The Effects of Speech Compression on Recall in a Multimedia Environment

Marcelite E. Dingle Johnson

Dissertation submitted to the Faculty of
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Curriculum and Instruction

Glen A. Holmes, Chair
John K. Burton
Jimmie C. Fortune
Barbara B. Lockee
Joyce Williams-Green

April 29, 1998
Blacksburg, Virginia

Keywords: Audio, Compressed Speech, Multimedia
Copyright 1998, Marcelite D. Johnson
The Effects of Speech Compression on Recall in a Multimedia Environment

Marcelite E. Dingle Johnson

(ABSTRACT)

Typically, instructional designers introduce audio in multimedia environments when a) it appears to be necessary -- for example, to provide feedback; or b) accessibility, availability, and/or hardware issues are not important -- in other words, It's there -- -- It sounds good -- It should work -- Why not use it? However, rarely is a decision to use audio based upon a thorough understanding of why, from a learning perspective, it is appropriate or optimal to do so. Clearly a lack of such an understanding can, and often does, lead to the inappropriate use of audio as a useful instructional medium.

This study is designed: (1) to investigate the educational value of compressed speech in a multimedia environment, and (2) to evaluate how rate of speech may influence learning in a multimedia setting. It is also concerned with ascertaining whether: a) the recall of information at various levels of compressed speech decreases when audio delivery-rate (words-per-minute) increases; b) there is an interaction between task completion time and recall of information at various levels of compression; and c) relative to recall, there is an interaction between audio delivery-rate and increasing exposure to an audio stimulus.

One hundred and ninety-two undergraduate students enrolled in business courses in two southeastern regional universities located in North Carolina and Virginia were randomly assigned to one of three groups. The three experimental test groups were normal speech, 125 words per minute, compressed speech: 175 words per minute and 200 words per minute.

In all groups, participants were asked to listen to the solution of the puzzle, given in one of the three presentation rates of speech and they were given three different puzzles to solve. All data were collected at each assigned computer workstation. Data analysis revealed a difference between the three treatment groups (125 wpm, 175 wpm, 200 wpm) in recall scores due to the rate of compression, no interaction between completion times and the recall of information based on compression rate, and a difference in recall scores between the three treatment groups and the amount of exposure to the audio stimulus.
Dedication

To the loving memory of my parents, Jane and Leon Dingle, who were always loving, supportive and stressed the importance of education, I dedicate my dissertation to you.

To my wonderful husband and the most important person in my life, Melvin, I dedicate my degree to you. Your love, constant support, and encouragement made me believe that I could accomplish anything. Your tremendous love for learning and exceptional professional accomplishments have forced me to excel academically as well as professionally. It meant so much for the two of us to take this journey together. You are the BEST and I love you.

To my children: my daughter DeAndra who exercised great patience, lots of love and understanding while earning this degree; my daughter Monet who added many wonderful events to my life like her wedding and my soon-to-be grandchild; my son M. Roschaun whose competitive spirit kept me working hard, since we both were struggling in graduate school together, his love and technical expertise; and my son-in-law Leotis who provided continual encouragement, I thank each of you.

Finally, I thank the rest of my family for their continual love. Special recognition goes to my aunt and uncle, Virginia and Edward, whose weekly telephone calls provided me with love, kind and encouraging words. Additional consideration goes to my mother-in-law and father-in-law, Vernedia and William who continually provided their love and support. I am truly blessed with the greatest family ever.
Acknowledgments

Many people have contributed to the completion of this document and I express my deepest appreciation to them all. However, I would like to recognize special people.

Glen Holmes, advisor and chair of my committee, guided me through this process and spent countless hours working with my model. Exposure to his technical expertise throughout my graduate years at Virginia Tech became the inspiration for further research. Thank you for the supportive hand. I would like to thank the other members of my committee: John K. Burton who always had encouraging words and was very supportive; Barbara Lockee who gave good academic advice and support about getting the work completed; Jim Fortune who offered research expertise which was so crucial to completing this dissertation; and Joyce Williams-Green whose critical eye was so important to completing this dissertation and her supportive and kind words. I owe special thanks to Terry Davis, my friend and supporter who made sure all appointments were made. She made life enjoyable during my stay at Virginia Tech in the Education Technology Lab. Special thanks to Julia for her friendship and editing expertise and Carolyn for the camaraderie and commiserating.

I will be forever grateful for the assistance, guidance, and support I have received from Dean Quiester Craig of the School of Business and Economics and Vice Chancellor Harold Martin at North Carolina A&T State University. I also want to thank the professors at A&T for their willingness to help with my experiment: Claude Barnes, Edna Johnson, Olenda Johnson, Alonzo Redmon, Janice Smith, and Sharon White. This dissertation would not be possible without the students who participated in this study.

Three persons deserve special recognition. Juone, “JB” Brown, who offered her assistance with my experiment (her professional voice), friendship, and encouragement. Paula Jeffries who always provided meaningful advice and kept me sane during numerous difficult days. Thanks to Glenda Scales for opening her heart, mind and home to me and for those fabulous dinners.

Finally, to my friend Joe Stetina owner of Starbucks Coffee Shop in Greensboro, thanks for the words of encouragement and tremendous “care package” enjoyed by my committee and I.
Table of Contents

Abstract ........................................................................................................................................... ii
Dedication ......................................................................................................................................... iii
Acknowledgments ........................................................................................................................... iv
Table of Contents ............................................................................................................................. v
List of Tables ..................................................................................................................................... vii
List of Figures .................................................................................................................................... viii

Chapter I .......................................................................................................................................... 1
  Introduction ............................................................................................................................... 1
  Rationale for the Study ............................................................................................................... 1
  Purpose of the Study .................................................................................................................. 2
  Statement of the Problem ......................................................................................................... 2
  Definition of Terms .................................................................................................................. 2

Chapter II ......................................................................................................................................... 5
  Review of Literature .................................................................................................................. 5
  Digital Audio in Multimedia Production ................................................................................... 8
  Foundations of Multimedia ....................................................................................................... 11
  Compressed Speech ................................................................................................................ 16

Chapter III ....................................................................................................................................... 19
  Methodology ............................................................................................................................ 19
  Research Hypotheses ............................................................................................................... 19
  Subjects ..................................................................................................................................... 19
  Experimental Design .............................................................................................................. 19
  Method of Analysis ................................................................................................................ 22

Chapter IV ...................................................................................................................................... 23
  Results ...................................................................................................................................... 23
  Discussion ............................................................................................................................... 35
  Impact of Study ....................................................................................................................... 36
  Summary ................................................................................................................................... 36
  Suggestions for Future Research .............................................................................................. 37

Reflections on Conducting Research ............................................................................................ 39

References ....................................................................................................................................... 40
  Appendix A: Pilot Study ........................................................................................................... 45
  Appendix B: Screens Used in the Module ................................................................................. 46
  Appendix C: Script for Study .................................................................................................... 60
List of Tables

Table 1  Frequency and Dynamic Ranges .......................................................... 7
Table 2  Recording Options ............................................................................. 8
Table 3  Type of Microphones ........................................................................ 10
Table 4  Statistical Values (Attempts): Descriptive ......................................... 24
Table 5  Repeated Measures ANOVA for Attempts .......................................... 25
Table 6  Tukey HSD Multiple Comparisons of Treatment Means ..................... 26
Table 7  Statistical Values (Time): Descriptive ............................................... 27
Table 8  Repeated Measures ANOVA for Time ............................................... 28
Table 9  Statistical Values (Correct Attempts): Descriptive ............................. 29
Table 10 Repeated Measures ANOVA for Correct Number of Continuous Attempts ..... 30
Table 11 Tukey HSD Multiple Comparisons of Treatment Means – Correct Attempts 31
Table F 1 Ethnicity ......................................................................................... 63
Table F 2 Gender ............................................................................................ 64
Table F 3 GPA ................................................................................................. 65
Table F 4 Age ................................................................................................. 66
Table F 5 Play Computer Games ..................................................................... 67
Table F 6 Academic Status ............................................................................. 68
Table F 7 Treatment Groups .......................................................................... 69
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Means – Number of Attempts per Treatment Group per Trial</td>
<td>32</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Means – Total Completion Time per Treatment Group per Trial</td>
<td>33</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Means – Correct Continuous Attempts per Treatment Group per Trial</td>
<td>34</td>
</tr>
<tr>
<td>Figure B 1</td>
<td>Introductory Screen</td>
<td>46</td>
</tr>
<tr>
<td>Figure B 2</td>
<td>Demographic Screen</td>
<td>47</td>
</tr>
<tr>
<td>Figure B 3</td>
<td>Insert Diskette Screen</td>
<td>48</td>
</tr>
<tr>
<td>Figure B 4</td>
<td>Phonetic Alphabet Screen</td>
<td>49</td>
</tr>
<tr>
<td>Figure B 5</td>
<td>Instructions &amp; Practice Puzzle Screen</td>
<td>50</td>
</tr>
<tr>
<td>Figure B 6</td>
<td>Beginning Puzzle Screen</td>
<td>51</td>
</tr>
<tr>
<td>Figure B 7</td>
<td>Hear the Solution Screen</td>
<td>52</td>
</tr>
<tr>
<td>Figure B 8</td>
<td>Solve the Puzzle Screen</td>
<td>53</td>
</tr>
<tr>
<td>Figure B 9</td>
<td>Correct Response Screen</td>
<td>54</td>
</tr>
<tr>
<td>Figure B 10</td>
<td>Incorrect Response Screen</td>
<td>55</td>
</tr>
<tr>
<td>Figure B 11</td>
<td>End of Puzzle 1 Screen</td>
<td>56</td>
</tr>
<tr>
<td>Figure B 12</td>
<td>End of Puzzle 2 Screen</td>
<td>57</td>
</tr>
<tr>
<td>Figure B 13</td>
<td>Thank You Screen</td>
<td>58</td>
</tr>
<tr>
<td>Figure B 14</td>
<td>Diskette label</td>
<td>59</td>
</tr>
</tbody>
</table>
Chapter I

Introduction

Background

‘Fast or slow’ … on the verge of destroying the string and the package, ‘words mean the same thing. Meaning is independent of speed.’ ‘It’s just that you don’t read the words---you swallow them. You have to savor words. You have to let them melt in your mouth.’

(Antonio Skarmeta, The Postman)

Several computer-based learning tools have emerged during the past decade. Examples include: computer-based instruction (CBI), computer-assisted instruction (CAI), computer-managed instruction (CMI), computer-based teaching (CBT), computer-aided learning (CAL), CD-ROMs, and most recently, web-based instruction (WBI). These tools have been often cited by instructional designers for their ability to facilitate learning. These multimedia environments allow designers and educators to incorporate text, video, animation, graphics and audio to create a rich learning environment (Kozma, 1991).

Instructional systems designers are inclined to use audio, especially speech, in their products, however, they face a variety of problems. Very little is known about the use of audio concerning processing aural information, events of instruction, and digital audio characteristics. Traditional research in audio is usually conducted within the audiovisual realm. Consequently, there is a scarcity of research on the use of audio in a multimedia environment. This trend is enigmatic considering that the auditory sense has proven to be an important element in the learning process (Aarntzen, 1993).

Rationale for the Study

Little empirical research has been done to indicate whether or not educational gains can be achieved by incorporating speech into computer software (Sales & Johnston, 1993). While there has been much research conducted on the usage of various sounds as educating tools, including music and speech, few guidelines exist to indicate how audio can be optimized in a computer-learning environment (Aarntzen, 1993). Therefore, not much is known about the use of audio, especially speech, in courseware. Designers have used audio, but only when necessary, or when it is easy to implement. For example, sound effects are often used to gain attention (Sales & Johnston, 1993). Audio is also a common feature in language and speech training.

Instructional designers frequently complain about the lack of guidelines for incorporating digital audio in computer-based instruction. Additionally, there is controversy surrounding the instructional payoffs relative to the associated time, effort, and expense involved in adding digital audio to computer-based instruction (Aarntzen, 1993).

Instructional designers still must decide if any form of audio should replace, enhance or use redundancy on a computer screen (Aarntzen, 1993). Additionally, Aarntzen (1993) states
that redundancy factors and multichannel transmission of communication can determine the effectiveness of audio CBI on students’ achievement and help provide guidelines for the instructional designer.

Audio is an important means of communicating and it acts as a source of information transfer. Communicators and instructional materials developers make countless decisions about audio messages each day (Jaspers, 1994). Thus, it is important to investigate the research conducted on audio in a traditional audio-visual format and in current multimedia environments because voice narration is a component of audio (Jaspers, 1994).

As technology filters downward, high quality audio computer cards and computers with built-in audio capabilities have allowed educators and trainers to realize the potential of using audio for CBI and other multimedia applications (Rehaag & Szabo, 1994).

Purpose of the Study

The purposes of this study are: (1) to investigate if there is educational value of using compressed speech in a multimedia environment, (2) to conduct an empirical study of how the rate of speech, words per minute (wpm), may influence learning in a multimedia setting and (3) to determine if it is practical to incorporate compressed speech into instruction.

This study will foster an awareness of the use of audio, particularly speech, in a multimedia environment and its effect on learning. The author hopes to create one particular section of the new audio design model by investigating the relevant literature and conducting relevant experimentation of audio characteristics (including digital audio production) and the use of speech, multiple channel communication, redundancy and compressed speech. This study specifically looks at how the rate of speech (words per minute) may influence learning in a multimedia setting.

Statement of the Problem

This study focuses on the educational value of compressed speech in a multimedia environment. Furthermore, this investigation will establish whether or not the use of audio in a multimedia environment substantially affects achievement as measured by recall scores of college-level students.

Definition of Terms

The following definitions are included to provide a common frame of reference for the terms used in this investigation.

Amplitude.

The strength or intensity of a sound or audio signal.

Analog.

In audio terms, a type of electrical signal whose frequency and amplitude vary continuously in direct relation to the original acoustic sound wave that it represents.
Analog-to-digital converter.  
A circuit that changes the continuously-fluctuating voltages of an analog audio signal into a series of numbers. Abbreviated as A/D converter or ADC.

DAT.  
Acronym for digital audio tape; a recording medium that uses a small cassette – much like a miniature video cassette – to store sounds in a digital format.

Decibel (dB).  
A unit measuring a signal’s strength and is useful when comparing the loudness of two sounds.

Digital.  
In audio terms, using numerical values in discreet steps to represent the continuously varying voltages of an audio signal.

Digital-to-analog converter.  
A circuit that changes digital data into a continuously fluctuating voltage. Commonly abbreviated as D/A converter or DAC.

Dynamic range.  
The difference between the loudest and softest sounds that a device can accurately record and/or reproduce.

Equalization (EQ).  
The manipulation of tone by increasing or decreasing frequency ranges through tone controls, filters, or equalizers.

Frequency.  
Scientific measurement of the number of vibrations per second; expressed in hertz (hz).

Hertz (hz).  
The frequency of sound measured in units. 1Hz equals 1 cycle per second. 1kHz equals 1,000 cycles per second.

Redundant audio.  
Audio that repeats verbally the text on the screen.

Sample.  
A digitally recorded sound.

Sampling Rate.  
In a digital audio device, the rate at which the incoming audio signal is examined to produce numbers representing its instantaneous amplitude levels. The higher the sampling rate, the higher the sound quality.
Signal.
The term for information, such as sound, when it’s been transformed from its original form – molecules bumping into each other in the air – into an electrical version that can be saved, manipulated, and played back.

Sibilance.
The emphasis of “s” and “sh” sounds.

Sound.
A vibration that is propelled through air, courtesy of air molecules that passes the vibration along, to our ears.

WAVE.
Standard sound files format (abbreviated .WAV) for Windows applications.

Waveform.
The shape of the wave. The graphic representation of a sound wave showing amplitude along the vertical axis and time along the horizontal axis.
Chapter II

Review of Literature

Overview of Digital Audio

Research on visual learning is abundant; however, the literature relative to audio has been seriously neglected. Jaspers and Ji-Ping (1991) remarks that the lack of discipline of audio design makes it necessary to investigate a number of topics concerning the qualities of audio and the use of speech in mediated instruction. When using audio, the message designers should be aware of cultural values that determine their own appreciation of audio as well as that of the public’s.

In previous years, storing and delivering audio for training was expensive and of limited capacity. Today, microcomputers have the capability of storing, processing and delivering audio at a substantial reduction of cost. The amount of storage space has increased based on much larger internal disk drives, external drives such as (Zip® and Jazz®), virtual memory, and greater random-access memory (RAM). Microcomputers are almost commonplace with greater than one gigabyte of storage and thirty-two megabytes of memory. Thus, it is more feasible to add digital audio to CBI than it was several years ago.

With the improvement of audio editing software, it is now an effortless task to record, digitize, edit, and insert audio into CBI. This provides a way of incorporating audio into instruction for designers and developers. Thus, digital audio sampling and audio editing software bring manipulation of the audio track within reach (Jaspers & Ji-Ping, 1991).

Three methods are employed to produce computerized speech: (1) text-to-speech synthesis, (2) linear predictive coding, and (3)-digitized sound. These methods vary tremendously in quality, cost, applications, and hardware requirements. Today, digitized speech is the preferred technique since it is extremely realistic and natural sounding, and its educational potential is enormous. Digitized sounds are recorded, with the use of a sound card, and stored as files on a hard drive or compact disk (CD) (Barron, 1991).

An enormous amount of storage space is required when digitizing sound, however, due to large drive sizes and the availability of external drives, the limitation of hard drive space is no longer a roadblock for designers. Never the less, developers are still concerned with the dynamic range and frequency of sounds and their resultant effects on the computer memory requirements. The developer should have as much RAM as possible to handle the range requirements (Jaspers & Ji-Ping, 1991).

Sound

The brain is aware of the object that emits the sound and the environment. The sonic quality of a sound allows the listener to make judgments about the spatial location, relative size, and environment of the sound (Huber & Runstien, 1997). When combined with visuals, these characteristics can enhance or confuse the visual message depending on how the aural and visual
attributes relate to one another. All sound elements share the characteristics that define location. The spatial location of a sound is perceived by tone quality, relative volume, and amount of reflections, known as echoes/reverb, (Huber & Runstien, 1997).

Tone quality refers to the brightness or dullness of the sound. Sounds that are farther away have less high frequency information, called treble, than those that are closer. Sounds farther away contain more low frequencies, bass. Our hearing apparatus is most sensitive to sounds in what is known as the presence range. By amplifying a sound’s spectral content in this 2kHz range, the audio appears to be closer, louder and more intelligible. Additionally, sounds that are farther away tend to have a lower pitch than those that are closer (Rossing, 1982). Closer sounds are louder than those at a distance. When placing a sound/image source at a distance it should be audible but somewhat softer than those placed in the foreground.

As noted above, sound sources that are at a distance tend to have more reflections or echoes than those nearby (Rossing, 1982). In addition to defining space, these same characteristics also identify location and movement. In order to define movement, stereo sound is desirable. However, the intensity of the sound changes as its source moves laterally. Even in mono, a change in volume, tone quality and pitch can enhance the perception of movement (Rossing, 1982).

When using sound effects in multimedia, many authors overlook the importance of matching the aural and visual characteristics of space and location (Daniels, 1994). Similarly, moving objects whose sound is static can confuse the viewer. This is especially problematic when cognitive information is being supplied by the audio track. Often in animations depicting physical, mechanical, or scientific properties, the sound of the action or process can serve as a valuable cue to the information being presented. However, the sound of a malfunctioning part is often as important as the image itself (Daniels, 1994).

Pitch and echoes usually intimate the size of a sound source. A deep pitched voice with many of echoes would imply a large, perhaps intimidating being while a high pitched human sound conjures up images of children rather than adults. Thus it is important that when a sound is placed inside an environment, the reflections should match the environment (Daniels, 1994).

Thus, it is important to remember that sound is described as an auditory sensation in the ear and that the complexity of the human auditory system is remarkable in function. This system responds to a wide range of stimuli. The hearing function, the ear, identifies the pitch and quality of sound and even the direction of the source. Recent research emphasizes how much hearing depends on the data processing that occurs in the central nervous system as well (Rossing, 1982).

**Frequency and Dynamic Ranges**

Pitch and frequency are closely related. A pitch is perceived from the frequency of a wave. The human ear can detect pitches ranging in frequencies of 20 Hertz to 20,000 Hertz (20 kHz) (Moscal, 1994). In addition, the ear can handle volumes (dynamic ranges) from one decibel to 120 decibels. However, duplicating huge variations of frequencies and volume is a problem for a digital audio system (Davis, 1989).

Decibels (dB) are used for describing sound pressure levels because of their similarity to how humans hear. Developers must, when apply to audio, remember certain rules about decibels. A one-decibel change in sound pressure is impossible for most people to detect. The
average human hears loudness differences at 3 dB increments. This means in order to generate a slight change in perceived loudness, power must be doubled. In order to double the perceived loudness, sound pressure level must be increased by 10 dB, which requires 10 times the power (Moscal, 1994).

Davis (1989) outlined frequency and dynamic ranges provided by various systems in Table 1.

Table 1
Frequency and Dynamic Ranges

<table>
<thead>
<tr>
<th>Audio System</th>
<th>Dynamic Range</th>
<th>Frequency Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Telephone</td>
<td>35 dB</td>
<td>100 Hz to 4 kHz</td>
</tr>
<tr>
<td>AM Radio</td>
<td>45 dB</td>
<td>40 Hz to 4 kHz</td>
</tr>
<tr>
<td>FM Radio</td>
<td>60 dB</td>
<td>20 Hz to 16 kHz</td>
</tr>
<tr>
<td>Compact discs</td>
<td>95 dB</td>
<td>20 Hz to 22 kHz</td>
</tr>
<tr>
<td>Digital Audio (PC)</td>
<td>(variable)</td>
<td>20 kHz to (variable)</td>
</tr>
<tr>
<td>Human ear</td>
<td>120 dB</td>
<td>20 Hz to 20 kHz</td>
</tr>
</tbody>
</table>

Because files generated by audio are proportional to the frequency range and dynamic range of the recording, these factors must be constrained to keep files to a manageable size. When recording audio files, care must be taken to keep the amplitude of the digitized signal within the range of the audio card or distortion will occur (Barron, 1991).

Sampling Rates

The resolution of a digital audio recording is crucial to sound quality and the sampling rate is equally significant. At a sampling rate of 40 kHz, 40,000 measurements are generated for each second of sound. The higher the sampling rate, the wider the frequency response of the recording. The upper frequency limit is slightly less than half the sampling rate. A rate of 44.1 kHz is the current standard for compact discs (Davis, 1989).

Selecting an optimal sampling rate usually involves a trade-off; the higher the sampling rate, the more natural the speech, but the more computer storage space required. Shown in Table 2, Rubin (1995) represents some common recording options.
Table 2
Recording Options

<table>
<thead>
<tr>
<th>Resolution</th>
<th>Sample Rate (kHz)</th>
<th>Stereo/ Mono</th>
<th>File Size for 1 minute</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>16-bit</td>
<td>44.1</td>
<td>Stereo</td>
<td>10.5 MB</td>
<td>CD Quality</td>
</tr>
<tr>
<td>16-bit</td>
<td>44.1</td>
<td>Mono</td>
<td>5.25 MB</td>
<td>Good for dialog and most sound effects</td>
</tr>
<tr>
<td>16-bit</td>
<td>22.05</td>
<td>Stereo</td>
<td>5.25 MB</td>
<td>Lack higher frequencies</td>
</tr>
<tr>
<td>16-bit</td>
<td>22.05</td>
<td>Mono</td>
<td>2.6 MB</td>
<td>Use for dialog</td>
</tr>
</tbody>
</table>

Digital Audio in Multimedia Production

Multimedia developers have employed digital audio as another powerful resource. Digital audio simply refers to music and other sounds, such as speech and sound clips, that have been converted, during recording, from analog into digital data and then reconverted into sounds again for listening.

Since the introduction of Microsoft Windows®, the market for PC sound cards has skyrocketed. Numerous sound card vendors bring sound and music to the desktop. However, the standard was set by Creative Labs Sound Blaster®. Sound cards can double the number of voices the computer can play back simultaneously, bringing richer, more realistic sound to the PC. Many sound cards range from less than $200 to over $1000. They all provide a means for recording sounds and saving them to the hard drive as sound files. In fact, for a sound card to qualify as an MPC (Multimedia PC) compatible card it must have, in addition to MIDI, digital audio recording and playback capability (Rubin, 1995).

The high-end sound cards have advanced sound input and output features that make it more appropriate for people interested in composing digital music rather than just listening to it and digitizing speech in a more natural state. Some cards even come with onboard memory to assist in processing complicated sound. In addition, the high-end cards add 64-voice technology, which produce up to 64 “voices” or individual sounds, simultaneously.

The Macintosh® (Mac) computer systems include built-in digital audio playback capability. Because this component exists on all Macintoshes®, their sound card market is more limited. Third party sound cards are available for the Macintosh®, but are mainly professional level, CD-quality sound. Prices usually start in the neighborhood of about $1000 and go up considerably for professional recording-studio-level equipment (Rubin, 1995).

A high level sound card and audio editing software are needed to work with digital audio. The editing software should be able to save sound files in different formats. The most common formats for the Macintosh® are AIFF and SND. For Windows-based PCs, the WAVE (.WAV) format is by far the most common. In addition, the editing software and sound card allow
developers to compress the recording. These compression rates are expressed as ratios such as 4:1 or 8:1. In general, the greater the ratio, the poorer the sound quality (Rubin, 1995).

Digital audio editors range from $100 to as high as $800. High level audio editors usually include a large waveform display showing the left and right channels (for stereo). A number of small icons are used for editing, viewing and recording functions. Other important features include time compression/expansion, a compressor/limiter, and a parametric equalizer. A compressor/limiter is to restrict the dynamic range of a sound file and a parametric equalizer regulates the sample rate, bandwidth, boost/cut, and channels (Rubin, 1995).

Use of Speech

It would be difficult to deny that speech can play a central role in the delivery of instruction. Speech appears to be the most common source of information (Jaspers, 1991). Most modern models of instruction assume the extensive use of spoken or written words. For example, each of the events of instruction (Gagne, Briggs et al. 1982), from informing the learner of the objectives to enhancing retention and transfer, can be accomplished at some level through spoken words.

Speech is an important part of human communication and can be used effectively to transmit information. Johnston (1987) states that one advantage of using natural speech is the power of the human voice to persuade. Dean and Whitlock (1989) concluded that another advantage is that speech can potentially eliminate the need to display large amounts of text on the screen.

Speech can function as either narration or dialogue. Narration, like text, is often used to deliver concrete information. Narration can be used to eliminate the need for long sections of on-screen text, and increase understanding by using a second channel of communication (Johnston, 1987).

Researchers have expressed concern over techniques and strategies for the use of speech in CBI. Sanders, Benbasset and Smith (1976), for example, argue that for speech should not replace printed material such as manuals on the computer terminal to add significantly to learning.

Research on the use of speech in CBI has produced mixed results. For example, studies conducted with young children learning phonics and vocabulary found no achievement or attitude differences for students in several treatments involving different levels of speech and “motivational” sound. The authors suggest several explanations including the low level of students’ metacognitive skills and the redundant presentation of critical information through text (Sales & Johnston, 1993).

One study examining the different uses of audio in CBI was conducted with special needs students. This study produced dramatically different findings. Chiang (1983) found that audio designed to reward and motivate students was actually detrimental to learning. He argued that audio served as a distracter, causing students to require more study time to achieve the same level of learning.

The use of speech is particularly important when a learner is required to attend to more than one chunk of information at a time. Using an audio explanation allows the learner to focus attention on the information or data displayed. Langdon (1973) suggests that additional learning
effectiveness and efficiency can be expected when speech is used for tasks that require the user to attend to more than one chunk of content concurrently.

In summary, a number of studies have reported mixed results with speech and other auditory elements of CBI. The limited agreement of these findings is in part due to the limited amount of research conducted. The rapid change in technologies and the commercial application of speech in CBI are two reasons for the inadequate amount of research done in this area (Sales & Johnston, 1993).

**Speech Technology**

The tone quality of the narration can have an effect on the listener’s perception. A narration that is bright is perceived to be closer and, therefore, more intimate and trustworthy as opposed to speech that sounds dull and distant (Alten, 1990). Also, a narration track with an amplified presence range will be more intelligible and require less volume to be heard over music or sound effects (Woram, 1989).

The spoken word should be recorded with a consistent, clean sound. The tone quality, average recording level, average pitch, and average tempo of the voice should not change noticeably throughout the recording, except for special effect. Several factors that affect digitized audio: recording level, microphone choice, microphone placement, text position, announcer’s position, noise reduction, EQ, and even the announcer’s voice itself (Bartlett & Bartlett, 1992). Table 3, (Bartlett & Bartlett, 1992) represents four types of microphones commonly used for voice recording.

**Table 3**  
**Type of Microphones**

<table>
<thead>
<tr>
<th>Type of Microphone</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top-quality Lavalier condenser</td>
<td>Clips onto the speaker’s shirt or jacket.</td>
</tr>
<tr>
<td>Flat-response cardioid</td>
<td>Full lows and detailed highs. High-end. Constant distance from microphone.</td>
</tr>
<tr>
<td>Ribbon</td>
<td>Warm and smooth sound. Constant distance from microphone.</td>
</tr>
<tr>
<td>Multiple-D dynamic microphone</td>
<td>Unidirectional to compensate for proximity effect, with a bass response that varies only slightly with distance.</td>
</tr>
</tbody>
</table>

The microphone placement for speech recording affects the pickup of room acoustics, breath and lip noises, table thumps, and sound reflections. A typical distance is 8 to 12 inches.
between the microphone and the speaker. To produce quality sound, the following must be considered: minimize table thumps, minimize sound reflections, control the speaker’s position and voice, and reduce sibilance (Bartlett & Bartlett, 1992).

Overall, the quality of the speech used is an important issue in multimedia applications. Quality production of speech can add an important dimension to a multimedia program.

Foundations of Multimedia

Introduction

The term multimedia refers to a computer mediated environment that incorporates two or more media types such as images (still or moving), text, graphics, sound and other data (Daniels, 1994). The effectiveness of multimedia as an instructional medium is based on the theory of multiple-channel communication. Hannafin and Hooper (1993) conclude that two conditions influence the effectiveness of using multiple channels: (a) information presented in each modality must be congruent (similar and not contradictory), and (b) identical presentations of words in sound and text should be avoided.

Multiple Channel Communication

Audio is a more effective channel than print when the information presented is simple and easily understood by the subjects and for illiterates and semi-illiterates (e.g., children) regardless of the difficulty of information. Print shows increasing advantage over audio for literate subjects roughly proportional to the increasing difficulty in their comprehension of the material (Hartman, 1961).

According to Moore, Burton, & Myers (1996), multimedia is more commonly thought of as a computer environment that interacts with the user. The efficacy of multimedia instruction is based on the assumption that adding a channel of media to transmit a message will effectively increase the amount of communication (Dwyer, 1978). Multiple channel communications can be divided into three theories: the single-channel theory, the cue-summation theory, and the dual-code theory.

Single-channel theory - There is only one channel to higher centers of the brain; therefore, dual-channel transmission can be equal to, but not greater than, the single-channel processing. In fact, if both stimuli arrive at the same time, information jamming may occur and cause the dual-channel effectiveness to be less than either of the single-channel’s (Broadbent, 1958).

Cue summation theory - Dual-channel presentations result in more learning than single-channel presentations because the number of stimuli or cues is increased (Severin, 1967).

Dual-code theory - two types of information (verbal and imaginal) are encoded by separate subsystems, one specialized for sensory images and the other specialized for verbal language (Paivio, 1991).

In 1958, Broadbent declared that there is only one channel of communication to the brain (Broadbent, 1958). Travers (1964) followed by hypothesizing that if there was only one channel to the brain, there would be no advantages to using two channels of information (print and audio, for example). Both Broadbent (1958) and Travers (1964) found what they interpreted as
jamming of the human system when information in two channels were presented through a single-channel link, resulting in a reduction of communication (Severin, 1967).

A body of research seems to support the single channel theory. Dwyer, summarizing the results of nearly fifty studies on visual-verbal presentations, notes that the addition of cues in a second channel (or even excessive cues within the same channel) can be distracting and detrimental to learning (Dwyer, 1978).

Barton and Dwyer (1987) report mixed results in achievement based on audio-redundancy. They found that adding audio helped high IQ students in only one of four tests, and it had a negative effect on low IQ students. Additionally, Muraida and Spector (1992) failed to demonstrate treatment effects for various combinations of visually and auditory presented text. However, Hartman (1961) found that multiple-channel presentation was effective for learning. He evaluated seventeen studies concerning simultaneous audio-print and print. Hartman (1961) concluded that “information simultaneously presented by the audio and print channels [was] more effective in producing learning than [was] the same information in either channel along” (p. 42).

Cue summation theory has also been a part of multichannel communication. The benefits of adding additional media channels when communicating is offered by cue summation theory (Severin, 1967). This theory asserts that learning will be increased when stimuli that share information are presented because they reinforce each other. Recent studies have supported similar findings in that recognition and recall of textual information were increased when adding redundant audio (Emerson & Tumey, 1984).

In an attempt to reconcile the two theories: the single-channel theory and the cue-summation theory, Hsia (1977) hypothesized that communication through multiple channels could be more effective as long as the central nervous system was not overloaded. Moore, Burton, and Myers (1996) suggest that when highly related cues are summated across channels, multiple-channel presentations are superior to single channel presentations.

In 1982, Nugent’s study challenged previous single channel findings. It indicated that students exposed to audio and imagery had increased achievement but not as much with audio and text. Using television and film materials, he criticized previous research for the use of non-representative treatments. Paivio (1971,1991) added to this theme with his dual coding theory that is another integral component of multiple-channel communication. Paivio (1971) found that visual images are usually coded by two referential systems, information that is presented in both the aural and visual modes should increase recall and retention. This theory supports the effectiveness of multiple-channel communications.

Finally, within the last few years, two studies, Aarntzen (1993) and Jaspers & Ji-Ping (1991), have emerged who have worked from Nugent’s (1982) and Paivio’s (1971) research and have focused on the usage of audio in a CBI format. Questions addressing the issues of interactivity and its implications for audio are currently under investigation. Because the student can essentially control every aspect of learning in this new environment, including pace and content, new paradigms need to be developed in regard to audio in multimedia (Aarntzen, 1993). While most of their work has been theoretical, its importance will be ascertained as empirical research is conducted (Jaspers & Ji-Ping, 1991; Aarntzen, 1993).
Levie and Lentz (1982) contend that the attributes and information across the two channels reinforce each other and enhance both recall and comprehension. Daniels (1996) concluded in his study that when across channel redundancy or relevance is high, and processing capacity demands are not exceeded, multiple-channel communication appears to be superior to single-channel communication. The method of presenting information to the learner is a critical factor in how it is encoded, elaborated, stored, and retrieved. Educators and designers also must be aware of the cognitive load that message complexity and media type place on the processing demands of the learner (Daniels, 1996).

Several possible explanations can be made for the discrepant findings on the effectiveness of multi-channel communication. Much of the early research is plagued by poor methodology, sampling error, faulty analysis, and test channel bias (Barron & Kysilka 1993, Dwyer 1978). Many of Travers’ studies presented verbal information across modalities (i.e. text and speech). According to the dual-code hypothesis, this should not necessarily result in dual-coding and improved memory. Barron’s and Atkins’ 1994 study examined the redundancy of text and speech. They concluded that since the redundant channel is between two linguistic formats, no reinforcement of information occurs (Barron & Atkins, 1994).

In summary, information presented across channels should be highly correlated to improve learning and avoid inter-channel interference. Additionally, the multimedia/instructional designer should ensure that images or verbals presented across channels are not in conflict (Daniels, 1994). Moore, Burton and Myers (1996) concluded “the goal for instructional designers is to take advantage of suggestions from multiple-channel research in order to facilitate cognitive processes particularly in the development of multimedia presentations” (p. 868).

**Redundancy**

Research on the interplay between the auditory/verbal and visual/pictorial channels of information processing of electronic media presentations is determined not only by the clarity of the message being presented but also by the associations made between the two cognitive channels during processing (Ottaviani & Black, 1994).

Effectively placed audio can improve a multimedia presentation. Redundancy is one main reason for incorporating audio to ensure that the goals and objectives of a multimedia program are met. As an additional communication channel for multimedia, audio can improve a user’s ability to understand and learn the information presented.

Hsia (1977) shows that redundancy is the key to better communication. He defines redundancy as the transmission of the same or closely related information to the receiver or learner through two sensory channels usually the aural and visual channels. The capacity of the human information system is limited; and, therefore, information loss and error do occur. Redundancy is probably the key to minimizing these negative factors. Thus, presenting information via more than one communication channel increases the chances that the message will be attended to, understood, and remembered.

Several studies unveil the potential increase in learning performance based on a multichannel delivery. However, few studies exist that correlate multichannel effectiveness and redundancy level of computer-based instruction (Barron & Atkins 1994). Barron and Kysilka (1993) indicate that there is no statistical difference in comprehension between a text-based
tutorial program and a combination of text/audio program. They also show that the completion
time of a text/audio lesson is significantly less than the comparable text-based program.

A study conducted by Shih and Alessi (1996) conducted an investigation of the relative
effects of voice versus text on learning spatial and temporal information on overall learning and
on learners’ preferences. They did not find any significant differences of the three presentation
methods (text, voice, or text and voice) on learning different types of information or on overall
learning.

Reese (1984) examined recall and retention of information in television segments with
treatment conditions that consisted of separate audio or video presentations, multiple channel
presentations, and varying levels of redundancy between channels. He found that the
presentation of redundant audio-visual communication was superior. Other studies (Drew &
Grimes, 1987) and (Grimes 1990) found redundant audio and visual communication more
effective for comprehension and problem solving than single channel presentations.

Aarntzen (1993) concluded that presenting a stimulus auditively as well as visually would
lead to better memory performance than presenting it only visually or only auditively. Barron
and Atkins (1994) asserted that employment of redundant channels of information would result
in increased learning. They reported studies that also indicate that visual information dominates
audio when they are presented together and that redundancy theories had a positive effect on
learning (Barron & Atkins, 1994).

Aarntzen (1993) states that, according to redundancy theories, speech accompanied by
the same text would enhance learning. Dwyer (1978) and Nugent (1982) state that when audio
redundancy is introduced into the audio/visual environment, its intent is to assist learners in
further reducing their dependence on the printed instruction.

An abundance of research has been accomplished in the area of textual redundancy, most
Little research, however, is available that investigates the effects of audio-redundancy on
learner’s comprehension in a computer-assisted environment (Lang 1995, Drew & Grimes 1987,
Barron & Atkins 1994).

With respect to redundancy, Paivio (1971) posed the following hypothesis on
redundancy. Memory increases directly with the number of alternative memory codes available
for an item. This would mean that coding an item in two ways would enhance memory more
than coding an item in one way only.

In another study, Severin (1967) designed a series of treatment conditions: audio only,
print only, audio and print, audio and related pictures, and audio and unrelated pictures. The
sample population consisted of 201 seventh-grade students. Results demonstrated that the audio-
print combination was not significantly superior to the audio-alone treatment, causing Severin
(1967) to conclude, “a redundant multichannel communication was not significantly better than a
single-channel communication” (p. 392).

Heuvelman (1985) stated that processing redundant auditive and visual signals would be
faster than processing one-channel signals. He claimed that redundant signals reinforce each
other mutually. The effect of redundancy only takes place when there is little to no background-
oise, or other information that is not redundant with the auditive signal on the screen.
Heuvelman (1985) concluded that when this condition is met, the effect of redundancy would be so strong that information processed by the viewer will significantly improve.

Barron and Atkins (1994) asserted that employment of redundant channels of information would result in increased learning. Reviews of research in this area, however, indicated that studies in multichannel communication typically produce mixed results. They reported that these studies also indicate that visual information dominates audio when they are presented together but this redundancy has a positive effect on learning. The rationale cited by Barron and Atkins (1994) for contradictions between multiple channel communications studies included variations in the definition of redundancy.

Aarntzen (1993) also reported that, according to redundancy theories, speech accompanied by the same text would enhance learning. She stated that redundancy would be important here because a simultaneous presentation of the same message auditorily as well as visually might intensify the processing and storage of the information presented.

The effectiveness of the multimodal approach in improving student achievement is dependent on how the information is organized and synchronized (Dwyer 1978, Nugent 1982). Visualization needs to be redundant and embedded precisely in the instruction where students are experiencing difficulty in acquiring information from the print mode alone. The use of visualization in this fashion has the tendency to reduce students' dependence on the printed instruction in their effort to acquire understanding. When audio redundancy is introduced into the audio/visual environment, its intent is to assist learners in further reducing their dependence on the printed instruction.

Reese (1984) demonstrated that when documents were summarized verbally, while simultaneously being shown verbatim, viewers inevitably read the document and ignored the verbal summary. This appeared to be an involuntary response. Subjects reported they 'can't help' but read the enclosed document (Reese, 1984). The same results were found by van Parreren (cited in Aarntzen, 1993). Only exact duplicates on the screen of what is being said in the narration appear to lead to unimpeded processing of the intended message (Grimes, 1990). Aarntzen stated that this is important with respect to redundancy (Aarntzen, 1993).

When a voice is paraphrasing the text on the screen, it is quite possible that only the text on the screen will be read but the information given by the voice will be ignored. This led to the conclusion that the auditive information should be an exact duplicate of the visual information in order to facilitate information processing. Aarntzen (1993) stated that when the two messages are not duplicates, they might even hinder information processing which applies to spoken and written text.

The literature on multiple channel communication is inconsistent. However, Daniels (1996) concluded that there are some general observations that can be made:

- Regardless of how or why, excessive cues (information) can exceed the processing capacity of the human system.

- When across channel redundancy or relevance is high, and processing capacity demands are not exceeded, multiple-channel communication appears to be superior to single-channel communication.
- Cue-summation theory, dual-code theory, and the concept of a limited capacity of processing contribute to the visual superiority effect (p. 21).

Instructional designers find these generalizations important to the design of multimedia. How information is presented to the learner is a critical factor in how it is encoded, elaborated, stored, and retrieved. Educators and designers also must be aware of the cognitive load that message complexity and media type place on the processing demands of the learner (Daniels, 1996).

Compressed Speech

Goldstein (1940) conducted a study of presenting material at different rates of speed where he compared visual and auditory methods of presenting materials. He presented speeds of 100, 137, 174, 211, 248, 285, and 322 wpm. A trained professional read the materials at each of the rates and phonograph recordings were prepared for the experiment. For both the auditory and the visual presentations, the rate of presentation was found a significant factor influencing comprehension. The auditory presentation proved to be superior at the slower rates of presentation, but the faster rates of auditory and visual conditions produced negligible differences. He concluded that no single sensory modality was preferable because of the negligible differences between the auditory and visual conditions (Eckhardt, 1970).

One of the several propositions by Cheatham (1950) quoted by Hartman (1961), in a review of single and multiple-channel communications, which is applicable to the study of compressed speech is that the rate of transmission of speech is limited to the speaking rate; whereas, visual presentation can be faster. This proposition is mentioned because it reflects the basic assumptions of the present research.

The reviewed literature forms the theory base for this study. Research on the usefulness of compressed speech as a viable instructional innovation focuses on the effects of compression on the intelligibility and comprehension of the material. The ability of the listener to accurately repeat single words or brief phrases is employed as a measure of intelligibility (Foulke & Sticht, 1969).

Additional reviews are now reported on studies dealing expressly with the investigation of intelligibility and comprehension of compressed speech. For example, Johnston (1987) stated that although listening can be as effective as reading for many purposes, transmitting information orally can be slow. The reason stated is that learners can read or comprehend information faster than a person can speak. Thus, an important consideration in the use of speech to present information is whether normal speech will be too slow and cause users to lose interest (Fulford, 1993).

Several studies were conducted over the past twenty years that focused on the comprehension and intelligibility of compressed speech. Factors affecting intelligibility have effects on comprehension as well. There are, however, consistent findings that intelligibility declines less rapidly with greater degrees of compression than does comprehension (Foulke & Sticht, 1969). Diehl et al. (1959) found no change in listening comprehension within the compression rate of 126 -272 wpm. Without reference to actual word rates, Nelson (1948) and Harwood (1955) found slight but insignificant loss in listening comprehension as wpm increased.
A study by Harwood (1955), conducted seven years later, resulted in similar findings. Fairbanks et al. (1957) found little difference in comprehension of listening selections at 141, 201, and 282 wpm. Thereafter, comprehension, as indicated by the percentage of text questions correctly answered, declined from 58% at 282 wpm to 26% at 470 wpm.

Foulke and Sticht (1967) found a 6% loss in comprehension between 225 and 325 wpm and a 14% loss from 325 to 425 wpm. They reported that they found an initial moderate linear decline in comprehension, followed by an accelerating decline. A rapid decline in listening comprehension that commenced beyond a word rate of approximately 275 wpm, regardless of the percentage of compression required to achieve that rate, was found by Fairbanks, et al. (1940), and Goldstein (1940).

Foulke (1964) investigated the effect upon the listener’s comprehension when different speakers recorded the same material. He compressed the same material read by three different readers to 275 wpm. The administration of a listening comprehension test over this material showed that significant differences were associated with the speaker variable and with the word-rate variable. However, the speaker’s effect on listening comprehension did not depend upon the word rate at which the selection was presented (Foulke, 1964).

Fairbanks et al. (1940), Goldstein (1940), and Nelson (1948) found a positive relationship between intelligence and the ability to comprehend accelerated speech. A significant positive correlation between reading rate and the ability to comprehend accelerated speech was reported by Goldstein (1940) and Orr et al. (1965). They also found that practice in listening to compressed speech resulted in improved reading rates (Goldstein, 1940; Orr, Friedman et al. 1965). Goldstein (1940) and Travers (1964) found simultaneous reading and listening at 350 wpm resulted in better comprehension than either mode alone. Thames and Rossiter Jr. (1972) reported significant gains in reading rates with the use of compressed speech as a pacing device.

Watts (1969) conducted an experiment where he presented recorded lectures compressed approximately 37% to senior U.S. Air Force Officers attending an instructor course. He found that students listening to compressed speech achieved approximately the same as students listening to live lectures. He also found significant advantage for compressed materials. The word rate reported by Watts (1969) for his experiment was approximately 250 wpm. He concluded that adult students had good attitudes toward using compressed speech materials and he observed achievement for the participants. Watts (1969) restricted the generalization of compressed speech to “adults,” since he defined his sample as military adults, virtually all were college educated.

Short (1972) found no significant differences in post test scores of college students in a nutrition class using speech compressed to 20%, 30% and 40% when compared with scores of students in the same class who listened to normal rate recordings. She also reported that students spent 40% to 50% more time in the carrel area when listening to compressed tapes – over the original recording time. Conversely, students listening to non-compressed tapes spent about 100% more time in the carrel area than the original recording time. When students were allowed to choose their compression rates, the greatest percentage (41% of the students) chose the 20% compression rate, while 28% chose the 30% compression rate, 6% chose 40% compression while the rest of the students preferred normal rate recordings.

A more recent study by Tripp and Roby (1996) found that people can comprehend speech faster than narrators can speak. In addition, learner characteristics (sex, age, reading ability, etc)
may influence the comprehension of compressed speech. Junor (1992) found time compressed speech may provide a valuable increase in presentation rate as well as an increase in comprehension.

Natural speech, the communication medium of lectures and taped talks, is a very time-wasting medium for ideal transmission in comparison with print. The common production rate of a speaker is around 120-170 words per minute (wpm) but the reading rate of a high school graduate will be in the vicinity of 300-350 wpm when silently reading a newspaper (Junor, 1992).

An investigation by Carver (1973) showed that listeners retained information better than readers up to approximately 240 wpm. After this, listener performance dropped dramatically, but readers continued to function well up to 450 wpm. It was found that when students were allowed to determine their own delivery rate for presentations by manipulating the tape machines, their chosen speed of 207 wpm was substantially above that of the normal delivery rate of the speaker (Carver, 1973).

Fulford (1993) investigated the use of systematically designed text alone, text with normal speech, and text with compressed speech (262 wpm). She reports that text with compressed speech was as effective for mastery of objectives, as efficient as text alone, and more efficient than text supported by normal speech. She also suggests that systematically designed materials those that are comprised only of “need-to-know” information, with little or no “nice-to-know” content, may have a high density of information and thus be “over-compressed” when speech compression is used. This is an important consideration if speech compression is contemplated as a way of increasing the efficiency of speech delivery. Thus, incorporating carefully designed compressed speech with text and visuals can increase motivation and improve the effectiveness of a multimedia program (Fulford, 1993).

It is evident from this research that learners have the potential to deal with rates of information which are much higher than the normal delivery rate due to the brain’s information handling capacity. The major limiting factor appears to be a lack of capability to cognitively process audio information, which is delivered at too high a rate (Tripp & Roby, 1996).

Research Questions

Several questions arise from the review of the literature concerning the rate of presentation and recall. Questions under consideration are:

1) How does the rate of speech influence learning in a multimedia setting?
2) How practical is it to incorporate compressed speech in this setting?
3) What is the educational value of compressed speech in this environment?
4) Does the use of audio in a multimedia environment affect learning as measured by recall scores of college-level students?
Chapter III

Methodology

The problem under consideration in this study is whether or not compressed speech could be used as a technique for increasing recall of college students. The determination of the affects of presentation rates of speech on recall in a multimedia environment is the primary objective of the present study. In designing the investigation, the assumption was made that the presentation of compressed speech improves recall of information. A second assumption was made that the human mind can only utilize a limited amount of information, seven plus or minus two, in a specific period of time (Miller, 1956).

Research Hypotheses

Based on the assumptions a research hypothesis was developed stating that compression rates of speech have differential effects on recall. The null hypotheses were tested at the .05 level of significance. It is hypothesized that an experimental research study examining the research questions stated above would produce the following results:

Research Hypothesis 1

The recall of information at various levels of compressed speech decreases when audio delivery-rate (words-per-minute) increases.

Research Hypothesis 2

There is an interaction between task completion times and recall of information based on words-per-minute rate of presentation.

Research Hypothesis 3

Relative to recall, there is an interaction between audio delivery-rate and increased exposure to an audio stimulus.

Subjects

This study was conducted at two locations: North Carolina A&T State University in Greensboro, North Carolina and Virginia Tech in Blacksburg, Virginia.

The population of this study consisted of undergraduates majoring in business at both institutions. In order to obtain the student’s demographic information such as age, race, sex, classification, and grade point average, an on-line questionnaire was designed. Each student was required to complete this information before proceeding.

Experimental Design

There were three test experimental groups: normal speech, 125 words per minute; compressed speech, 175 words per minute, 40% compression; and, compressed speech, 200 words per minute, 60% compression. A sample population of 192 students was randomly assigned to one of three treatment groups.
The variables used in this study are:

**Dependent variable:**
- Recall
- Time

**Independent (treatment) variable:**
- Rate of delivery of the instruction.
  This is controlled by the rate of words per minute.

The experimental design consists of three treatment groups (64 students in each group, randomly assigned, using random numbered diskettes located in a box at the entrance to the computer lab). Identical computer-based instruction is developed in three alternate designs:

- **Group I—E1** Normal speech, 125 words per minute
- **Group II—E2** Compressed speech, 175 words per minute, 40% compression
- **Group III—E3** Compressed speech, 200 words per minute, 60% compression

**Experimental Design Paradigm**

<table>
<thead>
<tr>
<th>R</th>
<th>X1</th>
<th>O1—X1</th>
<th>normal speech, 125 words per minute</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>X2</td>
<td>O1—X2</td>
<td>compressed speech, 175 words per minute, 40% compression</td>
</tr>
<tr>
<td>R</td>
<td>X3</td>
<td>O1—X3</td>
<td>compressed speech, 200 words per minute, 60% compression</td>
</tr>
</tbody>
</table>

- **R** = Random Assignment
- **X** = Treatment Variable
- **O** = Outcome Variable

The recordings for the audio instruction were made by the same professional voice as recommended by Utz (1970), whose reading rate is an average of 125 words per minute. The audio was digitized using a sampling rate at or near 7 kHz (Barron, 1991). The dynamic range of the narration was kept within the 40 dB. This rate determined by Nichols and Stevens (1978) as the average conversational rate was used for the purposes of this study. There has been an accumulation of experimental results which support the general conclusion that speech may be presented at a rate of 126-224 words per minute with the expectation of satisfactory comprehension (Foulke, 1971). Foulke and Stitch (1967) found a 6% loss in comprehension between 225 and 325 words per minute. Compressed speech rates of 175 wpm, (40% compression rate) and 200 wpm (60% compression rate) were used.
Experimental Environment

Two multimedia-equipped student computer laboratories were used for this experiment. The student laboratories were located in Merrick Hall in the School of Business & Economics and Fort IRC building on A&T’s campus and WMH 220 on Virginia Tech’s campus. The labs were closed during the experimentation period to control for outside distractions. Each student was randomly assigned an individual computer workstation.

Experimental Apparatus Used

Windows-based personal computers with the following specifications: Pentium® 166 processor, 32 MB RAM and 2 GB hard drive was used to conduct this experiment. The audio instruction treatment groups used headphones to minimize interference and distraction between treatments. Control of the audio volume was allowed while viewing the instruction by using headphones with volume control. Each computer had the mouse speed controls set up at the same rate to avoid different mouse speeds. CoolEdit Pro® was used to compress speech and to manipulate sound files. Toolbook® was the application for the experimental unit.

Procedures

Diskettes marked with the group assignment number were located in a box at the front entrance to the lab. Each student randomly selected a diskette before proceeding to the computer workstation. The treatment group was decided by the first number of the assigned four-digit number found on the label of the diskette. The researcher gave brief instructions about starting the instruction. All other instructions were received on-line. Each group received the instructions in normal conversational voice.

The student began with the introductory screen (see Figure B 1) and then proceeded to complete the information on the demographic screen (see Figure B 2). The student must insert the diskette (see Figure B 3). The next screen allowed the student to hear the phonetic alphabet used in the module (see Figure B 4). The phonetic alphabet could be repeated as many times as necessary by the student. The practice and instructions screen was viewed next (see Figure B 5 (see Figure B 6). The student was given the redundant audio and textual instructions. The student was allowed to practice moving the pieces of the puzzle. He or she could practice as long as necessary. Upon completion of the practice session, the students were allowed to begin the main module, which was the puzzle. It was necessary for the student to shuffle the puzzle pieces by clicking on the shuffle button; thus ensuring random selection. Then, he or she must click the click hear for solution button (see Figure B 7). The solution was given audibly at a rate based on the treatment condition. During the audible solution, the puzzle was covered with a white screen to keep the student from using visual cues and patterns. Group E1 received each solution to the module with digitized recordings at a speech rate of 125 words per minute. Group E2 received each solution to the module as Group E1 but with compressed speech rate of 175 words per minute. Group E3 received the same with a compressed speech rate of 200 words per minute. After completing the third puzzle, the student was dismissed and thanked for participating. The whole procedure could be completed within ten minutes.
Method of Analysis

All data were collected within the computer program and analyzed using SPSS™. One hundred and ninety-two students were used in this study. To determine the sample size and the power of the study, only 90 students were needed for a precision of 10% where the confidence level was 95% and P=0.05.

This experiment is designed to test hypotheses regarding the number of attempts, time, and number of correct continuous attempts within and between students. Results of the study based on empirical data are derived from descriptive statistics and General Linear Model - repeated measures. The level of significance was set at .05.
Chapter IV

Results

The results of the analyses performed on the data are presented in this chapter based on the performance of 184 students divided into three audio treatment groups. The dependent measures included (1) total attempts, (2) amount of time (hh:mm:ss) required to complete the puzzle, and (3) continuous correct recall scores.

Additional student information included: estimated GPA, gender, age, ethnicity, academic status, hearing disorder, and amount of exposure to computer games. The tables of frequencies and descriptive statistics can be found in Appendix F.

Of the one hundred ninety-two participants, eight were eliminated. Two students answered “yes” to a hearing disorder and six students were graduate students, leaving one hundred eighty-four participants.

Every student was placed one of three experimental treatment groups focused on the recall of information. Trial is the repeated measure or within-student effect, across three levels; they are trial 1, trial 2, and trial 3. The between-group effect represents the presentation rate of speech that each student is assigned to, and they are: treatment group 1 - 125 wpm, treatment group 2 - 175 wpm, and treatment group 3 - 200 wpm.

Research Hypothesis 1
The recall of information at various levels of compressed speech decreases when audio delivery-rate (words-per-minute) increases.

Tests of the first hypothesis were conducted using repeated measures ANOVA with total attempts as the measure during three separate trials.

Table 4 contains the descriptive statistics for the trial 1, trial 2, and trial 3 within treatment group 1 (n=63), treatment group 2 (n=66), and treatment group 3 (n=55).

In Table 5, the results of the repeated measures ANOVA show that the students in treatment group 1 performed better, lower number of attempts to solve the puzzle, than students in treatment groups 2 and 3. It appears that student attempts in treatment group 1 decreased in trial 2 but regressed in trial 3, however the student attempts in treatment groups 2 and 3 decreased for each trial.

The Tukey HSD multiple comparisons of means was employed to identify those means that were different. The data for the multiple comparisons of the measure, the number of attempts, based on estimated marginal means is found in Table 6. The only meaningful finding dealt with attempts in Trial 2 between 125 wpm - 175 wpm and between 125 wpm - 200 wpm at the .05 level of significance.
Table 4

Statistical Values (Attempts): Descriptive

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Attempts-Trial 1</th>
<th>Attempts-Trial 2</th>
<th>Attempts-Trial 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>125 wpm</td>
<td>Mean</td>
<td>41.3810</td>
<td>32.6984</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>63</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>10.1718</td>
<td>12.8699</td>
</tr>
<tr>
<td>175 wpm</td>
<td>Mean</td>
<td>42.0000</td>
<td>39.0606</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>66</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>9.3348</td>
<td>10.6336</td>
</tr>
<tr>
<td>200 wpm</td>
<td>Mean</td>
<td>42.8727</td>
<td>37.6364</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>55</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>7.9210</td>
<td>8.1431</td>
</tr>
<tr>
<td>Total</td>
<td>Mean</td>
<td>42.0489</td>
<td>36.4565</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>184</td>
<td>184</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>9.2129</td>
<td>11.1163</td>
</tr>
</tbody>
</table>

Note. Students (n = 184)
Table 5
Repeated Measures ANOVA for Attempts

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>dF</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>1231.664</td>
<td>2</td>
<td>615.832</td>
<td>5.036</td>
<td>.007*</td>
</tr>
<tr>
<td>Error</td>
<td>22133.668</td>
<td>181</td>
<td>122.285</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>dF</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial</td>
<td>4120.275</td>
<td>2</td>
<td>2060.138</td>
<td>27.393</td>
<td>.000*</td>
</tr>
<tr>
<td>Trial x</td>
<td>836.114</td>
<td>4</td>
<td>209.028</td>
<td>2.779</td>
<td>.027*</td>
</tr>
<tr>
<td>Error(Trial)</td>
<td>27224.745</td>
<td>362</td>
<td>75.206</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Based on observed means. The error term is Error.

* - The significance level is set at .05.
Table 6
Tukey HSD Multiple Comparisons of Treatment Means

Tukey HSD

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>(I) Treatment</th>
<th>(J) Treatment</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responses 1</td>
<td>125 wpm</td>
<td>175 wpm</td>
<td>-.6615</td>
<td>1.618</td>
<td>.912</td>
</tr>
<tr>
<td></td>
<td>200 wpm</td>
<td></td>
<td>-2.0439</td>
<td>1.699</td>
<td>.451</td>
</tr>
<tr>
<td></td>
<td>175 wpm</td>
<td>125 wpm</td>
<td>.6615</td>
<td>1.618</td>
<td>.912</td>
</tr>
<tr>
<td></td>
<td>200 wpm</td>
<td></td>
<td>-1.3823</td>
<td>1.699</td>
<td>.694</td>
</tr>
<tr>
<td></td>
<td>200 wpm</td>
<td>125 wpm</td>
<td>2.0439</td>
<td>1.699</td>
<td>.451</td>
</tr>
<tr>
<td></td>
<td>175 wpm</td>
<td></td>
<td>1.3823</td>
<td>1.699</td>
<td>.694</td>
</tr>
<tr>
<td>Responses 2</td>
<td>125 wpm</td>
<td>175 wpm</td>
<td>-6.8308*</td>
<td>1.883</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>200 wpm</td>
<td></td>
<td>-5.9670*</td>
<td>1.976</td>
<td>.007</td>
</tr>
<tr>
<td></td>
<td>175 wpm</td>
<td>125 wpm</td>
<td>6.8308*</td>
<td>1.883</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>200 wpm</td>
<td></td>
<td>.8638</td>
<td>1.976</td>
<td>.900</td>
</tr>
<tr>
<td></td>
<td>200 wpm</td>
<td>125 wpm</td>
<td>5.9670*</td>
<td>1.976</td>
<td>.007</td>
</tr>
<tr>
<td></td>
<td>175 wpm</td>
<td></td>
<td>-.8638</td>
<td>1.976</td>
<td>.900</td>
</tr>
<tr>
<td>Responses 3</td>
<td>125 wpm</td>
<td>175 wpm</td>
<td>-.6769</td>
<td>1.492</td>
<td>.893</td>
</tr>
<tr>
<td></td>
<td>200 wpm</td>
<td></td>
<td>-2.3983</td>
<td>1.567</td>
<td>.276</td>
</tr>
<tr>
<td></td>
<td>175 wpm</td>
<td>125 wpm</td>
<td>.6769</td>
<td>1.492</td>
<td>.893</td>
</tr>
<tr>
<td></td>
<td>200 wpm</td>
<td></td>
<td>-1.7214</td>
<td>1.567</td>
<td>.515</td>
</tr>
<tr>
<td></td>
<td>200 wpm</td>
<td>125 wpm</td>
<td>2.3983</td>
<td>1.567</td>
<td>.276</td>
</tr>
<tr>
<td></td>
<td>175 wpm</td>
<td></td>
<td>1.7214</td>
<td>1.567</td>
<td>.515</td>
</tr>
</tbody>
</table>

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.
Research Hypothesis 2

There is an interaction between task completion times and recall of information based on words-per-minute rate of presentation.

Tests of the second hypothesis found no interaction between treatment groups based on completion times and the recall of information. It was expected that students who completed the puzzle quicker would have improved responses, which could not be confirmed.

Table 7 shows the descriptive statistics for completion times for each trial across the three treatments groups.

In Table 8, repeated measures ANOVA was performed. It appears that the students in all treatment groups finished their responses faster between each trial but without improved accuracy.

Research Hypothesis 3

Relative to recall, there is an interaction between audio delivery-rate and increased exposure to an audio stimulus.

Table 9 summarizes the descriptive statistics for the longest strand of correct attempts chunked by the students. It appears that students in treatment group 1 improved in trial 2 but the correct attempts decreased in trial 3. Treatment group 2 improved in trial 2 but showed no change in trial 3. Treatment group 3 improved in both trials 2 and 3.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Completion Time-Trial 1</th>
<th>Completion Time-Trial 2</th>
<th>Completion Time-Trial 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>125 wpm</td>
<td>Mean</td>
<td>0:02:14</td>
<td>0:01:17</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>63</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>0:00:54</td>
<td>0:00:37</td>
</tr>
<tr>
<td>175 wpm</td>
<td>Mean</td>
<td>0:02:10</td>
<td>0:01:20</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>66</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>0:00:53</td>
<td>0:00:30</td>
</tr>
<tr>
<td>200 wpm</td>
<td>Mean</td>
<td>0:01:59</td>
<td>0:01:15</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>55</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>0:00:51</td>
<td>0:00:23</td>
</tr>
<tr>
<td>Total</td>
<td>Mean</td>
<td>0:02:08</td>
<td>0:01:17</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>184</td>
<td>184</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>0:00:53</td>
<td>0:00:31</td>
</tr>
</tbody>
</table>

Note. Students (n = 184)
Table 8
Repeated Measures ANOVA for Time

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>dF</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>1400.731</td>
<td>2</td>
<td>700.366</td>
<td>.277</td>
<td>.758</td>
</tr>
<tr>
<td>Error</td>
<td>457134.0</td>
<td>181</td>
<td>2525.602</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>dF</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial</td>
<td>390967.5</td>
<td>2</td>
<td>195483.7</td>
<td>198.771</td>
<td>.000*</td>
</tr>
<tr>
<td>Trial x</td>
<td>3162.937</td>
<td>4</td>
<td>790.734</td>
<td>.804</td>
<td>.523</td>
</tr>
<tr>
<td>Error(Trial)</td>
<td>356014.1</td>
<td>362</td>
<td>983.464</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Based on observed means. The error term is Error.

* - The significance level is set at .05.
Table 9
Statistical Values (Correct Attempts): Descriptive

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Mean Max Contiguous Responses Trial #1</th>
<th>Mean Max Contiguous Responses Trial #2</th>
<th>Mean Max Contiguous Responses Trial #3</th>
</tr>
</thead>
<tbody>
<tr>
<td>125 wpm</td>
<td>1.9524</td>
<td>2.9524</td>
<td>2.3651</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>63</td>
<td>63</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>Std. Deviation</td>
<td>Std. Deviation</td>
</tr>
<tr>
<td></td>
<td>.8314</td>
<td>1.3729</td>
<td>1.2221</td>
</tr>
<tr>
<td>175 wpm</td>
<td>Mean 1.7273</td>
<td>2.1061</td>
<td>2.1061</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>66</td>
<td>66</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>Std. Deviation</td>
<td>Std. Deviation</td>
</tr>
<tr>
<td></td>
<td>.7348</td>
<td>1.0096</td>
<td>.9136</td>
</tr>
<tr>
<td>200 wpm</td>
<td>Mean 1.6909</td>
<td>2.0000</td>
<td>2.1273</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>55</td>
<td>55</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>Std. Deviation</td>
<td>Std. Deviation</td>
</tr>
<tr>
<td></td>
<td>.7422</td>
<td>.7935</td>
<td>.9633</td>
</tr>
<tr>
<td>Total</td>
<td>Mean 1.7935</td>
<td>2.3641</td>
<td>2.2011</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>184</td>
<td>184</td>
<td>184</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>Std. Deviation</td>
<td>Std. Deviation</td>
</tr>
<tr>
<td></td>
<td>.7758</td>
<td>1.1702</td>
<td>1.0443</td>
</tr>
</tbody>
</table>

Note. Students (n = 184)
Table 10 shows repeated measures ANOVA for the correct number of continuous attempts. The students chunked information better in treatment group 1 than treatment groups 2 and 3 across all trials. It confirms that recall improves between trials 1 and 2 as indicated by the number of consecutive correct attempts. Although, it was expected that chunking would have occurred in the seven plus or minus two range (Miller, 1956), students chunked at a mean of three or less. A possible observation is that Miller’s research did not address chunking with compressed speech.

The data for the Tukey multiple comparisons of the measure continuous correct attempts based on estimated marginal means are found in Table 11. The only noted improvements are the attempts in trial 2 between treatment groups 1 and 2, and between treatment groups 1 and 3.

Table 10
Repeated Measures ANOVA for Correct Number of Continuous Attempts

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>dF</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>27.680</td>
<td>2</td>
<td>13.840</td>
<td>14.417</td>
<td>.000*</td>
</tr>
<tr>
<td>Error</td>
<td>173.762</td>
<td>181</td>
<td>.960</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>dF</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial</td>
<td>29.823</td>
<td>2</td>
<td>14.911</td>
<td>15.827</td>
<td>.000*</td>
</tr>
<tr>
<td>Trial x</td>
<td>17.809</td>
<td>4</td>
<td>4.452</td>
<td>4.725</td>
<td>.001*</td>
</tr>
<tr>
<td>Error(Trial)</td>
<td>341.065</td>
<td>362</td>
<td>.942</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Based on observed means. The error term is Error.

* - The significance level is set at .05.
# Table 11

**Tukey HSD Multiple Comparisons of Treatment Means – Correct Attempts**

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Mean Difference (l-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>125 wpm</td>
<td>.1846</td>
<td>.136</td>
<td>.362</td>
</tr>
<tr>
<td>200 wpm</td>
<td>.2194</td>
<td>.143</td>
<td>.273</td>
</tr>
<tr>
<td>175 wpm</td>
<td>-.1846</td>
<td>.136</td>
<td>.362</td>
</tr>
<tr>
<td>200 wpm</td>
<td>3.5E-02</td>
<td>.143</td>
<td>.968</td>
</tr>
<tr>
<td>125 wpm</td>
<td>-.2194</td>
<td>.143</td>
<td>.273</td>
</tr>
<tr>
<td>175 wpm</td>
<td>-3.E-02</td>
<td>.143</td>
<td>.968</td>
</tr>
<tr>
<td>125 wpm</td>
<td>.8923*</td>
<td>.188</td>
<td>.000</td>
</tr>
<tr>
<td>200 wpm</td>
<td>1.0926*</td>
<td>.198</td>
<td>.000</td>
</tr>
<tr>
<td>175 wpm</td>
<td>-.8923*</td>
<td>.188</td>
<td>.000</td>
</tr>
<tr>
<td>200 wpm</td>
<td>.2003</td>
<td>.198</td>
<td>.569</td>
</tr>
<tr>
<td>125 wpm</td>
<td>-1.0926*</td>
<td>.198</td>
<td>.000</td>
</tr>
<tr>
<td>175 wpm</td>
<td>-.2003</td>
<td>.198</td>
<td>.569</td>
</tr>
<tr>
<td>125 wpm</td>
<td>.2308</td>
<td>.183</td>
<td>.418</td>
</tr>
<tr>
<td>200 wpm</td>
<td>.1903</td>
<td>.192</td>
<td>.584</td>
</tr>
<tr>
<td>175 wpm</td>
<td>-.2308</td>
<td>.183</td>
<td>.418</td>
</tr>
<tr>
<td>200 wpm</td>
<td>-4.E-02</td>
<td>.192</td>
<td>.976</td>
</tr>
<tr>
<td>125 wpm</td>
<td>-.1903</td>
<td>.192</td>
<td>.584</td>
</tr>
<tr>
<td>175 wpm</td>
<td>4.0E-02</td>
<td>.192</td>
<td>.976</td>
</tr>
</tbody>
</table>

Based on observed means. The error term is Error.

* The mean difference is significant at the .05 level.
Figure 1. Means – Number of Attempts per Treatment Group per Trial

Number of Attempts per Treatment Group per Trial

![Bar chart showing number of attempts per treatment group per trial.](chart.png)
Figure 2. Means – Total Completion Time per Treatment Group per Trial
Figure 3. Means – Correct Continuous Attempts per Treatment Group per Trial
Discussion

Introduction

Based on the review of literature presented in this study, several observations seem justified which influenced the design of this study and the interpretation of results: (1) students prefer listening rates in the range 175 to 200 wpm (Lovitt, 1974); (2) comprehension of multimedia materials is not seriously affected by moderate compression rates (Challis, 1973); (3) additional learning effectiveness and efficiency can be expected when speech is used for tasks that require the user to attend to more than one chunk of content concurrently (Langdon, 1973). Finally, compressed speech in a multimedia learning environment can save time (Short, 1972).

This study examined learning outcomes of participants who interacted with one of three treatment groups. The research questions under investigation for this study were:

1) How does the rate of speech influence learning in a multimedia setting?
2) How practical is it to incorporate compressed speech in this setting?
3) What is the educational value of compressed speech in this environment?
4) Does the use of audio in a multimedia environment substantially affect learning as measured by recall scores of college-level students?

Today’s audio editing software make compression tools available and inexpensive to the instructional designer. With the use of a microcomputer with sufficient memory and storage, incorporating audio into instruction is a seamless task for today’s developers and designer. These tools can produce high quality audio.

The audio clips for this model were produced by a professional voice with a speaking rate of 125 wpm. The auditory instructions, in normal speech, were accompanied by the same text which, according to redundancy theories, should enhance learning. Barron and Atkins (1994) assert that employment of redundant channels of information would result in increased learning. In addition, an audio explanation allows the learner to focus attention to the information displayed (Langdon, 1973).

Conclusion

Based on the analysis of the first hypothesis, mean scores of recall were higher for treatment group 1 compared to treatment groups 2 and 3. The students performed better, which means that they had a lower number of attempts in treatment group 1. This analysis failed to support the findings of Diehl (1987) who found no change in comprehension with changes in word rate within the compression rate of 126-272 wpm. Based on these findings, it could be assumed that the students became familiar with their rate of compressed speech and the puzzle. In addition, they may have been paying closer attention during the second trial and really had the desire to solve the puzzle. However, on the third trial, they seemed to be guessing or became anxious about finishing the experiment.

From the analysis of the second hypothesis, the results support the findings of Challis (1973) that there are no significant differences between completion times and the recall of information based on compression rates. Challis (1973) observed that the compressed versions
saved time. Although the data indicated a decrease in time as the rate increases, the differences between the groups were not meaningful. This would suggest an opportunity to save substantial instructional time without recall lost, however, this research experiment did not show such timesaving efficiencies. It is interesting to note that although students completed the puzzle faster by the third trial, their attempts did not improve.

An analysis of the data for hypothesis three indicated that there is an interaction between the three treatment groups and the amount of exposure to the audio stimulus measured by recall scores. Treatment group 1 attempts changed the most between trials. They seem to have done better the second trial than the first and third. In trial 3, only treatment group 3 showed a slight increase. There were negligible differences between the recall scores of treatment group 2 and 3 in any trials. Results in this experiment confirm the earlier findings that recall of information decreases as the rate of presentation increases. Based on the mean scores of the continuous correct attempts, the students chunked less information than was expected. One possible explanation is that the puzzle was difficult to solve and for many students this experiment was their first exposure to the phonetic alphabet.

It is evident from the research that learners have the potential to deal with audio-delivery rates, which are much higher than the normal delivery rate due to the brain’s information handling capacity that supports Tripp & Roby’s study (1996). The conclusion is reached that within the limits of this research, compressed speech can safely be used in a multimedia environment.

**Limitations of Study**

As with the preceding discussions, the implications were drawn exclusively from the analysis and evaluation of those data resulting from the present investigation. The population of this study was limited to business students enrolled in two southeastern universities during the spring semester of the academic year 1997-1998. Generalizations to populations, learning tasks, or environments outside those reported in this study are neither stated nor implied. The study only encompassed one type of learning, immediate recall.

**Impact of Study**

The problem under scrutiny in this study dealt with the presentation of information via the auditory channel in various compressed forms. The impact of this study, therefore, rests in its attempt to determine whether or not compressed speech influences recall in a multimedia setting. To increase effectiveness, the use of compressed speech models should incorporate more than one trial preceded by practice sessions. Instructional designers can develop applications incorporating compressed speech at 175 wpm without serious degradation of recall.

**Summary**

There is ample research to show a need to conduct an experiment by looking at the general uses of audio, the uses of speech, and compressed speech. By doing so, new guidelines should evolve in this area. In the meantime, instructional designers will continue to include speech in a multimedia learning environment to make it more appealing to the learner. As Sales (1993) noted, “regardless of what our research finds, we may not have a choice but to include speech in instructional software”. Without many established design principles, this seems to be a precarious position. Designers need to continually ask pertinent questions and study and test new ideas in order to avoid this complacent attitude.
Although little research exists on audio instruction, the literature revealed that a distinct relationship exists between compressed speech, its effects on recall, and its influence on the message that is heard. There exists a need to conduct further research regarding the use of compressed speech in a multimedia environment. Consideration given to the saving of student and instructor time should be investigated to justify the production of instruction at the compressed rates, although this study found no significant differences.

Implementing audio into computer-based instruction for instructional designers and developers is becoming more important each day. Some of the possible benefits of incorporating compressed speech into instruction include the ease in manipulating compressed speech using the user friendly editing features of the audio editing software in today’s software markets. These tools can be used on desktop computers and is extremely cost effective.

A systematic approach needs to be developed to implement more effective acquisition on the part of learners (Barton & Dwyer, 1987). Dwyer (1978) and Nugent (1982) verify that visualizing alone is not always sufficient to enable learners to achieve at predetermined levels of accomplishment (Barton & Dwyer, 1987).

According to redundancy theories, speech accompanied by the same text should enhance learning (Aarntzen, 1993; Heuvelman, 1985; Paivio, 1971). Recognition and recall of verbal information increased when adding redundant audio (Emerson & Tumey, 1984; Barron & Atkins, 1994). In addition, the focus on redundancy factors and multiple channel transmission of communication evaluated the effectiveness of audio CBI on learner achievement (Severin, 1967; Hsia, 1968, Dwyer, 1978; Broadbent, 1957).

Research relating to the instructional uses of compressed speech could be accurately described as promising. To realize its full instructional potential, considerably more research and experimentation is needed, especially in a multimedia environment. Instructional designers and developers should give serious consideration to the application of compressed speech in multimedia application.

Suggestions for Future Research

During several phases of this study, to include planning, conducting of the experiment, and analysis of data, additional hypotheses relating to using compressed speech in a multimedia learning environment continued to surface. The findings of this study strongly suggest that further research be conducted to confirm or reject the hypothesis that compressed speech influences recall in a multimedia setting.

Future studies need not be confined to these presentation rates of speech. However, consistent with previous studies, the investigator should keep the rates between 124 and 224 words per minute for intelligibility.

It is recommended that in replicating this study, the investigator consider adding a repeating option to the audible solution to the puzzle. The purpose of the study would be to investigate whether the addition of the repeat option will produce greater recall of information.

Further research could investigate the effect of learner control over the various levels of compressed speech on recall. Is there an advantage in allowing the student to choose the rate of speech? Will recall improve when students are allowed to adjust the compression rate?
It is recommended that parallel versions of the model with the sex of the narrator as an independent variable should also be investigated. What is the comparative effectiveness of a female voice versus a male voice for the audible instructions and solutions?

Additionally, it is recommended that further research be conducted on the interaction between the ability to recall information and the various levels of speech compression using treatment-aptitude models.

It is also recommended that an investigation be performed on how to apply compressed speech on other learning attributes such as cognitive complexity.

A final recommendation is to study the effects of visuals versus the various rates of compressed speech. These studies should analyze the relative advantages of visuals and compressed speech and report the results of an experimental study.
Reflections on Conducting Research

The reflections on conducting research can be found in Appendix H. The researcher discusses the framework used to successfully complete and defend this dissertation. This approach is highly recommended for consideration by future doctoral students.
References


Goldstein, H. (1940). “Reading and listening comprehension of various controlled rates.” Teachers College Contributions to Education 821.


Israel, Glenn D., Determining sample size, Fact Sheet PEOD-6, 1992, University of Florida, Gainesville.


Appendix A: Pilot Study

Thirty-one people were asked to evaluate the module for this experiment during the week of March 17. During the pilot study the researcher carefully observed each participant. Initial instructions were given, then they were asked to follow the instructions given on each screen. First, the student was given a diskette with an assigned four-digit number on it. The first digit of the number represented what treatment was administered to the student and the remaining digits were the student control number. Screen One included demographic questions that had to be answered before continuing. On Screen Two the students were given a chance to listen to the phonetic alphabet with a repeat option. Screen Three included the visual instructions with audio redundancy and a practice session. The student was allowed to repeat the visual instructions as many times as needed. At this point the student had a chance to practice clicking and moving the puzzle pieces an unlimited amount of time. Each puzzle piece has the phonetic alphabetic engrained on it. On Screen Four, the student had to shuffle the puzzle pieces, listen to the solution based on the treatment number, and solve the puzzle. There are nine positions on the puzzle of which one position is kept open at all times. Based on Miller’s 1956 theory that the human mind can only utilize a limited amount of information – that is seven plus or minus two— in a specific period of time. During the audible solution, the puzzle is covered with a white screen to keep the student from using visual cues and patterns. Each puzzle piece is active so the student has one in eight chance for each move. The student completed three different versions of the puzzle, randomly selected by the computer. The students were required to move the puzzle pieces into the order given in the audible solution.

After the student completed the module, the researcher interviewed the participant to find out if there were problems with the experimental module or any other part of the experiment. Some of the comments were: more solutions to the puzzle are needed to avoid getting the same puzzle two out of three times, scores on the first puzzle should not be counted, and repeating the solution should be allowed for the compressed speech. Some of these comments were used to adjust the experimental materials accordingly.
Appendix B: Screens Used in the Module

Figure B 1. Introductory Screen
Figure B 2. Demographic Screen
Figure B.3. Insert Diskette Screen
The "Phonetic Alphabet" is used in this experiment. Even if you are familiar with it, please take time to listen to each letter's pronunciation as demonstrated by the narrator.

**Click To Hear Phonetic Alphabet - Repeat If Desired**

- Alpha = A
- Bravo = B
- Charlie = C
- Delta = D
- Echo = E
- Foxrot = F
- Golf = G
- Hotel = H

**Figure B 4.**  Phonetic Alphabet Screen
Figure B 5. Instructions & Practice Puzzle Screen
Figure B.6. Beginning Puzzle Screen
Figure B 7.  Hear the Solution Screen
Figure B 8. Solve the Puzzle Screen
Figure B.9  Correct Response Screen
Figure B 10. Incorrect Response Screen
Figure B 11.   End of Puzzle 1 Screen
Figure B.12. End of Puzzle 2 Screen
Figure B 13. Thank You Screen
The Effects of Speech on Recall in
A Multimedia Environment

by

Marcy E. Johnson

Assigned No. 1001

A&T VT

Figure B 14. Diskette label.

First Digit of Assigned No. Treatment Number

1  125 wpm, 0% Compression Rate
2  175 wpm, 40% Compression Rate
3  200 wpm, 60% Compression Rate

A&T North Carolina A&T State University
VT Virginia Tech
Appendix C: Script for Study

Script for Informing Participants of the Nature of the Study

Good Day! My name is Marcy Johnson. I’m a doctoral student in Instructional Technology at Virginia Tech conducting a study on how compressed speech influences learning and its educational value in a multimedia environment. In addition, I am determining if it is practical to incorporate compressed speech in an academic setting. Student volunteers are needed. The experiment will take about 20 minutes but you may leave at any time during the experiment. The demographic information requested will be needed for this experiment only. Please come to the main student lab located in Merrick Hall or Fort IRC building located on North Carolina A&T State University campus or War Memorial Hall 225 on Virginia Tech’s campus during the weeks of April 8 –15.
Appendix D: Phonetic Alphabet Used in the Module

<table>
<thead>
<tr>
<th>Word</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha</td>
<td>A</td>
</tr>
<tr>
<td>Bravo</td>
<td>B</td>
</tr>
<tr>
<td>Charlie</td>
<td>C</td>
</tr>
<tr>
<td>Delta</td>
<td>D</td>
</tr>
<tr>
<td>Echo</td>
<td>E</td>
</tr>
<tr>
<td>Foxtrot</td>
<td>F</td>
</tr>
<tr>
<td>Golf</td>
<td>G</td>
</tr>
<tr>
<td>Hotel</td>
<td>H</td>
</tr>
</tbody>
</table>
Appendix E: Script of Instructions

Beginning on the next screen, you will be asked to solve three different puzzles. First, you shuffle the puzzle (the puzzle will be hidden behind a white screen while this occurs). Second, you should listen to the audible solution to the puzzle. The solution consists of a series of alphabetically referenced steps that you will be expected to follow.

The simple goal of a puzzle is to move each alphabetically labeled button to the puzzle's currently vacant cell and according to the sequence prescribed in the audio-delivered solution. A solution will only be presented once for each puzzle. Afterwards, you must recall the solution from memory. Accordingly, you should concentrate and listen very carefully.
Appendix F: Demographic Data Analyses

Table F 1

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>156</td>
<td>84.8</td>
<td>84.8</td>
<td>84.8</td>
</tr>
<tr>
<td>Asian American</td>
<td>1</td>
<td>.5</td>
<td>.5</td>
<td>85.3</td>
</tr>
<tr>
<td>Hispanic American</td>
<td>1</td>
<td>.5</td>
<td>.5</td>
<td>85.9</td>
</tr>
<tr>
<td>Native American</td>
<td>3</td>
<td>1.6</td>
<td>1.6</td>
<td>87.5</td>
</tr>
<tr>
<td>White</td>
<td>19</td>
<td>10.3</td>
<td>10.3</td>
<td>97.8</td>
</tr>
<tr>
<td>Other</td>
<td>4</td>
<td>2.2</td>
<td>2.2</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>184</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>184</td>
<td>100.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table F 2

Gender

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>102</td>
<td>55.4</td>
<td>55.4</td>
<td>55.4</td>
</tr>
<tr>
<td>Male</td>
<td>82</td>
<td>44.6</td>
<td>44.6</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>184</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>184</td>
<td>100.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table F 3

<table>
<thead>
<tr>
<th>GPA</th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 2.00</td>
<td>2</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
</tr>
<tr>
<td>2.00 - 2.49</td>
<td>39</td>
<td>21.2</td>
<td>21.2</td>
<td>22.3</td>
</tr>
<tr>
<td>2.50 - 2.99</td>
<td>75</td>
<td>40.8</td>
<td>40.8</td>
<td>63.0</td>
</tr>
<tr>
<td>3.00 - 3.49</td>
<td>48</td>
<td>26.1</td>
<td>26.1</td>
<td>89.1</td>
</tr>
<tr>
<td>3.50 - 4.00</td>
<td>20</td>
<td>10.9</td>
<td>10.9</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>184</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>184</td>
<td>100.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table F 4

<table>
<thead>
<tr>
<th>Age</th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 18</td>
<td>7</td>
<td>3.8</td>
<td>3.8</td>
<td>3.8</td>
</tr>
<tr>
<td>19 - 22</td>
<td>142</td>
<td>77.2</td>
<td>77.2</td>
<td>81.0</td>
</tr>
<tr>
<td>23 - 28</td>
<td>26</td>
<td>14.1</td>
<td>14.1</td>
<td>95.1</td>
</tr>
<tr>
<td>29 - 34</td>
<td>7</td>
<td>3.8</td>
<td>3.8</td>
<td>98.9</td>
</tr>
<tr>
<td>&gt; 35</td>
<td>2</td>
<td>1.1</td>
<td>1.1</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>184</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>184</td>
<td>100.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valid</td>
<td>Frequency</td>
<td>Percent</td>
<td>Valid Percent</td>
<td>Cumulative Percent</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
<td>---------</td>
<td>---------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Never</td>
<td>23</td>
<td>12.5</td>
<td>12.5</td>
<td>12.5</td>
</tr>
<tr>
<td>Seldomly</td>
<td>71</td>
<td>38.6</td>
<td>38.6</td>
<td>51.1</td>
</tr>
<tr>
<td>Occasionally</td>
<td>71</td>
<td>38.6</td>
<td>38.6</td>
<td>89.7</td>
</tr>
<tr>
<td>Frequently</td>
<td>19</td>
<td>10.3</td>
<td>10.3</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>184</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

Table F 5

Play Computer Games
### Academic Status

<table>
<thead>
<tr>
<th>Academic Status</th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid Freshman</td>
<td>14</td>
<td>7.6</td>
<td>7.6</td>
<td>7.6</td>
</tr>
<tr>
<td>Sophomore</td>
<td>48</td>
<td>26.1</td>
<td>26.1</td>
<td>33.7</td>
</tr>
<tr>
<td>Junior</td>
<td>63</td>
<td>34.2</td>
<td>34.2</td>
<td>67.9</td>
</tr>
<tr>
<td>Senior</td>
<td>59</td>
<td>32.1</td>
<td>32.1</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>184</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>184</td>
<td>100.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table F 7

#### Treatment Groups

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Age</th>
<th>Play Computer Games</th>
<th>GPA</th>
<th>Ethnicity</th>
<th>Gender</th>
<th>Academic Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>125 wpm</td>
<td>Mean 2.1846</td>
<td>2.4000</td>
<td>3.1846</td>
<td>1.8462</td>
<td>1.3538</td>
<td>2.8154</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation .6347</td>
<td>.9152</td>
<td>.8820</td>
<td>1.8220</td>
<td>.4819</td>
<td>.9665</td>
</tr>
<tr>
<td></td>
<td>Minimum &lt; 18 Never 2.00 - 2.49</td>
<td>3.50 - 4.00</td>
<td></td>
<td>African American</td>
<td>Female</td>
<td>Freshman</td>
</tr>
<tr>
<td></td>
<td>Maximum &gt; 35 Frequently 3.50 - 4.00</td>
<td></td>
<td></td>
<td>Other</td>
<td>Male</td>
<td>Senior</td>
</tr>
<tr>
<td>175 wpm</td>
<td>Mean 2.1538</td>
<td>2.5846</td>
<td>3.2462</td>
<td>1.5846</td>
<td>1.4615</td>
<td>3.0000</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation .5371</td>
<td>.7884</td>
<td>.9359</td>
<td>1.7038</td>
<td>.5024</td>
<td>.9186</td>
</tr>
<tr>
<td></td>
<td>Minimum &lt; 18 Never &lt; 2.00</td>
<td></td>
<td></td>
<td>African American</td>
<td>Female</td>
<td>Freshman</td>
</tr>
<tr>
<td></td>
<td>Maximum 29 - 34 Frequently 3.50 - 4.00</td>
<td></td>
<td></td>
<td>Other</td>
<td>Male</td>
<td>Senior</td>
</tr>
<tr>
<td>200 wpm</td>
<td>Mean 2.3148</td>
<td>2.4074</td>
<td>3.3148</td>
<td>1.7037</td>
<td>1.5370</td>
<td>2.9074</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation .6956</td>
<td>.8131</td>
<td>1.0429</td>
<td>1.7225</td>
<td>.5033</td>
<td>.9370</td>
</tr>
<tr>
<td></td>
<td>Minimum &lt; 18 Never &lt; 2.00</td>
<td></td>
<td></td>
<td>African American</td>
<td>Female</td>
<td>Freshman</td>
</tr>
<tr>
<td></td>
<td>Maximum &gt; 35 Frequently 3.50 - 4.00</td>
<td></td>
<td></td>
<td>White</td>
<td>Male</td>
<td>Senior</td>
</tr>
<tr>
<td>Total</td>
<td>Mean 2.2120</td>
<td>2.4674</td>
<td>3.2446</td>
<td>1.7120</td>
<td>1.4457</td>
<td>2.9076</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation .6217</td>
<td>.8422</td>
<td>.9466</td>
<td>1.7458</td>
<td>.4984</td>
<td>.9392</td>
</tr>
<tr>
<td></td>
<td>Minimum &lt; 18 Never &lt; 2.00</td>
<td></td>
<td></td>
<td>African American</td>
<td>Female</td>
<td>Freshman</td>
</tr>
<tr>
<td></td>
<td>Maximum &gt; 35 Frequently 3.50 - 4.00</td>
<td></td>
<td></td>
<td>Other</td>
<td>Male</td>
<td>Senior</td>
</tr>
</tbody>
</table>
Appendix G: Consent Form

Informed Consent for Study Participants

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY

Informed Consent for Participants of Investigative Projects

Title of Project: *The Effects of Speech on Recall in a Multimedia Environment*

Investigator(s): *Marcelite E. Dingle Johnson, Dr. Glen A. Holmes*

**The Purpose of this Research/Project**

The purpose of the experiment is to study how compressed speech influences learning, to investigate if there is educational value of compressed speech in a multimedia environment and to determine if it is practical to incorporate compressed speech in an academic setting. Approximately one hundred students will be involved in this study.

**Procedures**

You will come into the computer lab at any point within a week’s time period and be instructed on how to use the program on a windows-based personal computer. The computerized card you receive will randomly assign you to one of three experimental groups. The computer will request this assigned number and it determines your treatment group. The study will take approximately ten minutes.

**Risks**

There are minimal risks to the student. No personal information will be used against you, and our answers will not be used in any other class or academic situation outside this study.

**Benefits of this Project**

There are no direct benefits to you (other than the extra-credit points discussed in the compensation section), but your participation in this research might help us improve college teach.

**Extent of Anonymity and Confidentiality**

You will remain anonymous throughout the study. You will be randomly assigned an ID number by the computer and the only information we will receive from you will be your gender, ethnicity, educational level, age, grade point average, and major. Neither your name nor social security number will be used as an identifier, therefore once you leave the computer; no method of matching you to the information given is possible.

**Compensation**

Students participating in the study will received a small percentage of extra credit, as agreed upon by your immediate teacher. Students who withdraw from the study can receive extra credit through other means, also as agreed upon by your immediate teacher.

**Freedom to Withdraw**
You are free to withdraw from participation in this study at any time. Just inform the researcher or call one of the others listed at the bottom of this page.

By signing below, you indicate that you have read and understood the informed consent and conditions of this project, that you have all of your questions answered, and that you give your voluntary consent for participation in this project.

If I participate, I may withdraw at any time without penalty. I agree to abide by the rules of this project.

-------------
Signature       Date
-------------

Should I have any questions about this research or its conduct, I may contact:

Marcy Johnson     231-5587     Investigator
Dr. Glen Holmes   231-6387     Faculty Advisor
Tom Hurd          231-5281     Chair, IRB Research Division
Appendix H: Reflections on Conducting Research

1) Formation of the research team
   a) A research team consisting of the Committee Chair, Educational Research Advisor, fellow doctoral student, and the researcher was formed in September 1997. The team established a timeline up to preliminary examinations.
   b) The purpose was to create an efficient approach to designing and conducting dissertation research.
   c) The team started with developing a generalized study consisting of:
      i) Topics, Concerns, Groups
         (1) Strategies
            (a) Redundancy
            (b) Feedback
            (c) Pace
            (d) Timing
         (2) Content
            (a) Science
            (b) Mathematics
            (c) Language
            (d) Engineering
      ii) Age Groups
         (a) Infants
         (b) Adolescents
         (c) Teens
         (d) Young Adults
         (e) Older Adults
      iii) Purpose/use
         (a) Explanation
         (b) Instruction
         (c) Remediation
         (d) Reinforcement
         (e) Simple Presentation
      iv) Media
         (a) Multimedia
         (b) Computer-Based Instruction
         (c) CD ROM
         (d) Audio Tapes
         (e) Live presentations
         (f) Web-based
v) Ethnic/racial
   (a) African American
   (b) Native American
   (c) Pacific Islander
   (d) Asian American
   (e) White
   (f) Hispanic American
vi) Learning theories
   (a) Social
   (b) Behavior
   (c) Cognitive
   (d) Motivational
vii) Gender
    (a) Male
    (b) Female
viii) Methodology
    (a) Lecture
    (b) Programmed instruction
    (c) Workbook
    (d) Demonstration
    (e) Inquiry
ix) Channel
    (a) Visual
    (b) Auditory
    (c) Tactile
x) Dimensionality
   (a) Simple
   (b) Multiple
   (c) Color Depth
xi) Cognitive styles
    (a) Linearity
    (b) Circular
    (c) Non-linear
    (d) Field Dependent
    (e) Field Independent

2) The team selected two to three topics from topics, concerns, and groups in order to narrow the scope of the research design. This provided the framework for conducting the literature review.
3) Each doctoral student defined a set of research questions after searching the literature.
4) The chair and research advisor provided the following overall guidance.
   i) Know the statistical basics
   ii) Define your research skills
   iii) Exhibit professional behavior
5) Not only was this framework helpful, it created an efficient environment for this researcher to successfully complete and defend this dissertation. This approach is highly recommended for consideration by future doctoral students.
Curriculum Vitae

MARCELITE DINGLE JOHNSON

3206 Greenhollow Drive
Greensboro, North Carolina 27410
marcyj@ncat.edu

EDUCATION:

**Doctor of Philosophy**, Instructional Systems Development, April 1998. Virginia Polytechnic Institute and State University, Blacksburg, VA

**Graduate Systems Management Program**, University of Southern California, Stuttgart, West Germany, 1985

**Master of Science (Education)**, Indiana University, Bloomington, Indiana, 1979

**Bachelor of Arts (Mathematics)**, University of Colorado, Denver, Colorado, 1972

EXPERIENCE:

06/97-present  **Senior Technical Consultant**, Academic Affairs, North Carolina A&T State University, Greensboro, North Carolina

08/96 – 05/98  **Graduate Assistantship**, Education Technology Lab, Teaching and Learning Department, Virginia Tech, Blacksburg, Virginia

06/92  - 05/97  **Manager**, Academic Computing Department, Computer Center, North Carolina A&T State University, Greensboro, North Carolina

08/92  - 05/95  **Adjunct Assistant Professor**, College of Engineering, North Carolina A&T State University, Greensboro, North Carolina

10/91  - 06/92  **Assistant to the Dean/Lecturer**, College of Arts and Sciences, North Carolina A&T State University, Greensboro, North Carolina

01/91  - 12/91  **Lecturer**, Computer Science Courses, University of North Carolina at Greensboro, Greensboro, North Carolina

06/88  - 08/90  **EAME Training Director**, Wang Europe, Africa and Middle East (EAME)
Headquarters, Brussels, Belgium.

11/85 - Business Services Manager, Wang Laboratories, Inc.,
06/88 Federal Systems Division, Stuttgart, West Germany

10/83 - Computer Systems Manager, Greater Stuttgart
11/85 Military, Community Civilian
          Personnel Office, Department of the Army, Stuttgart
          West Germany.

09/83 - Instructor, City Colleges of Chicago, Stuttgart, West
10/85 Germany

09/81 - Systems Manager, Penrose Cancer Hospital, Colorado
08/83 Springs, Colorado.

09/82 - Systems Design and Acquisition Consultant, Norfolk
08/83 State University, Norfolk, Virginia

11/79 - Software Engineer, Hewlett-Packard, Colorado
08/81 Springs, Colorado.

05/80 - Honorarium Instructor, University of Colorado,
08/80 Colorado Springs, Colorado.

09/76 - Senior Technical Support Programmer, Indiana
07/79 University, Bloomington, Indiana.

04/74 - Software Engineer, Ford Aerospace Communications,
07/76 Colorado Springs, Colorado.

05/72 - Programmer, Delco Electronics, Kokomo, Indiana.
05/74

RESEARCH GRANT: North Carolina A&T State University Multimedia Laboratory,
Sponsor: AT&T ($125,000) October 1994

PROFESSIONAL SERVICES:
Member, Distance Learning Task Force, NC A&T SU
Search Committee For Vice Chancellor Of Academic Affairs, NC
A&T SU
Chair, Academic Computing Committee
Board Of Directors, Tylane, Inc., Kokomo, Indiana

PROFESSIONAL AFFILIATIONS:
Association for Educational Communications and Technology
International Visual Literacy Association
Eastern Educational Research Association
American Educational Research Association
Association of Computing Machinery


HONORS: PHI KAPPA PHI
ALPHA KAPPA MU
WORLD WHO’S WHO OF WOMEN (1987)
WHO’S WHO OF AMERICAN WOMEN (1988)
WHO’S WHO OF EMERGING LEADERS IN AMERICA (1988)
WHO’S WHO IN FINANCE AND INDUSTRY (1988)