2005

Assistive technologies based on synthetic speech : interface compatibility

Jose R. Concepcion
San Jose State University

ASSISTIVE TECHNOLOGIES BASED ON SYNTHETIC SPEECH: INTERFACE COMPATIBILITY

A Thesis

Presented to

The Faculty of the Department of Industrial and Systems Engineering

San Jose State University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science in Human Factors and Ergonomics

by

Jose R. Concepcion

May 2005

APPROVED FOR THE HUMAN FACTORS AND ERGONOMICS PROGRAM

Kevin Corker, Ph. D.

Eric Christierson

Kevin Jordan, Ph. D.

APPROVED FOR THE UNIVERSITY

ABSTRACT

ASSISTIVE TECHNOLOGIES BASED ON SYNTHETIC SPEECH: INTERFACE COMPATIBILITY

by Jose R. Concepcion

A growing debate about functional gaps between graphical user interfaces and screen access applications has surfaced. Assumptions made about the topic point at limits in *synthetic speech intelligibility* and deficiencies at the GUIs' *levels of interpretation.* In order to determine the proportions in which both constructs were affecting human-computer interaction, the author evaluated different aspects of performance while using JAWS® V 5.0 Screen Reader on Microsoft® Windows XP. Measurements included perception of verbal material while controlling *acoustic* and *contextual features* as well as *task efficiency.* Data collected in the form of words recalled, comprehension test scores, and time for tasks such as searching, editing, and transferring information were used to infer about these topics. Nine blind subjects *(n=9)* participated in the experiment. The author did not find a significant effect while controlling *synthetic speech rate* and *noise* as contextual factor. However, a significant effect on task dependence was found within the range tested.

ACKNOWLEDGEMENTS

The preparation this thesis project required the support and encouragement of many people. Thanks to the thesis committee for their outstanding contributions and professionalism. I addition, I feel extremely grateful to those who, on behalf of the California School for The Blind and The Peninsula Center for The Blind and The Visually Impaired, valued the contributions made through this research paper.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

People with visual and motor impairments are often disengaged from daily-life activities, with computing being one of the most critical. Throughout the years, rehabilitation engineers have devised artificial means to minimize the effect of the "digital divide," thus returning some productivity and autonomy to users. The connection to any form of Assistive Technologies has given the disabled a chance to overcome obstacles that would otherwise impede their full integration and productivity.

In general, Assistive Technologies can be described as adaptive resources designed to minimize the effects of motor and cognitive impairments (Cook & Hussey, 2002). From an Ergonomics and Human Factors standpoint, assistive technologies applications should increase, maintain, or improve task efficiency. In this study, an attempt was made to validate this assumption through observations of performance while using screen readers.

Screen-access applications are a good example of adaptive functionality, currently used to help the blind read from the computer screen. Although the blind use such applications as speech-based interfaces, they are forced to use the same interaction techniques as on picture-based interfaces. As argued by subject-matter experts, "disparity" between the speech-based context and visually oriented interfaces can affect performance in the ability to navigate, error rate, use of idle time, and orientation (Cook, 1997; Stephanidis, et al., 1997; Lazarro, 2001; Barraga, & Erin, 2001; Edwards, Mynatt, & Stockton, 1995).

A lack of experimental evidence in this matter motivated the author to observe some of the factors affecting interaction with speech-based interfaces. One of the aspects considered was the importance of synthetic speech intelligibility on perception (Venkatagiri, 1994; Miranda and Beukelman, 1987; Pratt, 1987; Hjelmquist, Dahlstrand, & Hedelin, 1992). Therefore, the author decided to manipulate constructs such as *speech rate*, *listening conditions,* and *tasks* to observe users' potential to recall words, facts, and procedures. Likewise, interaction was analyzed based on performance outcomes while completing three different tasks, trying to find quantitative and qualitative data to support the notion of "context disparity."

Assumptions

Some of the assumptions found in this thesis could be used as a theoretical basis for further investigations. However, the author decided to limit his hypotheses to the constructs previously cited, that is, *synthetic speech perception* and *task efficiency* while using screen readers over graphical user interfaces.

More specifically, one can hypothesize that *rate*, as an acoustic feature of synthetic speech, has no impact on the recall of verbal material as measured by the *number of words correctly recalled* and *errors*. Within the range tested, this could be the first indicator of "good intelligibility," meaning that in between a speech rate of 40 and 80 words per minute, the quantity and quality of word recall were the same.

Similarly, the author hypothesized that changes in the *listening conditions* had no impact on the comprehension of verbal material. In other words, blind individuals are able to retain verbal information whether in the presence or absence of *noise*, even when noise has the same magnitude as the meaningful source.

On the other hand, the author's third and last hypothesis points at a significant effect on task dependence, that is, differences on *time* while completing different tasks including searching, editing, or transferring information in the electronic media.

Assuming that JAWS® V 5.0 Screen Reader's prosodic attributes have reached an almost-perfect level of intelligibility (Venkatagiri, 1994; Miranda and Beukelman, 1987; Pratt, 1987; Hjelmquist, Dahlstrand, & Hedelin, 1992), is a matter of finding specific evidence to defend that flaws in performance could be attributed to the poor integration of the platform's level of interpretation.

Delimitations

Through this research project, the author measured aspects such as synthetic speech perception, intelligibility, contextual factors, and task efficiency while using screen-access applications. It is important to mention that the author did not intend to measure or discuss cognitive aspects of performance such as processing, retrieval, retention, memory workload, and the like. Similarly, the author did not aim to measure any effect of acoustic features such as pitch, volume, punctuation, or synthetic speech personification.

Definition of Terms

a) Adaptivity: Interface characteristic that refers to the process of selecting aspects of the user interface in a dynamic fashion, according to situations that are detected while in use. Therefore, an "adaptive interface" must be able to accommodate different interaction patterns and respond to emerging user requirements (Stephanidis, et al, 1997).

b) Unified Interface: Interface with the potential to tailor itself either for the individual user, interaction, or any particular situation of use (Stephanidis, et al, 1997).

c) Lexical level of interpretation: A collection of lines, dots, and text on a computer screen (Edwards, Mynatt, Stockton; 1995).

d) Syntactic level of interpretation: Groups of elements at the lexical level of interpretation converted into buttons, text-entry fields, scrollbars, and so forth (Edwards, Mynatt, Stockton; 1995).

e) Semantic level of interpretation: Describes an interface in terms of the operations to be performed to generate a desired output (Edwards, Mynatt, Stockton; 1995).

II. LITERATURE REVIEW

Characterizations and Assessment of Assistive Technologies

Assistive Technologies are an important resource for task aiding, especially for the visual, hearing, and motor impaired. Applications of adaptive technologies include a broad range of devices and information systems conceived to minimize the effect of physical disabilities (Cook & Hussey, 2002). In general, any device or *assistive* mechanism must increase or at least maintain human functional capabilities.

Among the existing *assistive technology characterizations* (e. g., low and high technologies, assistive and rehabilitative technologies, minimal and maximal technologies, tools, appliances, etc.) *hard* and *soft technologies* is one of the segments that deserves more applied research. Hard technologies are those readily available components that can be purchased or assembled into assistive systems (Odor, 1984). On the other hand, *soft technologies* can be viewed as any application that allows impaired individuals to interact with desktop technologies (e.g. browsers, word processors, spreadsheets, and the etc).

Creating "assistive" applications demands a human-centered approach at both the design and assessment phases. Similarly, researchers and developers may understand the importance of the human-environment relationship and how it can be studied in order to sustain users' functional capabilities.

For instance, the *Human Activity Assistive Technology Model* or HAAT provides a solid base suitable for development and assessment of adaptive resources (Cook & Hussey, 2002). The HAAT model considers interaction with assistive technologies through four elements in a close loop, namely *humans, activities, applications,* and *context*. In this model, the *activity* component is defined by the overall goal of the assistive technology system. On the other hand, *humans* represent the individual with a disability who is operating the system through any *intrinsic enabler*, that is, any sensory input mechanism, central processing system, or effector (Christiansen & Baum, 1997).

Context, at any one of the three levels of environmental interaction, is represented by all contextual factors affecting the system either at a *mesosystemic level* (directly), *microsystemic level* (immediately), or at *macrosystemic level* (indirectly). Any assessment model may consider the use of assistive technologies at these different levels.

A complete and comprehensive assessment model may explore each one of the previous factors to discover underlying flaws in-between and within platforms, always considering the HAAT model as the basis of the analysis. Part of the author's research effort will be geared towards the use of some the HAAT elements for the evaluation of screen-access applications.

Screen-Access Applications

Synthetic Speech is, without a doubt, one of the most used forms of assistive soft-technologies. Common applications include accessing interfaces of word processors, databases, spreadsheets, and web browsers. Among other important advantages, the low cost and rapid implementation have made them an indispensable resource for the blind.

Basically, speech-synthesis engines are used in assistive applications to translate electronic information into human-sounding speech. Conversion is based on the same components of human speech, that is, text into phonemes. Normally, means for synthetic voice output are provided via hardware such as soundcards, microprocessors, digital signal processors, and digital-to-analogue converters.

Technically, a *screen reader* uses a speech synthesizer and mathematical algorithms to provide real time verbalization of every text portion displayed on the screen. Screen readers drive the computer's soundcard to speak aloud letters and words. With most of the screen readers available on the market, users with any kind of visual impairment are able to "read the screen," search for icons, text, and send commands directly to the running application.

As a way to improve perception of synthetic speech, *acoustic features* such as volume, rate, pitch, and punctuation level can be controlled directly from the application. General commands were designed to control voice output, that is, to speak *characters*, *words*, *lines*, and even whole *text blocks*. In addition, speech synthesis applications can speak current window titles, speak the top and bottom line of windows, speak the status line, speak highlighted text, speak the default button of dialog box, speak tab controls of dialog box, search text strings, and speak cursor coordinates.

However, some affordable systems are not perfectly intelligible despite their capacity to pronounce an ample vocabulary (Venkatagiri, 1994). This is because of two primary factors: the lack of human *intonation* and the *robotic* sound of synthetic speech. Nonetheless, top applications have reproduced almost perfectly the acoustic features at the voice output, giving it a little or no difference in intelligibility (Venkatagiri, 1994; Miranda and Beukelman, 1987; Pratt, 1987; Hjelmquist, Dahlstrand, & Hedelin, 1992).

Mainstream Limitations

Contextual differences between graphical user interfaces and assistive applications could be the origin of many interaction flaws. In addition, the fast evolution of technology has impeded the expansion of adapted-based solutions (Stephanidis, Paramythis, Savidis, Sfyrakis, Stergiou, Leventis, Maou, Paparoulis, & Karagiannidis, 1997).

Certainly, new applications in the field have been conceived following the needs of "average" users groups, limiting people with disabilities and physical impairments from accessing such applications (Stephanidis, Savidis & Akoumianakis, 1997).

From the author's standpoint, the greatest limitation for the application of adaptive solutions has been the *graphical user interface,* a computer display modality that is part of most standard software applications. Since navigation within GUIs requires the use of vision and hand-eye coordination, people with visual impairments could be affected.

In fact, new developments on visual interfaces have stalled the progress made by screen-reading technology (Edwards, Mynatt, & Stockton, 1995; Boyd, Boyd, & Vanderheiden, 1990). The dominance of *picture-based interfaces* based on *direct manipulation* and hand-eye coordination has limited the effectiveness of *speech-based systems* (Christian, 2000).

Another limiting aspect to be considered is the interface's *level of interpretation* (Edwards, Mynatt, & Stockton, 1995). Truly adaptive interfaces tend to concentrate more on the *semantic levels* rather than the *lexical* and *syntactic levels*. Transforming an interface into abstract terms is also a compelling approach in creating usable and efficient assistive solutions (Edwards, Mynatt, & Stockton, 1995).

Nonetheless, one of the solutions applied to minimize the impact of platform differences are the so-called *unified interfaces* (Stephanidis, *et al.* 1997). Interfaces qualified as *unified* are capable of supporting special goals through accessible features, alternative access systems, interaction techniques, and by redefining information content and structure. *Unified interfaces* utilize added functionality that can realize specific patterns on the basis of externally acquired knowledge (or user oriented knowledge) and criteria such as simplicity, error tolerance, and error recovery.

Similarly, user-oriented adaptivity is another resource to reduce cross-platform limitations. Providing *user-oriented adaptivity* implies conjoining five elements as part of the design platform: *sensory abilities* (e.g., visual functions, visual field deficits, range, auditory functions, etc.), *interests, preferences, domain knowledge,* and *competence* in computer-based applications (Fink, Kobsa, & Nill, 1997).

<u>Factors Limiting Information Acquisition</u>

User productivity and efficiency regarding assistive technologies are not only limited by the use of an alternative perceptual channel, but also by the capacity to sustain information processing under specific circumstances. As a consequence of contextual differences and limited channel capacity, performance could be affected in aspects such as reasoning, retrieval, and classification of information (Van Reusen & Head, 1994).

In terms of *learning tactics* and *knowledge representations,* sightless individuals are at a relative disadvantage in reading text due to the lack of visual cues from the information source. This could have an impact on the development of metaphors, which are the basis of interpretation on graphical user interfaces. As a result of such a visual barrier, the blind are also at a disadvantage while attempting to locate items of interest and scan selected parts on the screen in a timely manner.

Likewise, blind individuals are at a disadvantage with regards to *reading performance* and the *interpretation* of picture-based interfaces. As pointed out, blind individuals rely mostly on the re-reading of text, the use of different tactics to access information, and the elaboration of different representations as basis of interpretation. In addition, the ability to identify phonological characteristics among words represents another factor of interest in creating interfaces based on synthetic speech (Gillon & Young, 2002).

Visually impaired individuals are also disengaged from the immediate environment in some extent (Rusalem, 2001). Prolonged disengagement tends to limit both social and cognitive development. This could reduce the individuals' readiness for interaction and even limiting the possibilities for vocational rehabilitation.

Finally, emotional aspects of blindness can affect interaction and information acquisition. Behavioral manifestations such as fear, anxiety, frustration, helplessness, desperation, and withdrawal can influence performance (Rusalem, 2001). Although blind individuals face the same emotional traits as sighted people, coping with the unseen environment triggers new sensations capable of reducing emotional strength and stability (Wagner, 1961).

Perception and Cognition On Speech-based Interfaces

Factors such as *sensory abilities* and *memory capacity* have an impact on interaction with speech-based interfaces. It is commonly accepted that cognition involves the use of senses because the mind cannot process, store, or retrieve anything that has not been perceived. Sightless individuals follow a similar scheme, relying instead on the residual vision or on any alternative channel such as hearing or touch.

In respect to *hearing*, spoken information has almost immediate access to verbal working memory, even with the same input quality as the visual channel. However, *discriminability* at the auditory channel is sometimes very limited, especially when performing a *dichotic listening task* (Wickens, 1984; Broadbent, 1958; Cherry, 1953; Treisman, 1969; Proctor & Dutta, 1995). The abilities to note differences or similarities between two or more audible stimuli while performing such kinds of tasks are reduced by the *masking effect* phenomena (Barraga & Erin, 2001). Nevertheless, *sound discrimination* can be physically improved based on variations of *frequency, intensity, rhythm,* and *duration* of the auditory signal.

In terms of *attention*, the auditory channel is also different from the visual in many aspects. The auditory sense can take input from any direction as soon as it surpasses sound thresholds. However, since most auditory input is transient, the *pre-attentive characteristics* of auditory processing are more critical in audition than in vision (Wickens, & Hollands, 2000). Even in the absence of *consciousness*, information in the unattended channel can make contact with *long-term memory.* Moreover, one can attend selectively to auditory messages coming from similar locations. As mentioned before, the greater the differences between two messages along a given dimension, the easier it will be to focus on one source by reducing attention to the other (Ward, 1994).

Not only the blind can take advantage from selective attention, but also from a poli-sensory capacity by integrating alternative perceptual channels. Nevertheless, the auditory is the only one that can interpret information without the need of pattern recognition (as happens with Braille reading). In fact, the auditory input is processed more efficiently due to its autonomy as channel. According to Liberman and Mattingly (1985), such channel autonomy is based on the existence of a distinct brain module, which is responsible for the perception of phonetic information.

It could be valuable to investigate whether all previous phenomena occur under interfaces based on synthetic speech. For instance, failures in the process of perceiving auditory information could be attributed to less-than-perfect audio transmissions. A*coustic features* have been demonstrated to be key elements on speech perception (Ainsworth, 1972). Pitch, duration, and amplitude as *prosodic features* (Miller, 1981), and *language elements* could also affect speech perception (Cole & Jakimik, 1978; Marslen-Wilson & Welsh, 1978; Massaro, 1987; Cook & Hussey, 2000).

Given that voice output can be manipulated in most screen readers, it is important in experimental matters to observe if blind individuals depend continuously on such functionality to improve perception and how effective this resource is in helping them to form accurate mental images from spoken language.

Advantages and Disadvantages of Multimodal Perception

The use of an alternative sensory channel, such as touch, is another way to improve information acquisition (Cook & Hussey, 2002). Haptic processing may offer new advantages that can be utilized to reinforce perception if coupled with the auditory input provided by *speech user interfaces*. Through the tactile channel, the blind take advantage of specific haptic information such as shape, hardness, texture, and size (Klatzky, Lederman, & Reed, 1988).

In fact, cross-modal time-sharing could be more advantageous than the intra-modal type (Wickens & Holland, 2000). The relative advantage of cross-modal time-sharing over the intra-modal kind is not only caused by the use of separate perceptual resources, but by peripheral factors that place intra-modal conditions at a disadvantage (e.g., *confusion* and *masking*).

Functional differences have a crucial impact on interaction. For instance, it has been demonstrated that visual reading is more efficient than tactile. The average reading speed of an expert Braille reader is 95-115 wpm if compared to 250-300 wpm for sighted readers (Rosa, Huertas, & Simon, 1994). The main reason for such differences can be found at the beginning of the reading process, specifically on the amount of information that can be processed through the sensory channels.

Individuals using Braille as their reading media employ the same transformation mechanisms to generate "phonological" representations (Pring, 1982). As indicated by Kirman (1973), Braille readers must identify and remember all the letters in one word and then integrate them to represent a meaningful unit. In addition, reading Braille could demand sustained focused attention, implying the use more cognitive resources for information processing.

The serial nature of Braille characters, the absence of *fixation* over a particular stimulus, timing of the input, and differences in lexical processing could affect comprehension (Carreiras & Alvarez, 1999; Mousty & Bertelson, 1992). Indeed, Braille reading has been recognized as the highest level of tactual-kinesthetic development due to the complexity of pattern discrimination. This means that a more abstract level of *perceptual-cognitive association*, *tactual imagery*, and *haptic interpretation of context* could be required for information acquisition (Barraga & Erin, 2001).

Attention, Workload, and Intra-Modal Time-Sharing

Divided attention is a common cognitive phenomenon found while interacting with electronic information systems. Consequently, the same condition is likely to occur while depending on speech synthesis applications for information acquisition. Because demanding environments often impede narrowing attention on a specific source, *divided attention* is perhaps the only strategy available to sustain effective performance.

However, the author doubts that by using the same perceptual channel it may be possible to talk about *perfect parallel processing*. As a result, *dual task interference* and *effortful performance* could be probable while interacting with screen readers in resource-demanding contexts.

One of the most important consequences of picture-based interfaces interpretation through parallel processing is *mental workload* (Wickens & Holland, 2000). Cognitive demands while interacting with graphical user interfaces are fulfilled using limited resources in the case of the blind. Based on such a premise, memory burden is likely to occur if performing different tasks at the same time, thus resulting in loss of momentum, hesitation, and even task abandonment.

Performance Evaluation

To evaluate interaction with screen-access applications, one could manipulate a variety of constructs such as perception, cognition, workload, transparent access, context, functionality, and adaptivity. Each one of these elements can be analyzed using both qualitative and quantitative assessment models (Cook & Hussey, 2002).

Assessment of assistive applications allows the researcher to answer specific questions about performance. In addition, applied research designed to measure the effectiveness of assistive applications must be representative of the situations, tasks, and stimuli commonly found in an everyday environment.

Assessing human performance on adaptive systems also implies considering the individuals' cognitive abilities as well as functional skills. Indeed, disregarding these components of performance will leave the assessment equation incomplete since it is not quite clear if *loss of information* was caused by the poor quality of the interface or by any of the individual's limitations.

Cognitive assessment in the field of assistive technology could be viewed differently if compared with the conventional evaluation models. Although there are standardized test batteries for primary assessment of cognitive skills (Cook & Hussey, 2002; Duchek, 1991), they are inapplicable, too complex, or too case specific for field observations.

Because of the characteristics of naturalistic settings, test batteries are not the most suitable tools to evaluate users' cognitive skills and interaction with assistive technologies. However, test batteries that measure *preferences* and *satisfaction* in a multidimensional and qualitative way could be of interest (Demers, Weiss-Lambrou, and Ska, 1996).

In addition, researchers must consider other aspects such as language, system knowledge, speed, systems' usability, accuracy, error tolerance, and preferences regarding specific aspects of the screen access application (Stephanidis, *et al.* 1997). Likewise, familiarity with specific tasks, the ability to navigate, and common strategies are important aspects to be considered in the assessment model.

Finally, when considering performance based on changes in the environment, the researcher must identify any source that could affect interaction directly. For instance, in moderately noisy environments, speech synthesis tends to be obscured or masked (Osborn, Erber, & Galletti, 2000). While sighted individuals can solve this setback by relying more on other visible cues to compensate (Sumby & Pollack, 1954), sightless individuals are more susceptible to fail because it is difficult for them to compensate through other stimuli. Thus, listening conditions are another key aspect that may be observed on speech-based interaction.

III. METHODS SECTION

Participants

To conduct the experiment, a group of 9 legally blind participants were recruited within the California's South Bay area. The recruitment process was carried out through specialized training centers for visually handicapped people. The *Peninsula Center for the Blind and Visually Impaired* (Palo Alto, CA) and *The California School for the Blind* (Fremont, CA) were the selected private institutions for recruitment purposes. Both institutions offer comprehensive educational services to the visually impaired community.

Basic information about impairment origin was processed following the Standard Classification of Causes of Blindness (Rusalem, 2001), which refers to specific aspects such as site, type of agent, and underlying causes of the impairment (See Appendix A). Three groups of causes were considered, namely *conditions of unknown cause* (e.g., glaucoma, senile cataracts, myopia), *conditions associated with general diseases* (e.g., diabetes, vascular problems), and *prenatal causes.*

In addition, knowledge of Microsoft Windows XP® operating system was considered within the selection rules. Familiarity with JAWS® V 5.0 screen reader was an obligatory requirement to participate in the experiment (at least six months experience). All participants were required to have English as their first language and to report normal hearing abilities (Sharit, Czaja, Nair, & Lee, 1999). Access to the experiment was granted after signing an informed consent document (see Appendix B).

Apparatus

Devices and methods used in the experiment have been used in the assessment of similar assistive technologies. The main experimental instrument included a standard screen reader, in this case JAWS® V 5.0 (by Freedom Scientific, Blind/Low Vision Group; St. Petersburg, FL) and Windows XP® Home Edition as graphical user interface (by Microsoft Corporation, Redmond, WA). JAWS® V 5.0 Screen Reader uses synthetic voice output to emulate human voice through the computer's soundcard. Once the screen reader is running, highlighted information on the computer's screen is "read aloud" by the application allowing visually impaired users to get access to computer-based information.

Hardware selected to conduct the experiment includes a Dell Inspiron Laptop 8200 (by Dell Computer Corporation, Austin, TX), built on an Intel® Pentium® 4M CPU 1.7GHz processor, and a Crystal WDM Audio Codec soundcard (by Cirrus Logic, Inc. Austin, TX). Alternate and primary audio was generated through a Dell Inspiron Laptop 8200 built-in sound card and transmitted through a set of Labtec® Elite™ 820 Headphones (by Labtec Inc., Fremont, CA).

Windows Media Player® V 8.0. (by Microsoft Corporation, Redmond, WA) was used to convey the noise source. Sound samples were provided by recordings on the compact disc "Sounds of Machines, Movement, and Transportation" (by Nesak® International; East Hanover, NJ). In addition, a Panasonic RN3053 Micro Cassette Recorder was used to collect information from the participants at some of the experimental sessions (like those requiring recall of verbal material).

Materials

Materials to be used in the experiment were classified into two groups. The first group was considered as *pre-interaction inquiries,* aimed to gather basic information from the participant. This section covered issues such as the screening form used to obtain demographic and clinical data from each participant (see Appendix A) and the consent form (see Appendix B). Every *subject profile* as obtained from the screening document included information about platform proficiency, sensory and perceptual abilities (including vision and hearing), social interaction, learning, and behavior (see Appendix A).

A second group or *post-interaction inquiries* materials were used to gather both qualitative and quantitative data about dependent variables, tasks, instructions, and other key performance indicators that must be collected throughout the experimental process (see Appendix C).

In addition, auditory reading passages and comprehension tests were used as indicated in the experimental protocols (see Appendix C). Each of the selected passages contained about 1000 words, organized into 12 to 15 short paragraphs to ease reading (see Appendix D). Content of each passage was evaluated to confirm *narrative* characteristics (Carreiras & Alvarez, 1999; Hensil, & Whittaker, 2000), that is, suitable for conveying information about episodes, outcomes, activities, and events related to a given character or event.

On the other hand, "auditory reading comprehension" will be measured through *oral quizzes* (or comprehension tests) comprised by a set of ten multiple-choice questions concerning the content of each of three reading passages (see Appendix E). Questions were designed based on content memorization, that is, by measuring the probability of recalling words, facts, ideas, or thoughts under the assigned experimental condition (Wetzel & Knowlton, 2000; Carver, 1981).

Finally, each questionnaire was developed on the basis of textual analysis of the article, which consisted of breaking down each sentence into the number of thoughts or facts contained and then turning them around in the form of questions (Wetzel & Knowlton, 2000).

Independent Variables

In order to collect experimental evidence to validate the author's hypotheses, synthetic speech *rate* and *listening conditions* were manipulated to observe users' potential to recall words and facts while depending on JAWS® V 5.0 synthetic speech output as an interface. Hence, the first independent variable addressed perception of *synthetic speech* while controlling its *rate* on three different levels, namely, 40, 60, and 80 words per minute.

On the other hand, a second independent variable was included to measure "auditory reading comprehension" under two different *listening conditions*: *ambient noise present* and *ambient noise absent.*

Finally, a third independent variable was manipulated to measure *task efficiency* while depending on synthetic speech output while interacting with graphical user interfaces. At this point, interaction was analyzed based on performance outcomes while completing three different tasks: *editing, searching,* and *transferring* of data in the electronic media. This variable was introduced trying to find quantitative and qualitative evidence to support the notion of "context disparity" (based on navigation problems, error rate, and loss of momentum).

Dependent Variables

Four dependent variables were used to measure any plausible effect on perception and interaction. With regards to the first independent variable (synthetic speech rate), the *total of words correctly recalled* and *errors* (or number of words not associated with the passage content) were used as metrics for the intentional recall of verbal material under three different speech rate levels.

On the other hand, "auditory reading comprehension" as measured by changes in the listening conditions was measured through *comprehension test scores* (Carver, 1981). Such parameters were used as key performance indicators to determine if under such experimental conditions (ambient noise present or absent) participants were able to recall at least 75% of the passage content based on a multiple-choice test score (see Appendix D). Finally, to measure task efficiency while using JAWS® V 5.0 Screen Reader on the selected graphical user interface, *time-on-task* was considered as a dependent variable.

Procedures

The evaluation process started once the subject agreed to participate in the experiment by signing the *consent form* (see Appendix A). Then each participant was asked to sit in front of the Dell Inspiron Laptop 8200, which was placed at an approximate height of 30" and a horizontal distance of about 28" maximum.

Since participants were required to understand basic keystrokes commands, each one was trained for approximately about 20 minutes in the use of a lap top keyboard to recognize general functions. Basically, they were practicing "keystroke navigation" using the Dell Inspiron Laptop 8200 keyboard, identifying keys such as TAB, ENTER, ESC, BACKSPACE, and ARROW KEYS, all used in substitution of the graphical input device.

JAWS® V 5.0 users were required to demonstrate knowledge of keystrokes combinations such as Help (F1), Open the Start Menu (WINDOWS LOGO KEY or CTRL+ESC), Open Windows Explorer (WINDOWS LOGO KEY+E), Open Run Dialog (WINDOWS LOGO KEY+R), Find File or Folder From Desktop (F3), Minimize All Applications (WINDOWS LOGO KEY+M), Switch Between Applications (ALT+TAB), Quit the Active Application (ALT+F4), and the like.

Once the hardware was tested and site requirements confirmed (e.g., posture, distance, comfort), the researcher instructed the participant about the tasks to be completed. Each session was conducted as follows:

Experimental Session I: To measure human potential to recall words and facts while depending on synthetic speech output, participants were instructed to *intentionally recall* as many words as possible from an experimental passage read aloud using JAWS® V 5.0 screen reader. At this experimental session, *synthetic speech rate* was increased by 20 words per minute. More specifically, at session I-A, rates were set at 40 wpm; for session I-B to 60 words per minute, and for session I-C to 80 words per minute respectively. For each recall task, three reading passage were used randomly (see Appendix D).

For each one of the three sub-sessions (either I-A, I-B, or I-C), the researcher used Microsoft Word® V 9.0 to access the passage. Content on every passage was highlighted to allow the screen access application to scan and read. Participants listened to the material using a set of Labtec® Elite™ 820 Headphones. At the end of each sub-session, subjects had three minutes to dictate to the researcher all recalled words. Both *total of words correctly recalled* and *errors* or *number of words not associated with the passage content* were registered on audiotapes and at the data collection form (see Appendix F).

Experimental Session II: The second session was intended to measure "auditory reading comprehension" under changes in the *listening conditions* (Osborn, Erber, Galletti, 2000; Wickens & Hollands, 2000; Posner & Boies, 1971). Ten minutes after finishing the first session, subjects returned to the evaluation site to complete the new exercise. In order to assess the *effects of environmental noise* (or ambient noise), subjects were requested to answer a multiple-choice test after listening the auditory reading passage (see Appendix E).

Basically, this session consisted of two parts. In the first one (II-A), there was no *background noise* while reading the passage. Conversely, in the second sub-session (II-B), a noise source was played back at the same time the passage was read aloud (using Windows Media Player V 8.0 and JAWS® V 5.0 screen reader at the same time). Environmental noise or "street noise" was continuously replayed until task completion using the Labtec® Elite™ 820 Headphones. Noise samples were provided from the

"Sounds of Machines, Movement, and Transportation" compact disc (by Nesak® International, East Hanover, NJ).

At the end of each sub-session, participants completed an oral multiple-choice test for every passage read, which reflect to some extent the *auditory comprehension level* based on the number of questions correctly answered.

Experimental session III: The third and final session was intended to measure *task efficiency* while depending on synthetic speech output to interact with graphical user interfaces.

The first sub-session (III-A) consisted of observing participants completing an "editing" task which consisted in changing specific audio settings on the Microsoft Windows XP® operating system. Participants were indicated to "adjust the volume of the soundcard" (see Appendix I). Voice output from the screen access application was used as feedback source, that is, every command or key pressed to navigate was read aloud by the screen-access application to orient the subject. In general, successful task completion can be achieved by using a common menu path, that is, Windows Menu + All Programs + Control Panel + Sound and Audio Devices + Volume + Device Volume + High. *Time-on-task* was measured from the first keystroke until task completion.

On the second sub-session (III-B), subjects were requested to "transfer" a specific file *from* and *to* different locations within the Dell Inspiron Laptop 8200 hard drive (see Appendix I). The file "Test Document. txt" was saved at the location C:\Documents and Settings\ISE299-Participants\Desktop. Subjects were requested to "move" or "copy" the current file to a second location, that is, C:\Documents and Settings\ISE299-

Participants\My Documents\My Computer\Destination Folder. At this point, successful task completion can be achieved by using basic keystrokes such as" Control + C" to copy the file, browsing through the folders using Arrow Keys, and "Control + V" to paste the file on the desired location. *Time-on-task* was used as metrics.

Finally, a third sub-session (III-C) used a "searching" or "browsing" task. In this case, subjects were requested to "search" a specific file within the Dell Inspiron Laptop 8200 hard drive (see Appendix I). Participants were requested to "search" and "browse" the system's folders structure to find the "Target File.txt."

To do so, participants should locate the file by using *arrow keys* and *Tab key*. Starting from the *Desktop* (C:\Documents and Settings\ISE299 Participants\Desktop\Target Folder 5), participants were oriented to browse from a first folder (Target Folder 5) to the destination folder (Target Folder 1) where the "Target File.txt" was located. In general, successful task completion can be achieved by using a common path like C:\Documents and Settings\ISE299-Participants\Desktop\Target Folder 5, C:\Documents and Settings\ISE299-Participants\Desktop\Target Folder 4, C:\Documents and Settings\ISE299-Participants\Desktop\Target Folder 3, and so on until reach the final folder. *Time-on-task* was considered in this sub-session as well.

<u>Analysis of Data</u>

Obtained data was analyzed using both descriptive and inferential statistics. Each analysis was carried out for every independent variable. A Single-Factor Within-Subject Design Analysis of Variance (ANOVA) was used to determine synthetic speech perception based on three levels of synthetic speech rate, that is, 40, 60, and 80 words per minute respectively. The same method was used to analyze auditory reading comprehension under different listening conditions.

Finally, a third Single-Factor Within-Subject Design Analysis of Variance (ANOVA) was used to assess task efficiency based on the time cumulated while completing tasks such as editing, transferring, and searching.

To avoid experimental errors at the third session, some of the intervening variables were held constant. For instance, the number of steps needed to complete any of the three tasks assigned was controlled to approximately 24 keystrokes and a maximum time-on-task of 55 seconds (even if allowing the screen access application to read aloud every utterance). Besides, each task should be completed in approximately 5 steps if following the general instructions (see Appendix C).

IV. RESULTS

This section contains the results for each experimental session. Gathered data was analyzed using both descriptive and inferential analyses. For instance, at the first session, the mean of words correctly recalled and standard deviation were calculated from a group of nine participants (see Appendix H, Table 1a). Treatment means indicated no differences in the number of words correctly recalled whether passages were "read aloud" at 40, 60, or 80 words per minute.

A Single-Factor Within-Subjects Analysis of Variance (ANOVA) was conducted showing no significant effects on the number of words correctly recalled as function of synthetic speech rate, F (2,16)=0.44, n. s. (see Appendix H, Table 1b). Such findings are a clear indicator that synthetic speech intelligibility, within the range tested, is not yet a negative factor in perception.

Likewise, the mean of *errors* (or number of words recalled but not contained in the passage) and standard deviation were calculated (see Appendix H, Table 2a). In this case, treatment means indicated that participants made the same amount of mistakes despite of *rate* levels. A Single-Factor Within-Subjects Analysis of Variance (ANOVA) was conducted indicating no treatment effects based on the number of errors or words incorrectly recalled, F (2,16)=1.34, n. s. (see Appendix H, Table 2b).

Now turning to the second experimental session, the mean of comprehension test scores and the standard deviation were calculated showing no significant differences in the comprehension test scores (see Appendix H, Table 3a). A Single-Factor Within-Subjects Analysis of Variance (ANOVA) was conducted revealing no significant effect,

$\underline{F}$ (1,8)=2.22, n. s. (see Appendix H, Table 3b). Results showed that, within the range tested, environmental noise does not represent a negative factor for synthetic speech comprehension.

A third and last construct, task efficiency while depending on speech-based interfaces, was assessed as well (see Appendix H, Table 4a). Data gathered as means of time-on-task indicated that the task involving "file transferring" took longer ($\underline{M}$=4.95) than those conducted for "searching" ($\underline{M}$=2.32) or editing ($\underline{M}$=2.14). Similarly, the standard deviation for the transferring task was notably higher followed by the scores obtained at the searching and editing task respectively.

Finally, a Single-Factor Within-Subjects Analysis of Variance (ANOVA) was conducted revealing a significant effect on task dependence, $\underline{F}$ (2,16)=6.96, $P<0.05$, R^2 =0.30 (see Appendix H, Table 4b). Although task knowledge was confirmed before each trial, the author observed periods in which the participants were not interacting with the system. This could be attributed to issues such as loss of momentum, lack of guidance, problems in differentiating system's landmarks (i. e., "Desktop" from the "Task Bar"), one-way navigation tendency, accidental use of applications, and hesitation. Further experiments are needed to confirm these speculations. Finally, the author did not observe major flaws on performance caused by synthetic speech intelligibility.

V. CONCLUSIONS

This research project was intended to evaluate perception and cross platform interaction while using speech-based interfaces and picture-based interfaces. More specifically, the author's concern about interaction with screen readers was addressed in terms of perception of verbal material while controlling prosodic and contextual features.

Likewise, task efficiency was considered trying to find quantitative and qualitative evidence to explain "context disparity" based on navigation problems, error rate, and loss of momentum (Cook, 1997; Stephanidis, et al., 1997; Lazarro, 2001; Barraga, & Erin, 2001; Edwards, Mynatt, & Stockton, 1995). More specifically, differences in time to complete tasks such as searching, editing, and transferring files pointed at platform limitations to sustain interaction, thus affecting task efficiency.

The author did not find any significant effect while controlling synthetic speech rate as prosodic feature, neither in the forms of words correctly recalled nor errors. Changes on rate, whether at 40, 60, or 80 words per minute, had no impact within the range tested.

Similarly, the author did not find any significant effect on noise as contextual factor, at least, within the range tested. Means of comprehension test score demonstrated that perception of verbal material while using synthetic speech output was the same either under ambient noise present or in the absence of such condition.

Results at this point can be interpreted as evidence of good synthetic speech intelligibility. As demonstrated previously, an intelligible output is a key factor in the perception of auditory reading material. Despite of the assumptions made by other authors criticizing this aspect on screen-access applications (Venkatagiri, 1994; Miranda and Beukelman, 1987; Pratt, 1987; Hjelmquist, Dahlstrand, & Hedelin, 1992), JAWS® V 5.0 Screen Reader has a reliable output source.

Although technological aspects of synthetic speech seem not to have a significant impact on interaction, further research is necessary to conclude over specific characteristics such as the *intonation quality* and the *robotic appearance* of JAWS's synthetic speech output.

In considering "context disparity" and cross platform flaws, the author found a significant effect based on the current experimental approach. Means of time-on-task were significantly different for ach one of the tasks including editing, searching, and file transferring. More specifically, the author found that transferring files while depending on synthetic speech output took almost twice as long as the searching and editing tasks respectively.

However, within the transferring and searching scenarios, there were common syntactic characteristics that could be the origin of navigation problems, error rate, poor error recovery, and loss of momentum. For instance, at the transferring and searching tasks, each folder was opened at a different location. This made the participants "jump" between windows in order to proceed with the next step. The author could speculate that due to an increasing demand of motor skills to perform keystroke combinations and to the need to hold tasks instructions on short-term memory, more time was used to complete the task. However, new experimental efforts are necessary to collect specific evidence in this regard.

In addition, the author observed that participants were more hesitant while trying to complete specific tasks, that is, difficulties trying to "jump" from window to window, navigated only in up-and-down direction, and looped into the same error as if trying to recover. Probably, there was a memory burden at the time. Nonetheless, further research should be conducted on this matter to quantify such phenomena.

What one could infer based on these observations is that the picture-based interface in this study (JAWS® V 5.0 by Freedom Scientific, Blind/Low Vision Group; St. Petersburg, FL) does not provide the blind user of any means to improve error recovery, momentum, or navigation. Despite of the different interaction strategies used to maintain performance, the interface did not respond or accommodate any emerging user requirements.

A meaningful content at semantic level (Edwards et al., 1995) and simplicity at the lexical level of interpretation could be the right combination to augment performance. It could be interesting to explore if navigation could be improved by redefining the content at such level, that is, by reducing the iconic dependence at the lexical level of interpretation (as in command line interfaces).

In addition, the use of alternative senses, both in the form of input or output, could improve navigation and task efficiency (Cook & Hussey, 2002). For instance, haptic interpretation may offer great advantages. Based on differences in channel capacity, it is even possible to have some advantages in *redundancy*. Incorporating information such as hardness, texture, and size into hard-assistive technologies could make the interface more reliable for the blind.

Although this research project was a basic attempt to understand human-computer interaction at speech-based interfaces and graphical user interfaces, the author believes that these findings could be a valuable contribution in the development of soft-assistive applications. Perhaps *adaptivity* and *compatibility* can be achieved by creating an assistive interface that works as portal of the GUIs functions but not as an instrument to translate its graphic content. In other words, it is like reversing the actual equation, which means depending more on the functions of the assistive interface rather than on a changing the GUI.

VI. REFERENCES

Ainsworth, W. A. (1972). Duration as a cue in the recognition of synthetic vowels. *Journal of the Acoustical Society of America, 51,* 648-651.

Baddeley, A. D. (1986). *Working memory.* Oxford, UK: Oxford University Press.

Bayley, R. W. (1989). *Human Performance Engineering.* Englewood Cliffs, NJ: Prentice Hall.

Barraga, N., Erin, J. (2001) *Visual Impairments and Learning.* Austin, TX: Pro-Ed.

Boyd, L. H.; Boyd, W. L.; Vanderheiden, G. C. (1990). The graphical user interface: Crisis, danger, and opportunity. *Journal of Visual Impairment and Blindness, 4,* 496-502.

Broadbent, D. E. (1958). *Perception and communications.* London: Pergamon Press.

Carreiras, M., Alvarez, C. (1999). Comprehension Processes in Braille Reading. *Journal of Visual Impairment and Blindness, 93, 589-595.*

Cole, R. A.; Jakimik, J. (1978). Understanding speech: How words are heard. In G. Underwood (Ed.), *Strategies of information processing* (pp. 67-116). London: Academic Press.

Cook, A. (1997) Assistive Technology: North American Perspectives. In G. Anogianakis et al. (Eds.) *Advancements on Assistive Technology* (pp. 263-267), US: IOS Press.

Cook, A., Hussey, S. (2002). *Assistive Technologies, Principles and Practices.* St. Louis, Missouri: Mosby, Inc.

Cherry, C. (1953). Some experiments on the perception of speech with one and two ears. *Journal of the Acoustical Society of America, 25,* 975-979.

Christian, K. (2000), *Design of Haptic and Tactile Interfaces for Blind Users.* Retrieved on August 15, 2003 from http://www.otal.umd.edu/UUGuide/kevin.html.

Christiansen, C., Baum, C. (1997). Person-environment occupational performance: a conceptual model for practice. In Christiansen C., Baum, C. (Eds.). *Occupational therapy: enabling function and well being.* Thorofare, NJ: Slack.

Demers, L., Weiss-Lambrou, R., Ska, B. (1996). Development of the QUEBEC User Evaluation of Satisfaction with Assistive Technology (QUEST). *Assistive Technologies, 8(3).* 3-11.

Demers, L. Monette M., Lapierre Y., Arnold D., Wolfson C. (2002). Reliability, validity, and applicability of the Quebec User Evaluation of Satisfaction with assistive Technology (QUEST 2.0) for adults with multiple sclerosis. *Disability Rehabilitation. 24(1-3),* 21-30.

Duchek, J. (1991). Assessing cognition. In Christiansen C., Baum, C. (Eds.). *Occupational therapy: overcoming human performance deficits.* Thorofare, NJ: Slack.

Edwards, W. K.; Mynatt, E. D.; Stockton, K. (1995). Providing Access to Graphical User Interfaces –Not Graphical Screens. *Interactions, 1,* 56-67.

Fink, J., Kobsa, A., Nill, A. (1997). Adaptable and Adaptive Information Access for All Users, Including the Disabled and the Elderly. In A. Jameson, C. Paris, and Tasso, C. (Eds.). *User Modeling: Proceedings of the Sixth International Conference UM97* (pp. 171-173) Wien, New York: Springer.

Fougeyrollas, P., Gray, D. B. (1998). Classification systems, environmental factors and social change. In Gray, D. B., Quatrano, L. A., Liberman, M. L. (Eds.). *Designing and using assistive technology: The human perspective.* Baltimore: Paul H. Brookes Publishing.

Gillon, G., Young, A. (2002). The Phonological-Awareness Skills of Children Who Are Blind. *Journal of Visual Impairment and Blindness, 96, 38-49.*

Hensil, J., Whittaker, S. (2000). Visual Reading Versus Auditory Reading by Sighted Persons and Persons with Low Vision. *Journal of Visual Impairment and Blindness, 94, 762-770.*

Hjelmquist, E., Dahlstrand, U., Hedelin, L. (1992). Visually impaired persons' comprehension of text presented with speech synthesis. *Journal of Visual Impairment and Blindness, 19, 426-428.*

Koenig, J., Holbrook, C. (1989). Determining the reading medium for student with visual impairments: A diagnostic teaching approach. *Journal of Visual Impairment and Blindness, 83, 296-302.*

Lazarro, J. (2001) *Adaptive Technologies for Learning and Work environments.* Chicago: American Library Association.

Liberman, A. M.; Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition, 21,* 1-36.

Marslen-Wilson, W.; Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10,* 29-63.

Massaro, D. W. (1984a). Building and testing models of reading processes. In P. D. Pearson (Ed.), *Handbook of reading research* (pp. 111-146). New York: Longman.

Massaro, D. (1987). *Speech Perception by Ear and the Eye: A paradigm for Psychological Inquiry.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Miller, G. A. (1981). *Language and speech.* San Francisco, CA: Freeman.

Miranda, P., Beukelman, R. (1987). A comparison of speech synthesis intelligibility with listeners from age three age groups. *Augmentative and Alternative Communication, 3, 120-128.*

Mousty, P., & Bertelson, P. (1992). Finger movements in Braille reading: The effect of local ambiguity. *Cognition, 43, 67-84.*

Odor, P. (1984). Hard and Soft Technology for Education and Communication For Disabled People. *Proceedings of The International Computational Conference.*

Osborn, R., Erber, N., Galletti, A. (2000). Effects of Background Noise on the Perception of Speech by Sighted Older Adults and Older Adults with Severe Low Vision. *Journal of Visual Impairment and Blindness, 94, 648-653.*

Pratt, L. (1987). Quantifying the performance of text to speech synthesizers. *Speech Technology, 5, 54-63.*

Pring, L. (1982). Phonological and tactual coding of Braille by blind children. *British Journal of Psychology, 73, 351-359.*

Proctor, R., Dutta, A. (1995). Performance of Multiple Tasks. In Lyle E. Bourne, Jr. (Ed.), *Skill Acquisition and Human Performance* (pp. 133-165). Thousand Oaks, CA: SAGE Publications.

Rosa, A., Huertas, J., Simon, C. (1994). Braille reading. In A. Rosa and E. Ochaita (Eds.). *Cognitive Psychology of the Blindness.* Madrid: Psychology Alliance.

Rusalem, H. (2001) *Coping With the Unseen Environment: An Introduction to the Vocational Rehabilitation of Blind Persons.* New York: Teachers College Press, Teachers College, Columbia University.

Simon, C., Huertas, J. (1998). How Blind Readers Perceive and Gather Information Written in Braille. *Journal of Visual Impairment and Blindness, 92, 322-330.*

Sharit, J.; Czaja, S.; Nair, S., Lee, C. (1999). Effects of Age, Speech Rate, and Environmental Support in Using Telephone Voice Menu Systems. *Human Factors (45)2, 218*-234.

Stephanidis, C., Savidis, A., Akoumianakis, D. (1997) Unified Interface Development: A step towards User Interfaces for All. In G. Anogianakis et al. (Eds.) *Advancements on Assistive Technology* (pp. 29-33), US: IOS Press.

Stephanidis, C., Paramythis, A., Savidis, A. Sfyrakis, M., Stergiou, A., Leventis, A., Maou, N., Paparoulis, G., Karagiannidis, C. (1997) Developing Web Browsers Accessible to All: Supporting User Adapted Interaction. In G. Anogianakis et al. (Eds.) *Advancements on Assistive Technology* (pp. 233-238), US: IOS Press.

Sumby, W. H.; Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America, 26,* 212-215.

Treisman, A. (1969). Strategies and models of selective attention. *Psychological Review, 76,* 282-299.

Van Reusen, K., Head, N. (1994). Cognitive and metacognitive interventions: Important trends for teachers of students who are visually impaired, *Re: view, 25, 153-162.*

Venkatagiri, S. (1994). Effect of sentence length and exposure on the intelligibility of synthesized speech. *Augmentative and Alternative Communication, 10, 96-104.*

Ward, L. M. (1994). Supermodal and modality-specific mechanisms for stimulus-driven shifts of auditory and visual attention. *Canadian Journal of Experimental Psychology, 48,* 242-259.

Wetzel, R., Knowlton, M. (2000). A Comparison of Print and Braille Reading Rates on Three Reading Tasks. *Journal of Visual Impairment and Blindness, 94, 146-154.*

Wickens, C., Hollands, J. (2000). *Engineering Psychology and Human Performance.* Upper Saddle River, NY: Prentice-Hall.

Appendix A: Consent Form

Jose R. Concepcion
701 Curtner Ave, Apt. 339. San Jose, CA 95125
Phone: 408-806-2174, E-mail: jrcq_sjsu@yahoo.com, jrcq_sjsu@hotmail.com

---

AGREEMENT TO PARTICIPATE IN RESEARCH

You have been asked to participate in a research study investigating aspects such as *compatibility* and *adaptivity* between screen-access applications and graphical user interfaces. As part of the interface evaluation, you will be asked to complete a series of tasks using JAWS® V 5.0 screen reader as main interface. Hardware to be used throughout the evaluation process include a Dell Inspiron Laptop 8200 (by Dell Computer Corporation) and a Labtec® Elite™ 820 Headphones (by Labtec Inc., Fremont, CA).

As described on the protocols, participants will be requested to complete the following:

A- Intentionally word recall from three passages read at different synthetic speech rates (40, 60, and 80 wpm respectively).

B- Listening to two passages read under different listening conditions (namely, under ambient noise present and ambient noise absent) and then complete two multiple-choice tests based on the content of each passage.

C- Edit system's settings, search, and transferring files within a given folder structure.

There are no foreseeable risks involved in this study, that is, physical and/or emotional discomforts. As participant, you will be rewarded for your contributions to this study. You will receive $50.00 dollars in cash after the completion of the third session due to the amount of time invested on the evaluation –around one hour and thirty minutes.

No service of any kind, to which you are otherwise entitled, will be lost or jeopardized if you choose to "not participate" in the study. Your consent is being given voluntarily. You may refuse to participate in the entire study or in any part of the study. If you decide to participate in the study, you are free to withdraw at any time without any negative effect on your relations with San Jose State University or with any other participating institutions or agencies. At the time that you sign this consent form, you will receive a copy of it for your records, signed and dated by the investigator.

Questions or complaints about this study may be addressed to Kevin Corker Ph.D., at the Industrial & Systems Engineering Department, San Jose State University (One Washington Square San Jose, California 95192; Telephone: 408-924-3988, E-mail: kcorker@email.sjsu.edu). In addition, queries about research subjects' rights, or research-related injuries may be presented to Pamela Stacks, Ph.D., Interim Associate Vice President, Graduate Studies and Research, at (408) 924-7029.

1. The signature of a subject on this document indicates agreement to participate in the study.
2. The signature of a researcher on this document indicates agreement to include the above named subject in the research and attestation that the subject has been fully informed of his or her rights.

_______________________________________ ________________
Participant's Signature Date

_______________________________________ ________________
Investigator's Signature Date

Appendix B: Screening Form

Screening Form
(Adapted from Cook and Hussey, 2002 p.129)

Instructions: This questionnaire is aimed to gather information from the participant to verify that the individual match with the screening criteria:

General Information

A1/Name:

___

A2/Age:

___

A3/Native language:

___

A4/Ethnic group:

___

A5/Residence (select from the following):
1. Live with family,
2. Alone,
3. With a caregiver,
4. Others

A6/Impairment origin:
- □ Unknown cause
- □ Conditions associated with general diseases
- □ Prenatal causes

A7/Educational level:

___

A8/Employment status:

___

Platform Proficiency

B1/Years operating computers:

___

B2/Type of hardware used:

___

B3/ Product exposure (time using screen readers):

___

B4/ Type of screen reader used (if used):

___

B5/General abilities with Microsoft Windows® Operating System (based on self-assessment):

- ☐ Low
- ☐ Sufficient
- ☐ High

B6/Current use (select one or more from the following):

- ☐ Browsing
- ☐ Communication
- ☐ Word processor
- ☐ Others

Medical And Health Information

C1/Origin of blindness or visual disability:

- ☐ Cataracts
- ☐ Trachoma
- ☐ Glaucoma
- ☐ Vitamin "A" deficiency
- ☐ Onchocerciasis or "river blindness"
- ☐ Leprosy
- ☐ Diabetic retinopathy
- ☐ Injuries
- ☐ Age-related macular degeneration
- ☐ Prenatal causes
- ☐ Unknown cause
- ☐ Others
- ☐ Refused to state

C2/Anticipated course of condition:

- ☐ Stable
- ☐ Improving
- ☐ Deteriorating
- ☐ Fluctuating

Sensory And Perceptual Abilities

Vision:
D1/ Completely blind? (YES/NO)
D2/ Wear glasses or corrective lenses? (YES/NO)
D3/ Does light affect vision quality? (YES/NO)
D4/ Can fixates vision on a stationary object? (YES/NO)
D5/ Can fallows a moving object? (YES/NO)

Hearing:
D6/ Have any kind of hearing loss? (YES/NO)
D7/ Wear hearing aids? (YES/NO)
D8/ React rapidly to sounds? (YES/NO)
D9/ Can understands speech? (YES/NO)

Social Interaction, Learning, And Behavior

E1/Level of alertness:
- ☐ No interest in surroundings
- ☐ Little interest in surroundings
- ☐ Sometimes
- ☐ "Observant"
- ☐ Very Alert

E2/Social play:
- ☐ Unoccupied behavior
- ☐ Onlooker behavior
- ☐ Solitary
- ☐ Little interaction
- ☐ In constant interaction with others

Appendix C: Experimental Protocols

EXPERIMENTAL PROTOCOLS

Confirm subject's profile

Candidates must comply with the following demographic and clinical criteria:

Age between 20 and 50 years old, good hearing (that is, not depending on any hearing aid), have English as first or native language, knowledge of Microsoft Windows XP® operating system (or at least one year of experience), knowledge of basic keystrokes (such as ESC, TAB, SPACE, BACKSPACE, ENTER, etc), and basic knowledge or entry-level experience with screen access applications or any other soft-assistive technology. Disregard for recruitment purposes issues such as gender, ethnic group, residence, impairment origin, educational level, employment status, and platform use.

Set up and testing

Instructions: Install and test hardware and software. Fill out the CONSENT FORM and gather demographic information through SCREENING DOCUMENT. Archive files for further consultation. Then, confirm the following environmental and human attributes:

- ☐ Posture
- ☐ Distance (approximate height: 30"; horizontal distance of about 28" maximum)
- ☐ Comfort (as stated by the subject, use dichotomous response)
- ☐ Sound (~60 dB)
- ☐ Number of observers (no more than one excluding the experimenter)
- ☐ Can understand concept of direction and location?
- ☐ Room temperature

Finally, Go to DATA COLLECTION FORM and insert this information.

Experimental Session I

Experimenter: Have the passage ready on the screen. REMEMBER to use a different passage for each session. Reading passages must be assigned randomly to every participant. Finally, collect the following data:

1. Words recalled
2. Total of words correctly recalled (# words)
3. Synthetic speech rate (wpm)
4. Passage title (Colonial-Indian Relations, The First Americans, or Early Settlements)

Finally, Go to DATA COLLECTION FORM and insert this information.

Participant: You are about to begin the first session of the experiment, aimed to observe the effects of synthetic speech rate over perception. In other words, we will analyze part of JAWS® V 5.0 synthetic speech capabilities. Your goal as participant will be recalling as many words as possible from the following passages. Each passage (three in total) will be read aloud by the screen access application.

Once every passage is read at three different synthetic speech rates, that is, the first passage at 40wpm, the second passage at 60 wpm, and the third passage at 80wpm, you will either write down or dictate as many words you can recall from the passage. Right after each exercise, please inform the experimenter once you are done.

Experimental Session II

Experimenter: Wait for 10 minutes before the start this session. Prepare to launch the second session. Set JAWS® V 5.0 default synthetic speech rate. Open Windows Media Player V 8.0, and playback the file "Heavy Traffic, Auto Horns, Voices.mp3", if assigned condition is ambient noise present. Then, open the reading passage and let JAWS® V 5.0 "scan and read aloud" its content. REMEMBER to use a different passage for each session, reading passages to be assigned randomly. Conduct the first session (ambient noise absent) and then the second one (ambient noise present). At the end of both sections, the subject may complete the comprehension test orally, for each one of the sessions. Finally, collect the following data:

1. Comprehension test scores (%)
2. Synthetic speech rate (wpm): JAWS® V 5.0 default
3. Passage title (Aviation, Architecture, or Armed Forces)

Finally, Go to DATA COLLECTION FORM and insert this information.

Participant: You are about to begin the second session of the experiment, intended to measure synthetic speech perception under changes in the listening conditions. In other words, we will analyze part of JAWS® V 5.0 synthetic speech capabilities along with ambient noise. In this exercise, pay close attention to each of the following passages. Each one (two in total) will be read aloud by the screen access application. At the end of the reading, you will be requested to complete a brief multiple-choice comprehension quiz of only 10 questions, read out to you by the experimenter.

The first passage will be simply read by the screen access application without any interfering noise. Nevertheless, conditions will change on the second passage, where the experimenter will play back an audio file resembling street noise (cars, horns, buses, etc). In both scenarios, the participant must complete the multiple-choice test whatsoever. Test grades will be used for experimental purposes only, and will remain in confidentiality until the conclusion of the third session.

Experimental Session III:

Experimenter: Wait for 10 minutes before the start this session. Prepare to launch the third session. Set JAWS® V 5.0 default synthetic speech rate. Instruct or review with the participant steps to be taken in order to reach the assigned goal; make sure the participant can complete the task using previous knowledge. For instance:

A) To edit setting of the soundcard:

E. G: Windows Menu + All Programs + Control Panel + Sound and Audio Devices + Volume + Device Volume + High

B) To transfer a file from a different location:

E. G.: Control + C, Control + V,

C) Browse files and folders.

The experimenter is not allowed to assist during any of the assigned task. Finally, collect the following data:

1. Time-on-task.
2. If the task was successfully completed or not (or number of tasks successfully completed).

Finally, Go to DATA COLLECTION FORM and insert this information.

Participant: You are about to begin the third and final session of the experiment, intended to measure compatibility and adaptivity of screen readers with graphical platforms or GUIs. This session comprises of three parts. First, you will be requested to "adjust the volume of the computer's sound card" (or a simple volume adjustment) using only certain keystrokes for navigation, for instance, TAB, SPACE, ESC, ENTER, ARROW KEYS, and CTRL. Since no visual contact with the display is allowed for those partially blind, you will only depend on the audible feedback emitted by the screen reader while trying to achieve such goal. As participant, you must inform the experimenter whenever you think you have achieved the task on discussion.

The second task is in somehow similar to the first one, since requires keystrokes to navigate through folders in the operating system. As participant, you will be requested to "copy" or "move" a text file named "Test Document" from location A, the computer's desktop, to location B, the "Destination Folder", located right after "My Computer". Once again, you must inform the experimenter whenever you think you have achieved the task on discussion.

Finally, a third task will also require of keystrokes in order to complete the exercise successfully. In this case, you as participant will be requested to "search" or "browse" a specific file named "Target File" located at the "Target Folder 1". Target folders can be found on the computer's desktop, ranging from "Target Folder 5" to "Target Folder 1". Remember to inform the experimenter whenever you think you have achieved the task on discussion.

Appendix D: Narrative Passages (Experimental Session I)

NARRATIVE PASSAGE: THE FIRST AMERICANS

Fragment by U.S. Department of State's Bureau of International Information Programs

"At the height of the Ice Age, between 34,000 and 30,000 B.C., much of the world's water was contained in vast continental ice sheets. As a result, the Bering Sea was hundreds of meters below its current level, and a land bridge, known as Beringia, emerged between Asia and North America. At its peak, Beringia is thought to have been some 1,500 kilometers wide. A moist and treeless tundra, it was covered with grasses and plant life, attracting the large animals that early humans hunted for their survival.

The first people to reach North America almost certainly did so without knowing they had crossed into a new continent. They would have been following game, as their ancestors had for thousands of years, along the Siberian coast and then across the land bridge.

Once in Alaska, it would take these first North Americans thousands of years more to work their way through the openings in great glaciers south to what is now the United States. Evidence of early life in North America continues to be found. Little of it, however, can be reliably dated before 12,000 B.C.; a recent discovery of a hunting lookout in northern Alaska, for example, may date from almost that time. So too may the finely crafted spear points and items found near Clovis, New Mexico.

Similar artifacts have been found at sites throughout North and South America, indicating that life was probably already well established in much of the Western Hemisphere by some time prior to 10,000 B.C.

Around that time the mammoth began to die out and the bison took its place as a principal source of food and hides for these early North Americans. Over time, as more and more species of large game vanished -- whether from over hunting or natural causes -- plants, berries and seeds became an increasingly important part of the early American diet. Gradually, foraging and the first attempts at primitive agriculture appeared. Indians in what is now central Mexico led the way, cultivating corn, squash and beans, perhaps as early as 8,000 B.C. Slowly, this knowledge spread northward.

By 3,000 B.C., a primitive type of corn was being grown in the river valleys of New Mexico and Arizona. Then the first signs of irrigation began to appear, and by 300 B.C., signs of early village life.

By the first centuries A.D., the Hohokum were living in settlements near what is now Phoenix, Arizona, where they built ball courts and pyramid-like mounds reminiscent of those found in Mexico, as well as a canal and irrigation system."

Appendix D: Narrative Passages (Experimental Session I)

NARRATIVE PASSAGE: COLONIAL-INDIAN RELATIONS
Fragment by U.S. Department of State's Bureau of International Information Programs

"By 1640 the British had solid colonies established along the New England coast and the Chesapeake Bay. In between were the Dutch and the tiny Swedish community. To the west were the original Americans, the Indians.

Sometimes friendly, sometimes hostile, the Eastern tribes were no longer strangers to the Europeans. Although Native Americans benefited from access to new technology and trade, the disease and thirst for land which the early settlers also brought posed a serious challenge to the Indian's long-established way of life.

At first, trade with the European settlers brought advantages: knives, axes, weapons, cooking utensils, fishhooks and a host of other goods. Those Indians who traded initially had significant advantage over rivals who did not.

In response to European demand, tribes such as the Iroquois began to devote more attention to fur trapping during the 17th century. Furs and pelts provided tribes the means to purchase colonial goods until late into the 18th century.

Early colonial-Indian relations were an uneasy mix of cooperation and conflict. On the one hand, there were the exemplary relations, which prevailed during the first half century of Pennsylvania's existence. On the other were a long series of setbacks, skirmishes and wars, which almost invariably resulted in an Indian defeat and further loss of land.

The first of the important Indian uprisings occurred in Virginia in 1622, when some 347 whites were killed, including a number of missionaries who had just recently come to Jamestown. The Pequot War followed in 1637, as local tribes tried to prevent settlement of the Connecticut River region.

In 1675 Phillip, the son of the chief who had made the original peace with the Pilgrims in 1621, attempted to unite the tribes of southern New England against further European encroachment of their lands. In the struggle, however, Phillip lost his life and many Indians were sold into servitude.

Almost 5,000 kilometers to the west, the Pueblo Indians rose up against the Spanish missionaries five years later in the area around Taos, New Mexico. For the next dozen years the Pueblo controlled their former land again, only to see the Spanish retake it. Some 60 years later, another Indian revolt took place when the Pima Indians clashed with the Spanish in what is now Arizona.

The steady influx of settlers into the backwoods regions of the Eastern colonies disrupted Indian life. As more and more game was killed off, tribes were faced with the difficult choice of going hungry, going to war, or moving and coming into conflict with other tribes to the west. The Iroquois, who inhabited the area below Lakes Ontario and Erie in northern New York and Pennsylvania, were more successful in resisting European advances."

Appendix D: Narrative Passages (Experimental Session I)

NARRATIVE PASSAGE: EARLY SETTLEMENTS

Fragment by U.S. Department of State's Bureau of International Information Programs

"The early 1600s saw the beginning of a great tide of emigration from Europe to North America. Spanning more than three centuries, this movement grew from a trickle of a few hundred English colonists to a flood of millions of newcomers. Impelled by powerful and diverse motivations, they built a new civilization on the northern part of the continent.

The first English immigrants to what is now the United States crossed the Atlantic long after thriving Spanish colonies had been established in Mexico, the West Indies and South America. Like all early travelers to the New World, they came in small, overcrowded ships. During their six- to 12-week voyages, they lived on meager rations. Many died of disease; ships were often battered by storms and some were lost at sea.

Most European emigrants left their homelands to escape political oppression, to seek the freedom to practice their religion, or for adventure and opportunities denied them at home. Between 1620 and 1635, economic difficulties swept England. Many people could not find work. Even skilled artisans could earn little more than a bare living. Poor crop yields added to the distress. In addition, the Industrial Revolution had created a burgeoning textile industry, which demanded an ever-increasing supply of wool to keep the looms running. Landlords enclosed farmlands and evicted the peasants in favor of sheep cultivation. Colonial expansion became an outlet for this displaced peasant population.

The colonists' first glimpse of the new land was a vista of dense woods. The settlers might not have survived had it not been for the help of friendly Indians, who taught them how to grow native plants -- pumpkin, squash, beans and corn. In addition, the vast, virgin forests, extending nearly 2,100 kilometers along the Eastern seaboard, proved a rich source of game and firewood. They also provided abundant raw materials used to build houses, furniture, ships and profitable cargoes for export.

Although the new continent was remarkably endowed by nature, trade with Europe was vital for articles the settlers could not produce. The coast served the immigrants well. The whole length of shore provided innumerable inlets and harbors. Only two areas -- North Carolina and southern New Jersey -- lacked harbors for ocean-going vessels.

Majestic rivers -- the Kennebec, Hudson, Delaware, Susquehanna, Potomac and numerous others -- linked lands between the coast and the Appalachian Mountains with the sea. Only one river, however, the St. Lawrence -- dominated by the French in Canada -- offered a water passage to the Great Lakes and into the heart of the continent. Dense forests, the resistance of some Indian tribes and the formidable barrier of the Appalachian Mountains discouraged settlement beyond the coastal plain. Only trappers and traders ventured into the wilderness. For the first hundred years the colonists built their settlements compactly along the coast."

Appendix D: Narrative Passages (Experimental Session II)

NARRATIVE PASSAGE: AVIATION
Fragment by Crouch, Tom D.

"Americans have always been fascinated by the possibility of flight. On June 24, 1784, only seven months after Pilatre de Rozier and the Marquis D'Arlandes became the first human beings to fly, thirteen-year-old Edward Warren rose above the streets of Baltimore aboard a balloon constructed by Peter Carnes, a lawyer and tavern keeper from Maryland. During the next century, balloons became a familiar sight, but the gaily-decorated gasbags were captives of the wind. Navigating in air with the freedom of the birds came only with the invention of the airplane.

During the 1890s Octave Chanute and Samuel Pierpont Langley helped set the stage for achieving winged flight. In 1896, Langley, the secretary of the Smithsonian Institution, launched a series of large, steam-powered model aircraft on flights of up to three-quarters of a mile over the Potomac River. Several months later, Chanute, a civil engineer, led a band of assistants into the dune country east of Chicago, where they flew a series of manned gliders, including an advanced biplane.

Wilbur and Orville Wright, the proprietors of a bicycle shop in Dayton, Ohio, wrote to Langley and Chanute in 1899-1900, requesting information on aeronautics and announcing their decision to conduct their own tests. They made the world's first powered, sustained, and controlled flights with a heavier-than-air flying machine at Kitty Hawk, North Carolina, on December 17, 1903. Unwilling to risk unveiling their technology without the protection of a patent and a contract for the sale of airplanes, the Wrights did not make their flights in public until 1908. By that time, photographs and descriptions of their machine had inspired other pioneers to follow their lead.

Glenn Hammond Curtiss, a motorcycle builder from New York, emerged as their most important American rival. Flying in a competition in France in 1909, Curtiss won the first James Gordon Bennett trophy competition with a speed of forty-six miles per hour. In spite of the Wrights' legal efforts to curb his activity, Curtiss had, by 1914, established himself as the most successful of all American aircraft manufacturers.

American aeronautical hegemony was short-lived, however. With war looming, European leaders were quick to recognize the military potential of the technology and to encourage its development by sponsoring speed, altitude, and distance competitions, establishing aerial units in their armed forces, and creating laboratories to conduct research and development programs.

During World War I, the nation that had given birth to the airplane only fourteen years before scarcely qualified as a third-rate aeronautical power. American pilots flew into combat aboard airplanes designed and, for the most part, manufactured in Europe. In spite of some success in the production of training craft and engines, the performance of the fledgling aircraft industry was disappointing.

Postwar congressional investigations underscored the problems of a limited market and high research and development costs faced by airframe and engine manufacturers. Recognizing the growing importance of the airplane to national defense and prestige, federal officials took a series of steps between 1915 and 1940 designed to strengthen and regulate the aviation industry.

Established by Congress in 1915, the National Advisory Committee on Aeronautics (NACA) conducted programs of research and development that by 1925 had demonstrated the value of basic research. Technical reports issued by the agency introduced U.S. aircraft designers to a host of improvements, including revolutionary airfoils; improved propellers, engines, and instruments; and various streamlining techniques. Specialists experimented with wing flaps and other high-lift devices and explored innovative construction techniques and new materials.

American engineers made use of the information provided by the NACA, university researchers, and organizations dedicated to flight research. By the 1930s, a new generation of low-wing, streamlined, all-metal airplanes were flowing off their drawing boards. Aircraft like the Boeing 247, the Douglas DC-3, and the Sikorsky, Martin, and Boeing flying boats marked the return of the United States to a position of world aeronautical leadership.

Congressional leaders had taken steps to ensure that there would be a market for the new airplanes. The Kelly Air Mail Act of 1925 authorized the use of private companies for the delivery of airmail. Most American airlines trace their lineage back to contract mail carriers; postal subsidies were an important source of income during the years when paying passengers were few and far between.

The government also regulated commercial aviation. The Air Commerce Act of 1926 created a Bureau of Aeronautics within the Commerce Department, which had limited regulatory authority and was charged with establishing aids to aerial navigation. The Civil Aeronautics Act of 1938 and the Civil Aeronautics Board and Civil Aeronautics Administration (1940) worked to improve passenger safety, route markings, and air traffic control systems.

The time between the wars was the golden age of American aviation. The products of companies like Lockheed, Boeing, Douglas, and Northrop were instantly recognizable by small boys from coast to coast. The pilots who flew higher, faster, and farther—fliers like Charles Lindbergh, Amelia Earhart, Jimmy Doolittle, Wiley Post, Richard Byrd, and Howard Hughes—were the heroes of what everyone referred to as the air age.

The airplane, an instrument of commerce, also gave birth to total war during the years 1939-1945. Traditional definitions of the battlefield lost their meaning in an age when fearful destruction could be rained on the enemy's heartland. Attacks from the sky directed against Guernica (1937), Nanking (1937), Warsaw (1939), Pearl Harbor (1941), and a hundred other places climaxed with the destruction of Hiroshima and Nagasaki by atomic bombs in 1945. From the great carrier battles of the Pacific to the fierce combat fought four miles up in the sky over Europe, the products of American aircraft builders carried the day.

Traditional piston-engine, propeller-driven aircraft technology reached its height during the Second World War. But far more revolutionary was the turbojet engine, which opened the way to much higher speeds. After the war the pressure of international tension between the United States and the Soviet Union led to increased defense spending and a drive for supremacy in the field of aerospace technology."

Appendix D: Narrative Passages (Experimental Session II)

NARRATIVE PASSAGE: ARCHITECTURE
Fragment by Wright, Gwendolyn

"From its colonial and Native American origins unto the present day, American architecture has been exceptionally complex, both in the multiple traditions from which it has drawn and in the variations of style and public response it has produced. Architecture has also been caught up in the commercial domain. Major commercial buildings—business blocks and department stores in the nineteenth century, office towers and shopping malls in the twentieth—dominate the public domain of most cities.

There has been no official patronage of certain firms for governmental commissions. Instead architects compete for most jobs; they must sell their skills and new ideas to the public; and they constantly vie with builders, who still produce the bulk of new housing and commercial buildings in this country.

Efforts to promote hegemonic styles and master architects, though continually asserted, have never fully prevailed. When the first European settlers arrived in the New World, some two hundred nations of Native Americans had already developed their own architectural traditions. Likewise, the colonialists' Spanish mission and the clapboard-sided New England dwelling of the seventeenth century endure as romantic prototypes in many parts of the country, suggesting a synthesis of local cultures into an idealized homogeneous society—a recurring theme in American cultural history.

The New England and Spanish colonies at first required settlements to be grouped in towns, close to a central plaza or meetinghouse square for governmental and religious structures, reinforcing social and architectural homogeneity. In time the abundance of land and the social constraints of these societies spun off new settlements and eventually isolated farmsteads. All the same, architecture remained a significant social act that reinforced community ties, even in the sparsely settled southern colonies. The planning and construction of a major building, whether an aristocratic home or a parish church, involved considerable deliberation, time, and effort on the part of many people.

By the time of the Revolution the American colonies still had no trained architects. Journeymen carpenters designed and built most structures, but gentlemen amateurs often undertook their own estates, following the English tradition that regarded architecture as one of the refined arts. Both groups looked to English architectural books and builders' guides for counsel.

They adapted the fashion of neoclassicism to their own circumstances, usually building in wood and simplifying the detailing in a style we have come to call Federal. Especially admirable was the work of Boston's Charles Bulfinch.

Thomas Jefferson envisioned his major architectural works—Monticello (1772, 1789-1809), the Virginia State Capitol (1785-1796), and the University of Virginia (1817-1826)—as prototypes for the new nation. Looking to Enlightenment ideals of social reform, Jefferson believed that the right environment could uplift minds and promote civic virtue. The nation's first professional architect, the English émigré Benjamin Latrobe, aspired to similar ideals, hoping that moral purpose would enhance professional prerogatives. He designed the Bank of Pennsylvania (1798-1800) and the Baltimore Cathedral (1804-1821), but he was even more appreciated for his engineering skills, most evident in the Philadelphia water system (1798-1801).

The early nineteenth century witnessed an extraordinary rate of urbanization, with a plethora of new banks, exchanges, public markets, and commercial buildings. Despite the absence of any form of public regulation, distinct districts appeared, including elegant blocks of colonnaded row houses and the first multifamily tenements and lodging houses. A greater homogeneity of styles was visible, in both dwellings and public buildings; yet variations in details, materials, and even the arrangement of blocks maintained the particularity of cities such as Savannah, Charleston, Baltimore, Boston, and Cincinnati.

Among the buildings that attracted the most attention were asylums—prisons, orphanages, almshouses, mental hospitals, and the like—now removed to the outskirts of cities where it was possible to have a carefully controlled environment. John Haviland's Eastern State Penitentiary in Philadelphia (1823-1826), with its radial arrangement of cellblocks around a central control station, was one of the first instances of American architecture influencing European design.

A belief in reform through the environment also fueled the park movement, beginning with the rural cemeteries of the 1830s and culminating in Frederick Law Olmsted's majestic Central Park (1857-1880). The parks provided a respite within the commercial city, a place designed for leisure, recreation, nature, and social intermixing. Olmsted, a major figure in American intellectual and reform circles, elaborated the idea of landscaped "parkways" to connect a coordinated park system, most notably the Emerald Necklace for the city of Boston, designed with Charles Eliot (1889).

Some of Olmsted's numerous other commissions suggest the wide range of late-nineteenth-century residential settings which saw moral purpose in planned natural beauty: Riverside, Illinois (1868), a bucolic suburb outside Chicago; Riverside Drive, New York (1888); Stanford University (1888); and Vandergrift, Pennsylvania (1895), a planned factory town for the Apollo Steel Company.

What Olmsted accomplished in site planning, others parlayed into architectural design. A distinctive and original American architectural fashion—later named the Shingle Style—emerged in the suburban and resort work of such architects as Henry H. Richardson, William Ralph Emerson, and Willis Polk. Residences were self-consciously individualized, combining diverse materials and architectural elements with a rambling floor plan. Post-Civil War industrialization allowed this style to proliferate, with the factory production of building materials and ornament, inexpensive pattern books with abundant Figures, and electric streetcars to facilitate commuting.

Yet stylistic diversity persisted. These same decades also produced museums, libraries, and sumptuous homes in the manner of Renaissance palazzi or medieval castles—the most grandiose by Richard Morris Hunt and the firm of McKim, Mead, and White. Business districts were being transformed by another architectural innovation, similarly the product of technological advances and artistic creation: the skyscraper. Beginning in the 1880s, architects and engineers in Chicago and New York began to experiment with new framing systems to achieve greater height, as well as an appropriate elevation for these taller buildings—at first only ten and then soon twenty stories high."

Appendix D: Narrative Passages (Experimental Session II)

NARRATIVE PASSAGE: ARMED FORCES
Fragment by Weigley, Russell F.

"The most important historical influence shaping the armed forces of the United States has been the abruptness of their transition from a merely peripheral involvement in international politics to a center-stage role of dauntingly large responsibilities.

Through most of its history, the U.S. Army and the colonial forces that preceded it were not instruments of foreign policy but tools for the domestic task of protecting settlements against the North American Indians. When World War II broke out in Europe in 1939, the army was still mainly deployed for constabulary tasks, patrolling the Mexican border and policing colonial subjects in the Philippine Islands.

Until nearly the end of the nineteenth century, the navy similarly was less an instrument of foreign policy than a device for showing the flag in support of private American business ventures around the world. Although the navy began to take on diplomatic significance sooner than the army, neither of the armed forces was prepared by its history for sudden elevation during World War II to international preeminence.

The military institutions of the United States have their roots in the militia systems of the British colonies before 1776. Because of dangers posed by the Indians and the rival colonial powers, all the colonies that were to form the United States except Quaker Pennsylvania compelled their citizens to become part-time soldiers under universal military service laws applying to males of appropriate age. (Even Pennsylvania created a voluntary militia in 1755 and made military service virtually compulsory in 1775.)

The deficiencies of a part-time soldiery for campaigns extended in time or geography led to supplementing the militia with the British regular army, beginning on an important scale in 1755, early in the French and Indian War.

Thus the United States inherited from the colonial era a dual military tradition of citizen-soldiers and regulars. The United States established a standing army modeled on the British regulars, first in the Continental army of the revolutionary war, growing out of legislation of the Second Continental Congress of June 14, 1775. This army was almost completely disbanded by the Confederation Congress on June 2, 1784, but the next day Congress authorized the creation of a new, albeit small force that the government of the Constitution inherited in 1789 and that became the nucleus of the Regular Army. The United States also accepted the compulsory-service militia legacy, particularly with the Second Amendment to the Constitution, which went into effect in 1791.

For defense against foreign enemies, it was intended that the militia would be mobilized to reinforce the Regular Army, which remained small throughout the nineteenth century; it numbered fewer than twenty-five thousand as late as the eve of the Spanish-American War of 1898. Much of the time, however, Americans perceived the Regular Army and the militia less as complementing each other than as rivals.

But Indian troubles along with friction with Great Britain and France during the French Revolution and the Napoleonic Wars prevented the anti-standing-army tradition from prevailing altogether. Furthermore, the compulsory-service militia declined before the Civil War because a rhetorical preference for citizen-soldiers over professionals became an insufficient motive for enforcing universal military training once the Indian frontier had receded westward. The rivalry between citizen-soldiers and professionals remained alive, however, as volunteer militia companies sprang up all over the United States in the first half of the nineteenth century. These companies achieved a remarkable vitality, and their drill competitions became a vehicle for expressing rivalries among towns and cities before the heyday of organized athletics.

Possessing such military forces, both the secessionist and the loyal states had the means to begin the Civil War in 1861, the volunteer companies providing the core of both armies. They were later supplemented by conscription, in what became the first American war of mass armies, over 2 million men eventually serving in the Union forces and some 750,000 in the Confederate forces. The Regular Army of the United States, only about 16,000 strong when the Civil War began, found itself playing only an inconspicuous role in the war.

The role was not unimportant, however, for the Regular Army had been growing increasingly professional. The U.S. Military Academy at West Point, founded in 1802, had after the War of 1812 added to what was essentially an engineering curriculum a solid foundation of schooling in tactical practice and strategic thought. The West Point influence encouraged officers to continue their studies after graduation, and some visited Europe to observe armies there. By the Civil War, both armies were commanded, organized, and administered with considerable professional skill largely by officers drawn from the Regular Army.

After the war, the emphasis on educated professionalism as a criterion of officership intensified. Gen. William Tecumseh Sherman, commanding general of the army from 1869 to 1883, was a vigorous exponent of military education and encouraged the formation of schools for the various combat arms to advance the professional education of their officers beyond West Point's undergraduate curriculum.

The navy also took steps toward refining officer professionalism. At the U.S. Naval Academy, founded at Annapolis, Maryland, in 1845, the technical mastery of seamanship took priority over military study, which the navy had little need for until the Civil War. Its single-ship duels in the Quasi-War with France in 1798-1800 and the War of 1812, its commerce-raiding in the latter war, and its major peacetime function of showing the flag required little strategic or military insight. But the blockading and capturing of Confederate ports and, after the war, the first intimations of American participation in world politics provided the impetus for establishing the Naval War College at Newport, Rhode Island, in 1884.

The War College helped transform the navy from a loose collection of individual ships into coordinated squadrons organized around battleships and designed to contend for command of the seas against potential enemies. Although the navy could not yet challenge the preeminence of the British Royal Navy, its transformation had progressed far enough by 1898 to achieve spectacular successes in the war with Spain."

Appendix E: Comprehension Tests

COMPREHENSION TEST
Aviation
Fragment by Crouch, Tom D.

Participant #: ___
Session: A/B

1.1) What was the name of the city in which thirteen-year-old Edward Warren did his first flight aboard a balloon?
A/ Baltimore*
B/ Mississippi
C/ Kentucky

1.2) In what year was launched the steam-powered model aircraft that flew over the Potomac River?
A/ 1850
B/ 1896*
C/ 1910

1.3) When was Wilbur and Orville Wright (known as the Wright brothers) made the world's first powered, sustained, and controlled flight at Kitty Hawk, North Carolina?
A/ On Dec. 17, 1903*
B/ On August 20, 1902
C/ On June 5, 1897

1.4) Who was the first winner of the James Gordon Bennett trophy competition held in France in 1909?
A/ Glenn Curtiss*
B/ Orville Wright
C/ Wilbur Wright

1.5) European leaders encouraged the development of military technology by sponsoring competitions aimed to:
A/ To increase aircraft size
B/ To increase flight safety
C/ To increase speed, altitude, and distance*

1.6) Most of the combat planes flew by American pilots during World War one were manufactured at:
A/ Asia
B/ Europe*
C/ United States

1.7) Congressional efforts to strengthen and regulated the aviation were mostly made between:
A/ 1915 and 1940?*
B/ 1900 and 1912?
C/ 1978 and 1985?

1.8) The National Advisory Committee on Aeronautics –created to conduct programs of research and development- was established by the congress on:
A/ 1989?
B/ 1915?*
C/ 1920?

1.9) By the 1930s, a new generation of low-wing, streamlined, all-metal airplanes were flowing off their drawing boards. Which was one of the following was one of the most famous models?
A/ The F-18 Hornet
B/ The F-4 Phantom
C/ The Douglas DC-3*

1.10) Congressional leaders took steps to ensure that there would be a market for new airplanes built. In that order, the Kelly Air Mail Act of 1925 authorized:
A/ The creation of more customs?
B/ The use of private delivery of airmail?*
C/ The construction of more research facilities

* Correct answers

Appendix E: Comprehension Tests

COMPREHENSION TEST
Architecture
Fragment by Wright, Gwendolyn

Participant #: ___
Session: A/B

2.1) From colonial origins, the American architecture has been considered:
A/ Minimalist?
B/ Rudimentary?
C/ Exceptionally complex?*

2.2) At the moment of the arrival of the first European settlers, native Americans had already developed:
A/ Their own architectural traditions?*
B/ Their own public transportation system?
C/ Transmission lines for electricity?

2.3) How were the New England and Spanish Colonies grouped in?
A/ In towns, close to a central plaza or meetinghouses*
B/ In big cities
C/ They were not organized at all

2.4) By the time of the Revolution of the American colonies, who built most of the home structures at the time?
A/ Journeymen carpenters*
B/ Skilled architects
C/ Civil engineers

2.5) Looking to the ideals of social reform, he believed that the right environment could uplift minds and promote civic virtue:
A/ Abraham Lincoln
B/ Frank Lloyd Wright
C/ Thomas Jefferson*

2.6) Despite the absence of any form of public regulation, distinct districts appeared, including which of the following developments:
A/ Official buildings?
B/ Elegant blocks of colonnaded houses, multifamily tenements, and lodging houses?*
C/ Amusement parks?

2.7) Which of the following is part of the first instances of American architecture that influenced the European design of the time?
A/ The Golden Gate Bride in San Francisco
B/ The John Haviland's Eastern State Penitentiary in Philadelphia*
C/ The Eiffel Tower in France

2.8) Parks provided a respite within the new commercial cities. Who elaborated the idea of landscaped "parkways" to connect coordinated park system such as the Emerald Necklace of the city of Boston?
A/ Frederick Law Olmsted*
B/ Francis Allen
C/ Samuel Baker

2.9) Beginning in the 1880s, business districts were transformed by another architectural innovation as the product of new technological advances and artistic creation. It was:
A/ The skyscraper?*
B/ The subway system
C/ Four-lane avenues

2.10) The contrast among a diversity of buildings soon produced a new way of experiencing cities, captured in the term *skyline,* first used in:
A/ In 1896 to describe Lower Manhattan?*.
B/ In 1945 to describe Chicago's financial district?
C/ In 1978 to describe downtown San Francisco?

* Correct answers

Appendix E: Comprehension Tests

COMPREHENSION TEST
Armed Forces
Fragment by Weigley, Russell F.

Participant #: ___
Session: A/B

3.1) The U.S. Army and preceding colonial forces have been tools for domestic tasks like:
A/ Protect settlements against the North American Indians*
B/ Protect the U.S. against natural disasters
C/ Persuade citizens to vote in the presidential elections

3.2) The military institutions of the United States have their roots in:
A/ The Indian militia?
B/ The Arabian militia?
C/ The militia systems of British colonies?*

3.3) Deficiencies of part-time soldiery for extended campaigns led to supplementing the militia with the British regular army, beginning in:
A/ 1755?*
B/ 1975?
C/ 1850?

3.4) For defense against foreign enemies, it was intended that the militia would be mobilized to reinforce:
A/ Troops at the regular Army?*
B/ Peacekeepers of the United Nations?
C/ Ships of the U.S. Coast Guard?

3.5) The U.S. Military Academy at West Point, founded in 1802, was essentially an engineering curriculum a solid foundation of schooling in:
A/ Paralegal
B/ Tactical practice and strategic thought*
C/ Field medicine

3.6) The West Point influence encouraged officers to continue their studies after graduation, and some visited Europe to:
A/ Get married with European women
B/ Observe European Armies*
C/ Steal warfare equipment from the Nazis

3.7) He was a Commanding General of the Army from 1869 to 1883, a vigorous exponent of military education who encouraged the formation of schools:
A/ William Tecumseh Sherman*
B/ Martin Luther King
C/ Collin Powell

3.8) The Navy also took steps toward refining professionalism. At the U.S. Naval Academy, the technical mastery of seamanship took priority over military study because:
A/ It was more important in order to fight terrorism
B/ They had little need of military training by the time*
C/ They will never participate in a conflict

3.9) The Naval War College at Newport, Rhode Island, was established in 1884 after:
A/ The attack at Pearl Harbor
B/ The great depression
C/ The blockading and capturing of Confederate ports*

3.10) The War College helped transform the Navy from a loose collection of individual ships into:
A/ An unorganized fleet of vessels
B/ Coordinated squadrons organized around battleships which defended the seas against potential enemies*
C/ The new merchant marine.

* Correct answers

Appendix F: Data Collection Form

Participant #: ____________________
Date: ____________________

Sub-session I Results

Sub-session I-A	Sub-session I-B	Sub-session I-B
Name of the passage	Name of the passage	Name of the passage
Words recalled:	Words recalled:	Words recalled:
Notes from sub-session I		

Sub-session II Results

Comprehension test scores

Sub-session II-A (ANA)				Sub-session II-A (ANP)				Notes from sub-session II
Q1		Q6		Q1		Q6		
Q2		Q7		Q2		Q7		
Q3		Q8		Q3		Q8		
Q4		Q9		Q4		Q9		
Q5		Q10		Q5		Q10		

Sub-session III Results

Task A	Task B	Task C
Time-on-task (in minutes):	Time-on-task (in minutes):	Time-on-task (in minutes):
Task completed (Yes/No):	Task completed (Yes/No):	Task completed (Yes/No):
Notes	Notes	Notes

Appendix F: SJSU-IRB Human Subjects Approval Letter

San José State
UNIVERSITY

Office of the Academic Vice President

Academic Vice President Graduate Studies and Research

One Washington Square
San José, CA 95192-0025
Voice: 408-283-7500
Fax: 408-924-2477
E-mail: gradstudies@sjsu.edu
http://www.sjsu.edu

To: Jose R. Concepcion
701 Curtner Avenue, #339
San Jose, CA 95125

From: Pam Stacks, Pam Stacks
Interim AVP, Graduate Studies & Research

Date: September 2, 2004

The Human Subjects-Institutional Review Board has approved your request to use human subjects in the study entitled:

"ASSISTIVE TECHNOLOGIES BASED ON SYNTHETIC SPEECH: INTERFACE COMPATIBILITY AND ADAPTIVITY."

This approval is contingent upon the subjects participating in your research project being appropriately protected from risk. This includes the protection of the anonymity of the subjects' identity when they participate in your research project, and with regard to all data that may be collected from the subjects. The approval includes continued monitoring of your research by the Board to assure that the subjects are being adequately and properly protected from such risks. If at any time a subject becomes injured or complains of injury, you must notify Pam Stacks, Ph.D. immediately. Injury includes but is not limited to bodily harm, psychological trauma, and release of potentially damaging personal information. This approval for the human subjects portion of your project is in effect for one year, and data collection beyond September 2, 2005 requires an extension request.

Please also be advised that all subjects need to be fully informed and aware that their participation in your research project is voluntary, and that he or she may withdraw from the project at any time. Further, a subject's participation, refusal to participate, or withdrawal will not affect any services that the subject is receiving or will receive at the institution in which the research is being conducted.

If you have any questions, please contact me at (408) 924-2480.

cc: Kevin Corker, Ph.D.

Appendix M: Tables

Table 1a.

Mean Number of Correct Words Recalled as a function of Synthetic Speech Rate

	Number of words correctly recalled		
Synthetic Speech Rate	n	M	SD
40 words per minute	9	23.66	8.60
60 words per minute	9	21.44	9.72
80 words per minute	9	22.66	10.44

Table 1b.

An Analysis of Variance (ANOVA) Summary Table

Source	*SS*	*df*	*MS*	*F*
Speech Rate	23.41	2	11.70	.44*
Subjects	1659.29	8	207.41	
Error	423.94	16	26.44	
Total	2106.63	26		

*P>.05

Table 1c.

Mean Number of Errors as a Function of Synthetic Speech Rate

	Number of errors		
Synthetic Speech Rate	n	M	SD
40 words per minute	9	6.78	6.85
60 words per minute	9	11.44	11.21
80 words per minute	9	12.55	11.59

Table 1d.

An Analysis of Variance (ANOVA) Summary Table

Source	*SS*	*df*	*MS*	*F*
Speech Rate	119.19	2	59.59	1.34*
Subjects	2239.85	8	279.98	
Error	708.15	16	44.25	
Total	3067.19	26		

*P>.05

Table 2a.

Mean Comprehension Levels as a Function of Changes in the Listening Conditions

	Comprehension Test Scores		
Listening conditions	n	M	SD
Ambient noise absent	9	80.00	15.00
Ambient noise present	9	62.22	25.87

Table 2b.

An Analysis of Variance (ANOVA) Summary Table

Source	*SS*	*df*	*MS*	*F*
Comprehension	272.22	1	272.22	2.22*
Subjects	2177.78	8	272.22	
Error	977.78	8	122.22	
Total	3427.78	17		

*P>.05

Table 3a.

Mean Task Efficiency as a Function of Three Different Tasks Performed

	Time-on-task (TOT)		
Task performed	n	M	SD
Editing	9	2.14	1.22
Transferring	9	4.95	2.95
Searching	9	2.35	1.62

Table 3b.

An Analysis of Variance (ANOVA) Summary Table

Source	*SS*	*df*	*MS*	*F*
Task efficiency	44.19	2	22.09	6.96*
Subjects	51.91	8	6.48	
Error	50.77	16	3.17	
Total	146.87	26		

*P<.05

Appendix I: Figures

Figure 1.

Information Processing Outcomes Based on Changes at Synthetic Speech Rate: Mean and Standard Deviation of Number of Words Correctly Recalled

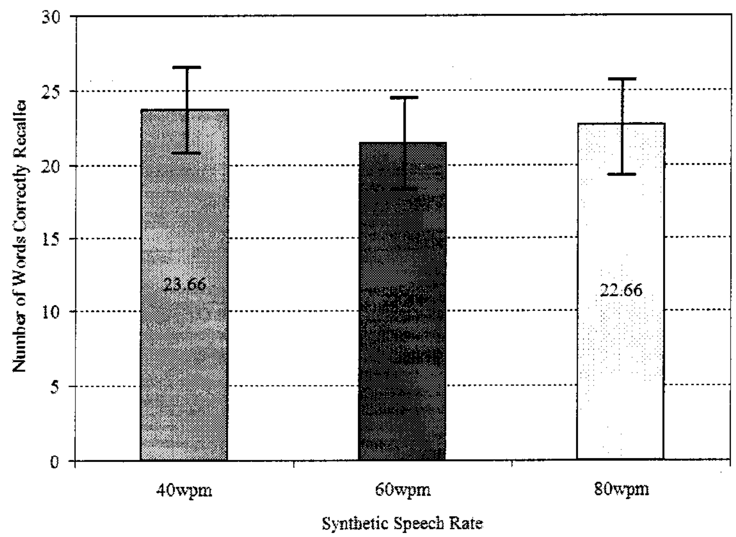

Figure 2.

Information Processing Outcomes Based on Changes at Synthetic Speech Rate: Mean and Standard Deviation of the Number of Errors

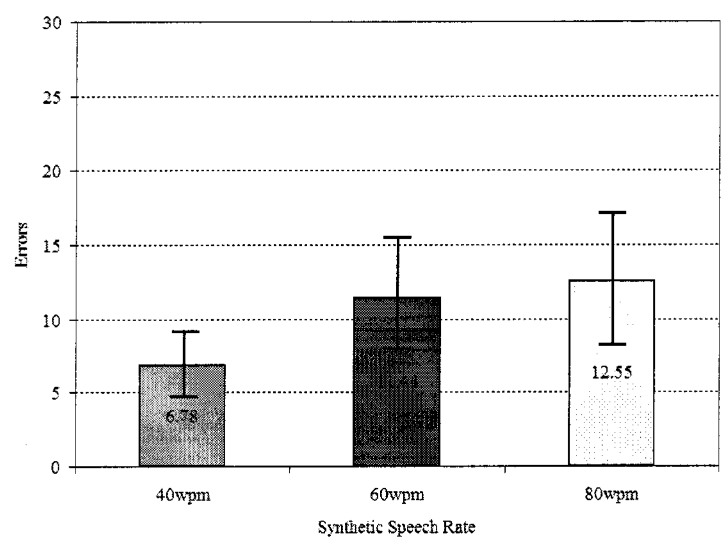

Figure 3.

Comprehension of Synthetic Speech Based on Different Listening Conditions: Mean and Standard Deviation of Comprehension Test Scores

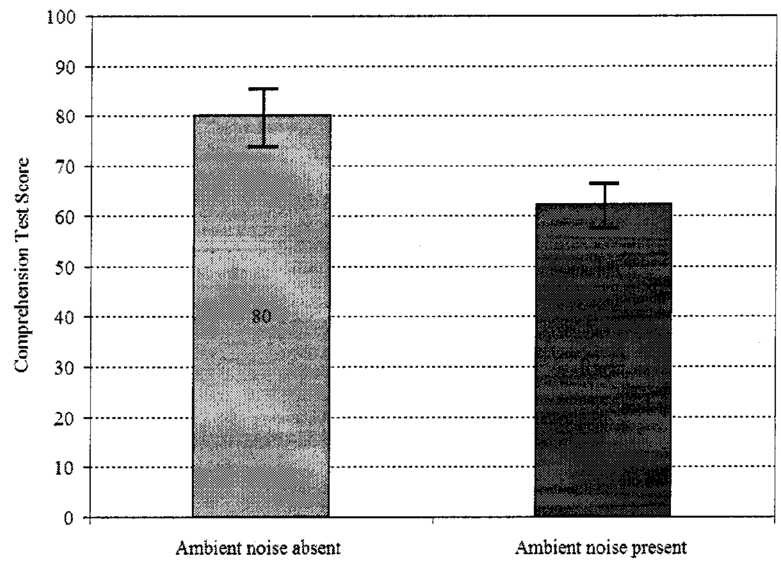

Figure 4.

Mean and Standard Deviation of Time-On-Task at the Editing, Transferring, and Searching Tasks

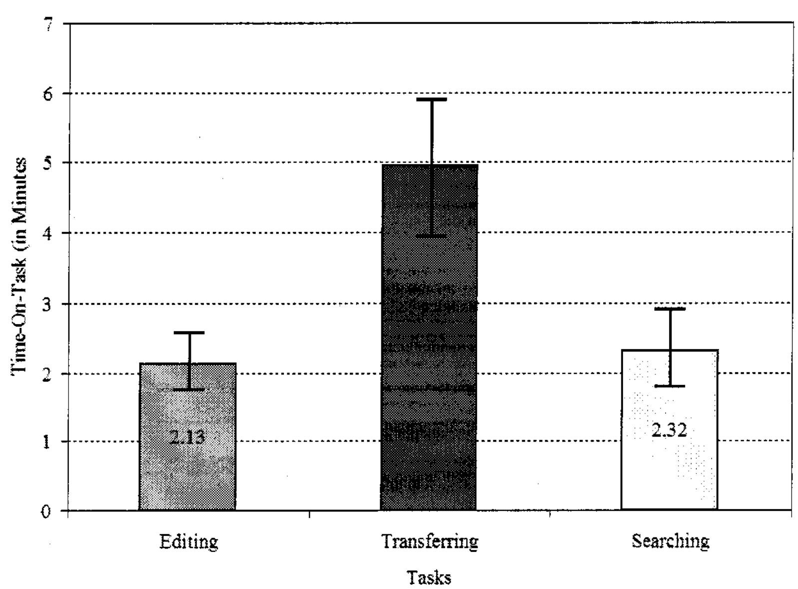

Figure 5. Sequence for the “Editing” Task used at Experimental Session III

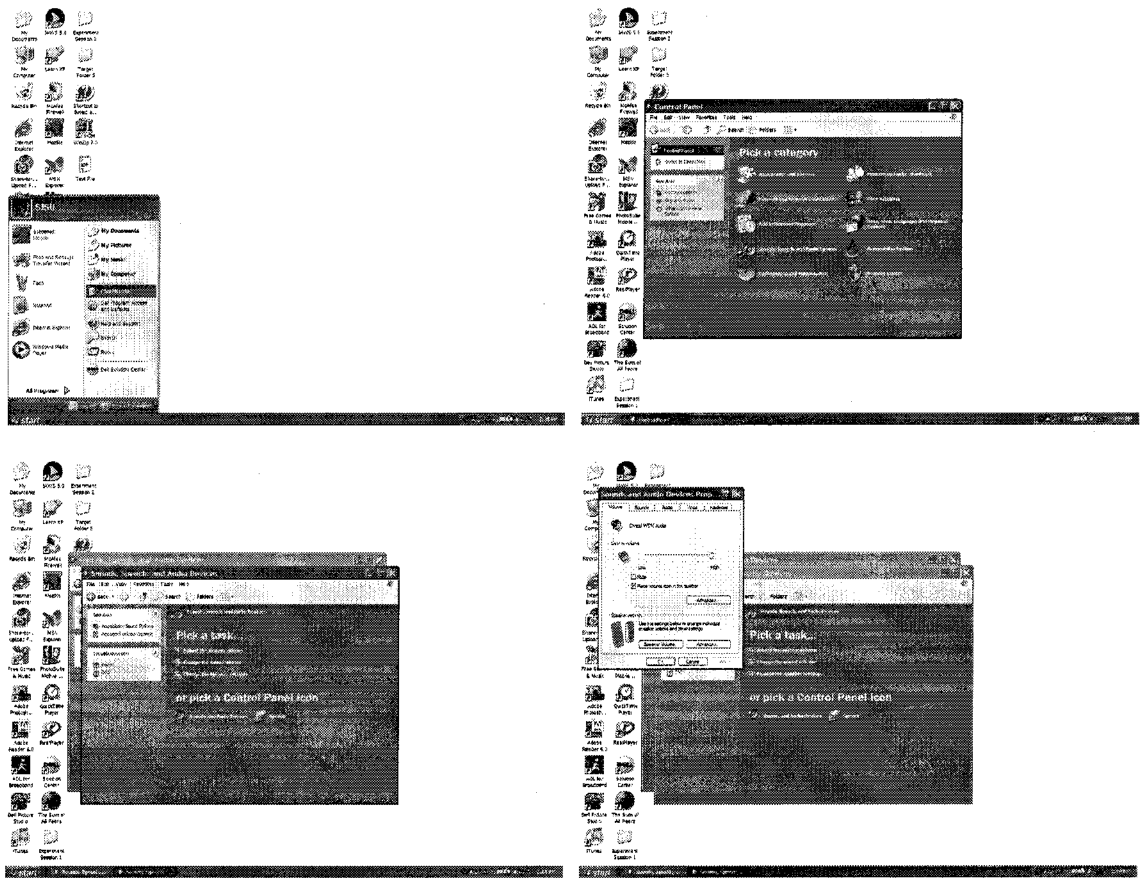

Figure 6. Sequence for the "Transferring" Task used at Experimental Session III

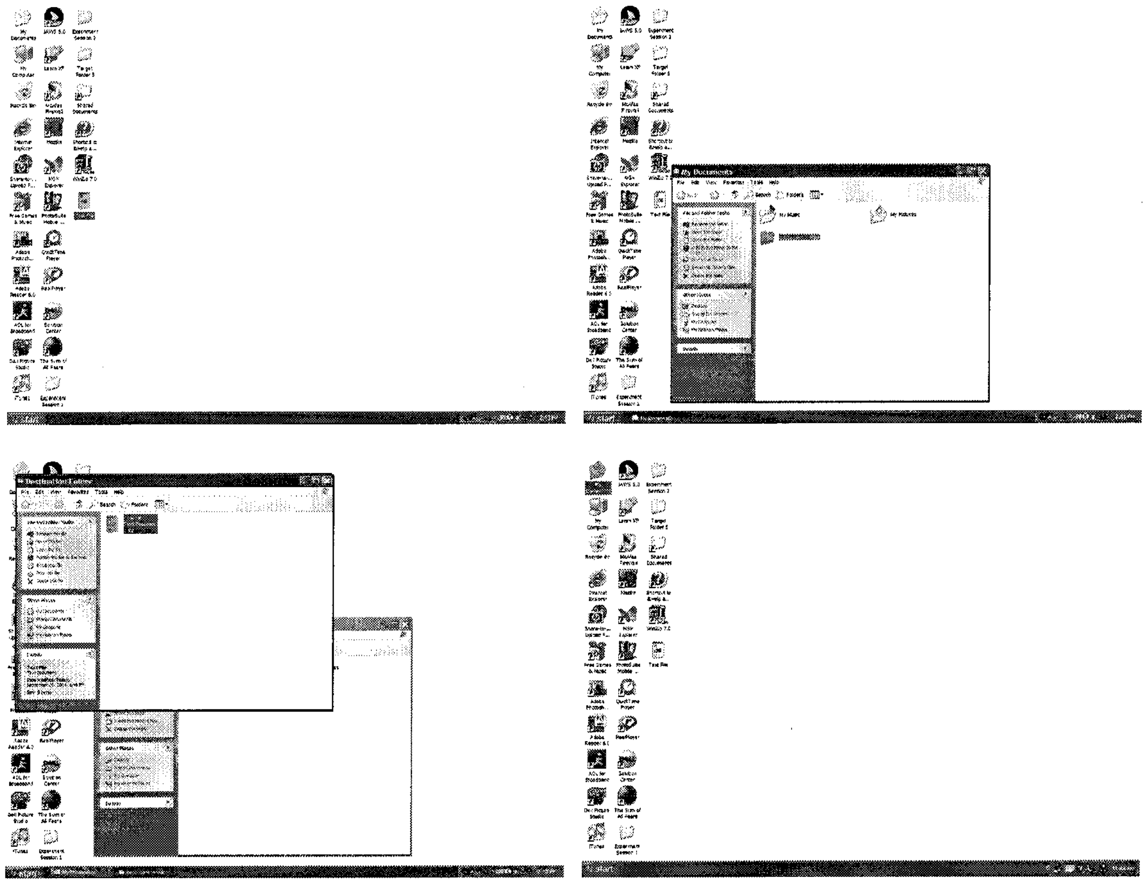

Figure 7. Sequence for the "Searching" Task used at Experimental Session III