DESIGN FOUNDATIONS FOR CONTENT-RICH ACOUSTIC INTERFACES: INVESTIGATING AUDEMES AS REFERENTIAL NON-SPEECH AUDIO CUES

Mexhid Adem Ferati

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the School of Informatics,
Indiana University

August 2012

Accepted by the Faculty of Indiana University, in partial
fulfillment of the requirements for the degree of Doctor of Philosophy.

Mark Pfaff, Ph.D., Chair

Davide Bolchini, Ph.D.

Doctoral Committee

Mathew Palakal, Ph.D.

June 26, 2012

Amy Shirong Lu, Ph.D.

DEDICATION

To my parents.

ACKNOWLEDGMENTS

It is a pleasure to thank those who have made this dissertation possible. First and foremost, I would like to acknowledge the contribution of my advisor, Davide Bolchini, without whose help and guidance it would have been next to impossible to write this dissertation. He has been more than an advisor; he has also been a friend who has elevated my confidence in the days when it was hard to see a light at the end of the tunnel. I would like to show my gratitude to Steve Mannheimer and Mathew Palakal for introducing me to the research of sounds, which motivated my interest in research and gave me the opportunity to pursue a career in academics. Also, I am truly indebted and thankful to Mark Pfaff and Amy Shirong Lu for guiding me throughout the process of completing this dissertation, and for their cooperative spirit and valuable suggestions.

I am obliged to many of my colleagues at the School of Informatics USER Lab who have supported me on this magnificent journey. Their comments and suggestions have sparked many ideas that are part of this dissertation. Especially, I would like to thank Tao Yang and Romisa Rohani Ghahari for taking the time to listen, even when I was just complaining.

Special thanks goes to the students and staff of the Indiana School for the Blind and Visually Impaired, whose collaboration has strengthen my ideas and made me believe in my research. I would also like to acknowledge the support of the Nina Mason Pulliam Charitable Trust and the National Science Foundation.

Last, but not least, I would like to thank my family and friends for their constant source of encouragement and moral support. I thank them all as regardless of their whereabouts—whether close or thousands of miles away—they were always next to me. I owe sincere and earnest thankfulness to my wife, Regaipe, for her love and patience. I hope she will learn to love this dissertation, which was a source of loneliness to her for many years.

ABSTRACT

Mexhid Adem Ferati

DESIGN FOUNDATIONS FOR CONTENT-RICH ACOUSTIC INTERFACES: INVESTIGATING AUDEMES AS REFERENTIAL NON-SPEECH AUDIO CUES

To access interactive systems, blind and visually impaired users can leverage their auditory senses by using non-speech sounds. The current structure of non-speech sounds, however, is geared toward conveying user interface operations (e.g., opening a file) rather than large theme-based information (e.g., a history passage) and, thus, is ill-suited to signify the complex meanings of primary learning material (e.g., books and websites). In order to address this problem, this dissertation introduces audemes, a new category of non-speech sounds, whose semiotic structure and flexibility open new horizons for facilitating the education of blind and visually impaired students.

An experiment with 21 students from the Indiana School for the Blind and Visually Impaired (ISBVI) supports the hypothesis that audemes increase the retention of theme-based information. By acting as memory catalysts, audemes can play an important role in enhancing the aural interaction and navigation in future sound-based user interfaces. For this dissertation, I designed an Acoustic EDutainment INterface (AEDIN) that integrates audemes as a way by which to vividly anticipate text-to-speech theme-based information and, thus, act as innovative aural covers. The results of two iterative usability evaluations with total of 20 blind and visually impaired participants showed that AEDIN is a highly usable and enjoyable acoustic interface.

Yet, designing well-formed audemes remains an ad hoc process because audeme creators can only rely on their intuition to generate meaningful and memorable sounds. In order to address this problem, this dissertation presents three experiments, each with 10 blind and visually impaired participants. The goal was to examine the optimal combination of audeme attributes, which can be used to facilitate accurate recognitions of audeme meanings. This work led to the creation of seven basic guidelines that can be used to design well-formed audemes. An interactive application tool (ASCOLTA: Advanced Support and Creation-Oriented Library Tool for Audemes) operationalized these guidelines to support individuals without an audio background in designing well-formed audemes. An informal evaluation conducted with three teachers from the ISBVI, supports the hypothesis that ASCOLTA is a useful tool by which to facilitate the integration of audemes into the teaching environment.

Mark Pfaff, Ph.D., Chair

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

In our everyday experiences, we recognize sounds and have the ability to associate them powerfully with an object or event. For instance, hearing a short song on a bus or in a restaurant may bring back memories of the setting in which you heard the song for the first time. Sometimes, strong emotions are elicited by the song, which can make us holistically recall the same setting. Sounds have the ability to instantaneously elicit memories of events, places or people of a recent or distant past.

Given that sounds facilitate the recall of events, what would be the benefit of taking advantage of this ability to memorize educational content? Memorizing content is the most important aspect of learning. Typically, in an education setting, learning takes place by reading textbooks and memorizing their content for the purposes of recognition (e.g. multiple choice questions) and recall (e.g. essay type questions).

The blind and visually impaired are a community that would benefit from using sounds for content memorization. Traditionally, education has been very vision-centric and the consumption of digital information through computers is typically based on the Graphical User Interface (GUI). This puts the blind and visually impaired (BVI) community at a serious drawback as they mainly rely on the auditory channel to consume digital content. This tradition is likely to continue, as we recently witnessed the introduction of the newest version of iBooks (http://www.apple.com/education/ibooks-textbooks); a very vision-centric consumption of content, which promises to be the future of textbooks.

Of the approximately 448,000 school-age Americans with blindness or severe visual impairment (Adams, Hendershot, & Marano, 1999), only 19% progress through the educational levels required to pursue professional careers (Wagner, Newman, Cameto, Garza, & Levine, 2005). Most blind and visually impaired students are overwhelmed by the logistical and emotional obstacles that confront them in standard, vision-centric educational curricula, textbooks and support materials. This limitation has been particularly true in STEM (science, technology, engineering and math) education, which is heavily dependent upon visual pedagogy. In recent years, there has also been increased international concern for improving access of the blind and visually impaired to and engaging in STEM education, especially in K-12 grades. A government study in

Australia cited considerable evidence that resources intended at stimulating student engagement in STEM need to capture the imagination of students in the K-12 level grades (Tytler, Osborne, Williams, Tytler, & Clark, 2008).

Being aware of the limitations that GUI presents to the education of the blind and visually impaired community, the importance of utilizing sounds in GUI computer user interfaces has been continuously raised by researchers for the past three decades (Bly, 1982; Lunney et al., 1983; Mansur, Blattner, & Joy, 1985; Mezrich, Frysinger, & Slivjanovski, 1984; Brewster, 1994; Brewster, Wright, & Edwards, 1995; Conversy, 1998; Edwards, 1989; Kantowitz & Sorkin, 1983; Sanders & McCormick, 1987). For example, Walker, Nance and Lindsay (2006) argued that, with the increasing number of mobile devices where the screen real estate is significantly reduced, the auditory channel has begun to assume much greater importance. This study suggests that sound can be leveraged to present information in devices with limited display. Not limited to only portable devices, Brewster, Wright, and Edwards (1993) pointed out the potential of non-speech sounds as a useful complement to visual output in order to increase the amount of information communicated to the user or to reduce the stress on the user in regard to focusing his attention on smaller or non-stationary visual targets.

Non-speech sounds are computer-generated synthetic sounds used to represent objects and events in a user interface by mimicking sounds from nature or abstract structured tones. Inspired by daily events, such as people’s ability to recognize the difference in pitch when automobiles are close to a breakdown, researchers have manipulated sound dimensions, such as pitch, timbre and register, to convey messages about activities or tasks in user interfaces (Brewster, 1994). Other researchers (Gaver, 1989) have used natural sounds to map the interpretation of natural events to computer events, such as the sound of a crumbling paper to represent when a file is being deleted. These scenarios indicate examples in which a sound from the nature is used as a metaphor to represent similar actions in a computer user interface.

1.1 Problem Statement

The introduction of non-speech sounds and their usage in user interfaces have made computers more accessible to the blind and visually impaired community.

Consequently, blind and visually impaired users can efficiently utilize the auditory channel for the purpose of consuming digital content. Such a practice, that emphasizes the importance of designing an artifact based on a user's ability rather than disability, is known as the ability-based design principle (Wobbrock, Kane, Gajos, Harada, & Froehlich, 2011). This paradigm maintains that a design should be driven by a user's capabilities. In the case of blind and visually impaired users, who typically have heightened auditory performance, ability-based design calls for taking advantage of this dimension. Utilizing non-speech sounds is one step in this process.

Non-speech sounds could be used to deliver digital content in educational settings. Educational content is typically organized around a single topic or theme, which I refer to as theme-based content. Such an example could be an essay of around 500 words discussing the central theme of slavery from U.S. history books. Non-speech sounds could be used to signify the complex meanings of theme-based information. The current structure of non-speech sounds, however, is geared toward conveying user interface operations (e.g. opening a file) rather than large theme-based information and, thus, is ill-suited to signify the complex meanings of primary learning material (e.g., books and websites). In order to address this problem, this dissertation introduces a new category of non-speech sounds whose semiotic structure and flexibility open new horizons for facilitating the education of the blind and visually impaired students. Complementary to the role of conveying theme-based content, these novel non-speech sounds could be used in content-rich interfaces. These are the types of interfaces that typically contain theme-based information, which could be delivered visually (e.g. reading) for sighted users or aurally (e.g. text-to-speech) for blind and visually impaired users.

While diverse types of non-speech sounds exist, no guidelines exist specific to their design. Therefore, the design of non-speech sounds has been done based on intuition rather than following any theoretical framework (Pirhonen, Murphy, McAllister, & Yu, 2007). Similarly, researchers have argued that sound designers should not only rely on a random selection of arbitrary sounds, but, instead, should create and follow guidelines that could guide sound designers to generate sounds (Mustonen, 2008). Mynatt (1994) claims that the process of designing non-speech sounds used in a user interface is more of an art than a science. Lumsden and Brewster (2001) pointed out that the fact that designers do not have guidelines for the use of sound is one reason why sound lags behind in the competition of its implementation in user interfaces.

1.2 Dissertation Goal

The goal of this dissertation is to address issues in three different areas. First, it investigates a new category of non-speech sounds – audemes – intended to convey educational theme-based content through content-rich interfaces. This study of audemes will present an effort to mitigate the challenges that the blind and visually impaired community face in the visual-centric education environment. Consequently, audemes will ease access to theme-based educational content. Second, this dissertation empirically derived design guidelines for generating well-formed non-speech sounds, specifically audemes. Finally, this dissertation demonstrates ways and tools by which to support the process of creating well-formed non-speech sounds for use in content-rich interfaces.

The invention of audemes will fill a gap in the variety of non-speech sounds available. Additionally, devising guidelines for the effective design of audemes will make them the first type of non-speech sounds with empirically grounded design guidelines. Further, these design guidelines will be operationalized in an interface tool that will enable users without any knowledge in sound design to create effective audemes.

1.3 Research Questions

The research questions for this dissertation are as follows:

Q1: When a new category of non-speech sounds (audeme) is played along with theme-based information, does it help the user to better memorize the theme-based content?

Q2: What is the function of audemes in content-rich interfaces for the blind and visually impaired?

Q3: What characteristics of audemes help blind and visually impaired users recognize audeme meaning?

1.4 Research Aims

Three aims guide this research:

Aim 1: To introduce audemes, a new category of non-speech sound for content-rich interactions. Audemes are based on two theoretical frameworks: semiotics and modes of listening. The advantages of audemes when compared to other non-speech sounds will be investigated through empirical studies and a novel acoustic interface prototype.

Aim 2: To define the characteristics of the new category of non-speech sounds. The principles used when defining the perceptual/cognitive characteristics of audemes will be investigated. Guidelines on designing effective audemes will be identified through a set of empirical studies involving blind and visually impaired participants.

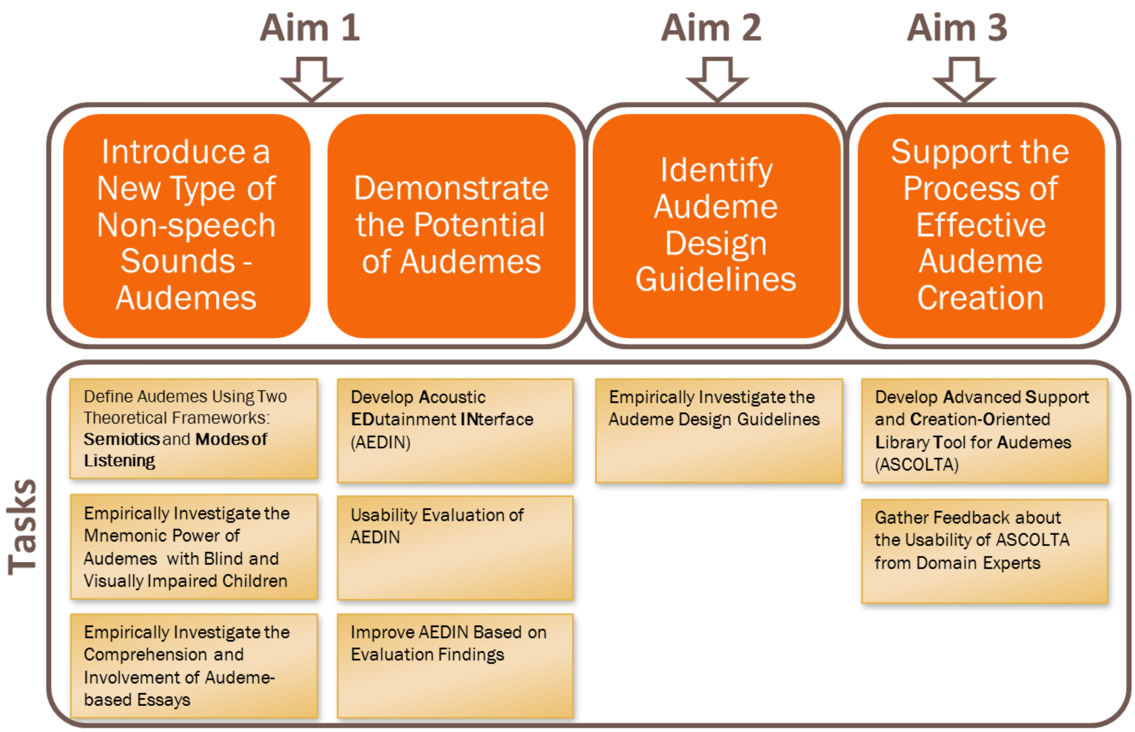

Figure 1.1. Dissertation aims and tasks.

Aim 3: To support the process of the effective generation of non-speech sounds. Based on the guidelines defined in Aim 2, a tool will be designed and evaluated that will operationalize the guidelines. These guidelines will be used to guide teachers in K-12 schools for the blind and visually impaired in generating effective non-speech sounds. These guidelines will facilitate the process of creating and utilizing non-speech sounds for use in education and acoustic interfaces.

An overview and detailed description of the aims is shown in Figure 1.1.

The overall contribution of this dissertation is that it represents the first work that demonstrates audemes as a new category of non-speech sounds aimed at conveying theme-based information. Additionally, audemes are presented here as the first type of non-speech sound, the creation of which is based on guidelines empirically derived from numerous experiments. Finally, the operationalization of the audeme design guidelines into a tool interface contributes to making audemes easily accessible and able to be designed by individuals without any audio background. This dissertation will discuss seven experiments conducted to achieve aims.

1.5 Significance of the Dissertation Contribution

The significance of this research lies in broadening the use of content-rich non-speech sounds by making them easier to design. Lessons from this research should directly benefit user interface designers who need to implement non-speech sounds into their applications. Particularly, I believe this research will be beneficial to the education of the blind and visually impaired community by making their user interfaces more accessible and utilizing the audio channel to a higher potential.

1.6 Overview of the Dissertation

The remaining chapters are organized as follows:

Chapter 2, *Review of background,* gives an overview of the research previously conducted in investigating knowledge acquisition models as part of education and the different types of non-speech sounds that exist. Understanding existing non-speech sounds is important for the development of a new category of non-speech sound aimed at effectively communicating theme-based content to the blind and visually impaired community. Also, an overview of the research previously completed in regard to designing acoustic interfaces is presented as a way by which to demonstrate the characteristics and functions of different non-speech sounds.

Chapter 3, *Theoretical frameworks: Audemes and models of non-speech sounds,* explores two existing theoretical frameworks that were used to model the characteristics of the new category of non-speech sounds. The semiotic approach was used as the theoretical framework by which to model the non-speech sounds and analyze the process of meaning generation by treating them as audio signs. The listening modes theory was used to define non-speech sounds in terms of the four listening modes established in the audio community. Finally, the audemes are introduced as a novel category of non-speech sounds.

Chapter 4, *Understanding the potential of audemes for content recognition and signification*, empirically explores the characteristics of the audemes as innovative non-speech sounds. Using the empirical findings from two experiments, this chapter highlights the strengths of the audemes when heard along with theme-based content. The participants for the experiments included blind and sighted participants.

Chapter 5, *Audeme-based acoustic interface: AEDIN*, demonstrates an Acoustic Edutainment Interface (AEDIN), which was developed in order to validate the power of audemes into a real application. Usability testing using blind and visually impaired participants with two versions of AEDIN was conducted. The key design principles are highlighted as part of the findings from the usability evaluation.

Chapter 6, *Toward non-speech sound design guidelines for well-formed audemes*, describes three experiments conducted in order to investigate the guidelines by which to create effective audemes. The first experiment investigates how audemes perform over time in terms of content recognition. The second experiment explores information recognition using audemes without reinforcement, while experiment three examines the audeme cognitive load. The outcome of these experiments is a set of guidelines that need to be followed in order to create well-formed audemes.

Chapter 7, *Advanced support and creation-oriented library tool for audemes (ASCOLTA)*, demonstrates a tool developed to guide the teachers from a K-12 school for the blind and visually impaired in generating effective non-speech sounds. The guidelines established in Chapter 6 are internalized into a tool that will be used to facilitate the process of creating well-formed audemes. The findings of an evaluation conducted with teachers at the Indiana School for the Blind and Visually Impaired (ISBVI) are discussed.

2. REVIEW OF BACKGROUND

In order to understand the important elements in the process of knowledge acquisition as part of the education, I initially investigated existing models serving this purpose. Thus, in this section, I review previous literature concerning types of learning, the role of non-speech sounds in the learning process and types of non-speech sounds.

2.1 The Role of Sounds in Multimedia Learning

The process of learning with multimedia material (Mayer & Moreno, 2003) has been conceptualized by contrasting verbal and non-verbal representations as elements which are encoded and stored in different substructures of short- and long-term memory (Clark & Paivio, 1991; Kosslyn, 1994; Baddeley, 1999). Verbal representations, such as words, are encoded in the form of propositional representations, while non-verbal representations, such as pictures, are encoded and stored in analogical representations. This phenomenon is explained by the dual-coding principle, which assumes that each channel in the human information-processing system has a limited capacity. Because only a limited quantity of mental processing can occur either in the verbal or non-verbal medium at any one time, learning is positively influenced by introducing words and non-verbal illustrations jointly (Bernard, 1990; Glenberg & Langston, 1992; Guri-Rozenblit, 1988; Waddill, McDaniel, & Einstein, 1988).

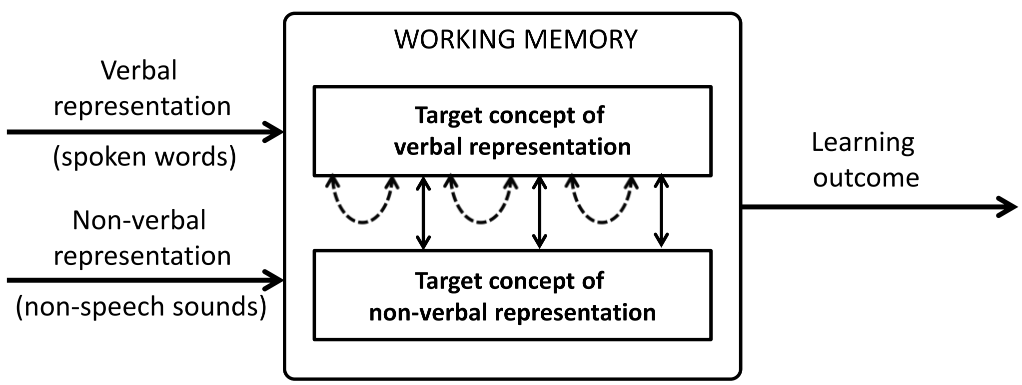

Figure 2.1. A dual-coding model of multimedia learning (adapted from Mayer and Moreno (1998)).

By replacing the visual channel with the auditory channel, *non-speech sounds* act as a type of non-verbal representation that can be leveraged by the blind and visually impaired community to enhance the learning process. In fact, as we integrate non-speech sounds in the dual-coding model (see Figure 2.1), the learning outcome can be considered as formed through an interaction between the target concepts of the verbal (spoken words) and non-verbal representations (non-speech sounds). Thus, the generated balance of the cognitive load between the target concepts of both verbal and non-verbal representations can help memorize concepts more effectively for those who have no access to visual content. Based on this idea, I have investigated different types of non-speech sounds that could help blind and visually impaired individuals learn educational content.

2.2 Non-speech Sounds

Non-speech sounds have been researched for more than 30 years in the field of human factors and later in human-computer interaction (Deatherage, 1972; Kantowitz & Sorkin, 1983; Sanders & McCormick, 1987). Different types of non-speech sounds have been developed by different researchers to represent objects and events in user interfaces, whether as the sole medium or in conjunction with graphics. They have been used as the main channel, along with text-to-speech (TTS) in aural applications, when serving the needs of the blind and visually impaired community (Parente & Bishop, 2003). As an additional communication channel, non-speech sounds have been used to complement visual interfaces in semi-aural or eyes-free context applications, such as driving or walking (Li, Baudisch, & Hinckley, 2008). In these contexts, sounds can serve as a valuable asset by which to communicate information when the visual channel is limited or non-existent.

Other researchers have found that sound or music can be a powerful cue for memory enhancement (Sánchez & Flores, 2003) and, to this extent, sound has been used to teach blind children math (Sánchez & Flores, 2005; Stevens, Brewster, Wright, & Edwards, 1994). In this area, non-speech sounds have also been used as reminders or notifications delivered in the home environment to inform individuals when to take medication or about an upcoming appointment (McGee-Lennon, Wolters, & McBryan 2007; McGee-Lennon, Wolters, McLachlan, Brewster, & Hall, 2011).

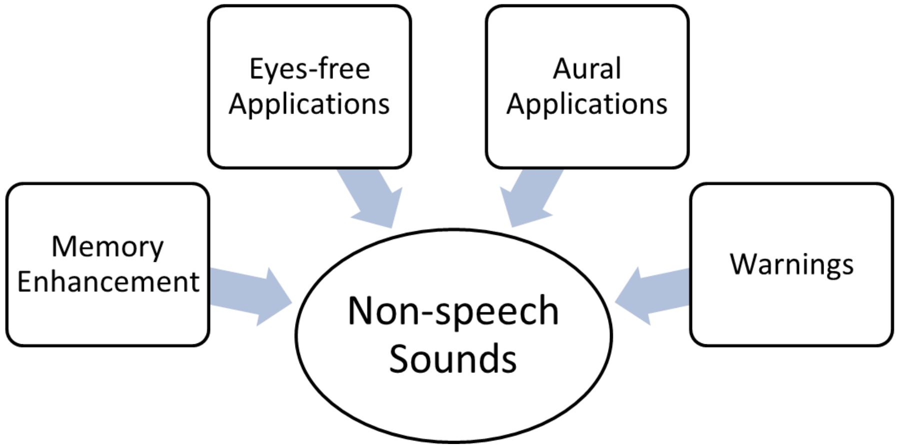

Figure 2.2. Application domains for non-speech sounds.

Of course, symbolic sounds have been present for a long time. For example, sirens and alarms have been used to communicate warnings and suggest an increase in attention to important information or events. Likewise, in computer user interfaces, sounds have been used to get the user's attention to an event or object on the screen. Using sounds to indicate event failures (e.g., a file failing to be saved); however, is insufficient because of the amount of information that is left unexpressed (e.g., what prevents the file from being saved). Consequently, an effort needs to be made to use sounds to communicate more complex information to complement the visual channel, or provide the most accurate interaction to communities (such as the blind or visually impaired) that can only utilize the audio channel. This need was noted by different researchers who have taken different approaches in regard to implementing and evaluating non-speech sounds in computer interfaces.

The use of sounds in these different circumstances (Figure 2.2) suggests the increasing importance of sounds and efforts to perfect their communication abilities. Some of the most established types of non-speech sounds will be briefly reviewed in the next section.

2.3 Types of Non-speech Sounds

In this section, I systematically review the structure of the most established and important types of non-speech sounds as identified in the literature.

Earcons are a type of non-speech sounds used in computer applications. They provide feedback and information to the user about objects and events within the interface. They are created using brief, rhythmic sequences of pitches with variable timbres, registers and intensities. Blattner, Sumikawa and Greenberg (1989) explained that earcons can "describe" objects (e.g., files and menus) and operations (editing and executing) by conventionally assigning to them a meaning that must be memorized. In a user interface, earcons are used as a technique by which to add context to a menu, helping users understand in which level of the menu they are currently positioned. For example, earcons could deliver navigational cues in a telephone-based interface (Brewster, 1998; Brewster, Raty, & Kortekangas, 1996).

Earcons need to be learned and remembered in order to be useful. Brewster, Wright and Edwards (1993) conducted an evaluation of earcons used with desktop computer and menu screens and established that earcons present an effective resource of communication between the user and computer interface. Earcons have demonstrated superior performance when compared to other less systematic sounds (Brewster, Wright, & Edwards, 1994) and when compared to no sound at all used on the desktop (Brewster, 1997) and in mobile (Leplatre & Brewster, 2000; Brewster & Cryer, 1999) devices.

Auditory icons are another type of non-speech sound that replicate or utilize actual recordings of every day sounds, such as a barking dog, a car passing by, a glass breaking, etc. Auditory icons are used to represent information about events and actions in a user interface by similarity with everyday natural events. They were developed by Gaver (1986) for two auditory interfaces: SonicFinder (Gaver, 1989) and ARKola (Gaver, Smith, & O'Shea, 1991). As auditory icons are sounds from the environment, they have a connection to the objects or actions that produced them, as well as objects or actions that may be easily perceived as semantically connected. This attribute makes auditory icons potentially easy to recognize and learn by users.

The relevance of auditory icons lies in their similarity of sounding just like the real object it represents, making them appropriate to be used in applications where the easy learnability of the relationship between the sound and object is critical. A limitation, however, is the difficulty to create a proper auditory icon for menu items, such as *search*, because no sound exists in the everyday environment that will inherently represent the notion of a search (Palladino & Walker, 2007; Palladino & Walker, 2008). Successful auditory systems for visually impaired users were designed in order to enable efficient navigation through a hypermedia system using 24 commands for which auditory icons or earcons were associated (Morley, Petrie, & McNally, 1998). An extensive overview of earcons and auditory icons is provided by Absar and Guastavino (2008).

Spearcons are a type of non-speech sound generated by speeding up a spoken word of the object or event that we want to represent until it is not recognized as speech anymore. First, the spoken word is generated as an audio file using TTS, then it is accelerated until it is incomprehensible. Walker et al. (2006), claims that since spearcons are created from the actual name of the object, the mapping is not arbitrary and, as such, learning spearcons should be easier. Spearcons can be understood as audio fingerprints that uniquely identify the parts of information contained in the original words. Spearcons are intended to improve the performance and usability of computer interfaces and make them more accessible to diverse groups of users, such as the sighted and visually impaired.

Walker et al. (2006) have evaluated spearcons in comparison with earcons and auditory icons. In a menu-only interface, the authors instructed the subjects to find items in a 5 x 5 menu set. According to this study, the spearcons resulted in better performances when compared to the earcon and auditory icon results. A recent study (Dingler, Lindsay, & Walker, 2008) compared the learnability of earcons, auditory icons, spearcons and speech and showed that spearcons are as learnable as speech, followed by auditory icons, and earcons.

In recent years, two additional types of non-speech sounds have been introduced: *auditory scrollbars* and *spindex*. *Auditory scrollbars* (Yalla & Walker, 2008) represent sounds that convey information about the length of information presented in a user interface. Similar to how visual scrollbars are used to communicate the length of the information presented on screen, or the number of items presented in a menu (Brewster, 1998), auditory scrollbars present the auditory counterpart that provides an overview of the information presented through auditory cues. Yalla and Walker's (2008) study describes auditory scrollbars as sounds generated by tones varying in pitch played preceding the TTS information of the menu items. This study shows the benefits of auditory scrollbars, through which visually impaired users were able to evaluate the contact lists within mobile phone interfaces. The dissimilarity between the visual and auditory scrollbars is that the former has two functions, display and control, while the latter is used only as display function. Visual scrollbars, as seen in most desktop computer applications, can be used to show the amount of information on the screen, and can also be used to control and navigate through the content information. Auditory scrollbars are only used in the display function, giving the sense of information presented, such as, the length of the content, without any means by which to control it. Recently, an example of visual scrollbars that only use the display function was found in mobile phone interfaces, such as iPhone's Safari browser. *Spindex* (i.e., a speech index), another type of non-speech sound, is generated by coupling an auditory cue with each menu item. The cue represents the utterance of the leading letter of each menu item (Jeon & Walker, 2009). For example, in order to represent an item named "Book," a spindex would use a sound created by the spoken sound "B." Spindex cues are similar to visual index tabs found in telephone books, which enable users to flipping to the correct section of the book.

Musicons are another emerging type of non-speech sound aimed at providing reminder messages around the home environment (McGee-Lennon et al., 2011). Musicons are short excerpts of music used to remind users of activities around the home, such as, when to take a medication or go to an appointment. The optimal length of a musicon is 0.5 seconds. In previous experiments, researchers have established that musicons are useful as private reminder messages because they have the ability to make users create strong, meaningful links between the music excerpt and task. For example, a musicon created from the theme of the popular TV medical drama *House* could remind the user to take medication, while the same sound heard by others in the house would not have the same meaning. Compared to other non-speech sounds, musicons similar to auditory icons, can be learned with little or no training. Musicons are similar to earcons in terms of being private, as those can be easily personalized to mean a specific thing to the user, while remaining unrecognizable by other individuals.

2.4 Non-speech Sounds and Human Memory

Foundational work in psychoacoustics (Back & Des, 1996) raised questions about how speech and non-speech stimuli proceed from short- to long-term memory. With the advent of the personal computer in the 1980s, exploratory work in the use of acoustic cues for graphic interfaces was performed by researchers such as Gaver (1989), Brewster (1994), Blattner and Glinert (1996) and Edwards (1989). This research helped promote work with sound- and speech-based interfaces, to be used by blind and visually impaired people, or in eyes-free situations, such as driving (Edwards, 1989; Stevens, 1996). One conceptual debate in this arena concerned the relative value of speech vs. non-speech sound cues to supplement graphic-textual displays. Smither (1993) suggested that synthetic speech is generally less memorable than natural speech. A statement that Brewster (1994) agreed with.

Further debate concerns the relative value of abstract sound (beeps, blips, et al.) vs. natural sound (also called metaphoric or iconic) that refers to a topic (e.g., the sound of rain to signify rain, a weather report or meteorology). Gaver (1989) suggested that iconic sounds are more memorable cues for content, both more long-lasting in memory and better able to conjure a range or depth of content associations. Conversy (1998) suggested abstract or synthesized sounds can signify concepts such as speed or waves. As suggested by Back and Des (1996), popular media strongly influences how we expect the natural world to sound. As we know from movies "…thunder must crack, boom, or roll, and seagulls must utter high lonesome cries or harsh squawks…" (Back & Des, 1996, p. 2). In their workshop, Frohlich and Pucher (2005) stated that the systematic integration of sound and speech played an important role in some projects, in which promising design ideas of auditory systems were presented.

Previous studies have strongly suggested that sound can be a powerful prompt for memory (Sanchez & Flores, 2003). In a study performed with blind and visually impaired students (Doucet et al., 2005), the findings showed that their performance in accurately identifying the position of the sound source was superior to that of sighted students, perhaps due to a relative lack of acoustic acuity in sighted children (Sanchez, Lumbreras, & Cernuzzi, 2001). Other researchers have created games that enhanced children's short-term memory (Eriksson & Gardenfors, 2005). Previous work on non-speech sounds has focused on short-term associations with relatively simple content or meanings, and has not explored their long-term potential to encode and cue relatively large amounts of theme-based content.

2.5 The Need for a New Category of Non-speech Sounds to Signify Theme-based Content

A study of the literature on this topic has revealed that existing non-speech sounds have been used to represent objects or events in user interfaces. These objects and events typically represent brief information, which is conveyed by the majority of non-speech sounds, such as earcons, auditory icons, spearcons, auditory scrollbars and spindex. Similarly, musicons are typically used to represent brief, simple reminders in the home environment. This review indicates that existing non-speech sounds are not suitable to represent large information, such as educational content.

Content used for educational purposes is typically large and well-organized around a single theme. In order to signify such theme-based content, a new category of non-speech sounds needs to be designed and developed. The limitation posed by the existing non-speech sounds has been overcome by inventing audemes suitable for the signification and retention of educational theme-based content.

2.6 Auditory Interfaces for Blind and Visually Impaired Users

Research related to this dissertation lies at the intersection of two areas: accessible design for the blind and visually impaired and aural interactive communication.

Computer user interfaces typically offer visual feedback for the user's actions, which are obviously inappropriate for the blind and visually impaired population. In order to support accessible design, feedback solutions providing at least an equivalent user experience (to the one available to sighted users) should be developed and validated. Although this principle is affirmed in working standards for accessibility (www.w3c.org), the guidelines are very general in regard to helping create accessible content, thus much research needs to be done to create viable design strategies.

Audio is generally the easiest and cheapest solution for semantically rich feedback when the visual channel cannot be used. Audio feedback improves pointing accuracy for objects on the screen, especially small targets (Brewster, 1999). As mentioned previously, one of the seminal works in using audio feedback in interactive environments was the development of earcons (Blattner et al., 1989). Further work has been done in combining sound-based feedback with TTS. In a study reported by Parente and Bishop (2003), for example, users effectively explored a map of Roman Britain with a mouse, an on-screen cursor, trackball or touchscreen, with their movements followed by auditory icons and TTS synthesis. For example, mousing over the ocean elicited a sound of crashing waves, clicking on a city announced its name and repeated presses of the button elicited spoken facts. A similar project, the Slide Rule (Kane, Bigham, & Wobbrock, 2008), lets blind and visually impaired users explore multi-touch finger gestures on smartphones in order to read the names of items and functions on the screen without showing any visual objects (Figure 2.3). In a different vein, Sanchez and Aguayo (2006) developed a mobile messenger system for blind and visually impaired users using nine squares on a screen each containing three letters that could be selected like a cell phone keypad with the letters announced in a synthesized voice. The spaces between the buttons elicited a different sound (Figure 2.4).

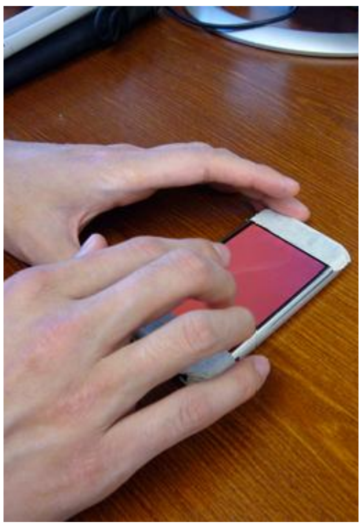

Figure 2.3. Slide Rule, a multi-touch application using audio output

Figure 2.4. Mobile messenger for the blind

2.7 The Need for Non-speech Sounds to be used in Content-rich Interfaces

A common factor of the above described acoustic interfaces is the small amount of information that they are designed to represent. This design decision has been inherited by the types of non-speech sounds these interfaces typically utilize. Existing acoustic interfaces use earcons and other non-speech sounds of the same type that are only suitable to convey brief information about an object or event. This limitation indicates that existing non-speech sounds cannot be used in acoustic content-rich interfaces.

The design of acoustic, content-rich interfaces requires a new category of non-speech sounds that can represent theme-based content. In order to advance the current state-of-the-art acoustic interfaces, audemes could be used as a solution to unleash the full potential of aural communication to support blind and visually impaired users. Such interfaces will provide users with a richer and more engaging user experience when browsing a large collection of content.

2.8 Data Sonification

An important research using sounds extensively in a user interface is the data sonification, which is the use of non-speech sounds to transform large sets of data into acoustic signals for conveying information and extracting knowledge (Kramer et al., 2010). The main purpose of sonification is to understand complex and abstract data and ease the process of decision-making and information discovery.

A successful project involving data sonification includes the mapping of data to an auditory signal in the dynamic process of monitoring tasks, such as at anesthesiology workstations (Fitch & Kramer, 1994). Other avenues of research investigated complex and dynamic auditory patterns in music and speech (Bregman, 1994; McAdams & Bigand, 1993) and detecting trends in data streams (Kramer, 2000). Sonification was used in various scientific domains, such as medicine (Kramer & Walker, 1999), seismology (Dombois, 2001), education (Bonebright, Nees, Connerly, & McCain, 2001), stock trading (Janata & Childs, 2004) and genetics (Won, 2005). More recent work has been conducted in order to create music through the sonification of large amounts of georeferenced data (geographic data that uniquely identifies a specific location on earth) (Park, Kim, Lee, & Yeo, 2010) and emotional hand gesture data collected by an accelerometer (Fabiani, Dubus, & Bresin 2010). In order to help visually impaired users of touch screen mobile devices and tablets, sonification has been used in the process of map exploration (Delogu et al., 2010) and image perception (Yoshida, Kitani, Koike, Belongie, & Schlei, 2010).

Sonification research is relevant to the increasingly vast availability of media technologies. This trend has generated ever-expanding amounts of data that need to be interpreted and understood in different ways.

With respect to the sonification community, this dissertation is concerned with issues of learnability and the retention of information delivered in the audio channel, whereas the data sonification field is concerned with means of translating massive data sets into something that might be processed more effectively with the ear than the eye. In other words, this dissertation augments existing verbal information with additional non-speech sounds with the goal of increasing the impact and effectiveness of user interfaces, while the goal of sonification is to enable the perception of the data in new ways, through the process of transforming data into the aural channel in contrast to the more traditional visual representations of data.

3. THEORETICAL FRAMEWORK; AUDEMES AND MODELS OF NON-SPEECH SOUNDS

Non-speech sounds are entities that contain signification (i.e., an ability to convey meaning). In order to explain the process of signification in non-speech sounds, I will use two existing theoretical frameworks: semiotics and modes of listening. Semiotics is used as a framework to investigate the process of meaning generation, while the modes of listening framework is used to investigate the capabilities of non-speech sounds by analyzing them as sound entities.

Using these two theoretical frameworks, I will also analyze and explain a new category of non-speech sounds—audemes. The benefit of using these theoretical frameworks is to uncover the qualities of non-speech sounds and, most importantly, the audemes, which will help in the process of establishing the full potential of the audemes.

3.1 Semiotic Traits of Non-speech Sounds

Non-speech sounds, in general, and audemes, in particular, can be considered to be audio signs as they are used to convey meaning. Using this definition, Peirce's (1908) sign model can be used as a theoretical framework for understanding the structure of non-speech sounds and the mechanisms through what they produce meaning. Peirce defined the process of interpreting signs as referring to their objects, called semiosis, which involves the following three entities (Figure 3.1):

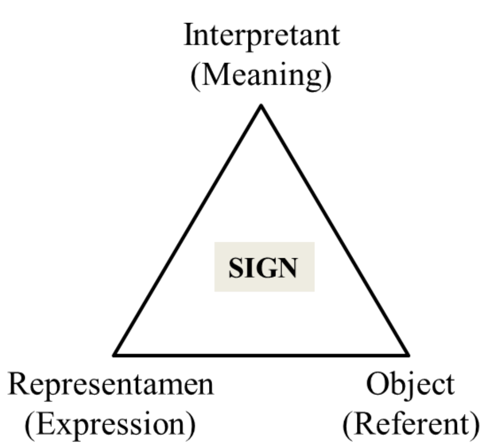

Figure 3.1. Semiosis triadic system.

Representamen (Expression) – the signifier (e.g., written word, smoke as a sign of fire);

Object (Referent) – words signified, the object to which the written word is attached (e.g., the fire signified by the smoke);

Interpretant (Meaning) – a sign signifies only in being interpreted (e.g., the meaning of danger created out of seeing smoke and fire). In other words, the meaning of a sign is manifested in the interpretation that it generates in sign users.

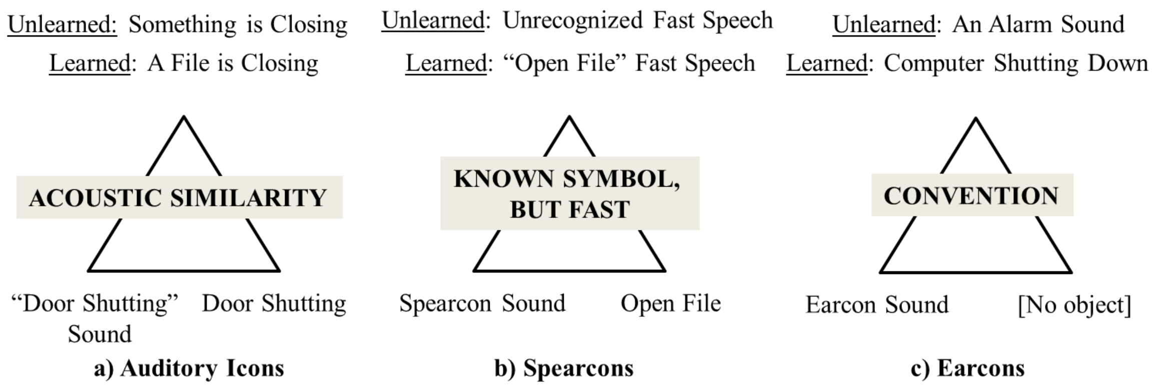

Figure 3.2. The schema of the semiosis triadic system for different non-speech sounds. Comparing auditory icons, spearcons and earcons.

Commonly, when people generate meanings about surrounding objects or events, they go through the process of semiosis, which lets them identify the interpretant by identifying the object that the representamen is linked to. In this sense, auditory icons have the advantage of being inherently recognized, meaning that the interpretant is expected knowledge that people have from their daily activities. For instance, an auditory icon of a *door shutting* can be used to signify, in a user interface, the event of a file closing. The link between the sound of the door shutting (representamen) and the event of a file closing (interpretant) may not be in the user’s knowledge yet, but would be quickly established by reflecting on the actual event of a door shutting (object). In other words, in auditory icons, the representamen is the direct, acoustic reflection of the object, which helps establish the meaning (Figure 3.2a).

Similarly, the mapping of spearcons in a user interface is not arbitrary (Walker et al., 2006) and it is expected that the interpretant should be the same to every user. For instance, in a menu-based interface, a spearcon associated with the *open* file menu option would be quickly identifiable and meaningful because the sound itself is a TTS representation of the words “open file.” In other words, in spearcons the interpretant is based on a known verbal symbol (speech), but said much more *quickly* than normal (Figure 3.2b).

Earcons are different from auditory icons and spearcons as they lack one of the components of the semiosis process: the object. Earcons, being synthetic non-speech sounds created of manipulating the sound attributes, such as the timbre, register and pitch, have no object to which they refer. They typically do not resemble anything from everyday objects and events. Considering this, the object component is absent, which halts the semiosis process and, consequently, the meaning cannot be achieved. In this case, the meaning should be created by explicitly stating the interpretant component, that is, create a direct link between the representamen and interpretant and bypassing the need for an object. For instance, an earcon created by a few notes in a low register has no object as there is nothing that people can relate to that sound when they hear it. As a result, the interpretant could be different for different people. In order to avoid this confusion, earcon designers explicitly tell people the meaning, that is, they “lock” the interpretant so that it assigns the same meaning to each person, independent of the user or context. For example, an earcon created of simple tones rising in pitch could be learned to represent the process of opening a file. In other words, in earcons the interpretant is assigned through a “convention” achieved between the designer and user (Figure 3.2c).

3.2 Modes of Listening

Imagine a group of people listening to the sound of a roaring lion. When asked to describe the sound, a variety of answers are given, such as a lion, approaching danger or a rough, deep sound. These answers are all different, despite the fact that the individuals are listening to the same sound. This difference in answers is because they were listening to the sound on different levels or modes.

There are four modes of listening: *causal*, *semantic*, *reduced* and *referential*. The first three modes were identified by Chion, Gorbman and Murch (1994), while the referential mode was introduced by Sonnenschein (2001). These different modes apply differently to different types of non-speech sounds and each (earcons, auditory icons and spearcons) will be briefly revisited here as each mode of listening is explained.

Causal mode is listening to a sound in order to gather information for its cause or source. For example, when experiencing a sound generated when writing on paper, the causal listening would be the pen and paper that are making the sound. Additionally, the causal mode is also known as the everyday or ecological listening because people tend to describe sounds by their source or the event that caused them (e.g., hearing footsteps, car passing by). Typically, we make efforts to recognize who or what generates sounds in order to take action (e.g., identify danger in order to defend ourselves from it). Further, this source-oriented mode of listening provides cues about the object generating the sound, such as its composition, mass and vibration. This mode of listening strongly applies to auditory icons because identifying the cause of the sounds is crucial in regard to generating their meanings. On the other hand, the causal mode of listening does not apply to earcons or spearcons because both lack the source of sound as an attribute of their creation: the former is a synthetic sound with variation in pitch, timbre and intensity, while the latter is a fast TTS sound.

Semantic mode refers to a code or language used to interpret a message, while concentrating on the linguistic meaning that the sound conveys, rather than the sound traits. Language is a good example of semantic listening because by convention the sounds or symbolic codes are learned to have specific meaning. Typically, anything that creates a meaning for a sound that needs to be learned by convention is an example of the semantic mode of listening (e.g., Morse code). Only earcons and spearcons can be experienced in the semantic mode of listening as they gain meaning only after they have been learned and remembered to represent a particular concept. As earcons and spearcons are generated typically from synthetic sounds, no real-world phenomenon exists that could be perceived as the cause of the sound (causal mode). Auditory icons are not experienced in the semantic mode because their meaning is not generated by convention, but by identifying the cause of the sound (causal mode).

Reduced mode of listening focuses on the traits of the sound itself independent of its cause and meaning (e.g., analyzing the sound in terms of pitch, intensity and timbre). *Pitch* is a characteristic of a sound that represents the perception of the frequency that the sound is played. For example, sound with a higher frequency is perceived as being in high pitch. *Intensity* is the power of the sound, typically measured in decibels. *Timbre* is the quality of the sound that enables people to distinguish the different notes in terms of the number of harmonics present, their order and relative intensities. The reduced mode is similar to Gaver's (1989) definition of *musical listening*, which describes sounds by listening them musically, concentrating on their abstract qualities. In other words, this mode of listening is concerned with how the sound *sounds* rather than how it has been produced or what it means. This mode of listening applies only to earcons because concentrating on and memorizing the different traits of sound is crucial to generating their meaning. Auditory icons and spearcons typically are not experienced in the reduced mode because the "sound traits" have less relevance in the generation of their meaning.

The *referential mode* of listening typically relies on the causal mode, but expands further by building an awareness of the context, such as the surrounding where the sound is experienced, and it links to the emotional connotation of the sound. For example, the sound of a foghorn might convey isolation or, by extension, loneliness, while the roar of a lion might convey fierceness or exposure to danger. The context is also achieved by adding another sound to the background or just playing it next to the first sound. For example, the sound of a key jangling could be a reference to a house, car or, even, security. Additionally, this mode of listening is socio/cultural-dependent because sounds could invoke different feeling to individuals of different cultures or societies. For example, the above mentioned example of the sound of roaring lion might convey danger for people in western societies, but, in Africa or India, lions might be seen as tribe protectors.

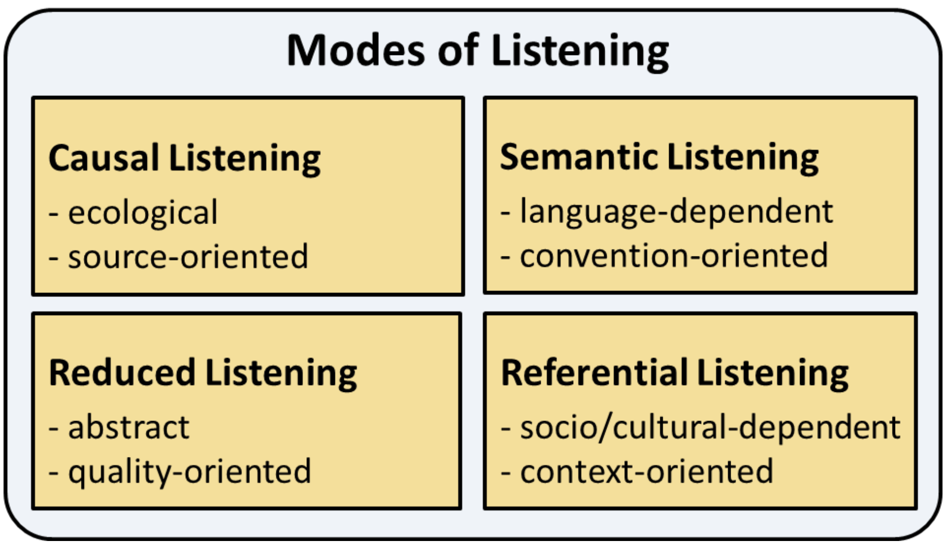

Figure 3.3. Four modes of listening and their main attributes (modified from Tuuri, Mustonen, & Pirhonen, 2007).

In summary, causal listening is an ecological mode of listening that is experienced by understanding the source that generates the sound. Semantic listening is a language-dependent mode of listening that is experienced by learning and understanding the convention applied to that sound. Reduced listening is an abstract mode of listening that is experienced only in terms of the qualities of the traits of the sound. Referential listening is a socio/cultural mode of listening that is experienced by understanding the context of the sound and the different meanings it might generate based on the culture of the listener. Figure 3.3 shows these four modes of listening and their main attributes.

3.3 Audemes as Innovative Non-speech Sounds

In addition to the previously discussed non-speech sounds, this dissertation introduces a new category of non-speech sounds called audemes.

> Audemes are short, non-speech sound symbols, under seven seconds, and are comprised of various combinations of sound effects referring to natural and/or artificial, man-made context, abstract sounds and even snippets of popular music (Mannheimer, Ferati, Bolchini, & Palakal, 2009, p. 109).

Audemes differ from previously discussed non-speech sounds in that they are semantically more flexible. While other non-speech sounds tend to be univocal (having one meaning only) for an object, audemes could generate different meanings depending on the number of sounds or shorter audemes they are created and in what order. For example, audeme (or sound) “a” (representing a key jangle) references “keys.” Adding audeme (or sound) “b” (representing car engine revving) steers the meaning of the audeme toward the process of driving or a trip. This process could continue by adding audeme (or sound) “c” (representing seagulls and surfing), which additionally steers the meaning toward “vacation and trip to the beach” (Figure 3.4).

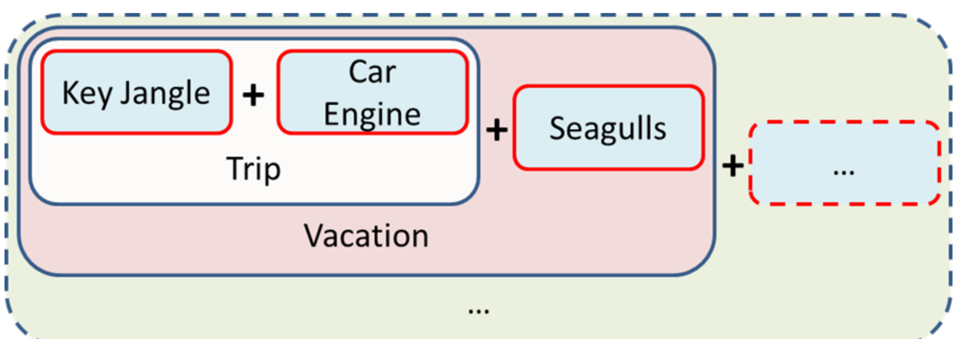

Figure 3.4. The process of audeme meaning generation by the concatenation of sounds.

While this process shows the multivocality (having many meanings) of audemes, the meaning is not completely open and arbitrary. In reality, the meaning starts wide and narrows with the introduction of additional sounds or audemes. This process indicates that, eventually, audemes will merge into a single meaning (Figure 3.5). Interviews with blind and visually impaired students from the Indiana School for the Blind and Visually Impaired (ISBVI) indicated that, for the purpose of coherent meaning creation, the maximum number of sounds in an audeme should be five (Manhheimer, Ferati, Huckleberry, & Palakal, 2009).

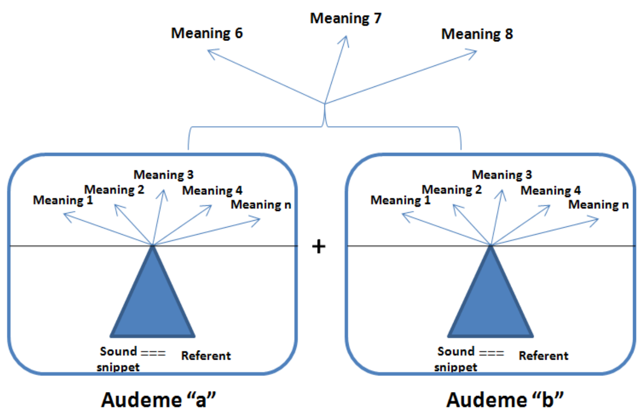

Figure 3.5. Semiotic structure of an audeme (Mannheimer, Ferati, Bolchini, & Palakal, 2009).

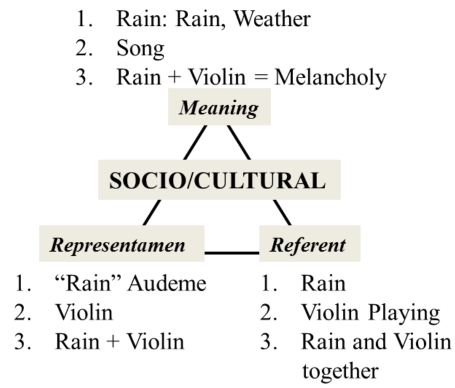

Figure 3.6. Audeme semiotics: examples.

This dissertation comes as a response to Turnbull, Barrington, Torres and Lanckriet's (2006) claim that the semantic power of non-speech sounds is underexplored. For this reason, within this study, audemes are analyzed using the two theoretical frameworks discussed above (semiotics and modes of listening). In terms of the semiotics triadic system, audemes are similar to auditory icons in the sense that when an audeme is heard, the object is clearly defined for the listeners. For instance, when an audeme of "rain" is heard, the sound is recognizable and the object clearly defined. As explained earlier, when the object is known, it is easier to develop a meaning. However, the meaning (interpretant) changes if audemes are heard concatenated to (or interleaved with) another audeme (e.g., if "rain" is interleaved with "violin playing," it could mean melancholy) (Figure 3.6). This process of audeme meaning creation shows an important feature of audemes in that their meanings change when using a concatenation of sounds.

Despite the semiotic similarity of audemes with auditory icons, these two types of non-speech sounds differ in two key aspects. First, while auditory icons consist of a single sound, audemes are generated using a concatenation of sounds to convey a meaning. For instance, an auditory icon of a "door shutting" could be used to represent a file closing on a user interface. An audeme could be generated by using this sound and adding the sound of a “key jangle” to represent a file closing and it being password protected. Essentially, in this case, an audeme represents a concatenation of two auditory icons. However, audemes do not always represent a concatenation of two or more auditory icons, which is the second reason why audemes differ from auditory icons. Auditory icons are mainly sounds occurring in the natural environment. Audemes are created from sounds occurring in nature as well as abstract and musical sounds, such as excerpts from songs. In other words, auditory icons are a subset of audemes in terms of the source of the objects.

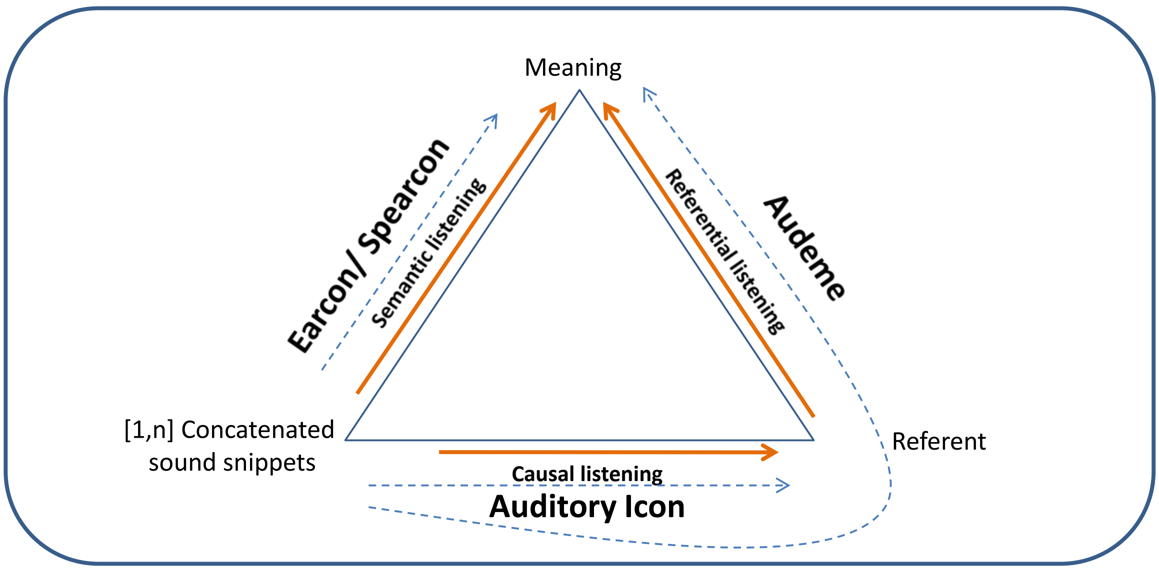

Figure 3.7. Overlaying modes of listening over the semiotic structure of non-speech sounds.

The unique difference between audemes and other non-speech sounds is particularly revealed when analyzed in terms of modes of listening. For earcons and spearcons, the meaning is created through semantic listening (i.e., the meaning is created directly by learning the sound and what it represents without any reference to the cause of the sound) (Figure 3.7). In auditory icons the meaning is created through causal listening (i.e., the meaning is created by identifying the reference to the cause of the sound). In audemes, however, the meaning is created by first identifying the cause of the sound generation, and then through the referential mode, the definite meaning is formed (depicted by a dashed curved line in Figure 3.7). For instance, an audeme of a “hammer hitting a nail,” will first trigger the causal mode by identifying the objects generating the sound (referent). Once this link is established, the referential mode can be activated and the final meaning generated could be “building construction,” “bringing things into order” or even “designing architecture.” This ability to generate a meaning through the referent (i.e., through the causal and referential modes of listening), makes audemes evocative rather than iconic non-speech sounds (such as earcons, spearcons and auditory icons). Earcons, spearcons and auditory icons typically convey a message about an object or event, and usually the meaning is simple and straightforward. Audemes, on the other hand, generate more complex semantics that can be used in conveying content-rich information, such as essays.

Essays are typically short in content with a unifying or a dominant idea and are around 500 words. Audemes are used to convey meanings that are essential to the topic of the essay. For example, an audeme consisting of the sound of a gavel would symbolize court. When this sound is added to the sound of a quill pen writing and the Star Spangled Banner, the conveyed meaning is an essay on the U.S. Constitution that describes the judiciary system in the U.S., how it was established and important amendments included in the Bill of Rights.

This dissertation does not explore the causal mode of listening as the link between the sound and referent (the object causing the sound) as it is fairly straightforward in the case of audemes. Instead, this dissertation analyzes the link between the referent and meaning through the referential mode of listening. For example, through the causal mode of listening, a sound of the “wind” reveals the wind as its referent, while a sound of a “human kiss” reveals kissing as its referent. This process is straightforward. It gets interesting when these two sounds generate an audeme, which analyzed in the referential mode of listening could mean “sending kisses,” “wind lover” or even “love is like flying.”

4. UNDERSTANDING THE POTENTIAL OF AUDEMES FOR CONTENT RECOGNITION AND SIGNIFICATION

4.1 Overview of the Experiments

The purpose of these two experiments was to evaluate audemes and their effect on information retention when associated to large content collections. Experiment one was conducted in order to test the factors pertaining to the recognition of the essay content when associated with the audemes. Experiment two was a replica of experiment one, conducted to investigate further the effects of audemes on the different types of essays.

In these two experiments, three audemes were used: radio, the U.S. constitution and slavery. The radio audeme was the sound of a radio dial being twisted through the stations. The U.S. constitution audeme combined the sound of a gavel (symbolizing courts), the sound of a quill pen writing and the Star Spangled Banner. The slavery audeme combined a short passage of a choir singing “Swing Low, Sweet Chariot” followed by the sound of a whip crack. For each of these audemes, an approximately 500 word essay was created. The essays contained information on the evolution of radio in the U.S., the U.S. constitution and slavery.

4.2 Experiment 1: Exploring Information Recognition Using Audemes

4.2.1 Overview and Purpose

The aim of this experiment was to identify the potential of audemes to enhance the recognition of contents presented as an essay. Three audemes, along with three essays for radio, the U.S. constitution and slavery, were used in this experiment. The independent variables were the three audemes and three essays. The dependent variable was the recognition rate of the content associated with the audeme (Table 4.1). In order to evaluate long-term retention of the information, the participants in this experiment were tested two weeks and five months after their exposure to the essays. This experiment has been approved by IRB #IRB-1009001883.

Table 4.1. Independent and dependent variables for experiment 1.

Independent Variables	**Dependent Variables**
Audemes Essays	Content Recognition Rate

4.2.2 Method

4.2.2.1 Participants

Twenty one participants (8 male and 13 female, 10 blind and 11 visually impaired) were recruited for the experiment. The participants were paid for their participation. The participants were students from the ISBVI and were all between the ages of 10 and 18. For the experiment, blind and visually impaired participants were recruited because audemes were created to help them. Additionally, other studies have shown that, in similar experiments, blind and visually impaired children outperformed sighted children (Doucet et al., 2005; Sánchez et al., 2001).

The students of ISBVI were notified of the study via announcements from their teachers. All 21 students who volunteered took part in the experiment. Informed consent documents were given to the participants' parents prior to their involvement in the research.

4.2.2.2 Stimuli

The three audemes, radio, the U.S. constitution and slavery, were generated by combining sounds obtained from a library of sounds that professional audio designers use. The Soundtrack Pro (http://www.apple.com/finalcutpro) software was used to generate the audemes. Each audeme was four to six seconds long. The essays used were created from online sources, such as Wikipedia, and were suitable for the participants. The themes of the three essays, radio, the U.S. constitution and slavery, fall into U.S. history, which was used as the area of interest because it was relatively common knowledge across all of the participants as it is part of their standard curriculum in the school. Complete stimuli are included in Appendix B.

4.2.2.3 Apparatus

The participants collectively listened to the audemes. Sound amplifiers were used in order to play the sounds at a higher volume. The test was conducted using Braille and large-font printed sheets and the participants used crayons to mark their answers. Meetings with the participants took place at the ISBVI once a week.

Radio essay:

(760 words)

Radio is a technology that uses electro-magnetic signals, sometimes called radio waves that travel through the air and many solid materials and can be detected by receivers, generally called radios. Although the basic principle of radio is simple, developing the complex technology for sending and receiving radio waves took almost a century to evolve. There are many military, communication and entertainment uses for radio, but most of us think of radio in terms of music and information broadcast by radio stations.

As the 19th century progressed with many advances in science and technology building on previous efforts, it was clear to many inventors that wireless communication was possible. Many people working in different countries throughout the century in different countries contributed various pieces of the complex technology that would become modern radio. Some of the more famous people were Michael Faraday, James Clerk Maxwell, Thomas Alva Edison, Nikola Tesla, Ernest Rutherford and Guglielmo Marconi, who was awarded a British patent for radio in 1896. Marconi received an American patent, and in 1901, he conducted the first successful transatlantic experimental radio communication. In 1909 Marconi won the Nobel Prize for his work with radio.

In 1912, the famous ocean liner Titanic sank. This event made it clear to the public that lives might have been saved if radio communications had been available and monitored. After this, radio transmitters quickly became universal on large ships. In 1913, the International Convention for the Safety of Life at Sea produced a treaty requiring shipboard radio stations to be manned 24 hours a day.

At the same time the United States government decided to regulate the use of broadcast radio and the airwaves in America. Congress created the Federal Radio Commission to control broadcasting, and many years later this government agency evolved to become the Federal Communication Commission. In the beginning, most radio users were amateur broadcasters who built their own radios and communicated with other amateur users. When the government began to regulate broadcasting, this helped promote the growth of the commercial radio.

On February 17, 1919, station 9XM at the University of Wisconsin in Madison broadcast the first human speech to the public at large. That station is still on the air today as WHA. But the first for-profit commercial radio station in the US is generally thought to be KDKA in Pittsburgh, Pennsylvania, which in October 1920 received its broadcasting license and went on the air as the first US licensed commercial broadcasting station.

However, in the early years of radio, broadcasting was not yet supported by advertising or listener sponsorship. The stations owned by radio manufacturers and department stores were established to sell radios and stations owned by newspapers were used to promote the sales of newspapers and express the opinions of the owners.

Figure 4.1. An excerpt from the radio essay depicting the respective audeme being played between paragraphs.

4.2.2.4 Procedure

The participants were randomly assigned to three different groups, however, it was necessary to keep certain criteria balanced for each group (i.e., ages, learning abilities and level of visual impairment). In order to establish a baseline, all of the participants first took a pretest questionnaire that contained 10 questions covering the content of the essay. During the same session, each of the participants heard the essay content twice, which was read aloud by one of the researchers. Two weeks later, the participants took a posttest that contained the same questions from the pretest, but in a randomized order, as well as three additional questions serving as noise (in order to give the impression that we are not using the same questionnaire as in the pretest). The questions in the pretest and posttest were read aloud by the researcher, while the students marked their answers on Braille or large-font printed sheets. For instance, the researcher would read the questions (indicated by numbers one through ten) and the possible answers (indicated by letters A through E), while participants read the same question and answers on their sheets. The answers would appear, for example, as 1A through 1E. Then, the students were to mark one answer for each question. The questions and possible answers were read to the participants twice.

The three groups (each consisting of seven participants) were created in order to track the effects of the audemes in the encoding and recognition of the essay content. The first group – the Control Group – was not exposed to the audemes at all, and simply listened to the essays and questions. The second group – the Encode Group – listened to the audemes along with the essay, but was not exposed to the audemes when taking the posttest questionnaire. The respective audeme for each of the essays was played five times in the course of reading the essay content, between each paragraph or approximately every one minute (Figure 4.1). The researcher would pause the reading, play the audeme, and then resume reading the essay. The third group – the Encode+Recognition Group – listened to the audemes along with the essay and also when taking the posttest questionnaire. Similar to the Encode Group, the Encode+Recognition Group listened to the respective audeme five times while the essay was being read to them. When taking the posttest, the participants of the Encode+Recognition group listened to the audeme in between each question (Figure 4.2). This division into groups made it possible to track how well the participants remembered the essay content by itself after two weeks (Control Group) as well as if the audemes enhanced encoding (Encode

Group) and how well the audemes enhanced both encoding and recongition (Encode+Recognition Group).

In order to avoid overloading the participants, they were tested on different essays on different weeks, thus stretching the experiment out to six weeks.

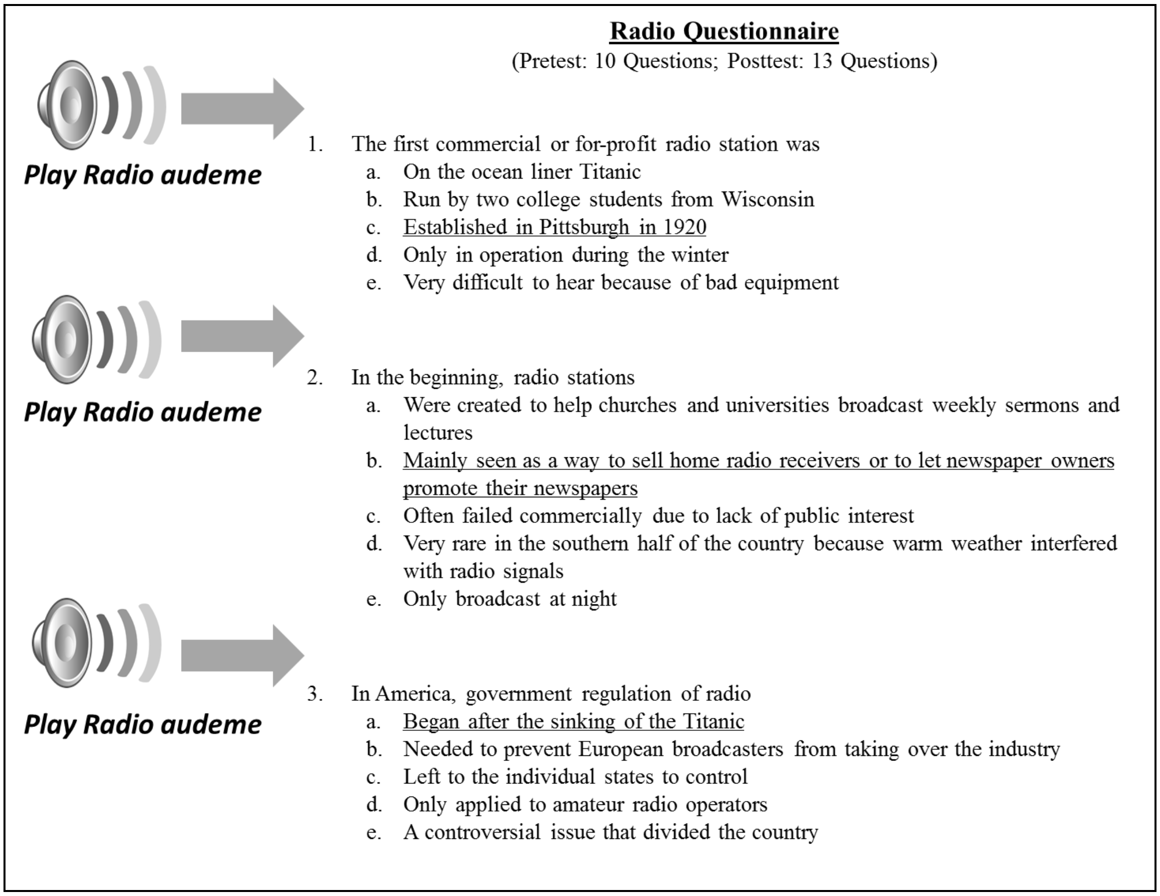

Figure 4.2. An excerpt from the radio questionnaire depicting the respective audeme being played between each question. The correct answers are underlined.

4.2.2.5 Results

The results of this experiment show that the audemes improved the participants' abilities in regard to encoding and recognition. A mixed repeated measures ANOVA was conducted in order to identify any significance between the three groups (Control, Encode and Encode+Recognition). An overall significant difference was found, $F(2,18) = 8.33$, $p < .005$, $\eta^2 = .481$. A Tukey HSD post hoc test revealed significant differences between the Control and Encode+Recognition Groups ($p < .01$) and between the Encode and Encode+Recognition Groups ($p < .05$), but not between the Control and Encode

Groups (p > .05). The sample means of the gain (posttest - pretest) are shown in Figure 4.3.

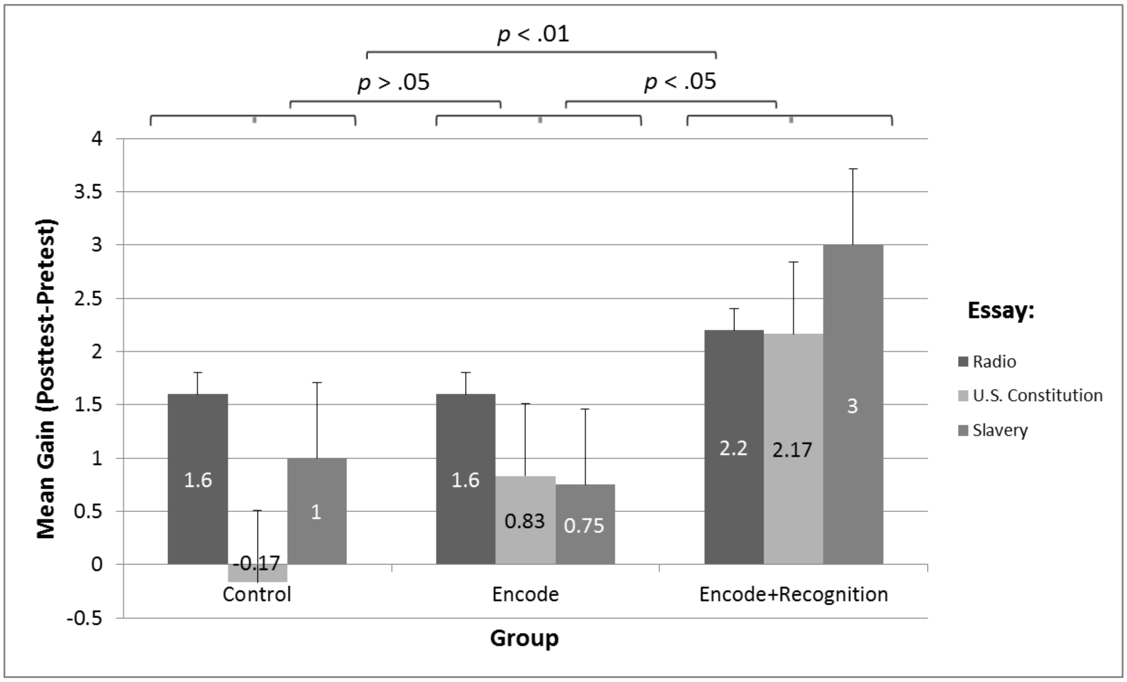

Figure 4.3. Audemes facilitate encoding and recognition of educational essays.

These findings demonstrate that exposure to audemes in conjunction with the text increased encoding and recollection of the associated content. The Encode+Recognition Group for the radio essay showed a 52% increase in tested knowledge (from 4.2 correct answers in the pretest to 6.4 correct answers in the posttest), factored against pre-knowledge. For the U.S. constitution essay, there was a 65% increase (from 3.3 correct answers to 5.5 correct answers), while for the slavery essay, there was an 80% increase (from 3.75 to 6.75 correct answers). The Encode Group showed a 38% increase for the radio essay (from 4.2 to 5.8 correct answers), a 16% increase for the U.S. constitution essay (from 5.16 to 6 correct answers) and a 12% increase for the slavery essay (6.25 to 7 correct answers). For the Control Group, there was a 47% increase for the radio essay (3.4 to 5 correct answers), a 3.6% decrease for the U.S. constitution essay (4.67 to 4.5 correct answers) and a 20% increase for the slavery essay (5 to 6 correct answers).

Five months later, I re-tested the participants on all three essays. The order of the questions and answers was changed and three additional questions were added to serve as noise. The average results for all three essays for the Control Group showed a 7.2% *decrease* compared to the pretest results. For the Encode Group, there was 0.78%

increase, while for the Encode+Recognition Group there was a 3.8% increase. The sample means of the gain (posttest - pretest) are shown in Figure 4.4.

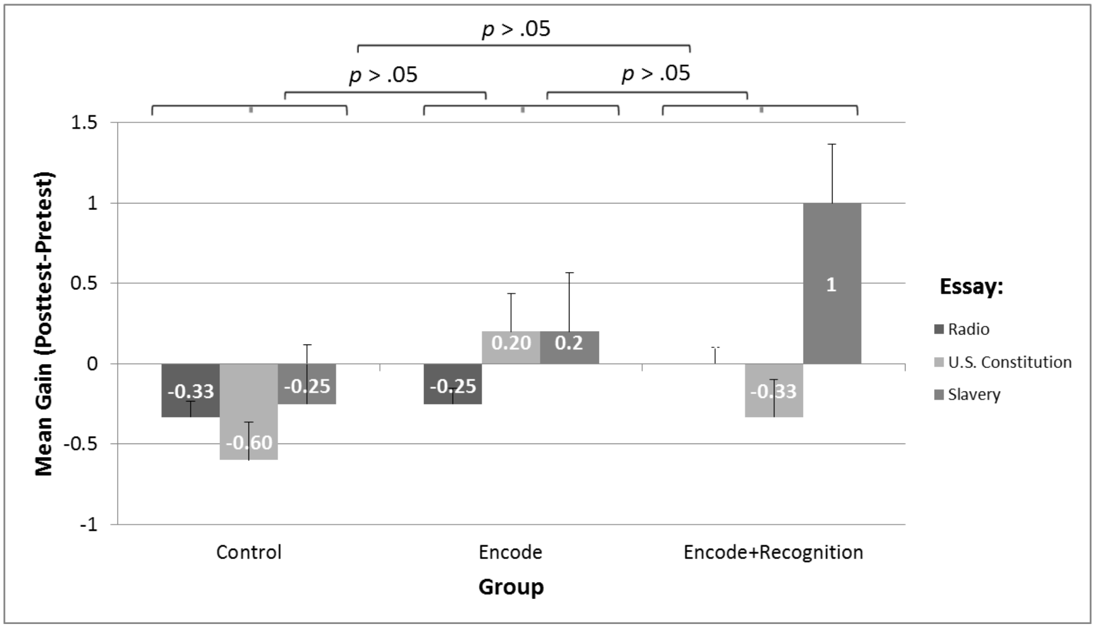

Figure 4.4. After five months audemes guard against memory erosion.

Although, no significant difference existed between the groups, $F(2, 18) = .75$, $p>.05$, $\eta^2 = .077$, it is important to note that consistently the Encode+Recognition Group (whose participants heard the audeme along with the essays as well as when taking the posttest) performed better than the Encoding Group (whose participants heard the audeme only when hearing the essay, but not when taking the posttest) and the Control Group (whose participants were not exposed to the audeme). In order to further investigate these findings, I compared the three experimental groups regardless of the essay presented for the two time periods, after two weeks and five months. The repeated measures of the ANOVA revealed the main effect for the time periods (two weeks and five months), $F(1,18) = 20.66$, $p < .001$, $\eta^2 = .535$, and the experimental groups (Control, Encode and Encode+Recognition), $F(2,18) = 7.67$, $p < .005$, $\eta^2 = .460$. A Bonferroni pairwise comparison revealed significance between the two week and five month periods ($p < .001$) and between the Control and Encode+Recognition Groups ($p < .005$). No significant difference was found between the Control and Encode Groups ($p > .05$) or the Encode and Encode+Recognition Groups ($p = .08$). There was also no interaction found between the time period and the three Groups, $F(2,18) = 1.11$, $p >.05$, $\eta^2 = .11$. The sample means are shown in Figure 4.5. The experiment raw data set is included in Appendix A.

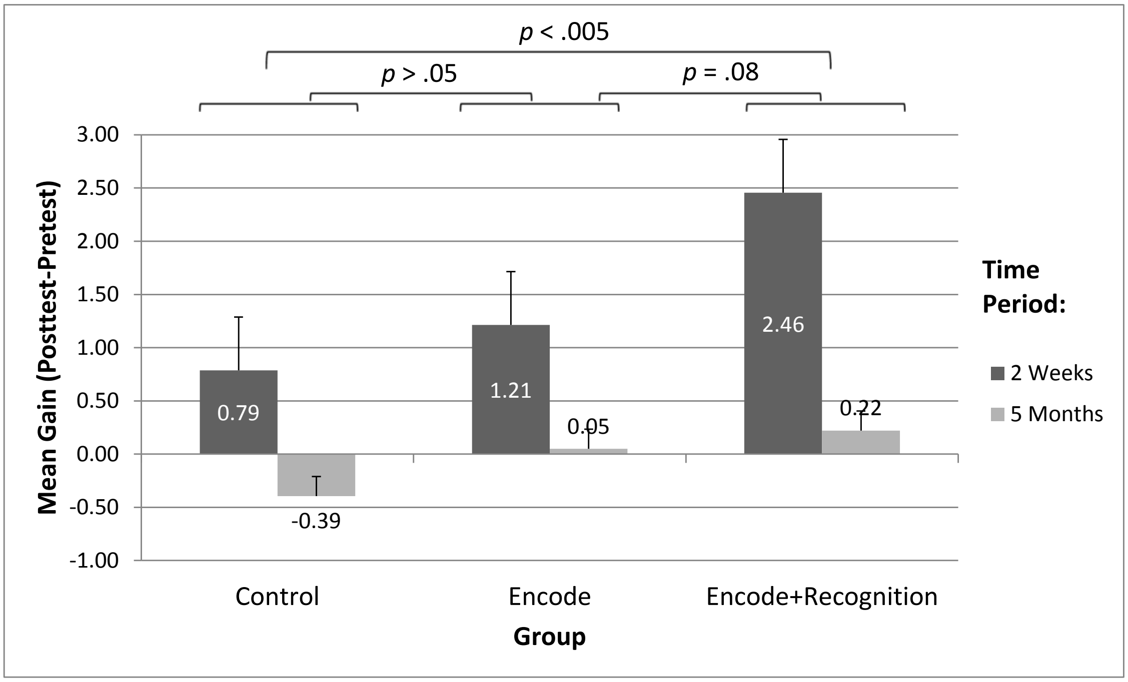

Figure 4.5. Longitudinal comparison reveals identical trends for the tests conducted after two weeks and five months.

4.3 Experiment 2: Studying Comprehension and Involvement with Audeme-based Essays

4.3.1 Overview and Purpose

The aim of this experiment is to evaluate the essays used in the previous experiment and discover the factors that pertain to different effects found between the three essays. For example, the slavery essay in the Encode+Recognition Group had the largest effect in comparison to the other two essays. Hence, this follow-up experiment was conducted in order to investigate the nature of the essays used in experiment one. This experiment has been approved by IRB #IRB-11060057641.

4.3.2 Method

4.3.2.1 Participants

Fifteen sighted participants (8 male and 7 female) were initially recruited for the experiment. One participant, for medical reasons, failed to show up to one of the sessions, thus making the data incomplete for that participant. Therefore, the results only take into consideration the data from 14 participants (8 male and 6 female). All of the participants were Indiana University-Purdue University Indianapolis (IUPUI) graduate students. They received information about the project via email. The first 15 participants who volunteered were included in the experiment and were given gift cards for their participation.

4.3.2.2 Stimuli

The same audemes and essays that were used in experiment one were also used in this experiment. Complete stimuli can be found in Appendix B.

4.3.2.3 Apparatus

The same apparatus that was used in experiment one was also used in this experiment. The experiment was conducted in the User Simulation and Experience Research (USER) Lab on the IUPUI campus.

4.3.2.4 Data Collection

For this experiment, four measures in a questionnaire format were used. First, a knowledge gain questionnaire (with the same questions used in experiment one) were used to estimate the level of knowledge that the participants gained pertaining to the essay content. Second, a narrative transportation (Green & Brock, 2000) questionnaire was used in order to assess the level of emotional attachment each participant experienced with the essay content. Third, a content comprehension (Macedo-Rouet, Rouet, Epstein, & Fayard, 2003) questionnaire was used in order to assess the level of cognitive load necessary for each participant to understand the essays. Finally, a cognition scale (Lord & Putrevu, 2006) was used in order to assess personality traits of the participants based on their willingness to engage in complex and high cognitive demanding tasks. These questionnaires were answered using a seven-point Likert scale. A complete list of questions can be found in Appendix B.

4.3.2.5 Procedure

In order to counterbalance the exposure order of essays, the participants were randomly assigned to three groups; however, an effort was made to keep the groups gender balanced. A within subject design was conducted with all of the participants taking part in all of the experimental conditions. The experiment took two sessions spread out over three weeks.

In the first session, all of the participants took a pretest questionnaire in order to establish their knowledge baseline for the first essay. Afterward, the participants listened to the first essay twice in a recorded audio format spoken by a native English speaking person. While listening to the essay, between each paragraph the participants listened to the respective audeme for that essay, just as in experiment one (Figure 4.1). Upon completion of listening to the essay, the participants filled out a questionnaire measuring three factors that could influence retention of the essay's content (i.e., narrative transportation, content comprehension and a need for cognition dimensions). The same strategy was followed for other two essays in the same session. The entire session took approximately 60 minutes.

In the second session, two weeks after the first session, the participants took three multiple choice questionnaires (one for each essay). The questions were the same as those on the pretest questionnaire with the additional of three noise questions. Thus, the questionnaires each contained 13 questions. The same audemes, respective to the essays, were played between each question (Figure 4.2).

4.3.3 Results

The slavery essay invoked higher involvement and was better comprehended: The main goal of this experiment was to assess the content of the essays provided along with the audemes. The results showed that the participants had significantly higher ratings for the slavery essay when compared to the radio and U.S. constitution essays for the narrative transportation and content comprehension questionnaires. The results were analyzed using a repeated measure ANOVA, within-subjects design.

For the narrative transportation questionnaire, the analysis showed an overall significant difference among the essays' content, $F(2, 26) = 5.69$, $p < .01$, $\eta^2 = .304$. The Bonferroni pairwise comparison revealed a significant difference between the slavery and U.S. constitution essays ($p < .05$) and an approaching significance between the slavery and radio essays ($p = .07$), but a non-significant difference between the radio and U.S. constitution essays ($p > .05$). The sample means of the narrative transportation rating per essay are shown in Figure 4.6.

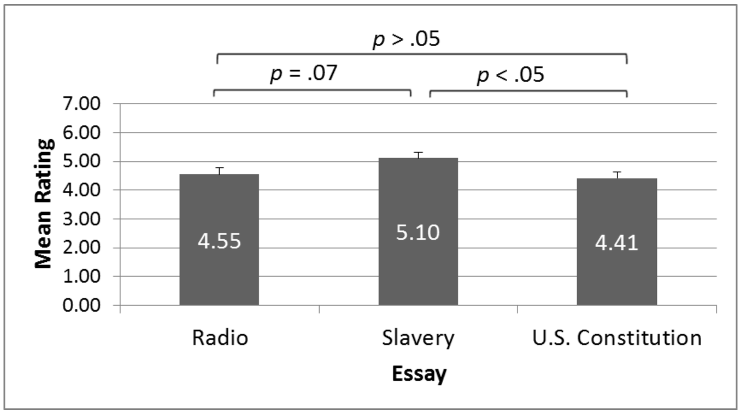

Figure 4.6. The slavery essay invoked the highest involvement.

Similarly, for the content comprehension questionnaire, the analysis showed an overall significant difference among the essays' content, $F(2, 26)=3.94$, $p < .05$, $\eta^2= .233$. The Bonferroni pairwise comparison revealed a significant difference between the slavery and radio essays ($p < .05$), and a non-significant difference between the slavery and U.S. constitution ($p > .05$) and the radio and U.S. constitution essays ($p > .05$). The sample means for the content comprehension rating per essay are shown in Figure 4.7.

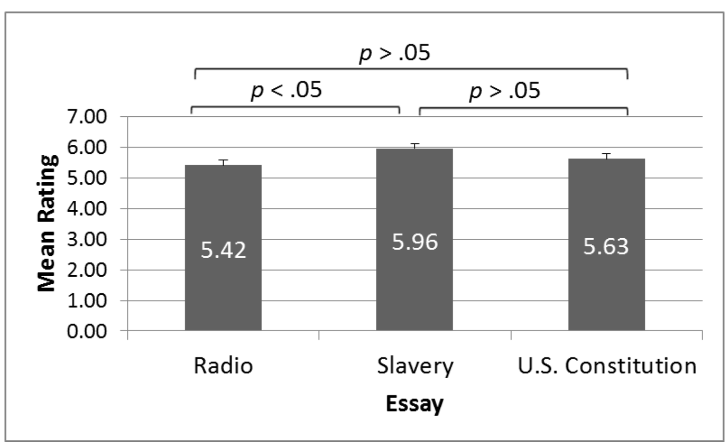

Figure 4.7. The slavery essay was better comprehended.

Strong correlation between involvement and content comprehension: A strong correlation was identified between the narrative transportation and content comprehension scores across all of the essays. A significant, positive correlation was found between the narrative transportation and content comprehension for the radio ($r = .569$, $n = 14$, $p < .05$), U.S. constitution ($r = .763$, $n = 14$, $p < .01$) and slavery ($r = .501$, $n = 14$, $p = .05$) essays. This result indicates that an increase in user's involvement with the essay's content was correlated with an increase in score in regard to understanding the essay.

The U.S. constitution essay had the highest knowledge gain: In this experiment, it was expected that the highest gain will result for the slavery essay because the slavery essay had the highest scores on the narrative transportation and content comprehension questionnaires. The U.S. constitution essay, however, had the highest knowledge gain, although not significantly. The repeated measures ANOVA revealed an approaching overall significance among the three essays, $F(2, 26) = 2.88$, $p = .07$, $\eta^2 = .181$. The sample means of knowledge gain per essay are shown in Figure 4.8. The experiment raw data set is included in Appendix A.

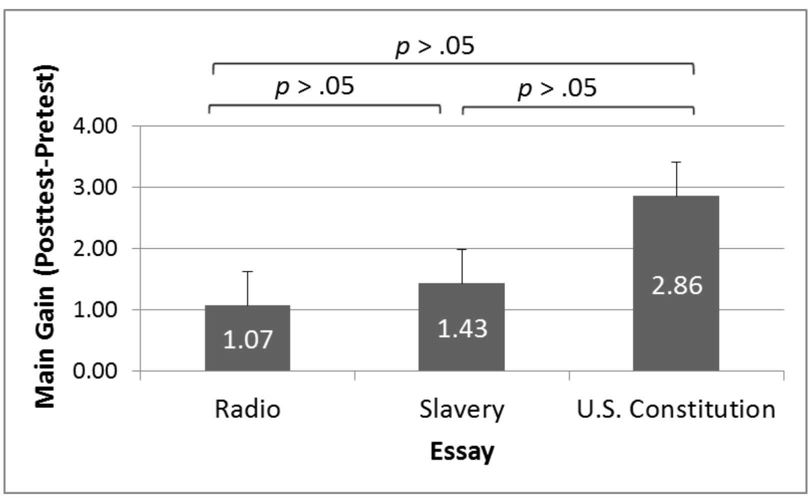

Figure 4.8. The U.S. constitution essay scored the highest in terms of knowledge gain.

4.4 Discussion of the Two Experiments

One of the main results of the first experiment indicated that the audemes increased knowledge gain for the associated text. This result was observed with the Encode+Recognition Group, which heard the audeme along with the essay and also during the posttest questionnaire after two weeks. In addition, the follow-up questionnaire taken after five months indicated the same trend across all of the essays for this group. A particularly interesting observation was that for both questionnaires, the test after two weeks and also five months, the slavery essay had the highest knowledge gain (although not significant) when compared to the other two essays. A biased content (e.g., an emotionally loaded narrative) presented as part of the essay was suspected to be the cause of these differences.

With this possible explanation in mind, a follow up experiment was conducted with the purpose of understanding the differences among the essays. This experiment was conducted by using the same audemes and essays, but introducing instruments to assess the essays' content within two dimensions: narrative transportation and content comprehension. Both of these two dimensions revealed that the essays presented consisted of biased content. The slavery essay was judged to contain emotionally laden content and, therefore, was being comprehended better than the other essays. This finding explained the high knowledge gain observed for the slavery essay in experiment one. This finding, however, does not explain the high knowledge gain scores for the U.S. constitution essay seen in experiment two.

In order to investigate the possible reasons to this cause (the slavery essay scoring highest in experiment one with the blind and visually impaired participants and the U.S. constitution essay scoring highest in experiment two with the sighted participants), a closer exploration of the pretest and posttest data for all of the essays was conducted. The comparative results between the sighted and blind and visually impaired participants revealed similar results across all of the essays, except for the pretest scores for the slavery essay. As shown in Figure 4.9, in the slavery essay pretest, the sighted participants ($M = 5.36$, $SD = 1.55$) scored significantly higher than the blind and visually impaired participants ($M = 3.75$, $SD = 1.86$). The t-test analysis demonstrated this significant difference, $t(19) = 2.01$, $p < .05$. Moreover, the posttest scores for the slavery essay for groups were almost identical, perhaps an indication of a ceiling effect reached (6.8 on the scale of 10). Thus, the highest scores for the U.S. constitution essay for the sighted participants were not due to the audemes being more effective, rather, for some reason, the blind and visually impaired participants had significantly less prior knowledge about slavery compared to the sighted participants.

A possible explanation for the sighted participants having a higher prior knowledge about slavery is the users' profiles. The blind and visually impaired participants were elementary and middle school students, while the sighted participants were graduate students. Perhaps, since the sighted participants had a higher education level, their general knowledge about slavery was higher than that of the blind and visually impaired participants.

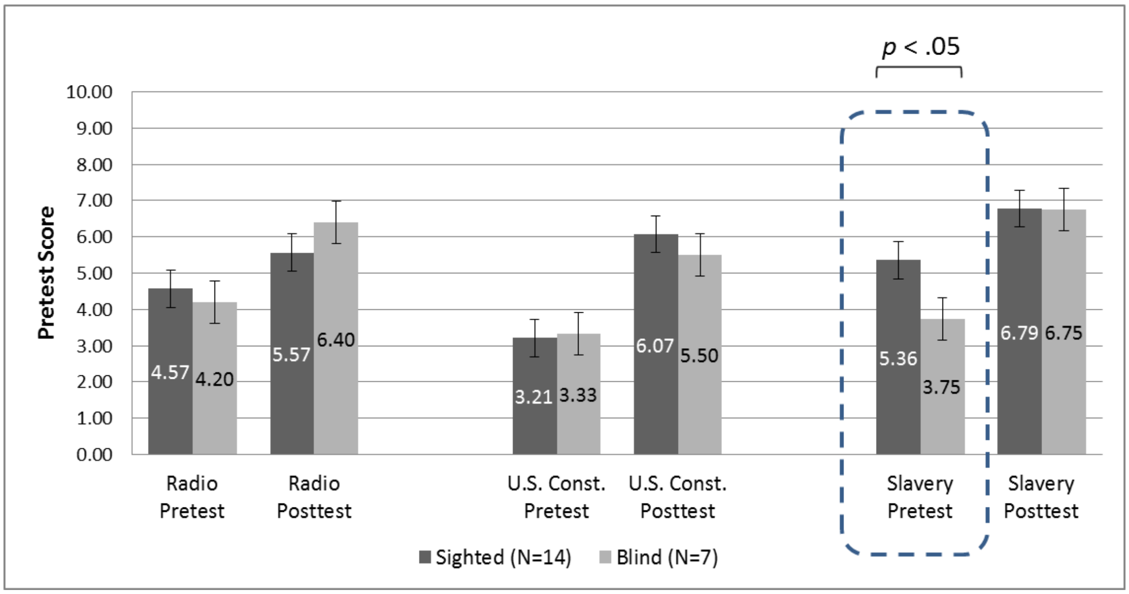

Figure 4.9. The largest difference was seen in the slavery pretest scores.

Another factor that might have influenced the differences between these two types of users is the double sample size of the sighted participants (N = 14) compared to the blind and visually impaired participants (N = 7). The small number of blind and visually impaired participants could be seen as a limitation to this experiment, although this number is consistent with most of experiments conducted with blind and visually impaired participants (i.e., four to six participants (Stevens, 1996; Challis & Edwards, 2000)). It is difficult to obtain additional blind and visually impaired participants due to mobility issues and skill qualifications.

Another important factor is the significant positive correlation found between the narrative transportation and content comprehension dimensions across all of the essays. This correlation is an indication that the comprehension of the essay content and the user's involvement with the content had a reciprocal positive effect.

5. AUDEME-BASED ACOUSTIC INTERFACE: AEDIN

5.1 Overview and Conceptual Design

This chapter explains the design, implementation and evaluation of an acoustic interface called AEDIN (Acoustic EDutainment INterface). AEDIN is an auditory and touchscreen interface built to support education for blind and visually impaired students. AEDIN was developed at the School of Informatics at Indiana University-Purdue University Indianapolis (IUPUI), and tested at the Indiana School for the Blind and Visually Impaired (ISBVI). It functions as a knowledge base of aural educational content (recorded essays and quiz questions about these essays) navigated through touch input with audio output or feedback (audemes, as well as navigational feedback sounds).

Following the steps of a typical user-centered design, AEDIN was built in the principle of ability-based design (Wobbrock et al., 2011), which uses a user's abilities rather than disabilities to guide the design process. Considering that blind and visually impaired individuals navigate the surrounding environment by hearing sounds and touching objects, AEDIN is primarily based on auditory feedback and touchscreen input (Figure 5.1).

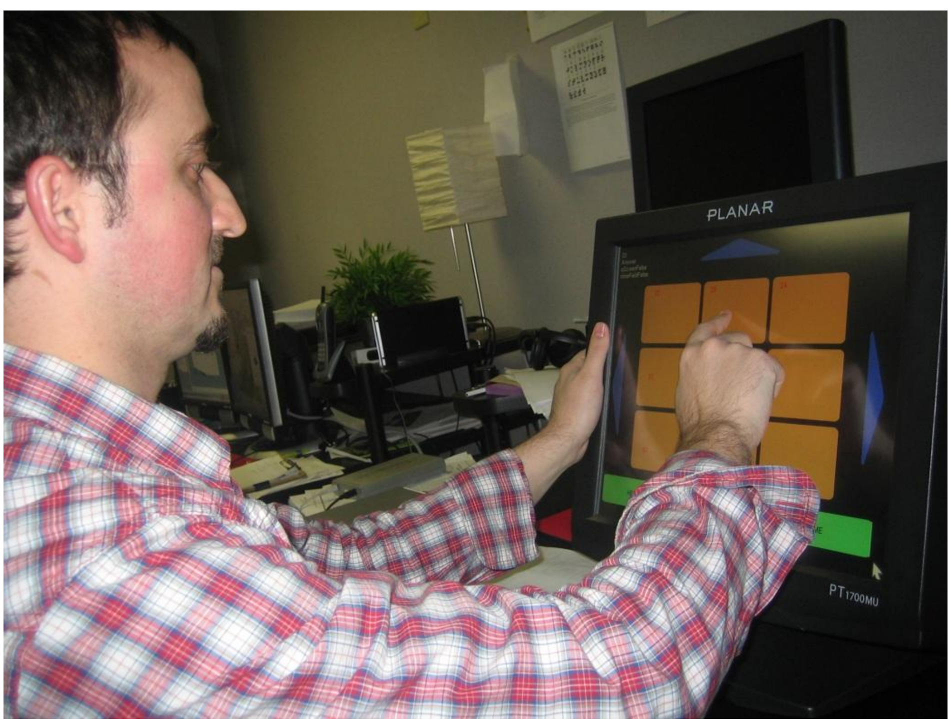

Figure 5.1. Interacting with AEDIN.

5.1.1 The Purpose of Designing AEDIN

The aim of designing and implementing AEDIN was to develop an application able to utilize audemes. In AEDIN, audemes are used as *aural covers* to anticipate large content information, such as TTS essays. In order to understand the notion of aural covers, an example can be seen in existing image and video sharing web sites. In image sharing sites (e.g., www.flickr.com), an image thumbnail is presented to anticipate the actual larger image that is shown when the thumbnail is clicked upon. Similarly, in video sharing sites (e.g., www.dailymotion.com), a thumbnail with rotating sample screenshots of the video content is presented to anticipate the video content. These techniques help users anticipate the content and make decisions before engaging it.

Serving a similar purpose, audemes—created with concatenated sounds—are used to anticipate concepts covered in the content delivered as a TTS essay. Figure 5.2 depicts a scenario in which users engage the AEDIN interface by hearing an audeme on economics (created by the sound of cash register and coins) to anticipate an essay about economics.

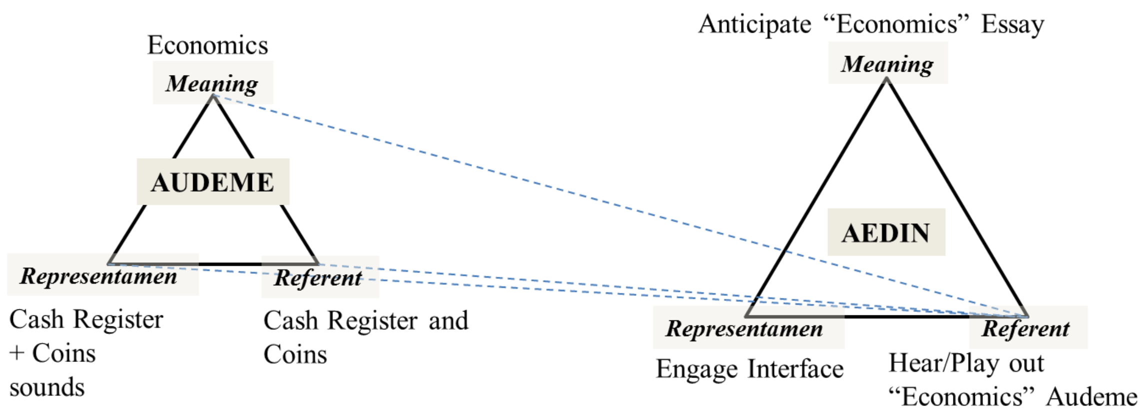

Figure 5.2. The semiotics of audemes used as aural covers to anticipate TTS essays in AEDIN.

Forty-nine audemes associated to 49 essays are built in to AEDIN. Figure 5.3 depicts an excerpt of three audemes (radio, economics and slavery) being used to anticipate respective TTS essays (radio, economics and slavery essays). Only a fragment of the content of each essay is shown.

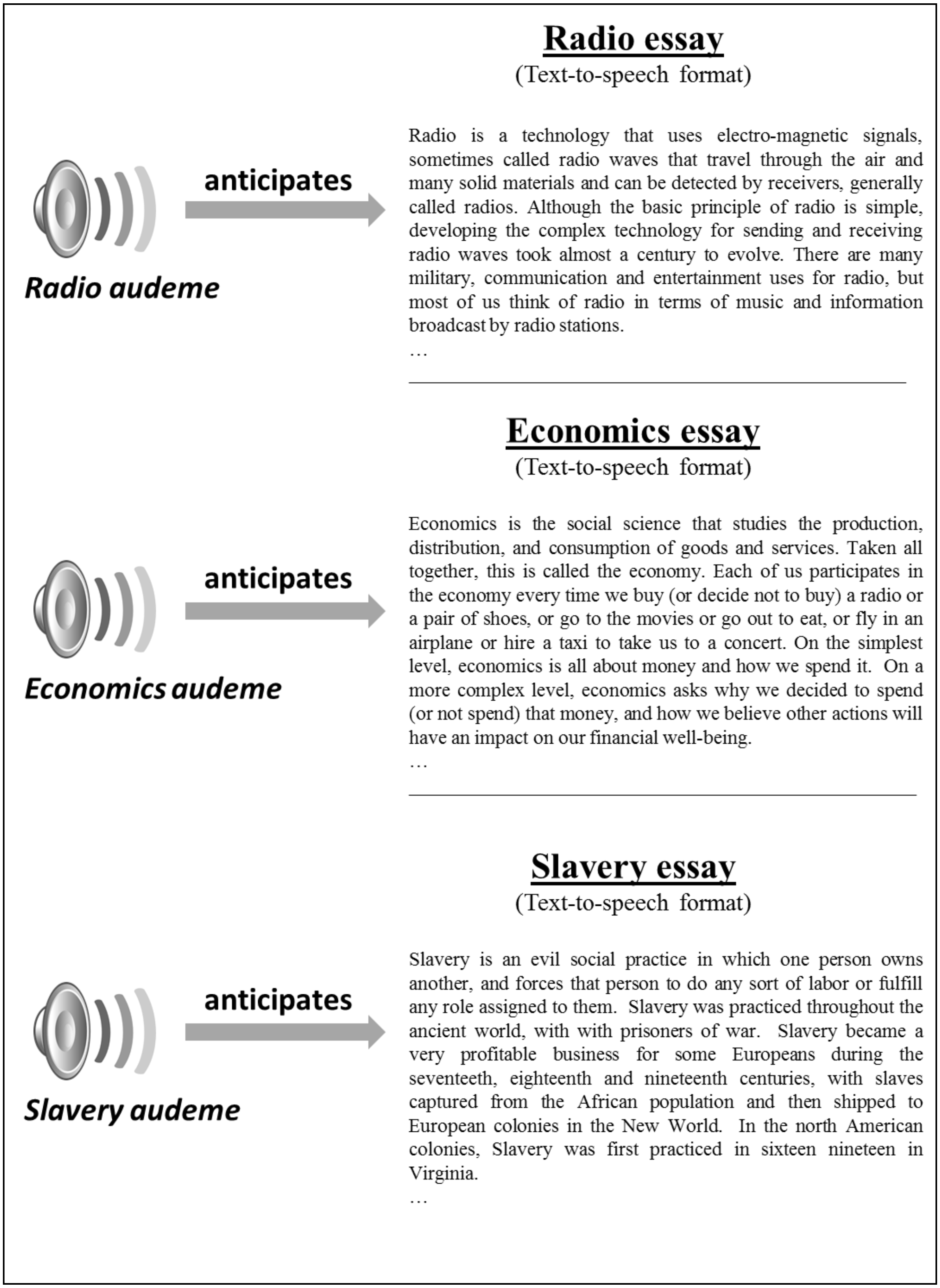

Figure 5.3. An excerpt of three audemes anticipating three respective TTS essays as part of the AEDIN interface.

5.1.2 Auditory Feedback and Touchscreen Input

AEDIN follows examples of successful interfaces based on auditory feedback. For instance, SonicFinder (Gaver, 1989) and ARKola (Gaver, 1990) represent interfaces where auditory icons were used to convey messages about objects or events on the screen.

Ballas (1994) and Walker and Kramer (1996) discussed fundamental design principles for auditory feedback interfaces, focusing mainly on how auditory interfaces can adapt effectively to human perception and cognition. Studies by Shinn-Cunningham (1998) and Brungart (1998) discussed spatial hearing in the realm of multi-talker speech interfaces. Fernstrom and McNamara (2005) investigated the ability of using several streams of sound for browsing, which could be applicable to various types of human-computer interfaces. Frauenberger, Putz, Holdrich and Stockman (2005) proposed design patterns for interaction with auditory interfaces.

The use of auditory feedback has been an important element in designing accessible interfaces for the blind and visually impaired or applications used by sighted users in eyes-free contexts, such as driving. Brumby, Davies, Janssen, and Grace (2011) conducted a study that tested drivers using auditory and visual interfaces. The results of this study demonstrated better lane performance with auditory interfaces. Another successful use of auditory feedback can be found in BlindSight, an application for mobile devices in which users interact using a phone’s keypad and auditory feedback, without looking at the screen (Li, Baudisch, & Hinckley, 2008).

Lately, more successful applications are being developed showcasing the use of auditory feedback on touchscreen devices. In the Slide Rule (Kane et al., 2008) interface, blind and visually impaired users read the names of screen objects and functions on the mobile screen by employing finger gestures. Similarly, No-Look Notes is an eyes-free text-entry application that leverages the iPhone’s multi-touch feature to enhance the accessibility of using touchscreens for the blind and visually impaired (Bonner, Brudvik, Abowd, & Edwards, 2010).

5.1.3 AEDIN: An Acoustic and Touch Interface

AEDIN's purpose is to demonstrate another successful integration of touchscreen input and auditory output in an interface suitable for blind and visually impaired users. The overall interactive structure of AEDIN offers a game-like learning experience as users progress through a variable series of audemes and their associated essays; encounter, at unpredictable times, simple yes/no quiz questions based on the essays already heard; and accumulate points by correctly answering those questions (Ferati, Mannheimer, & Bolchini, 2009). The goal of the initial game is to accumulate as many points as possible. AEDIN's content consists of 49 audemes, each associated with and serving as a trigger for an individual essay. These audemes are arranged in a virtual, 7-by-7 grid of 49 squares (called the bookshelf), although AEDIN could easily be expanded to a bookshelf of 64, 81 or 100 squares. The entire 49 square bookshelf is not displayed on the screen, but instead is accessed through a screen designed with nine squares (the window). The window can be virtually moved throughout the bookshelf to access any of the 49 squares.

AEDIN features three modes of operation. The primary mode, called the Home Screen, features all of the audemes in a set arrangement on the bookshelf. In Shuffle Mode, the users employ a semi-circular left-to-right finger swipe on any square to trigger a *reshuffling* of the 49 audemes to display the eight most semantically relevant to the swiped audeme, which automatically relocate to the upper left corner of the window. The final mode is the Quiz Screen, which is triggered at random whenever the user has listened to two or more essays.

5.1.4 Auditory Interactive Dialogue Model (AIDM)

In order to support the design activity in screen- and vision-based applications (e.g., most common Websites), several design languages, notations and tools have been developed to describe and facilitate these applications at all levels, from the information architecture to the navigation and graphic interfaces (Ceri, Fraternali, & Bongio, 2000; Schwabe & Rossi, 1998). However, there is currently no way for designers to express design decisions for auditory applications at the proper level of abstraction. For example, although we can diagram the workflow of a visual interface, how can we describe the interactions possible through the auditory channel or through a mix of touch-based and auditory interaction? What concepts and notations could we use to represent sound-based feedback and auditory communication during the design process? As it is for Web applications, there are several benefits of having a design modeling language, including the possibility for designers to describe and master the complexity of the applications with a few, understandable concepts in order to discuss the overall requirements and strategic design decisions before locking in the interface details and to perform early validation before committing resources to a specific implementation or prototype.

For the process of designing AEDIN, a design modeling language for auditory applications was created by extending an existing interactive application design model (IDM, Interactive Dialogue Model (Bolchini, Colazzo, Paolini, & Vitali, 2006)). IDM was developed within the Web and hypermedia design community and, for the purposes of supporting the conception and representation of the interactive structure of systems based on the auditory channel, new primitives, terms and notations have been added. This new modeling language, called Aural IDM (AIDM), provides a set of concepts, terms and notation to describe the nature of the content, information architecture, navigation patterns and interaction mechanisms of the auditory applications (Ferati, Bolchini, & Mannheimer, 2009).

The set of primitives (i.e., notations from which other notations are derived) included in the AIDM are as follows:

Content Duration. This primitive provides information about the length of the content presented in audio format. To the user, the content is implicitly known by looking at an interface (for GUI) and linearly listening to the audio content (for AUI – Auditory User Interface). Most of the time, it is important to properly annotate this information in the design phase in order to calculate the overall content load. In other terms, aural time length can be specified for each unit of consumption (screen, page or dialogue act in IDM) and each topic or group of topics (structured information object). At AEDIN, the essays are a type of content synthesized through TTS and are generally approximately 1 minute long.

Location Sound Feedback. For a touchscreen auditory interface, this event is crucial. For a visual interface, the location of the mouse cursor is sufficient as a cue, but a need exists for a substitute when a mouse cursor is not used or cannot be seen. In this case, a touch-and-hear interaction achieves comparatively the same goal. The schema in Figure 5.3 depicts this primitive to be found in every subject of conversation between the user and interactive application. It is an event (sound) that notifies the user as to which part of the screen is being navigated or engaged. At AEDIN, the users tap on the screen with their fingers and, depending on the location of the finger tap, hear a different sound or synthesized TTS word. For example, if the user taps on the "Home" command button, he would hear a synthesized word saying "home." However, if he taps on content, a specific TTS would be heard that represents that content.

Content Change Sound Feedback. A dynamic change in content is an inherent process of digital artifacts for visual interface applications. In auditory interfaces, in which visual cues are absent or act as the secondary type of feedback, a need exists by which to devise auditory strategies to signal content change. This primitive presents an event (sound) that notifies the user that the content on the screen or page has changed. The content change could be due to a user's action, such as navigating to new content, or a system action, such as a screen or page time-out or pop-up information. The primitive is specified with each IDM introductory act, which is the list of topics belonging to the dialogue, and transition act, which represents the shift within the list of possible topics. Within AEDIN, this dynamic shift is represented as a sound played anytime the content on the screen is changed. This sound is very short, less than a second, and is an essentially abstract, computer-generated, yet pleasant sound (e.g., the blip sound used to announce the arrival of an email).

Background Sound. This aural primitive indicates the type of content for each unit of consumption (screen, page or dialogue act in IDM) for each topic or group of topics. This primitive is presented as a monotone sound usually with a lower volume that serves to easily identify the type of content presented. The sound has a neutral affect and represents acoustical non-action, such as a light wind sound. For this study, within AEDIN, different background sounds were used for different types of page content, in an effort to make users easily identify the screen or page on. For the main screen, the "wind" sound was used, but for the Quiz screen, the "time ticking" sound was used. This primitive will help users to know the general content of the page just by hearing the background sound. This feature is particularly helpful if the user is interrupted or moves away from the computer and, upon return, is able to immediately know the state of the application.

Interaction Meta-Information. Affordance is the quality of the visual interfaces that indicate or imply the type of action or interaction featured on the interface. The interaction meta-information primitive translates this feature for auditory interfaces. Each unit of consumption (screen, page or dialogue act) for each topic or group of topics needs to have this primitive in order to explain the transition acts and transition strategies that lead to a list of possible new topics or explain the way in which the user can access all of the new topics. The interaction meta-information primitive is the information on demand given to the user about the overall contents of the screen and the possible interactions within it. Often, users will forget what page or screen they are working on. In addition to the help provided by the feedback from the background and location sounds, there is still more that the user might need to know, such as available interactions on the screen, available commands and ways by which to interact with the available content. Within AEDIN, this information is provided through the Help button, which, in audio format, tells the user the general screen content, such as the buttons, essays and questions as well as possible interactions within the page, such as how to access the provided content.

Dialogue Unit Timeout. A typical feature in games is the constraint of the amount of time allowed for a certain goal to be achieved. Similarly, most account driven Website applications use the session timeout trait to enforce security. If the unit of consumption (screen, page or dialogue act) for a topic or group of topics needs to evoke a transition act or access new topics within a limited time, then this primitive is used as a notation. The unit timeout primitive provides a notation if the page or screen is due to expire after a certain amount of time. For dynamic contents, this notation is useful in that it lets the designer communicate the time length before the page will expire. For instance, AEDIN contains a pop-up quiz screen, which has a time-out of the total time the question is played plus 15 seconds within which the user has to submit the answer. Otherwise, the quiz screen will expire and the user will get bounced back to the main screen.

Conditional Navigation. This primitive notates a page as accessible only if certain conditions are fulfilled. Branching is a fairly well-known concept with applications, and as such, should be acknowledged in the design phase. This new primitive describes the category of possible subjects of conversation between the user and interactive application that may or may not occur depending upon the underlying interaction logic of the application. AEDIN has a feature called Landmine, in which the screen is populated with hidden "landmines" after the user has heard at least two essays. After the "landmine" has been engaged, the pop-up question is read to the user and he/she must submit the answer. Table 5.1 lists all of the AIDM primitives described above and a brief explanation about each one.

The diagram in Figure 5.4 depicts the overall interaction logic of AEDIN. It represents the three types of content: audemes, essays (TTS content related to the audeme) and questions (derived from the essay content). Each audeme is related to eight other audemes semantically similar. Each audeme is associated with one essay. Each essay has three questions associated with it, which are accessed only when the user listens to two or more essays.

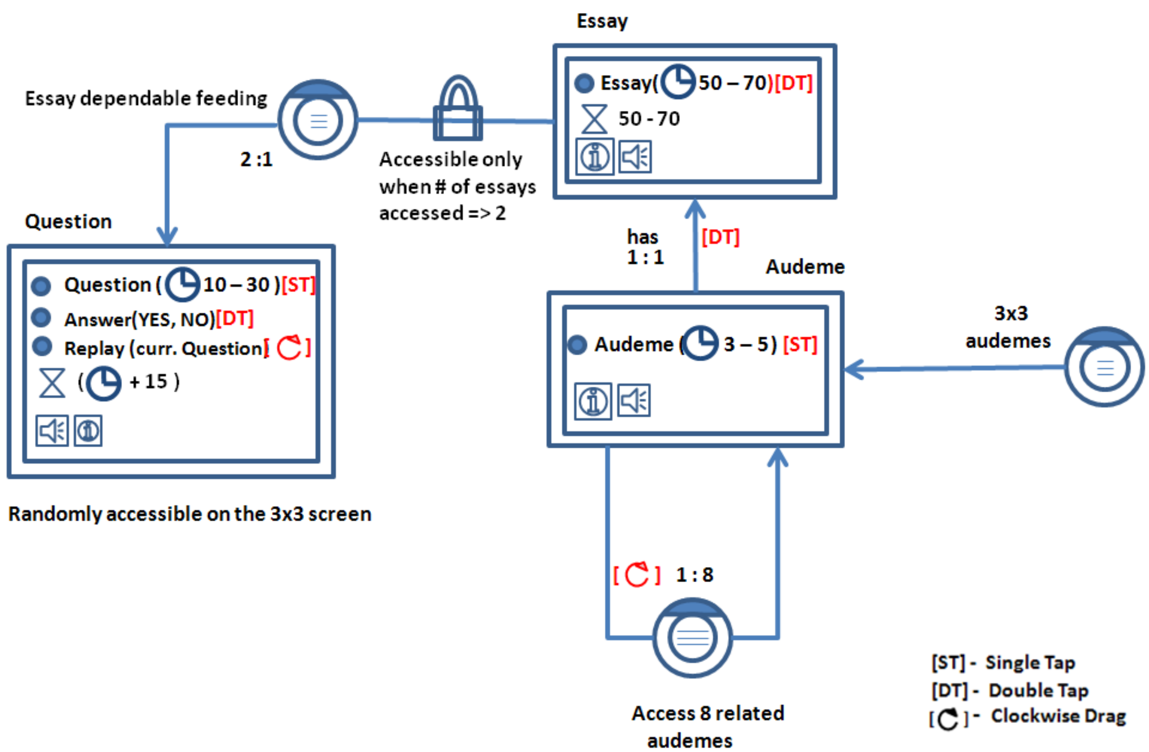

Figure 5.4. The AIDM dialogue map and legend for AEDIN.

Table 5.1. AIDM primitives and their explanations.

AIDM Primitive	Explanation
Content Duration	Provides the length of the content in audio format.
Content Change Sound Feedback	An audio-based notification that occurs when a change in the current content node occurs.
Background Sound	Sound cues to generally signify the type of content presented.
Interaction Meta-Information	Overview of the possible interactions available from the current application status.
Dialogue Unit Timeout	Annotates the automatic expiration of a dialogue unit.
Conditional Navigation	The navigational path available under certain conditions.
Location Sound Feedback	Gives feedback on the type of interface element activated by the user.

For each of the AIDM primitives in Figure 5.4 an explanation of the notation used is given in Table 5.2. The three notations marked with a star (Dialogue Act, Multiple Topic and Multiple Introductory Act) are inherited from the original IDM notations. The rest are new and correspond to the seven AIDM primitives.

Table 5.2. The list of the AIDM primitives used in AEDIN. (*Inherited from IDM).

Symbol	Description	Symbol	Description
	Sound Duration in Seconds		Conditional Navigation
	Content Change Sound Feedback		Page Information
	Background Sound		Dialogue Act*
	Page Time-out in Seconds		Multiple Introductory Act*
	Location Sound Feedback		Multiple Topics*

5.2 Designing the Acoustic Interface

5.2.1 AEDIN Navigation Modes: Bookshelf, Shuffle and Quiz Modes

AEDIN offers three main modes of navigation: the *bookshelf, shuffle* and *quiz* modes (Ferati, Manheimmer, & Bolchini, 2011). The *bookshelf* consists of audeme squares arranged in a 7-by-7 grid or virtual "bookshelf" of 49 "books" and is accessed through a “window” screen of nine books manipulated via a touchscreen monitor. Each book comprises an audeme, plus a recorded TTS essay of educational content. Put simply, the audemes function as aural covers and the essays function as the verbal content for the books.

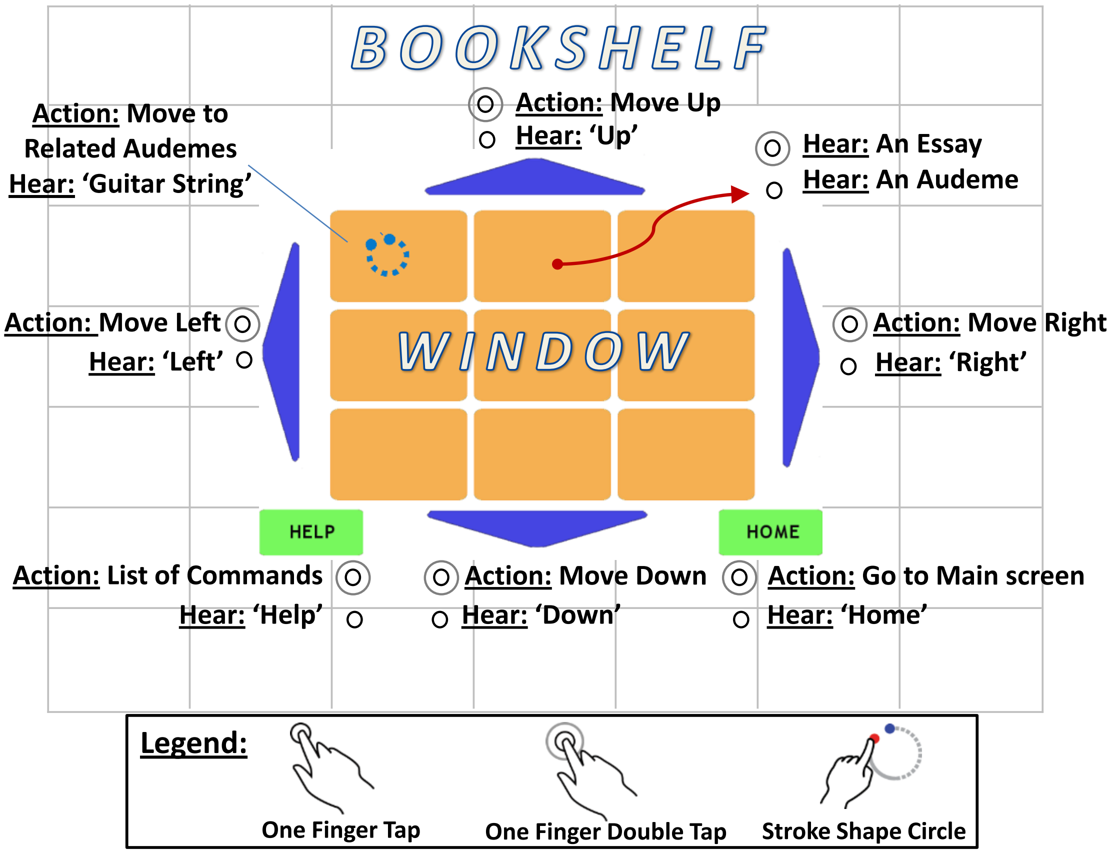

Figure 5.5. AEDIN interface; touch-based browsing through audemes.

AEDIN’s bookshelf is accessed and navigated through a 9 book window that occupies most of the monitor screen (Figure 5.5). The default position of the window is centered in the bookshelf, a position called *home*. There are also four movement-command buttons, one on each edge of the screen, graphically depicted as triangles. To manage AEDIN, the users may tap once on a square to hear its audeme. A quick double-tap elicits the attached essays. Similarly, a single tap on those movement buttons elicits TTS words for *up*, *down*, *left* and *right*. A quick double-tap on a button moves the window one row or column in that direction. Imagining the bookshelf as immobile, users may move the window throughout the entire bookshelf, a process requiring no more than four double taps to reach the farthest corner (e.g., up-up, left-left to reach the upper left-hand square). The arrangement of books in the bookshelf is fixed, but arbitrary, without chronological or alphabetical order. Only through book-by-book exploration can users familiarize themselves with the total layout. After moving the window, the user may use the *home* command button (always present in the lower right corner of the monitor screen) to instantly restore the window to the centered position. Successful bookshelf navigation relies on the user's "positional memory" for book location within this virtual space.

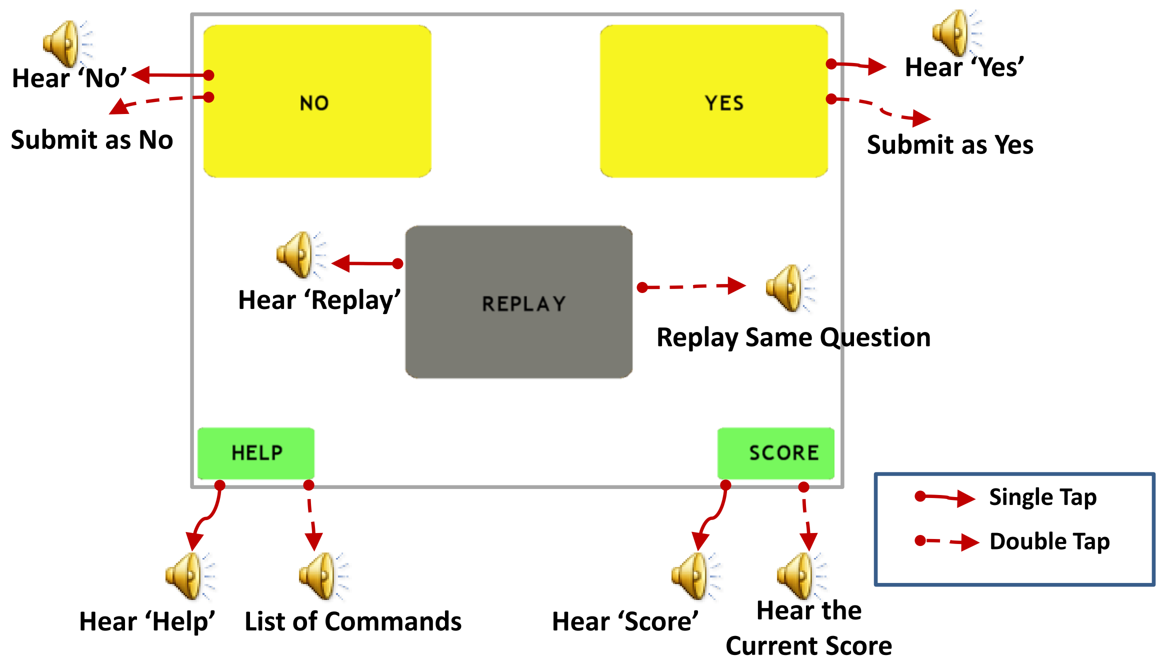

Figure 5.6. AEDIN quiz screen.

AEDIN also offers a semantically-based *shuffle* mode. In this mode, the metaphoric bookshelf is not needed because the window offers variable sets of books. The user may shuffle the window by first choosing a book as a starting point, then "shuffling" the window with a quick sweep of the finger across that book square. This touch command re-populates the window with other books semantically linked to that first book. This process may be repeated over and over with any book from any new population. To illustrate: The user chooses the book for "Automobile Industry" and shuffles the window with a finger sweep. AEDIN automatically repositions the chosen book in the upper-left square of the window. In the remaining eight squares, AEDIN generates a fresh arrangement of books thematically related to the Automobile Industry book, including "Transportation," "Interstate System," "Luxury," "Cities and Urban

Culture" and "Industrialization." The shuffle mode may be engaged at any time. However, once it is engaged, it continues to operate until the user disengages it by double-tapping the *home* button. The movement commands used to move the window in the bookshelf mode do not operate in shuffle mode.

These two navigational modes (bookshelf and shuffle mode) currently available in AEDIN are designed to engage distinct types of associational memory. Bookshelf navigation engages positional memory, while shuffle mode draws upon the user's memory of thematic relationships.

After the user has heard at least two essays, AEDIN's *quiz* mode (Figure 5.6) self-activates at unpredictable intervals. Announced by an explosion sound ('landmine'), the quiz mode reads a true or false statement based on the heard essays. The users respond by double-taping *yes* at the upper right (evoking a *right* answer) or *no* at the upper left (evoking a *wrong* answer). A large center button allows for question replay. Correct answers elicit more questions and a positive score. Incorrect answers send users to the past screen with no score. Additional AEDIN mockups are included in Appendix C.

5.2.2 AEDIN Sound Categories: Content Sounds, Feedback Sounds

AEDIN contains two sound categories depending on the context or meaning they convey: *content sounds* and *feedback sounds*.

5.2.2.1 Content Sounds

The great majority of sounds in AEDIN convey content to users through *audemes* or recorded TTS *essays* and *quizzes*.

The *audemes* in AEDIN are used to signify the content or theme of the associated verbal essay by offering users an iconic, metaphoric or micro-narrative representation of that verbal content. Although experienced audeme listeners can often anticipate the content associated with new audemes by decoding them quickly, audemes are generally used as mnemonic prompts that allow users to more easily remember the associated content after an audeme is heard (Manheimmer, Ferati, Huckleberry, & Palakal, 2009).

Essays are TTS recordings lasting between one and two minutes at 160 words per minute. The essay themes are taken from U.S. history lessons appropriate for the early high school level. Each essay contains a mixture of important figures, facts and analyses consolidated from reliable websites and standard textbooks.

Quizzes, presented as TTS, challenge users with true or false statements intended to test the users' knowledge of the heard essays. Each of the 49 essays offers three quiz statements drawn from essay content (total 147). Users hear these statements and then must answer the basic question: Is this true? Yes or no.

5.2.2.2 Feedback Sounds

These are the sounds that AEDIN uses to communicate information about the position of the objects on the interface or the running state of the interface. These feedback related sounds could be *positional sound feedback* or *background* sounds.

Positional sound feedback has the role of communicating to the user the object on which they just tapped. In that sense, audemes can function as positional sounds to inform users of their location within the bookshelf. In other words, when users tap anywhere on the screen, they will hear a sound revealing the underlying object. If the user taps on a functional command button, they will hear the aural label (TTS spoken name) of the command (e.g., *up* or *left* to indicate the intended movement of the window) (Figure 5.5).

Background sounds reveal the screen mode that the interface is running at that moment. Each screen features a different continuous background sound playing at lower volume whenever the user is not engaging anything on the screen. Once the user taps on an audeme or a button, the background sound stops. The sound resumes once the audeme or aural button label ends. Since the background sound runs continuously, it can get annoying quickly, thus the sounds were carefully selected to present soothing and calming experiences. For example, the home screen has a soft wind sound playing, the shuffle screen has a dripping water sound and the quiz screen plays the audeme related to essay that the question being presented at that moment belongs. Using the specific audeme as a background sound for the quiz screen helps the users perform better on the quiz. This decision is based on the results in experiment 1 showing that audemes can help users remember content (Manheimmer et al., 2009).

5.3 System Architecture

5.3.1 Platform

AEDIN was built using Microsoft's XNA Game Studio 3.1. The main reason behind using a game engine was the requirement that the interface should have the minimal possible latency to the user's interaction. Considering that the main content of AEDIN is sounds in *wav* and *aiff* format, playing these relatively large files requires computational resources, which could make the interface lag. A game engine, being a platform typically built to handle games with heavy graphics, was the logical choice within which to build the AEDIN interface that was loaded with sounds of different sizes.

In addition, XNA incorporates an audio authoring tool called Cross-platform Audio Creation Tool (XACT), which is used by XNA to organize sounds into *wave banks* and *sound banks*. *Wave banks* are files holding several wav files. *Sound banks* are files holding instructions on executing the sounds in the wave banks. The wave and sound banks are invoked by XACT from within AEDIN.

All non-speech sounds, such as audemes and background sounds, were created with Soundtrack Pro software. TTS content, such as essays, quizzes and aural labels for command buttons, were created using the Balabolka software (www.cross-plus-a.com/balabolka.htm), a free application that transforms text into a wav sound file.

One of the drawbacks of using XNA as a platform with which AEDIN was built was the relatively long time it took to develop, compared to, for example, Flash, which provides an interface for faster development. Another limitation is that AEDIN can only be installed and used on the Windows platform. Despite these drawbacks and limitations, the speed that AEDIN operates makes XNA a satisfactory solution for the target users, platform and initial goal (i.e., to evaluate the initial design decisions and identify general design principles through users testing).

5.4 Improving the Usability of AEDIN: Designing AEDIN v.2

AEDIN was improved based upon usability testing results and user comments given while being trained on how to use the interface (in detail discussed in Section 5.5). The usability evaluation was conducted at the ISBVI. Quick changes were done within one week on the speed of the essays, button design and improved feedback sounds.

5.4.1 *Increased Speed of the Essays*

For the initial AEDIN implementation, the TTS essays were recorded at a speed of 160 words per minute, a rate roughly equivalent to a lively conversation. ISBVI participants, however, considered this speed too slow, and requested an increase in the speed in order to make the listening less tedious. The speed of the TTS essays was increased to around 250 words per minute based on previous research showing that 275 words per minute is a rate that does not significantly affect comprehension or retention (Foulke, Amster, Nolan, & Bixler, 1962). In addition, this rate also conforms to research claiming that blind users can understand spoken content at a speech rate that is 1.6 to 2.5 times faster than sighted users (Asakawa, Takagi, Ino, & Ifukube, 2003; Moos & Trouvain, 2007).

5.4.2 *Designing for Overlapping Affordances*

In the first design iteration of AEDIN, the buttons were designed to act as the following:

- When single-tapped, would provide verbal feedback (e.g., a voice saying “up” for the up arrow).
- When double-tapped, would perform the expected operation and provide aural feedback to indicate its completion.
- Would strongly contrast visually with the background in order to aid the visually impaired users.
- Could be activated by touching the visual button.

It was discovered that all of the visually impaired users easily located targeted buttons. Most of the blind users, however, struggled in precisely locating buttons. Although remembering the general locations, they could not easily find the exact button locations without finger explorations. Significantly increasing the size of these buttons would have also changed their shapes and disrupted the design layout.

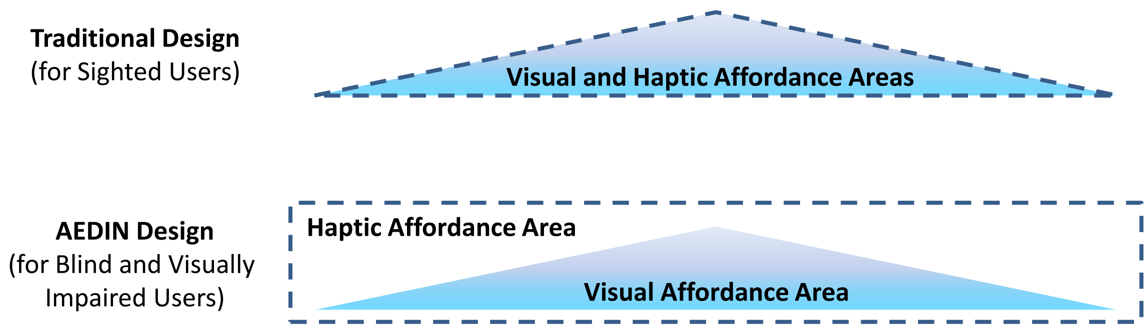

Figure 5.7. Overlapping, but not congruent visual and haptic affordance communication.

In order to solve this problem, it was discovered that the haptic affordance of the button (the area that can be activated through touch) does not need to coincide with the visual affordance (the area which is visibly a button). The visual affordance should be easily located by visually impaired users, whereas the haptic affordance should be easily reached by the blind. The two affordance areas would necessarily overlap for some parts, but the haptic one (although not visible) could take more screen real estate (Figure 5.7).

Therefore, in the second design iteration of AEDIN, the haptic and visual affordances were designed to be *overlapping, but not congruent*. A larger rectangular area around the triangular arrow button, for example, was available through touch (haptic affordance), although only the triangular button (visual affordance) was actually visible on the screen for the visually impaired users.

When the users were observed while experiencing the new design, it was noticed that all of the blind users were able to quickly locate and activate the buttons; however, the sighted and visually impaired users did not see any change in the interface and continued to easily use the buttons.

5.4.3 Improved Feedback Sounds

Another domain where a major change was conducted was within the feedback sounds, which, in some cases, seemed to annoy the participants. Initially the TTS voice conveyed the score each time the participant answered a question. This feature proved to annoy the participants and, thus, in the next design, the "Your Score is [number]" was left as an on demand feature accessible through the score button or by pressing the space bar on the keyboard.

Initially, most of the aural labels for the command buttons featured a robotic-sounding synthesized voice announcing the function of each (e.g., home, up, left). This voice was chosen in an attempt to create a humorous or distinctive atmosphere for the entire interface. The participants disliked this voice quality, so these labels were re-recorded with more humanlike voices.

In the initial design, I unintentionally created an ambiguous aural label for the button that triggered the replay of the question, labeling it simply “Replay.” The participants suggested that I change this aural label to “Replay Question,” a phrase which more clearly identified its function.

5.5 Experiment 3: Usability Evaluation of AEDIN

5.5.1 Overview, Hypotheses and Purpose

AEDIN was carefully designed following guidelines from existing acoustic interfaces created for blind and visually impaired users (e.g., McGookin, Brewster, & Jiang, 2008). For the evaluation of AEDIN, two usability tests were conducted at ISBVI. AEDIN v.1 was used in the first usability test. With the data gathered from the first usability test, AEDIN v.2 was designed by performing quick improvements, which were then used to conduct the second usability test (Figure 5.8). AEDIN was installed in ten touchscreen computers in an ISBVI lab. In both usability tests, the participants were given a 60 minute training session to familiarize them with the respective versions of AEDIN. Two days later, with no interim opportunity to practice, each participant was tested for 30 minutes on their ability to smoothly operate all aspects of AEDIN.

The purpose of the evaluation was to substantiate the AEDIN design decision and help identify key design principles for the effective design of acoustic interfaces. To this end, I will compare two versions of AEDIN and discuss the relationship between the two. This experiment has been approved by IRB #IRB-1009001883.

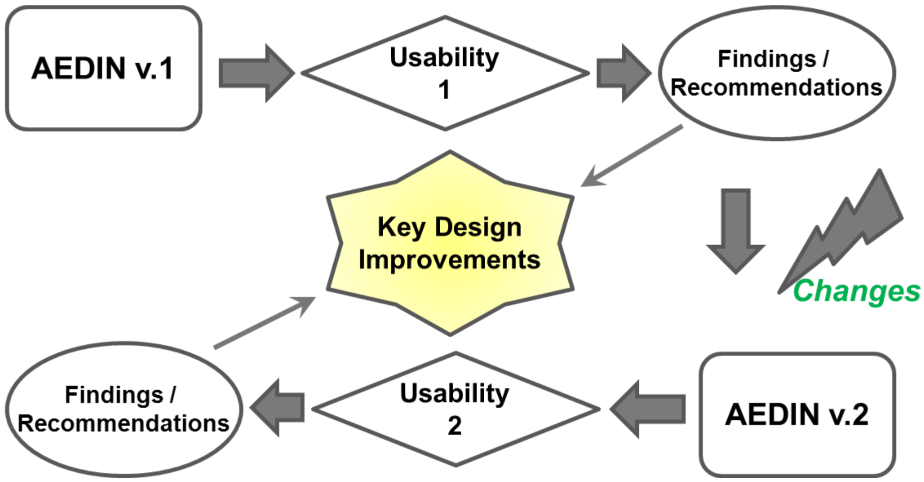

Figure 5.8. AEDIN design and evaluation flow.

5.5.2 Method

5.5.2.1 Participants

A total of 20 participants (10 male, 10 female; 7 blind, 13 visually impaired) took part in this study. Nine of the participants (4 male, 5 female; 2 blind, 7 visually impaired) took part in the usability test of AEDIN v.1. Eleven new participants (6 male, 5 female; 5 blind, 6 visually impaired) took part in the second usability test using AEDIN v.2.

A call for volunteers was distributed via the ISBVI administration to all middle- and high-school students. The students who volunteered for the research were randomly assigned to one of two groups. Initially, both groups were comprised of 10 students, but as one student assigned to participate in the first usability test missed the training session, this student participated in the second usability test. All of the participants were paid.

5.5.2.2 Design

For this study, a mixed methods design using quantitative and qualitative methods was used. The quantitative methods were the user satisfaction ratings, while the qualitative methods included recorded user comments that occurred while performing the tasks, open-ended questionnaires and observation notes. This type of design was used for both usability tests.

5.5.2.3 *Procedure*

The meeting with the participants for the first usability test took place in the ISBVI computer lab. The first session consisted of training and familiarization with AEDIN v.1. The nine participants took the hour-long training session at the same time. Two days later, the same participants underwent individual usability testing, four with one researcher in one room and the other five with another researcher in a different room. Each participant spent approximately 30 minutes performing eight tasks (see the list of tasks in Table 5.3) and completed questionnaires (Table 5.4) verbally with the answers recorded by the facilitator.

Table 5.3. Tasks used for AEDIN usability testing.

Task	Task Description
1	Play/hear three audemes
2	Go to the other audemes related to last audeme played
3	After you listen to an audeme, go to its essay
4	Listen to another essay containing a different audeme
5	Continue listening to different audemes until you hit a pop-up question; then, answer the question
6	Tell me your score
7	Describe the available command on the current screen
8	Explore the breadth of the grid by navigating UP, DOWN, LEFT and RIGHT.

Based on the observations of the training session and the participants' comments while using AEDIN v.1, improvements were made to create AEDIN v.2. This version was used during the following week to conduct a second usability test with the remaining 11 participants. The session tasks and questionnaire, however, remained the same as those used in the first usability test.

5.5.2.4 *Data Collection*

The following four types of data were collected in both usability tests performed with a total of 20 participants:

Participant Comments: Extensive notes were taken about the participants' comments during the training session with AEDIN v.1. The changes implemented for AEDIN v.2 were mainly derived from these comments.

Satisfaction Questionnaire: In both usability tests, after performing the tasks, each participant completed a satisfaction questionnaire based on their experience. There were 19 questions that the participants used to rate AEDIN on a five-point Likert scale (Strongly Disagree, Disagree, Neutral, Agree and Strongly Agree). In order to make the process smoother and avoid participant fatigue, the questions were read by the researchers and the participants answered verbally with the appropriate value. For example, a researcher read the question "Using the interface is enjoyable" and asked the participant to rate this from 1 to 5, with 1 indicating Strongly Disagree and 5 indicating Strongly Agree.

Open-Ended Questionnaire: After the task performance sessions and satisfaction questionnaires, a set of five open-ended questions was administered for each usability test. The goal was to gather broader insights into each participant's experience using more specific questions to uncover issues that were not anticipated earlier.

Observation: Researchers took notes on the participants' behaviors during the training sessions and while they performed the tasks as part of both usability tests. Important usability issues were uncovered through these observations. A complete list of questions can be found in Appendix B.

5.5.3 Results

5.5.3.1 Factor Analysis

Initially, in order to detect the relationships between the questions presented in both usability tests, the factorability of the 19 satisfaction rating questions was examined. Ten of the 19 questions correlated with each other, suggesting reasonable factorability (Table 5.4). The Keiser-Meyer-Olkin test of sampling adequacy was .51 and Bartlett's test of sphericity was significant, χ^2 (45) = 99.12, $p < 0.001$. Cronbach's Alpha was .731.

Four factors were extracted (see Table 5.5 for the question loadings for each factor). The first factor was the enjoyability of AEDIN, which explained 30.83% of the total variation. The second factor was the meaningfulness of AEDIN, which explained 22.68% of the total variation. The third factor measured the easiness of using AEDIN as a touchscreen interface, which explained 16.89% of the total variation. Finally, the fourth factor measured the appropriateness of the feedback sounds and explained 13.86% of the total variation.

Table 5.4. Extracted factors from satisfaction post-hoc questions.

Factor	Questions
Factor 1. Enjoyability of AEDIN	Q16. The interface is fun to use. Q15. Using the interface is enjoyable. Q17. I would use this interface again.
Factor 2. Meaningfulness of AEDIN	Q8. The audemes were meaningful. Q5. The bookshelf metaphor makes sense. Q4. Our explanation was sensible after you experienced the interface.
Factor 3. Ease of using the touchscreen	Q19. Using the touchscreen was comfortable. Q18. Using the touchscreen was easy.
Factor 4. Appropriateness of the feedback sounds	Q11. The feedback sounds were meaningful. Q13. The feedback sounds were short enough.

Table 5.5. Question loadings for each factor.

Rotated Component Matrix[a]

	Component / Factor			
	1	2	3	4
Q16	.932			
Q15	.855			
Q17	.811			
Q8		.915		
Q5		.850		
Q4		.815		
Q19			.951	
Q18			.950	
Q11				.866
Q13				.855

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 5 iterations.

5.5.3.2 Consistent Improvement

A comparison of the satisfaction rating for both versions of AEDIN revealed a trend in that AEDIN v.2 was rated higher than AEDIN v.1, especially on factor 4 (F4: Appropriateness of the feedback sounds), which demonstrated a significant difference ($p < .05$). For this factor, the participants using AEDIN v.1 ($M = 3.06$, $SD = 1.24$) gave lower ratings compared to participants using AEDIN v.2 (M = 4.05, SD = 0.88), t(18) = 2.09, p < .05, d = .92. The satisfaction rating for each factor is shown in Figure 5.9.

An interesting finding is the high rating given to both versions of AEDIN for factor 3 (F3: Ease of using the touchscreen), which is an indication that AEDIN's designed division of space for screen objects was successful. This success results from the physical size and placement of the designed buttons and affordances, the ability of users to easily orient the position of their hands to easily locate these buttons, and the easily-learned processes of AEDIN's functionality. The lowest rating, although still above average, was given to factor 1 (F1: Enjoyability of the AEDIN interface). The reason for this low rating can be found in the post-hoc open-ended questions in which the users commented that the interface provided too few and too easy quiz questions, too few audemes and essays, and a general lack of challenge in the game logic.

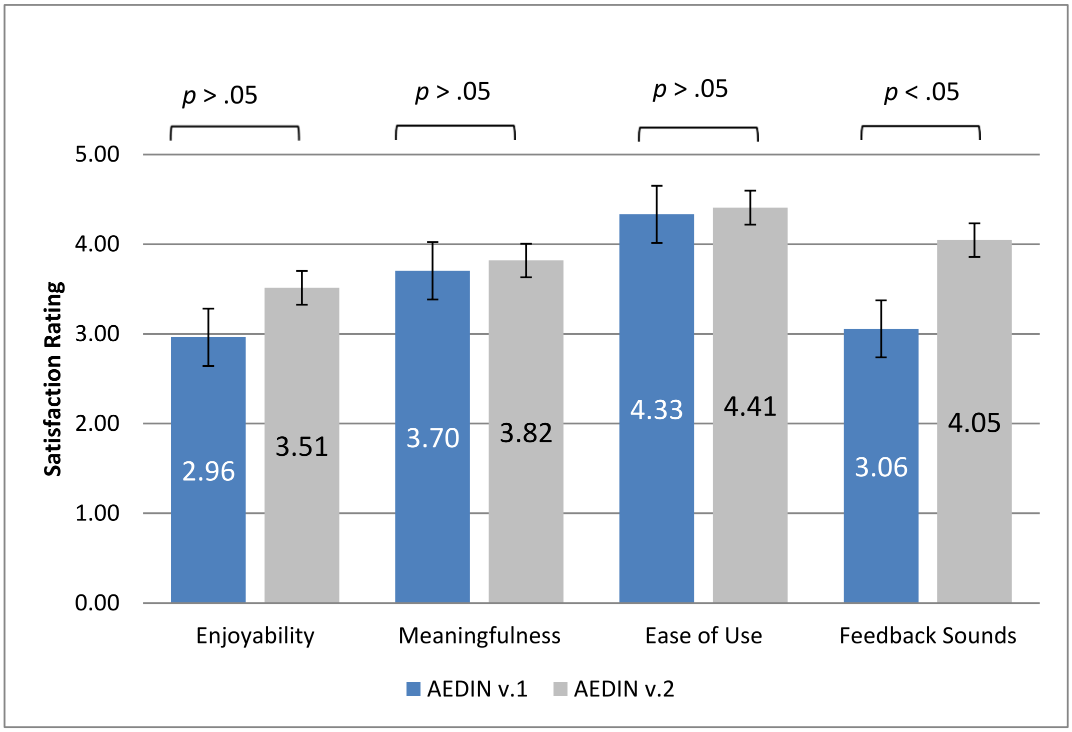

Figure 5.9. Satisfaction ratings for all four factors in both versions of AEDIN.

In order to understand the underlying reasons for the significant difference found in factor 4 (F4: Appropriateness of feedback sounds), a further investigation was conducted on Q11 (the feedback sounds were meaningful) and Q13 (the feedback sounds were short enough). In the further exploration of Q11, it was found that participants who conducted the usability testing with AEDIN v.2 (M = 4.27, SD = 1.01) rated AEDIN's feedback sounds significantly higher than participants using AEDIN v.1 (M = 3.00, SD = 1.50), t(18) = 2.26, p < .05, r = .47. This result is consistent with the fact that majority of improvements done from AEDIN v.1 to AEDIN v.2 were focused on the feedback sounds (see Section 5.4.3). The satisfaction rating for Q11, when comparing the two versions of AEDIN, is shown in Figure 5.10.

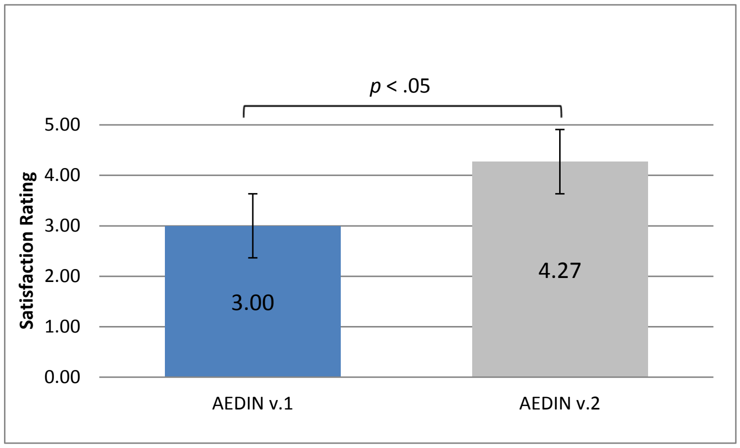

Figure 5.10. The feedback sounds in AEDIN v.2 were rated significantly higher than AEDIN v.1.

Although only one week was available to do quick improvements on AEDIN, the improvements resulted in a significant increase of enjoyability of the interface. Figure 5.11 shows that for Q15 (using the interface is enjoyable), a significant difference existed between the participants using AEDIN v.1 (M = 2.78, SD = 0.83) and the participants using AEDIN v.2 (M = 3.64, SD = 0.92), t(18) = 2.16, p < .05, r = .45. This result shows that AEDIN still needs improvement, but that significant improvements can be achieved with relatively minor changes.

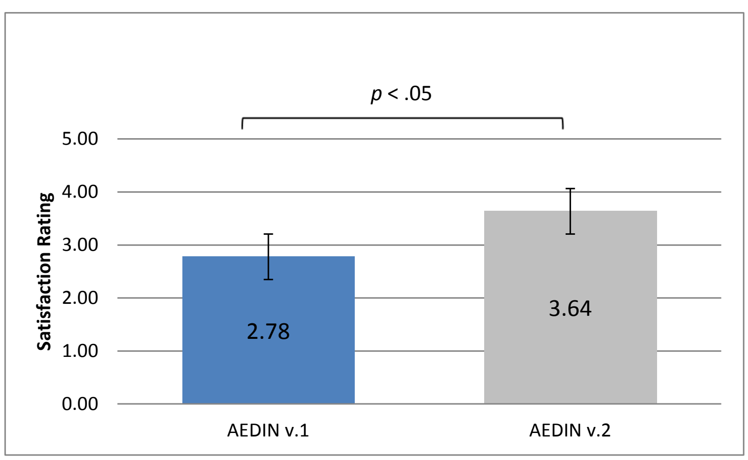

Figure 5.11. AEDIN v.2 was rated significantly more enjoyable than AEDIN v.1.

5.5.3.3 Gender Effect

The usability evaluation results show that the female participants (M = 4.05, SD = 0.88) rated the appropriate length of the feedback sounds higher than the male participants (M = 2.79, SD = 0.92), t(18) = 2.16, $p < .05$, r = .45. It appears that the female participants liked the sounds to be short and succinct rather than long. The satisfaction rating for Q13 (feedback sounds were short enough) when comparing the two versions of AEDIN for females and males are shown in Figure 5.12.

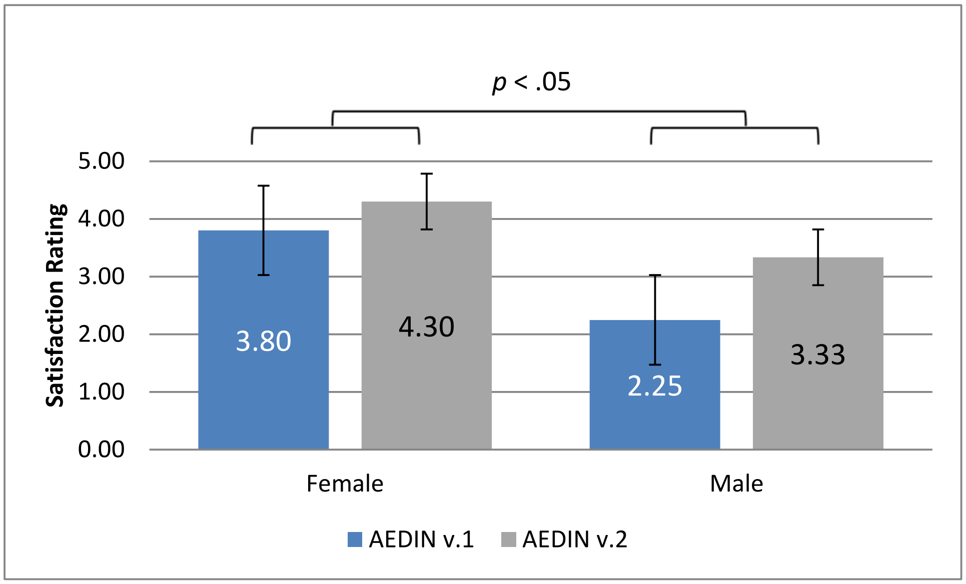

Figure 5.12. Female participants rated the length of the AEDIN feedback sounds significantly higher than the male participants.

5.5.3.4 Impairment Effect

For Q15 (using the interface is enjoyable), a significant difference between the blind and visually impaired was noticed in the perception of AEDIN being enjoyable. Overall, the blind participants (M = 3.7, SD = 1.00) rated the enjoyability of AEDIN significantly higher than the visually impaired participants (M = 2.86, SD = .69), t(18) = 3.05, $p < .05$, r = .58. Figure 5.13 shows that the blind participants noted an improvement in their enjoyability of AEDIN v.2. This result is in line with the changes implemented in overlapping affordances, where the goal was to increase the usability for the blind, but not negatively impact the usability for the visually impaired (see Section 5.4.2). Usability testing raw data set is included in Appendix A.

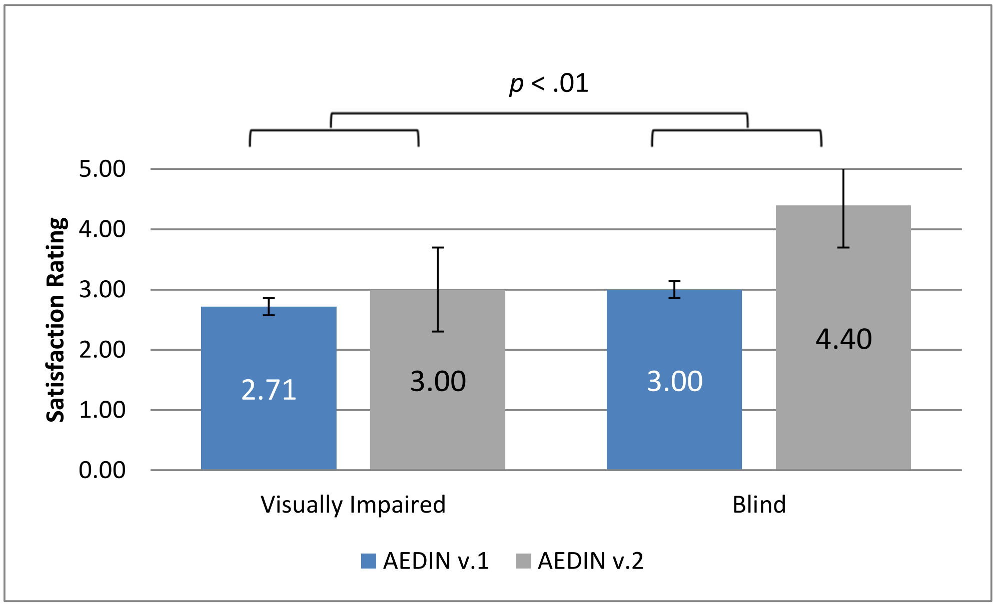

Figure 5.13. Blind participants rated the enjoyability of the AEDIN significantly higher than the visually impaired participants.

5.5.4 Discussion of Experiment 3

5.5.4.1 Simple Design Changes Yield Significant Improvements

One of the main findings of the usability tests indicated a consistent usability improvement from AEDIN v.1 to AEDIN v.2. The highest improvement was found in the use of the feedback sounds, which is not surprising considering that the greatest improvements were made in this part of the AEDIN design between v.1 and v.2.

5.5.4.2 The Touchscreens are Highly Usable for the Blind and Visually Impaired Users

Contrary to what may be a "commonsense" belief that touchscreens represent an interaction modality inaccessible and unsuitable to blind users, the results of this experiment reveal the highest satisfaction rating for touchscreen modality of interaction for blind and visually impaired. Although touchscreens inherently lack the tactile dimension of a haptic technology, such as Braille, they do offer a proprioceptive dimension (awareness of bodily positions) that can be strengthened by audio feedback. This same positional/audio combination is employed in playing the piano, where all of the surfaces are essentially indistinguishable except by virtue of their physical location relative to the user's body and the distinct sounds produced when different objects are engaged. A similar study (Kane et al., 2008) highlighted the importance of using the physical edges and corners of a touchscreen as useful landmarks for blind users. With its carefully designed feedback sounds and simple, logical layout, the AEDIN design leverages both the audio and positional dimensions.

5.5.4.3 Underdeveloped Game Logic

Compared to the overall satisfaction rating, the enjoyability of the AEDIN showed the lowest score, although it was still above average. This result addressed the underdeveloped design of AEDIN's game rules and logic, which did not provide enough mental challenges or playful experiences for this group of users. The participants commented that the quiz questions were too easy. In addition to adding more questions and making them more complex, a few of the students suggested penalizing players by subtracting points for incorrect answers. Further, based on several of the participants' comments that they too-easily familiarized themselves with the bookshelf layout and contents, AEDIN could be improved by populating the bookshelf with more audemes and essays, expanding it either two-dimensionally or even three-dimensionally in layers. This expansion is to suggest that there may be a natural "field of comprehension" or content map that users can easily hold in their minds for sets of ideas. Conceptual territories smaller than this field may be perceived as too simple and, thus, boring. Territories larger than this field may provide greater opportunities for a serendipitous experience of knowledge discovery or rediscovery. I speculate that further study of this idea could be enhanced by correlation to research into mental maps as well as domains, such as urban design. I noted that the blind and visually impaired users were accustomed to navigating large spaces based on aural cues. I also speculate that an inverse reciprocal relationship exists between the complexity of the "playing field" and its game logic (i.e., if one is perceived as too simple, then the other must assume greater complexity or variability in order to sustain user enjoyment).

5.5.4.4 The Female Users Rate Feedback Sounds Higher

An interesting finding is the gender difference in the perception of the AEDIN feedback sounds. Compared to the male users, the female users more positively rated the length of the feedback sounds as appropriate. I have no strong opinion as to what this finding may demonstrate, and neither the observations nor the qualitative data indicate an explanation for this result. Although previous research has produced relevant findings that female users prefer audio experiences more strongly than male users in navigational contexts (Siochos & Eisenhauer, 2006), how this finding correlates to a preference for shorter aural cues remains an open question. Further study should be dedicated to this matter.

5.5.4.5 Designed for an Aural Rather than Visual Experience

The goal when designing AEDIN was to make it usable for the blind and visually impaired, although the differences between these groups were not fully anticipated in terms of their expectations for screen-based user experiences. The initial design (AEDIN v.1) was rated as average in terms of being enjoyable in its use. The visually impaired users, who could discern the overall graphic design of the interface, commented that the interface was dull and not enjoyable to look at. This outcome was not surprising as, while designing AEDIN, the main focus was on providing a highly aural user experience and, as such, graphically simplistic and undetailed visual strategies and elements were chosen. The graphic quality of AEDIN will be addressed in future versions, with simple enhancements to the color scheme as well as other changes to the graphic elements. Again, I speculate that making the game logic or content infrastructure more complex or variable may compensate for the perceived simplistic quality of the visual design.

The dissatisfactions were different for the blind participants. It was enough to spend a few minutes watching them use AEDIN v.1 to discover the problems that existed with the size of buttons, which were not always easily accessible to them. In AEDIN v.2, the buttons were designed in such a way that the visual and haptic affordances were overlapping, but not congruent (see Section 5.4.2). This modification increased the enjoyability of the interface for the blind, while not making any negative impact for the visually impaired.

5.5.4.6 Further AEDIN Improvements

The usability findings gave me a clear indication on several ways in which the AEDIN interface could be improved. First, the game logic might be further developed to meet users' demands in regard to providing challenging experiences. For example, the levels of the challenges could be introduced in order to accommodate both the novice and expert users and provide appropriate complexity. Second, in order to address the needs of the visually impaired users, a more interesting visual interface with colors, graphic elements and images should be provided. One participant suggested providing an image to go along with each audeme/square. I recognize the appeal of such a multi-modal approach, but also appreciate the difficulty of introducing effective visual signifiers that would reinforce, but not cognitively distort the complex metaphoric impact of audemes. Third, the virtual dimensions of the bookshelf could be expanded to offer users more variety, complex discovery, retrieval challenges and serendipity. Fourth, the AEDIN infrastructure could be applied to knowledge/content domains other than education and, thus, offer experiences that appeal to different aspects of young lives, such as popular culture, music or social networking.

In order to ensure the usage of well-formed audemes in AEDIN, audemes were pretested in previous experiments conducted with the participants from ISBVI. While this process seemed viable for a limited set of audemes, the appropriate way to create well-formed audemes would be to establish guidelines that might ensure their effective generation for accurate meaning recognition. The next chapter describes the experiments conducted in order to identify the guidelines created to help design future well-formed audemes.

6. TOWARD NON-SPEECH SOUND DESIGN GUIDELINES FOR WELL-FORMED AUDEMES

6.1 Overview of Experiments

The design of non-speech sounds has been done based on intuition rather than following any guidelines (Pirhonen, Murphy, McAllister, & Yu, 2007). Audemes are not an exception to this tradition and thus far have been created in an ad hoc manner. In order to create design guidelines for the design of effective audemes in terms of information retention, three experiments were conducted. The first experiment (Experiment 4: Audeme-based Content Recognition over Time) investigated which audeme attributes (discussed in Section 6.1.1) made the most effective audemes. This experiment was also a longitudinal study helping to investigate the understanding and retention of information over a two-month period, testing participants on the first, fourth and seventh week after the initial exposure to the audemes. The second experiment (Experiment 5: Information Recognition without Reinforcement) was a manipulation of the first experiment in order to examine the effects on audeme meaning recognition of the first test after the exposure, which served as a reinforcement, thus positively influencing recognition for the fourth week test. The third experiment (Experiment 6: Audeme Learning Load) was conducted in order to establish the optimal number of facts that can be reliably associated with an audeme. These experiments have been approved by IRB #IRB-10070474B.

6.1.1 General Experimental Design and Procedure

We investigated the structure of audemes as composed of source, semiotic and syntactic attributes.

Source attributes are the types of sounds used to create audemes. Based on the typical components of non-speech sounds found in the literature, two broad groups of sound types were distinguished:

- **a. Music:** segments of songs, including instrumental songs, of varying genres: classic, rock, etc.
- **b. Sound effects:** pre-recorded or artificially created sounds. This group of sounds consisted mainly of (i) abstract sounds, which do not normally occur

in nature, and (ii) everyday listening sounds, which are sounds occurring in nature and common enough to be familiar to listeners.

Semiotic attributes are the different modes of listening that influence the types of meaning audemes that can convey. The three modes of listening (causal, referential and reduced) can be applied to audemes, but the referential mode is the one that gives audemes an important strength by utilizing an individual's emotions as they are related to the object to which the audeme is attached. The semantic mode of listening does not apply to audemes (at least not yet), since they are not fully established to mean one specific thing, in the way that a specific word does.

a. Causal listening: Sounds used for the creation of audemes were classified according to the listening mode targeted. For instance, if a sound of a chicken was used to represent a concept named *chicken*, then a note was made that the sound was used in the causal mode, since it directly represented the *object* causing the sound, in this case, the chicken itself.

b. Referential listening: If the same sound of a chicken, however, was used to represent a concept named *farm*, then it addresses the referential mode of listening, since it did not directly represent the object, but what the object could represent, in this case, a farm where chickens might be found.

c. Reduced listening: Sounds were classified as being used in the reduced mode when the characteristics of the sound (rhythm, pitch and intensity) gave insights about the concept. For instance a sound with fast rhythm indicated a fast moving object, such as *fast cars,* while a sound with a slow rhythm indicated a slow moving object, such as an *old man walking.*

Syntactic attributes are the combinations of two or more sounds used to create an audeme. Similar to the study conducted by Brewster et al. (1995b), in which compound earcons were created with the concatenated of individual earcons, I examined audemes created by the concatenation of individual sounds. The difference between the studies is that Brewster et al. created compound earcons from earcons, while I created audemes from individual sounds, not audemes. Hence, two types of combinations are distinguished:

a. Serial*:* the combination of two or more sounds played one after another.

b. **Parallel*:*** the combination of two or more sounds played at once.

These three audeme attributes are important in regard to testing the recognition of audeme meanings as those attributes represent the most atomic level that can be manipulated in order to create an audeme. Table 6.1 lists the audeme constructs to be used to depict the manipulation of the source (music, sound effect) and syntactic (serial, parallel) attributes. The semiotic (causal, referential or reduced) attributes were manipulated by grouping all of the sounds used in creating the audemes into one of three listening modes.

The concepts associated with the audemes used in these experiments were based on content from a variety of academic topic suitable for K-12 school students, including history and geography. In order to generate the audemes, a collection of sounds was used that were previously used in studies involving blind and visually impaired students from the ISBVI. Five audemes for each combination were created in order to increase the data points and generate additional statistical power.

Table 6.1. Audeme combinations with syntactic, source and semiotic attributes used in the three experiments. The number of audeme created for each attribute is provided in parenthesis.

<table>
<tr><th>Syntactic Attribute</th><th>Source Attribute</th><th>Semiotic Attribute</th></tr>
<tr><td rowspan="4">Serial (20)</td><td>Music + Music (M+M) (5)</td><td>Referential-Referential (3)
Reduced-Reduced (2)</td></tr>
<tr><td>Music + Sound-Effect (M+SFX) (5)</td><td>Referential-Causal (3)
Referential-Referential (2)</td></tr>
<tr><td>Sound-Effect + Music (SFX+M) (5)</td><td>Causal-Referential (2)
Referential-Referential (3)</td></tr>
<tr><td>Sound-Effect + Sound-Effect (SFX+SFX) (5)</td><td>Referential-Causal (4)
Causal-Referential (1)</td></tr>
<tr><td rowspan="3">Parallel (15)</td><td>Music + Music (M+M) (5)</td><td>Reduced-Reduced (4)
Referential-Referential (1)</td></tr>
<tr><td>Music + Sound-Effect (M+SFX) (5)</td><td>Referential-Referential (4)
Referential-Causal (1)</td></tr>
<tr><td>Sound-Effect + Sound-Effect (SFX+SFX) (5)</td><td>Referential-Referential (3)
Referential-Causal (1)
Reduced-Reduced (1)</td></tr>
</table>

6.1.2 Research Questions

The following research questions are central to this study.

Q1: How well do audemes aid in recognizing information?

The goal of this question is to establish the level of effectiveness that audemes have in relationship to recognizing any textual content that is associated with them. Once this information is established, it will be important to determine how long an audeme can be effective, so that reinforcement can be applied in order to maintain the effective link between the audeme and the content it is representing.

Q2: Which combination of audemes would be the best for accurate information recognition?

Audemes are created from different attributes (as discussed in Section 6.1.1) and it is important to understand which combination is the most effective in regard to the recognition of the content associated to them.

The outcome of these research questions will help to understand the nature of audemes along with their strengths and weaknesses. Moreover, it will help derive a set of initial guidelines for designing effective audemes to be used by acoustic interface designers.

6.2 Experiment 4: Audeme-based Content Recognition over Time

6.2.1 Overview

The aim of this experiment was to identify the characteristics of an audeme that will positively impact the correct recognition of the meaning associated with it. In order to record these effects, three tests were conducted over a period of two months. All seven audeme combinations listed in Table 6.1 were used in this experiment. The independent variable was the types of audeme combinations. The dependent variable for this experiment was the recognition score of the content associated with the audeme (Table 6.2).

Table 6.2. Independent and dependent variables for experiment 4.

Independent Variables	Dependent Variables
Audemes	Content Recognition Score

The following three hypotheses guided this experiment.

H1: Audemes created from a serial combination of sounds will yield a higher meaning recognition when compared to audemes created from a parallel combination of sounds.

This hypothesis is based on a pattern recurrently seen in past experience with audemes, in which audemes created from a serial combination of sounds were more effective when compared to audemes created from parallel sounds.

H2: The recognition of the meaning of audemes will decrease with each additional testing due to the weakening of memory.

No matter the combination methods used (serial or parallel), it is expected that audemes will decrease in performance with the passing of time. In other words, since, in this experiment, the participants were only exposed to the concepts that they need to remember once, it is expected that the recognition of the meaning of the audemes will decrease over time due to the weakening of memory.

H3: Audemes created from sounds in the causal and/or referential modes of listening should yield higher meaning recognition when compared to audemes created from sounds in the reduced mode of listening.

It is expected that audemes created from sounds in the causal and/or referential modes of listening will generate better meaning recognition. The expectation is based on the fact that sounds in the causal and referential modes of listening, contradictory to sounds in the reduced mode of listening, reveal information that can be directly associated to the concept that the audeme represents.

6.2.2 Method

6.2.2.1 Participants

Eight participants (4 male and 4 female) were recruited for the study and were paid for their participation. These were students from the ISBVI and were between the ages of 12- and 20-years-old. Five of the participants were blind and three were visually impaired, but they all used Braille.

A notification via the teachers was sent to all of the students at the ISBVI asking for their participation in the research. Informed consent was given to participants' parents prior to their involvement in the research.

6.2.2.2 Stimuli

Along with the seven groups of audemes, all of the students were given three concepts to remember for each audeme. It was decided to use three concepts in order to provide rich and various semantics for each audeme. The concepts were thematically close to the audeme. For instance, the three concepts *burning the barn*, *crazy farm* and *vandals at the zoo* were associated with a SFX/SFX audeme consisting of *breaking glass* and *crying sheep* sounds.

Using a professional-grade sound effects library, different sounds were combined into an audeme. The audemes were generated using the Soundtrack Pro (http://www.apple.com/finalcutpro) software. All of the generated audemes were five seconds long. This length represents an intermediate value in the range of audeme lengths used in previous studies, which were between three to seven seconds (Mannheimer, Ferati, Huckleberry, & Palakal, 2009). No space was left between the sounds constituting the audeme.

Over the years, my strategy of audeme creation has become more consistent as those were continually pre-tested with students at ISBVI who confirmed the well-formed audemes and helped discard confusing ones. This strategy proved to be true in the domain of U.S. history. This domain was decided upon because it was relatively common knowledge across all of the participants as part of their standard curriculum in the school. Complete stimuli can be found in Appendix B.

6.2.2.3 Apparatus

The participants collectively listened to the audemes during their initial exposure, as well as for all subsequent tests of recognition. The audemes were played using a laptop with amplified speakers. The test was conducted using Braille or large-font printed test forms, on which the participants marked their answers using crayons. The meetings with the participants took place in one of the classrooms at the ISBVI location (Figure 6.1).

The concepts were read out loud by the researcher, while the students marked their answers on Braille or large-font printed sheets. For instance, the researcher would say: "The concept for 1B is a *car racing*." On the students' sheets, it only read "1B."

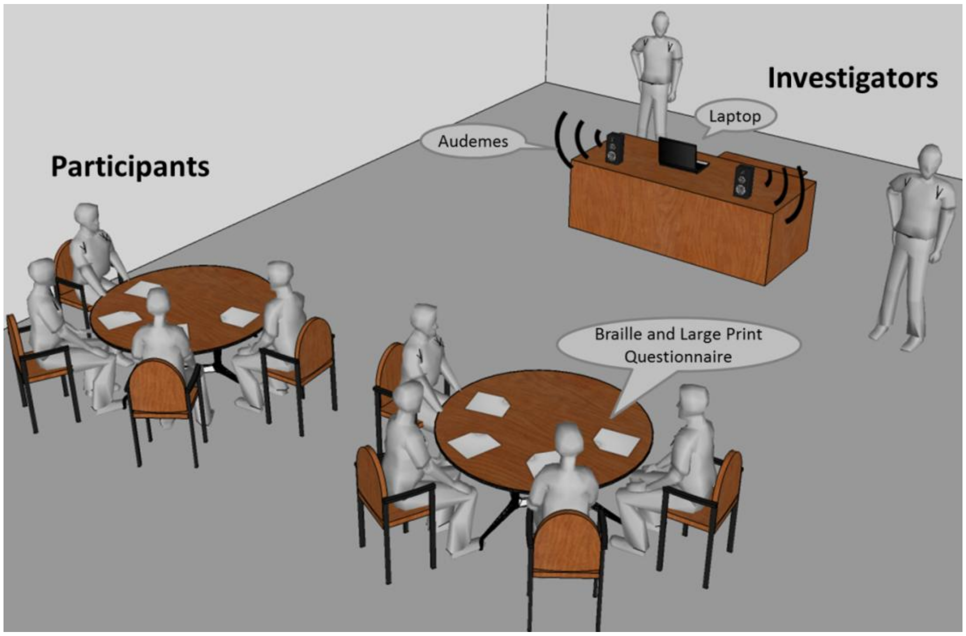

Figure 6.1. The physical setting in which the experiments were conducted at the ISBVI.

6.2.2.4 Procedure

Covering all seven of the audeme combinations (35 audemes) would have taken longer than the time that the students were available for a session (one hour). Hence, the audeme combinations were exposed to the student over a two weeks period (four combinations during the first meeting and the remaining three combinations in the second week). Initially, the participants were exposed to 35 audemes and 105 concepts (across the two weeks). The testing on the exposed audemes and concepts was conducted three times: after one week (T1), four weeks (T2) and seven weeks (T3) (Figure 6.2). The experiment sessions at the ISBVI were held once a week for an hour.

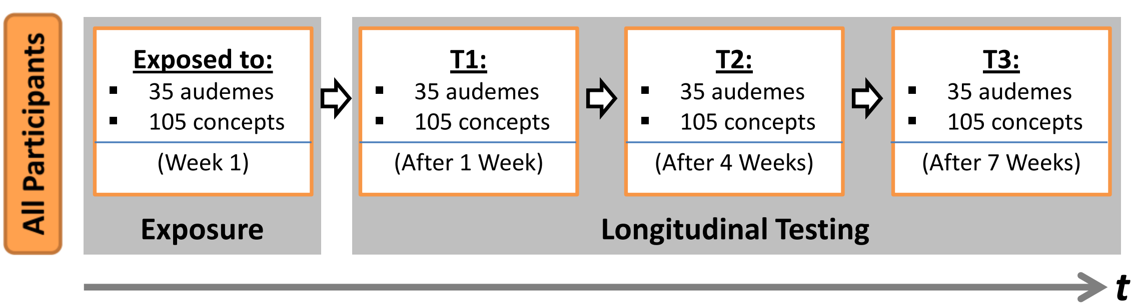

Figure 6.2. 35 audemes presented across four periods of time.

For each exposure, an audeme was played and then three thematically related concepts were read. This process was repeated (the audeme and three concepts) another time, giving the participants a better chance at creating an association between the audeme and concepts. The testing procedure consisted of using a multiple choice questionnaire on which the students had to mark the three correct concepts among the nine given, six of which were used as noise (Figure 6.3).

Music+Music Serial 1 audeme

Concepts

1. Many seeds make a heap
2. Dance as an ongoing social customs
3. Discovering science
4. Road assistance
5. Regression
6. New wave trendy to old-fashioned
7. Healthy environment
8. Blast from the Past
9. Future to past

Music+Music Serial 2 audeme

Concepts

1. Windowless building
2. Vivid imagination
3. Action into calm
4. Sweating while dancing
5. Retaining wall
6. Angelic movement
7. Stop light
8. Swinging - into the air
9. Gone with the wind

Music+Music Serial 3 audeme

Concepts

1. Cloudy sky
2. A modern king
3. Groovy Cool Establishment
4. Peeking into the laboratory
5. Prince Harry (young, hip Prince of England)
6. Satellite radio
7. Red truck
8. Majestic-hipster young guy
9. Outdoors and indoors

Figure 6.3. A depiction of the three audemes each representing nine concepts, in which the underlined concepts are the correct ones, while the six other concepts are given as noise. A complete list of all audemes and concepts can be found in Appendix B.

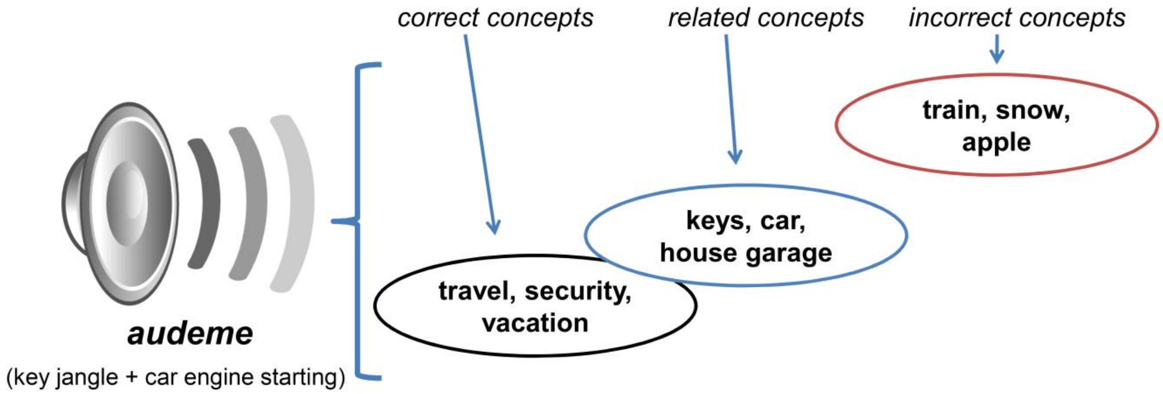

Figure 6.4. An example of the three levels of concept groups associated to an audeme in regard to testing recognition rates.

Out of the six concepts used as noise, three were related to the initial correct concepts, while the other three were entirely different from the initial concepts. An example of concepts is shown in Figure 6.4, in which the concepts *travel*, *security* and *vacation* were the given concepts associated with an audeme consisting of the sounds for *key jangle* and *car engine starting*. The concepts of *keys*, *house garage* and *car* were used as the related concepts because, semantically, they relate to the audeme and the correct given concepts. In such fashion, the concepts of *train*, *snow* and *apple* were given as the incorrect concepts because they do not relate to the audeme or the correct concepts. This division was done in order to understand not only when the participants answered correctly or incorrectly, but also when they understood the general meaning of the audeme, although they might forget the specific wording of the concept associated with it.

This method of testing the concepts was repeated three times (testing periods) in order to investigate the longitudinal effects on the memorization of the concepts associated to the audemes. In every repetition of the test, the three correct concepts to which the participants were initially exposed were kept the same, but shuffled with six *new* noise concepts consisting of related and incorrect concepts. The participants were not exposed to audemes or concepts between testing sessions and were never given the correct answers to the audeme meanings.

6.2.3 *Results*

Learnability: A common struggle in education is learnability and a learner's ability to memorize information. In this experiment, the audemes have shown a potential to help overcome this issue. Over a period of eight weeks, the participants showed an ability to remember the meaning of the presented audemes. Although the expectation was that with each additional testing the participants would forget the meanings of the audemes, the opposite was found to be true as they recognized the concepts more accurately with each week of testing.

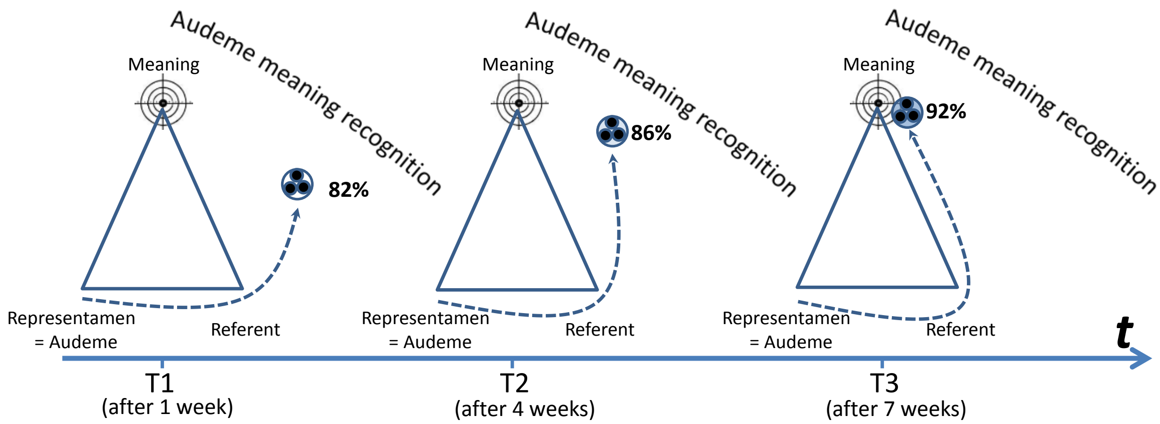

Figure 6.5. The recognition of the target meaning of the audemes increased over time.

As shown in Figures 6.5 and 6.6, the overall increase from T1 (the test after one week) to T2 (the test after four weeks) was 4.67% (from 2.45 to 2.59), and from T2 (the test after four weeks) to T3 (the test after seven weeks) was 6% (from 2.59 to 2.77).

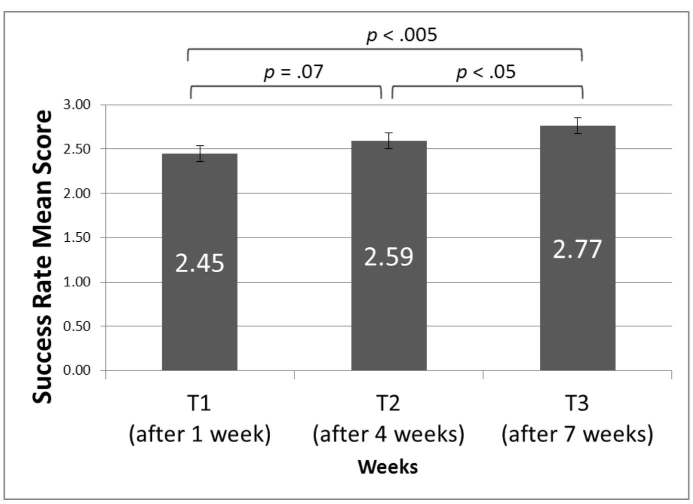

Figure 6.6. The mean recognition score of the audeme meaning for the tests after one (T1), four (T2) and seven (T3) weeks.

The results were analyzed using a repeated measure ANOVA within-subjects design. The analysis showed an overall significant increase, $F(2, 14) = 17.32$, $p < .001$, $\eta^2 = .71$. The Bonferroni pairwise comparison revealed a significant increase between T1 and T3 ($p < .005$) and T2 and T3 ($p < .05$). It also showed an approaching significance between T1 and T2 ($p = .07$).

The differences among the audeme types became lesser with time: It was interesting to see that the findings of the test after one week (T1) revealed that the differences among the seven types of audemes were high. However, the subsequent tests after four (T2) and seven (T3) weeks revealed that these differences among the audeme types became less with the passing of time. The repeated measures ANOVA revealed significant differences among the seven audeme types for T1, $F(6, 36) = 3.54$, $p < .01$, $\eta^2 = 0.38$, and T2, $F(6, 36) = 4.24$, $p < .01$, $\eta^2 = 0.41$, but no significance for T3, $F(6, 36) = 1.59$, $p > .05$, $\eta^2 = 0.21$. The sample means for the audeme types per the three tests are shown in Figure 6.7.

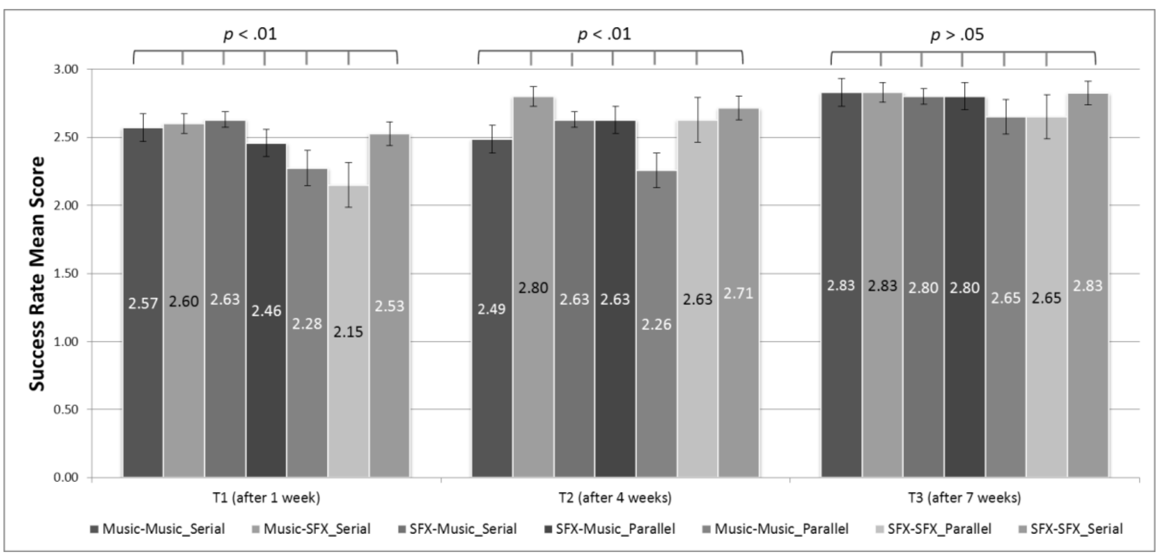

Figure 6.7. The mean recognition scores of the audeme meanings for the different audeme types, as they vary over the three tests (T1, T2 and T3).

It was also interesting to find the best and worst audeme types. The repeated measures ANOVA revealed an overall significance among the seven audeme types across all weeks, $F(6, 36) = 3.12$, $p < .05$, $\eta^2 = 0.34$. Further, the Bonferroni test revealed individual differences between the two types of audemes, music-music parallel and music-SFX serial, showing that the worst audeme type was music-music parallel and the best was music-SFX serial ($p < .01$), as shown in Figure 6.8.

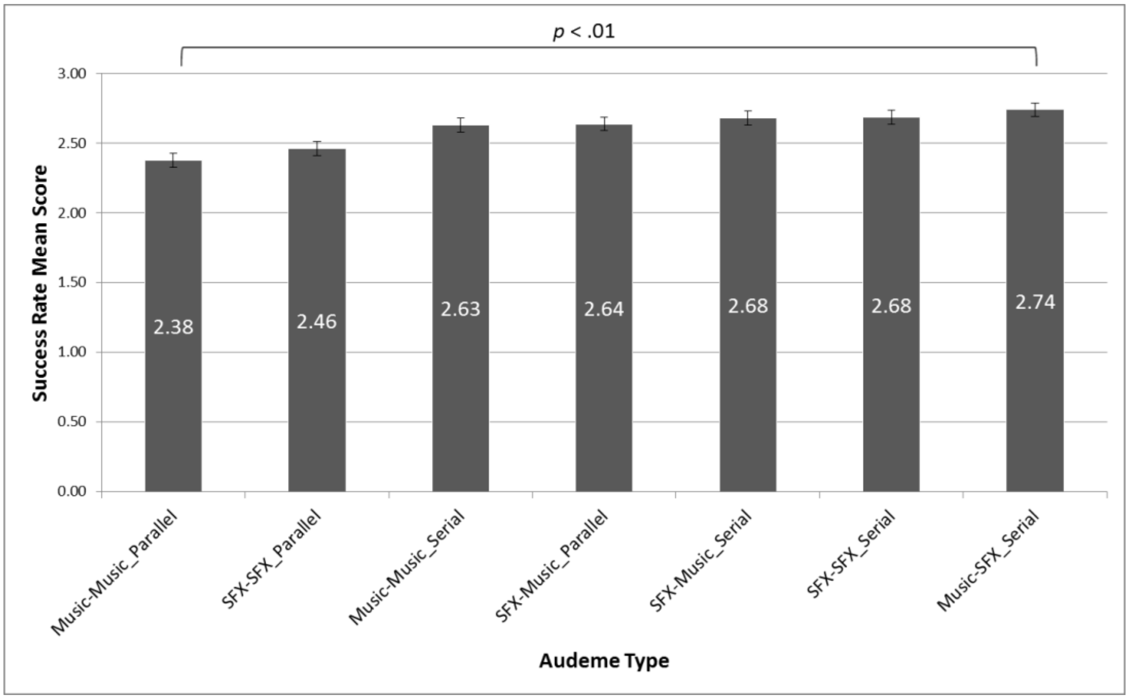

Figure 6.8. The mean recognition score of the audeme meanings for all of the audeme combinations, showing the best (MS_Serial) and worst (MM_Parallel) audeme combinations.

Differences among the audeme sound orders became less with time: Similar trends of differences among the audeme types were observed with audeme sound order (Music-Music, Music-SFX, SFX-Music and SFX-SFX) as shown in Figure 6.9. The repeated measures ANOVA revealed significant differences among the audeme order for T1, $F(3, 18) = 4.15$, $p < .05$, $\eta^2 = .41$, and T2, $F(3, 18) = 7.32$, $p < .01$, $\eta^2 = .55$, but not for T3, $F(3, 18) = 0.76$, $p > .05$, $\eta^2 = .11$. This result is an indication that with the passing of time, due to learnability effects, the order of sounds used in the audeme combinations become less effective. In other words, all of the combinations are equally good.

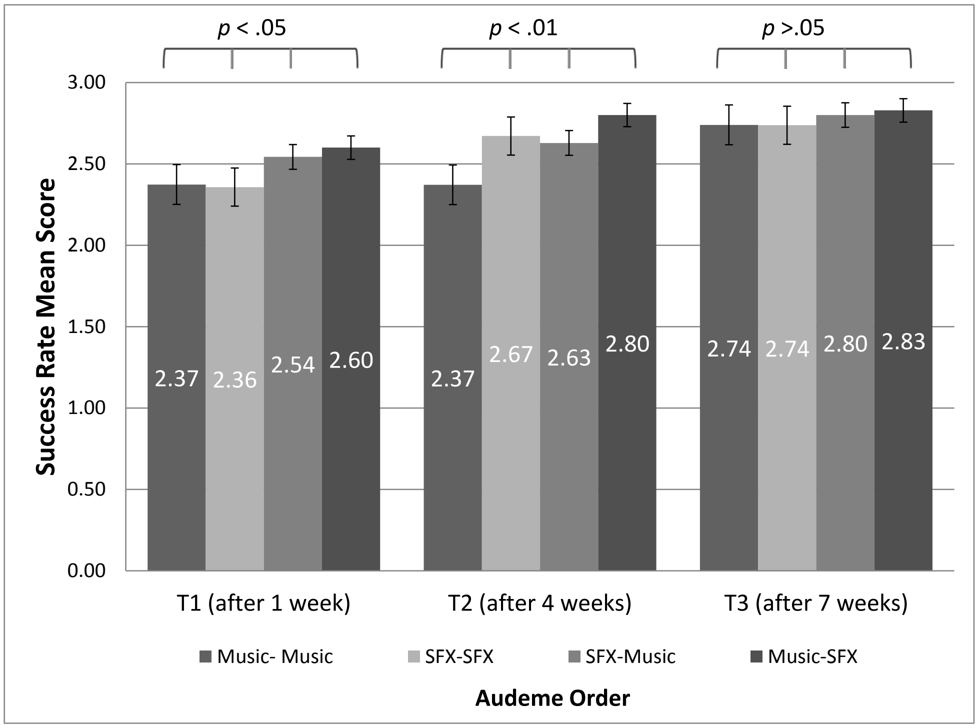

Figure 6.9. The mean recognition scores of the audeme meanings for the audemes created with a different order of source attribute sounds as they varied over the three tests (T1, T2 and T3).

Females performed better than males: Fortunately for this experiment I had an equal number of participants divided by gender (4 male, 4 female), which helped in tracking any gender effects. The independent measures t-test analysis across all three tests revealed that the female subjects (M = 2.76, SD = 0.12) recognized more concepts than the male subjects (M = 2.43, SD = 0.16). This difference was significant, t (6) = 3.62, p < .05, r = .83. The sample means for the males and females are shown in Figure 6.10.

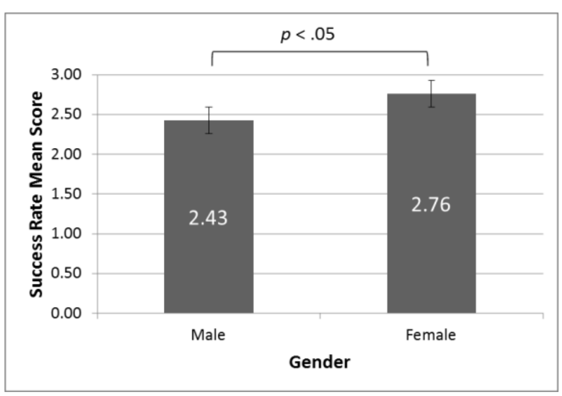

Figure 6.10. The mean recognition scores for the audeme meanings as they varied by males and females.

Serial combination is better than parallel: Throughout this longitudinal experiment, it was consistently observed that serial combinations of audemes were better than parallel combinations. The repeated measures t-test showed that the serial audemes (M = 2.69, SD = 0.20) helped in the recognition of concepts more than the parallel audemes (M = 2.50, SD = 0.30). This difference was significant t(7) = 4.71, p < .005, r = .87. The sample means for the parallel and serial combinations are displayed in Figure 6.11.

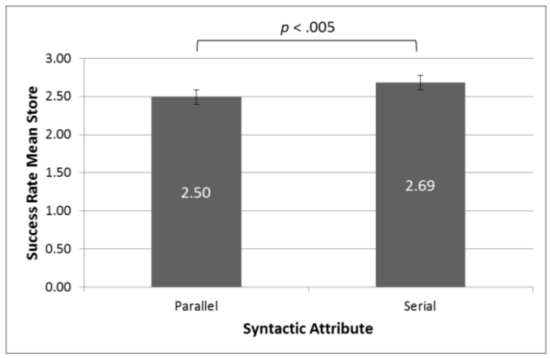

Figure 6.11. The mean recognition scores for the audeme meanings as they vary for the audemes created with parallel and serial combination of sounds.

Mixing music and SFX yields better recognition: Audeme combinations were also analyzed in a higher layer, that is, just in terms of their combination, whether constituted of similar (Music-Music, SFX-SFX) or different (Music-SFX, SFX-Music) types of sounds (Figure 6.12). The repeated measures t-test showed that the audemes created from different types of sounds (M = 2.69, SD = 0.25) helped the participants to recognize the right concepts with a higher accuracy than the audemes created from the same type of sounds (M = 2.54, SD = 0.22). This difference was significant t(6) = 3.95, p < .01, r = .85.

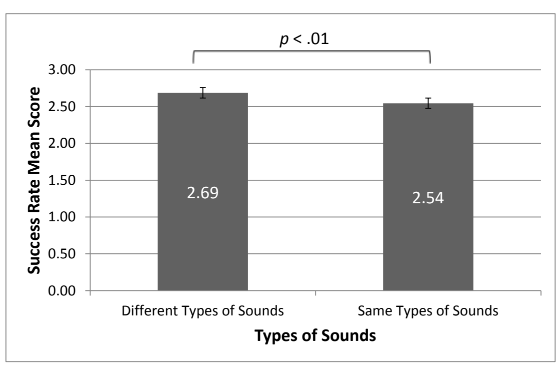

Figure 6.12. The mean recognition scores of the audeme meanings for the audemes created with different and same types of sounds.

Causal and referential modes of listening yielded better recognition of the audeme meanings: The results indicated that when the sounds composing the audemes were used in the causal and referential modes, then the recognition of the concept was significantly better, while the reduced mode of listening proved to be the least effective. For the first sound (out of the two composing the audeme), the repeated measures ANOVA revealed an overall significance, $F (2, 14) = 22.25$, $p < .001$, $\eta^2 = .76$. Further, the Bonferroni test revealed individual differences among the three modes of listening: referential and reduced ($p < .005$), reduced and causal ($p < .005$), and a non-significant difference between referential and causal ($p > .05$) (Figure 6.13).

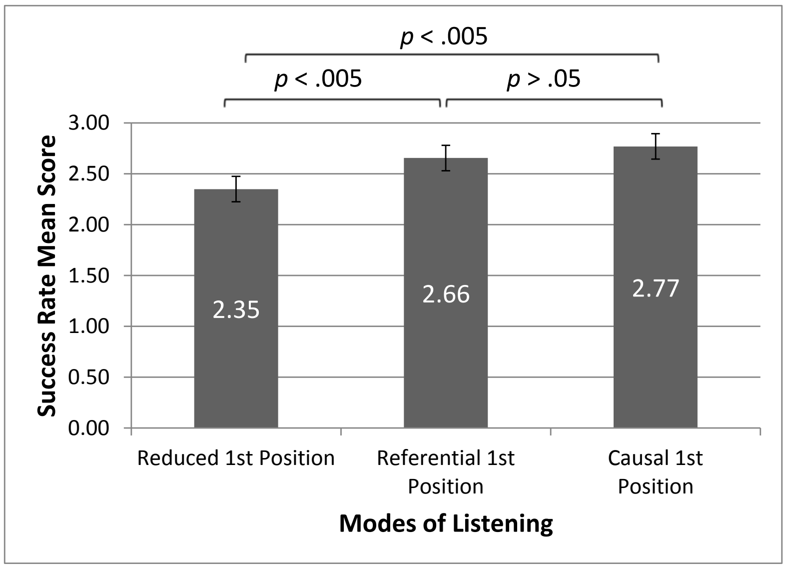

Figure 6.13. The mean recognition score of the audeme meanings for the audemes created with sounds in the different modes of listening in the first position.

Similar results were found for the second sound, $F (2, 14) = 25.87$, $p < .001$, $\eta^2 = .79$. The Bonferroni pairwise comparison revealed significant differences for the referential and reduced ($p < .05$) and reduced and causal ($p < .001$), but not for the referential and causal ($p > .05$) (Figure 6.14).

An interesting finding was that in order to be the most effective, the audemes consisting of different modes of listening were better (Figure 6.15). A repeated measures t-test was conducted in order to track the audemes that constituted of sounds with same ($M = 2.56$, $SD = 0.24$) or different ($M = 2.69$, $SD = 0.24$) modes of listening. The results revealed a significant difference between the two groups, $t(7) = 5.32$; $p < .001$, $r = .89$.

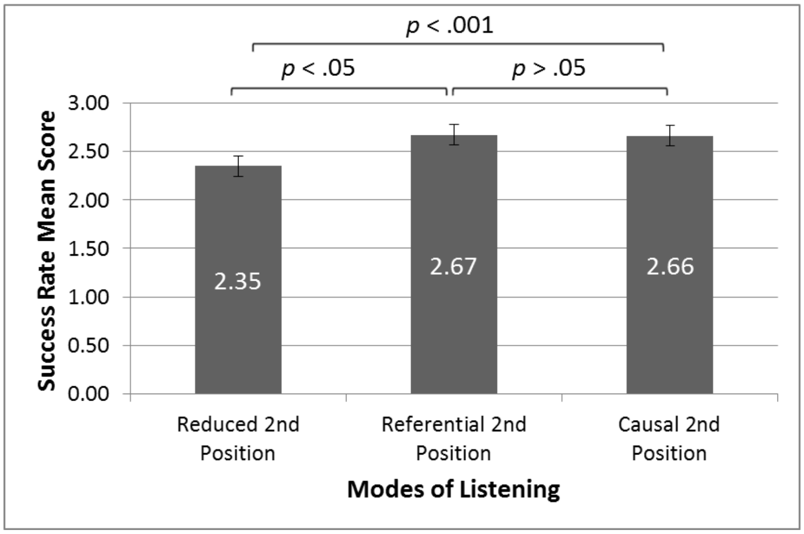

Figure 6.14. The mean recognition scores of the audeme meanings for the audemes created with sounds in the different modes of listening in the second position.

The rhythm and timbre of sounds did not yield any differences in the recognition of the audemes: In this experiment, the audemes, particularly the sounds used, were analyzed in terms of two additional dimensions: the rhythm and timbre of the sounds. In terms of rhythm, the sounds were classified into two groups, fast (e.g., high rhythmic drums) and slow (e.g., walking footsteps), and their performance was tracked. In terms of timbre, the sounds were classified into two groups, harsh (e.g., metal scraping, electric guitar, etc.) and smooth (e.g., light wind, violin, etc.), and their performance was tracked.

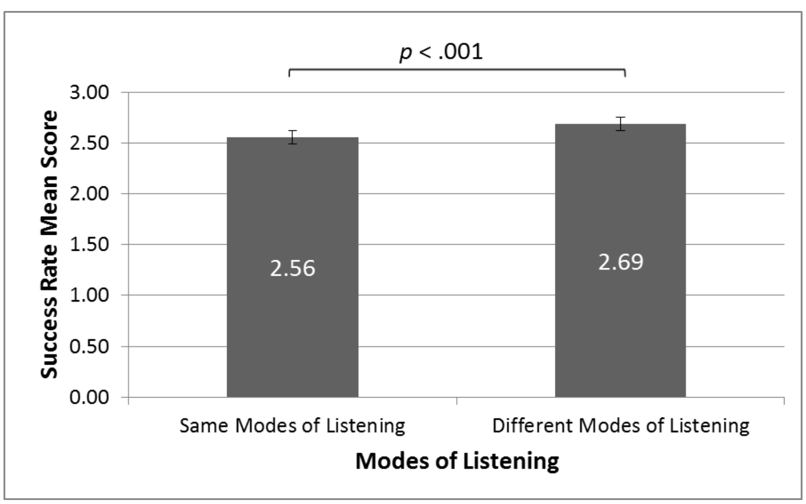

Figure 6.15. The mean recognition scores of the audeme meanings for the audemes created with sounds from the same and different modes of listening.

The t-test analysis revealed no significant differences between the sounds of different rhythms and timbres. In addition, audemes created with different or same rhythm sounds yielded no significantly different recognition rates. Similar results were found for the audemes created with different or same sound timbres. The experiment raw data set is included in Appendix A.

6.3 Experiment 5: Audeme Meaning Recognition without Reinforcement

6.3.1 Overview

While experiment four was used to assess recognition over time, the goal of experiment five was to identify the amount of reinforcement effect that the T1 test had after one week in regard to the recognition of the concepts. With this idea in mind, a decision was made to skip the T1 test after one week, and only conduct the T2 test after four weeks. In order to explore this condition, experiment four was replicated with all its elements, and the T2 test after four weeks was performed, but not T1 and T3, the tests after one and seven weeks, respectively. This experiment was conducted three months after experiment four. The independent variables were the different types of audeme combinations and the prior test serving as reinforcement. The dependent variable for this experiment was the recognition score of the content associated to the audeme (Table 6.3).

Table 6.3. Independent and dependent variables for experiment 5.

Independent Variables	Dependent Variables
Audemes Reinforcement Test	Content Recognition Score

The decision to conduct this experiment came after noticing that each testing in experiment four increased the meaning recognition of the audemes. This finding led to the investigation of whether the first test actually reinforced the meanings of the audemes. Hence, the following was hypothesized.

H4: The test in the first week, in addition to being a test, also served as an additional exposure to the audemes and concepts.

6.3.2 Method

6.3.2.1 Participants

The same participants from experiment four participated in experiment five.

6.3.2.2 Stimuli

A new set of audemes and concepts were generated for this experiment in order to avoid any memorization effects carried over from experiment four. All of the stimuli are provided in the Appendix B.

6.3.2.3 Apparatus

The same apparatus was used as in experiment four (see Section 6.2.2.3).

6.3.2.4 Procedure

The procedure for this experiment was the same as used in experiment four, but without the testing after the first week (T1). The difference between experiments four and five is that the period between the exposure and testing was one week for experiment four, but four weeks for experiment five.

6.3.3 Results

The main goal of this experiment was to measure the reinforcing effects of repeated exposures through a comparison of these results with experiment four. As described earlier, this experiment is a replica of experiment four minus the T1 test. Since a continual increase in the recognition of the concepts in experiment four was seen, it is important to establish how much testing after one week accounts for that learnability effect.

A repeated measures t-test was conducted in order to compare the scores of the test after four weeks (taken from experiment four—T2E4) with the additional test after one week ($M = 2.59$, $SD = 0.27$) and scores of the test after four weeks (taken from experiment five—T2E5) without the additional exposure to the audemes provided by T1 ($M = 2.04$, $SD = 0.19$). The results show that the participants had significantly better recognition when tested after four weeks with the preceding test after one week (with reinforcement) when compared to being tested after four weeks without the preceding test (without reinforcement), t (6) = 6.25, p < .001, r = .93. The sample means are shown in Figure 6.16.

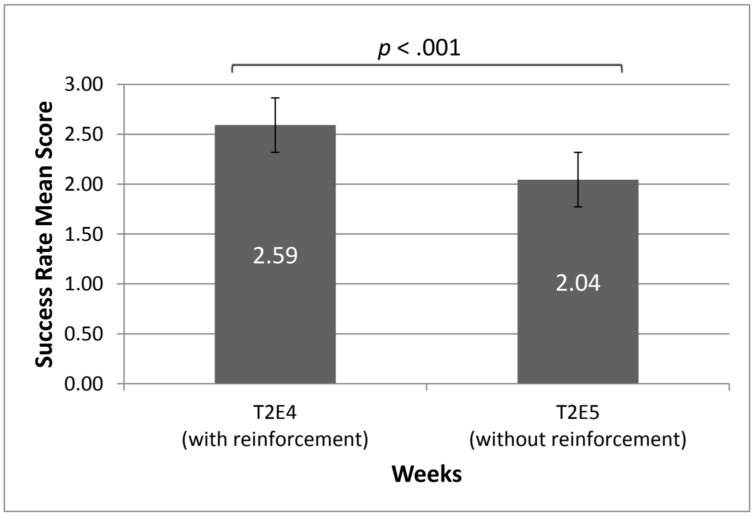

Figure 6.16. A comparison of the mean recognition scores of the audeme meanings for week four with prior testing (with reinforcement) and without prior testing (without reinforcement).

The results of this experiment show that the testing in the first week after the initial exposure had a great impact on the recognition of the concepts. This finding concludes that the first test after one week, in addition to serving as a test, also inadvertently serves as reinforcement of the meaning of the audemes. The experiment raw data set is included in Appendix A.

6.4 Experiment 6: Audeme Learning Load

6.4.1 Overview

The aim of this experiment was to investigate the ability of an audeme to help retain facts, in particular the goal was to determine the optimal number of facts that can be correctly associated with one audeme. In order to increase the scope of meaning in this experiment, statements or sentences instead of concepts were articulated. The reason for this decision was that statements are true or false, while concepts are not. In essence, statements establish relationships between concepts. Therefore, a fact was considered to be the crucial information derived from a single sentence (e.g., *Shenandoah River and its* *valley are in Virginia*). In addition, the sentences chosen for this experiment were determined by the nature of the sounds available in the school's digital library.

For this experiment, a new set of audemes was created, but covered all possible combinations as shown in Table 6.1. The new variable was the number of facts, and each audeme was presented along with three, five or seven facts. The number of facts for each audeme was randomly assigned in order to ensure an even distribution of three-, five- and seven-fact audemes. The independent variables were the different types of audeme combinations and the number of facts associated with them. The dependent variable was the recognition score of the content associated with the audeme (Table 6.4).

Table 6.4. The independent and dependent variables for experiment 6.

Independent Variables	Dependent Variables
Audemes Number of Facts	Content Recognition Score

I maintain that audemes will be most effective when associated with a low number of facts and as the number of facts associated to it increases, the meaning recognition of the audemes will decrease. Therefore, in this experiment, the number of facts is systematically increased until a threshold is reached that establishes the number of facts that an audeme can be associated with.

6.4.2 Method

6.4.2.1 Participants

Ten participants (5 male, 5 female) from the ISBVI took part in this experiment. Eight of the participants had taken part in experiments four and five, while two new participants were recruited. The participants were between 12- and 20-years-old. They were paid for their participation. Six of the participants were blind and four visually impaired, but all used and were proficient in Braille.

6.4.2.2 Stimuli

Along with the audemes, the students were given sentences with single facts to remember. For purposes of this test, a single fact was established to be a single sentence holding a fact from U.S. history. In such manner, the same audeme combination (e.g., M+M means music and music in Serial) was presented along with three, five or seven sentences. For this experiment, a total of 21 audemes were played: 3 (levels of the number of facts) x 7 (levels of the combination of audemes), and the total number of sentences read was 105. All of the stimuli are presented in the Appendix B.

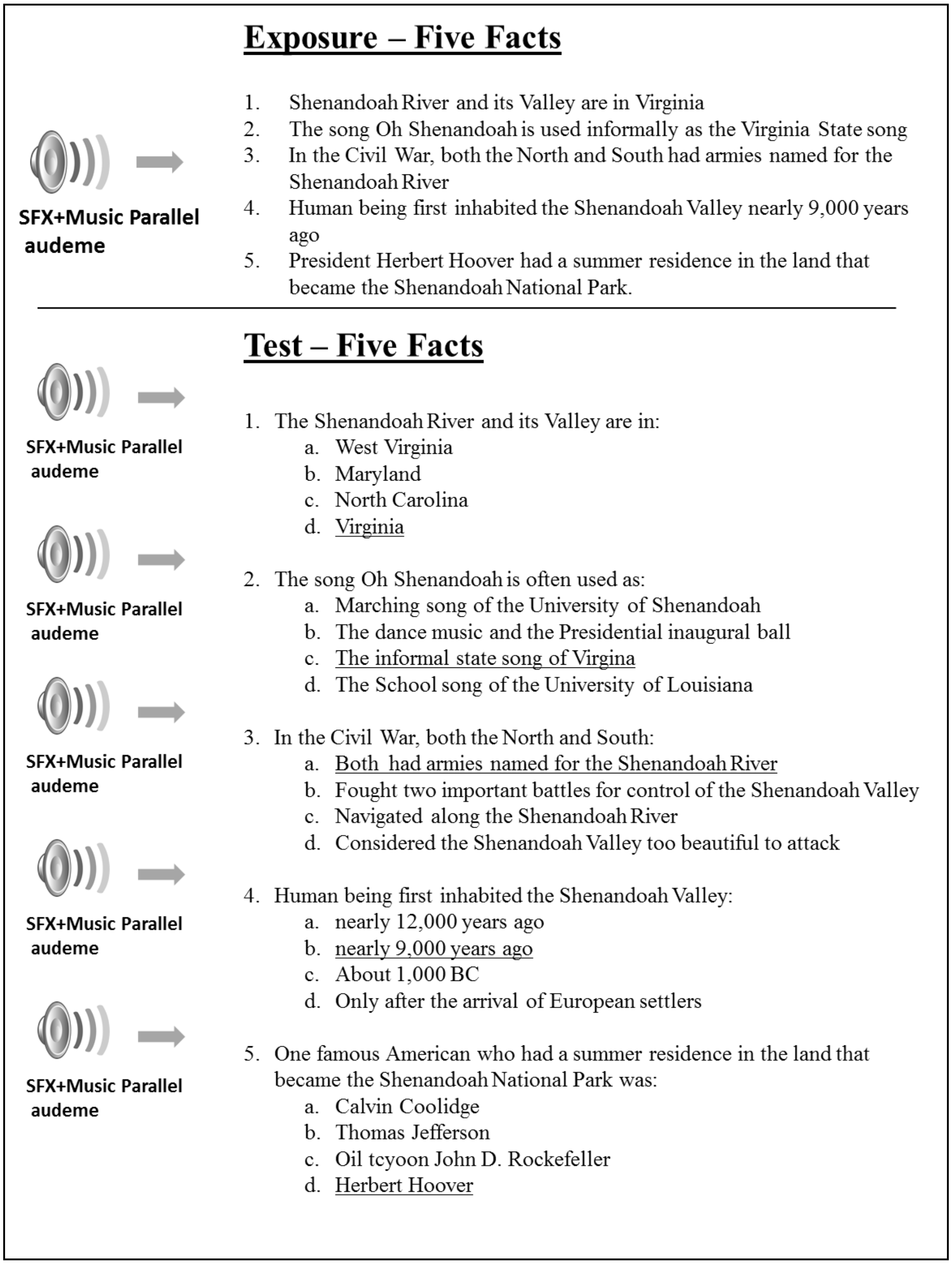

Figure 6.17. A depiction of an audeme being presented to signify five facts. After a week, the same audeme was played between each of the five questions (a question per fact). The correct answer is underlined.

6.4.2.3 Apparatus

The same apparatus was used as in experiments four and five (see Section 6.2.2.3).

6.4.2.4 Procedure

The fulfillment of this experiment was stretched over a six week period, where half of the weeks were used to expose the students to sentences and the following week to test for their recollection of the facts (Figure 6.17).

6.4.3 Results

The results revealed that audemes are best when associated to five facts, as participants scored best with those audemes. A repeated measures ANOVA showed an overall significant difference among the three groups, $F(2, 18) = 6.14$, $p < .01$, $\eta^2 = .41$. The Bonferroni pairwise comparison revealed a significant difference only between the audemes associated with five and seven facts ($p < .05$), while the other differences were not significant. The sample means are shown in Figure 6.18.

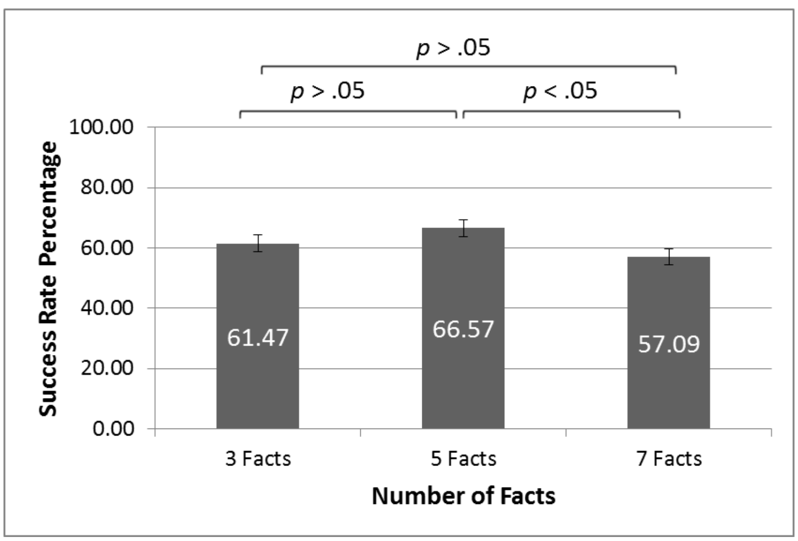

Figure 6.18. The mean recognition scores of the audeme meanings for different audeme combinations as they vary by the number of facts associated with the audeme (three, five or seven facts).

Gender Effect: Consistent with the previous experiment, the female participants again performed significantly better than the male participants. A repeated measures factorial ANOVA was performed in order to determine the effect of gender on the three groups of facts (three, five and seven facts). A quadratic main effect for the number of facts, $F(1, 8) = 10.15$, $p < .05$, $\eta^2 = .559$, and linear main effect for gender were found, $F(1, 8) = 15.75$, $p < .005$, $\eta^2 = .66$; however, no interaction was found to exist between the number of facts and gender, $F(1,8) = .003$, $p > .05$, $\eta^2 = .01$. The sample means for gender per number of facts are shown in Figure 6.19. The experiment raw data set is included in Appendix A.

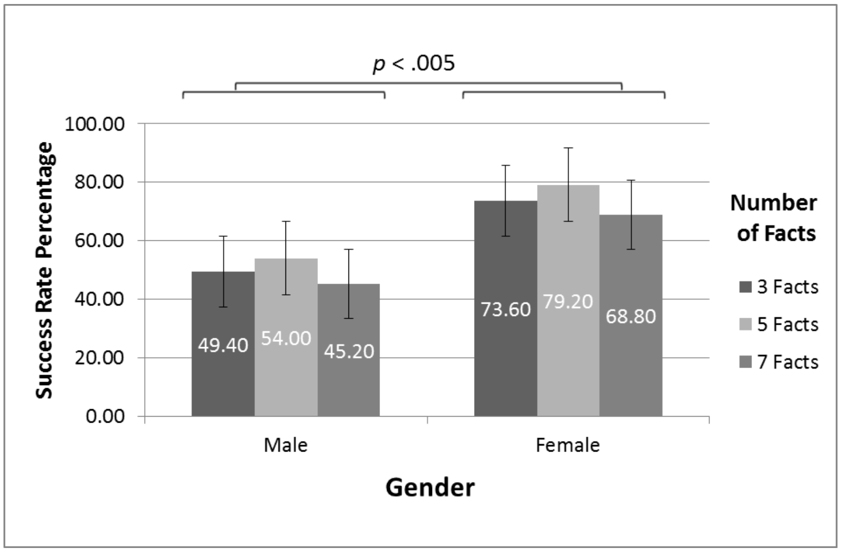

Figure 6.19. The mean recognition scores of the audeme meanings depicting the gender effect for audemes associated with three, five and seven facts.

6.5 Synopsis of the Findings

The following is a synopsis of the findings derived from the three experiments.

Learnability of the audemes: Audemes can be successfully learned and used to help students remember the information associated with them. In the beginning of the process of audemes getting learned by participants, variations in their attributes (source, semiotic and syntactic) are essential, but this effect fades once they are learned. In answering the first research question (Section 6.1.1), the audemes were shown to be highly effective in aiding the students' abilities to remember and recognize any textual content associated with the audemes.

Best audeme combination: The empirical data showed that the best audeme combination was the serial concatenation of music and sound effect with the first and second sounds being in different modes of listening (causal or referential). Moreover, the optimal number of facts that can be effectively associated to an audeme is five facts. This finding answers the second research question (Section 6.1.1).

Gender effect: In these experiments, the female participants outperformed the male participants in terms of recognizing the concepts associated with the audemes.

The effect of the audemes' syntactic attributes: Audemes created from serially combining two sounds yielded better recognition results when compared to the audemes created from the parallel combination of sounds.

A mixture of audeme source attributes yielded better audeme meaning recognition: Audemes created from different source attributes (music and sound effect) yielded better recognition results when compared to audemes created from same source attributes (music and music, or sound effect and sound effect).

A mixture of audeme semiotic attributes yielded better audeme meaning recognition: Audemes created from sounds used in different modes of listening (causal and referential) yielded better recognition results when compared to audemes created from sounds in the same modes of listening (casual and causal, or referential and referential). The reduced mode of listening in any combination yielded the lowest recognition for the audemes and associated concepts.

The causal and referential modes of listening yielded better audeme meaning recognition: Audemes created from sounds relying on the causal or referential modes of listening were more effective than those created from sounds relying on the reduced mode of listening.

Audeme content recognition is effective with a wide range of rhythms and timbres: Audemes created from sounds with different rhythms (fast, slow) and timbres (harsh, smooth) did not yield different results in terms of recognizing the concepts associated with those audemes.

6.6 Discussion of the Three Experiments

This study confirmed the following hypotheses that:

- audemes created of serial combinations of sounds will yield higher meaning recognitions when compared to parallel combinations of sounds (H1);

- audemes created from sounds in the causal and/or referential modes of listening yielded a higher meaning recognition when compared to audemes created from sounds in the reduced mode (H3); and
- frequent exposure to audemes has a high impact in their learnability (H4).

However, the hypothesis that the meaning of audemes will decrease with each additional testing due to the weakening of memory is not supported (H2). This section discusses the implications of the findings from these three experiments.

6.6.1 Can Audemes be Learned?

One of the main results of this study indicates the learnability of audemes. Experiment four shows an increase in concept recognition as time passes. This improvement is attributed to the testing conducted, which served as a reinforcement of the concepts. Since the tests were multiple choice answers in which the participants could hear the concepts along with the audemes, each test in itself reinforced the relationship between the audemes and concepts. This claim is supported by the findings in experiment five, in which we see a decreased level of concept recognition when not performing the test after one week. When comparing the results of the test conducted after four weeks from experiment four and five, it is seen that the students scored significantly higher when they took the first test after the first week as it served as additional exposure to the audemes and concepts.

In experiment six, the optimal number of facts that can be correctly associated with an audeme was investigated. The findings indicated that audemes associated with five facts could be the ideal load threshold for an audeme, compared to audemes associated to three and seven facts. This conclusion is based on the finding that audemes associated to three and five facts performed almost with the same effectiveness (the difference was not significant), while the effectiveness of the audemes dropped significantly when associated with seven facts.

These findings suggest that audemes require a single exposure (in which meaning is explicitly assigned) to achieve 82% accuracy in meaning recognition. With three exposures over a period of 8 weeks the meaning of audemes can be learned with 92% accuracy.

6.6.2 Which Combination of the Audeme Attributes Maximize Information Recognition?

The audemes were investigated as created from three attributes: syntactic (serial or parallel concatenation of sounds), semiotic (causal, reduced or referential sounds), and source attributes (music or sound effect sounds).

The experiment findings revealed that audemes constructed from serial combination of sounds were more effective than parallel combinations. This effect is attributed to the linear nature of the sound as it feels more natural to experience sound combinations one after another. Contradicting results about this effect are found in Brewster et al. (1995b) who, in their evaluation of earcons, discovered no difference in serial or parallel earcons in terms of their effectiveness. Similarly, Gerth (1992) conducted experiments that presented several sounds to participants at once. The results showed that the recognition score for three sounds playing in parallel was 90%.

In terms of modes of listening, the results were analyzed under the assumption that an audeme was perceived as being in one of the three modes: reduced, referential or causal. As previously described in the findings section, audemes used in the causal and referential modes helped the participants to more accurately remember the given concepts. In one case, the causal mode was slightly better than the referential mode, and this fact is intuitive because the causal audeme revealed more about the object of the concept and the connection was naturally established. In the same way, the referential mode took advantage of the contextual and inherent connection between the audeme and concept. The reduced mode did not help the participants recognize the concepts, which is an indication that a greater cognitive power is needed to relate the audeme to an object when they only share some common attributes (e.g., fast rhythm for a fast running animal, such as a rabbit). In relation to the modes of listening, it is an interesting finding that audemes composed of sounds with different modes helped the participants to recognize concepts better than when the audemes were composed of sounds with the same mode of listening.

In terms of the source attributes, experiment four revealed that combining the different types of sounds (Music-SFX; SFX-Music) was better than using the same type (Music-Music; SFX-SFX), which lends support for the idea that music and SFX offer complementary or even synergistic conceptual contributions. Along these lines, the findings show that audemes are most effective for correct concept recognition when a music is followed by a sound effect.

Among the findings in the third experiment are the effects of rhythm and timbre in relation to the level of recognition of the concepts associated with the audemes. The tests showed no significant impact in regard to the sound's rhythm and timbre for the effectiveness of the audemes. In essence, the rhythm and timbre analysis falls into the reduced listening mode, and providing that no effect was found for the reduced mode of listening, this finding makes the claim stronger.

6.6.3 Does the Relevance of the Audeme Attributes Fade with Time?

The guidelines devised for generating the best combination of audeme attributes are essential for the initial exposure to the audemes. The longitudinal study, however, shows that with the passing of time, the effectiveness of the audemes relies less on the sound attributes (source, semiotic and syntactic) that constitute them. While this conclusion is true for all of the attributes, an interesting fact was observed with the semiotic attributes.

Like the other attributes, the importance of the semiotic attributes became less with time, indicating that when we experience audemes in the referential mode, we rely more on the sound attributes of the audeme. This reliance fades away once the connection between the audeme and concepts gets stronger, that is, we start experiencing audemes in the semantic mode. This shift is an indication that the learnability of audemes could have a similarity with how we learn languages.

6.6.4 Could Audemes be Experienced as a Language?

Experiment four revealed the trend that, with repeated exposure, most of the audeme types become equally effective in regard to concept recognition. In terms of the modes of listening, the indication is that audemes are *passing* from the referential to the semantic mode.

In the first few weeks of audeme exposure, the participants are perhaps struggling with creating and remembering the references between the audeme and concepts. In this initial phase, multiple connections are built between the audeme and concepts, with some stronger than others. However, after few exposures through the tests, the participants inherently learned to discard the weak connections and promoted only the strong ones. This process of narrowing down the concepts and meanings of the audemes is similar to how language works, hence it represents a strong demonstration of the semantic mode of listening.

The process of the audemes being experienced in the semantic mode of listening suggests the possibility that audemes could be used to create a language. The audeme language could possibly be a sonic analogue to American Sign Language. Dorner and Hagen (1994) explained that American Sign Language uses fewer words than English, which is achieved through eliminating redundant words and phrases such as articles, interjections, dummy subjects and linking verbs. Their example is that the sentence "The car has a flat tire" in sign language could be expressed as "Car Flat Tire."

In order to address the problem that it would be impossible to find an audeme for each concept or word (a critique often addressed to auditory icons, too), we recall that sign language has already solved this issue. Klima and Bellugi (1979) explained that the signs for *summer*, *ugly* and *dry* are the same in terms of hand shape and movement, but differ only in the position of the body where the signs are made, in this case on the forehead, nose or chin, respectively. Similarly, the audemes signifying *travel*, *vacation* and *security* have the key jangle as the common sound, but the meaning differs only when played next to a sound of a *car*, *seagulls and sea waves*, or *typing on a computer keyboard.* The success of this method of representing information shows that similarities can be drawn between sign language and how audemes are designed and used.

6.6.5 Is There a Gender Effect in Interpreting Audemes?

The effect of gender in all of the experiments is interesting, and reinforces the results found in experiment three (Section 5.5.3). I tend to attribute this result to the females' innate abilities to process and use emotions better than males (Barrett, Lane, Sechrest, & Schwarts, 2000; Seidlitz & Diener, 1998). It is a well-established fact that emotion is a crucial part of memorization (Bower, 1981; Laird, Wagener, Halal, & Szegda, 1982), and I believe that the use of audemes invokes powerful emotional experiences. Similarly, female superiority over males has been found in second language acquisition (Ehrman & Oxford, 1988; Oxford, 1993).

6.6.6 How Does the Learnability of Audemes Compare to Other Non-speech Sounds?

The audemes in this study have been investigated in a framed context, which is the educational content related to U.S. history and geography. In this context, the audemes were correctly recognized 92% of the time over the course of four exposures (one initial exposure and three tests conducted with two weeks in between each test). The lack of context could have caused the audemes to be ambiguous and confusing, as was the case of auditory icons in two different studies investigating those with and without context. In the study reported by Mynatt (1994), auditory icons were identified only 15% of the time when the context was absent. In another study (Fernstrom & Brazil, 2004), it was reported that the participants were successful in identifying auditory icons with objects or actions for a specific context 70% of the time.

In terms of learnability, existing non-speech sounds are effective in recalling objects or actions for various cases. Earcons showed an 80% accuracy in recall to hierarchical navigation menus (Brewster et al., 1996). Auditory icons showed high accuracy in recall of objects and actions in a user interface (Lucas, 1994) and in identifying links and providing information in web navigation (Morley, Petrie, & McNally, 1998). A study (Absar & Guastavino, 2008) that presented a systematic comparison between earcons and auditory icons claimed that earcons were highly structured, but users required specified training in order to understand earcons. Auditory icons, on the other hand, are easy to learn, but need to be used as individual sounds because their combination is not possible.

A recent study (Dingler, Lindsay, & Walker, 2008) compared the learnability of earcons, auditory icons, spearcons and speech and showed that spearcons were as learnable as speech, followed by auditory icons. The results also showed that earcons were much more difficult to learn. A study conducted by McGee-Lennon et al. (2011) claims that musicons were as learnable as auditory icons. Spindex as spearcons required little or no training to be learned by users (Jeon & Walker, 2009).

6.7 Guidelines for Effective Audeme Creation

Based on the findings of the above described experiments, a set of initial guidelines for designing audemes that can maximize the recognition of their meanings is proposed.

Serially concatenate sounds. Audemes with sounds serially concatenated (juxtaposed in sequence) are more effective in eliciting meaning recognition than audemes made of parallel sounds (different sounds played together or overlapping).

Mix different sounds. Use different types of sounds in the audeme combination for the most effective recognition of the associated meanings: Music-SFX and SFX-Music.

Follow a music with a sound effect. Audemes that consist of two parallel music sounds are the least effective, while those having music first and then a sound effect in a serial fashion are the most effective.

Use the sound of the real object causing it. Use audemes that consist of sounds in causal and referential modes, rather than reduced mode of listening for the best results. For instance, use audemes created from sounds that refer to the object that is causing the sound, such as, the sound of a chicken to represent the chicken itself or the farm in general. On the other hand, avoid using audemes created by sounds that only refer to the low level qualities of the sound, such as rhythm, pitch and timbre. An example would be using a sound with fast rhythm to indicate a fast car, which is not an effective audeme in regard to associating meaning with the sound.

Frequent exposure has a large impact on the recognition of the audeme meaning. For effective recognition of the audeme meaning, expose the audience to the audemes at least twice a week. Longer periods between two exposures could drastically affect the audemes' learnability and recognition.

Recognition of the audeme meaning works well with a broad range of rhythms. For the best results, use audemes of any pace or rhythm (fast, slow). Using audemes created from fast or slow rhythms has the same effect in the recognition of the meaning associated with the audeme.

Recognition of the audeme meaning works well with a broad range of timbres. For the best results, use audemes of any timbre (harsh, smooth). Using audemes created from harsh sounds (metal scraping, electric guitar, etc.) or smooth sounds (light wind, violin, etc.) have the same effect in the recognition of the meaning associated with the audemes.

Defining guidelines will help acoustic interface designers make decisions in regard to where and for what to use audemes in their applications. These guidelines will enable the designers to rely on a framework, rather than an ad hoc manner of selecting sounds for acoustic interfaces. A summary of the guidelines described above and a brief description of each one is given in Table 6.5.

Table 6.5. Summary of the guidelines and their descriptions.

Guideline Name	**Guideline Description**
G1. Serially concatenate sounds	Sounds used to create an audeme should be concatenated in serial (played one after the other), rather than in parallel (played at once) fashion.
G2. Mix different sounds	Use different types of sounds (music-sound effect; sound effect-music), rather than sounds of the same type (music-music; sound effect-sound effect).
G3. Follow a music with a sound effect	An audeme should start with music and followed by a sound effect sound.
G4. Use the sound of the real object causing it	Sounds that are experienced in the causal and referential modes of listening should have higher priority to be included into the audeme compared to sounds experienced in the reduced mode of listening.
G5. Frequent exposure has great impact on the recognition of the audeme meaning	Expose the audience to audemes twice a week to maintain its correct meaning recognition.
G6. Recognition of the audeme meaning works well with a broad range of rhythms	Sounds of any rhythm (fast, slow) can be used in generating an audeme.
G7. Recognition of the audeme meaning works well with a broad range of timbres	Sounds of any timbre (harsh, smooth) can be used in generating an audeme.

7. ADVANCED SUPPORT AND CREATION-ORIENTED LIBRARY TOOL FOR AUDEMES (ASCOLTA)

7.1 Overview

This chapter focuses on the design, implementation and initial evaluation of the Advanced Support and Creation-Oriented Library Tool for Audemes (ASCOLTA) (Italian: *to listen*) prototype. The ASCOLTA is an interactive application that enables individuals without an audio design background to create effective audemes. The aim of designing and implementing the ASCOLTA is three-fold.

First, to propose a design that offers an interface that guides users in the process of generating well-formed non-speech sounds (audemes). The created audemes then can be used in different scenarios, including being used as supplementary teaching materials. For example, teachers can utilize audemes to teach their lectures by playing them multiple times while they explain the subject material.

Second, to implement a design that operationalizes the guidelines derived empirically in the experiments reported in chapter 6, which will ensure the creation of well-formed audemes. The ASCOLTA will generate audemes by automatically concatenating sounds based on sound attributes, such as the type of sound (music or sound effect) and listening mode (causal, referential or reduced).

Third, to facilitate the process of generating well-formed audemes for a non-technical audience. While sound design has become a relatively easy task in the recent years, it still requires basic skills in regard to using commercially available sound design tools, such as Soundtrack Pro or Sound Forge. ASCOLTA facilitates this process by automatically generating well-formed audemes and offering to the users a simple interface to modify, edit and save the created audemes.

7.2 Conceptual Design

7.2.1 User Requirements and Target Users

The ASCOLTA development was employed within a user-centered design process. The initial step of the design process was to brainstorm an intuitive interface to be used to address the needs of the potential users, who are teachers within a K-12 school. I observed these teachers in their typical settings and noted the materials that they used while lecturing on a given subject. The teachers typically had at least one book per subject being taught. At the end of each chapter, the book offered a set of keywords to be used to describe the content of the lecture. The teachers could use these keywords to generate audemes by entering them into the ASCOLTA interface, and then use the audemes generated by the ASCOLTA in class.

7.2.2 The ASCOLTA Content

The ASCOLTA application contains brief sounds and descriptions about them. Currently, 41 sounds exist in ASCOLTA, of which 12 are music and 29 are sound effects. In terms of listening modes, 36 sounds are in the causal and/or referential modes, while five are in the reduced listening mode.

Sounds: The sounds are very brief, typically 1-3 seconds in a *wav* file format. The length of each individual sound was kept as short as possible in order to keep the audemes within a reasonable length. The goal was to create audemes that did not exceed 10 seconds in order to preserve their thematic cohesion.

The selection of sounds was based on the keywords taken from a U.S. history book (Clayton, Perry, Reed, & Winkler, 2000) used as part of the standard curriculum of K-12 schools. Initially, three chapters from the book were selected: Chapter 8 - The Market Revolution, Chapter 9 - Religion and Reform and Chapter 12 - The Civil War. These particular chapters were chosen based on the availability of sounds in the digital library. Subsequently, the specific sounds for each chapter were chosen based on the keywords provided at the end of each chapter.

Sound Description: Each sound is defined by its type (music or sound effect), the modes of listening (causal, referential or reduced) and meta names. The type and modes of listening attributes are required in order to enforce the creation guidelines when audemes are created using the ASCOLTA. Meta names are used to describe each sound and mainly cover, but are not limited to, the keywords provided at the end of the book chapter. For example, a music excerpt from the song *Dixie* in the referential mode of listening consisted of the following meta names: *civil war*, *dixie*, *confederacy* and *south*. Typically, three to five meta names were included for each sound. The meta names were assigned by three researchers in order to ensure the correct mapping between the meta names and sounds. An excerpt of the sounds with their attributes and meta names is given in Table 7.1. The full list of 41 sounds is provided in Appendix B.

Table 7.1. An excerpt of sounds and their description used in the ASCOLTA.

#	Sound Name	Meta Names	Sound Type	Listening Mode
1	"Dixie" song	dixie, civil war, confederacy, south	Music	Referential
2	"Battle Hymn of the Republic" song	battle hymn of the republic, civil war	Music	Referential
3	Shooting guns	gun, fight, war, military	Sound effect	Causal and Referential
4	Signing on a piece of paper	emancipation, proclamation, decree, presidential, amendment	Sound effect	Causal and Referential
5	Swords	swords, fight, war, military	Sound effect	Causal and Referential

7.2.3 Interface Design

The design of the ASCOLTA user interface has gone through multiple iterations with screen mockups. The initial mockups were generated using MS PowerPoint to promote quick feedback, which was gathered using peer review and design critique discussions. Figure 7.1 shows a sample of the screen mockups that lists the audemes retrieved using matching keywords entered by the user.

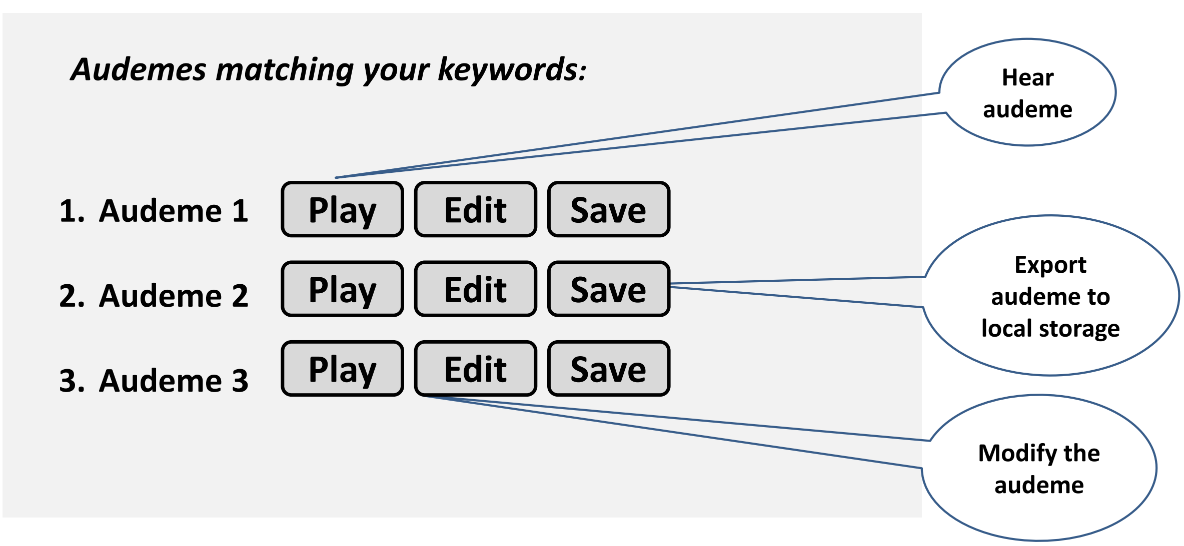

Figure 7.1. Screen mockup of the ASCOLTA audemes.

Figure 7.2 shows the sample screen mockup that lists the sounds constituting a given audeme.

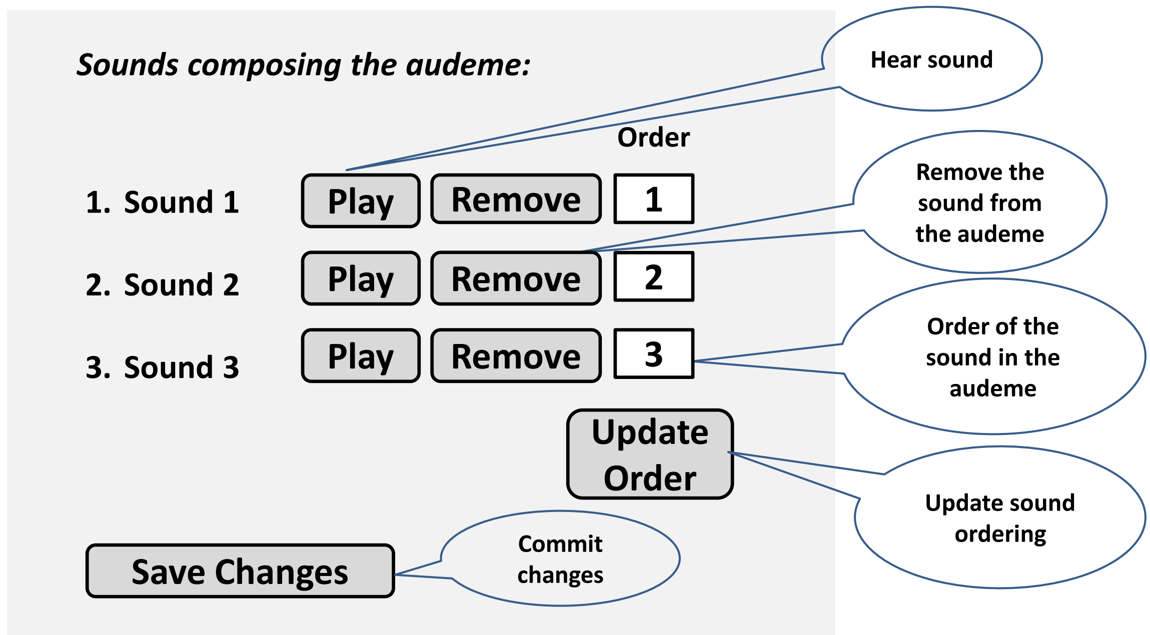

Figure 7.2. Screen mockup of the ASCOLTA sounds.

The goal of the iterative design was to develop an intuitive and usable interface for the type of users identified in the User Requirements Section (7.2.1). Additional ASCOLTA mockups are included in Appendix C.

7.2.4 Core Functionality

In this section, the functionalities of ASCOLTA will be outlined in detail.

7.2.4.1 Applying the Guideline

At the core of ASCOLTA lies the ability to generate audemes based on the guidelines described in Section 6.7. There are four guidelines operationalized in the ASCOLTA:

Guideline 1: Serially Concatenate Sounds: Sounds that are retrieved based on a match with the keywords given by the user are concatenated in a serial fashion. This guideline is based on the results of experiment four that showed that the meanings of audemes created from serially concatenated sounds were more often recognized than the meanings of audemes created from parallel sounds (Figure 7.3).

Figure 7.3. Serial concatenation of sounds.

Guideline 2: Mix Sound Types: The sounds retrieved are combined in such fashion that sounds of the same type are not concatenated closely together, but mixed with another type. For example, if three sounds are retrieved (two sound effects and one music), then the music will be placed in-between the two sound effects (Figure 7.4). This guideline is based on the results of experiment four that showed that the meanings of audemes created from different sound types (music and sound effect) were more often recognized than the meanings of audemes created from the same type of sound (music and music, or sound effect and sound effect).

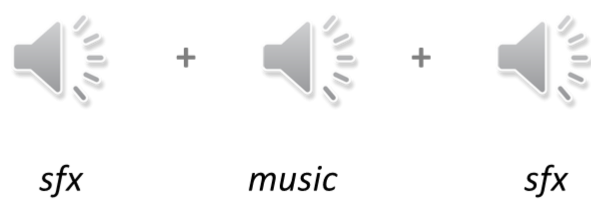

Figure 7.4. Concatenating different mixes of sounds.

Guideline 3: Music First: In the event that music and sound effects are retrieved, the music should be followed by the sound effect. For example, music + sound effect (Figure 7.5). This guideline is based on the finding in experiment four, in which the meanings of audemes created from music followed by a sound effect were more often recognized than the meanings of audemes created from combinations where the sound effect was put before the music.

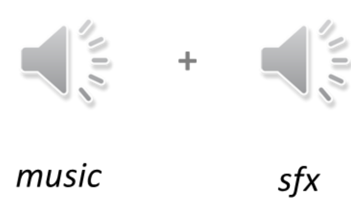

Figure 7.5. Music concatenated before the sound effect.

Guideline 4: Causal and Referential Sounds Have Priority: When multiple sounds are retrieved, the sounds that are in the causal and referential modes of listening take precedence over the sounds in the reduced mode of listening when creating the audeme (Figure 7.6). This guideline is based on the finding of experiment four, which showed that the meanings of audemes that consist of sounds in the causal and/or referential modes of listening were more often recognized than the meanings of audemes that consist of sounds in the reduced mode of listening.

Figure 7.6. Causal and referential sounds have priority when concatenated.

Looking closely at the four guidelines, a circumstance can be identified when guidelines 3 (music first) and 4 (causal and referential sounds have priority) will be tied. For instance, let us consider a situation in which two sounds are retrieved. In terms of their type, sound 1 is music, while sound 2 is a sound effect. In terms of their modes of listening, sound 1 is reduced, while sound 2 is causal. If applying guideline 3, which states that music should be followed by a sound effect, the sounds should be concatenated in the following fashion, sound 1 + sound 2. On the other hand, if applying guideline 4, which states that sounds in the causal and referential modes of listening have priority compared to sounds in the reduced mode of listening, the sounds should be concatenated in the following fashion, sound 2 + sound 1. In order to resolve such a conflict, a closer look at the findings for these two guidelines should be undertaken. These results can be found in Figures 6.6, 6.10 and 6.11. After reviewing the results, it can be seen that guideline 4 was more influential, therefore, should be followed when in conflict with guideline 3.

It is important to note here that the audemes used in the experiments for this study (Chapter 6) through which the guidelines were derived consisted of only two sounds. In the ASCOLTA, however, audemes could consist of up to five sounds. Thus, applying guidelines to five-sound audemes that were derived from two-sound audemes might produce slightly different results in terms of the generated audemes being effective.

7.2.4.2 Letting Users Explore and Create Audemes

While the audemes provided by the ASCOLTA are well-formed based on the guidelines, the users should be given a way to explore their creativity and make decisions in the process of generating an audeme. Therefore, the users will be able to play and hear the created audeme, as well as edit it.

The ASCOLTA will enable users to conduct two types of edits on the audemes. First, the users can remove one or more of the sounds making up the audeme. For instance, if the given audeme is created from three sounds, the user can choose to remove one or more of the sounds and leave the audeme with only the remaining sounds. Second, the users will be given the option to change the order of the sounds in an audeme. For instance, if an audeme consists of three sounds (1, 2 and 3), the user can change the order as desired. For example, 2, 3 and 1. It is possible that when users commit changes to an audeme, they violate one of the guidelines; thus, ASCOLTA will keep track of such actions and warn users before they apply the changes.

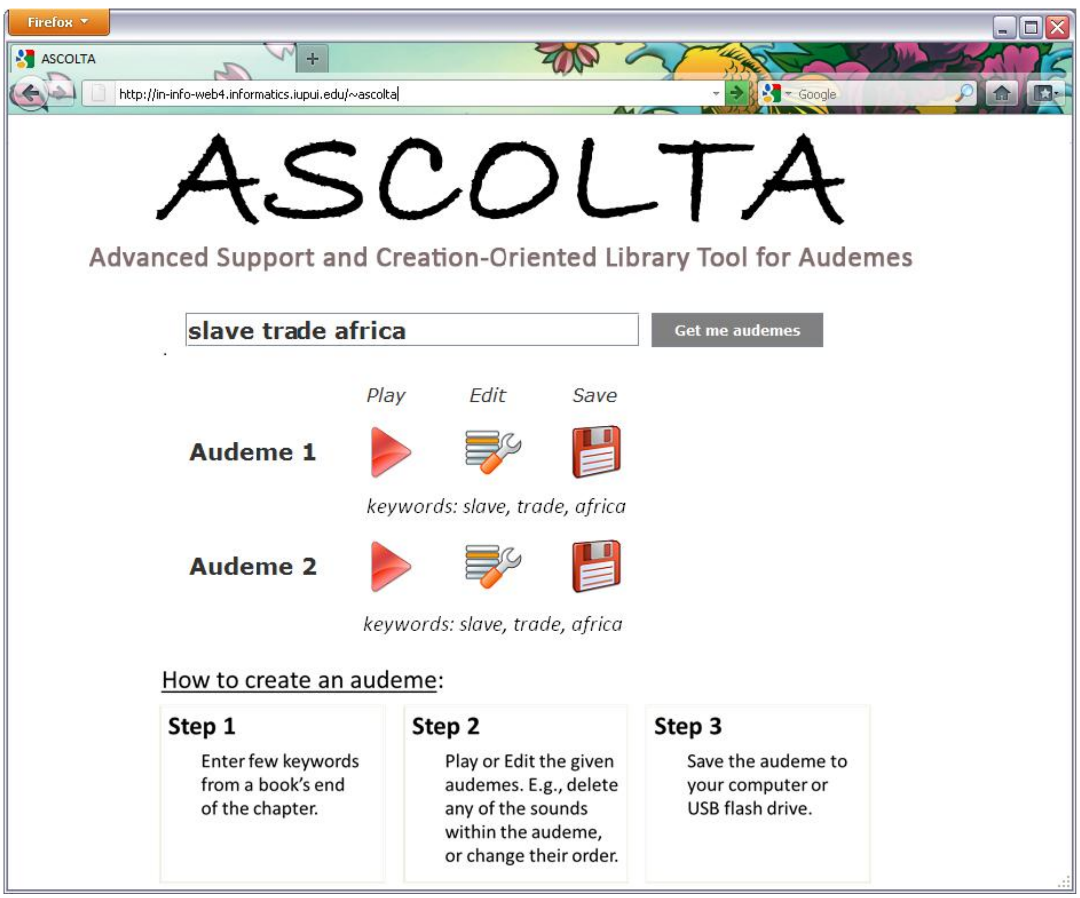

Figure 7.7. ASCOLTA retrieving two audemes based on three keywords.

7.2.4.3 Exporting an Audeme

Once a user has built his desired audeme, it is crucial that the ASCOLTA provide a way to let him export it to local storage. This feature will enable users to play the audeme independently of the ASCOLTA interface. The audeme will be saved as a generic wav sound file, which can be played on any computer machine.

7.2.5 Usage Scenario

A usage scenario is given here to illustrate how the ASCOLTA works. Bill is a teacher at the K-12 School for the Blind and Visually Impaired. He teaches history and has an upcoming lecture on slavery. Bill already knows that playing audemes to his students when he lectures helps them to remember the topic with higher accuracy. Now, he needs to create an audeme on slavery, but has no knowledge of sound design. Hence, he turns to the ASCOLTA application to create the desired audeme.

Bill opens the ASCOLTA. Once it loads, the main page requires him to enter a few keywords about the subject of the audeme. Bill opens his history book and flips to the last page of the slavery lecture. He selects the keywords *slave*, *trade* and *Africa* to input into the ASCOLTA interface. Bill then hits the “Get me audemes” button and the interface returns two audemes matching the given keywords (Figure 7.7). He prefers one of the audemes, which was created from three sounds: a short passage of a choir singing *“Swing Low, Sweet Chariot”* (music), the sound of *a whip crack* (sound effect) and the sound of *a person screaming* (sound effect). While he likes the first two sounds, the third sound he judges to be too harsh for the age of the students. Thus, he hits the “Edit” button, which shows the details of the sounds constituting the audeme (Figure 7.8a). He clicks on the “Remove” button for the third sound, which is then removed (Figures 7.8b and 7.8c). He confirms the change has been done by clicking on the “Preview Changes” button, which will play the audeme without the third sound. While he is still on this page, he decides to swap the order of the remaining two sounds (Figure 7.8d). At this point, the ASCOLTA interface shows him a warning stating that he is violating a guideline and might be compromising the effectiveness of the audeme (Figure 7.8e). Specifically, the ASCOLTA tells Bill that is he trying to violate guideline 3 by moving the sound effect before the music. Bill then decides not to proceed with the change and clicks on the “No” button. He saves the changes done to the audeme by clicking on the "Save Changes" button. Finally, he saves the audeme to his removable disk in order to be able to take it with him and play it on the classroom computer (Figure 7.8f).

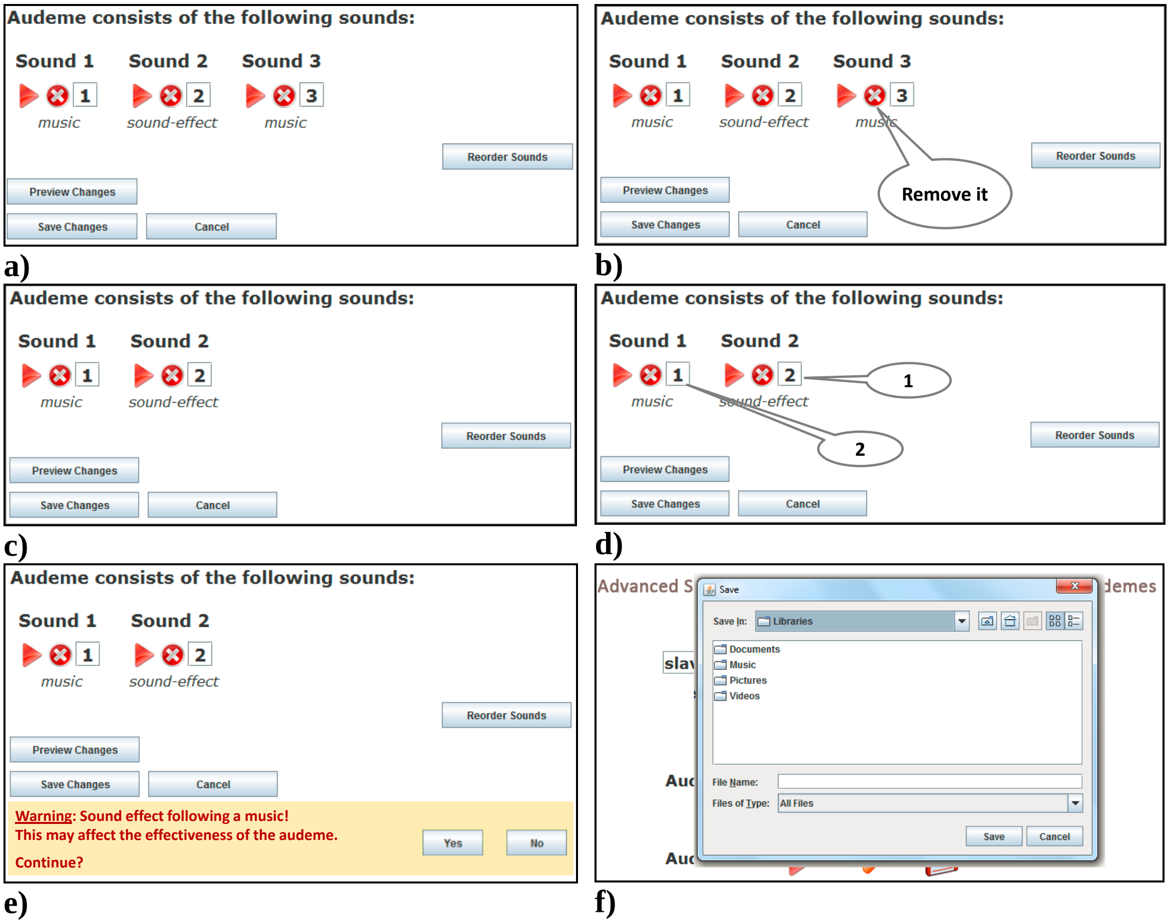

Figure 7.8. a) the ASCOLTA interface showing the three sounds that constitute the audeme, b) clicking on the red X button removes that sound, c) the third sound is removed, d) reordering the sounds can be accomplished by entering a number into the text field, e) the warning displayed when a guideline is about to be violated and f) the audeme is saved to local storage.

7.3 System Architecture

The implementation of the ASCOLTA was done on the Java Applet platform for two main reasons. First, java libraries offer an easy concatenation of sounds available in the wav format. Second, the applet can be easily integrated into an HTML page to be available as a web page.

The back end of the ASCOLTA is run by the MySQL database management system, which organizes the sounds and retrieves them based on a given criteria (i.e., keywords provided by the user). The database consists of a single table organizing the sounds. The table is named "Sounds" and consists of the following attributes (fields):

SoundID – A unique identification number for each sound.

SoundName – A name for each sound.

SoundType – A statement of the sound's type (i.e., music or sound effect).

SoundMode – A statement of the sound's mode of listening (causal, referential or reduced).

SoundURL – The path where the sound is located on the server.

SoundMeta-names – Keywords describing the sound.

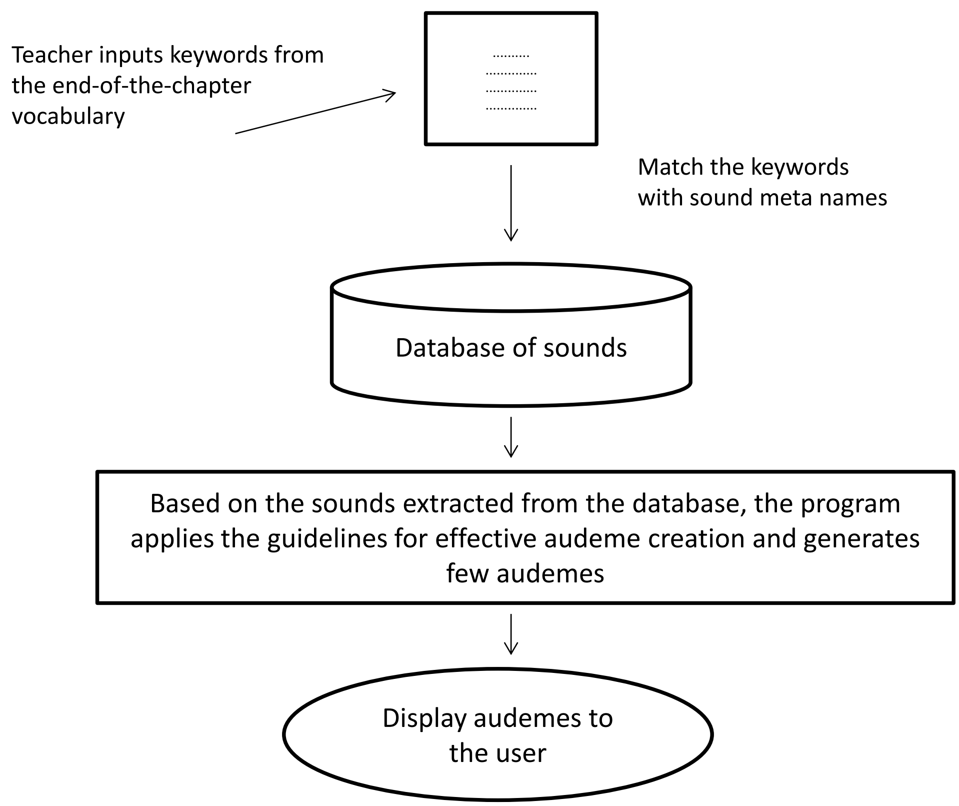

Figure 7.9. The ASCOLTA diagram depicting the process of audeme creation from keywords.

The process of retrieving audemes is illustrated in Figure 7.9. When a user enters keywords into the ASCOLTA interface, the application checks for any matches found between these keywords and the sound meta names in the database. If multiple matches are found, then sounds that contain more than one keyword match get higher priority. For example, if there are three given keywords and two keywords match for sound 1 and only one matches for sound 2, then sound 1 will have a higher priority in regard to being a potential sound for the audeme. Up to a total of five sounds may be retrieved for a single audeme, thus, an audeme cannot exceed the length of more than five sounds. Once the sounds are retrieved, the algorithm for applying the guidelines is used to combine the sounds to create a well-formed audeme. Finally, the application will generate an audeme as a single wav sound.

The sounds used to populate the database were from the Indiana University School of Informatics' digital audio library.

7.4 Experiment 7: Usability Evaluation of the ASCOLTA

7.4.1 Overview, Hypotheses and Purpose

The purpose of this experiment was to understand the value of the ASCOLTA and gain additional feedback on the utility and importance of the tool from potential users. The ASCOLTA prototype was shown to participants and their comments concerning the value of the tool, including the usefulness and usability of the tool, were collected. This experiment has been approved by IRB #IRB-1203008346.

7.4.2 Method

7.4.2.1 Participants

Three participants (2 male, 1 female) were recruited for the experiment and paid for their participation. The participants were high school teachers from the ISBVI. An email was sent to a few teachers that the School of Informatics had prior collaboration with in order to solicit for their participation in this study.

7.4.2.2 Design

This experiment was designed as an informal interview in which the ASCOLTA prototype was demonstrated to the participants and their comments on the relevance of the tool and ways to improve it were recorded. The only data collected was the participants' comments about the ASCOLTA.

7.4.2.3 Procedure

Individual meetings with the participants took place at the ISBVI. First, the participants were briefed about the role of ASCOLTA and then saw a demonstration covering the important functionalities of the tool. After the demonstration, the participants were encouraged to explore the ASCOLTA using four simple tasks (Table 7.2). Finally, the participants were asked 10 questions concerning the usefulness of the ASCOLTA (Table 7.3). The entire experiment lasted one hour and was audio recorded for later analysis.

Table 7.2. The list of tasks used to evaluate the ASCOLTA.

#	Task Description
1	Enter a few keywords for the given content that you want the audeme to focus on.
2	Play the audeme(s) that is returned as a result.
3	Use the application to change one of the audemes. For example, remove one of the sounds.
4	Save one of the audemes onto the computer.

Table 7.3. The list of interview questions used to evaluate the ASCOLTA.

#	Interview Question
1	Based on your understanding on what audemes are, do you think that audemes can be an integral part of your teaching process? If so, how?
2	Do you think that the ASCOLTA could be a tool that you might use in your teaching process to find and create audemes? If so, how would you use it?
3	If the ASCOLTA was available right now, would you use it? Why or why not? If yes, how?
4	Imagine a situation in which the ASCOLTA would be useful to you?
5	Can you summarize three good and three bad things about the ASCOLTA?
6	Based on your recent experience, how can the ASCOLTA be improved to be *more* useful?
7	Does finding audemes based on keywords make sense to you?
8	Is the way that the audemes are displaying meaningful to you?
9	Is the way that the sounds that constitute an audeme are displayed meaningful to you?
10	Do you have any other comments or observations about ASCOLTA that you would like to share with us?

7.4.3 Results and Discussion for Experiment 7

7.4.3.1 Simple User Interface, but with Limited Accessibility

The results of experiment 7 showed that ASCOLTA is a desired tool to create well-formed audemes. The general conclusion is that the ASCOLTA is an easy-to-use interface and provides the user with enough freedom to leverage their creativity when creating audemes. One participant pointed out the similarity between the ASCOLTA interface and Google. Typically, the interface familiarity communicates the modality of interaction with it. In this sense, users immediately understand what is expected of them in regard to submitting keywords in order to creating audemes.

Creating audemes based on few keywords was preferred by all of the participants. Two of the participants also commented that only displaying two resulting audemes makes the process of choosing an effective audeme easy. P2: “It’s simple; it’s good that you get just few audemes as a result, not too many.”

One participant suggested ranking the audemes displayed in terms of their match with the keywords and guidelines. P2: “It’s not good that all of the audemes look identical. It would be helpful to rank the audemes by providing the percentage on how accurately the audeme represents the keywords and guidelines.” The ASCOLTA could be improved by displaying a percentage of effectiveness next to each audeme. For example, a 100% value could be displayed next to the audeme that has a positive match on all of the keywords and the guidelines. This improvement would provide the users with visual feedback about the relevance of the listed audemes, and enable them to make preliminary decisions without playing the actual audeme.

Although the interface is intuitive, providing three easy steps describing its usage on the main screen was judged to be beneficial by all of the participants. The description was short and to the point, giving an overview of the features provided by the ASCOLTA. P3: “I quickly understood that given audemes can be tweaked as desired based on reading the three steps description.” One participant added that it might also be useful if a 30 seconds video was included on the ASCOLTA main screen showing the process of creating an audeme.

All of the participants commented that the interface was visually pleasing and easy-to-use. They suggested, however, that the interface be upgraded so that it is accessible to blind and visually impaired users. Initially, the tool was created to be used only by sighted teachers and, therefore, accessibility was not a requirement. The participants stated that there are several blind and visually impaired teachers at their school who would be interested in using the tool. More importantly, they indicated that the availability of the tool should not be limited to teachers only, but also to blind and visually impaired students who could use it as an in-class exercise.

Making the ASCOLTA accessible would expand the groups of users that would be able to use it, such as the blind and visually impaired. A big step toward achieving this goal would be to make the interface screen reader friendly. Additionally, providing ways to control all of the interface buttons and fields using the keys on the keyboard would be important in regard to achieving accessibility as blind and visually impaired individuals do not use the mouse to interact with an interface.

7.4.3.2 Freedom to Explore Users' Creativity, but Limited Collaboration

The participants appreciated the fact that the audemes were generated based on empirically derived experiments. However, the possibility to modify the audemes to the level of deleting or reordering the individual sounds was preferred by the users. P3: "I like the choice of being able to change the individual sounds." This feature gave the users the freedom to explore their imagination by trying different sound combinations, even at the cost of breaching one of the audeme design guidelines. For instance, putting two sound effects next to each other violates the guideline that sounds of the same type should not go together, but the users wanted to create different audemes by trial and error.

However, they suggested several improvements to be made to the audeme edit feature. First, the participants requested that the sounds constituting the audeme be given a more identifiable names, rather than generic names. P3: "Maybe it will make sense to name the sounds, such as a *whip crack*, instead of Sound 1." This improvement would provide the users with quick visual feedback and enable them to make decisions without being forced to listen to the sounds. Second, all of the participants asked for the possibility to manually add new sounds during the editing process. They suggested that sounds could be added by browsing the ASCOLTA database, or by uploading a sound from the users' local storage. P1: "I would like to be able to add more sounds. Adding a browse button would let me pull sounds from my computer." This feature stresses the importance of individuals' wishing to use their sound resources whenever they generate audemes. Perhaps, people collect sounds into their personal libraries, which they would want to utilize when creating audemes. Finally, one participant suggested applying a drag metaphor to reordering the sounds, instead of inserting numbers into the sound order fields, which is less intuitive.

A feature that all of the participants judged as important to include in the ASCOLTA was the ability to store and share the created audemes. P1: "Could you make it possible for the users to share the created audemes with their friends?" Through this collaborative feature, the participants claimed that they would be able to evaluate their audemes with feedback from other users and, in that way, learn to create better audemes. One participant added that if the ASCOLTA were to be used by the students, they would be very interested in this feature, which would make them compete with their peers as to who could create the best audeme.

Collaboration seems to be important for two reasons: evaluation and competition. In their interviews, the users suggested that it is always important to evaluate created audemes with other teachers or students and get their feedback. This suggestion is an indication that users need the audemes assessed by others regardless of the guidelines enforced within the ASCOLTA. Competition between students is another reason that makes collaboration an important feature. In addition to the entertainment factor, competition with peers would improve the students' skills in regard to creating effective audemes.

7.4.3.3 Users Anticipate using the ASCOLTA in Teaching their Classes

Although the participants noticed that the ASCOLTA has limitations, they would greatly appreciate having the tool at their disposal. P1: "It would be good to have this right now, so that the students can use it." Up until the ASCOLTA was shown to the participants, they had no free access to audemes and always relied on researchers from IUPUI to generate audemes for them based on their requests.

The three participants described scenarios in which they would use the ASCOLTA. The first participant wished to use ASCOLTA for a class in which blind and visually impaired students would use the tool to build audemes to represent the concepts and topics covered in the class. P1: “We can use the ASCOLTA to offer a nine-week rotation class in which students will create audemes. Students then would compete to see who could create the best audeme, which would be judged by their peers.” This participant noted that whenever the students get involved in generating audemes, they remember the topic and concepts lectured on in the class better.

The second participant, who teaches a travel skills class, claimed he would use the ASCOLTA to generate audemes to simulate different outdoor environments on the days when class cannot be held outside due to a bad weather. P2: “Audemes from the environment are very useful, so with the ASCOLTA I can create audemes to explain to my students what a residential environment is.”

The third participant would like to use the ASCOLTA to create audemes to teach the lessons in her class. In the participant’s year-long experience of using audemes while teaching a class, she has used audemes to test and review book chapters. The participant was using ready-made audemes, but now with the ASCOLTA she can create audemes in the class as she develops the lesson with students. P3: “I could use this tool to create audemes while I teach the lesson. The audeme will be created with suggestions from the kids. Essentially, this makes the students a part of the experience.” A complete semantic transcription of participants’ survey answers is included in Appendix A.

8. GENERAL DISCUSSION AND CONCLUSION

The primary aim of this dissertation was to establish a new category of non-speech sounds – audemes – to be used to convey meaning for theme-based contents (Figure 8.1: Aim 1). The need exists for this kind of non-speech sound in regard to supporting the education of the blind and visually impaired community. The learning process in an educational setting typically requires students to retain pieces of information from an extended theme-based content. Existing non-speech sounds were only suitable to convey meaning about objects or events in a Graphical User Interface (GUI), which limitation made them unfit to be used for the retention of educational content. This gap was filled by the invention of audemes.

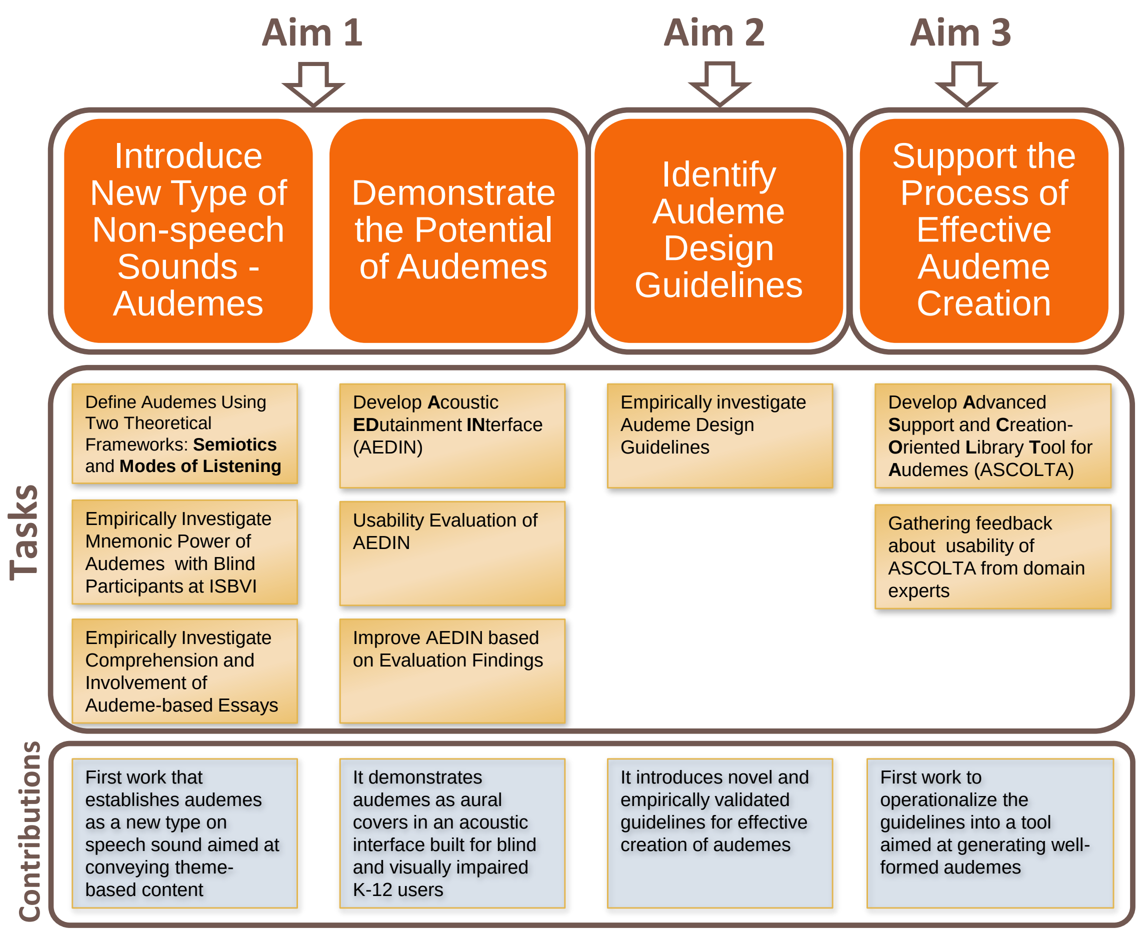

Figure 8.1. Revisiting the aims that guided this dissertation.

In order to establish the audemes, this dissertation investigated the guidelines that need to be applied to generate well-formed audemes (Figure 8.1: Aim 2). Previously, guidelines only existed for the generation of earcons, which offered guidelines to be followed in the process of combining musical notes. These guidelines, however, were unsuitable for the creation of audemes because the meaning of audemes is not created from the combination of discrete musical tones. Audemes, being similar to auditory icons, are created by combining different sounds to generate a meaning. In order to ensure that the audeme will convey the meaning required to represent a concept or extended theme-based content, a set of guidelines are empirically identified in Aim 2. Once the audeme design guidelines were established, a need for an interface aimed at their easy creation was the next viable step. To this end, an interface was built (ASCOLTA) that would enable individuals interested in audemes, and in particular teachers involved in schools for blind and visually impaired students, to find and create well-formed audemes to be used in their teaching processes (Figure 8.1: Aim 3).

Audemes could effectively serve two purposes. This dissertation finds that audemes as a novel type of non-speech sound convey meaning about theme-based content (e.g., a 500 word essay on slavery). In addition, this dissertation demonstrates that audemes can also serve as aural covers to anticipate TTS essays presented in an aural user interface. Essentially, this work has identified two important *functions* of audemes: audemes as memory catalysts and audemes as content anticipators.

As *memory catalysts*, audemes helped participants remember (experiment 1) and recognize (experiment 4) accurately theme-based content. Therefore, audemes can be used in combination with content, such as essays, to facilitate the learning of education material through the process of increasing the retention of theme-based content.

As *content anticipators*, audemes served as aural covers when integrated into an acoustic content-rich interface. Therefore, audemes can be used to anticipate the associated content in order to help users decide which TTS essay to listen to.

These two functions of audemes are similar in the sense that they require an established context in order for audemes to help retain content or anticipate it. For example, the U.S. constitution audeme could be associated to the U.S. constitution essay within the context of U.S. history. Hence, there are three entities involved for audemes to be effective for content recognition or anticipation: the *audeme* itself, the *content* associated with that audeme and the *context* within which the evoked audeme meaning should be directed.

The context could be set or initiated by the content (when audemes are used as memory catalysts) or the audeme (when audemes are used as content anticipators). When audemes are used as memory catalysts, the context is set by the content that the audeme is linked to (Figure 8.2). Initially, individuals hear an audeme, which could evoke various plausible meanings, but only when they hear the content following the audeme, the number of meanings narrows down and a more direct relation between the audeme and content is created. For instance, playing the U.S. constitution *audeme* (consisted of the sound of a gavel, the sound of a quill pen writing and the Star Spangled Banner) just before a lecture about the U.S. constitution (*context*), could initially evoke in blind and visually impaired students various meanings, including a court, writing, and the U.S. hymn. Multiple meanings are evoked because of the loose (or missing) context by just playing the audeme alone. When the students will listen to the U.S. constitution *content* (e.g., the teacher explaining or reading an essay or a book passage), the context will begin to be generated and the students could more accurately associate the meaning *U.S. constitution* to the audeme heard at the beginning of the class. This process of interaction between the content and the audeme reinforces the memorization of the content and makes audemes act as memory catalysts.

On the other hand, when audemes are used as content anticipators, the framing of the *context* goes in the opposite direction; audemes generate the context for the content that follows (Figure 8.2). Essentially, audemes give the users a cue about the content they should expect. For instance, playing the slavery audeme (consisted of a short passage of a choir singing *Swing Low, Sweet Chariot* followed by the sound of a whip crack) can set the context for upcoming essay on slavery.

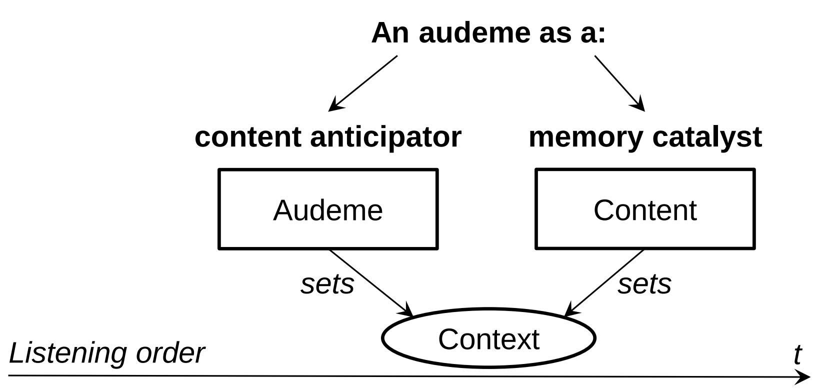

Figure 8.2. The process of context generation depending on the functions audemes serve.

It is also important to define the scope in which audemes could be useful. Unlike other non-speech sounds (earcons, auditory icons and spearcons), which are used typically in a GUI, audemes can be used beyond this realm. Earcons, auditory icons and spearcons contain, in their names, the word "icon," which is an indication that they are similar to icons and are a metaphor used to name objects in a GUI. Thus, it can also be assumed that these non-speech sounds suggest their usage must be limited to a GUI.

On the other hand, audemes can be experienced independent of the user interface, for instance, using any sound playing device, audemes can be used to teach lectures in a class. The naming chosen for this novel non-speech sound—audeme—is an additional suggestion that audemes are designed to be used beyond GUI (i.e., the audeme does not contain the word “icon”). The name *audeme* suggests the smallest entity of sound used to convey a meaning, similar to a *phoneme*, which is the smallest phonetic component in a language sufficient to communicate a distinction in the semantics of a word (Mannheimer, Ferati, Huckleberry, & Palakal, 2009).

8.1 Research Questions Revisited

The first research question for this study was to discover whether audemes help users to better memorize content when played along with theme-based content. In order to answer this question, experiments were conducted with blind and visually impaired K-12 students. The findings showed that the content was significantly better remembered when audemes were played along with it. Also, audemes helped in reducing memory erosion as, even after five months, the content was remembered better along with an audeme than without it.

The second research question was to discover the function of audemes when integrated into an aural content-rich interface used by the blind and visually impaired. In order to answer this question, an aural interface was built in which audemes were used as anticipators for essays presented in a TTS fashion. In the usability evaluation of the interface with blind and visually impaired students, audemes were used as aural covers to help the users anticipate the underlying TTS essay content associated with it. Essentially, the users recognized the content that was about to be presented to them based on the audeme that they heard.

Table 8.1. Synopsis of the research questions and answers.

Research Question	Answer
Q1. When a new category of non-speech sound (audeme) is played along with theme-based content, does it help the users to better memorize that content?	Audemes help users to better memorize theme-based content when they are played along with the content.
Q2. What is the function of audemes in content-rich interfaces for the blind and visually impaired?	Audemes serve as content anticipators when used in content-rich interfaces for the blind and visually impaired.
Q3. What characteristics of audemes help blind users recognize their meanings?	Numerous audeme characteristics, such as the serial concatenation of sounds and a mixture of different sound types, help blind users recognize the meaning of the audemes. The full list can be found in Section 6.7.

The third research question focused on the characteristics of audemes that help blind and visually impaired users recognize their meaning. In order to answer this question, it was first necessary to understand the process of meaning generation in audemes. To this end, audemes were analyzed using two theoretical frameworks: semiotics and sound modes of listening. Using these two theories to frame and conduct experiments, it became evident that audemes would be most effective if certain guidelines were followed upon their generation. These empirically derived guidelines were then integrated into an interface that enables individuals interested in using audemes to generate them to be effective in terms of their meaning recognition. A synopsis of the three research questions and their answers is given in Table 8.1.

8.2 Limitations

Every effort was made to ensure that the information gathered for this dissertation was a reliable and valid reflection of those users taking part in the experiments. Some limitations exist, however, that need to be listed here as the boundaries of the extent that these results can be generalized. First, a limited sample size was available for the experiments, with a range of 8 to 21 participants for the different experiments. When running experiments with blind and visually impaired participants, it is very common to have a limited sample. Studies suggest that most experiments conducted with blind participants typically have only four to six participants (Stevens, 1996; Challis & Edwards, 2000). Reasons for this are typically the difficulties to obtain large number of blind participants due to mobility issues and skill qualifications.

With respect to the participants, the second limitation of this dissertation is the population from which the sample participants were drawn. For the experiments, typical blind and visually impaired students from a K-12 U.S. Midwest school were used. On one hand, using this type of participants is an advantage for the study of audemes since audemes are sounds and studies have shown that blind participants have a heightened ability in processing sounds compared to sighted participants (Doucet et al., 2005; Sánchez et al., 2001). On the other hand, audemes can possibly be useful to a wider range of the population and not limited to the blind and visually impaired community. However, the use of audemes outside of the blind and visually impaired community was not investigated in this study.

The third limitation pertains to the content presented as stimuli in the experiments. Both the audemes and essays were created within the theme of U.S. history keyed to the Indiana K-12 education standard. The reason for this decision was to present content suitable for the level of the participants used in experiments. In general, the outcomes from this dissertation could apply to broad spectrum of content; however, these limitations are discussed here in order to bring awareness that using audemes in other scenarios might yield different results.

8.3 Future Research Directions

8.3.1 Audemes Used to Communicate a Process

There are few directions in which the current study could be expanded. One possibility is that audemes could be studied using a different population (i.e., sighted college students or older populations). While the current study exclusively studied audemes using blind and visually impaired participants, future studies could investigate the effect of audemes in the education of the sighted population, which could lead to a finding showing that audemes can not only be used as a memory catalyst, but also to communicate a process or procedure. For example, one of the avenues in which audemes could be investigated is in the military emergency drills or natural disasters. In these scenarios, whenever an alarm sounds, the only information conveyed through the siren is that a danger is approaching; however no information is given as to what procedure to follow to escape or protect oneself from the danger. Typically, procedures to be followed in the case of an emergency, are taught in special training, but the order of the actions needed to be taken (e.g., to seek shelter) could be forgotten over time. Audemes could be used along with the standard alarm sound to convey the procedure that needs to be followed. For example, when a tornado alarm is heard, in addition to the siren, an audeme could be played with the sounds for basement, bathtub and table, indicating the places where shelter is commonly sought (i.e., the basement of the house, bathtub or large table when the first two are not available). This combination would remind people as to the actions that they should undertake when a tornado is approaching.

Similar to this situation, audemes could be used to convey procedures that could be helpful to parents who typically, every morning, need to remind their kids what to do when going to the bathroom. An audeme could be created that includes the sound of water getting flushed down to a toilet, of the sound of brushing teeth and the sound of washing hands. This audeme could be used to remind kids of the order of the things they need to do when inside the bathroom.

8.3.2 Devising Further Guidelines for the Audeme Design

Another direction for future audemes studies could be to conduct experiments to identify additional guidelines that will help the process of forming audemes. One of these studies could be to investigate whether adding a pause between the sounds used to create an audeme would have an effect on the effectiveness of conveying its meaning. The introduction of a pause could be used to communicate audio gestalt, which is the perception of sounds as a group based on proximity and similarity.

Along these lines, future studies could also investigate the length of audemes (i.e., how many seconds an audeme should be in order to achieve the desire result).

8.3.3 Collaborating through Audemes

With respect to ASCOLTA, which finds audemes based on user defined keywords, the interface could be advanced to adopt a collaboration mode in which users share their created audemes and hear audemes created by other users, all within the interface itself. In this sense, users could compete to generate the best audeme for a given topic. The process of generating and co-creating audemes could enhance students' engagement with a topic of study. Along these lines, the ASCOLTA back-end could be enhanced using data mining techniques that would make it possible to automatically generate audemes for an entire book. This application could advance the practice of audemes to where, perhaps, one day, textbooks for K-12 schools could be accompanied by a CD containing audemes for each chapter to be used by teachers in their classes.

8.4 Closing Remarks

This dissertation highlights four novel contributions to the field of non-speech sounds for human-computer interaction. First, it presents the first work that introduces audemes as a new category of non-speech sound aimed at conveying theme-based content. Second, it demonstrates the use of audemes as novel aural covers in an acoustic interface (AEDIN) built for blind and visually impaired K-12 school level users. Third, it introduces novel and empirically validated guidelines for the creation of non-speech sounds that can effectively be used for the retention or anticipation of theme-based content. Finally, it is the first work to operationalize the guidelines into a tool (ASCOLTA) aimed at enabling individuals who do not have an audio background to easily create audemes.

9. APPENDICES

9.1 Appendix A. Experiment Results

9.1.1 Experiment 1: Exploring Information Recognition using Audemes

Subject	Impairment	Gender	Group	Pretest_Radio	Posttest_Radio	Pretest_U.S.Constitution	Posttest_U.S.Constitution	Pretest_Slavery	Posttest_Slavery
1	Blind	F	Control	1	5	7	6	6	3
2	Visually Impaired	F	Control						
3	Blind	M	Control			5	3		
4	Visually Impaired	M	Control	6	5	3	5		
5	Blind	F	Control	1	2	5	5	2	5
6	Visually Impaired	M	Control	3	6	4	4	4	6
7	Blind	F	Control	6	7	4	4	8	10
8	Visually Impaired	F	Encode	7	9	9	8	9	10
9	Blind	M	Encode						
10	Blind	M	Encode			8	7		
11	Visually Impaired	F	Encode	1	5	2	6	5	7
12	Blind	M	Encode	3	3	1	0	3	2
13	Visually Impaired	M	Encode	6	8	8	10	8	9
14	Blind	F	Encode	4	4	3	5		
15	Visually Impaired	F	Encode+Recognition	6	6	7	6		
16	Visually Impaired	F	Encode+Recognition			2	5		
17	Visually Impaired	M	Encode+Recognition	4	5	1	4	2	8
18	Visually Impaired	F	Encode+Recognition						
19	Visually Impaired	F	Encode+Recognition	4	5	0	5	1	5
20	Blind	F	Encode+Recognition	2	6	2	4	6	5
21	Blind	F	Encode+Recognition	5	10	8	9	6	9

9.1.2 Experiment 2: Studying Comprehension and Involvement with Audeme-based Essays

Slavery_Q14_I could easily write a summary of the essay.	4	6	4	7	6	4	5	7	7	7	5	7	6	4
Slavery_Q13_Listening to the essay made me confused.R	7	6	6	7	7	6	7	7	7	7	6	7	5	5
Slavery_Q12_The content presented in the essay makes sense.	6	6	5	7	7	5	7	6	7	7	5	7	5	3
Slavery_Q11_I had trouble understanding the details of the essay. R	7	4	7	7	7	5	7	7	7	7	6	7	6	4
Slavery_Q10_I had trouble understanding the main ideas of the essay. R	7	6	7	7	7	6	7	7	7	7	6	7	6	4
Slavery_Q9_It was clear for me the role of each detail information in the essay as a whole.	6	6	4	7	7	4	3	3	7	7	5	7	5	4
Slavery_Q8_Sometimes I felt lost when I was listening to the essay. R	6	6	5	7	6	6	6	4	6	4	6	7	5	5
USConst_Q14_I could easily write a summary of the essay.	4	5	2	5	6	4	5	5	4	7	5	6	4	4
USConst_Q13_Listening to the essay made me confused.R	7	7	7	7	7	6	7	5	6	6	7	6	5	6
USConst_Q12_The content presented in the essay makes sense.	6	6	5	7	7	4	6	6	6	7	5	7	6	4
USConst_Q11_I had trouble understanding the details of the essay. R	7	6	7	2	7	6	7	5	3	7	7	7	6	6
USConst_Q10_I had trouble understanding the main ideas of the essay. R	7	6	7	7	7	6	7	5	6	7	7	6	7	6
USConst_Q9_It was clear for me the role of each detail information in the essay as a whole.	5	6	5	6	7	3	5	3	6	5	6	7	5	4
USConst_Q8_Sometimes I felt lost when I was listening to the essay. R	3	7	4	6	6	5	7	3	3	5	6	5	3	4
Radio_Q14_I could easily write a summary of the essay.	5	5	3	6	7	3	5	7	3	7	5	6	3	4
Radio_Q13_Listening to the essay made me confused.R	7	6	6	7	7	5	7	7	6	6	6	7	3	5
Radio_Q12_The content presented in the essay makes sense.	6	6	5	6	7	5	7	6	6	6	5	7	4	3
Radio_Q11_I had trouble understanding the details of the essay. R	7	4	4	7	7	4	7	6	5	6	7	7	3	5
Radio_Q10_I had trouble understanding the main ideas of the essay. R	7	6	6	7	7	6	7	3	6	7	7	7	6	6
Radio_Q9_It was clear for me the role of each detail information in the essay as a whole.	6	5	5	6	7	4	5	4	5	6	4	7	3	3
Radio_Q8_Sometimes I felt lost when I was listening to the essay. R	2	4	4	7	6	4	7	5	2	2	2	7	4	6
Slavery_Q7_I found my mind wandering while listening to the essay. R	6	4	6	7	6	5	7	1	6	7	6	7	5	5
Slavery_Q6_The essay affected me emotionally.	5	6	6	6	5	5	5	2	4	2	2	6	6	4
Slavery_Q5_After finishing the essay, I found it easy to put it out of my mind. R	3	4	6	7	6	3	2	6	6	3	3	1	5	4
Slavery_Q4_I was mentally involved in the essay while listening to it.	6	6	5	7	6	4	5	1	7	7	5	7	6	4
Slavery_Q3_I could picture myself in the scene of the events described in the essay.	5	4	5	7	7	4	5	2	4	6	2	6	6	4
Slavery_Q2_While I was listening to the essay, activity going on in the room (lab) around me was on my mind. R	6	6	6	7	3	4	6	1	6	7	5	7	6	6
Slavery_Q1_While I was listening to the essay, I could easily picture the events in it taking place	6	6	5	7	7	5	6	2	7	7	4	7	6	4
USConst_Q7_I found my mind wandering while listening to the essay. R	3	4	4	7	5	5	7	6	3	5	7	4	2	4
USConst_Q6_The essay affected me emotionally	2	5	5	1	3	3	4	2	2	3	2	6	2	3
USConst_Q5_After finishing the essay, I found it easy to put it out of my mind	2	4	5	7	5	5	4	5	2	4	4	1	2	5
USConst_Q4_I was mentally involved in the essay while listening to it.	5	6	5	7	6	5	5	4	5	7	5	6	6	4
USConst_Q3_I could picture myself in the scene of the events described in the essay	5	5	3	6	6	3	5	2	2	4	2	4	3	2
USConst_Q2_While I was listening to the essay, activity going on in the room (lab) around me was on my mind. R	6	6	3	7	6	5	7	1	6	7	7	4	4	5
USConst_Q1_While I was listening to the essay, I could easily picture the events in it	6	6	4	6	6	3	5	2	5	3	4	5	6	3
Radio_Q7_I found my mind wandering while listening to the essay. R	6	4	4	7	6	4	7	1	6	4	6	5	4	5
Radio_Q6_The essay affected me emotionally	1	5	5	4	2	3	4	1	2	2	1	5	1	3
Radio_Q5_After finishing the essay, I found it easy to put it out of my mind. R	3	2	6	7	4	4	3	5	3	2	6	1	3	5
Radio_Q4_I was mentally involved in the essay while listening to it	6	4	5	7	7	5	7	6	6	6	5	7	5	4
Radio_Q3_I could picture myself in the scene of the events described in the essay	5	2	4	6	4	4	7	2	5	6	3	6	5	3
Radio_Q2_While I was listening to the essay, activity going on in the room (lab) around me was on my mind. R	7	6	4	7	5	4	6	2	6	2	5	6	5	6
Radio_Q1_While I was listening to the essay, I could easily picture the events in it taking place.	6	5	6	6	7	5	6	3	5	4	5	6	4	3
Q7_I feel relief rather than satisfaction after completing a task that required a lot of mental effort. R	4	5	6	4	5	4	6	5	5	7	3	1	4	4
Q6_I prefer to think about small daily projects to long-term ones R	4	6	4	4	5	4	3	4	3	1	4	1	3	4
Q5_I am attracted to situations where there is a likely chance I'll have to think in depth about something	2	6	5	7	7	6	6	7	6	7	6	7	5	4
Q4_I would prefer doing something that challenges my thinking abilities rather than do something that requires little thought.	4	2	3	7	7	7	6	7	6	7	6	7	4	2
SlaveryPosttest	5	3	8	6	8	8	9	9	10	8	3	4	5	9
SlaveryPretest	4	3	5	5	5	4	8	6	7	6	4	6	4	8
USConstPosttest	3	6	6	4	7	7	8	7	10	7	4	9	1	6
USConstPretest	2	2	4	3	5	2	5	4	5	4	1	0	5	3
RadioPosttest	6	6	4	5	9	6	8	6	7	3	4	5	3	6
RadioPretest	6	4	5	5	4	2	7	6	5	5	3	2	4	6
Gender	F	M	F	M	M	F	M	M	F	M	M	F	F	M
Participants	1	2	3	4	5	6	7	8	9	10	11	12	13	14

9.1.3 *Experiment 3: Usability Evaluation of AEDIN*

Parcticipant	Group (AEDIN v1 or AEDIN v2)	Impaiment	Gender	Q1.Understanding to use the interface was easy.	Q2.Understanding to use the buttons was easy.	Q3.The size of the buttons is appropriate.	Q4.Our explanation was sensible after you experienced the interface.	Q5.The bookshelf metaphor makes sense.	Q6.The structure of the interface makes sense.	Q7.The layout of the screen makes sense.	Q8.Audemes were meaningful.	Q9.Audemes were long enough.	Q10.Audemes were short enough.	Q11.Feedback sounds were meaningful.	Q12.Feedback sounds were long enough.	Q13.Feedback sounds were short enough.	Q14.Overall sound quality was good.	Q15.Using the interface is enjoyable.	Q16.The interface is fun to use.	Q17.I would use this interface again.	Q18.Using the touchscreen was easy.	Q19.Using the touchscreen was comfortable.
12	Version 1	Visually Impaired	M	2	4	4	4	1	3	4	1	3	3	4	3	3	2	2	2	1	4	4
13	Version 1	Visually Impaired	F	4	5	5	4	4	2	5	4	5	2	1	4	4	2	2	2	3	4	4
14	Version 1	Blind	M	5	5	5	5	4	5	5	4	4	3	2	5	1	4	2	2	3	5	4
15	Version 1	Visually Impaired	M	3	2	5	2	2	4	4	2	4	4	2	2	2	4	3	4	4	5	5
16	Version 1	Blind	F	4	2	2	5	5	5	4	4	4	3	2	3	2	5	4	4	4	4	4
17	Version 1	Visually Impaired	F	5	5	5	5	4	5	5	3	5	2	5	5	5	4	3	4	4	4	4
18	Version 1	Visually Impaired	F	5	5	5	4	4	4	5	4	3	3	4	4	4	4	3	3	4	4	4
19	Version 1	Visually Impaired	F	5	5	5	5	5	5	5	4	4	4	5	4	4	5	2	2	1	5	5
20	Version 1	Visually Impaired	M	4	4	4	4	4	4	4	3	5	1	2	4	3	4	4	4	4	5	5
1	Version 2	Visually Impaired	M	5	5	5	5	3	5	5	4	4	4	4	4	4	4	3	3	4	5	5
2	Version 2	Visually Impaired	M	2	5	5	4	4	4	5	4	5	2	5	4	2	4	2	2	2	4	4
3	Version 2	Blind	F	5	5	5	3	4	4	5	2	5	2	3	4	4	5	4	2	1	5	5
4	Version 2	Blind	F	5	5	5	5	5	5	5	5	5	4	5	5	5	2	4	4	4	4	4
5	Version 2	Visually Impaired	M	4	4	4	4	4	4	4	3	3	3	4	4	4	4	3	3	3	4	4
6	Version 2	Visually Impaired	M	5	5	5	5	5	3	4	5	5	1	5	5	3	5	4	4	5	5	5
7	Version 2	Blind	F	5	5	4	4	4	4	4	3	5	2	5	5	4	4	4	4	3	4	4
8	Version 2	Blind	F	5	5	3	3	3	4	4	3	4	4	5	5	5	3	5	5	4	4	4
9	Version 2	Visually Impaired	M	4	4	4	3	2	4	4	3	4	2	2	4	2	5	3	3	4	4	4
10	Version 2	Blind	F	3	4	5	4	4	3	4	4	5	5	4	4	4	5	5	4	4	5	4
11	Version 2	Visually Impaired	M	5	5	5	5	4	4	5	3	5	5	5	5	5	5	3	3	5	5	5

9.1.4 *Experiment 4: Audeme-based Information Recognition over Time*

Week4Q16.SFX-Music-Parallel 1	3	3	3	3	3	2	3	
Week4Q15.SFX-Music-Serial 5	3	3	3	3	2	3	3	
Week4Q14.SFX-Music-Serial 4	3	1	3	3	2	2	2	
Week4Q13.SFX-Music-Serial 3	3	3	3	3	2	3	2	
Week4Q12SFX-Music-Serial 2	2	3	3	3	3	2	2	
Week4Q11.SFX-Music-Serial 1	3	3	3	3	2	2	3	
Week4Q10.Music-SFX-Serial 5	3	3	3	3	2	3	2	
Week4Q9.Music-SFX-Serial 4	3	3	3	3	2	2	3	
Week4Q8.Music-SFX-Serial 3	3	3	3	3	2	3	3	
Week4Q7.Music-SFX-Serial 2	3	3	3	3	2	2	3	
Week4Q6.Music-SFX-Serial 1	3	3	3	3	3	3	3	
Week4Q5.Music-Music-Serial 5	3	3	1	3	2	2	2	
Week4Q4.Music-Music-Serial 4	2	3	2	2	2	1	3	
Week4Q3.Music-Music-Serial 3	3	3	1	3	3	2	3	
Week4Q2.Music-Music-Serial 2	2	3	2	3	3	3	1	
Week4Q1.Music-Music-Serial 1	3	3	3	3	3	3	3	
Week1Q35.SFX-SFX-Serial 5	2	3	2	3	2	1	1	2
Week1Q34.SFX-SFX-Serial 4	2	2	3	3	3	3	3	3
Week1Q33.SFX-SFX-Serial 3	3	3	3	3	2	2	2	3
Week1Q32.SFX-SFX-Serial 2	3	3	3	3	2	2	2	3
Week1Q31.SFX-SFX-Serial 1	3	2	2	3	3	2	3	3
Week1Q30.SFX-SFX-Parallel 5	2	2	3	3	2	2	3	3
Week1Q29.SFX-SFX-Parallel 4	2	2	1	3	1	2	2	2
Week1Q28.SFX-SFX-Parallel 3	2	2	3	3	1	3	2	2
Week1Q27.SFX-SFX-Parallel 2	1	3	3	3	0	2	0	3
Week1Q26.SFX-SFX-Parallel 1	2	3	2	3	1	2	2	3
Week1Q25.Music-Music-Parallel 5	3	3	3	2	2	2	2	3
Week1Q24.Music-Music-Parallel 4	2	0	2	3	1	1	1	3
Week1Q23.Music-Music-Parallel 3	2	2	2	3	2	1	3	3
Week1Q22.Music-Music-Parallel 2	3	3	3	3	2	1	2	3
Week1Q21.Music-Music-Parallel 1	2	2	2	3	3	2	3	3
Week1Q20.SFX-Music-Parallel 5	2	2	1	3	0	3	2	
Week1Q19.SFX-Music-Parallel 4	3	3	3	3	2	2	2	
Week1Q18.SFX-Music-Parallel 3	3	2	2	3	3	2	3	
Week1Q17.SFX-Music-Parallel 2	3	3	3	3	3	2	2	
Week1Q16.SFX-Music-Parallel 1	3	3	2	3	2	2	3	
Week1Q15.SFX-Music-Serial 5	3	3	3	3	2	3	2	
Week1Q14.SFX-Music-Serial 4	1	1	3	3	2	2	2	
Week1Q13.SFX-Music-Serial 3	3	3	3	3	2	3	2	
Week1Q12.SFX-Music-Serial 2	3	3	3	3	3	2	2	
Week1Q11.SFX-Music-Serial 1	3	3	3	3	3	3	3	
Week1Q10.Music-SFX-Serial 5	3	3	3	3	3	2	3	
Week1Q9.Music-SFX-Serial 4	2	3	3	3	3	1	2	
Week1Q8.Music-SFX-Serial 3	2	1	3	3	1	3	3	
Week1Q7.Music-SFX-Serial 2	3	3	3	3	2	2	3	
Week1Q6.Music-SFX-Serial 1	3	3	3	3	2	3	2	
Week1Q5.Music-Music-Serial 5	3	2	2	3	2	2	2	
Week1Q4.Music-Music-Serial 4	2	3	2	2	3	2	3	
Week1Q3.Music-Music-Serial 3	3	3	2	3	3	3	3	
Week1Q2.Music-Music-Serial 2	2	3	2	3	3	3	2	
Week1Q1.Music-Music-Serial 1	3	2	2	3	3	3	3	
Impairment (Blind,Visually Impaired)	B	B	VI	B	B	VI	B	VI
Gender	F	F	M	F	M	M	M	F
ParticipantID	1	2	3	4	5	6	7	8

Week7Q35.SFX-SFX-Serial 5	2	3	3	3	3	3	2	3
Week7Q34.SFX-SFX-Serial 4	2	3	3	3	2	3	3	3
Week7Q33.SFX-SFX-Serial 3	3	3	3	3	2	2	2	3
Week7Q32.SFX-SFX-Serial 2	3	3	3	3	3	3	3	3
Week7Q31.SFX-SFX-Serial 1	3	3	3	3	3	3	3	3
Week7Q30.SFX-SFX-Parallel 5	2	2	3	3	2	3	3	3
Week7Q29.SFX-SFX-Parallel 4	2	2	3	3	2	3	2	3
Week7Q28.SFX-SFX-Parallel 3	2	3	3	3	2	3	3	3
Week7Q27.SFX-SFX-Parallel 2	2	3	3	3	2	3	2	3
Week7Q26.SFX-SFX-Parallel 1	2	3	3	3	3	3	2	3
Week7Q25.Music-Music-Parallel 5	2	3	3	2	3	3	3	3
Week7Q24.Music-Music-Parallel 4	3	2	2	3	1	3	1	3
Week7Q23.Music-Music-Parallel 3	3	3	3	3	2	3	3	3
Week7Q22.Music-Music-Parallel 2	3	3	3	3	3	2	2	3
Week7Q21.Music-Music-Parallel 1	2	2	3	3	3	2	3	3
Week7Q20.SFX-Music-Parallel 5	3	3	3	3	2	3	3	
Week7Q19.SFX-Music-Parallel 4	3	3	3	3	3	2	2	
Week7Q18.SFX-Music-Parallel 3	3	3	2	3	3	3	3	
Week7Q17.SFX-Music-Parallel 2	3	3	3	3	3	3	3	
Week7Q16.SFX-Music-Parallel 1	3	3	2	3	2	2	3	
Week7Q15.SFX-Music-Serial 5	3	3	3	3	3	3	3	
Week7Q14.SFX-Music-Serial 4	3	3	3	3	2	3	1	
Week7Q13.SFX-Music-Serial 3	3	3	3	3	2	3	3	
Week7Q12.SFX-Music-Serial 2	3	3	3	3	3	3	2	
Week7Q11.SFX-Music-Serial 1	3	3	3	3	3	2	2	
Week7Q10.Music-SFX-Serial 5	3	3	3	3	3	3	3	
Week7Q9.Music-SFX-Serial 4	3	3	3	3	2	2	3	
Week7Q8.Music-SFX-Serial 3	3	3	3	3	2	3	3	
Week7Q7.Music-SFX-Serial 2	3	3	3	3	1	2	3	
Week7Q6.Music-SFX-Serial 1	3	3	3	3	3	3	3	
Week7Q5.Music-Music-Serial 5	3	3	3	3	3	2	3	
Week7Q4.Music-Music-Serial 4	3	3	2	3	2	3	3	
Week7Q3.Music-Music-Serial 3	3	3	2	3	3	3	3	
Week7Q2.Music-Music-Serial 2	3	3	3	3	2	3	2	
Week7Q1.Music-Music-Serial 1	3	3	3	3	3	3	3	
Week4Q35.SFX-SFX-Serial 5		3	2	3	3	3	1	3
Week4Q34.SFX-SFX-Serial 4		3	3	3	2	3	3	3
Week4Q33.SFX-SFX-Serial 3		3	2	3	2	3	2	3
Week4Q32.SFX-SFX-Serial 2		3	2	3	2	2	3	3
Week4Q31.SFX-SFX-Serial 1		3	3	3	3	3	3	3
Week4Q30.SFX-SFX-Parallel 5		3	3	3	2	3	3	3
Week4Q29.SFX-SFX-Parallel 4		2	3	3	1	3	2	2
Week4Q28.SFX-SFX-Parallel 3		3	3	3	2	2	3	3
Week4Q27.SFX-SFX-Parallel 2		3	3	3	1	3	1	3
Week4Q26.SFX-SFX-Parallel 1		3	3	3	2	3	3	3
Week4Q25.Music-Music-Parallel 5		3	2	2	2	1	2	3
Week4Q24.Music-Music-Parallel 4		2	1	2	1	2	2	2
Week4Q23.Music-Music-Parallel 3		3	2	3	1	2	2	3
Week4Q22.Music-Music-Parallel 2		3	3	3	3	2	2	3
Week4Q21.Music-Music-Parallel 1		2	3	2	2	2	3	3
Week4Q20.SFX-Music-Parallel 5	2	3	2	3	1	2	2	
Week4Q19.SFX-Music-Parallel 4	3	3	3	3	2	2	2	
Week4Q18.SFX-Music-Parallel 3	3	3	2	3	3	3	3	
Week4Q17.SFX-Music-Parallel 2	3	3	3	3	3	2	2	

9.1.5 Experiment 5: Information Recognition without Reinforcement

Week4Q35.SFX-SFX-Serial 5	3	3	2	3	3	3	3
Week4Q34.SFX-SFX-Serial 4	1	2	2	3	2	1	3
Week4Q33.SFX-SFX-Serial 3	1	2	2	1	2	2	2
Week4Q32.SFX-SFX-Serial 2	1	2	1	2	1	2	2
Week4Q31.SFX-SFX-Serial 1	3	2	2	2	2	1	2
Week4Q30.SFX-SFX-Parallel 5	2	3	3	3	3	2	3
Week4Q29.SFX-SFX-Parallel 4	2	2	2	2	2	3	2
Week4Q28.SFX-SFX-Parallel 3	2	2	2	3	1	2	3
Week4Q27.SFX-SFX-Parallel 2	3	2	2	2	3	2	3
Week4Q26.SFX-SFX-Parallel 1	2	2	2	2	2	3	2
Week4Q25.Music-Music-Parallel 5	2	3	2	2	2	2	3
Week4Q24.Music-Music-Parallel 4	1	1	1	1	2	1	2
Week4Q23.Music-Music-Parallel 3	2	2	1	2	2	3	2
Week4Q22.Music-Music-Parallel 2	1	2	3	2	2	1	2
Week4Q21.Music-Music-Parallel 1	1	0	2	2	1	0	3
Week4Q20.SFX-Music-Parallel 5	3	3	3	3	2	3	2
Week4Q19.SFX-Music-Parallel 4	2	2	3	3	2	2	3
Week4Q18.SFX-Music-Parallel 3	3	3	2	3	2	1	2
Week4Q17.SFX-Music-Parallel 2	1	2	1	3	1	1	2
Week4Q16.SFX-Music-Parallel 1	2	2	2	2	1	2	1
Week4Q15.SFX-Music-Serial 5	2	3	2	2	2	3	2
Week4Q14.SFX-Music-Serial 4	2	2	2	3	3	2	3
Week4Q13.SFX-Music-Serial 3	2	1	3	3	2	2	2
Week4Q12SFX-Music-Serial 2	2	3	2	3	3	2	1
Week4Q11.SFX-Music-Serial 1	3	2	2	2	2	2	2
Week4Q10.Music-SFX-Serial 5	2	1	1	2	1	3	2
Week4Q9.Music-SFX-Serial 4	2	2	2	3	1	2	2
Week4Q8.Music-SFX-Serial 3	2	2	2	2	1	2	2
Week4Q7.Music-SFX-Serial 2	2	2	2	2	2	2	3
Week4Q6.Music-SFX-Serial 1	3	3	2	1	1	1	2
Week4Q5.Music-Music-Serial 5	3	2	2	3	2	1	2
Week4Q4.Music-Music-Serial 4	2	2	2	3	2	2	2
Week4Q3.Music-Music-Serial 3	2	2	2	3	2	2	2
Week4Q2.Music-Music-Serial 2	1	1	1	2	1	3	2
Week4Q1.Music-Music-Serial 1	2	1	1	3	1	1	2
Impairment (Blind,Visually Impaired)	B	B	VI	B	B	VI	B
Gender	F	F	M	F	M	M	F
ParticipantID	1	2	3	4	5	6	7

9.1.6 *Experiment 6: Audeme Learnability*

Q21_SFX_SFX_Serial_7facts_	0.43	0.43	0.71	0.86	0.29	0.29	0.57	0.29	0.71	0.14
Q20_SFX_SFX_Serial_5facts_	0.8	1	0.6	1	0.8	1	0.8	0.6	0.6	0.4
Q19_SFX_SFX_Serial_3facts_	0.67	0.67	0.33	0.67	0	0.67	1	1	0.67	1
Q18_Music_SFX_Serial_7facts_	0.86	0.43	0.14	0.86	0.71	0.57	0.71	0.57	0.57	0.57
Q17_Music_SFX_Serial_5facts_	1	0.4	0.6	0.8	0.4	0.8	0.8	0.6	0.6	0.8
Q16_Music_SFX_Serial_3facts_	0.67	0.33	0	1	0.33	0.67	0.67	0.33	1	0.67
Q15_SFX_Music_Serial_5facts_	1	0.8	0.8	1	1	0	1	0.8	0.8	0.4
Q14_SFX_Music_Serial_3facts_	1	0.67	0.33	1	0.67	0.67	1	0.67	1	0.67
Q13_SXF_Music_Serial_7facts_	0.57	0.57	0.29	0.57	0.71	0.57	0.14	0.71	0.57	0.57
Q12_SFX_SFX_Parallel_7facts_	0.71	1	0	1	0.86	0.43	0.71	0.71	0.71	0.43
Q11_SFX_SFX_Parallel_5facts_	1	0.8	0.2	1	1	0	1	0.6	1	0.4
Q10_Music_Music_Parallel_7facts_	1	0.71	0	1	0.86	0.43	0.57	0.43	0.57	0.57
Q9_Music_Music_Parallel_3facts_	0.67	0.67	0	0.67	0.33	0	0.67	0.33	0.33	0
Q8_SFX_SFX_Parallel_3facts_	0.67	1	0.67	1	0.67	0.33	0.33	0.33	0.33	0.33
Q7_SFX_Music_Parallel_3facts_	0.67	1	0.67	1	1	0.33	0.67	0.33	0.33	0.33
Q6_Music_Music_Serial_5facts_	0.8	0.6	0.4	1	0.8	0.8	1	0.6	0.4	0.2
Q5_Music_Music_Sserial_7facts_	0.43	0.86	0.29	1	0.86	0.43	0.71	0.71	0.57	0.57
Q4_Music_Music_Serial_3facts_	0.67	0.67	0.33	1	1	0.33	1	0.67	1	0.67
Q3_SFX_Music_Parallel_5facts_	0.6	0.8	0.2	1	1	0	0.6	0.8	0.4	0.4
Q2_SFX_Music_Parallel_7facts_	0.57	0.71	0.14	0.57	0.71	0.29	0.57	0.57	0.43	0.29
Q1_Music_Music_Parallel_5facts_	0.2	0.6	0.6	0.6	0.6	0.6	0.2	0.4	0.6	0.8
Impairment (Blind,Visually Impaired)	B	B	VI	B	VI	VI	B	B	VI	B
Gender	F	F	M	F	F	M	F	M	M	M
Participant_ID	1	2	3	4	5	6	7	8	9	10

9.1.7 Experiment 7: Evaluation of ASCOLTA

QUESTION1. Based on your understanding on what audemes are, do you think audemes can be an integral part in your teaching process? How?

PARTICIPANT1. Yes. IT helps with the scores, but also makes them engaged; everybody is excited about using audemes in class.

PARTICIPANT2. Teaches “Travel Skills” class. So audemes from environment are very useful, because I teach my class mostly outside. E.g. to explain students what is a residential environment, in which we could play audemes to explain.

Some teacher who use audemes in their teaching process, see effect in students performing better with audemes, especially, lower functioning students.

PARTICIPANT3. Yes, audemes are useful for the teaching process.

I create my terms (keywords), and not always rely on the keywords provided by the book.

Audemes are powerful, in any settings, not only blind, because blind people remember songs.

QUESTION2. Do you think ASCOLTA can be a tool to find and create audemes to be used in your teaching process? How?

PARTICIPANT1. Yes.

PARTICIPANT2. Yes, it’s very basic. Also, the freedom to let users edit the audeme is good, although audemes are created while applying the guidelines of a good design.

Stating in the beginning and also when you give the warning to write that there has been a research.

PARTICIPANT3. Yes, this tool is valuable. Because the goal is get audemes in the hands of other people, so they can create audemes themselves.

QUESTION3. If ASCOLTA was available right now, would you use it? Why? How?

PARTICIPANT1. I would like to use it with blind users and see how they like the audemes I create with ASCOLTA.

Be able to add more sounds, by adding a Browse related to sounds in DB or pull from the computer.

Make it collaborative, by enabling students created audemes, then share with friends.

Machine gun not good example for civil war. Musket cannon ball, instead.

PARTICIPANT2. Yes, I can use to create an audeme to represent the environment outside. Or even explain the indoor and outdoor with the help of audemes.

I would also like to choose the individual sounds and add those to the audeme in the process of editing it.

PARTICIPANT3. I would use it.

QUESTION4. Imagine a situation in which ASCOLTA will be most useful to you?

PARTICIPANT1. Audemes help in digesting, reviewing a content. Or even as a game, e.g. jeopardy.

PARTICIPANT2.*[participant answered this through previous questions]*

PARTICIPANT3. Review the material that I cover in a class. Throughout the year create audemes to test and review, and at the end, use audemes as a game.

Possibly, use the tool and create audemes right in the class as I teach the lesson. The audeme will be created with the suggestion from the kids. Essentially, makes student part of the experience.

QUESTION5. Can you summarize three good and three bad things about ASCOLTA?

PARTICIPANT1. Is it accessible with screen readers? Because we might have a blind teacher, or perhaps students using it.

Add and choose a different sound.

Bring my own sounds and use. Browse my own sounds, but limited to 3 seconds. Perhaps, provide a way to trim the sounds to 3 sec.

Instead of keywords, use whole paragraphs that can be inserted into the input field.

We can use ASCOLTA to offer a nine week rotation class, in which students will create audemes.

Collaborative; make it possible people to exchange audemes created.

PARTICIPANT2. Good: It's very easy to use. It's simple. It's good that you get just few audemes as a result, not too many.

Bad: BE able to add new sounds. Explain why the rule is there. E.g. If you are interested, please read here the research showing how we came up with rules.

PARTICIPANT3. It looks easy. It is very user friendly.

I like the fact that keywords are shown below the audeme.

QUESTION6. Based on your recent experience, how can ASCOLTA be improved to be more useful to you?

PARTICIPANT1. *[participant answered this through previous questions]*

PARTICIPANT2. *[participant answered this through previous questions]*

PARTICIPANT3. *[participant answered this through previous questions]*

QUESTION7. Does finding audemes based on keywords make sense to you?

PARTICIPANT1. Yes. In addition, maybe provide through a paragraph to enter.

PARTICIPANT2. Yes, it makes sense.

PARTICIPANT3. This is good. I like this way of putting keywords.

QUESTION8. Is it meaningful the way audemes are displayed?

PARTICIPANT1. Google type, very friendly interface. The interface makes sense.

PARTICIPANT2. Yes. Differentiate the audemes, because right now they look identical. Rank the audemes by providing the percentage on how accurate the audeme represents the keywords and rules applied.

PARTICIPANT3. Yes, its good.

QUESTION9. Is it meaningful the way sounds constituting an audeme are displayed?

PARTICIPANT1. Yes. Buttons make sense. Maybe add a 30 sec video how to use ASCOLTA.

PARTICIPANT2. Maybe it will make sense to make the sounds, instead of Sound 1, use Gettysburg.

For reordering, it would be cool to just drag the sounds around and change their order.

Reordering by changing the numbers is not that intuitive. Maybe add a sentence before the Reorder button to clarify how the reordering by entering the numbers works.

PARTICIPANT3. Yes, it is good. I like the choice of being able to change individual sounds.

QUESTION10. Do you have any other comment or observation on ASCOLTA that you would like to share with us?

PARTICIPANT1. Make the interface to be suitable to generate an audeme for low-performing students (causal) and high-performing students (referential).

It has the Google Search like interface, which is familiar.

It would be good to have this right now, so that students can use it.

Also, have cloud storage where created audeme will be put.

I suggested showing the keywords, or just like in Google, auto fill the word with words from the database.

PARTICIPANT2. E.g. should not be capitalized.

Providing guidelines like Steps is very good.

Make it accessible, because there are blind teachers working at our school. Or even make it accessible to enable students use ASCOLTA.

PARTICIPANT3. One of the students said audemes were mind triggers. But, also audemes are entertaining.

9.2 Appendix B. Experiment Stimuli

9.2.1 Experiment 1: Exploring Information Recognition using Audemes

ESSAYS

Radio essay:

[PLAY RADIO AUDEME]

Radio is a technology that uses electro-magnetic signals, sometimes called radio waves that travel through the air and many solid materials and can be detected by receivers, generally called radios. Although the basic principle of radio is simple, developing the complex technology for sending and receiving radio waves took almost a century to evolve. There are many military, communication and entertainment uses for radio, but most of us think of radio in terms of music and information broadcast by radio stations. As the 19th century progressed with many advances in science and technology building on previous efforts, it was clear to many inventors that wireless communication was possible. Many people working in different countries throughout the century in different countries contributed various pieces of the complex technology that would become modern radio. Some of the more famous people were Michael Faraday, James Clerk Maxwell, Thomas Alva Edison, Nikola Tesla, Ernest Rutherford and Guglielmo Marconi, who was awarded a British patent for radio in 1896. Marconi received an American patent, and in 1901, he conducted the first successful transatlantic experimental radio communication. In 1909 Marconi won the Nobel Prize for his work with radio.

[PLAY RADIO AUDEME]

In 1912, the famous ocean liner Titanic sank. This event made it clear to the public that lives might have been saved if radio communications had been available and monitored. After this, radio transmitters quickly became universal on large ships. In 1913, the International Convention for the Safety of Life at Sea produced a treaty requiring shipboard radio stations to be manned 24 hours a day.

At the same time the United States government decided to regulate the use of broadcast radio and the airwaves in America. Congress created the Federal Radio Commission to control broadcasting, and many years later this government agency evolved to become the Federal Communication Commission. In the beginning, most radio users were amateur broadcasters who built their own radios and communicated with other amateur users. When the government began to regulate broadcasting, this helped promote the growth of the commercial radio.

[PLAY RADIO AUDEME]

On February 17, 1919, station 9XM at the University of Wisconsin in Madison broadcast the first human speech to the public at large. That station is still on the air today as WHA. But the first for-profit commercial radio station in the US is generally thought to be KDKA in Pittsburgh, Pennsylvania, which in October 1920 received its broadcasting license and went on the air as the first US licensed commercial broadcasting station. However, in the early years of radio, broadcasting was not yet supported by advertising or listener sponsorship. The stations owned by radio manufacturers and department stores were established to sell radios and stations owned by newspapers were used to promote the sales of newspapers and express the opinions of the owners.

[PLAY RADIO AUDEME]

When commercial radio became widespread in 1930's many predicted the end of records. Radio was a free medium for the public to hear music for which they would normally pay. Radio ownership grew from 2 out of 5 homes in 1931 to 4 out of 5 homes in 1938. Meanwhile sales of recorded music fell from $75 million in 1929 to $26 million in 1938 (with a low point of $5 million in 1933). Nowadays, instead of competing, the music and radio industries are seen as reinforcing each other.

Before World War Two, radio was a very important part of American popular culture. During the day, people would listen to radio soap operas, dramatic shows which got their name because soap companies were often the sponsors. In the evening, families would gather around the radio to listen to music, comedy or adventure shows in much the same way that modern Americans gather around the television.

[PLAY RADIO AUDEME]

After the war, as car ownership become widespread, radio increasingly focused on musical entertainment. It is estimated that 80 percent of radio listening in America happens when people are in the car. Today, along with music, radio is also an important medium for news and information, as well as the shows known as "talk radio" in which well-known radio personalities comment on current events. There are two main types of radio stations: A.M. and F.M. Some stations have very powerful transmitters and can be received many miles away, while others are relatively low-powered and heard only for a few miles.

As of August 2007, the Federal Communication Commission had issued licenses for nearly 14,000 radio stations in the United States. These include almost 4,800 AM stations and over 9,000 FM stations. Among these FM stations, there are approximately 6,000 commercial and 3,000 educational stations.

Slavery essay:

[PLAY SLAVERYAUDEME]

Slavery is a social-economic system under which certain persons — known as slaves — are deprived of personal freedom and compelled to perform labor or services for the people who own them. Slavery was common throughout the ancient world, and many people became slaves when their countries were conquered during war. In modern times, slavery is considered a terrible, inhuman practice. Although outlawed in nearly all countries today, slavery is still secretly practiced in many parts of the world. There are an estimated 27 million victims of slavery worldwide.

[PLAY SLAVERY AUDEME]

Nowadays, slavery is illegal in the United States. Unfortunately, this was not always the case. Slavery was part of the economic system of colonial America. It was first practiced in 1619 soon after the English colonists first settled in Virginia. Just like in ancient times, many early American slaves were used in agriculture and cultivation. Because this was very hard work, slaves were valued for their physical strength, and given enough food to keep them reasonably healthy. Many slaves worked on large farms or plantations to grow cotton, sugar and tobacco. However, they were not provided with the any of the normal comforts that other people could enjoy. They often lived in very poor housing, and wore only simple clothes. Their lives were completely controlled by their owners, and they had no legal rights. Slave owners had the legal right to whip or beat their slaves if they wanted. Although some slaves could marry and have families, their owners could sell family members away to other places if they wanted to. In many states it was even illegal to teach slaves how to read and right.

[PLAY SLAVERY AUDEME]

In the very early decades of the American slavery institution, some slaves were indentured servants, a kind of slavery which typically lasted a period of four to seven years for white and black people alike. By the end of the 17th century slavery was far more common in the Southern colonies than in the North, and almost all slaves were people of African descent.

From about the 1640s until 1865, people were forcibly taken or kidnapped from their homes in Africa, locked into chains and transported across the Atlantic Ocean to be sold to people in America. Most of these slaves were owned legally by people, but also by a number of American Indians and free blacks. The vast majority of this slaveholding was in the southern United States; approximately one Southern family in four held slaves prior to war. And 95% African Americans of lived in the South, making up one-third of the population. In the North, only 1% of the population was African Americans.

[PLAY SLAVERY AUDEME]

Many people, particularly in the northern United States, believed that slavery was an evil institution, and worked hard to have it legally abolished. These people were called Abolitionists. When southern slaves managed to escape from their owners, many of them journeyed to the north along secret routes that stopped at homes and farms owned by Abolitionists. These routes were called the Underground Railroad. Some of the more important routes along the Underground Railroad came through Indiana.

The national debate over slavery came to a boil after the national election of 1860, when anti-slavery candidate Abraham Lincoln was elected president. Eleven southern states decided to withdraw from the United States. They became the Confederate States of America, or the Confederacy. The north was called the Union. The two sides fought a bitter war from 1861 to 1865, killing millions of Americans on both sides.

[PLAY SLAVERY AUDEME]

In 1863, President Lincoln issued the Emancipation Proclamation ending slavery in the Confederacy. In 1865, the Union armies prevailed in the Civil War and re-united all of the states in the United States of America. After the war was over, new state legislatures and a new Congress were elected. At that time, the country ratified the 13th amendment to the United States Constitution, forever ending slavery in our country.

U.S. Constitution essay:

[PLAY U.S. CONSTITUTION AUDEME]

In the United States, we have important documents and a complex legal system that govern our how our society operates. Taken all together, we call this the Law. Only legislatures may enact laws, but the courts interpret these laws and decide when someone has broken a law and when they are innocent.

The law of the United States was originally largely derived from English common law that was in force at the time of the Revolutionary War. Today, the supreme law of the land is the United States Constitution. All other laws, whether from Congress or any of the state legislatures, are less powerful or subordinate to that document.

[PLAY U.S. CONSTITUTION AUDEME]

The United States Constitution was adopted in its original form on September 17, 1787 by the Constitutional Convention in Philadelphia, Pennsylvania, and later ratified by conventions in each state in the name of "the People." Most legal scholars believe that the U.S. Constitution is the oldest written national constitution in the world. The handwritten original document is on display at the National Archives in Washington, D.C.

[PLAY U.S. CONSTITUTION AUDEME]

The US Constitution begins with these words in a section called the Preamble:

We the People of the United States, in Order to form a more perfect Union, establish Justice, insure domestic Tranquility, provide for the common defense, promote the general Welfare, and secure the Blessings of Liberty to ourselves and our Posterity, do ordain and establish this Constitution for the United States of America."

[PLAY U.S. CONSTITUTION AUDEME]

The preamble is a basic statement of purpose for the constitution. The first three words ("We the people"), is one of the most-quoted sections of the Constitution, and emphasizes the idea that the Constitution and all the laws that flow from it have been created by the people of the United States, and the people remain as the highest authority in our legal system.

If the people agree, the Constitution can be changed by adding an amendment in Congress and having three-quarters of the state legislatures approve the idea. So far, there have been over 10,000 amendments proposed for the Constitution, but only 27 have passed. These include ideas such as abolishing slavery, giving women the right to vote, and limiting the number of times the President can be elected.

[PLAY U.S. CONSTITUTION AUDEME]

The first 10 amendments to the Constitution were passed by congress and the state legislatures in 1791 right after the Constitution was enacted. Collectively, these 10 amendments are called the Bill of Rights. Some of the more important ideas contained in the Bill of Rights include: freedom of religion, freedom of speech, freedom of the press, freedom of assembly, and freedom of petition, the right of the people to keep and bear arms, freedom from searches, arrests, and seizures of property without a specific warrant; the right of someone accused of a crime to keep silent and not answer questions at a trial; the right to a speedy trial by a jury of peers; a prohibition against cruel and unusual punishment, and generally the idea that any right or power not specifically given to the federal government must remain with the people.

<u>QUESTIONNAIRES (participants need to pick one answer; bold indicates correct answer)</u>

<u>Pre-test for Radio essay</u>

1. Radio works by:
    a. Making very loud noises that can be heard for miles
    b. The vibration of musical notes that is transmitted along metal wires
    c. Underground transmission of sound
    d. Invisible electro-magnetic waves that travel through buildings and space
    e. Electro-magnetic particles created by musical instruments

2. Radio was invented by
    a. Many scientists and inventors all working for the same large company owned by Tesla and
    b. Many scientists working at different times in different countries to develop various parts of the complex technology of radio.
    c. Thomas Alva Edison and his company in New Jersey
    d. The United States government in the 1890s
    e. A team of scientists from Harvard
3. The first commercial or for-profit radio station was
    a. On the ocean liner Titanic
    b. Run by two college students from Wisconsin
    c. Established in Pittsburgh in 1920
    d. Only in operation during the winter
    e. Very difficult to hear because of bad equipment
4. In the beginning, radio stations
    a. Were created to help churches and universities broadcast weekly sermons and lectures
    b. Mainly seen as a way to sell home radio receivers or to let newspaper owners promote their newspapers
    c. Often failed commercially due to lack of public interest
    d. Very rare in the southern half of the country because warm weather interfered with radio signals
    e. Only broadcast at night
5. In America, government regulation of radio
    a. Began after the sinking of the Titanic
    b. Needed to prevent European broadcasters from taking over the industry
    c. Left to the individual states to control
    d. Only applied to amateur radio operators
    e. A controversial issue that divided the country
6. The popularity of radio

a. **Grew dramatically during the 1930s, with nearly four out of five American homes owning a radio by the end of the decade**
b. Grew slowly during the 1930s because most homes didn't have electricity
c. Was deliberately kept small by government regulators
d. Was greater among teenagers than older adults
e. Grew dramatically in the 1960s because of the rise of folk music broadcasts from college stations.

7. Before World War Two, radio
   a. Was mostly seen as a way for businessmen to entertain themselves when working late into the evening
   b. Was mainly a way to entertain young children while their mothers did housework during the day
   c. **An important family activity when everyone gathered around the radio in the evening to hear comedies and adventure shows**
   d. Mostly thought of as an educational service to help college students to study for tests
   e. Only popular among young people who liked jazz.
8. After World War Two, radio stations
   a. **Became more focused on popular music**
   b. Increasingly devoted airtime to classical music and commentary by famous philosophers
   c. Were forbidden to play any recorded music for fear that it would hurt record sales.
   d. Never broadcast during the time when TV stations were also broadcasting
   e. Concentrated on comedy and dramatic serials.
9. Most Americans listen to radio
   a. Before breakfast
   b. When they go to bed
   c. On vacation
   d. **In the car**
   e. Only on weekends

10. Today, there are nearly

    a. **5,000 radio stations in North America**
    b. Over 9,000 high-powered radio stations owned by colleges and universities
    c. More than 30,000 radio FM and AM stations
    d. Approximately 14,000 radio stations
    e. Only one major radio station for each American state

<u>Post-test (Pre-test + 3 noise questions) for Radio essay</u>

1. Daytime radio dramas were called soap operas because
    a. Because the plots of these dramatic stories were always very clean with no violence or immoral activities.
    b. They were originally developed to entertain workers in large soap factories along the east coast.
    c. Women listened to them while taking long baths
    d. The sad stories often made people cry like they had gotten soap in their eyes.
    e. **The shows were sponsored by soap companies who believed this kind of stories appealed to stay-at-home women.**
2. The first American patent for radio technology was awarded to
    a. Nicola Tesla and Ernest Rutherford
    b. The US Navy
    c. Edison
    d. **Marconi**
    e. John D. Rockefeller and a group of businessmen
3. In the early years, the interaction between the radio industry and the music industry was
    a. Always very good
    b. Something that developed only in the 1960s when Baby Boomer teenagers started driving around listening to rock-and-roll.

c. A real problem because no one would buy radios as long as they could listen to records at home.
d. A real problem because the record companies feared that radio listeners wouldn't buy records
e. Only a problem in the late 1990s when the Internet became a popular way to download music.

<u>Pre-test for Slavery essay</u>

1. Slavery was first practiced:
    a. In medieval Europe
    b. In prehistoric China
    c. In many ancient cultures
    d. In France and Germany during the 17th century
    e. In Africa during the Renaissance
2. The Underground Railroad was
    a. The system of tunnels and caves where Union soldiers hid guns and ammunition
    b. The system of night time transportation the Confederate soldiers used to smuggle supplies from Canada to the South
    c. The railroad built under the Appalachian Mountains for settlers moving from the eastern colonies to the west.
    d. A network of homes and farms that runaway slaves used to hide while they journeyed north away from the south.
3. One of the most important causes of the Civil War was
    a. The growing fear that slaves would become more successful than whites.
    b. The election of Abraham Lincoln as United States Senator and his strong vocal opposition to slavery.
    c. Fear that rich northern industrialists would buy all the southern plantations
    d. The election of Abraham Lincoln as President and the fear in the south that he would try to make slavery illegal.

   e. The failure of the cotton crop in the south and the decision by many slaves to leave to find work in the north.
4. The Civil War ended
   a. In 1863 when the Confederacy agreed to outlaw slavery and rejoin the Union
   b. In 1863 when the Emancipation Proclamation made slavery illegal
   c. When Abraham Lincoln was assassinated.
   d. In 1865 when the Northern armies overcame the Confederates.
   e. Only after the 13th Amendment to the Constitution was ratified.
5. Approximately what percentage of African Americans lived in the South?
   a. 25-30 percent
   b. 50 percent
   c. 65 percent
   d. 75-80 percent
   e. 95 percent
6. Today around the world slavery
   a. Is still practiced secretly but only in the southern hemisphere
   b. Completely gone from all countries
   c. Unfortunately still being practiced in many places, with an estimated slave population of over 25 million slaves
   d. A good business in many Pacific islands
   e. Only practiced in along the Pacific coast of South America
7. In Colonial America, slavery was most common
   a. In the Wild West
   b. In the southern colonies with agricultural economies
   c. In colonies with many Native Americans inhabitants
   d. During the summer months
   e. In the northern industrial regions
8. People who spoke out against slavery were called
   a. Abolitionists
   b. Confederates

   c. Apologists
   d. Northern industrialists
   e. Colonials
9. Life for slaves was
   a. Generally OK but without a lot of free time during working hours.
   b. Often as nice as their owners except slaves were not allowed to have pets or own guns.
   c. Much better in the winter when there wasn't a lot of hard agricultural work.
   d. **Miserable because they could be whipped, beaten or sold.**
   e. Good for those slaves who had special skills like doctors or accountants.
10. Most slaves in America
    a. Came from China
    b. **Came from Africa**
    c. Worked as slaves for only a few years before they were allowed to become free.
    d. Were slaves until they reached the age of 21
    e. Were men because the agricultural work was considered too hard for women.

<u>Post-test (Pre-test + 3 noise questions) for Slavery essay</u>

1. Slavery was outlawed in England
   a. after the American Civil War
   b. many years before the American Revolution
   c. Onlty after the French Revolution
   d. **In the early 19th century**
   e. Because of the collapse of the cotton industry.
2. In the very early years of American slavery, some slaves were called indentured servants, a kind of slavery
   a. Which lasted four years for whites but seven years for blacks.
   b. In which slaves were purchased for their teeth.
   c. Which only lasted until slaves reached their 21st birthday.

d. **Which typically lasted a period of four to seven years for white and black people alike.**

e. Which only applied to women.

3. In the 17th and 18th centuries, most American slaves

    a. Were brought over from China.

    b. First trained in England and France before sold to Americans.

    c. Worked in factories in northern colonies.

    d. **Were kidnapped in Africa and shipped across the Atlantic Ocearn to America.**

    e. Volunteers from Africa hoping to find a better life in America.

<u>Pre-test for US Constitution essay</u>

1. Under the US Constitution, laws can only be enacted by

    a. **legislatures**

    b. courts

    c. the President

    d. town hall meetings

    e. the mayor

2. The law of the United States was originally largely derived from

    a. French civil law

    b. Religious rules from the Bible

    c. **English common law that was in force at the time of the Revolutionary War**

    d. Laws of the Chinese Empire

    e. Native American customs.

3. Today, the supreme law of the land can be found in

    a. The laws passed by the US Congress

    b. Laws passed by state legislatures

    c. **The laws and principles in the United States Constitution.**

    d. The interpretation of existing laws by local courts and judges.

e. The rules and announcements from elected officials like the governor, the mayor and the president.

4. When was the United States Constitution adopted in its original form?
    a. On September 17, 1776 by the Continental Congress.
    b. On September 17, 1787 by the Constitutional Convention.
    c. On July 4, 1776 in Washington D.C.
    d. On January 1, 1861 at the start of the Civil War.
    e. On July 4, 1826 on the 50th anniversary of the birth of the country.
5. Most legal scholars believe that the U.S. Constitution is
    a. The longest legal document in American history.
    b. The most beautifully written legal document in Western Civilization.
    c. Second only to the Declaration of Independence in defining our legal system.
    d. The oldest written national constitution in the world.
    e. Outdated and needs to be rewritten.
6. The preamble is
    a. A basic statement of purpose for the constitution.
    b. The display case in which the Constitution is shown to the people at the US Capitol.
    c. The opening statement of purpose heard before any courtroom procedure in the US legal system.
    d. A pedestrian pathway around the US Capitol.
    e. The first step in amending the Constitution.
7. The first three words ("We the people"), emphasize the idea that the Constitution
    a. Only applies to human beings and not to land, animals or any crops and natural resources.
    b. Only applies to people born in America.
    c. Only applies to people who agree to the Constitution.
    d. Can be changed only if all the people in the country agree that an amendment is needed.
    e. Has its authority only because it grows out of the will of the people.

8. Amendments are changes to the Constitution that
    a. Must be approved by all the state legislatures.
    b. Have made it impossible for anyone over the age of 60 to vote.
    c. Have given more rights to more people, including making it legal for women to vote.
    d. Have been successfully proposed and passed more than 10,000 times.
    e. Are proposed and approved by the President so that government can function more easily.
9. The first 10 amendments to the Constitution were passed by congress and the state legislatures in 1791. Collectively, these 10 amendments are called:
    a. The Great Books
    b. The 10 Commandments
    c. The Rules of the Road
    d. The Bill of Rights.
    e. The Great Freedoms.
10. Given all that you know about our Constitution and legal system, which of the following rights is probably NOT in the first 10 Amendments?
    a. Freedom of religion
    b. Freedom of speech.
    c. Freedom of the press.
    d. The right to a speedy trial.
    e. The right to be paid more than someone who was hired after you were.

<u>Post-test (Pre-test + 3 noise questions) for US Constitution essay</u>

1. Freedom of religion and speech are
    a. Guaranteed in the US thanks to Presidential proclamation.
    b. Guaranteed to American citizens even traveling in other countries.
    c. Are available to US citizens if they pay their taxes.
    d. Two of the important ideas in the Bill of Rights.

   e. Never discussed in Congress because they are controversial.

2. The Preamble of the US Constitution says that
   a. The President commands the army
   b. Congress is the highest authority
   c. the US is no longer part of England
   d. Everyone in this country can vote
   e. The Constitution was written to help form a more perfect union.
3. There have been how many amendments to the US Constitution?
   a. 27
   b. 109
   c. 16
   d. 10
   e. 34

9.2.2 Experiment 2: Studying Comprehension and Involvement with Audeme-based Essays

<u>Narrative transportation, content comprehension and need for cognition questionnaire for all essays</u>

All questions are rated with the following scale:

Not at all (1) (2) (3) (4) (5) (6) (7) Very much

<u>Narrative Transportation Questions:</u>

1. While I was listening to the essay, I could easily picture the events in it taking place.
2. While I was listening to the essay, activity going on in the room (lab) around me was on my mind. (R)
3. I could picture myself in the scene of the events described in the essay.
4. I was mentally involved in the essay while listening to it.
5. After finishing the essay, I found it easy to put it out of my mind. (R)
6. The essay affected me emotionally.
7. I found my mind wandering while listening to the essay. (R)

Content Comprehension Questions:

8. Sometimes I felt lost when I was listening to the essay. (R)

9. It was clear for me the role of each detail information in the essay as a whole.

10. I had trouble understanding the main ideas of the essay. (R)

11. I had trouble understanding the details of the essay. (R)

12. The content presented in the essay makes sense.

13. Listening to the essay made me confused.(R)

14. I could easily write a summary of the essay.

Need for Cognition Questions:

15. I would prefer doing something that challenges my thinking abilities rather than do something that requires little thought.

16. I am attracted to situations where there is a likely chance I'll have to think in depth about something.

17. I prefer to think about small daily projects to long-term ones (R)

18. I feel relief rather than satisfaction after completing a task that required a lot of mental effort. (R)

9.2.3 Experiment 3: Usability Evaluation of AEDIN

Tasks for Usability Testing

Task 1: Play/Hear three audemes.

Task 2: Go to the Relations of one audeme.

Task 3: After you listen to an audeme, go to its essay.

Task 4: Listen to another essay for a different audeme.

Task 5: Continue listening to different audemes until you hit a pop-up question, then answer the question.

Task 6: Tell me your current score.

Task 7: Tell me available commands on the current screen.

Task 8: Explore the breadth of the grid by navigating UP, DOWN, LEFT and RIGHT.

All questions are rated with the following scale:

Not at all (1)   (2)   (3)   (4)   (5)Very much

1. Understanding to use the interface was easy.
2. Understanding to use the buttons was easy.
3. The size of the buttons is appropriate.
4. Our explanation was sensible after you experienced the interface.
5. The bookshelf metaphor makes sense.
6. The structure of the interface makes sense.
7. The layout of the screen makes sense.
8. Audemes were meaningful.
9. Audemes were long enough.
10. Audemes were short enough.
11. Feedback sounds were meaningful.
12. Feedback sounds were long enough.
13. Feedback sounds were short enough.
14. Overall sound quality was good.
15. Using the interface is enjoyable.
16. The interface is fun to use.
17. I would use this interface again.
18. Using the touchscreen was easy.
19. Using the touchscreen was comfortable.

Open Ended Questions:

1. What did you think were the biggest problems with the interface?
2. What different features would you liked to be added? Why?
3. What did you like the most about the interface?
4. What did you like the least about the interface?
5. Did you get lost on the space? How did you recover from it?

9.2.4 Experiment 4: Audeme-based Information Recognition over Time

Experiment 4		
#	**Audeme Type**	**Concepts**
1	Music-Music-Serial 1	a. Future to past b. New wave trendy to old-fashioned c. Dance ongoing social customs
2	Music-Music-Serial 2	a. Swinging - into the air b. Angelic movement c. Action into calm
3	Music-Music-Serial 3	a. Majestic-hipster young guy b. Prince Harry (young, hip Prince of England) c. Groovy Cool Establishment
4	Music-Music-Serial 4	a. Calming the Rush b. Great excitement then relaxing c. Blizzard into Spring
5	Music-Music-Serial 5	a. Grandpa and young grandson b. Mice running away from Elephant c. Escape from the South
6	Music-SFX-Serial 1	a. Partying till the morning b. Urbanizing the village c. Neighbors complain about loud music
7	Music-SFX-Serial 2	a. Designing Architecture b. Bringing order to something c. Home building
8	Music-SFX-Serial 3	a. Earthquake b. Trip to the carnival c. Calm before the storm
9	Music-SFX-Serial 4	a. Trip Overseas b. End of the Adventure c. Disappearing Act
10	Music-SFX-Serial 5	a. Titanic b. Government waste c. Kid playing in the bathtub
11	SFX-Music-Serial 1	a. Birds flocking in the air b. Hitchcock c. Mother calms her complaining children
12	SFX-Music-Serial 2	a. Flying car b. Hybrid car c. Going on a date
13	SFX-Music-Serial 3	a. Chicago Fire

		b. Resolving an Emergency c. Post-Disaster Relief
14	SFX-Music-Serial 4	a. Downfall of the American auto industry b. Leaving in sorrow c. Running into Trouble
15	SFX-Music-Serial 5	a. Chinese sense of time b. Ancient Asian wisdom c. Tea time
16	SFX-Music-Parallel 1	a. Beach Party b. Happy Islands c. Summer refreshments
17	SFX-Music-Parallel 2	a. Middle Eastern turmoil b. Difficult to understand the Middle East c. North African military Raid
18	SFX-Music-Parallel 3	a. Economic cycles b. Africa's impoverished economy c. Street music
19	SFX-Music-Parallel 4	a. Immigrants crossing the ocean b. Slave trade c. Seaside village
20	SFX-Music-Parallel 5	a. Daily anxiety b. Racing for nothing c. Accelerating pace of life
21	Music-Music-Parallel 1	a. Everyday Life b. Stubborn argument c. Bus ride
22	Music-Music-Parallel 2	a. Secret admirer b. Scary walk c. Searching for clues
23	Music-Music-Parallel 3	a. Nervous calculation b. Hurrying to finish c. Busy bees
24	Music-Music-Parallel 4	a. Rainstorm b. Cleaning up a mess c. Frustrating discussion
25	Music-Music-Parallel 5	a. Falling into depression b. Walking with chained feet c. Malfunctioning machinery
26	SFX-SFX-Parallel 1	a. Burning the barn b. Crazy farm

		c. Vandals at the zoo
27	SFX-SFX-Parallel 2	a. Chat by the fire place b. Lost in the forest c. Play fighting with sticks
28	SFX-SFX-Parallel 3	a. Great expectations b. Victory c. Raising a champion
29	SFX-SFX-Parallel 4	a. Rusty movement b. Opening the dark gate c. 19th-century technology
30	SFX-SFX-Parallel 5	a. Transformation b. Eternal bliss c. Immortality
31	SFX-SFX-Serial 1	a. Taking a pill b. Swallowing sand c. Thirsty mirage
32	SFX-SFX-Serial 2	a. Sending kisses b. Wind lover c. Pardon me while I kiss the sky
33	SFX-SFX-Serial 3	a. Missing the target b. Pennies in the water c. Plink plunk, sink sunk
34	SFX-SFX-Serial 4	a. Announcing the fight b. Pots overheating c. Empty mailboxes after email
35	SFX-SFX-Serial 5	a. Washing the sins b. Burning a hole in the sky c. A cooling rain

9.2.5 Experiment 5: Information Recognition without Reinforcement

Experiment 5		
#	**Audeme Type**	**Concepts**
1	Music-Music-Serial 1	a. Evening Serenade b. Quiet contemplation c. Romantic Dinner
2	Music-Music-Serial 2	a. Celebrating Easter b. Emergence of a Hero c. Dancing after Church
3	Music-Music-Serial 3	a. Growing up to be a serious person

		b. Stretching to reach a cloud c. Anxious childhood
4	Music-Music-Serial 4	a. Leaving Home b. Sudden break-up c. Ruining the mood
5	Music-Music-Serial 5	a. Ducks gliding on the pond b. Traffic flowing along c. Resolving an argument
6	Music-SFX-Serial 1	a. Dancing Assassin b. Historical Change c. Ending the Party
7	Music-SFX-Serial 2	a. Colorful old car b. Powerful Machinery c. Surprising Speed
8	Music-SFX-Serial 3	a. Concert in the Park b. Stepping in an unexpected puddle c. Reading in the bath tub
9	Music-SFX-Serial 4	a. Waking from a deep sleep b. Can’t seem to relax c. Sad passing of the hours
10	Music-SFX-Serial 5	a. Dashing into battle b. Shocked by the earthquake c. Youthful optimism destroyed
11	SFX-Music-Serial 1	a. Dramatic battle b. Attacking the monster c. Death of the General
12	SFX-Music-Serial 2	a. Returning to the farm b. Trip to the countryside c. Harvesting the wheat
13	SFX-Music-Serial 3	a. Driving in Bombay b. Crowded Streets c. Going to the Wedding
14	SFX-Music-Serial 4	a. Baby at a Concert b. Lifetime Experience c. A baptism
15	SFX-Music-Serial 5	a. Locking up your emotions b. Becoming insensitive c. Keeping a secret
16	SFX-Music-Parallel 1	a. Dying beast b. Defeating Russia

		c. Sorrowful Fighter
17	SFX-Music-Parallel 2	a. Cattle Round-up b. Rural Parade c. Happy Cows from California
18	SFX-Music-Parallel 3	a. Modern Irish Farm b. Dance of the Animals c. Farm Karaoke
19	SFX-Music-Parallel 4	a. Lassie Come Home b. Shepherd's lament c. Handsome Stranger
20	SFX-Music-Parallel 5	a. Breaking into the Church b. Falling from Grace c. Overwhelmed by Sin
21	Music-Music-Parallel 1	a. Bluesy Hippo b. Grouch at the Party c. Bad Fat Dancer
22	Music-Music-Parallel 2	a. Mixing Drinks at the Bar b. Disconnected Stories c. Parents and Kids
23	Music-Music-Parallel 3	a. Tuning up the Performance b. Group Preparation c. Lazy Drunk
24	Music-Music-Parallel 4	a. Happy Neighborhood b. Life and Death c. Fast and Slow
25	Music-Music-Parallel 5	a. Musical Monk b. Dripping Cave c. Trying to Cheer you Up
26	SFX-SFX-Parallel 1	a. Bad Athlete b. Flushed with Victory c. Arguing in the Bathroom
27	SFX-SFX-Parallel 2	a. Playing Outside b. Shooting Birds c. Larry Bird
28	SFX-SFX-Parallel 3	a. Canine Hygiene b. Guarding against Cavities c. After the Hot Dog
29	SFX-SFX-Parallel 4	a. Open Up b. Drinking in Private c. Take it Easy

30	SFX-SFX-Parallel 5	a. Time for Snacks b. Crunched for Time c. Crazy for Potato Chips
31	SFX-SFX-Serial 1	a. Interrupting a Thought b. Meeting Deadline c. Coming up with a New Idea
32	SFX-SFX-Serial 2	a. Coughing Up b. Smoking in the Boys Room c. Horrible Smell
33	SFX-SFX-Serial 3	a. Marathon Winner b. Tired of Fame c. Running out of Tricks
34	SFX-SFX-Serial 4	a. Cute Toy b. Comedian's Trick c. Hidden Listener
35	SFX-SFX-Serial 5	a. Allergic to Guns b. Sniper with a Cold c. Winter Battle

9.2.6 Experiment 6: Audeme Learnability

Experiment 6			
#	**Audeme Type**	**Number of Facts**	**Sentences with Facts**
1	Music-Music-Serial	3	1. The American concept of the separation of Church and State was formalized in the Bill of Rights in the First Amendment to the Constitution 2. About one quarter of Americans consider themselves Catholics 3. Several of the original 13 American colonies were founded by immigrants who came to America to escape official government religions in Europe.
2	Music-Music-Serial	5	1. Modern science has made it possible to see trillions of miles into space, both with visual light and with radio waves emitted by stars. 2. The visual microscope was invented in the late 1500s in Holland. 3. The most famous early microscope scientist was Anton Van Leeuwenhoek (Lay-van-hook) who first reported the discovery of micro-organisms. 4. The Scanning Probe Microscope is another type of

			very powerful modern microscope that does not rely on lenses and direct visual observation 5. The largest orbiting telescope is called the Hubble Space Telescope after American Astronomer Edwin Hubble
3	Music-Music-Serial	7	1. The American Civil War was fought between the Confederacy and the Union. 2. The Confederacy was also known as the South, or the Confederate States of America, or Dixie 3. The Union was also called the North, or the United States of America 4. The South withdrew or seceded from the Union in 1861 after Abraham Lincoln was elected President 5. Perhaps the most famous battle of the Civil War was Gettysburg, a Union victory 6. The theme song of the North was the Battle Hymn of the Republic. 7. The War ended in April, 1865.
4	Music-SFX-Serial	3	1. Thanks to its tradition of wagon building dating from the Civil War era, and its location as the "crossroads of America," Indiana was home to over 70 independent car companies each hoping to capture a slice of rapidly growing market, including Marmon, Crosley, Cord, Blackhawk, Studebaker and the Stutz 2. Perhaps the most famous Indiana-based car was the Duesenberg, which was manufactured in Indianapolis and which in the late 1920s and into the 30s was THE status symbol and luxury car, and was owned by movie stars such as Clark Gable and Gary Cooper, as well as gangster Al Capone and the queen of Yugoslavia and the King of Italy. 3. Although the Duesenberg Company went out of business in 1937, vintage models are still operating and being sold at high-end auctions, including Duesenberg once owned by actress Greta Garbo, which sold recently for $1.4 million.
5	Music-SFX-Serial	5	1. Kissing is a complex behavior that requires significant muscular coordination; involving a total of thirty-four facial muscles and 112 postural muscles.

			2. The most important muscle involved is the orbicularis oris muscle, which is used to pucker the lips and informally known as the kissing muscle. 3. Diseases which may be transmitted through kissing include mononucleosis, allergic reactions to nuts and drugs, and herpes, when the infectious virus is present in saliva. 4. In a scientific experiment, increasing the frequency of kissing in marital and cohabiting relationships was found to result in a reduction of perceived stress, an increase in relationship satisfaction, and a lowering of cholesterol levels. 5. Kissing in Western cultures is a fairly recent development and is rarely mentioned even in Greek literature and as recently as the Middle Ages it was considered a sign of refinement of the upper classes rather than a common activity for everyday life.
6	Music-SFX-Serial	7	1. The first artificially carbonated beverages sold in the United States came from a Yale University chemistry professor named Benjamin Silliman, who began selling “soda water” in 1806 in Connecticut. 2. Originally, carbonated beverages were considered healthy, and were sold in pharmacies. During the 19th century, pharmacists began to add flavoring ingredients and sweeteners to improve the taste, and from this practice evolved the modern soft drink industry. 3. According to the National Soft Drink Association (NSDA), consumption of soft drinks is now over 600 12-ounce servings (12 oz.) per person per year. 4. Since the late 1970`s the soft drink consumption in the United States has doubled for females and tripled for males. 5. The highest consumption is in the males between the ages of 12 - 29; they average 1/2 gallon a day or 160 gallons a year.

			6. To meet the business goal of increasing soft drink consumption, soft drink companies have increased the size of the drinks from 61/2 ounces in the 1960`s to the 20-ounce bottle of today, and allowed many restaurants, both fast food and traditional restaurants, to offer free refills. 7. The United States ranks first among countries in soft drink consumption, with annual sales approaching $60 billion and per-capita consumption of soft drinks is in excess of 150 quarts per year, or about three quarts per week.
7	SFX-Music-Serial	3	1. The Great Lakes are a collection of five freshwater lakes located in eastern North America, on the Canada – United States border. 2. Consisting of Lakes Superior, Michigan, Huron, Erie, and Ontario, with Superior being the largest. 3. The Great Lakes form the largest group of freshwater lakes on Earth by surface, occupying over 90,000 square miles and containing 20 percent of the world's freshwater.
8	SFX-Music-Serial	5	1. The word "skyscraper" originally was a nautical term referring to a small triangular sail set above the skysail on a sailing ship. 2. The term was first applied to buildings in the late 19th century as a result of public amazement at the tall buildings being built in Chicago and New York City. 3. Although there are no exact or legal definitions of skyscraper, generally speaking a skyscraper is considered to be at least 500 feet high, while a structure taller than 1,000 feet may be called a "supertall." 4. In terms of construction technology, the term skyscraper refers to a very tall, multi-storey building that uses an internal skeleton of steel or reinforced concrete rather than thick, load-bearing walls, which was the traditional way large buildings were engineered in past centuries. 5. The tallest building the USA is the Willis Tower in Chicago – formerly called the Sears Tower, which is 1,451 feet high.

9	SFX-Music-Serial	7	1. On August 12, 1833, the Town of Chicago was incorporated with a population of 350. 2. In 1840, Chicago was the ninety-second most populous city in the United States with a population of over 4,000. 3. In the pivotal year of 1848, Chicago saw its first steam locomotives, and its first steam-powered grain elevators, and the arrival of the telegraph. 4. In 1860, Chicago was the ninth most populous city in the country and hosted the Republican Party National Convention that nominated Abraham Lincoln. 5. Because of its central geographic location, and its rail, water and road connections – and later its airport – Chicago became the most transportation hub in the country. 6. As a transportation hub, Chicago naturally became the headquarters for major mail order catalogue businesses such as Montgomery Ward and Sears and Roebuck. 7. By 1870 Chicago had grown to become the nation's second largest city, and one of the largest cities in the world.
10	SFX-Music-Parallel	3	1. The idea of Daylight Savings Time was first proposed in 1895 by the New Zealand entomologist George Vernon Hudson. 2. Mechanical clocks were perfected so that monks and priests could know when it was time to offer prayers at certain established hours. 3. The circadian rhythm operates even when there may be no direct perception of light and dark.
11	SFX-Music-Parallel	5	1. Shenandoah River and its Valley are in Virginia 2. The song Oh Shenandoah is used informally as the Virginia State song 3. In the Civil War, both the North and South had armies named for the Shenandoah River 4. Human being first inhabited the Shenandoah Valley nearly 9,000 years ago 5. President Herbert Hoover had a summer residence in the land that became the Shenandoah National Park.
12	SFX-Music-	7	1. The Erie Canal is an artificial river that runs about 360 miles.

	Parallel		2. The Erie Canal runs between Albany and Buffalo in Upstate New York 3. It was officially opened for traffic in 1825 4. The Erie Canal helped open Western New York State for immigration, 5. The Erie Canal is still used today for both recreational and commercial traffic. 6. Most of the traffic along the canal was barges towed by teams of horses or mules who walked along the shore connected by towlines. 7. The movement to build a national network of connecting canals came to an end when ta workable steam-powered railroad was perfected in the 1830s.
13	Music-Music-Parallel	3	1. Since 1970, the number of first- generation immigrants living in the United States has quadrupled, from 9.6 million to about 38 million in 2007. 2. Over one million persons were naturalized as U.S. citizens in 2008. 3. The leading countries of birth of the new citizens were Mexico, India and the Philippines.
14	Music-Music-Parallel	5	1. In official US Army history the quarter century from 1865 to 1890 is known as the Indian Wars. 2. There are over five hundred and sixty-one self-governed tribes, states, and ethnic groups of Native Americans in the US. 3. The largest Indian tribe today is the Cherokee 4. Gambling has become a leading industry for Native American communities. 5. The Battle of the Little Big Horn is often called Custer's Last Stand
15	Music-Music-Parallel	7	1. NASA, the National Aeronautics and Space Administration, was established on July 29, 1958. 2. Currently NASA is supporting the International Space Station and is developing a new manned spacecraft called Orion. 3. NASA's annual budget is approximately $18 billion. 4. NASA was created in response to the Soviet government's successful launch of Sputnik in October, 1957. 5. The Advanced Research Projects Agency (ARPA) was

			also created at this time to develop space technology for military application and later developed the Internet. 6. On May 5, 1961 NASA launched America's first manned rocket, Freedom 7 with astronaut Alan Shepard onboard, although the ship only achieved sub-orbital flight and did not circle the earth 7. America's first truly orbital space flight occurred on February 20, 1962 when astronaut John Glenn circled the globe three times during the flight of Friendship 7.
16	SFX-SFX-Parallel	3	1. About 93% of traffic accidents are due to human factors including intoxication and driver error 2. About 12 percent of fatal car crashes were caused by drunk driving 3. The first speeding ticket in World history was given to a movie actor in Dayton Ohio in 1904 for going 12 miles per hour.
17	SFX-SFX-Parallel	5	1. One of America's greatest poets of the 19th century was Emily Dickinson (1830-1886), who lived a sheltered life in small-town Amherst, Massachusetts and published very little during her lifetime. 2. Walt Whitman (1819-1892) was arguably THE greatest poet in 19th century America, who wrote "Leaves of Grass" and "Song of Myself." 3. One of America's greatest lyric poets of the early 20th century was Wallace Stevens, who made a living as an insurance executive. 4. American poet T.S. Eliot's masterpiece was a 434 line poem titled "The Wasteland" written in 1922. 5. Robert Lee Frost (1874 – 1963) was an American poet who won four Pulitzer prizes and read a poem at the inauguration of President John Kennedy.
18	SFX-SFX-Parallel	7	1. The earliest known mention of baseball in the United States was a 1791 Pittsfield, Massachusetts ordinance banning the playing of the game within 80 yards (73 m) of the town meeting house. 2. Prior to the Civil War, baseball competed for public interest with cricket and regional variants of baseball, notably town ball played in Philadelphia and the Massachusetts Game played in New England.

			3. By 1867, after the end of the Civil War, membership in the National Association of Base Ball Players, the first national league, ballooned to over 400 clubs, including some clubs from as far away as San Francisco and Louisiana. 4. One of those early teams was the Chicago White Stockings, a team that later became the Chicago Cubs and are the oldest team in American organized sports. 5. The first professional black baseball club, the Cuban Giants, was organized in 1885. 6. Professional baseball was desegregated in 1947 when Jackie Robinson joined the Brooklyn Dodgers and became the first African American to play in the previously all-white major leagues. 7. The three all-time career leaders in major league homeruns are Barry Bonds with 762, Henry Aaron with 755 and Babe Ruth with 714.
19	SFX- SFX- Serial	3	1. The term Helicopter Parents is a colloquial, early 21st-century term for a parent who pays extremely close attention to his or her child's or children's experiences and problems, particularly at educational institutions, and who hover closely overhead, rarely out of reach, whether their children need them or not. 2. In Scandinavia, this phenomenon is known as curling parenthood after the winter game where a player propels a large stone disk across the ice and then hurries out ahead of it to sweep all the slightest bits of ice and snow out of its path to allow it to glide as far ans as easily as possible. 3. The term was originally coined by Foster W. Cline, M.D. and Jim Fay in their 1990 book Parenting with Love and Logic: Teaching Children Responsibility but became very popular in recent years, because, as some experts believe, of the universal use of cell phones, which have been called "the world's longest umbilical cord".
20	SFX- SFX- Serial	5	1. Texas Chain Saw Massacre is the title of a 1974 movie one of the best known American horror films that might be categorized as a “slasher film” characterized by intense and generally psychopathic murders and

			mayhem, and may also be included in discussions of a related genre called "splatter films." 2. Although there are examples of mindless or psychopathic violence in movies as far back as the 1930s, the modern "slasher film" starts to appear in the 1960s with classics such as Alfred Hitchcock's "Psycho" from 1960. 3. Some experts say that the first true slasher film was "Black Christmas" from 1974, which established much of the genre's classic cinematic formula that included a mysterious stalker, a set of adolescent victims, a secluded location cut off from adult supervision, point-of-view shots showing the "killer's perspective", and graphic depictions of violence and murder. 4. Although slasher and splatter films have always generated controversy from concerned parents and citizen groups, the films seem to consistently attract a loyal following at the box office, inspiring producers to create series such as Nightmare on Elm Street, the Halloween series and the Child's Play series starring a murderous doll named Chucky. 5. Although most slasher movies are not blockbuster hits at the box office, some individual films such as Saw3 have earned more than $150 million worldwide, and the combined revenues for the most popular series of movies routinely rise into the hundreds of millions of dollars.
21	SFX-SFX-Serial	7	1. Pilgrim's Plunge at the Holiday World amusement park in Santa Claus, Indiana, is the world's tallest water ride at 135 feet high. 2. The tallest roller coaster in the world is the 456-ft high Kingda Ka at Six Flags Great Adventure in Jackson, New Jersey 3. The World's oldest operating roller coaster is Leap the Dips at Lakemont Park in Altoona Pennsylvania, which opened in 1902 and reaches a top speed of 10 mph. 4. A 200-ft high coaster is known as a "hypercoaster" and a 300-ft high coaster is called a "Giga-coaster" and a 400-foot coaster is called a "Rocket coaster." 5. There are 2,398 roller coasters in the world, but only

			164 made of wood while the rest are steel. 6. According to the Consumer Products Safety Commission, there are, on average, 4.4 deaths each years caused by amusement park rides in the US. 7. Amusement parks are a multi-billion-dollar industry worldwide, with approximately 300 million people each year visit the top 40 amusement theme parks in Europe and North America.

9.2.7 Experiment 7: Evaluation of ASCOLTA

#	Sound Name	Meta names	Sound Type	Listening Mode
1	Cash register	cash, register, economy, money, market, stock, entering a store	Sound effect	Causal and Referential
2	Coins	coins, money, market, economy	Sound effect	Causal and Referential
3	Urban cars honking	urban, city, cars, honking, busy, fast	Sound effect	Causal and Referential
4	Workers construction	workers, work, labor, labour, construction, noisy environment	Sound effect	Causal and Referential
5	Airplane	airplane, transportation, fast, take off, landing	Sound effect	Causal and Referential
6	Church bell	bell, religion, church, discrimination, old days	Sound effect	Causal and Referential
7	Car starting	car, transportation, engine	Sound effect	Causal and Referential
8	Drink song	drink, alcohol, temperance, movement, party		Causal and Referential
9	Horse	horse, transportation, rhythm, horse riding	Sound effect	Causal and Referential
10	I’ve been working on a railroad song	railroad, working, construction, happy	Music	Causal and Referential
11	Sipping drink	sipping, drink, alcohol, temperance, thirst	Sound effect	Causal and Referential
12	Sweet lord song	slavery, emancipation, abolitionist, abolition		Referential

13	Train	train, transportation, old steam locomotive	Sound effect	Causal and Referential
14	Whip crack	whip, slavery, slaves, abolition, abolitionist, hurt, whipping a stone	Sound effect	Causal and Referential
15	Dixie song	Dixie, Civil War, Confederacy, South, old movie	Music	Referential
16	Glory, glory song	battle hymn of the republic, civil war, army, marching	Music	Referential
17	Shooting guns	gun, fight, war, military, die, injury	Sound effect	Causal and Referential
18	Signing proclamation amendment	emancipation, proclamation, decree, presidential, amendment	Sound effect	Referential
19	Gettysburg Address	Gettysburg, address, Lincoln, speech, Abraham, presidential	Sound effect	Causal and Referential
20	Swords	swords, fight, war, military, victory	Sound effect	Causal and Referential
21	Swords and guns	swords, fight, war, military, guns	Sound effect	Causal and Referential
22	Ending with drums	rough ending, completion, abrupt change, disconnect , shock, drum	Music	Reduced
23	Ending classic	successful ending, completion, achievement,	Music	Reduced
24	Country music	Country, old, happy, backward, past, farm	Music	Causal and Referential
25	Ending piano	Ending, completion, sadness	Music	Reduced
26	Speeding music	Fast forward, quick, speed up, accelerated, start of a happy music, E.T., electron	Music	Reduced
27	Starting glamorous	Important, kingdom, presentation, glamour, great, concert, start	Music	Referential
28	US hymn	US hymn, patriot, countrymen, military, 4th of July, president	Music	Referential
29	Birds	Birds, safe environment, happy, springtime, forest	Sound effect	Causal and Referential
30	Clock ticking	Time, clock, ticking, last minute, under stress, broken	Sound effect	Causal and Referential

		clock		
31	Fireworks	Fireworks, holiday, 4th of July, celebration	Sound effect	Causal and Referential
32	People talking busy room	People talking, busy room, meeting, discussion, hot	Sound effect	Causal and Referential
33	School bell	Start, education, lecturing, school ending, lunch time, school bell, class break	Sound effect	Causal and Referential
34	Wind	Wind, cold, winter, tornado, snow	Sound effect	Causal and Referential
35	Clapping ovation	Success, clapping, ovation, achievement,	Sound effect	Causal and Referential
36	Flipping magazine pages	Reading, magazine, education, learning, change, advancement, boring afternoon, flipping a page	Sound effect	Causal and Referential
37	Heavy machinery working	Machine, factory, industry	Sound effect	Causal and Referential
38	Laser	Future, space, electron, robot, E.T.	Sound effect	Referential
39	Office typing chatting phone ringing	Busy office, machine typing, phone, ringing,	Sound effect	Causal and Referential
40	Phone dialing	Phone, dialing, reaching, connecting	Sound effect	Causal and Referential
41	Plunging into water	Plunging, water, unsuccessful, dead end, failure, thirst	Sound effect	Causal and Referential
42	Classic orchestra	Grand opening, great achievement,	Music	Reduced
43	River	River, flowing, progress, change,	Sound effect	Causal and Referential

9.3 Appendix C. Design Mockups

9.3.1 AEDIN

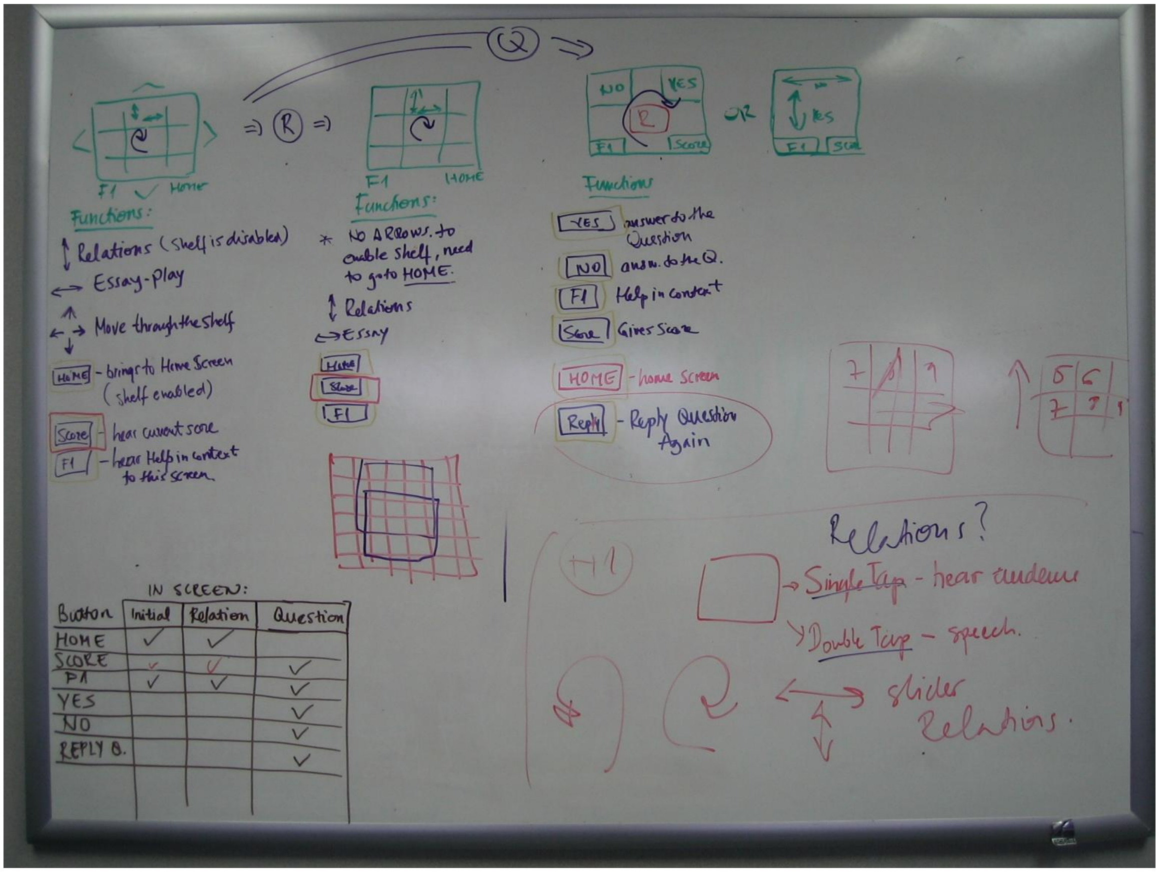

Figure 9.1. Initial sketching of AEDIN.

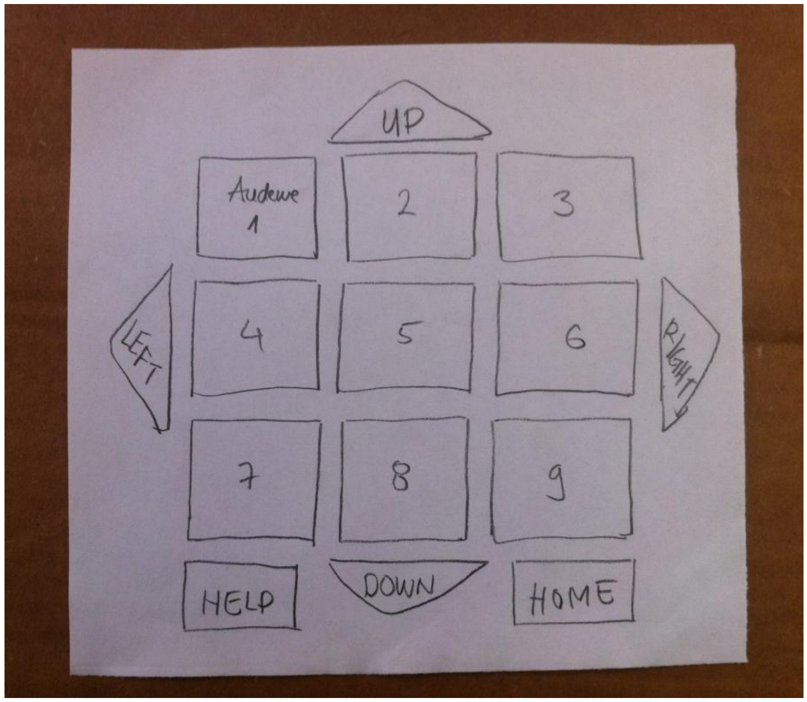

Figure 9.2. Paper prototype of AEDIN's main screen. The paper was pasted into a cardboard simulating the size of an actual desktop computer screen.

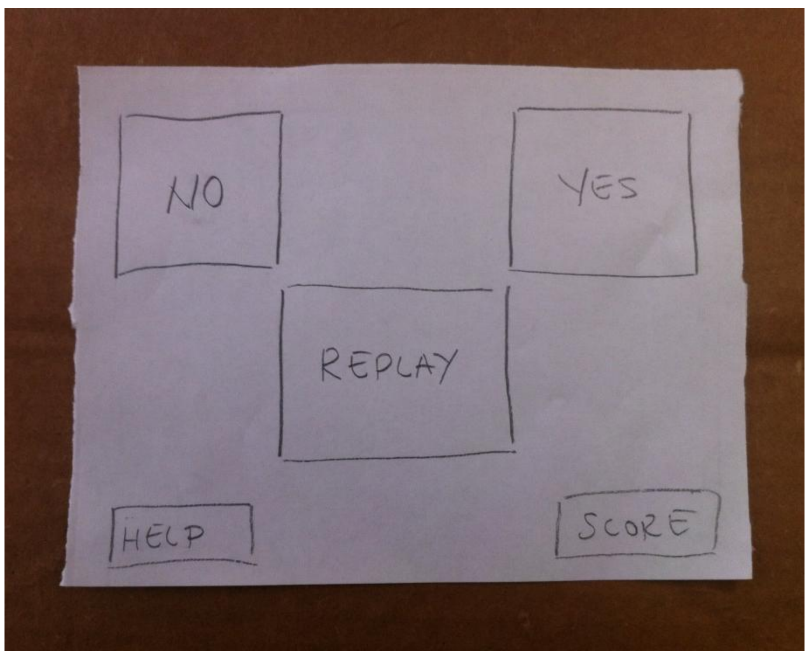

Figure 9.3. Paper prototype of AEDIN's quiz screen. The paper was pasted into a cardboard simulating the size of an actual desktop computer screen.

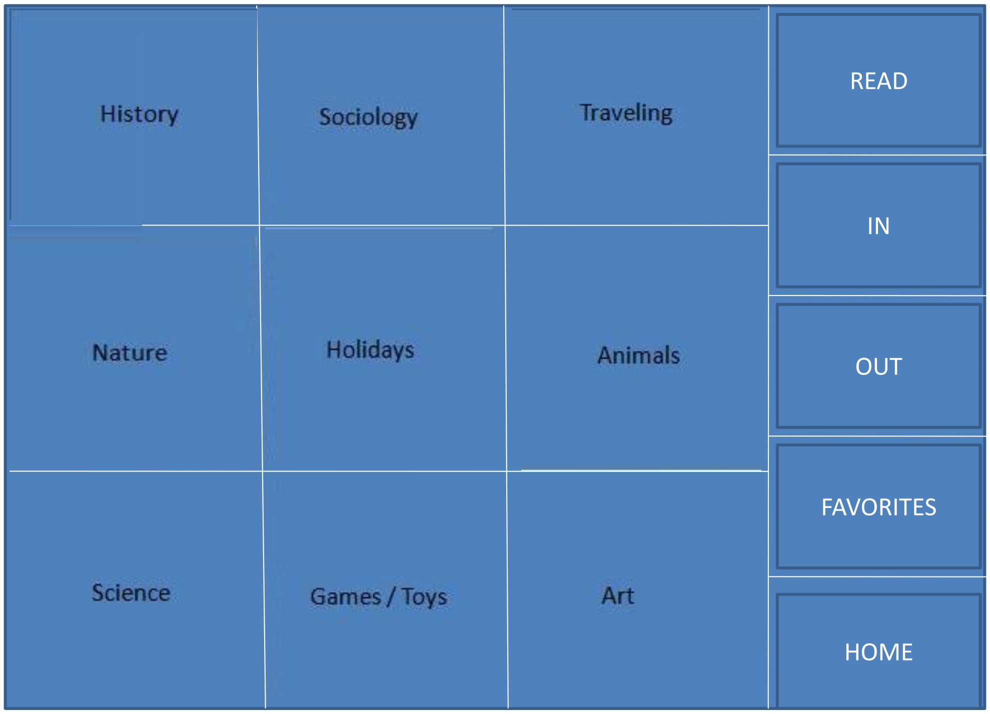

Figure 9.4. Early high-fidelity prototype of AEDIN. Main screen.

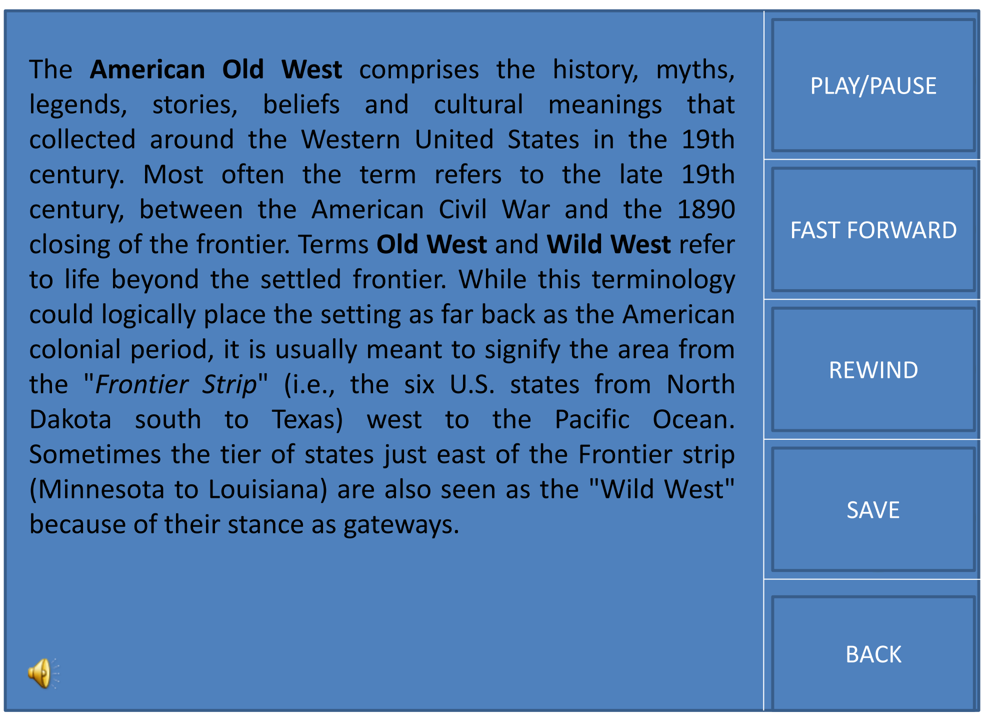

Figure 9.5. Early high-fidelity prototype of AEDIN. Listening to the TTS essays.

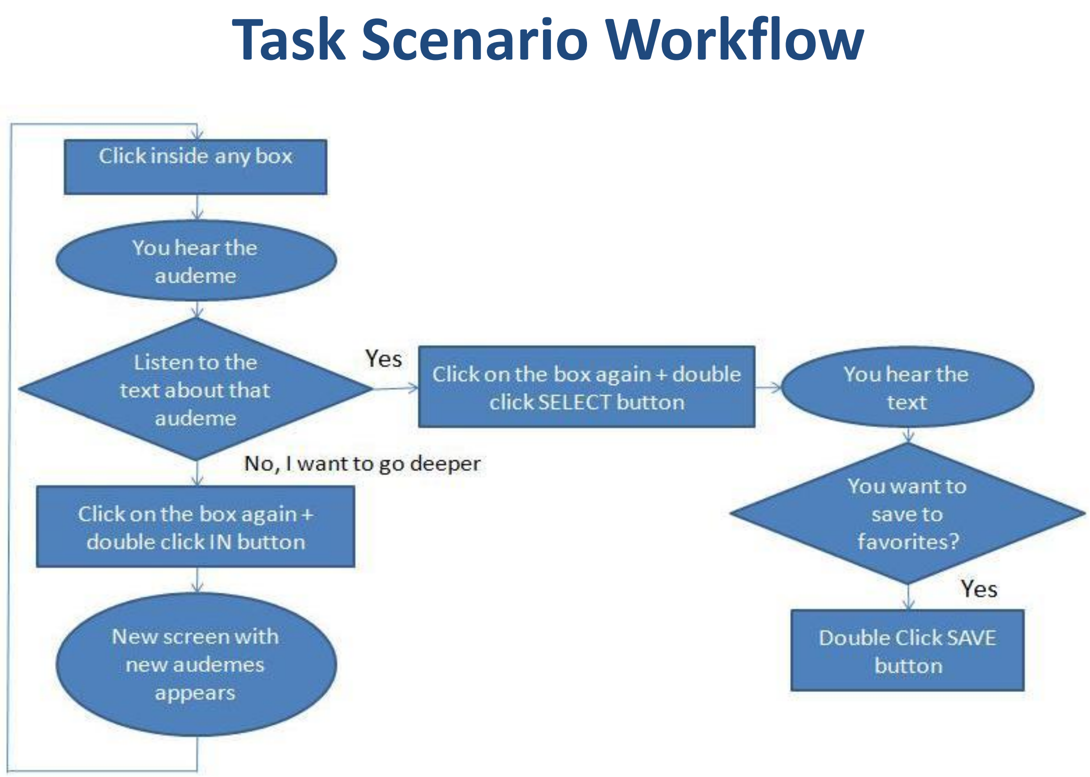

Figure 9.6. AEDIN's task scenario workflow.

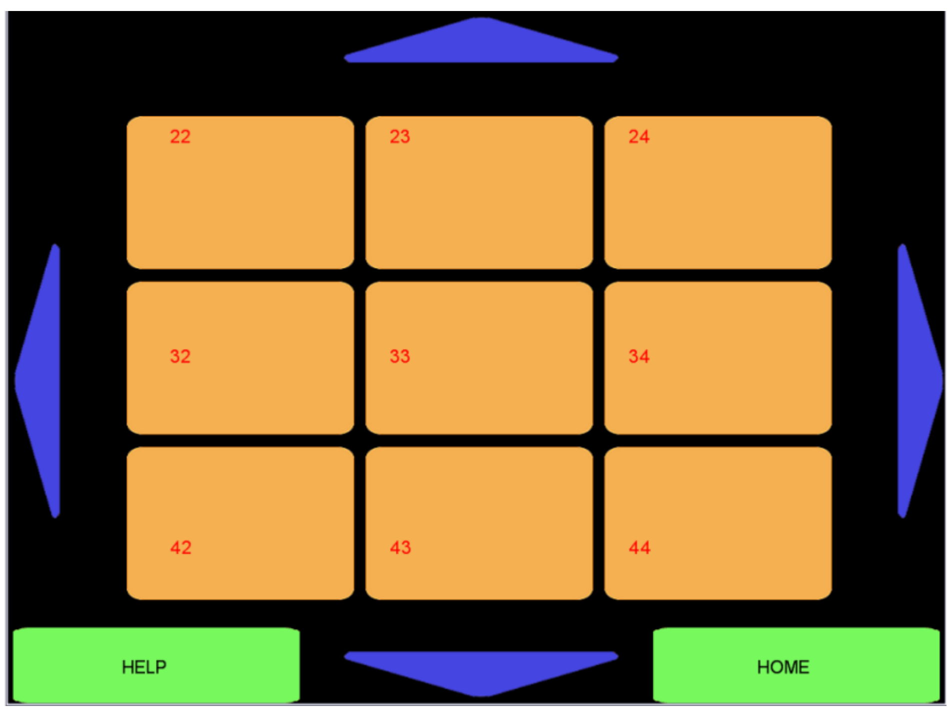

Figure 9.7. Later stage high-fidelity prototype of AEDIN main screen.

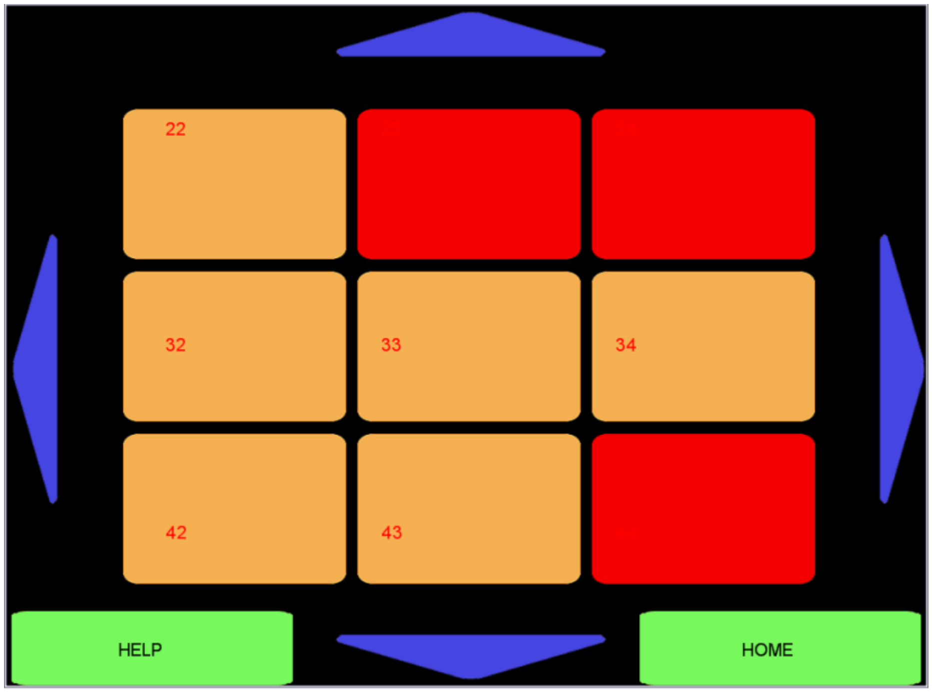

Figure 9.8. Later high-fidelity prototype of AEDIN main screen.

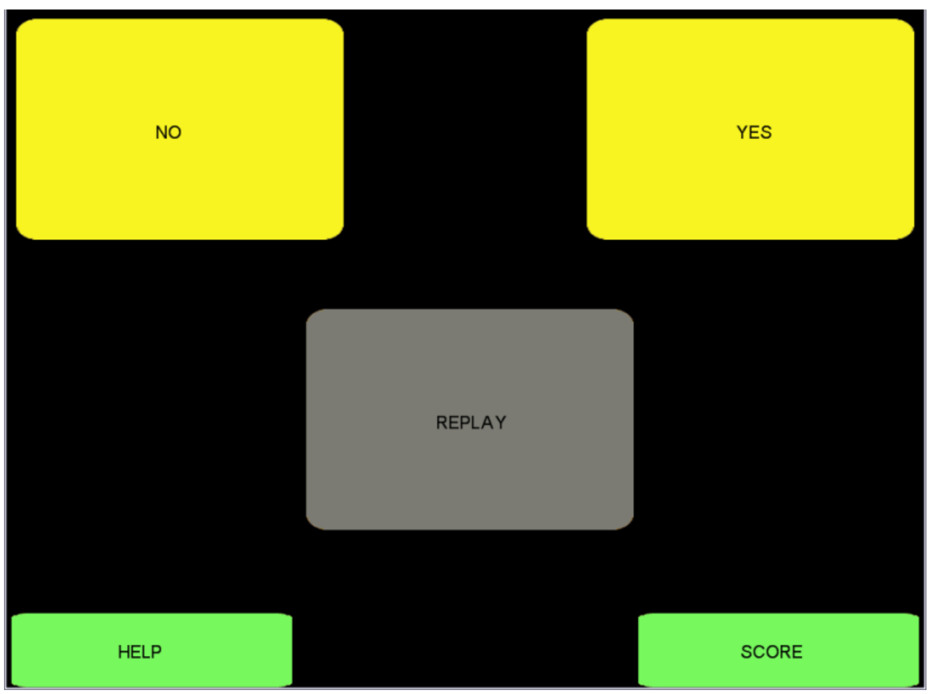

Figure 9.9. Later high-fidelity prototype of AEDIN quiz screen.

9.3.2 ASCOLTA

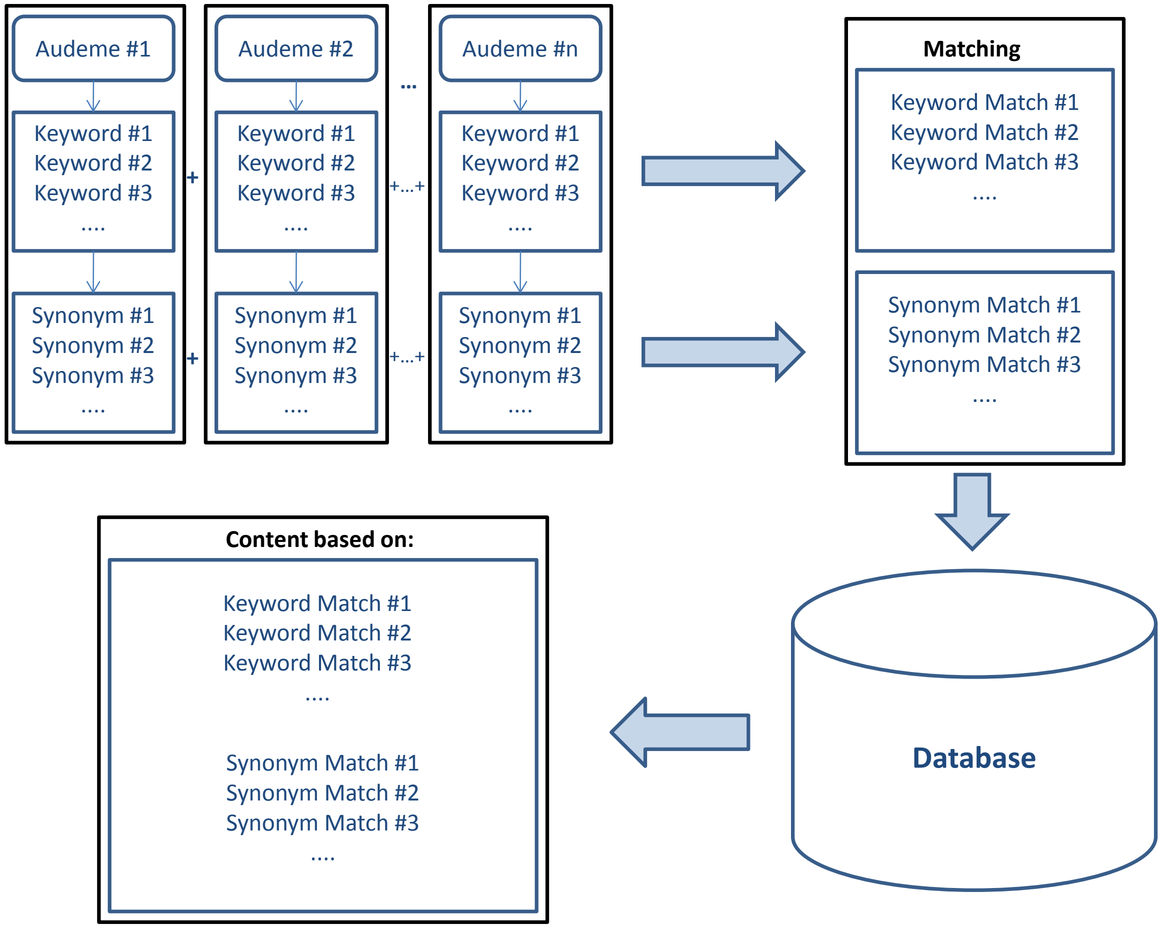

Figure 9.10. Early conceptual design of ASCOLTA.

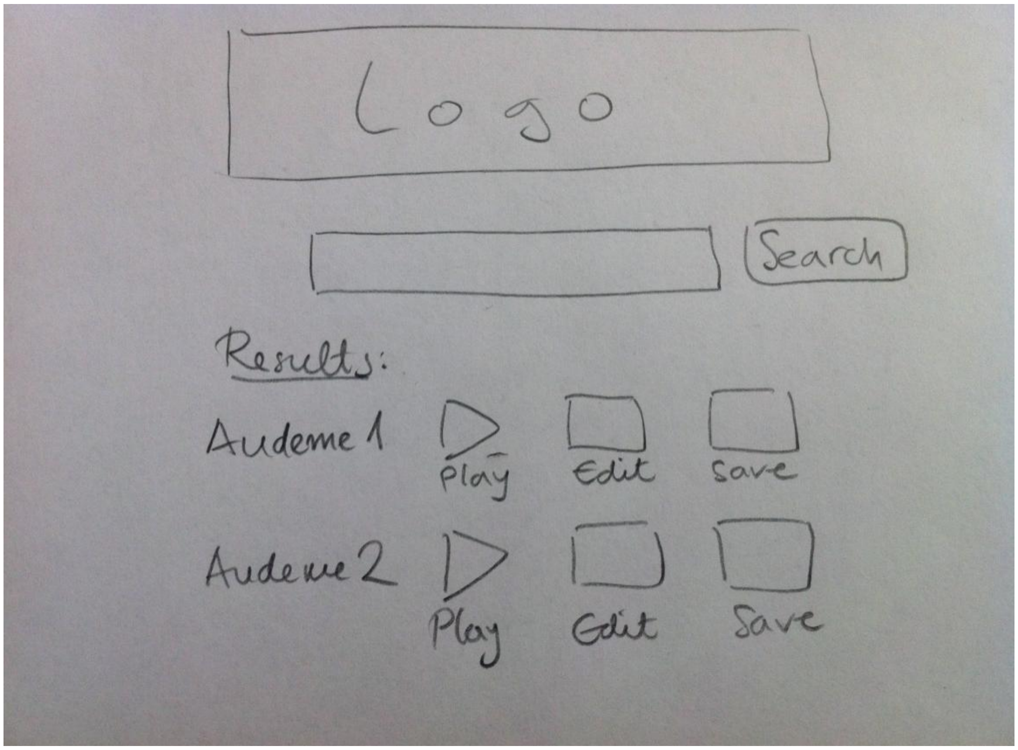

Figure 9.11. Early paper prototype of ASCOLTA. Audeme results screen.

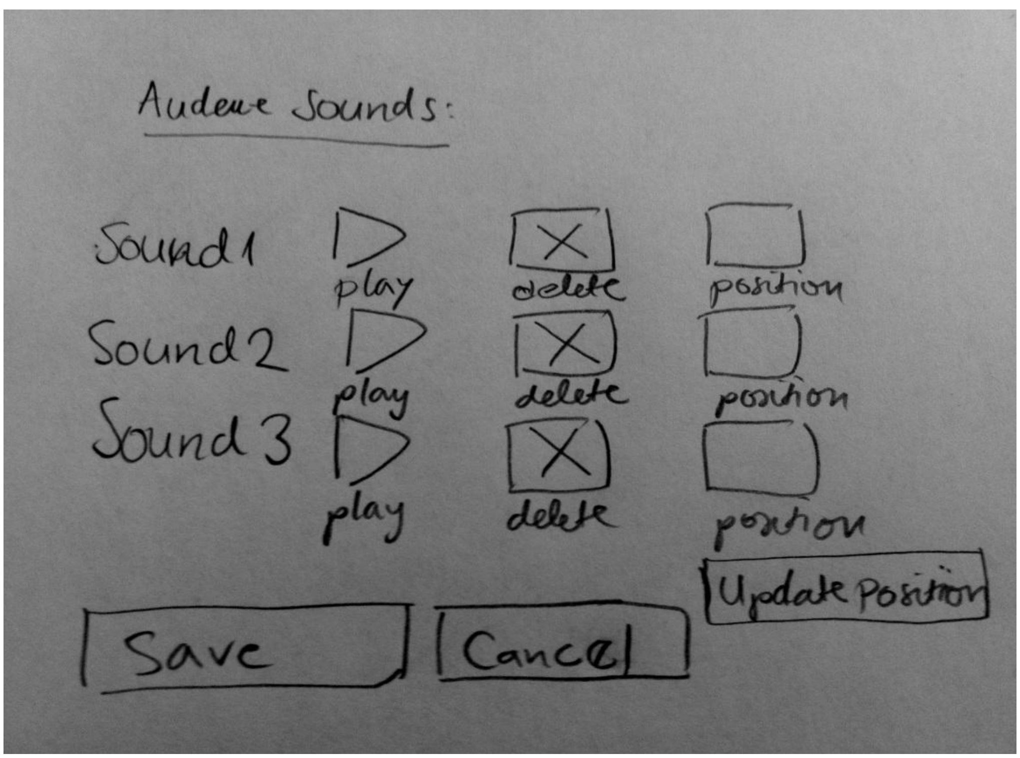

Figure 9.12. Early paper prototype of ASCOLTA. Sounds constituting an audeme screen.

REFERENCES

Absar, R., & Guastavino, C. (2008). Usability of non-speech sounds in user interfaces. In Proceedings of the International Conference on Auditory Display (ICAD'08), 1-8.

Adams, PF., Hendershot, GE., & Marano, MA. (1999). Current estimates from the National Health Interview Survey, 1996. Vital Health Stat 10; Oct: 1–203.

Asakawa, C., Takagi, H., Ino, S., & Ifukube, T. (2003). Maximum listening speeds for the blind. ICAD03, 276-279.

Back, M., & Des, D. (1996). Micro-narratives in sound design: Context, character, and caricature in waveform manipulation. ICAD96, IRCAM, Paris, France.

Baddeley, A. D. (1999). Essentials of human memory. Hove, UK: Psychology Press.

Balabolka. http://www.cross-plus-a.com/balabolka.htm Retrieved on December 2, 2011.

Ballas, J. (1994) Delivery of information through sound, in Auditory Display: Sonification, Audification and Auditory Interfaces, G. Kramer, ed. 79-94. Santa Fe Institute Studies in the Sciences of Complexity, Proc. Vol. XVIII. Reading, MA: Addison Wesley.

Barrett, L., Lane, R., Sechrest, L., & Schwartz, G. (2000). Sex differences in emotional awareness. Personality and Social Psychology Bulletin, 26(9), 1027-1035.

Bernard, R. M. (1990). Using extended captions to improve learning from instructional illustrations. British Journal of Educational Technology, 21, 215-225.

Blattner, M. M., & Glinert, E. P. (1996). Multimodal integration. Multimedia, IEEE, 3(4), 14-24.

Blattner, M., Sumikawa, D., & Greenberg, R. (1989). Earcons and icons: Their structure and common design principles. Human-Computer Interaction, 4(1), 11-44.

Bly, S. (1982). Sound and computer information presentation (UCRL-53282). Unpublished doctoral dissertation, Lawrence Livermore National Laboratory and University of California, Davis, CA.

Bolchini, D., Colazzo, S., Paolini, P., & Vitali, D. (2006) Designing Aural Information Architectures. In: Proc. ACM 24rd International Conference on Design of Communication (SIGDOC 2006), ACM Press, Myrtle Beach, 51–58.

Bonebright, T. L., Nees, M. A., Connerley, T.T., & McCain, G. R. (2001) Testing the Effectiveness of Sonified Graphs for Education: A Programmatic Research Project, In Proceedings of ICAD 2001, Espoo, Finland, 62-66.

Bonner, M., Brudvik, J., Abowd, G., & Edwards, W. (2010). No-Look Notes: Accessible eyes-free multi-touch text entry. Pervasive Computing, 409-426.

Bower, G. (1981). Mood and memory. American psychologist, 36(2), 129-148.

Bregman, A. (1994). Auditory scene analysis: The perceptual organization of sound: The MIT Press.

Brewster, S. A. (1998). The design of sonically-enhanced widgets. Interact. Comput. 11, 2, 211–235.

Brewster, S. (1994). Providing a structured method for integrating non-speech audio into human-computer interfaces. PhD Thesis.

Brewster, S., Wright, P., & Edwards, A. (1993). An evaluation of earcons for use in auditory human-computer interfaces. Proceedings of the INTERACT '93 and CHI '93 conference on Human factors in computing systems, 222-227.

Brewster, S., Wright, P., & Edwards, A., (1995a). Experimentally derived guidelines for the creation of earcons. Adjunct Proceedings of HCI 95, 155-159.

Brewster, S., Wright, P., & Edwards, A. (1995b). Parallel earcons: Reducing the length of audio messages. International Journal of Human Computer Studies, 43, 153-153.

Brewster, S. A., Raty, V.P., & Kortekangas, A. (1996). Earcons as a method of providing navigational cues in a menu hierarchy. In Proc. of the HCI'96. Springer, 167–183.

Brewster, S. A., Wright, P. C., & Edwards, A. D. N. (1994). A detailed investigation into the effectiveness of earcons. In Auditory Display (Ed, Kramer, G.) Addison-Wesley, Reading, MA, pp. 471-498.

Brewster, S. A. (1997). Using non-speech sound to overcome information overload. Displays, 17(3-4), 179-189.

Brewster,S.A., & Cryer, P. G. (1999). Maximising screen-space on mobile computing devices. In Pro-ceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'99). 224–225.

Brumby, D. P., Davies, S. C. E., Janssen, C. P., & Grace, J. J. (2011) Fast or safe? How Performance Objectives Determine Modality Output Choices while Interacting on the Move. Proc. CHI (2011), 473-482.

Brungart, D. S. (1998). Control of perceived distance in virtual audio displays. Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 1101-1104.

Ceri, S., Fraternali, P., & Bongio, A. (2000). Web Modeling Language (WebML): a modeling language for designing Web sites. Computer Networks, 33(1), 137-157.

Challis, B., & Edwards, A. (2000). Weasel: a computer based system for providing non-visual access to music notation. ACM SIGCAPH Computers and the Physically Handicapped (66), 1-12.

Chion, M., Gorbman, C., & Murch, W. (1994). Audio-vision: sound on screen: Columbia University Press.

Clark, J. M., & Paivio, A. (1991). Dual coding Theory and Education. Educational Psychology Review, Vol. 3, No. 3.

Clayton, A., Perry, E. I., Reed, L., & Winkler, A. M. (2000). America, Pathways to the Present, Upper Saddle River, NJ, Prentice Hall.

Conversy, S. (1998). Ad-hoc synthesis of auditory icons. ICAD'98.

Deatherage, B. H. (1972). Auditory and other sensory forms of information presentation. Human engineering guide to equipment design, 123-160.

Delogu, F., Palmiero, M., Federici, S., Plaisant, C., Zhao, H., & Belardinelli, O. (2010). Non-visual exploration of geographic maps: Does sonification help? Disability & Rehabilitation: Assistive Technology 5, Nr. 3: 164–174.

Dingler, T., Lindsay, J., & Walker, B. N. (2008). Learnability of sound cues for environmental features: auditory icons, earcons, spearcons, and speech. ICAD’08.

Dombois, F. (2001). Using audification in planetary seismology. Proceedings of the 7th International Conference on Auditory Display, Espoo, Finland, 227-230

Dorner, B., & Hagen, E. (1994). Towards an american sign language interface. Artificial Intelligence Review, 8(2), 235-253.

Doucet, M., Guillemot, J., Lassonde, M., Gagné, J., Leclerc, C., & Lepore, F. (2005). Blind subjects process auditory spectral cues more efficiently than sighted individuals. Experimental Brain Research, 160(2), 194-202.

Edwards, A. (1989). Soundtrack: An auditory interface for blind users. Human-Computer Interaction, 4(1), 45-66.

Ehrman, M., & Oxford, R. (1988). Effects of sex differences, career choice, and psychological type on adult language learning strategies. Modern Language Journal, 72(3), 253-265.

Eriksson, Y., & Gärdenfors, D. (2005). Computer games for children with visual impairments. International Journal on Disability and Human Development, 4(3), 161-168.

Fabiani, M., Dubus, G., & Bresin, R. (2010). Interactive sonification of emotionally expressive gestures by means of music performance. In Proc. ISon 2010 - Interactive Sonification Workshop.

Ferati, M., Mannheimer, S., & Bolchini, D. (2011). Usability Evaluation of Acoustic Interfaces for the Blind in Proc. of the 29th ACM International Conference on Design of Communication (SIGDOC), 9-16, October 3-5, Pisa, Italy.

Ferati, M., Mannheimer, S., & Bolchini, D. (2009). Acoustic interaction design through audemes: experiences with the blind. Proc. of the 27th ACM International Conference on Design of Communication (SIGDOC), Bloomington, IN, 23-28.

Ferati, M., Bolchini, D., & Mannheimer, S. (2009). Towards a Modeling Language for Designing Auditory Interfaces, in C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2009 Proceedings, Springer-Verlag LNCS 5616, 502–511, San Diego, CA, 2009.

Fernstrom, M., & Brazil, E. (2004). Human-computer interaction design based on interactive sonification hearing actions or instruments/agents. In Proceedings of the 2004 International Workshop on Interactive Sonification, (Bielefeld University, Germany), January 2004.

Fernstrom, M., & McNamara, C. (2005). After direct manipulation---direct sonification. ACM Transactions on Applied Perception (TAP), 2(4), 495-499.

Final Cut Pro. http://www.apple.com/finalcutpro Retrieved on February 17, 2012.

Fitch, W., & Kramer, G. (1994). Sonifying the body electric: Superiority of an auditory over a visual display in a complex, multivariate system. ICAD'98, 18, 307-307.

Foulke, E., Amster, C. H., Nolan, C. Y., & Bixler, R. H. (1962). The comprehension of rapid speech by the blind. Exceptional Children. 29, 134-141.

Frauenberger, C., Putz, V., Holdrich, R., & Stockman, T. (2005). Interaction patterns for auditory user interfaces. ICAD Proceedings, Limerick, Ireland, July, 6-9.

Frohlich, P., & Pucher, M. (2005) Combining Speech and Sound in the User Interface. ICAD 2005.

Gaver, W. (1986). Auditory icons: Using sound in computer interfaces. Human-Computer Interaction, 2(2), 167-177.

Gaver, W. (1989). The SonicFinder: An interface that uses auditory icons. Human-Computer Interaction, 4(1), 67-94.

Gaver, W., Smith, R., & O'Shea, T. (1991). Effective sounds in complex systems: The ARKola simulation. Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology. New Orleans, Louisiana, United States, 85-90.

Gerth, J. (1992). Performance based refinement of a synthetic auditory ambience: identifying and discriminating auditory sources. PhD Thesis.

Glenberg, A. M., & Langston, W. E. (1992). Comprehension of illustrated text: Pictures help to build mental models. Journal of Memory and Language, 31, 129-151.

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. Journal of personality and social psychology, 79(5), 701.

Growing Strong. http://bvi.growingstrong.org Retrieved on September 23, 2011.

Guri-Rozenblit, S. (1988). The interrelationship between diagrammatic representations and verbal explanations in learning from social science text. Instructional Science, 17, 219-234.

iBooks. http://www.apple.com/education/ibooks-textbooks/ Retrieved on March 3, 2012.

Janata, P., & Childs, E. (2004) Marketbuzz: Sonification of Real-Time Financial Data. In ICAD. Proceedings of International Conference on Auditory Display.

Jeon, M., & Walker, B. N. (2009). "Spindex": Accelerated initial speech sounds improve navigation performance in auditory menus. In Proceedings of the Annual Meeting of the Human Factors and Ergonomics Society (HFES09). 1081–1085

Kane, S. K., Bigham, J. P., & Wobbrock, J. O. (2008). Slide rule: making mobile touch screens accessible to blind people using multi-touch interaction techniques. Proc. ASSETS '08. New York: ACM Press, 73-80.

Kantowitz, B. H., & Sorkin, R. D. (1983). Human factors: Understanding people-system relationships: Wiley New York.

Klima, E., & Bellugi, U. (1979). The signs of language: Harvard University Press.

Kosslyn, S. M. (1994). Elements of graph design. New York: Freeman.

Kramer, G. (2000). Auditory Display: Sonification, audification and auditory interfaces: Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA.

Kramer, G., & Walker, B. (Eds.) (1999). Sonification Report: Status of the Field and Research Agenda. Santa Fe: The International Community for Auditory Display.

Kramer, G., Walker, B., Bonebright, T., Cook, P., Flowers, J. H., Miner, N., & Neuhoff, J. (2010). Sonification Report: Status of the Field and Research Agenda. Faculty Publications, Department of Psychology. Paper 444.

Laird, J., Wagener, J., Halal, M., & Szegda, M. (1982). Remembering what you feel: Effects of emotion on memory. Journal of personality and social psychology, 42(4), 646-657.

Leplatre, G., & Brewster, S. A. (2000). Designing non-speech sounds to support navigation in mobile phone menus. Proceedings of ICAD2000, Atlanta, USA.

Li, K.A., Baudisch, P., & Hinckley, K. (2008). Blindsight: eyes-free access to mobile phones. in CHI 2008, ACM, 1389-1398.

Lord, K. R., & Putrevu, S. (2006). Exploring the dimensionality of the need for cognition scale. Psychology and Marketing, 23(1), 11-34.

Lucas, P. (1994). An evaluation of the communicative ability of auditory icons and earcons. In Proceedings of the International Conference on Auditory Display (ICAD '94), 1994.

Lumsden, J., & Brewster, S. (2001). A Survey of Audio-related Knowledge Amongst Software Engineers Developing Human-Computer Interfaces. University of Glasgow, Department of Computing Science, Technical Report TR - 2001 - 97, September 2001.

Lunney, D., Morrison, R. C., Cetera, M. M., Hartness, R. V., Mills, R. T., Salt, A. D., & Sowerll, D. C. (1983). A Microcomputer-Based Laboratory Aid for Visually Impaired Students. Micro, IEEE, 3(4), 19-31.

Macedo-Rouet, M., Rouet, J. F., Epstein, I., & Fayard, P. (2003). Effects of online reading on popular science comprehension. Science Communication, 25(2), 99.

Mannheimer, S., Ferati, M., Huckleberry, D., & Palakal, M. (2009). Using Audemes as a Learning Medium for the Visually Impaired. Proceedings of HEALTHINF'09, Porto, Portugal, 175-180.

Mannheimer, S., Ferati, M., Bolchini, D., & Palakal, M. (2009). Educational Sound Symbols for the Visually Impaired. Universal Access in HCI, Part I, HCII 2009 Proceedings, Springer-Verlag LNCS 5614, San Diego, CA, 106-115.

Mansur, D. L., Blattner, M. M., & Joy, K. I. (1985). Sound graphs: A numerical data analysis method for the blind. Journal of Medical Systems, 9(3), 163-174.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. Educational Psychologist, 38, 43–52.

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. Journal of Educational Psychology, 90, 312–320.

McAdams, S., & Bigand, E. (1993). Thinking in sound: The cognitive psychology of human audition. Oxford University Press.

McGee-Lennon, M., Wolters M., & McBryan, T. (2007). Audio Reminders in the Home Environment. International Conference on Auditory Displays, Montreal, Canada.

McGee-Lennon, M.R., Wolters, M.K., McLachlan, R., Brewster, S., & Hall, C. (2011). "Name That Tune: Musicons as Reminders in the Home". In Proceedings of ACM CHI 2011, Vancouver, CA.

McGookin, D., Brewster, S., & Jiang, W. W. (2008). Investigating touchscreen accessibility for people with visual impairments. In Proceedings of NordiCHI, ACM, 2008, pp. 298-307.

Mezrich, J., Frysinger, S., & Slivjanovski, R. (1984). Dynamic representation of multivariate time series data. Journal of the American Statistical Association, 34-40.

Moos, A., & Trouvain, J. (2007). Comprehension of Ultra-Fast Speech–Blind vs.'Normally Hearing'Persons.

Morley, S., Petrie, H., & McNally, P. (1998). Auditory navigation in hyperspace: Design and evaluation of a non-visual hypermedia system for blind users. In Proceedings of the Annual ACM Conference on Assistive Technologies (ASSETS'98).

Mustonen, M. (2008). A review-based conceptual analysis of auditory signs and their design. ICAD Proceedings. IRCAM, Paris, France.

Mynatt, E. D. (1994). Designing with auditory icons: how well do we identify auditory cues?, in CHI '94: Conference companion on Human Factors in computing systems, (New York, NY, USA), pp. 269–270, ACM Press.

Oxford, R. (1993). Instructional Implications of Gender Differences in Second/Foreign Language (L2) Learning Styles and Strategies. Applied Language Learning, 4(1-2), 65-94.

Palladino, D. K., & Walker, B. N. (2007). Learning rates for auditory menus enhanced with spearcons versus earcons. In Proceedings of the International Conference on Auditory Display (ICAD'07), 274–279.

Palladino, D. K., & Walker, B. N. (2008). Efficiency of spearcon-enhanced navigation of one dimen-sional electronic menus. In Proceedings of the International Conference on Auditory Display (ICAD'08).

Parente, P., & Bishop, G. (2003). BATS: the blind audio tactile mapping system. ACMSE, Savannah, GA.

Park, S., Kim, S., Lee, S., & Yeo, W. S. (2010). Composition with path: musicla sonification of geo-referenced data with online map interface. In Proceedings of the International Computer Music Conference (ICMC), New York.

Peirce, C.S., "A Letter to Lady Welby" (1908), Semiotic and Significs, pp. 80-81.

Pirhonen, A., Murphy, E., McAllister, G., & Yu, W. (2007). Non-speech sounds as elements of a use scenario: a semiotic perspective. Proceedings of ICAD'06, 20-23.

Sánchez, J., & Flores, H. (2003). Memory enhancement through audio. ACM SIGACCESS Accessibility and Computing(77-78), 24-31.

Sánchez, J., Lumbreras, M., & Cernuzzi, L. (2001). Interactive virtual acoustic environments for blind children: computing, usability, and cognition. CHI '01 extended abstracts on Human factors in computing systems, 65 - 66.

Sánchez, J., & Flores, H. (2005). AudioMath: Blind children learning mathematics through audio. International Journal on Disability and Human Development, 4(4), 311.

Sanchez, J. H., Aguayo, F. A., & Hassler, T. M. (2007). Independent outdoor mobility for the blind. Virtual Rehabilitation 2007, September 27-29, Venice, Italy, pp. 114-120.

Sánchez, J., & Aguayo, F. (2006). Mobile Messenger for the Blind. In: Stephanidis, C., Pieper, M. (eds.) EP 2007. LNCS, vol. 4397, pp. 369–385. Springer, Heidelberg.

Sanders, M. S., & McCormick, E. J. (1987). Human factors in engineering and design: McGraw-Hill.

Schwabe, D., & Rossi, G. (1998). An object oriented approach to Web-based applications design. TAPOS, 4(4), 207-225.

Seidlitz, L., & Diener, E. (1998). Sex differences in the recall of affective experiences. Journal of personality and social psychology, 74(1), 262-271.

Shinn-Cunningham, B. G. (1998). Applications of virtual auditory displays. Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 1105-1108.

Siochos, V., & Eisenhauer, M. (2006). A user evaluation study of a multi-modal mobile navigation aid for pedestrians with respect to gender differences.

Smither, J. A. A. (1993). Short term memory demands in processing synthetic speech by old and young adults. Behaviour & Information Technology, 12(6), 330-335.

Sonnenschein, D. (2001). Sound design: the expressive power of music, voice, and sound effects in cinema: Michael Wiese Productions.

Stevens, R. D. (1996). Principles for the design of auditory interfaces to present complex information to blind people. University of York Department of Computer Science-Publications-YCST.

Stevens, R., Brewster, S., Wright, P., & Edwards, A. (1994) Providing an audio glance at algebra for blind readers. Proceedings of ICAD'94, 21-30.

Turnbull, D., Barrington, L., Torres, D., & Lanckriet, G. (2006). Modeling the Semantics of Sound. Department of Computer Science and Engineering, UCSD.

Tuuri, K., Mustonen, M., & Pirhonen, A. (2007). Same sound different meanings: A novel scheme for modes of listening. Proceedings of Audio Mostly, 13–18.

Tytler, R., Osborne, J., Williams, G., Tytler, K., & Clark, J. C. (2008). Opening up pathways: Engagement in STEM across the Primary-Secondary school transition. A review of the literature concerning supports and barriers to Science, Technology, Engineering and Mathematics engagement at primary–secondary transition. Canberra: Commissioned by the Australian Department of Education, Employment and Workplace Relations.

Waddill, P. J., McDaniel, M. A., & Einstein, G. O. (1988). Illustrations as adjuncts to prose: A test-appropriate processing approach. Journal of Educational Psychology, 80, 457-464.

Wagner, M., Newman, L., Cameto, R., Garza, N., & Levine, P. (2005). After high school: A first look at the postschool experiences of youth with disabilities. Menlo Park, CA: SRI International.

Walker, B., Nance, A., & Lindsay, J. (2006). Spearcons: Speech-based earcons improve navigation performance in auditory menus. Proceedings of ICAD'06, 63-68.

Walker, B. N., & Kramer, G. (2005). Mappings and metaphors in auditory displays: An experimental assessment. Proceedings of ICAD 96. International Community for Auditory Display.

Web Content Accessibility Guidelines. http://www.w3.org/TR/WCAG10/ Retrieved on November 21, 2011.

Wobbrock, J. O., Kane, S. K., Gajos, K. Z., Harada, S., & Froehlich, J. (2011). Ability-based design: Concept, principles and examples. ACM Transactions on Accessible Computing (TACCESS), 3(3), 9.

Won, S. Y. (2005). Auditory Display of Genome Data: Human Chromosome 21. Proc. of the International Conference on Auditory Display.

Yalla, P., & Walker, B. N. (2008). Advanced auditory menus: Design and evaluation of auditory scrollbars. In Proceedings of the Annual ACM Conference on Assistive Technologies (ASSETS'08). 105–112.

Yoshida T., Kitani K.M., Koike H., Belongie, S., & Schlei, K. (2011). EdgeSonic: image feature sonification for the visually impaired. Proceedings of AH. Tokyo, Japan.

CURRICULUM VITAE

Mexhid Adem Ferati

Education:

Indiana University, Indianapolis, IN, USA
PhD, Informatics – Human Computer Interaction (2007 – 2012)

Purdue University, Indianapolis, IN, USA
MSc, Computer Science (2005 – 2007)

South East European University, Republic of Macedonia
BSc, Computer Science (2002 – 2005)

Work Experience:

Non–academic:
Hardware/Software Technician, Kumanovo, Macedonia
"a3m Computers" (2004-2005)

Language Interpreter, Camp Bondsteel, Kosovo
U.S. Military Forces – KFOR (1999-2002)

Academic:
Research Assistant, School of Informatics, Indiana University, Indianapolis, IN, USA
Ph.D. Thesis Research (2007 – 2012)

Instructor, South East European University, Tetovo, Macedonia
CMCST04 – Human-Computer Interface (Spring 2012)

Co-Instructor, School of Informatics, Indiana University, Indianapolis, IN, USA
INFO I501– Introduction to Informatics (Fall 2011)

Guest Lecturer, School of Informatics, Indiana University, Indianapolis, IN, USA
INFO I624 – Advanced Seminar I in HCI (Fall 2011)

Guest Lecturer, School of Informatics, Indiana University, Indianapolis, IN, USA
NM 485/ INFO I590 – Psychology of Media (Fall 2011)

Teaching Assistant, School of Informatics, Indiana University, Indianapolis, IN, USA
INFO I590/I690 – Structured Conceptual Design for Interactive Applications (Summer 2011)

Teaching Assistant, School of Informatics, Indiana University, Indianapolis, IN, USA
INFO I624 – Advanced Seminar I in HCI (Fall 2010)

Guest Lecturer, School of Informatics, Indiana University, Indianapolis, IN, USA
INFO I624 – Advanced Seminar I in HCI (Fall 2010)

Co-Instructor, School of Informatics, Indiana University, Indianapolis, IN, USA
INFO I300– Human-Computer Interaction (Fall 2010)

English Language Tutor, South East European University, Macedonia
English Tutoring Center (2003-2004)

Presentations:

Speaker, Wright State University, Dayton, OH, USA
Brown Bag Lunch Talk Series, April 20, 2012.

Poster Presenter, Indianapolis, IN, USA
IUPUI Research Day, IUPUI Campus, April 13, 2012.

Speaker, West Lafayette, IN, USA
Independence Science: Learning a New Direction Conference for Disability (IsLAND'11), October 28-29, 2011.

Speaker, Pisa, Italy
29th ACM International Conference on Design of Communication (SIGDOC'11), October 3-5, 2011.

Speaker, Indianapolis, IN, USA
Usability Professionals' Association - Indiana Chapter, September 19, 2011.

Poster Presenter, Indianapolis, IN, USA
IUPUI Research Day, IUPUI Campus, April 8, 2011.

Presenter, Videoconference with Microsoft, Chicago, IL, USA
Imagine Cup 2011 Project Presentation, USER Lab, IUPUI, April 4, 2011.

Poster Presenter, Indianapolis, IN, USA
World Usability Day, IUPUI Campus, November 12, 2009.

Speaker, Bloomington, IN, USA
27th ACM International Conference on Design of Communication (SIGDOC'09), October 5-7, 2009.

Speaker, San Diego, CA, USA
13th International Conference on Human-Computer Interaction, July 19-24, 2009.

Poster Presenter, Indianapolis, IN, USA
World Usability Day, IUPUI Campus, November 13, 2008.

Reviewer:

ICWE, www.icwe2010.webengineering.org
10th International Conference on Web Engineering

ICWE, www.icwe2011.webengineering.org
11th International Conference on Web Engineering

Honors:

IUPUI Student Employee of the Year Nominee Certificate (2012)
Microsoft Imagine Cup 2011 Honorable Mention Certificate (2011)
Indiana University Scholarship (2007 – 2012)
USAID Scholarship (2005 – 2007)
British Council Scholarship (2004 – 2005)

Professional Membership:

ACM, www.acm.org
Association for Computing Machinery (2007 – present)

ICAD, www.icad.org
International Community for Auditory Display (2008 – present)

UPA, www.indiana-upa.org
Usability Professionals' Association – Indiana Chapter (2009 – present)

Publications:

Refereed:

Rohani Ghahari, R., Ferati, M., Yang, T., and Bolchini, D., Back Navigation Shortcuts for Screen-Reader Users. In Proceedings of ASSETS'12 *(submitted)*

Yang, T., Ferati, M., He, L., and Bolchini, D., Index and Guided Tour Navigation for Fact Finding. In Proceedings of SIGDOC'12 *(submitted)*

Ferati, M., Pfaff, M., Mannheimer, S., and Bolchini, D., Audemes at Work: Investigating Features of Non-speech Sounds to Maximize Content Recognition. International Journal of Human-Computer Studies *(submitted)*

Yang, T., Ferati, M., Liu, Y., Ghahari, R., and Bolchini, D., Aural Browsing On-The-Go: Listening-based Back Navigation in Large Web Architectures. CHI 2012, May 5–10, 2012, Austin, TX, USA.

Ferati, M., Mannheimer, S., Bolchini, D., Usability Evaluation of Acoustic Interfaces for the Blind. in Proc. of the 29th ACM International Conference on Design of Communication (SIGDOC), 9-16, October 3-5, Pisa, Italy, 2011.

Edwards, R.L., Stewart, J.K., Ferati, M., Assessing the Effectiveness of Distributed Pair Programming for an Online Informatics Curriculum. ACM Inroads, Vol.1, No.1, March 2010.

Ferati, M., Mannheimer, S., Bolchini, D., Acoustic Interaction Design through "Audemes": Experiences with the Blind. In Proc. of the 27th ACM International Conference on Design of Communication (SIGDOC), 23-28, Bloomington, IN, 2009.

Ferati, M., Bolchini, D., Mannheimer, S., Towards a Modeling Language for Designing Auditory Interfaces. In C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2009 Proceedings, Springer-Verlag LNCS 5616, 502–511, San Diego, CA, 2009.

Mannheimer, S., Ferati, M., Bolchini, D., Palakal, M., Educational Sound Symbols for the Visually Impaired. In C. Stephanidis (Ed.): Universal Access in HCI, Part I, HCII 2009 Proceedings, Springer-Verlag LNCS 5614, 106–115, San Diego, CA, 2009.

Mannheimer S., Ferati M., Huckleberry D., and Palakal M., Using Audemes as a Learning Medium for the Visually Impaired. In Proceedings of HEALTHINF 2009, 175-180, Porto, Portugal, 2009.

<u>Non-refereed (Exhibited Research Posters):</u>

Bolchini, D., Ferati, M., Rohani Ghahari, R., Yang, T., Navigating the Aural Web, poster presented at the 2012 IUPUI Research Day, Faculty and Community Showcase, IUPUI Campus Center, Indianapolis (IN), April 13, 2012.

Bolchini, D., Ferati, M., Liu, Y., Luebke, J., Rohani Ghahari, R., Tao Yang, How We May Navigate (the Aural Web), invited poster exhibit presented at the 2011 IUPUI Translating Research Into Practice (TRIP) Community Research Showcase, Indianapolis, IUPUI Campus Center, September 12, 2011.

Bolchini, D., Ferati, M., Liu, Y., Luebke, J., Rohani Ghahari, R., Tao Yang, How We May Navigate (the Aural Web), invited poster exhibit presented at the 2011 IUPUI Research Day, Community Research Showcase, Indianapolis, IUPUI Campus Center, April 8, 2011

Bolchini, D., Ferati, M., Rohani Ghahari, R., Liu, Y., Luebke, J., Yang, T., Navigating the Aural Web, poster presented at the World Usability Day in Indianapolis, IUPUI Campus Center, November 11, 2010.

Bolchini, D., Ferati, M., Rohani Ghahari, R., Liu, Y., Luebke, J., Yang, T., Navigating the Aural Web, poster presented at the Indiana TechPoint Innovation Summit, Indiana Convention Center, October 27, 2010.

Ferati, M., Bolchini, D., Mannheimer, S., Palakal, M. Acoustic Interaction Design for Blind and Visually Impaired Users poster presented at UPA World Usability Day 2009, IUPUI Campus, Indianapolis, November 12, 2009.

Ferati M., Bolchini D., Mannheimer S., Palakal M., Designing Aural Applications, poster presented at UPA World Usability Day 2008, IUPUI Campus, Indianapolis, November 13, 2008.

Computer Skills:
Languages: C#, XNA, Java, C/C++, HTML, PHP, ASP.NET, XML, CSS
Databases: MySQL, MS SQL
Statistics: SPSS, JMP
Miscellaneous: MATLAB, Photoshop, Axure