intelligibility of synthetic words for differing age groups and the recall of synthetic words and discourse as it relates to age while also comparing synthetic speech to natural speech. In this fashion the relationship between age and speech intelligibility could be examined, as well as the benefits of context. This research attempted to elucidate the source of the inconsistencies in the previous literature by examining the effects of age on the intelligibility and recall of synthetic speech (Drager & Reichle 2001a; Humes et al., 1991; Humes et al., 1993; Kangas & Allen, 1990; Smither, 1993). These factors were investigated using five hypotheses that follow from the aforementioned theory and research.

Research Hypotheses

The first hypothesis is concerned with the intelligibility of synthetic speech. The remaining four address synthetic speech recall and manipulations therein.

The first hypothesis is that the intelligibility of words is expected to be poorer for older adults. This would be consistent with theories on the increased workload imposed by synthetic speech in combination with theories of degraded working memory in older adults (Craig, 2000; Koul, 2003; Salthouse, 1994; Smither, 1993).

The second hypothesis is that recall is expected to be poorer for older adults across all conditions. This would reflect the increased cognitive demands of synthetic speech coupled with the diminished working memory capacity of older adults (Koul, 2003; Smither, 1993).

The third hypothesis is that the recall of discourse is expected to be better than the recall of isolated words. This would reflect the compensatory mechanism of context (Drager & Reichle, 2001a).

The fourth hypothesis is that the recall of natural speech is expected to be better than the recall of synthetic speech. The synthetic speech should increase cognitive demand, thus drawing resources away from the memory task (Sharit et al., 2003; Slowiaczek & Nusbaum, 1985).

The fifth hypothesis is that older adults will improve in their recall of discourse as opposed to words, and they will improve differentially more than younger adults (Hancock et al., 2001).

Method

Participants

The sample for the experiment consisted of two groups. The first group was comprised of college students from the Introduction to Psychology research pool available at N.C. State University. The age of this group ranged from 18 to 30 with a mean age of 20.3 years and a standard deviation of 2.37. These younger adults consisted of 12 males and 12 females. This group received course-related research credit for their participation. The second group consisted of community dwelling older adults. The older adults were recruited over the telephone via an existing list of older adults who have participated in past research. The age of this group ranged from 67 to 83 with a mean age of 73.1 years and a standard

deviation of 4.31. There were a total of 11 males and 13 females. The older adults were compensated for their time and participation. There were 48 total participants, equally divided between young and old participants. All participants were able to understand and fluently speak English.

Independent samples t-tests were used to examine the potential cognitive predictors of the participants. A significant difference was found between older and younger adult performance on the Shipley Institute of Living Scale, $t = 5.92, p < .001$. Older adults performed better on this scale of vocabulary knowledge. In addition, a significant difference was found for the measure of processing speed, the Digit Symbol Substitution task, $t = 8.67, p < .001$. Younger adults had a faster average processing speed than older adults. Finally, a significant difference was found for the measure of working memory, the Absolute Computation Span task, $t = 2.36, p = .02$. Younger adults recalled more during the task than older adults.

An independent samples t-test was also used to determine if hearing ability was different between younger and older adults. A significant difference was found, $t = 9.56, p < .001$. Younger adults had better hearing abilities on average.

Design

The study used a 2 (Age) x 2 (Task Type) x 2 (Speech type) mixed factorial design to address hypotheses two through five. Each participant completed word and discourse recall tasks to assess the effects of age, task type and speech type on recall. The first factor was the grouping variable of age (young vs. old). The second factor was the recall task type (word recall vs. discourse recall) which was observed within participants. The third factor was the type of speech (natural speech vs. synthetic speech) and was observed within participants.

These three factors were the independent variables, and the dependent variable was the percentage of words or discourse ideas that were correctly recalled.

In addition, each participant completed a word intelligibility task to assess the effects of age on speech intelligibility. The percentage of correctly identified words was used as the measure of speech intelligibility.

Materials / Apparatus

Each participant completed several tasks that used word lists as stimuli, including a speech intelligibility task and a word recall task. Two lists in the intelligibility task contained 50 words each, totaling 100 words (see Appendix A). Twelve lists in the recall task contained 15 words each, totaling 180 words (see Appendix B). All lists were adapted from lists in the Rey Auditory Verbal Learning Test (RAVLT; Schmidt, 1996). Words were compiled from existing lists of 50 words and 15 words in the RAVLT and repeated words were discarded and replaced with semantically or categorically similar words with similar syllable counts. For example, a duplicate word such as ‘plant’ might be replaced with the word ‘stem.’ (Toglia & Battig, 1978)

Participants were also asked to listen to four story passages during the discourse recall task. Two of these stories were synthetically produced while the remaining two were recorded by a male volunteer from North Carolina. One of the synthetic speech stories and one of the natural speech stories were taken from Wechsler Memory Scale, 3rd edition (see Appendix C; Wechsler, 1997). The other two stories were created using the Wechsler stories as a model (see Appendix D). Pacing and syllables counts were kept approximately equal, and all the stories revolved around a central character from a particular city that faces a

specific problem. Specifics properties of the stories can be found in their respective appendices.

All of the synthetic speech was produced by the DECtalk 5.0 demo in English, using an American accent. DECtalk's Perfect Paul was used for all instances of the synthetic speech condition. DECtalk has proven to be one of the very best synthesizers on the market, with the Perfect Paul voice being one of the most understandable voices in DECtalk (Greene et al., 1986). The synthetic speech was recorded and played by way of a standard pair of peripheral computer speakers. To mimic normal conversational sound levels, the sound level was approximately 70 decibels, as verified by a sound level meter.

Procedure

The experiment was administered in a lab room of North Carolina State University. Signage was posted outside the laboratory to inform others that there was an experiment in progress, and experimenters were instructed to minimize any possible noise while recorded speech was playing. Each participant was first presented with an informed consent form (see Appendix E) designed to describe the purpose of the study, and then given general instructions regarding task performance (See Appendix F). Afterwards, each participant was given a hearing test using an Earscan Audiometer (Micro Audiometrics Corp.) and asked basic demographic questions before the experiment began. In addition, each participant completed several individual difference measures, cognitive measures of processing speed, memory, and crystallized intelligence. The first of these measures was the computation span test, a test designed to assess working memory capacity. Furthermore, a digit-symbol substitution task was administered to the participants to gauge general cognitive processing speed. Finally, the Shipley Vocabulary assessment was administered to measure the

crystallized abilities of older and younger adults and ensure that the samples were representative of their respective populations (Shipley, 1986). As previously discussed, age-related deficits in working memory capacity and processing speed may be explanations for declines in word and discourse recall in older adults.

Participants were tested individually by a single experimenter. The intelligibility task consisted of the identification of synthetic or natural words as they were presented verbally. The participant listened to two lists of fifty words, one consisting of all natural words and one of all synthetic words. The participants were instructed to identify each word verbally. The experimenter recorded whether a word was correctly or incorrectly identified. The word recall and discourse recall tasks followed after.

Participants were asked to listen to twelve lists of 15 synthetic words during the word recall task. Participants verbally recounted all of the words that they could recall. The experimenter recorded any confabulations and tallied the number of words correctly recalled.

Participants were asked to listen to four stories in the discourse recall task. After listening to a story in its entirety, each participant was instructed to verbally recall as much about the story as possible. The experimenter was responsible for recording each idea unit that the participant correctly recalled. This within-subjects observation of task type allowed each participant to act as his or her own control.

The entire experiment required approximately one hour for each participant to complete. After the experiment, the participant was debriefed (see Appendix G) and dismissed.

Results

Several different analyses were undertaken to determine if a relationship between age and synthetic speech exists. The first such analysis focused on the intelligibility data, and whether age affected synthetic speech intelligibility. The second analysis was concerned with the effect of age and speech type on different memory tasks. Finally, correlations and regressions were performed to determine if any of the collected cognitive measures predicted intelligibility or memory performance. All analyses used an alpha level of .05

Intelligibility Analysis

A 2 (Age: younger vs. older adults) x 2 (Speech type: natural speech vs. synthetic speech) repeated measures analysis of variance was used to compare participants' natural speech intelligibility scores and synthetic speech intelligibility scores. This analysis was conducted to investigate the first hypothesis. A significant difference was found between the intelligibility of natural and synthetic words, indicating a main effect of speech type, $F(1,46) = 752.13, p < .001$. Synthetic words were more difficult to understand for all ages. A significant difference also existed between the intelligibility scores of older and younger adults, indicating a main effect of age, $F(1,46) = 20.70, p < .001$. Older adults had more difficulty understanding both natural and synthetic words as compared to younger adults. Means (out of 50) and standard errors can be seen in Table 1.

Finally, the interaction of age and speech type was significant, $F(1,46) = 20.58, p < .001$. This result supports the first hypothesis that synthetic speech is differentially more difficult to understand than natural speech for older adults as compared to younger adults.

Table 1

Mean scores and standard errors for word intelligibility

Age Group	Word Intelligibility	
	Natural	Synthetic
Younger adults	47.8 (.78)	34.1 (1.20)
Older adults	44.5 (1.06)	25.5 (1.22)

Recall Analyses

A 2 (Age: younger vs. older adults) x 2 (Task type: word recall vs. story recall) x 2 (Speech type: natural speech vs. synthetic speech) repeated measures analysis of variance was conducted to analyze the data pertaining to word and story recall and provide insight into hypotheses two through five. The first effect of note was a significant main effect of age, $F(1,46) = 4.46, p = .04$. Therefore, older adults had more trouble recalling both natural and synthetic speech in both tasks. This result corresponds with and supports the second hypothesis. The main effect of task type was also significant, $F(1,46) = 213.08, p < .001$. Performance during the recall of stories was significantly better than that of words regardless of age or speech type. This is consistent with the third hypothesis outlined. Finally, the main effect of speech type was also significant, $F(1,46) = 46.36, p < .001$. Natural speech was correctly recalled more often than synthetic speech regardless of age or task type. This final main effect supports the fourth hypothesis. Mean percentages recalled and standard errors (in parentheses) can be seen in Table 2.

Table 2

Mean scores and standard errors for word and story recall

Age Group	Word Recall		Story Recall	
	Natural	Synthetic	Natural	Synthetic
Younger adults	37.5% (.02)	29.7% (.01)	51.1% (.02)	47.0% (.03)
Older adults	26.6% (.02)	21.4% (.01)	50.6% (.02)	46.2% (.03)

There was also a significant first order interaction between age group and task type, $F(1,46) = 10.81, p = .002$. Older adults recovered more than younger adults between the recall of words and the recall of stories. That is, older adult performance improved more than younger adult performance through the presentation of story context. This interaction is consistent with the fifth hypothesis. All other effects were not significant.

Correlational Analysis

To further the examination of synthetic speech, a number of correlations were undertaken. These correlations were done as a preliminary step to regression analysis. The full table of correlations can be seen in Table 3.

Table 3

Correlations between measures of speech and cognitive and perceptual variables

Variables	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. Synthetic Intelligibility	-									
2. Natural Intelligibility	.76**	-								
3. Synthetic Word Recall	.60**	.41**	-							
4. Natural Word Recall	.58**	.44**	.81**	-						
5. Synthetic Story Recall	.34*	.28	.31*	.51**	-					
6. Natural Story Recall	.19	.07	.38**	.48**	.66**	-				
7. Crystallized Intelligence	-.31*	-.19	-.34*	-.27	.21	.10	-			
8. Processing Speed	.56**	.43**	.58**	.55**	.28	.11	-.39**	-		
9. Working Memory	.18	.13	.32*	.38**	.13	.23	.01	.30*	-	
10. Average Hearing Ability	-.84**	-.58**	-.58**	-.59**	-.14	-.15	.48**	-.69**	-.20	-
11. Age	-.63**	-.42**	-.55**	-.56**	-.05	-.04	.65**	-.80**	-.33**	.85**

* $p < .05$. ** $p < .01$.

The first result of note is that the speech measures themselves were interrelated. Intelligibility results were positively related to word recall results for both natural and synthetic speech. Synthetic story recall was positively related to synthetic intelligibility and synthetic word recall. Natural story recall was positively related to synthetic and natural word recall, as well as synthetic story recall.

From this analysis, it can also be concluded that the cognitive measures present in the study are generally related to one or more of the speech measurements. Particularly, processing speed was positively and significantly related to measures of speech intelligibility and the recall of individual words for both synthetic and natural speech. However, it was not related to story recall. Working memory was only related to individual word recall, an expected relationship. Finally, crystallized intelligence was negatively related to natural speech intelligibility and natural word recall. Therefore, as crystallized intelligence increased, natural speech performance declined. This result is likely an artifact of age. Crystallized intelligence increased with age and natural speech performance in intelligibility decreased with age.

Average hearing ability was related to many of the other measures. Average hearing ability was negatively related to the intelligibility and word recall measures, indicating those with poorer hearing had a more difficult time with speech intelligibility and word recall. However, hearing ability was not related to story recall. Finally, age was negatively related to the intelligibility and word recall results, but not the story recall. Age was positively related with Crystallized intelligence and negatively related to processing speed, results which match with previously discussed differences between the younger and older groups.

Regression Analysis

Following the correlational analysis, hierarchical regression analyses were done to determine the power of the recorded cognitive measures, age and hearing in predicting intelligibility performance.

First, a stepwise hierarchical regression was carried out to examine the effects of age, hearing, processing speed, crystallized intelligence and working memory on synthetic

intelligibility. The only two significant factors were age and hearing, accounting for 73% of the variance. The full analysis can be viewed in Table 4. Second, a similar procedure was done to predict natural intelligibility using the same predictors. Again, only age and hearing were significant predictors, accounting for 40% of the variance. It should be noted that age and hearing were much weaker predictors in the natural intelligibility model than they were in the synthetic intelligibility model.

Table 4

Intelligibility regressions

Regression of Predictor Variables on Performance		ΔR^2	Cumulative R^2	β
Synthetic Intelligibility	Age	.392	.392	-.626**
	Age + Hearing	.342	.734	-1.076**
	Age + Hearing + Processing Speed	.007	.741	.141
	Age + Hearing + Processing Speed + Crystallized Intelligence	.001	.742	.048
	Age + Hearing + Processing Speed + Crystallized Intelligence + Working Memory	.000	.742	.023
Natural Intelligibility	Age	.166	.166	-.408**
	Age + Hearing	.238	.404	-.852**
	Age + Hearing + Processing Speed	.000	.404	.257
	Age + Hearing + Processing Speed + Crystallized Intelligence	.000	.404	-.028
	Age + Hearing + Processing Speed + Crystallized Intelligence + Working Memory	.000	.404	.008

* $p < .05$. ** $p < .01$.

Next, additional regressions were conducted examining the effects of intelligibility and the cognitive factors on recall performance. The full analyses can be seen in Table 5. With respect to synthetic word recall both synthetic intelligibility and processing speed were significant predictors accounting for 47% of the variance. In the case of natural word recall natural intelligibility and processing speed were significant predictors accounting for 41% of

the variance. The model for synthetic story recall followed a different pattern compared to synthetic word recall with synthetic intelligibility being significant predictors, accounting for 11% of the variance. Finally, the natural story recall model only utilized crystallized intelligence as a predictor, accounting for 23% of the variance.

Table 5

Recall regressions

Regression of Predictor Variables on Performance		ΔR^2	Cumulative R^2	β
Synthetic Word Recall	Synthetic Intelligibility	.358	.358	.599**
	Synthetic Intelligibility + Processing Speed	.114	.472	.407**
	Synthetic Intelligibility + Processing Speed + Working Memory	.019	.491	.144
	Synthetic Intelligibility + Processing Speed+ Working Memory + Crystallized Intelligence	.005	.495	-.071
Natural Word Recall	Natural Intelligibility	.211	.211	.459**
	Natural Intelligibility + Processing Speed	.196	.405	.489**
	Natural Intelligibility + Processing Speed + Working Memory	.041	.446	.215
	Natural Intelligibility + Processing Speed+ Working Memory + Crystallized Intelligence	.001	.447	-.008
Synthetic Story Recall	Synthetic Intelligibility	.116	.116	.340*
	Synthetic Intelligibility + Processing Speed	.006	.122	-.097
	Synthetic Intelligibility + Processing Speed + Working Memory	.044	.166	.223
	Synthetic Intelligibility + Processing Speed+ Working Memory + Crystallized Intelligence	.038	.204	.221
Natural Story Recall	Natural Intelligibility	.005	.005	.070
	Natural Intelligibility + Processing Speed	.074	.089	.321
	Natural Intelligibility + Processing Speed + Working Memory	.003	.092	.064
	Natural Intelligibility + Processing Speed+ Working Memory + Crystallized Intelligence	.146	.238	.433**

* $p < .05$. ** $p < .01$.

Discussion

These results collectively indicate that older adults do have a more difficult time with synthetic speech when limited to disjointed words, as demonstrated by both the intelligibility task and the word recall task. These results agree with theory and observations by Ralston et al. (1991) and Smither (1993), who posited that cognitive workload would increase with the use of synthetic speech. Taken in conjunction with the assertion that listeners with degraded cognitive abilities have more difficulty with synthetic speech (Koul, 2003) and age-related declines in working memory have been observed (Craik, 2000; Salthouse, 1994), it seems plausible that age related degradation in a higher workload task leads to difficulty processing synthetic speech.

However, older adults can use synthetic speech as effectively as younger adults when the speech is used to convey whole meaning, as demonstrated by synthetic story recall. This contextual cueing result is congruent with previous findings (Perry & Wingfield, 1994; Wingfield et al., 1991) particularly research by Drager and Reichle (2001a), who found participants performed better when presented with a context enhancing story. This result also supports the idea of Universal design (Nayak, 1995; Vanderheiden, 1997). By designing to accommodate those with limited abilities, we can maintain or enhance performance in those without limitations.

Focusing on the other aspects of the intelligibility results, main effects of speech type and age were found. Synthetic word intelligibility was lower than natural and older adults had more difficulty with all words than younger adults. The main effect of speech type supports the results of previous research (Greene et al., 1986; Kangas & Allen, 1990; Mirinda & Beukelman, 1987). The main effect of age, a result which has previously been

supported (Drager & Reichle, 2001b; Kangas & Allen, 1990; Roring et al., 2007) and refuted (Drager & Reichle, 2001a; Humes et al., 1991; Humes et al., 1993), is the more poignant. This result seemed clearly linked to processing speed, a measure which was significantly and positively correlated with performance on the intelligibility task. Older adults were slower with respect to processing speed as demonstrated by a t-test, which may partly explain the performance of older adults on the intelligibility task. Finally, it should be noted that the intelligibility rate of synthetic speech was 71 percent for younger adults and 51 percent for older adults, a result which stands in contrast to the 95 percent rate found in Greene et al. (1986).

The memory tasks produced several main effects including task type, speech type and age. The main effect of task type is a consequence of context and follows from literature previously discussed (Nayak, 1995; Vanderheiden, 1997). The main effect of speech type is congruent with other studies and theory on workload with respect to synthetic speech (Sharit et al., 2003; Slowiaczek & Nusbaum, 1985; Smither, 1993). This stands in contrast to work done by Humes et al. (1993), where no main effect of age was found. Humes used monosyllabic words, which may have led to differing results. Finally, the main effect of age in recall stands in agreement with previous theories and research of working memory decline in older adults (Craik, 2000; Salthouse, 1994).

With respect to the correlations and regressions, the most engaging result is that processing speed as measured by the Digit Symbol Substitution task was significantly correlated with many of the measures of intelligibility and recall. Processing speed captured more of the variability than working memory or crystallized intelligence. These correlations were consistent with previous theory regarding age and recall (Craik, 2000; Salthouse, 1994).

In addition, the correlation of age to processing speed was consistent with Salthouse (1991). The correlation of processing speed to synthetic word recall was strong and consistent with Ralston et al. (1991). Synthetic speech word recall did impose high cognitive workload and was sensitive to each individual's processing speed.

More specifically, the results of the many stepwise regressions represent interesting findings. Both age and hearing were worthwhile contributors to scores in intelligibility. Crystallized intelligence was only related to the natural story recall, a result which was surprising but not unreasonable. Natural story recall may have represented the only task with a sufficient connection to semantic knowledge to be related to the Shipley task.

Taking all of these results into account, we believe that processing speed is having an effect on the recall of synthetic and natural speech words because of the increased cognitive load that is imposed when a message is presented without context. Natural recall of words and synthetic recall were all related to their respective intelligibility scores. However, natural story recall was not related to intelligibility, indicating that the effect of context may have been great in natural story recall, greater than the effect of context in synthetic story recall. This difference may have been due to the unfamiliarity of users with synthetic speech, a conclusion that follows from the differences in intelligibility between natural and synthetic speech. Finally, the significance of crystallized intelligence in the natural story recall seems to confirm such a theory.

Given these findings we present the following design recommendations for the use of synthetic speech and its designers.

- When information needs to be conveyed to a listener messages should be presented within context for all listeners and should not be presented in isolation.

- When ideas must be presented in isolation (sans context), designers should be careful to consider the limitations of working memory in older adults. Unrelated words, the number of words, and the rate of speech should be carefully monitored so as to not overload the cognitive abilities of the user.
- When presenting a message, one should allow for a warning to precede the message so that listeners have an opportunity to unload working memory, redirect attention and reduce cognitive load.
- The focus of TTS designers should continue to be prosodic cues, the subtle intonations and cues present in natural speech but not synthetic speech.

Limitations

With respect to the design, the experiment was limited by time as to the number of word lists and stories that could be utilized with each participant. An additional number of such materials may have yielded more accurate results. However, the use of over 250 words and four stories in the procedure makes it unlikely that a larger portion of stimuli would significantly improve the accuracy of the results. A larger pool of stimuli might also introduce a fatigue factor. In addition, the use of a single TTS device (DECtalk) may have biased results. However, DECtalk was chosen for its excellent performance (Greene et al., 1986), so this design represents a best case scenario for performance in TTS situations. A less intelligible TTS system might exaggerate the effects found in the present research.

With respect to the materials themselves, the new stories created for use during the story recall task may have unduly impacted those results. Measures were taken to make any such effects unlikely including the approximate equality in word count and the exact equality in idea units between the original stories and the created stories.

Future Research

Future researchers have many avenues that can be explored including mediation models, induced cognitive workload designs, speech rate, and the effects of background noise. More niche interests, such as domain specific application could also be investigated.

First, researchers might investigate the extent that processing speed may act as a mediator for age with respect to synthetic speech intelligibility and recall. To accomplish this, multiple measures of processing speed could be utilized to capture the construct. Other mediators might include a workload measure or working memory.

Second, designs that impose cognitive workload on participants might create greater differences in recall for older and younger adults. Particularly, these designs could impose additional workload by dividing attention, increasing speech rate, or creating distractors. Speech rate in particular is a design that should be explored further given the results of Roring et al. (2007), who found reduced speech rate hampered synthetic intelligibility. This unintuitive result may be due to the lack of natural prosody inherent in synthetic speech, making monosyllabic words awkward to listen to.

Third, designs that utilize background noise might create a more ecologically valid environment for experimentation. The effects of background noise are a common problem when audition is concerned in everyday life. By removing any background sound in the current design for the sake of removing confounds, a design aimed at the effects of background sound might yield interesting results and lend ecological validity.

Finally, a party interested in the domain specific effects of context could investigate those particulars. The domains of transportation information or synthetic speech via the phone are the areas that seem most ripe for such specific investigations.

By creating cognitive support for older adults via synthetic speech we can potentially improve the lives of older adults and younger adults alike. These design guidelines are ones that both designers and system users can follow to improve the related gains of synthetic speech. Human Factors researchers can continue to improve the usability of these systems.

References

- Allen, J., Hunnicut, S., & Klatt, D. (1987). *From text to speech, the MITTALK system*. Cambridge, MA: Cambridge University Press.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8). New York, NY: Academic Press.
- Bowles, N. (1993). Semantic processes that serve picture naming. In J. Cerella, J. Rybash, W. Hoyer, & M. L. Commons (Eds.), *Adult information processing: Limits on loss* (pp. 205-228). San Diego, CA: Academic Press.
- Caplan, D., & Waters, G. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, 22, 114-126.
- Craik, F. I. M. (2000). Age-related changes in human memory. In D. C. Park & N. Schwartz (Eds.), *Cognitive aging: A primer*. Philadelphia, PA: Taylor and Francis.
- Craik, F. I. M. (1994). Memory changes in normal aging. *Current Directions in Psychological Science*, 3, 155-158.
- Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagenforf (Eds.), *Human memory and cognitive capabilities: Mechanisms and performances*. Amsterdam: Elsevier.
- Craik, F. I. M. & Salthouse, T. A. (Eds.) (2000). *The handbook of aging and cognition* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Czaja, S. J., Guerrier, J. H., Nair, S. N., & Landauer, T. K. (1993). Computer communication as an aid to independence for older adults. *Behaviour and Information Technology*, 12, 197-207.

- Drager, K. D. R. & Riechle, J. E. (2001a). Effects of discourse context on the intelligibility of synthesized speech for young adult and older adult listeners: Application for AAC. *Journal of Speech, Language, and Hearing Research, 44*, 1052-1057.
- Drager, K. D. R. & Riechle, J. E. (2001b). Effects of age and divided attention on listeners' comprehension of synthesized speech. *Augmentative and Alternative Communications, 17*, 109-119.
- Greene, B. G., Manous, L. M., & Pisoni, D. B. (1986). Perceptual evaluation of DECtalk: A final report on version 1.8. (Research on speech perception, Progress Rep. No. 10). Bloomington, IN: Speech Research Laboratory, Psychology Department, Indiana University.
- Hahm, W., & Bikson, T. (1989). Retirees using email and networked computers. *International Journal of Technology and Aging, 2*, 113-123.
- Hancock, H. E., Fisk, A. D., & Rogers, W. A. (2001). Everyday products: Easy to use... or not? *Ergonomics in Design, 9*, 12-18.
- Hartley, J. (1994). *Designing instructional text*. London: Kogan Page Limited.
- Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.) *The psychology of learning and motivation* (Vol. 2). San Diego, CA: Academic Press.
- Humes, L. E., Nelson, K. J., & Pisoni, D. B. (1991). Recognition of synthetic speech by hearing-impaired elderly listeners. *Journal of Speech and Hearing Research, 34*(5), 1180-1184.

- Humes, L. E., Nelson, K. J., Pisoni, D. B., & Lively, S. E. (1993). Effects of age on serial recall of natural and synthetic speech. *Journal of Speech and Hearing Research, 36*, 634-639.
- Kangas, K. A., & Allen, G. D. (1990). Intelligibility of synthetic speech for normal-hearing and hearing-impaired listeners. *Journal of Speech and Hearing Disorders, 55*(4), 751-755.
- Kemper, S., & Mitzner, T. L. (2001). Language production and comprehension. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (5th ed.) (pp. 373-398). San Diego, CA: Academic Press.
- Kline, D. W., & Scialfa, C. T. (1997). Sensory and perceptual functioning: Basic research and human factors implications. In A. D. Fisk & W. A. Rogers (Eds.), *Handbook of human factors and the older adult* (pp. 125-149). Sand Diego, CA: Academic Press.
- Koul, R. (2003). Synthetic speech perception in individuals with and without disabilities. *Augmentive and Alternative Communication, 19*(1), 49-58.
- Light, L. L. (1992). The organization of memory in old age. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 111-166). Hillsdale, NJ: Erlbaum.
- Mayhorn, C. B., Rogers, W. A., & Fisk, A. D. (2004). Designing technology based on cognitive aging principles. In D. C. Burdick & S. Kwon (Eds.), *Gerotechnology: Research and practice in technology and aging* (pp. 18-41). New York, NY: Springer Publishing.
- Micro Audiometrics Corp. Earscan Audiometer. Murphy, NC

- Mirenda, P., & Beukelman, D. R. (1987). A comparison of speech synthesis intelligibility with listeners from three age groups. *Augmentative and Alternative Communication*, 3, 120-128.
- Morrell, R. W., & Echt, K. V. (1997). Designing written instructions for older adults: Learning to use computers. In A. D. Fisk & W. A. Rogers (Eds), *Handbook of human factors and the older adult* (125-149). Sand Diego, CA: Academic Press.
- Morrell, R. W., Mayhorn, C. B., & Bennett, J. (2000). A survey of world wide web use in middle-aged and older adults. *Human Factors*, 42(2), 175-182.
- Nayak, U. S. L. (1995). Elders-led design. *Ergonomics in Design*, 3(1), 8-13.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.
- Paris, C. R., Gilson, R. D., & Thomas, M. H. (1995). Effect of synthetic voice intelligibility on speech comprehension. *Human Factors*, 37(2), 335-340.
- Park, D. C. & Schwarz, N. (Eds.). (2000). *Cognitive aging: A primer*. Philadelphia, PA: Psychology Press.
- Perry, A. R., & Wingfield, A. (1994). Contextual encoding by young and elderly adults as revealed cued and free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 1185-1191.
- Ralston, J. V., Pisoni, D. B., Lively, S. E., Greene, B. G., & Mullennix, J. W. (1991). Comprehension of synthetic speech produced by rule: Word monitoring and sentence-by-sentence listening times. *Human Factors*, 33, 471-479.

- Rogers, W. A., Mayhorn, C. B., & Fisk, A. D. (2004). Technology in everyday life for older adults. In D. C. Burdick & S. Kwon (Eds.), *Gerotechnology: Research and practice in technology and aging* (pp. 18-41). New York, NY: Springer Publishing.
- Roring, R. W., Hines, F. G., & Charness, N. (2007). Age differences in identifying words in synthetic speech. *Human Factors*, 49(1), 25-31
- Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Salthouse, T. A. (1994). The aging of working memory. *Neuropsychology*, 8(4), 535-543.
- Schmidt M. 1996. Rey Auditory Verbal Learning Test--A Handbook. Los Angeles, CA: Western Psychological Services.
- Scialfa, C. T., Ho, G., & Laberge, J. (2004). Perceptual aspects of gerotechnology. In D. C. Burdick & S. Kwon (Eds.), *Gerotechnology: Research and practice in technology and aging* (pp. 18-41). New York, NY: Springer Publishing.
- Sharit, J., Czaja, S. J., Nair, S., & Lee, C. C. (2003). Effects of age, speech rate, and environmental support in using telephone menu systems. *Human Factors*, 45(2), 234-251.
- Shipley, W. C. (1986). *Shipley Institute of Living Scale*. Los Angeles: Western Psychological Services.
- Slowiaczek, L. M., & Nusbaum, El. C. (1985). Effects of speech rate and pitch contour on the perception of synthetic speech. *Human Factors*, 27, 701-712.
- Smith, A. D. (2002). Consideration of memory functioning in health care intervention with older adults. In W. A. Rogers & A. D. Fisk (Eds.), *Human factors interventions for the health care of older adults*. Mahwah, NJ: Lawrence Erlbaum Associates.

- Smither, J. A. (1993). Short term memory demands in processing synthetic speech by old and young adults. *Behaviour and Information Technology*, 12(6), 330-335.
- Sweller, J. (1993). Some cognitive processes and their consequences for the organization and presentation of information. *Australian Journal of Psychology*, 45, 1-8.
- Toglia, M. P., & Battig, W. F., (1978). *Handbook of Semantic Word Norms*. Oxford, England: Lawrence Erlbaum.
- U.S. Census Bureau (2004). Statistics. Available at www.census.gov. Retrieved March, 2005.
- Vanderheiden, G. C. (1997). Design for people with functional limitations due to disability, aging, or circumstances. In G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics* (pp. 2010-2052). New York: John Wiley & Sons, 1997.
- Wechsler, D. (1997). *Wechsler Memory Scale III*. (3rd Ed.). San Antonio, TX: The Psychological Corporation.
- Wickens, C. D., Lee, J. D., Liu, Y., & Becker, S. G. (2004). *An introduction to human factors engineering* (2nd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
- Wingfield, A., Aberdeen, J. S., & Stine, E. A. L. (1991). Word onset gating and linguistic context in spoken word recognition by young and elderly adults. *Journal of Gerontology: Psychological Sciences*, 46, 127-129.

APPENDICES

Appendix A

Word Lists for Intelligibility Task

List A, Synthetic Speech: 50 words

ring	tea	lawyer
shutter	mouse	rose
sock	stream	cloud
barn	towel	mail
ranger	shelf	stranger
mouth	bloom	garage
weather	shade	glasses
office	desk	stocking
hand	gun	shoe
pencil	crayon	teacher
home	church	stove
fish	eagle	nest
lunar	fountain	children
trunk	boat	tuba
balloon	hot	toffee
bird	boss	lamb
mountain	water	

List B, Natural Speech: 50 words

plant	peach	ocean
fireplace	log	palace
dawn	lane	lip
sprout	sheep	bar
shower	opal	dress
bottle	metal	steam
crib	brook	coin
chair	seed	rock
slug	county	army
malt	give	building
judge	grant	friend
plane	insect	storm
silver	wool	village
pool	zipper	cell
thumb	bold	ladder
bowl	meal	shield
lock	baby	

Appendix B

Word Lists for Word Recall Task

Synthetic Word Lists

List A

drum
curtain
bell
coffee
school
parent
moon
garden
hat
farmer
nose
turkey
color
house
river

List C

street
grass
door
arm
star
wife
window
city
pupil
cabin
lake
pipe
skin
fire
clock

List E

violin
tree
scarf
ham
suitcase
cousin
earth
stairs
dog
banana
town
radio
hunter
bucket
field

List B

book
flower
train
rug
meadow
harp
salt
finger
apple
chimney
button
key
dog
glass
rattle

List D

tower
wheat
queen
sugar
home
boy
doctor
camp
flag
letter
corn
nail
cattle
shore
body

List F

sky
dollar
valley
butter
hall
diamond
winter
mother
christmas
meat
forest
gold
plant
money
hotel

Natural Word Lists

List G

horn
screen
chime
soda
library
nanny
planet
yard
bracelet
singer
jaw
vulture
shape
shop
tide

List H

novel
tulip
bike
mat
lawn
flute
mint
neck
grape
closet
chain
code
roach
steel
blocks

List J

path
bush
gate
wing
hero
bride
porch
port
novice
shack
spring
vent
blood
wind
counter

List K

fort
potato
knight
candy
farm
lad
driver
tent
signal
report
crop
spike
bear
cliff
frame

List L

guitar
shrub
boots
bacon
sack
friend
dirt
sink
tiger
lemon
settlement
lamp
butcher
mop
plot

List M

breeze
wage
gulf
salad
lobby
gem
period
woman
spell
meeting
island
bronze
vine
treasure
guest

Appendix C

Logical Memory Stories from the Wechsler Memory Scale

“//” delineates story idea units

Story A presented in synthetic speech

Anna // Thompson // of south // Boston, // employed // as a cook // in a school // cafeteria, // reported // at the police // station // that she had been held up // on State Street // the night before // and robbed // of fifty-six dollars. // She had four // small children, // the rent was due, // and they had not eaten // for two days. //

The police, // touched by the woman's story, // took up a collection // for her.

Story B presented in natural speech

At 6:00 // on Monday // evening, // Joe // Garcia // of San Francisco // was watching television // as he dressed // to go out. // A weather bulletin // interrupted the program // to warn that thunderstorms // would move into the area // within the next 2 to 3 hours // and remain until morning. // The announcer // said the storm could bring hail // and up to 4 inches // of rain // and cause the temperature to drop by // 15 degrees.

Joe decided to stay home. // He took off his coat // and sat down // to watch old movies.

Story A: 65 Words, 25 idea units

Story B: 86 Words, 25 idea units

Appendix D

Logical Memory Stories Created for the Present Experiment

“//” delineates story idea units

Story C presented in Natural Speech using Story A as a model

Tami // Foster // of east // Detroit, //working // as a receptionist // at a computer // factory, // reported // at the payroll department // office // that she had lost // her paycheck // of 300 dollars // at home // earlier that day. // She had three // bills to pay, // a house to clean, // and needed to get home // before 6. //

The department, // concerned for their employee, // wrote a new paycheck // for her.

Story D presented in Synthetic Speech using Story B as a model

At 8:00 // on Wednesday // morning, // Fred // Miller // of Baltimore // was listening to the radio // as he took a shower // before work. // A radio announcement // about traffic conditions // cautioned drivers // to avoid the interstate // for the next hour // for their own safety. // The traffic reporter said // that a truck blocked the road, // 4 people // were injured, // and traffic had slowed // to a stand still. //

Fred thought of another way to work. // He finished dressing // and grabbed his keys // to lock the door. //

Story C: 63 Words, 25 idea units

Story D: 81 Words, 25 idea units

Appendix E

Informed Consent Form for Young (presented first) and Older adults

**North Carolina State University
INFORMED CONSENT FORM**

Reexamining Synthetic Speech

Principal Investigator: J. B. Hardee

You are invited to participate in a research study. The purpose of this study is to learn more about the ways in which people understand and remember synthetic speech. We are particularly interested in the circumstances that can cause someone to experience more difficulty when listening to synthetic speech. This information is very important because it can be used to inform designers and acoustic professionals of the needs of the people who use their products.

INFORMATION

In this study, you will be asked to read a consent form, complete several cognitive tests, complete a few tasks related to synthetic speech comprehension and recall, and read a debriefing statement. It should take approximately 1 hour for you to complete this study. If you have questions at any time during the study, please ask the experimenter.

RISKS

There are no foreseeable risks or discomforts associated with this study.

BENEFITS

Participation in this research as part of the Psychology 200 research requirement should aid your understanding of how research is conducted in the University environment. Because you are providing important information concerning your product usage and experience, other members of society may benefit through improvements in future design of new technologies

CONFIDENTIALITY

The information in the study records will be kept strictly confidential. Data will be stored securely and will be made available only to persons conducting the study unless you specifically give permission in writing to do otherwise. No reference will be made in oral or written reports which could link you to the study.

COMPENSATION

For participating in this study you will receive 2 course credits through the Experimetrix website. Other ways to earn the same amount of credit are described in your course syllabus. Specifically, psychology 200 students may choose to write a paper instead of participating in research. If you withdraw from the study prior to its completion, you will receive 1 course credit through Experimetrix.

CONTACT

If you have questions at any time about the study or the procedures, you may contact the faculty researcher, Christopher B. Mayhorn, at 728 Poe Hall, or 919-513-4856. If you feel you have not been treated according to the descriptions in this form, or your rights as a participant in research have been violated during the course of this project, you may contact Dr. Matthew Zingraff, Chair of the NCSU IRB for the Use of Human Subjects in Research Committee, Box 7514, NCSU Campus (919/513-1834) or Mr. Matthew Ronning, Assistant Vice Chancellor, Research Administration, Box 7514, NCSU Campus (919/513-2148)

PARTICIPATION

Your participation in this study is voluntary; you may decline to participate without penalty. If you decide to participate, you may withdraw from the study at any time without penalty and without loss of benefits to which you are otherwise entitled. If you withdraw from the study before data collection is completed your data will be returned to you or destroyed.

CONSENT

I have read and understand the above information. I have received a copy of this form. I agree to participate in this study.

Subject's signature _____

Date _____

Investigator's signature _____

Date _____

**North Carolina State University
INFORMED CONSENT FORM**

Reexamining Synthetic Speech

Principal Investigator: J. B. Hardee

You are invited to participate in a research study. The purpose of this study is to learn more about the ways in which people understand and remember synthetic speech. We are particularly interested in the circumstances that can cause someone to experience more difficulty when listening to synthetic speech. This information is very important because it can be used to inform designers and acoustic professionals of the needs of the people who use their products.

INFORMATION

In this study, you will be asked to read a consent form, complete several cognitive tests, complete a few tasks related to synthetic speech comprehension and recall, and read a debriefing statement. It should take approximately 1 hour for you to complete this study. If you have questions at any time during the study, please ask the experimenter.

RISKS

There are no foreseeable risks or discomforts associated with this study.

BENEFITS

Because you are providing important information concerning your product usage and experience, other members of society may benefit through improvements in future design of new technologies

CONFIDENTIALITY

The information in the study records will be kept strictly confidential. Data will be stored securely and will be made available only to persons conducting the study unless you specifically give permission in writing to do otherwise. No reference will be made in oral or written reports which could link you to the study.

COMPENSATION

For participating in this study you will receive 10 (ten) dollars per hour of participation. If you withdraw from the study prior to its completion, you will be compensated for the time that you spend participating in the study.

CONTACT

If you have questions at any time about the study or the procedures, you may contact the faculty researcher, Christopher B. Mayhorn, at 728 Poe Hall, or 919-513-4856. If you feel you have not been treated according to the descriptions in this form, or your rights as a participant in research have been violated during the course of this project, you may contact Dr. Matthew Zingraff, Chair of the NCSU IRB for the Use of Human Subjects in Research Committee, Box 7514, NCSU Campus (919/513-1834) or Mr. Matthew Ronning, Assistant Vice Chancellor, Research Administration, Box 7514, NCSU Campus (919/513-2148)

PARTICIPATION

Your participation in this study is voluntary; you may decline to participate without penalty. If you decide to participate, you may withdraw from the study at any time without penalty and without loss of benefits to which you are otherwise entitled. If you withdraw from the study before data collection is completed your data will be returned to you or destroyed.

CONSENT

I have read and understand the above information. I have received a copy of this form. I agree to participate in this study.

Subject's signature _____ Date _____

Investigator's signature _____

Date _____

Appendix F

Protocol and Instructions

Introduction

Hello! My name is XXXX (enter experimenter name). The experiment in which you are participating is designed to study the comprehension and recall of synthetic speech. You will complete a demographics questionnaire, several measures of cognitive ability, and several tasks with synthetic speech. You are free to leave at any time if you feel uncomfortable with participating in this study.

Informed Consent Form

Please read this informed Consent Form. It will tell you about this experiment and your rights and responsibilities as a participant. If you decide to participate in this study, please sign at the bottom of both forms.

Information

Before we begin, I need to get some information from you. Please take a few moments now to complete this questionnaire (hand out the demographic survey).

Assessments

Before we begin the actual experiment, I need you to participate in a few small tasks. This first task is a test of your hearing ability (use audiometer to assess hearing, and read relevant instructions). This next task is a vocabulary inventory (pass out task and read respective instructions). In this next activity you will be matching digits to symbols (pass out task and read respective instructions). In this final activity you will be hearing and repeating numbers (pass out digit-span task, forward then backward and read respective instructions). Now we can review how the main experiment will be conducted.

Overview

In this experiment you will complete three tasks. In the first task you will be presented with lists of words. Your job will be to say the word that you hear as soon as you hear it. After this task, you will complete another task that involves lists of words. Your job in this second task will be to listen to the entire list and then recall of all of the words that you can. Finally, your last task will be to listen to a few short stories. After you have heard them, I will ask you to recall as much about the story as you can. You will be listening to all the words and the stories from these computer speakers.

Procedure

Now here is how we are going to proceed. In this first task you will be repeating words that are played as soon as you hear them. When the experiment begins, I will say "Let's begin." At that time a series of words will begin to play. As soon as word is played you should tell me what it is. I will record what you say. I will tell you when we are done with the first set. I will give you a moment or two and then we will proceed to the subsequent set. After all the words have been done I will give you a brief break.

Do you have any questions so far?
Please make sure you are comfortable in your chair. OK. Do you have any questions at all?

Ok. Let's begin... (Participant completes Intelligibility task)

Now we are going to move on to our second task (Word and discourse tasks will be counterbalanced. The word task is presented second here for simplicity). When the experiment begins, I will say "Let's begin." At that time a list of words will begin to play. You should listen carefully to each word, but do not say the words aloud. If you say the words aloud you may miss subsequent words being played. After the entire list has been played, I will say "Recall the words." After I give you this signal, you should tell me all of the words that you heard in the most recent list that you can remember. I will record the words that you tell me. We will do this several times for several lists. Please remember that you should only repeat words after the entire list has been presented and only words that were on the most recent list. After all of the lists have been played I will give you another brief break.

Do you have any questions?

Ok. Let's begin... (Participant completes Word Recall)

Now we are going to move on to our final task. When the experiment begins, I will say "Let's begin." At that time a short story will begin to play. You should listen carefully to each part of the story, but do not say any of the story aloud. After the entire story has been played, I will say "Recall the story." After I give you this signal, you should tell me everything about the most recent story that you can remember. If you can remember the exact words that is fine, but if you can not then try to tell me whatever you can remember. I will record the details of the story that you tell me. We will do this several times with several stories. Please remember to wait for my signal before you tell me about the story. After all of the stories have been played, then you will be finished.

Do you have any questions?

Ok. Let's begin... (Participant completes Discourse Recall)

Wrap-up:

Thank you very much for your time. Please take a moment to look over this debriefing form. It contains more information about this study and contact information for the researchers involved. Do you have any other questions or comments? Thank you again!

Appendix G
Debriefing Form

Debriefing

Thank you for your participation in our survey. We could not conduct our research without your help.

In this study, you completed comprehension and recall tasks designed to measure how well people understand synthetic speech words and discourse. From the information you provided, we hope to better understand the difficulties that are caused by noise, fast speech rates, or other adverse conditions and how they can affect the ability of people to understand synthetic speech. By drawing on your experiences with synthetic speech, design suggestions can be made for future systems. The ultimate goal of this research is to improve synthetic speech by informing the next generations of devices and professionals.

We thank you again for participating; if you have any questions or concerns, please contact us.

Dr. Chris Mayhorn
NCSU Ergonomics and Aging Lab
919-513-4856