THE EFFECT OF ENHANCED NAVIGATIONAL AFFORDANCES ON COLLEGE STUDENTS' COMPREHENSION OF INFORMATIONAL AUDITORY TEXT, AND THE ROLE OF METACOGNITIVE AND MOTIVATIONAL FACTORS

A Dissertation

by

DAVID SEAN SWEENEY

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Ernest Goetz
Co-Chair of Committee,	Susan Pedersen
Committee Members,	Victor Willson
	Richard Furuta
Head of Department,	Shanna Hagan-Burke

August 2018

Major Subject: Educational Psychology

ABSTRACT

A proliferation of natural speech audio texts as well as improvements in synthetic text-to-speech technology have created new opportunities for learners. While many studies have examined factors affecting comprehension of print texts, few have examined factors affecting comprehension of audio texts and fewer still the effects of specific moderating variables. This study examines the effects of navigational affordance use on comprehension of informational audio texts. Factors of metacomprehension, including self-regulation and rehearsal, as well as motivational factors of interest, effort regulation, and test anxiety were studied for their relationship to the use of navigational affordances. The study utilized a mobile application distributed through the iTunes® store to administer the experimental procedure. Students enrolled in an introductory political science course at a large state university were solicited to participate. Participants were randomly assigned to either the experimental or control group. The experimental group (N = 74) had access to enhanced navigational affordances including pause and continue, forward by sentence, forward by paragraph, backward by sentence, and backward by paragraph. The control group (N = 11) only had access to pause and continue affordances. Results indicate that the presence of enhanced navigational affordances did not demonstrate a significant difference in comprehension between the experimental and control groups. However, there was a significant correlation between navigational affordance use and comprehension. The data indicate the relationship may be curvilinear meaning that affordance use is more frequent for learners with average comprehension, and less frequent for high and low comprehension learners. Metacomprehension and motivational factors were not significantly correlated with navigational affordance use. Motivational factors did positively correlate with comprehension

for both groups with an $F = 5.49$ and $\alpha = 0.002$. Beta weights for the three factors were 0.29 for interest, -0.35 for test anxiety, and 0.003 for motivation. Information on distractions during the study were also collected. Some participants demonstrated a pattern of skipping behavior when using navigational affordances in which they would quickly navigate through the audio text. The study platform could be used to administer other kinds of audio text comprehension experiments.

DEDICATION

I dedicate this work to my mother, Doris Mae Sweeney, who always believed that I would one day succeed. I wish she had lived to see it.

ACKNOWLEDGEMENTS

I would like to thank my committee chairs, Drs. Goetz and Pedersen for having infinite patience and faith in me. To Dr. Goetz, thank you for giving me a sense of joy in asking and answering the tough questions. To Dr. Pedersen, thank you for stepping up when I lost my first chair and keeping my feet on the ground. I would also like to acknowledge the support of my other committee members, Drs. Willson and Furuta. Dr Willson, you were always ready to give me advice and help me navigate the message. Dr. Furuta, you pointed me in the right direction when I was unable to find any literature that supported my topic. All of you had faith in me even when I didn't have faith in myself. You stuck with me through thick and thin. You are truly The Dream Team.

I would also like to thank my youngest son, Luke. Without his exceptional "computer hacking skills" in assessing the large set of navigational data detailed in this study, I would not have been able to finish. Thank you, Luke.

Finally, I'd like to thank my wife, Allyson, for making many sacrifices over many years to allow me to pursue this dream. Thank you, Ally.

CONTRIBUTORS AND FUNDING SOURCES

This work was supported by a dissertation committee consisting of Professors Emeritus Ernest Goetz [co-chair] and Victor Willson, Associate Professor Susan Pedersen [co-chair], all from the Department of Educational Psychology, and Professor Richard Furuta from the Department of Computer Science and Engineering.

Some of the data analyzed for Chapter IV were provided by Luke Sweeney. Figure 1 is based upon a figure from Douglas J. Hacker (1998). Instrument scale questions from the Motivated Strategies for Learning Questionnaire (Pintrich, 1991) are reprinted in Appendix C. Dr. Victor Willson was instrumental in the analyses depicted in Chapter IV. All other work conducted for the dissertation was completed by the student independently, with no outside financial support.

NOMENCLATURE

ANOVA	Analysis of Variance
App	Application, usually within the context of a mobile device application
BVI	Blind or Visually Impaired
ER	Effort Regulation (i.e., Motivation)
GUID	Globally Unique Identifier
I	Interest
iOS	Apple iPhone Operating System
JOL	Judgment of Learning
MANOVA	Multivariate Analysis of Variance
MSLQ	Motivated Strategies for Learning Questionnaire
POLS	Political Science
PTEW	Phonological Type Errors of Words
PTES	Phonological Type Errors of Sentences
RE	Rehearsal
RQ	Research Question
SR	Self-Regulation
SRL	Self-Regulated Learning
SSML	Speech Synthesis Markup Language
TA	Test Anxiety
TBI	Traumatic Brain Injury
TTS	Text-To-Speech

QR Code	Quick Response Code
XML	Extensible Markup Language

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

Multiple modalities of instruction are becoming increasingly available in today's learning landscape with an ever-widening variety of instructional media, from digital versions of traditional paper books, to instructor customized digital collections, to audio texts, to multimedia texts. A proliferation of natural speech audio texts as well as improvements in synthetic text-to-speech technology have created a rich information environment with easily accessible audio media. This situation has created expanded opportunities for learners who wish to use audio text, both those who choose to use them, such as students and audio book consumers, and especially those who cannot access printed text, such as visually impaired learners or those with reading disabilities. For these populations, audio text provides an essential medium for learning.

However, with increased availability comes increased scrutiny. Is learning from audio text as good as learning from printed text? Do audio text delivery systems provide effective learner/text controls and interactions? If so, do these controls (i.e., affordances) facilitate or hinder learning? Gibson (1977) defines affordances as environmental properties that affect the learner-environment interaction. In this sense, properties of the text or interface that the learner perceives and acts upon are affordances. Many researchers have studied affordances as they affect comprehension within the context of printed text, but few have studied the comprehension effects of affordances in the audio domain. Fewer have studied the effects of metacognitive and motivational factors on comprehension in the audio domain.

Among text affordances is navigation. Navigating print text is very different from navigating audio text. Print text, by nature both spatial and linear, affords navigation more easily and conveniently when compared to navigation of audio text, which by nature is only linear. As Paap and Cooke (1997) stated, "Audition is linear and vision is not… [Listeners] don't actively search for a target, rather they must wait and try to ambush it" (p. 567). This makes navigation within the audio text domain an issue. For example, if a learner wants to refer back (i.e., look back) to information in printed text, the eye can easily and efficiently track back to information that is within the visual field. Accomplishing the same feat in the audio domain is more complicated. Arons (1993) states, "…there is no natural way for humans to skim speech information because of the transient character of audio – the ear cannot skim in the temporal domain the way the eyes can browse in the spatial domain" (p. 187). Since audio text is linear, learners must use their temporal ability to first estimate the location of the needed audio text (i.e., how far back in time is the desired information) and then use a second ability such as manual dexterity or voice command to manipulate a navigational control. Besides the obvious increase in complexity, current navigational and orientation affordances in audio texts could be described as crude and inaccurate when compared to equivalent affordances inherent in printed text. This may be changing with the advent of more fine-tuned temporal affordances such as word, sentence, paragraph, and chapter navigation controls, audio location markers (e.g., audio icons, earcons, etc.), the use of variations in pitch and volume to connote meaning (i.e., prosody), and intelligent audio interfaces that can extract summary information from audio texts (i.e., summarization affordances) for orientation. Even as these technology affordances are being rapidly developed, they are outstripping the ability of research to keep up with them. As Wood and her colleagues recently stated (Wood, Moxley, Tighe, & Wagner, 2018), there is a lack of randomized control studies related to the effects of audio text on comprehension.

While navigational affordances are a topic of research interest, what affects the use of affordances? Hacker's model (1998) of self-regulated learning hypothesizes that comprehension is monitored at the metacognitive level. If a failure in comprehension is perceived, a control strategy may be used to correct the failure. Use of affordances is one type of control strategy (Zabrucky & Moore, 1994). Specifically, audio navigational affordances such as pausing, rereading, and selectively searching a text are affordances that can be used as a control strategy (Pressley & Afflerbach, 1995).This would suggest that metacognitive abilities may affect the use of affordances. Motivational factors have also been shown to positively affect text comprehension (e.g., Schiefele, 1988). Specifically, interest, effort regulation, and test anxiety have support in the literature as factors that may affect learning (cf. Pintrich & Smith, 1993). But do motivational factors affect the use of affordances?

Problem

Even with the growth in audio text domain affordances, there is a dearth of research on how they affect the comprehension of learners. Many early studies focused the effects of audio text reading rate on comprehension (e.g., Foulke, 1962; Foulke, 1964; Foulke, 1966). Some studies have focused on specific audio affordances and their effects on comprehension. But almost all of these studies examine navigation *of the interface*, for example, web menu navigation, rather than navigation of the instructional text (e.g., Walker et al., 2013). Several recent meta-analyses (e.g., Buzick & Stone, 2014; Li, 2014) have compiled effect sizes of audio text vs. written text on comprehension for both learners with and without disabilities. A recent large meta-analysis of the effects of audio text on comprehension (Wood et al., 2018) concluded that the effect size of between-subjects studies of text-to-speech on comprehension was $d = 0.48$ ($n = 2942$, $\rho < 0.05$). The effect size for all studies included was $d = 0.35$ ($n = 2942$, $\rho < 0.01$). Li concluded that the effect size was larger for populations of learners with disabilities. This supports a moderate positive effect of audio text on comprehension when compared to print text. But as Wood and her colleagues point out, more studies are needed to explore the moderating variables (e.g., specific affordances) of audio text for improving comprehension.

Purpose

The purpose of this study is to examine the use and effects of linear navigational affordances on comprehension within the audio text domain, and to examine the role of motivational factors and metacognitive strategies in the use of navigational affordances. These five affordances include: pause and continue, navigate backward by sentence and paragraph, and navigate forward by sentence and paragraph. For the purposes of this study, enhanced navigational affordances is defined as including all five navigational affordances specified, while basic navigational affordances include only the pause and continue option. All these affordances have been demonstrated within the interface design literature, but none have systematically been examined for their effects on comprehension.

Research Questions

Navigational affordances and comprehension

Do enhanced navigational affordances improve participants' comprehension of audio text? Examining this research question in more detail, two aspects become apparent. Does the *presence* of enhanced affordances improve participants' comprehension of audio text? To answer this question, the presence of enhanced affordances will be correlated with higher levels of comprehension. For the purposes of this study, this question will be labeled RQ1-A.

A second question also presents itself. Does the *use* of enhanced affordances improve participants' comprehension of audio text? To answer this question, increased affordance use will be correlated with higher levels of comprehension. For the purposes of this study, this question will be labeled RQ1-B.

Monitoring/control strategies and affordance use

What role do monitoring and control strategies play in the use of navigational affordances in audio text? If a participant metacognitively monitors learning, and a discrepancy is found between the goal and perceived comprehension, a control strategy may be used. Use of navigational affordances is one type of control strategy. To answer this question, the hypothesis is that higher levels of metacognitive monitoring and control will result in greater use of navigational affordances. In order to measure the participant's ability to monitor and control the learning process, an instrument that measures metacognitive self-regulation should be used. Specific cognitive control strategies such as rehearsal align well with navigational affordances within the audio domain. For the purposes of this study, this question will be labeled RQ2.

Motivational factors and affordance use

What role do motivational factors play in the use of navigational affordances in audio text? The literature identifies a number of intrinsic motivational factors that can affect the self-regulated learning skill of the participant including interest, effort regulation, and test anxiety. To answer this question, the hypothesis is that increased motivation will result in greater use of navigational affordances. For the purposes of this study, this question will be labeled RQ3.

CHAPTER II

REVIEW OF THE LITERATURE

Gibson (1979) defined affordances as environmental properties that influence the learner-environment interaction. In this sense, properties of the text or text interface that the learner perceives and acts upon are affordances. However, Norman (1988) points out that Gibson's definition only applies to affordances that are actionable by the actor (e.g., learner). What if an affordance is not perceptible? Gaver (1991) further refines the delineation of affordances into four distinct groups based upon two dimensions - the existence of the affordance, and the actor's perception of the affordance. Gaver identified four groups of affordances: false affordances (perceived but non-existent), perceptible affordances (perceived and present), hidden affordances (imperceptible but present), and correctly rejected affordances (non-existent). This point becomes important when text affordances are described for populations of learners who may not be able to perceive the affordance such as those with perceptual disabilities. Navigational affordances of print texts may be perceptible affordances to sighted learners but are always hidden affordances to blind and visually impaired (BVI) learners. Navigational affordances of audio texts may be perceptible affordances to both sighted and BVI learners. Thus, audio texts are, in a certain sense, more accessible to a wider population of learners. Consequently, I will define affordances using Gaver's perceptible affordance classification.

In order to understand the effects of navigational affordances on comprehension, a theory of comprehension must be identified that can explain why and how such affordances would affect learning. Since the focus of this study is audio texts, the importance of decoding in such a theory is diminished since decoding only applies to printed texts. A good candidate theory would seek to explain under what circumstances comprehension fails, how a learner recognizes comprehension failures, and what strategies may be employed to mitigate such failures. One such theory is the construction-integration model of comprehension.

The Construction-Integration Model of Comprehension

Kintsch (1998) first described the construction-integration model of comprehension. In this model, learners must comprehend the text at three different levels for effective comprehension to occur. The surface-level presentation of the text is the explicit linguistic level of comprehension and is the first hurdle for a learner to overcome. This refers to implicit and explicit attributes of the text; the fonts, colors, text size, etc., but also the complexity of the words and structures. Within the audio domain, it may also refer to analogs such as speech characteristics and dialect, gender, volume, inflection, timing, and other audio properties. The second level of the model is the text-base representation. This is the semantic meaning of the text as the author intended it. For example, anaphors, words and phrases that refer back to an earlier subject, exist at the text-base level. Paronyms, or words that are acoustically similar, also exist at this level. The last level in the construction-integration model is the situation level or mental model representation. This is the sense and meaning that the learner constructs from the text based upon his or her knowledge and experience. Comprehension at the situation level is the most integrated and meaningful.

Comprehension failures can occur at any level of the model. Kintsch points out that failures at the lower levels prevent comprehension at the higher levels. That is, a learner must successfully comprehend the text at the surface level before comprehending at the text-base level, and so forth. Disruptions such as unfamiliar words, complex sentence construction, anaphoric ambiguity, phonemic intelligibility, and lack of prior background knowledge can all cause disruptions. When a disruption occurs, the learner must recognize the disruption in order to take action to mitigate it. This failure in recognition is a mechanism of metacomprehension and will be treated in the next section.

Dunlosky and Rawson (2005) elaborate on comprehension failures by asserting that disruptions in text processing at the situation level are highly predictive of comprehension; "…processing disruptions will not only influence metacomprehension judgments but will also be predictive of performance on tests of comprehension" (p. 39). This is termed the levels-of-disruption hypothesis. In it, they assert that conditions supporting situation level processing will improve metacognitive awareness which, in turn, supports the use of metacognitive strategies when comprehension fails. What conditions support situation level processing? One possible answer is anything that reduces disruptions at lower levels, specifically, something that reduces failures of comprehension at the surface and text-base levels of comprehension. This poses an interesting question. If affordances of audio text can reduce comprehension disruptions at the lower levels, will comprehension at the higher levels be improved?

For example, perhaps we have a learner with a reading disability who needs to comprehend a text. Given her disability, the learner may have trouble comprehending printed text at the lower surface and text-base levels. Consequently, she is not given an opportunity to comprehend at the situation level because there was a failure at the lower levels. To mitigate this situation, audio text material is a common accommodation. Theoretically, since decoding is not necessary, audio text removes lower level disruptions caused by the reading disability and allows the student to process at a higher level more efficiently, thus increasing comprehension. Other affordances exist, including navigational affordances. Could navigational affordances reduce disruptions and increase comprehension? Should you discuss what this theory means for BVI learners? For example, you might look at what types of failures at different levels might be relevant to this group but not to sighted learners.

Audio Comprehension

Much research on audio comprehension was done in the 1960s, 70s, and 80s, mainly in the blind and visually impaired (BVI) community. Many of the studies focused on the reading rate of blind subjects. Not surprisingly, BVI learners who use audio texts on a regular basis develop greatly accelerated reading rates when using this medium. One of the landmark researchers in the area of audio reading rate (i.e., compressed speech) was Emerson Foulke. A number of early works (Foulke, 1962; Foulke, 1966; Foulke & Louisville Univ., 1967; Foulke, Amster, Nolan, & Bixler, 1962; Foulke, 1964; Foulke & Louisville Univ, KY Perceptual,Alternatives Lab, 1969; Foulke, Robinson, & Louisville Univ, KY Perceptual,Alternatives Lab, 1970; Foulke & Louisville Univ, KY Perceptual,Alternatives Lab, 1971) contain an array of experiments on comprehension of compressed speech texts. However, the focus of these studies is almost entirely on measuring the effects of audio reading rate on subjects' comprehension rather than on the design of audio text interfaces to enhance comprehension, navigational affordances, or the mechanisms of comprehension itself. This is understandable since the scope of audio technology at the time was limited to records and cassette tapes.

Several more recent studies have tangentially examined the effects of audio text affordances on comprehension, but generally focus on interface design efficacy rather than potential effects of the design on comprehension. For example, Davis and Hirschberg (1988) describe how "appropriate intonational features" can be represented in audio text by using prosodic affordances to give driving directions. They did not directly examine comprehension effects of these affordances. A number of other researchers have examined text prosody as an affordance (V. S. Argyropoulos, Sideridis, Kouroupetroglou, & Xydas, 2009; Xydas, Argyropoulos, Karakosta, & Kouroupetroglou, 2005). But although prosodic affordances in the audio domain affect usability of the text, they are not the interest of this study. Rather navigational affordances and their effects on comprehension are.

Rohani Ghahari and his colleagues (2012) explored back-navigation within the audio domain based upon topic and list levels with a small population of blind learners. They found that navigating by topic or list decreased the time it took to find relevant information and decreased perceived cognitive effort. Thus, they studied the efficiency of navigational affordances. Comprehension of the material was not tested. In general, most studies have focused on prosodic and organizational affordances rather than navigational affordances, and few studies have examined the relationship between navigational affordances in the audio domain and comprehension. This creates an opportunity to explore the effects of navigational affordances on audio comprehension.

Self-Regulated Learning

While improvements in audio affordances, whether interface or content-based, are extrinsic to the learner, intrinsic factors can also influence comprehension. One area in which there is a large body of research is self-regulated learning (SRL). According to Pintrich (2000), SRL is defined as "…an active, constructive process whereby learners set goals for their learning and then attempt to monitor, regulate, and control their cognition, motivation, and behavior, guided and constrained by their goals and the contextual features in the environment" (p. 453).

Pintrich goes on to describe SRL processes within four distinct phases; planning and goal setting, monitoring and metacognitive awareness, control and regulation, and reaction and/or reflection. These phases can further be delineated into four areas of regulation. The first three areas coincide with the common areas of psychological functioning; cognition, affect/motivation, and behavior (p. 455). The fourth, external context, examines how SRL attempts to monitor and control the external environment of learner. Of greatest interest to the current study are cognitive regulations within the phases of monitoring and control/regulation. This is because the use of navigational affordances is most closely associated with control/regulation within the audio domain.

Cognitive monitoring

Cognitive monitoring describes various processes of cognition which have traditionally been labeled as metacognition. Metacognition, or thinking about thinking, was first studied by John Flavell (1976) who stated, “Metacognition refers to one’s knowledge concerning one’s own cognitive processes and products or anything related to them” (p.232). Hacker (1998) describes metacomprehension as a person’s assessment of her own comprehension of text and is a central component of self-regulated comprehension. Kintsch and van Dijk (1978) explain that continuous evaluation and regulation of one’s understanding during reading through the use of strategies including rereading and searching previously read information is important for comprehension. If readers fail in their monitoring of comprehension, they will not devote the resources necessary to resolve failures in comprehension (Zabrucky & Moore, 1994).

Closely related to cognitive monitoring is the construct of judgements of learning, or JOLs (Connor, Dunlosky, & Hertzog, 1997). Pintrich (2000) describes one example of JOLs as an individual “becoming aware that they do not understand something they have just read or heard, or becoming aware that they are reading too quickly or slowly given the text and their goals” (p. 459). Implicit within the definition of JOLs is one or more goals or criteria against which comprehension is judged. Thus, comprehension goals become a strategy within the model of comprehension that allows the learner to take action (cognitive control) when a problem arises. Within this context, JOLs are a type comprehension monitoring that may be compared against a learning goal. For example, a learner sets a goal for a study task such as, “I want to understand this material well enough to make at least a ‘B’ in this course.” During study, the learner engages in JOLs to see if the goal is being met (monitoring). If the there is a perceived deficit in meeting the goal, the learner may engage in a control strategy such as pausing, slowing down, rereading, etc., to correct the deficit.

Cognitive control

When monitoring identifies a comprehension failure, the learner must take action to control or regulate the error. Pressley and Afflerbach (1995) catalog dozens of strategies to control comprehension failures. Among them are pausing during reading, rereading material, and selectively searching text to clarify meaning and increase comprehension. Although these navigational strategies can be accomplished in the print text domain with no affordances other than the eye and the printed page, a navigational affordance is required in the audio text domain. This is an important distinction because while the learner may make the cognitive decision to engage in a control strategy, action must be taken in the learning environment to affect the strategy in the audio domain. Thus, navigational affordances in the environment become a method of cognitive control, and consequently become a method for measuring a control strategy that is not as easy to measure in the print text domain.

Using Pintrich’s framework, one can imagine a learner setting a comprehension goal for a reading or listening task (planning), monitoring the state of their comprehension relative to his or her goal (cognitive monitoring), recognizing a discrepancy between his or her goal and actual state (i.e., JOL), and using heuristic strategies to correct deficits (regulation strategy). Within the context of printed text, these strategies are mainly internal (e.g., pausing, slowing down, rereading). But within the context of audio text, these same strategies require using external, contextual features in the environment. An affordance must be used because of the transient nature of audio, and these affordances may be measured. By following this process, one can see how the SRL framework can be applied to the comprehension of audio text.

Self-regulated comprehension

By describing the relationship between cognitive monitoring and control, a model for self-regulated comprehension begins to emerge (see Figure 1). Self-regulated comprehension is a multidimensional process consisting of evaluation and regulation (Hacker, 1998). Evaluation is the process of monitoring one's comprehension of text while regulation is controlling and adjusting the reading process to resolve problems of comprehension (i.e., disruptions in comprehension). Hacker states that looking back to prior text is an example of a monitoring process, but I respectfully disagree since the process of looking back occurs when a passage is not understood or integrated. It is a failure in comprehension, and the recognition of the failure (monitoring) that triggers the process of looking back. Thus, looking back is a metacognitive strategy to regulate or control comprehension (Alessi, Anderson, & Goetz, 1979). In other words, looking back is a process resulting from the comprehension failure detected by comprehension monitoring. This model supports the hypothesis that looking back is a metacognitive strategy to regulate disruptions in comprehension. The argument can be extended to the audio text domain if audio navigational affordances are analogous to pausing during reading, rereading material, and selectively searching text to clarify meaning and increase comprehension.

Thus, several characteristics of Hacker's model make it a good fit for describing the effects of metacognitive and motivational factors on comprehension in the audio text domain. First, Hacker's model de-emphasizes decoding, or rather treats comprehension at a level above it. Since audio comprehension differs from text comprehension in its elimination of decoding, this is an advantage. Second, Hacker's model demonstrates the role of monitoring and control in the comprehension process. Third, it explains how failures of comprehension represented by JOLs can be controlled by using a metacognitive strategy such as a navigational affordance.

Comprehension and Use of Affordances

Since regulation within the printed text domain is not readily observable, researchers have historically used errors in text to measure monitoring and regulation strategies. For example, Zabrucky and Moore (1994) conducted several experiments to examine monitoring and regulation strategies in adults. Their instruments contained "inconsistent" sentences that were in conflict with earlier propositions, delivered in a system that would display one sentence at a time. They would then measure the length of time a subject would dwell on a sentence, as well as the incidence of navigating back to a previously read sentence and pauses which could indicate a comprehension problem. They would then query the subjects on their recognition of the inconsistent sentences. They concluded that subject would take longer to read inconsistent information and would also reread it or a previous sentence that would confirm that the target sentence was incorrect. Using this model, they concluded that using navigational affordances was in response to a comprehension failure, and their use in the print text domain would increase comprehension of the text.

Roger Azevedo (2005) draws a link between self-regulated learning strategies and tool use in the hypermedia environment. Indeed, he indicates that self-regulatory skills are necessary for the effective use of tools. For example, "…students who lack key metacognitive and self-regulatory skills learn very little from these [technology mediated] environments" (p. 201). These studies suggest that utilizing affordances of the learning environment may result in greater comprehension. They further suggest that learners with low metacognitive ability may not effectively utilize affordances within technology mediated environments. Other studies suggest that the same may be true of learners with high metacognitive ability as the use of affordances may not be needed for comprehension. One may hypothesize that the relationship between metacognitive ability and affordance use may present as a nonlinear relationship (Credé & Phillips, 2011). While Credé and Phillips posit this, no subsequent research located by the author examined or supported this statement.

In summary, most of the audio comprehension literature has been focused on BVI learners. Cognitive and metacognitive factors affect comprehension, but most of the literature has focused on print text. The use of affordances has been linked to greater comprehension, but the types of affordances studied have focused on prosodic and organizational (e.g., menu navigation) affordances rather than audio text navigation, and these studies haven't examined comprehension. No study has examined a correlation of these factors in the audio domain, thus leaving a gap in the literature. One may hypothesize that these factors would affect comprehension. However, the use of navigational affordances would have to be tested. Learners' cognitive and metacognitive abilities, if measured with an appropriate instrument, could be correlated with affordance use and thus linked to comprehension.

Lookbacks

The term "lookback" has a respectable literature behind it. In most of the studies using the term, a lookback refers to the navigational affordance of referring back to text previously sensed, almost always within the printed domain. Since this term is equivalent to a type of navigational affordance (albeit in the textual domain), it is included in the discussion. A search of the term “lookback” in the ERIC database returns eleven sources with the term in the title and thirty results that use the term in some field. Four references target research on lookbacks for deaf learners. No study treats lookbacks for learners who are BVI, thus a gap in the literature is evident. This is understandable since no study to date has treated the topic of lookbacks in the audio domain. Since printed text is not accessible by BVI learners, one would not expect to find any, and there is not yet an equivalent term in the audio domain. One study (Yamada, 2009) connects Kintschian comprehension theory to the topic of lookbacks. A cursory review of these articles reveals differing definitions of the lookback term and also the measures of comprehension. Some actively train learners in lookbacks as a treatment (e.g., Walczyk & Taylor, 1996).

The earliest study to use the term lookback came from Alessi, Anderson and Goetz (1979): “Any reader will sometimes recognize that he or she has not understood a segment of text, and having recognized that will sometimes engage in rereading or some other fixup activities” (p. 198). In their study, the researchers defined three different types of comprehension failures that could occur when prerequisite information is in the text; the information was comprehended but later irretrievable, the information was never fully comprehended and integrated, or the information was too difficult for the learner to comprehend. They hypothesize that the third condition may not be corrected by a lookback, but later retract this hypothesis. Their comprehension failure scenarios are also supported by Craik and Lockhart (1972), who concluded that two reasons why a learner would reread a text segment include failure to engage in deep, meaningful processing, or incomplete/incorrect prerequisite knowledge, either outside of the text (i.e. prior or background knowledge) or failure to comprehend information already read in the text (i.e. subordinate information).

Alessi, Anderson, and Goetz contrasted those learners who needed few lookbacks versus many lookbacks. Learners who needed few lookbacks showed little difference in their comprehension of the text, while there was a significant difference for groups of learners who needed many lookbacks. Those who had access to lookbacks scored significantly better on the comprehension measure than those not having access to lookbacks. These findings indicate that lookbacks support increased comprehension of the text and are used more frequently when the learner's prior knowledge of the topic may be low (as evidenced by frequent lookbacks). Deficits in prior knowledge can be partially remediated by lookbacks (p. 208). Lookbacks significantly increased learner comprehension of the material, especially when initial comprehension of the material (as measured by frequent lookbacks) was poor. This finding is also supported in similar studies by Garner, Hare, Alexander, Haynes, and Winograd (1984), and Walczyk and Taylor (1996) in which learners were trained to use lookbacks when reading printed text material. This would support a hypothesis that navigational affordances in text are a desirable condition that increases text comprehension.

Baker and Anderson (1982) found that learners spontaneously looked back in printed text upon encountering inconsistencies in information that could indicate a problem with comprehension. Interestingly, they found that the percentages of recognized inconsistencies in main idea and detail of the text showed no significant difference. One would expect that inconsistencies in main idea (i.e. gist) would be recognized more frequently than detail inconsistencies. They also found that the between subjects condition of informing one group that there were inconsistencies in the text demonstrated no difference in results. Another finding of the study was that greater time on task is an indicator of the learner's awareness of an inconsistency, but not necessarily a lower level of comprehension.

Alvermann (1988) studied spontaneous and induced lookbacks in self-perceived high and low comprehenders. In her experiment, she created a graphic organizer that could be used by the learners to locate themes in the text read. She found that the use of the organizer by self-perceived low comprehenders had a positive effect on the comprehension, while there was no significant effect for self-perceived high comprehenders. Her conclusions also supported earlier discussions (cf. Alessi et al., 1979; Alvermann, 1988; Garner et al., 1984) that cited some learners' inability to organize text into a comprehensible structure, which may lead to an ineffective lookback strategy, as they may be distracted by irrelevant information. This condition matches Alessi, Anderson and Goetz's failure in comprehension due to the material being too difficult. Organizing text into a comprehensible structure suggests that metacognitive strategies will have an effect on comprehension.

More recent studies in lookbacks have focused on rapid eye movements (e.g., Burton & Daneman, 2007; Kinnunen & Vauras, 1995). These studies draw a distinction between saccades, which are small movements between six to nine character spaces, and regressions, which are larger, and generally between 15-20 character spaces. Reichle, Rayner, and Pollatsek (2003) conclude that these types of visual lookbacks may be the result of linguistic processing difficulties or oculomotor errors. The former would be problems characteristic of the linguistic comprehension process, while the latter would be problems with the decoding process, according to the Simple View of Reading comprehension model. Although this line of inquiry would not be useful for examining the topic of audio lookbacks, it is referenced here in the interests of a complete review of the literature on the topic of lookbacks.

Why Do Learners Look Back

As cited in Franks et al. (2013), three models of logical reasoning can be applied to explain why learners may use lookbacks as a metacognitive learning strategy. Since the author examines comprehension problems during the learner's logical construction, these explanations can be applied only to the linguistic comprehension processes of reading and not problems with decoding. In Moshman's model of metacognitive reasoning (1990), explicit awareness of one's own reasoning processes increases with age. This increase in metacognitive awareness would indicate that a strategy such as lookbacks should increase with age. The results of the experiment demonstrated a modest difference according to age, with college age learners looking back the most frequently. In a model proposed by Henry Markovits (1993), semantic memory as it affects reasoning is emphasized. In the model, information from long-term memory must be activated and integrated into the context of the material. This becomes problematic when indeterminate conclusions not supported by the text must be inferred, thus relying on prior knowledge. According to this model, the authors assert that indeterminate conclusions in the text would elicit greater lookbacks. This prediction was significantly demonstrated only for invalid (i.e. false) indeterminate conclusions. In the third model by Klaczynski and Daniel (2005), information processing takes place within the context of either experiential or analytical processes. Experiential processing, based upon prior knowledge and constructions, is relatively automatic, thus creating less dissonance and cognitive load than the analytical process, which requires greater cognitive effort and conscious control. Based upon this model, the authors predicted that indeterminate forms requiring analytical constructions would require more frequent lookbacks. This prediction was supported only when learners were asked to evaluate indeterminate invalid conclusions that would not be consistent with prior knowledge. Overall, college age learners look back more than younger learners. Questions about a text that ask the learner to evaluate rather than generate logical conclusions elicited more lookbacks. Walczyk and Taylor (1996) also found that those learners who had more efficient information in working memory used less lookbacks. They hypothesize that lookbacks may compensate for limited verbal working memory. Taken as a whole, these findings point to the use of lookbacks as a metacognitive strategy to increase comprehension and are instructive when designing an experimental instrument to measure effects of lookbacks on comprehension.

Comparison of Listening and Reading Comprehension

Since much of the literature is concerned with reading comprehension, it is important to consider whether or not this body of knowledge is applicable to listening comprehension. Danks (1980) takes up the question of equivalency between listening and reading comprehension. He proposes four areas that must be evaluated to answer this question; presentation conditions, learner characteristics, language materials, and comprehension measures. Assuming equivalency of language materials in both domains, and not taking up the topic of learner characteristics, the remaining two topics are discussed here.

Under the topic of presentation conditions, Danks emphasizes, besides obvious differences in the nature of printed and audio texts, suprasegmentals of audio (e.g., stress, intonation, pitch, and speed) when compared with text. In later studies, these affordances would be termed audio prosody. He asserts that punctuation cannot begin to convey the level of information that suprasegmental characteristics of audio convey. One may hypothesize that synthetic speech when compared to natural speech may eliminate some of the information conveyed in suprasegmentals, but this proposition creates other issues. Goudiras, Papadopoulos, Koutsoklenis, Papageorgiou, and Stergiou (2009) discuss learner's preference for natural speech over synthetic speech. Recently, advances in voice "fonts" have greatly increased the suprasegmental qualities of synthetic speech. Availability of tools to convert text to synthetic speech also creates a condition of ecological validity since it is likely that synthetic speech is more available than natural speech for a given text.

Danks also comments on the relative permanency of print text when compared with the ephemeral nature of audio text, a point previously made by Arons, that affects the nature and use of navigational affordances within these domains. Size of segments in both formats must also be considered. Research from Bever, Garret, and Hurtig (1973) concludes that integration happens at the end of clausal units, thus maintaining these units in the presentation would help comprehension. Danks asks if there is a transfer of listening comprehension to reading. According to Tunmer and Chapman (2012), we have evidence that there is, indeed, a positive interaction between listening comprehension and decoding, resulting in greater reading comprehension.

When discussing comprehension measures, Danks is cautious to state that even with a number of early studies (35) reporting correlations in comprehension measures ranging from .45 to .82, the distribution was not skewed and left a good deal of variance unaccounted for (p. 20). Danks also remarks that one problem with current comprehension measures is a delay in recall. He posits that what is needed is a method to measure comprehension simultaneously with input, a procedure similar to what Alessi, Anderson and Goetz (1979) did in their experiments by inserting the comprehension measures within the text, and thus controlling the delay in recall. One final point made about comprehension measures was the domain in which the responses were taken. According to Danks, the evidence for an interaction between the presentation and response modes is unclear.

Summary

Affordances are environmental properties that determine the learner-environment interaction. The navigational affordances of text media can affect comprehension. One type of navigational affordance, the lookback, has been studied primarily in the print text domain. Few studies have been conducted in the audio domain and no studies have examined BVI learners, thus identifying a gap in the literature.

Learners refer back to previously read or heard text for a variety of reasons. The three most widely cited reasons were; the information was comprehended but later irretrievable, the information was never fully comprehended and integrated, or the information was too difficult for the learner to comprehend. In any case, the learner must engage in self-regulation to identify the comprehension disruption and then take action (e.g. lookback) to correct the problem. Considering the audio domain, this process requires the use of a navigational affordance. Within this context, navigational affordances can be used as a metacognitive strategy to increase comprehension.

A Systematic Literature Review of the Effects of Navigational Affordances on the Comprehension of Audio Text by Blind and Visually Impaired Learners

Since much of the knowledge on audio text comprehension is contained in the blind and visually impaired learner literature, a systematic literature review was completed to further elucidate the relationship between audio text and comprehension, albeit in a specific population of learners. The BVI community was the first population of learners to use audio text as a significant learning modality.

Relevant factors for blind and visually impaired learners

Learning media

Goudiras et al. (2009) report that cassettes are the most likely medium for visually impaired learners (52%), followed by computer audio or screen magnified text (35%), while Braille is the least likely (28%). They state, "The popularity of cassettes can be explained by the participants' view that cassettes are perceived through the auditory sense, without much effort or training being required, and that they can be used as an easy and convenient means of recording of speech (p. 125)." They further conclude that those learners who used audio text formats were more satisfied that their daily, educational, and professional needs were met. The article also states that the "natural voice" of cassettes are generally preferred over the "computer voice."

Another notable area of study for BVI learners is access to non-linear or graphic learning materials such as tables, graphics, and illustrations. There are numerous studies that examine accessibility of such digital learning media. One notable source of information on this topic is the World Wide Web Consortium's Web Accessibility Initiative (World Wide Web Consortium, 2014). This resource provides comprehensive information about accessibility affordances of tabular and graphic materials. While there are some characteristics of navigational affordances in this body of knowledge, especially when dealing with table navigation, this topic falls outside the scope of the present study.

Population

The estimated population of adults aged 18-64 in the United States with a visual impairment is 1,862,889 (Cornell University, 2015). Considering reading medium statistics, reports for the year 2013 detailed the different modalities of text utilized by students with visual impairments. The American Printing House for the Blind (2015), reported that 19.5% of respondents are “auditory readers” which are defined as those students who primarily rely on audio, either natural or synthesized. This statistic reflects a 1.5% increase in two years. When added to the reported samples of pre-readers (18%) and non-readers (35%), the potential population who may benefit from audio texts increases to 72% of the population. This population is not a random sample given that they were only considering students registered with the American Printing House for the Blind. However, the sample size is large enough to use the figure as a rough estimate for those visually impaired learners who primarily use audio. Using these statistics, a total population size of visually impaired audio learners for the U.S. can be estimated at 334,000, and the potential population who may benefit from audio texts is 1.34M.

Ease of navigation

Goudiras et al. (2009) found that difficulty of navigation was reported by visually impaired learners for cassettes but not for computer audio. This would indicate that the navigational affordances of computer audio interfaces are more advanced than cassette tapes. It also infers that differences in navigational affordances can affect the learner's interaction with the text, and consequently may affect comprehension.

Audio text comprehension

Research in the area of audio text comprehension by BVI learners reached its peak in the 1960’s, 70’s and 80’s as evidenced by a review of the literature. Many of the studies focused on the reading rate of blind subjects. Not surprisingly, BVI learners who use audio texts on a regular basis develop greatly accelerated reading rates when using this medium. One of the landmark researchers in audio reading rate studies (i.e., compressed speech) was Emerson Foulke. A number of early works (e.g., Foulke, 1962; Foulke et al., 1962; Foulke, 1964; Foulke, 1966; Foulke & Louisville Univ., 1967; Foulke & Louisville Univ KY Perceptual Alternatives Lab, 1969; Foulke et al., 1970; Foulke & Louisville Univ KY Perceptual Alternatives Lab, 1971) contain an array of experiments on comprehension of compressed speech texts. However, the focus of these studies is almost entirely on measuring the effects of audio reading rate on subjects' comprehension rather than on the design of audio text interfaces to enhance comprehension, navigational affordances, or the mechanisms of comprehension itself. Another heavy focus of these studies was the mechanics of producing compressed speech texts. During this period, the technology used to produce compressed speech texts was in its infancy. Topics of interest included controlling the pitch of audio texts (not within the context of prosody), listening rate preferences, and measuring the reading rate more precisely (e.g., word rate vs. syllable rate).

To examine audio text comprehension by BVI learners, one must draw a distinction between intelligibility and comprehension. Within the audio text literature, intelligibility is classically defined as the listener's ability to recognize phonemes and words presented in isolation (Ralston, Pisoni, & Lively, 1991). This definition may be contrasted to that of comprehension, which involves understanding the underlying meaning from the acoustic signals of speech (Duffy & Pisoni, 1992). In more modern terminology such as that of the Simple View of Reading as applied to audio texts (Hoover & Gough, 1990), these factors of intelligibility and comprehension may be considered equivalent to encoding and linguistic comprehension. This distinction is necessary to begin a review of the audio text literature, and to clarify terms.

Method

Based upon these studies, two questions were examined for the systematic literature review. What challenges to comprehension do adult blind or visually impaired learners face when using audio text? What aspects of audio text interface design support learner comprehension? Systematic literature reviews were conducted for each of these questions limited to primary reports of quantitative studies.

Comprehension of audio text by BVI learners

To answer the first question, a systematic literature review was conducted on the topic of audio text comprehension by BVI learners. The review was designed to answer the first research question. The search was conducted using the EBSCO search engine and all available databases. The search focused on locating sources with three criteria: treatments of audio text (i.e., speech synthesis, synthesized speech, speech synthesizer, text-to-speech, audio text, or synthetic speech), comprehension (i.e., comprehension, speech perception, or auditory perception), and participants who were BVI (i.e., blind or visually impaired). Initially, terms that identified studies focused on text navigation were included. But it was determined that adding this constraint was too limiting to the results. Indeed, no studies were identified that treated the effects of navigational affordances on the comprehension of BVI participants. While the final search query did not explicitly define navigational affordances as a term, the remaining three search terms were thought most likely to lead to the identification of studies that could potentially address the first question.

A number of systematic review databases including the Cochrane Collaboration, Campbell Collaboration, and the Centre for Evidence Informed Policy & Practice in Education were searched for previous reviews on the topic. Only one review was found (Virgili, Acosta, Grover, Bentley, & Giacomelli, 2013), and it was not related to the topic at hand, instead focusing on a review of optical reading aids for visually impaired individuals. The initial search found sources in 16 unique databases (Ergonomics Abstracts, ERIC, Education Full Text (H.W. Wilson), OmniFile Full Text Mega (H.W. Wilson), Academic Search Complete, Health Source:

Nursing/Academic Edition, MEDLINE with Full Text, Psychology and Behavioral Sciences Collection, CINAHL Plus with Full Text, MasterFILE Premier, MEDLINE, Professional Development Collection, Information Science & Technology Abstracts (ISTA) Library, Information Science & Technology Abstracts with Full Text, Science & Technology Collection, and Business Source Complete). Ergonomics Abstracts proved to be the most fruitful database with ten sources, while ERIC returned a respectable nine sources. In all, 70 sources were found. Of these, 38 were found to be duplicates between databases, thus yielding an initial sample of 32 sources.

As detailed in Figure 2, during primary screening, 11 sources were eliminated based upon a reading of the title and abstract. There were four screening criteria for this phase of the review. Sources that primarily treated the variable of reading rate were eliminated, resulting in three articles culled (Asakawa, Takagi, Ino, & Ifukube, 2003; Hertrich, Dietrich, Moos, Trouvain, & Ackermann, 2009; Hertrich, Dietrich, & Ackermann, 2013). A second screening criterion eliminated sources that dealt primarily with non-audio formats. This screening criterion eliminated another two articles (Boulton, 1991; Levitt, 1995). A third screening criterion identified sources that treated populations other than BVI or treated BVI only briefly, resulting in the elimination of another two sources (Elkind, Black, & Murray, 1996; Haskins Labs., New Haven, CT., 1972). The last screening criterion identified sources that did not focus on comprehension. These included studies of interface design and technology development. Using this screening criterion, four sources were eliminated (Edwards, 1991; Malyan, Sunthankar, & Barnwell, 1987; Pavešic, Gros, Dobrišek, & Mihelic, 2003; Walker et al., 2013).

In the second phase of the review, sources were eliminated based upon reading the full text of the source. Another 10 sources were eliminated using the same screening criteria, including one source that treated non-BVI subjects (Smith, 1972), and six sources that did not examine comprehension of audio texts (Fellbaum & Koroupetroglou, 2008; Freitas & Kouroupetroglou, 2008; Ramloll, Brewster, Yu, & Riedel, 2001; Simpson, 1996; Simpson, 1999; Spiliotopoulos, Xydas, Kouroupetroglou, & Argyropoulos, 2005). Most of these studies focused on interface design issues rather than assessing audio text comprehension. Two studies focused on reading rate experiments (Foulke & Louisville Univ., 1967; Torihara, Nakamura, Ueda, Wada, & Ishizaki, 2006), while one study (Garcia, 2004) treated printed texts rather than audio texts. One source was eliminated because it was not available (Luxton, 1985).

Findings

Upon review of the remaining 10 sources, four themes emerged: intelligibility and comprehension, prosody, comparisons of natural and synthetic speech, and the effect of prior learner experience.

Intelligibility and comprehension

Argyropoulos, Papadopoulos, Kouroupetroglou, Xydas, and Katsoulis (2007) studied paronyms in blind and sighted populations of students. A paronym is a word that is acoustically very similar to another word, such as “affect” and “effect.” In this quantitative study, Argyropoulos, et al. studied 30 BVI and 30 sighted students, aged 21-24, in two experiments. In the first, 12 pairs of paronyms were offered in isolation to the subjects who then recounted what they had heard. Phonological type errors of words (PTEW) were measured and recorded including the category of error such as accentuation, phonemic substitution, or combinations of these. In the second experiment, 17 paronyms contained within sentences were presented and the subjects asked to recount what they heard. These types of errors were considered phonological type errors of sentences (PTES). The subjects were able to pause the audio material, but not rewind it. Synthetic speech texts were used rather than natural. While the authors confirm their use of synthetic speech texts to increase ecological validity, they commented on the associated cost in intelligibility. While there was no correlation between PTEW and PTES for BVI subjects, there was a moderate correlation for sighted subjects ($r_s = .368$, two-tailed, $p <= .05$). The authors offered no rationale for this result other than to remark, “It seemed that the rendering of the paronyms caused a sort of confusion in the sighted [subjects] rather than in the blind participants (p. 211).” When comparing between subjects’ results, the BVI students had significantly less errors than the sighted students in the PTEW test (Mann-Whitney U=237.0, Z=3.192, p<= .005). The authors attributed this result to the practice effect since BVI subjects had a much greater exposure to using synthetic speech text than sighted users. While the results of the experiments are not very surprising, what is notable is that phonemic errors of recognition occur and may result in confusion and lower intelligibility (i.e., recognition). Hypothetically, BVI subjects with experience in audio texts would have less phonemic errors of recognition and greater comprehension than those with less experience. When framed within Kintsch’s model, would paronyms create a disruption in comprehension? It is unfortunate that only a pause affordance was offered within the experiments. Would subjects have been able to resolve phonemic errors if a review affordance were available?

Similarly, a quantitative study by Papadopoulos, Argyropoulos, and Kouroupetroglou (2008) duplicated the procedure of the experiment by Argyropoulos et al. (2007) with different subjects but similar groups (BVI students aged 18-35, and sighted students aged 21-24). They confirmed the finding that there was no significant difference in PTES between BVI and sighted students ($t = 1.316$, $df = 54$, $p = .194$), while there was a significant difference in PTEW ($t = 3.049$, $df = 54$, $p =< .01$), with BVI subjects outperforming sighted subjects. The authors did additional analyses of the experimental data and concluded that gender, age, degree of vision loss, age at vision loss, and frequency of synthetic audio text use were not significantly correlated with the number or type of errors for either group studied.

The authors also did a second experiment to examine "comprehensibility" (i.e., comprehension). The subjects were given a structured synthetic audio text consisting of 304 words including 36 paronyms. The subjects were then given 10 comprehension questions about the text. The results indicated that there was no significant difference between the groups in comprehension ($t = 1.364$, $df = 54$, $p = .178$). This result is expected since the earlier experiment found no significant difference in PTES, which is the same context under which phonological type errors in sentences were tested in the earlier experiment. However, there was one interesting artifact from the experiment. Within the BVI subjects studied, when comparing those that scored in the lower half of the comprehension test with those in the upper half, the former made an average of 5.0 PTEW while the latter made an average of 2.95. While the difference was not statistically significant ($t = -1.956$, $df = 26$, $p = .061$), this finding is suggestive that BVI subjects may demonstrate a correlation between phonological errors and comprehension specifically related to paronyms. That is, subjects with a greater propensity of phonemic errors may have lower comprehension. What if confusion resulting from phonemic errors could be resolved? What types of affordances could affect resolving the confusion? While subjects were not afforded navigational controls other than the ability to pause the synthetic audio text in this experiment, it certainly begs the question.

Papadopoulos, Koutsoklenis, Katemidou, and Okalidou (2009) studied intelligibility and comprehension of natural and synthetic speech texts for BVI subjects. Twenty-five adults, aged 18 to 53, participated in three experiments; two were very similar to Argyropoulos et al.(2007)

and Papadopoulos et al.(2008), and examined intelligibility of words in isolation and sentences in isolation, but applied these conditions to both synthetic and natural audio texts for comparison. They found that intelligibility of both single words ($t = -7.667$, $df = 24$, $p < .01$) and sentences ($t = -3.165$, $df = 24$, $p < .01$) was significantly better when using natural speech rather than synthetic speech texts. But in the third experiment, which tested comprehension of a discourse text consisting of 10 sentences, they found that comprehension was not affected by the audio text format, even with the differences in intelligibility:

> It seems that context cues provided by the text assisted the participants in identifying and comprehending the text more effectively. If we take into consideration that the overall purpose of reading is comprehension, this finding has a unique practical value because it indicates that the use of [text-to-speech] systems by individuals with visual impairments does not affect the intended purpose of reading. (Papadopoulos et al., 2009, p. 411)

This finding would indicate that synthetic speech texts can be comprehended as effectively as natural speech if audio material is not presented in isolation.

Interestingly, the authors cited a number of studies in their review of the literature which indicated that background noise has a detrimental effect on comprehension of synthetic audio texts, for example (Fucci, Reynolds, Bettagere, & Gonzales, 1995). This indicates that future studies may need to control for background noise when studying comprehension. They further cited studies that found differences in intelligibility as a correlation of treatment conditions with “meaningful” vs. “anomalous” sentences. Although not the subject of this study, and not identified during the systematic literature review, one may take interest in how the differences in these conditions affected intelligibility and/or comprehension as they may inform the design of future inquiries.

In Story and Kuyk (1988), audio text recognition (i.e., intelligibility) was examined in a population of 20 sighted adults with a mean age of 37 across 3 synthetic audio text systems (Echo+, Slotbuster, and SynPhonix). None of the participants had any previous experience using the audio systems. The procedure measured recognition of letters, words, sentences and paragraphs. They concluded that there was a significant main effect difference in recognition between the default and user-selected system settings ($F = 12.36$, $p = .0009$) in the direction of the user-selected settings, but this was due primarily as a result of the difference observed in the Slotbuster system ($t = 3.736$, $p = .0006$). While this study did not use a BVI population of participants, it is included here because a later study (Kuyk & Story, 1988), focused on practice effect, references it. One useful inference that can be derived from this study is that the quality of the synthetic audio text system may affect intelligibility. While the study confirmed that user-selected settings affect intelligibility favorably, they do not necessarily translate to increased comprehension. Indeed, Hjelmquist et al. (1992) found that there was no increase in comprehension as a result of increased speed (i.e., a user preference), while Story and Kuyk didn’t differentiate between different types of user preferences. Also, in this very early study, the quality of the systems can be assumed to be significantly less than systems available today. Would effects be greater as a result of the increased audio quality?

Audio text prosody

Audio text prosody refers to affordances that vary audio pitch, speed, and/or volume to convey meaning. A distinction must be made between studies that examine prosody as a means of conveying meta-information (e.g., tabular structure, heading size, etc.) in order navigate a text, and those that use prosody affordances of text to examine their effect on comprehension. While one may argue that the former may affect comprehension, the focus of this inquiry is to examine affordances as they affect comprehension of primary text. In Xydas, Argyropoulos, Karakosta, and Kouroupetroglou (2005), as well as Argyropoulos, Sideridis, Kouroupetroglou & Xydas (2009), typographical attributes of bullets, bold and italics were conveyed using audio text prosody (i.e., varying pitch, speed and volume) for both BVI and sighted subjects.

In Xydas et al. (2005), the authors described the creation of an audio production system that parsed XML formatted documents and converted them into speech synthesis markup language (SSML) documents to be processed by a text-to-speech system. The system was designed to create an experimental testing environment capable of configuring different types of prosody (i.e., volume, pitch, speed, accent tones) for different typographical attributes (i.e., bold, italics, bullets). The authors then tested four experimental conditions which varied the types of prosody used across eight sighted and BVI subjects aged 22-25. For each condition, the dependent variable was recognition of the typographical attribute based upon the prosodic affordance marker provided. The findings of the study indicated that BVI subjects recognized differences in prosodic markers better than their sighted peers, but not significantly. The greatest effect between the samples was observed in the test condition of greatest pitch difference (45%) and greatest speed difference (30%) between bold and italics. Means and standard deviations of the test results were given. Calculating a t-difference, for this last group, we find $t = 1.894$, $df = 7$, $p = .95$. This is primarily due to the small sample size, and it was obvious that the main purpose of the study was to describe the system model, rather than run a rigorous quantitative analysis. In a second experiment, the authors asked the subjects to identify the type of typographical attribute that was being conveyed by the prosodic affordance. Again, the BVI subjects were able to identify the typographical attributes better than their sighted peers, but not significantly. Qualitative interviews of the subjects indicated that change in speed was least identifiable prosodic affordance when compared to pitch, volume, and audio icon affordances. One interesting observation was that the experimental condition with the greatest effect was not the preferred condition, according to the subjects.

In Argyropoulos et al.(2009), sixty students at Greek colleges, thirty BVI and thirty sighted, aged 20-25 were studied. The authors expanded on the earlier study by Xydas et al. (2005) by examining bold and italics typographical attributes in audio texts as indicated by prosodic markers. The authors further refined the italics typographical attribute into two distinct semiotics – emphasis and definition, while maintaining the bold semiotic to indicate strong emphasis. One control condition and four experimental conditions with different types of prosodic affordances were tested across both populations. The results indicated that BVI students performed better than sighted peers in both identification of typographical attributes due to prosodic affordances and recognition of the correct type of prosody being used in all four conditions ($F(3,174) = 3.060$, $p = .03$). Both populations performed better on test condition four in which bold was indicated by a 20% increase in pitch, definitions were indicated by a 30% increase in volume, and italics was indicated by a 15% decrease in pitch, when compared to the control condition (pitch=110Hz, speed=140 wpm, volume=100%). They confirmed the earlier finding that speed is not an effective prosodic accommodation when compared to changes in pitch and volume. During the qualitative interviews, the students indicated that changes in pitch and volume were easier to recognize than changes in speed. The authors assert that being able to identify typographical attributes in an audio text is "…tremendous because this information enriches comprehension and provides a fuller picture of the semantics of text… (p. 199)."

Aryropolous et al. conclude, “These studies need to increase and be extended in the area of comprehension to draw reliable inferences regarding the effectives of the various auditory manipulations (p. 199).”

Comparing natural and synthetic speech audio texts

In Hjelmquist, Dahlstrand, and Hedelin (1992), memory and comprehension of natural and synthetic audio texts were compared in groups of BVI participants varied by age and experience. Forty-eight participants were split into three groups. Group 1 (inexperienced) were adults aged 35-55 with no previous experience with synthetic audio texts. Group 2 (experienced) were adults aged 35-55 with at least one year of synthetic audio text experience, and Group 3 (old) were adults aged 65-80 with no experience in synthetic audio text. A four-factor design with repeated measures for one factor was utilized. Three between-subjects variables were speech mode (natural and synthetic), retention interval (immediate and 2 weeks), and sample (old, inexperienced, and experienced). The within-subjects variable was the text type used, which was one of four averaging 215 words in length. The authors observed a significant increase in performance for the natural group when compared to the synthetic group ($F(1, 36) = 14.33$, $p < .001$). There was also a large decrease in memory in the delayed group when compared to the immediate recall group ($F(1, 36) = 70.01$, $p < .0001$). Analyzed together, there was not a significant difference between speech mode and text type, although analyzed separately, there were significant differences when the text type had a larger portions of spatial information ($F(1, 36) = 4.48$, $p = .041$; $F(1, 36) = 15.13$, $p = .0004$; $F(1, 36) = 18.76$, $p = .0001$). The experienced group did not perform better than the inexperienced group when using synthetic audio texts, although they did perform better than the old group.

In general, their findings indicate that memory and comprehension are significantly better when natural speech texts are used when compared to synthetic speech texts across all subject groups, a finding in conflict with Papadopoulos et al (2009). They were able to replicate significant differences in memory and comprehension based upon text type such as literary, informational and spatial, though non-significant differences in comprehension of these text types were due to modality (i.e. natural vs. synthetic). This finding supports previous research indicating that comprehension is affected by text type.

Age and experience significantly affected the preferred speed of synthetic audio texts ($F(2,36) = 40.92$, $p = .0001$). Participants experienced with synthetic speech texts preferred a faster reading rate (159 wpm) while the old group preferred the slowest rate (108 wpm) and the inexperienced group fell in the middle (124 wpm). This may have affected the finding that experienced group didn't perform better when using synthetic audio text since the increase in speed may have decreased their comprehension. It also infers that comprehension can be achieved by experienced synthetic speech text users with much higher preferred reading rates, but compensated for with inexperienced participants by slowing the rate down to a normal speech rate. This finding has significance for the present study by confirming earlier work from the reading rate researchers such as Foulke (1967) that a synthetic speech delivery system testing comprehension in BVI participants should use a normal reading rate to reduce between participants variability. While this study supports that there is a tradeoff between the greater independence of synthetic speech and greater comprehension of natural speech, reading rate, though comprehensible at higher speeds by participants with experience, should be kept at normal rates.

Hansen, Lee, and Forer (2002), in a brief “research note” administered a “self-voicing” system to 17 BVI participants, aged 17 to 55, and asked them to rate the usefulness of the system for test-taking multiple-choice tests. The participants reported approval of the system for test taking when compared to human readers since it gave the participants greater independence in contrast to human readers. The most serious issue with the system was synthesized speech quality, specifically pronunciation and lack of prosodic intonation. The use of Microsoft’s® fourth version Speech Developer’s kit could have affected the findings as this was a young technology at the time of the article. Pronunciation and prosody in speech fonts have developed nicely in the intervening years. The lesson learned from this research source is that speech font quality is an important factor in a text-to-speech interface. It is also notable that the preference of synthetic speech texts when compared to human readers (i.e. natural speech) was tied to greater independence rather than a finding of superior intelligibility or comprehension. This finding supports greater ecological validity in an interface that uses synthetic speech as BVI learners would prefer it.

Learner experience

While many of the previously cited studies have remarked on the practice effect, and one (Hjelmquist & Others, 1992) confirmed a positive difference between BVI participants that have had experience with synthetic audio texts and those have not, none of these studies has focused on the effects of experience. Rhyne (1982) conducted one of the first inquiries into the comprehension of synthetic speech by BVI participants, focusing on the effects of practice. Using four children ranging in ages from 11 to 14, inexperienced in synthetic audio text use, she recorded audio 40 stories using the Kurzweill® Reading Machine as an audio source, and presented the stories to the participants, 4 per day, for 10 days. Dividing the trials into four groups, the researcher found that comprehension significantly increased linearly with the number of trials completed ($x^2 = 6.375$, $df = 3$, $p < .10$). The author also noted in her conclusion that use of the system allowed students to manage their own learning rather than rely upon readers, thus supporting the observation that synthetic audio texts are easier to obtain and use than natural audio texts which increases the ecological validity of studies using synthetic audio texts.

Kuyk & Story (1988) compared synthetic audio text recognition (i.e., intelligibility) between 6 BVI participants experienced in synthetic audio text use and 20 inexperienced sighted participants from a previous study (Story & Kuyk, 1988). The researchers reported no significant difference between the groups ($p > .10$). It may be noted that the previous study did not examine a population of BVI participants. There was also a significant difference in the sizes of the groups. These factors bring the authors' conclusion into question.

Audio Text Interface Design for BVI Learners

With an understanding of audio text comprehension challenges reviewed, a second systematic review was conducted on the topic of audio text interface design for BVI learners. A cursory review of the interface literature revealed an extremely large body of research on textual interface design, a respectable body on audio text interface design, and a relatively compact body on audio text interface design for BVI learners. No studies were found that directly addressed navigational interfaces as a treatment affecting comprehension for BVI learners, and so the question of interface design for BVI learners had to be studied more broadly.

A systematic literature review was conducted on the topic of audio text interface design for BVI learners. The search was conducted using the EBSCO Host search engine and all available databases. The initial search focused on locating sources with three criteria; studies that focused primarily on audio text (i.e., speech synthesis, synthesized speech, speech synthesizer, text-to-speech, audio text, or synthetic speech), examinations of navigation or user interfaces excluding sources that examined physical navigational interfaces, and participants who were BVI (i.e., blind or visually impaired). A number of databases, including the Cochrane Collaboration, Campbell Collaboration, and the Centre for Evidence Informed Policy & Practice in Education, were searched for previous reviews on the topic. Only one review was found (Virgili et al., 2013), but it was tangentially related to the topic focusing on a survey of BVI technology tools rather than audio text. The initial search found sources in 23 unique databases (Ergonomics Abstracts, Academic Search Complete, OmniFile Full Text Mega (H.W. Wilson), Applied Science & Technology Full Text (H.W. Wilson), Business Source Complete, CINAHL Plus with Full Text, Education Full Text (H.W. Wilson), ERIC, Internet and Personal Computing Abstracts, Library/Information Science & Technology Abstracts with Full Text, MasterFILE Premier, MEDLINE with Full Text, Business Abstracts with Full Text (H.W. Wilson), Computer Source, Gender Studies Database, Health Source: Nursing/Academic Edition, Information Science & Technology Abstracts (ISTA), Library Literature & Information Science Full Text (H.W. Wilson), MLA International Bibliography, Professional Development Collection, Readers' Guide Full Text Mega (H.W. Wilson), Science & Technology Collection, Vocational and Career Collection). Ergonomics Abstracts proved to be the most fruitful database with 28 sources. No other database returned more than five sources. In all, 69 sources were found. Of these, 30 were found to be duplicates within and between databases, thus yielding an initial sample of 39 sources.

As detailed in Figure 3, during primary screening, 27 sources were eliminated based upon a reading of the title and abstract. There were four screening criteria for this phase of the review. Sources that did not primarily examine audio text interface design were eliminated. Most of these were product announcements. This criterion eliminated four sources (Chen, 2005; Editorial, 1994; Laws, 2005; Raman, 1997). There were a large number of sources that examined audio text interface design, but in a non-textual context. Most of these articles focused on graphic user interfaces, web search interfaces, and appliance interfaces. Using this criterion, 18 sources were culled (Andronico, Buzzi, & Leporini, 2005; Andronico, Castillo, & Leporini, 2006; Andronico, Buzzi, Castillo, & Leporini, 2006; Asakawa & Takagi, 2001; Ebina, Igi, Miyake, & Takahashi, 1999; Edwards, 1992; Grammenos, Savidis, & Stephanidis, 2005; Gunderson, 1993; Hunsucker, 2013; Karshmer, 1997; Lazzaro, 1991; Martial & Dufresne, 1993; McNulty & TLT Group, Washington,,DC, 1994; Morrissey & Zajicek, 2000; Murphy, Kuber, McAllister, Strain, & Yu, 2008; Pontelli et al., 2004; Uslan, Eghtesadi, & Burton, 2003; Weber, 1992). There were also a number of sources that treated physical navigation interfaces, even though this topic was screened in the search construction. Four additional sources were eliminated using this criterion (Hub, Diepstraten, & Ertl, 2004; Klie, 2013; Marston & Church, 2005; Parslow, 2005). The final criterion eliminated sources that did not focus on audio interface design. Sources that treated audio text among other formats (e.g., Braille, voice recognition, etc.) fell into this category. Using this criterion, one source was eliminated because it examined haptic interfaces rather than audio (Wall & Brewster, 2006).

In phase two screening, a full reading of the articles further reduced the body to seven articles. One article was eliminated because it did not focus on audio interface design (Kochanek, 1992), and four articles were eliminated because they focused on some aspect of audio interface design in a non-textual context (Abu Doush, 2010; Asakawa, Takagi, & Itoh, 2001; Loyd, Phalangas, & Barner, 1999; Zajicek & Powell, 1997). No articles were eliminated with criterion three (physical navigation) or criterion 4 (non-interface studies).

At this point in the protocol, it was determined that the remaining seven sources were not sufficient to adequately treat the topic of audio text interface design. The limiting factor appeared to be the conditional search criterion that limited sources to those that studied BVI participants. So the search was modified to remove the BVI search limiter and run again. This change resulted in 314 sources including the original set. Many of the sources were in popular literature and trade publications and were not primary research reports, so the search was further limited to journal articles, leaving 118 sources. After eliminating duplicates, 63 sources were left, 41 of which were new sources not included in the initial search. Primary screening eliminated 12 sources that examined audio interfaces focused on a non-textual context (Archer, Head, Wollersheim, & Yuan, 1996; Editorial, 1997; Grams, Smillov, & Li, 1992; Grasso, Ebert, & Finin, 1998; Hakulinen, Turunen, & Raiha, 2005; Jeon, Walker, & Srivastava, 2012; Kim, Kim, & Hahn, 2006; Knott & Kortum, 2006; Moller, Krebber, & Smeele, 2006; Rezac et al., 2012; Vertegaal & Eaglestone, 1996; Yuan, Head, & Du, 2003), 10 sources that examined interfaces that did not focus on audio text such as robotic controllers, multimodal interaction controllers, gesture identification, and affective recognition interfaces (Alonso & Keyson, 2006; Lamell, Bennacef, Gauvain, Dartigues, & Temem, 2002; Marsic & Dorohonceanu, 2003; Maxwell et al., 2001; Noot & Ruttkay, 2005; Norén, Svensson, & Telford, 2013; Partala & Surakka, 2004; Perzanowski, Schultz, & Adams, 2001; Qiu & Benbasat, 2005; Qiu & Benbasat, 2009), and three sources that focused on physical navigation interfaces (Bretz, 2001; Editorial, 2007; Moller, Jekosch, Mersdorf, & Kraft, 2001). No sources were eliminated during this phase using the non-interface screening condition.

In phase two, sources were reviewed and eliminated based upon reading the full text of the source. Another 15 sources were eliminated using the same screening criteria, including three articles that were not focused on audio interface design (Balaji & Roche, 2003; Commarford, Wilson, & Stanney, 2002; Zue, 1995), eight sources that focused on some aspect of audio interface design in a non-textual context (Engwall, Balter, Oster, & Kjellstrom, 2006; Kamm, 1995; Lee, Jung, & Nass, 2011; Nass, Robles, Heenan, Bienstock, & Treinen, 2003; Palladino & Walker, 2008; Rosson, 1985; Schmandt & Roy, 1996; Sharit, Czaja, Nair, & Lee, 2003), and three sources that did not primarily treat audio text interfaces (Balaji & Roche, 2004; Bennett, Greenspan, Syrdal, Tschirgi, & Wisowaty, 1989; Boyle, Washburn, & Rosenberg, 2002). No sources were eliminated in this review based upon the physical navigation criterion. This procedure left two unique sources to add to the original seven for a total of nine sources focused on audio text interface design. This second review is also detailed in Figure 3.

Findings

Upon reviewing the remaining nine articles, none of the sources examined interface design as a treatment affecting comprehension in BVI learners. However, all of them drew conclusions about how affordances affected the perceived usability of the interface design, i.e., ease of use, or improved navigation as measured by speed, improved orientation, or more efficient navigation. Many of the conclusions were qualitative, though some provided quantitative data. Most were focused on the interface design rather than examining an experimental design. Two themes emerged: navigational affordances, and orientation affordances. Most sources treated at least one aspect or both of these themes. So rather than split sources based upon themes, they are described independently and subsequently summarized.

Navigation and orientation

Perhaps the most useful source reviewed was a system design article by Arons (1997) which also included a qualitative interview of participants. Arons developed an audio text browser interface named SpeechSkimmer that provided a number of navigational affordances for natural speech audio texts, including time compression, prosody, and skimming plans. One novel affordance of the system was the use of emphasis, pitch, and/or pause detection analysis used to extract summary information from the text to afford efficient skimming. However, the most intriguing affordance of the SpeechSkimmer system for the present study was backward skimming – an affordance which is the audio equivalent of the lookback affordance in the textual domain. In this affordance, audio text segments, with any of the process modifications already discussed, were played in reverse order. For example, if the selected text segment size was a sentence, the last sentence played was replayed, then the preceding sentence, and so on. Not all participants found backward skimming useful, but those that did made several interesting observations. For example, one subject commented, “Compared to a tape where you're constantly shuffling back and forth, going backward and finding something was much easier, since [while] playing backward you can still hear the words” (p. 25). Another subject commented that it would be useful to indicate when the recording was being played backward to distinguish it from forward (p. 26). One can imagine that an audio icon could fulfill this purpose.

The affordance of jumping was also described. Jumping allows textual navigation movements during playback. Jumping allows a learner to control their own browsing plan rather than being constrained by waiting on the interface. The interface also used auditory icons to indicate varying levels of browsing. Auditory icons are symbolic sounds in audio texts used to denote actions. For example, the sound of crumpling paper could symbolize deleting a file. The SpeechSkimmer increased the pitch of the auditory icons with higher (i.e., less granular) levels of skimming. There were also auditory icons for a number of situational events, such as a “boing” sound when the subject had reached the end of the audio text. One important finding of this study was that slowing down a segment during skimming is reportedly useful to participants, and begs the question if slowing down audio text upon review would increase comprehension. Another intriguing finding was that users preferred pitch-based skimming, which targeted conceptual boundaries, over isochronous skimming, which extracted and played back segments at regular intervals (e.g. 5 seconds every minute). One comment echoed by several participants was the desire to have a "bookmark" affordance where the subject could "drop" a bookmark during the playback so that they could refer back to that particular location.

Morley, Petrie, O'Neill, and McNally (1999) completed a small usability study which examined audio text interface design from the perspective of navigation and orientation. Among these were navigational, search, and summary affordances. They also incorporated audio icons into their design for orientation cues and command execution. They found that the most frequent affordances used for orientation were "Home," "Back," "Read Right," and "Where Am I?" in that order. Participants became disoriented infrequently and reported that the system gave sufficient orientation cues because of the audio icons.

Along similar lines, Reddy, Gupta and Karshmer (2005) studied audio navigation of mathematical markup for BVI participants. They suggest two basic guidelines (i.e., affordances) for effective navigation. First, the user should have the ability receive orientation feedback at any time during the playback. Second, the user should have the affordances of not only hearing what they are navigating over or through, but also be able to define "anchors" that can be returned referenced at any time, like bookmarks in the Arons study. They state, "full-fledged dynamic navigation […] necessitates that the listener be granted freedom to label a portion of the document with an audio identifier of their choice and later (during the same browsing session) return to that labeled portion by uttering the identifier" (p. 5).

Shajahan and Irani (2007) explored using prosodic attributes of pitch, pitch range, and speed of synthetic audio text objects to test participants' recognition of nodes in a hierarchy. In the first experiment, two hierarchical trees of nine synthetic audio text phrases each were created. Each tree had three nodes in the first layer and nine in the second. Each first layer node was a parent to three second layer nodes. The audio text phrases assigned to each node were short unrelated phrases that did not duplicate words and were meant to be neutral objects. For each tree, in the first layer of nodes, one of the prosodic attributes was varied while the other two were held constant. For example, in the first layer of the first tree, pitch and range were held constant while speed was varied across the three nodes. In the second layer of the first tree, speed was inherited from the parent node, while pitch and range were varied. A control tree with no prosodic attribute variation was also created. Twelve undergraduate students were given an unlabeled diagram of the tree hierarchy and then allowed to test each node by clicking on it and listening to the phrase and associated voice. Each participant was then randomly assigned to either a control group (with audio text nodes presented in a neutral voice only) or an experimental group (with prosodic attributes varied as described) and asked to identify six nodes selected at random from the tree. A Mann-Whitney test indicated that the participants performed significantly better using the two experimental trees where prosody was used for navigation when compared to the control tree ($p = .0001$ and $p = .005$ respectively). While the authors describe the hierarchy within the terms of a menu, one can also imagine applying the findings to other types of hierarchies such as document headings. One interesting observation from the authors was that the performance using the first hierarchical tree was better than the second because, "pitch is a better dimension than speech rate for identifying depth in a hierarchy… (p.

8).” This finding supports earlier findings by Xydas et al. (2005) that varying speed has the least effect on recognition and intelligibility.

In the second experiment, a complex hierarchical tree with three levels and 28 nodes was created using the same methodology. The authors added three more prosodic attributes of laryngealization, breathiness, and gain in volume. Laryngealization was defined as a “creaky voice” at the beginning and end of phrases or sentences. Breathiness was defined as adding “airy” sounds to the voice. Gain in volume was defined as giving certain words or phrases intensity by raising the volume of that element. Another condition was added to the experiment where two groups of participants were given different training on the system. One group was given training by the experimenter while the second group was given three minutes to learn the hierarchy on their own. Fourteen voices were randomly selected from the hierarchy, played for both groups, who were then asked to identify which node in the tree the voice represented on an unlabeled diagram. The overall recall rate for participants was 84%. Two nodes tested were not heard by the participants during training. The recall rate for these “unheard” nodes was 97%. The authors concluded that this was because the participants were able to understand the rules applied to the hierarchy and could predict where the nodes should be. The type of training was not found to be significant ($F = .043$, $df = 15$, $p = .8387$). The authors concluded from this finding that the hierarchical rules were easy to remember. It is notable that the authors did not examine the effects of different types of prosody other than their observation about the relative ineffectiveness of varying speed. The main finding of this research was that prosody can be used to orient a listener to an audio text with simple or complex structure.

Hjelmquist and Jansson (1990) analyzed the audio text reading behavior of 48 BVI persons between the ages of 23 and 79. Specifically, they studied incidence and type of keyboard navigational errors made while using a text-to-speech newspaper reading device. They compiled approximately 2,400 days of log data averaging 50 days per reader from the text-to-speech systems and analyzed them for patterns of errors and usage. Approximately 20% of all commands were used to read material sentence by sentence, 45% of the total commands were used to move to the next article, and 15% were navigational commands such as move forward, backward, up or down. They found that most errors were typing errors rather than issuing an incorrect command (syntactic error) and suggested that interface navigation should be as simple as possible to reduce user error. As unexciting as this finding was, there was one observation which was of interest to the present study. The incidence of a second error after a syntactic error was made (i.e., a navigational movement that was unintentional and therefore disorienting the subject), was very high. They concluded that once a subject becomes disoriented, it is difficult for them to find where they are. This type of circumstance would undoubtedly affect comprehension and would be interesting to confirm experimentally. The authors conclude that orientation within the text is crucial in the audio domain. Unfortunately, the authors did not examine comprehension as a variable. Linking this finding to others, designing an audio text interface that affords orientation markers such as audio icons to indicate location and command execution would have a positive effect on usability.

Sandweg, Hassenzahl, and Kuhn (2000) authored a case study of a telephone-based interface (TBI) designed to control a home automation system. The system used synthetic audio text as an output to the system. The audio text could be accessed at a keypad or via a phone interface. Input was done by entering touch-tone numbers on the keypad or phone. The auditory menus were created using seven guidelines. Each menu should have a title, body of options and an ending. The menus should be structured with more layers (depth) and less options (breadth) than typical visual menus use. This was to prevent overloading the user's memory. Options on the menus were ordered by decreasing frequency of use. Items followed a model of "function followed by action" such as, "For applications, press 6." Timeouts were built into the menus to allow users time to process responses. "Help" and "Cancel" commands were provided throughout so that users could get an overview of available commands from anywhere in the system. Finally, the system supported "dial ahead" so that experienced users could navigate the menu without waiting for the audio text menus to finish.

Six participants aged 20-56 participated in a Wizard of Oz methodology which used a researcher speaking in the place of the audio text menu while the participant pushed non-functional buttons on a keypad. Eight pre-programmed scenarios were given to the participants to "solve." The participants "talked aloud" their thoughts as they interacted with the system. The sessions were recorded. In this manner, the interface was mocked up as if it were live without the expense and time of building the system. The participants were then asked to rate the system on ergonomic quality, hedonic quality (i.e., non-task related quality), and appeal, and were also qualitatively assessed by observing their use of the system. A number of useful findings were reported in the source. The choice of menu wording was critical to its usability as there were several instances of confusion in use due to terminology. Menu order was also found to be important. In general, participants preferred to have the most common menu options presented first. The authors also found that it was important to give the participants a method of returning to the "home menu" when they became disoriented within a sub-menu. Their findings were in keeping with design recommendations from a number of authors (Halstead-Nussloch, 1989; Marics & Engelback, 1997; Paap & Cooke, 1997).

In Sodnik, Jakus, and Tomažic (2012), an audio text interface was created for BVI learners that provided spatial positioning of audio (as well as changes in pitch and tone) to indicate the format of text, such as bold, italics, indents and cell position within a table. Spatial positioning utilizes stereophonic audio to simulate a two-dimensional position of audio within the hearer's aural field. As such, spatial positioning of audio can be considered an extension of text prosody. Spatial positioning of audio was used to convey the typographical layout of the document. Response times to report the structure of the document were compared to a control group using conventional text-to-speech and refreshable Braille to accomplish the same structural reporting tasks. The researchers tested 12 BVI adults with a mean age of 33.4 years. The participants were selected because of their proficiency in using Braille, each having at least five years of experience, as well as familiarity with the JAWS screen-reading software and the ALVA refreshable Braille display. The two experimental conditions were the Braille condition consisting of the JAWS screenreader and the ALVA refreshable Braille display, and the spatialized speech condition which utilized the custom auditory interface being studied. The participants were given both conditions in random order and asked to read several text documents with a number of textual structures and formatting options including tables, various text alignments, and text attributes such as italic and bold. No instruction was given on the Braille interface since all participants were experienced JAWS and Braille users. General instructions were given on the spatialized speech interface, but the authors contended that the system was mainly intuitive to use. The participants were then asked to complete six tasks, three in each experimental condition. They were evaluated on task completion times, correctness text styles and alignment, correctness of tabular information, and a subjective evaluation of the interface. The authors found that processing the files was significantly faster in the spatialized speech interface when compared to the Braille interface ($F(1,70) = 391.5$, $p < .001$). Both interfaces resulted in comparable results for text attributes, but the spatialized speech interface demonstrated significantly more accuracy in helping to identify text alignment ($F(1,70) = 28.2$, $p < .001$). No significant difference was observed in identifying tabular structure ($F(1,70) = 1.045$, $p = .31$). Subjectively, the participants commented on the highly intuitive nature of the spatial positioning when describing text and tabular layouts.

A number of studies examined both speech- and non-speech-based audio orientation affordances. In Walker et al. (2013), the concept of spearcons is explored. Auditory icons are symbolic sounds in audio texts used to denote actions. For example, the sound of crumpling paper could symbolize deleting a file. Earcons are different in that they are not symbolic representations, but rather musically based embellishments. Earcons are frequently used as navigational markers. For example, when a rewind command is given to the interface, an earcon of a rising note could indicate that the text was being re-wound. Spearcons are essentially speeded-up text-to-speech representations of menu items. For example, “Save,” “Save As,” and “Save As Web Page” all have the same beginning sound but are different lengths and may not be recognizable as words. The utility of spearcons becomes apparent when a menu is translated into an audio interface. Spearcons may be created easily using the existing menu text and text-to-speech. No specialized audio files are necessary to create them.

In the first experiment, Walker et al. (2013) found that spearcons significantly reduce time-to-target menu when compared to auditory icons and earcons ($F(3, 24) = 177.14$, $p < .001$). None of the affordances were faster than text-to-speech alone indicating that, at least in menu design, there was no speed advantage. In the second experiment, spearcons resulted in users selecting menu items more accurately when compared to auditory icons (p = .041) and earcons (p = .038) In experiments three, four, and five, learning was gauged by using the experimental conditions in a series of tasks. The results indicate that spearcons significantly improved learning on the menu identification and navigation tasks, even though it lengthened the menu item audio label when compared with text-to-speech labels only.

Discussion

The first systematic literature review sought to answer the question, "What challenges to comprehension do adult blind or visually impaired learners face when using audio text?" The ten studies reviewed focused on several key issues related to audio text intelligibility and comprehension. The factor of intelligibility, or recognition of audio text at the surface level (to use Kintch's term), is equivalent to encoding in the printed text modality. As such, intelligibility can be considered to be a necessary condition for comprehension at the text base and situational model levels. One may apply Dunlosky and Rawson's (2005) levels of disruption hypothesis to assert that anything that disrupts intelligibility will decrease comprehension.

From the literature, we find that phonemic errors of recognition, prosodic elements, and the use of natural vs. synthetic audio text usage may affect intelligibility and comprehension. In Aryropolous et al. (2007) and Papadopolous et al. (2008), we find that phonemic errors of recognition such as paronyms can affect intelligibility at the word and sentence levels but seem to have little effect on comprehension at the passage level. Comparisons of BVI to sighted populations are the norm for these studies and none compare the effects of recognition errors on comprehension. Papadopolous et al. (2009) compare recognition errors in both natural and synthetic audio texts finding significant differences in word and sentence recognition between the two but no significant difference at the passage level indicating that synthetic audio text does not adversely affect comprehension when compared to natural audio text, at least in populations of BVI learners.

Using prosodic affordances such as varying audio text pitch, volume and speed, can effectively convey text attributes and aid in comprehension. Of the three most common prosodic affordances, speed demonstrates the least effect on text recognition. Both Xydas et al. (2005) and Argrylopolous et al. (2009) demonstrate that BVI populations are more effective at recognizing text attributes marked with prosodic affordances than their sighted peers.

Comparing natural to synthetic audio texts, Hjelmquist et al. (1992) found that memory and comprehension were significantly better in BVI populations when using natural audio texts rather than synthetic. Experience using synthetic audio texts attenuated this finding. Papadopolous et al. (2009) concluded that there was no difference in comprehension between natural and synthetic audio texts in BVI populations. While the Hjelmquist study provided a larger sample size compared to Papadopolous et al., there were significant differences in the synthetic audio text technology between the two and this may account for the difference.

Taken as a whole, the ten sources found addressing the first research question, i.e., what challenges to comprehension do adult blind or visually impaired learners face when using audio text, provide mixed and sporadic findings at best and tend to focus mainly on comparing BVI to sighted populations. None examine an experimental effect specifically designed to measure differences in comprehension due to affordances in BVI populations. Thus, a gap in literature was identified.

The second systematic literature review sought to answer the question, “What aspects of audio text interface design support learner comprehension?” After completing the systematic review procedure, only seven sources were identified. The search terms were updated to remove the BVI requirement. After applying the original screening criteria, two additional sources were added (Arons, 1997; Shajahan & Irani, 2007). In Arons (1997), a multitude of navigational affordances were designed and studied among a sighted population of participants. Participants found value in being able to select the level at which the system would navigate – by word, sentence, paragraph and section. They found high-speed skimming useful and were observed to use backward skimming to re-hear a passage. The concept of creating an audio bookmark was discussed by a number of participants. Reddy et al. (2005) also commented on the desirability of audio bookmarks. They went further by stating that navigational feedback indicating the current location in the text should always be available. Hjelmquist and Jansson (1990) came to this conclusion much earlier.

Prosodic markers such as audio icons, earcons and spearcons were also studied. Evidence from Walker et al. (2013) suggested that spearcons are most effective at helping learners recognize a variety of textual attributes and formats. Shajahan and Irani (2007) used prosodic markers to convey a hierarchical tree structure within the audio domain. Their study indicated that even a complex tree hierarchy can be more effectively recognized when using those prosodic markers to indicate position within the tree. Sandweg et al. (2000) qualitatively tested a number of design principles for increased usability and navigation. The most important for the current study was the affordance of giving the user a method of "returning home," that is, of resolving navigational disorientation. Sodnik et al. (2012) took navigational orientation to another level by creating an audio spatial positioning system used to convey text and table attributes effectively. It would seem that there are options for conveying some of the typographical attributes and structures effectively within the audio domain.

The two systematic literature reviews conducted suggest that a study of the effects of navigational affordances on comprehension of audio text of BVI learners could prove fruitful. Navigational affordances of audio text have been shown to affect comprehension of BVI and sighted learners in a variety of studies. A learner who demonstrates self-regulated comprehension may evaluate a comprehension failure when the information was comprehended but later irretrievable, the information was never fully comprehended and integrated, or the information was too difficult for the learner to comprehend. The learner may then regulate the failure by using metacognitive strategies to mitigate the failure. One mitigation strategy involves using navigational affordances to reprocess information. Properties and functions of navigational affordances of audio text have been studied, but mainly for sighted learners. No studies have treated the effects of navigational affordances on comprehension of BVI learners, and thus, there is a gap in the literature.

BVI learners have traditionally used Braille or audio media in learning. With the recent decrease in the use of Braille as a primary learning medium, a surge in audio learning materials, both natural and synthetic have increased. The ubiquity of systems to convert printed text to synthetic audio text encourages the use of such texts, and thus create ecological validity to a study utilizing synthetic audio texts. Comparative studies of natural and audio texts have had mixed results with text type serving as a confounding factor.

One significant finding of audio text is that prosody is indicated positively for both comprehension and learner experience. While the prosody of natural speech cannot be compared to that of synthetic speech (yet), great strides in synthetic speech engines have reduced the prosodic variability between the two media. While there is some indication that natural audio texts may be superior, the utility of synthetic speech for a variety of interfaces may outweigh this consideration.

Aspects of audio text interface design have been studied, but the vast majority of studies focus on non-disabled participants. Most research topics focus on properties and functions of navigation and/or orientation affordances and their efficacy. Text prosody as a property indicating navigation and orientation information to the learner has been found to be beneficial. Audio icons, earcons and spearcons have also been studied as affordances indicating navigation and/or orientation, to varying degrees of efficacy.

Research Conclusions

Taken as a whole, the literature reviews conducted indicate that research focusing on the effects of audio text affordances on comprehension, particularly among BVI learners, may prove fertile ground. Therefore, the purpose of this study is to examine the use and effects of linear navigational affordances on comprehension within the audio text domain, and to examine the role of motivation and metacognitive strategies in the use of navigational affordances.

To summarize the literature review and support this purpose, one can identify a number of salient points. Kintsch's theory (1998) proposed that comprehension is dependent upon the learner successfully processing information at the linguistic, meaning, and situation levels to fully understand the material. Dunlosky and Rawson (1997) extended this theory to assert that disruptions in comprehension affect the use of metacomprehension judgements and that disruptions at the situational level are highly predictive of comprehension. Hacker (1998) synthesized these findings into a theory of self-regulated learning in which failures in comprehension, if successfully monitored (i.e., judgements of learning; Connor, Dunlosky, & Herzog 1997), may result in the learner using a control strategy to correct the failure. Zabrucky and Moore (1994) stated that using affordances can be a control strategy resulting from successful monitoring of comprehension failures. Pressley and Afflerbach (1995) catalogued dozens of strategies to control comprehension failures. Among them are pausing during reading, rereading material, and selectively searching text to clarify meaning and increase comprehension. These affordances are the realm of audio navigational affordances.

Based upon these research findings, it would follow that learners listening to an informational audio text who successfully monitor failures in comprehension may control their learning through the use navigational affordances. And so, the first research question - do enhanced navigational affordances improve learners' comprehension of audio text – will be studied. But Gaver (1991) defines perceptible affordances as perceived environmental properties that determine the learner-environment interaction. In this vein, audio affordances can be considered perceptible if they are present in the learning environment and perceived by the learner. Gaver's distinction raises several interesting questions. Therefore, the first question can be decomposed: Does the presence of navigational affordances affect comprehension? Does the use of navigational affordances affect comprehension? Both questions will be studied.

Assuming that audio affordances affect comprehension, another question presents itself. What affects the use of audio navigational affordances? Returning to Hacker's theory (1998) of self-regulated learning, we found that when learners listen to audio text, comprehension is monitored at the metacognitive level. This is further supported by Kintsch and van Dijk (1978) who stated that successful monitoring results in regulation (i.e., control), which can take the form of using an affordance. Thus, successful metacognitive monitoring of comprehension is required to recognize a comprehension failure and implement a control strategy to mitigate it. This would indicate that metacognitive skills may affect the use of navigational affordances. And so, the second research question - what role do monitoring and control strategies play in the use of navigational affordances in audio text – will be studied.

Pintrich (2000) extended monitoring and control into his theory of self-regulated learning, where he further described metacognitive skills. Pintrich also asserted that motivational factors play a role in comprehension. He defined self-regulated learning as "…an active, constructive process whereby learners set goals for their learning and then attempt to monitor, regulate, and control their cognition, motivation, and behavior, guided and constrained by their goals and the contextual features in the environment" (p. 453). Thus, motivation may also play a role in the use of affordances as a controlling strategy. And so, the last question to be studied – what role do motivational factors play in the use of navigational affordances in audio text – comes into focus.

All the research referenced in this section so far was conducted with texts in the printed domain. From the second systematic literature review, we found that most of the audio text research, as it applied to comprehension, had been conducted in populations of BVI learners. From the reading rate studies (e.g., Foulke, 1962), we found that learning from audio with technology-based affordances is a skill to be learned. In general, BVI learners showed greater skill than sighted learners in this domain due to practice effect. Argyropoulos et al. (2007) confirmed this when they compared error recognition in the audio domain between sighted and BVI learners.

Regarding audio interface design considerations, Goudiras et al. found that navigation of audio text is easier with computer audio when compared to cassettes. Arons (1997) found that learners liked audio navigational affordances but reported difficulty when the affordances did not allow for orientation to the overall structure of the text. Morley et al. (1999) confirmed this issue of orientation. Reddy, Gupta and Karshmer (2005) identified two guidelines to audio text interface design. First, the user should have the ability receive orientation feedback at any time during the playback. Second, the user should have the affordances of not only hearing what they are navigating over or through, but also be able to define “anchors” that can be returned referenced at any time. This finding was similar to bookmarks in the Arons study.

Research on the equivalence of audio text to print text were inconclusive. For example, Papadopoulos et al (2009) found that comprehension was not significantly affected by the use of synthetic speech audio text when compared to natural speech. Yet Hjelmquist, Dahlstrand, and Hedelin (1992) found that memory and comprehension are significantly better when natural speech texts are used when compared to synthetic speech texts. Hjelmquist et al. (1992), Hansen, Lee, and Forer, (2002) as well as Argyropoulos et al. (2007) all remarked on the ecological validity of synthetic speech audio text in light of greater technology-mediated instruction in the future. These points have informed the design of an audio text interface.

CHAPTER III

METHODOLOGY

The present study utilized a method to measure audio navigational affordance use by participants listening to an informative audio text and correlate affordance use with comprehension. Pre-test questions measuring prior knowledge of the audio content as well as selected factors of self-regulation and motivation were given. A post-test measuring comprehension of the audio text material in addition to self-report questions querying the test conditions was also given.

Participants

The target population for the present study was college students. A convenience sample of students registered for an introductory political science course at Texas A&M University was studied. During the semester that data was collected, approximately 3,913 students were enrolled in this course. The pool of solicited students was approximately 1500 potential participants.

The study was administered via an Apple iOS mobile application. Based upon available data about students' use of wireless devices on campus (Sweeney, 2015), participants who own, or have access to an iOS device were calculated to be approximately 1,290 potential participants per 1500. The study sought to solicit approximately 3,000 participants, yielding a potential pool of approximately 2,580 participants. For the purposes of this study, 200-300 participants were desired to provide sufficient statistical power. The sample was self-selecting, and therefore not random. However, assignment to groups was randomly determined at runtime and should have resulted in approximately 20% of the sample being assigned to the control group, which, based upon the desired number of participants, should have been approximately 40 participants. The decision to assign the bulk of the students to the treatment group was based upon a greater desired sample size to provide sufficient statistical power for the metacognitive and motivational factor analyses with an expected number of approximately 160 participants. Because the assignment was truly random, the actual proportions of participants in each group were not exact.

Application Design and Data Collection

Since the use of navigational affordances was the focus of the study, it was designed to record all navigational responses from the participants. The study was administered via an application (app) designed specifically for this study, titled TactileTTS. The app was distributed via the Apple iTunes Store and could be freely installed by anyone within North America. The app was self-contained but connects the Qualtrics® survey engine to administer pre- and post-survey questions from within the app itself. This allows the researcher to utilize the powerful features of the Qualtrics® platform from within the tightly controlled environment of an iOS app.

The app was designed to be compatible on any iOS device running iOS 8.0 or newer. This includes all operating systems deployed by Apple since September of 2014. The app was developed in Apple's Swift programming language using the xCode development environment. A copy of the code base can be found at https://github.com/adaptiman/TactileTTS. The app utilized the built-in text-to-speech engine which has been standard in Apple operating systems for more than ten years. No other specialized software is needed. All components were taken from native Apple libraries. The app was designed to adjust its views (what few there were) to the device-specific display dimensions. Since the app is mainly an audio app, this feature is more to satisfy Apple design guidelines than to satisfy requirements of the study.

The app presented the participant with both survey questions and an audio interface that could be navigated with touch gestures as described. While some textual instruction was given during the orientation and training portion of the study, the only visual components available during the audio text portion were a graphic that reminded the participant of the five types of gestures available to navigate the text and a progress bar that indicated the relative position of the audio text at any given time. These were only available if the participant was assigned to the experimental group. Those assigned to the control group only had a tap/continue prompt and a progress bar during that portion of the study.

The app recorded participant metadata and all navigational gestures during the study, both in the training and the audio text portion of the study. Navigational gestures that were recorded include: tap (PC - pause/continue), swipe-left (F- forward by sentence), swipe right (B- back by sentence), two-finger swipe left (FP- forward by paragraph), and two-finger swipe right (BP- back by paragraph). For each gesture, the precise time and location within the text was recorded in the participant's log. For example, a log entry of "P,994,1454624457.34338" indicated that a tap gesture that paused the audio at character 994 of the text was executed on Thursday, February 4th, 2016 at exactly 22:20:57.48 Greenwich Mean Time. This time, which was recorded as a Unix timestamp, is accurate to the millisecond. Gestures performed during the training phase of the study were noted in the metadata log preceded with a "T". This created a method to discriminate between navigations performed during the training portion of the protocol and the audio text portion. Also recorded was the participant's uniquely generated identifier, the randomly assigned group, and the number of trials the participant completed.

Instructional Material

For this study, the instructional material was an informational text consisting of approximately 2,400 words on the topic of "The Executive Branch of State Governments" taken from the eleventh chapter of the textbook *Comparing the States and Communities* (Tucker & Luttbeg, 2015), which is used in the introductory political science course from which the participants were drawn (see Appendix C.1). This specific section was selected because it was presented at the end of the course, which made it less likely that participants would have read it since the experiment took place during the middle part of the semester. The material is conceptually rich with many hierarchical connections throughout. At normal conversational speed (approximately 160 words per minute), the text took approximately 17 minutes to go through. Within the experimental group, this could be extended to approximately 19-21 minutes to account for the use of navigational affordances.

Instruments

The Motivated Strategies for Learning Questionnaire (MSLQ) (Pintrich, 1991) is a well-known instrument for measuring metacognitive learning and self-regulation skills. Selected scales from this measure were used to assess participants' skills in these areas. To measure comprehension, questions were taken from tests given by the faculty member who authored the text material used in the study. Some questions were used from previous tests given by the instructor while others were created based upon the text material by the researcher.

The Motivated Strategies for Learning Questionnaire

The MSLQ, which has been used with populations of college-age students for a variety of subjects, assesses metacognitive skills (i.e., learning strategies) and motivation, and has been found to correlate reasonably well with academic success (e.g., Pintrich & Smith, 1993). The 31 items composing the learning strategies section of the instrument examine cognitive and metacognitive strategies including rehearsal, metacognitive self-regulation, and effort regulation. Credé and Phillips (2011) completed a large meta-analysis of the MSLQ and found moderate to strong relationships between academic success as measured by class grades and effort regulation ($\rho = 0.41$). Weaker relationships were found between academic success and metacognitive self-regulation ($\rho = 0.23$), rehearsal ($\rho = -0.18$), intrinsic goal orientation ($\rho = 0.20$), and extrinsic goal orientation ($\rho = 0.10$). The rehearsal learning strategy, which is hypothesized to be correlated with the use of navigational affordances for obvious reasons, has a low correlation ($\rho = 0.12$). Credé and Phillips hypothesize that this may be the result of a curvilinear relationship with high self-regulating participants comprehending material the first time without the need for rehearsal, while low self-regulating participants don't sufficiently monitor comprehension and thus do not use the strategy. The 31 items composing the motivation scale examine students' goals and value beliefs, ability to succeed, and anxiety about the course. The author states in the instrument manual that selected scales can be used to measure independently of each other.

Comprehension questions

Comprehension questions were provided by Dr. Harvey Tucker, one of the authors of the text material (Tucker & Luttbeg, 2015). Some of the provided questions had been used in previous years' tests and Dr. Tucker had collected reliability scores for these questions. These scores ranged from -0.01 to 0.42 ($n = 13$). While these questions were taken from the same chapter as the text object, they did not cover the section of the text ultimately used in the experiment. As a result, no reliability data for the final question set were available. New difficulty and reliability data were generated from participant responses (cf. Table 1 and Table 2).

Ten questions were selected as the pre-test comprehension set (Appendix C.4). The ten pre-test questions queried participants on material from the same chapter as the experimental text material but were not answered by experimental text material. This meant that the participant would not gain any information from the pre-test questions that would be useful in answering the post-test questions.

Treatment Conditions

A treatment condition of enhanced audio navigation affordances (experimental) versus basic navigational affordances (control) was used. Operationally, those participants in the experimental group had access to pause and continue as well as backward and forward navigation by sentence and paragraph, while the control group only had access to pause and continue.

Data Sources

In order to investigate whether the *presence* of enhanced affordances improved participants' comprehension of audio text (RQ1-A), the independent variable was group assignment, either control or experimental. The dependent variable was comprehension. The datum source for this variable were ten pre-test multiple choice questions entered as a covariate, and twenty post-test multiple choice comprehension questions over the audio text. Item analyses was completed on both the pre-and post-test question banks, and the resulting sets used for further analyses. The pre-test questions were used to control for prior knowledge as a covariate for RQ1-A. The post-test questions were used to operationally measure comprehension of the audio text. Once again, RQ1-A asks if the *presence* of enhanced affordances improve participants' comprehension of audio text.

In order to investigate whether the *use* of enhanced affordances improved participants' comprehension of audio text (RQ1-B), the independent variable was total affordance use. The data source for this variable was affordance use as measured by total navigations demonstrated during the study, excluding training navigations. Only participants assigned to the experimental condition were used. The dependent variable was comprehension as previously described. Again, RQ1-B asks if the *use* of enhanced affordances improve participants' comprehension of audio text.

Data representing independent variables for effects of monitoring and control strategies on affordance use (RQ2) and effect of motivational factors on affordance use (RQ3) were collected during the survey phase of the study prior to listening to the audio text. These questions were taken directly from the MSLQ (cf. Appendix C.3). Five scales representing both metacognitive skills and motivation were collected. Two scales were used to represent metacognitive skill – self-regulation (SR) and rehearsal (RE). The SR scale consisted of 12 questions, specifically 33, 36, 41, 44, 54, 55, 56, 57, 61, 76, 78, and 79. The RE scale consisted of four questions, specifically 39, 46, 59, and 72. Motivation was represented by three scales – interest (I), test anxiety (TA), and effort regulation (ER). The I scale consisted of eight questions, specifically 1, 16, 22, 24, 7, 11, 13, and 30. The TA scale consisted of five questions, specifically 3, 8, 14, 19, and 28. The ER scale consisted of four questions, specifically 37, 48, 60, and 74.

For both RQ2 and RQ3, the dependent variable was total affordance use. The data source for this variable was affordance use as measured by total navigations demonstrated during the study. Only participants assigned to the experimental condition were used.

Four control questions were administered after the audio text was given. These prompted the participant to disclose distractions during the test and asked whether the participant used notes during the study or had taken it on another device, thus potentially confounding the results for that participant.

Procedure

Students within the target population were solicited to join the study by a direct campaign supported by distribution of a quarter sheet flyer handed out during classes which explained the study. The researcher attended sections of an introductory political science course (POLS 207) to distribute the flyer and explained the nature of the study emphasizing that participation was voluntary and would not affect the student's grade in the course. For those students who were ineligible to join the study because they did not own an iOS device, information was also given in the flyer referencing a website where an audio text of the study material could be accessed. While navigational affordances would not be available in the web-based copy of the material, this option presented the material in an equivalent manner as the control group participants.

The researcher also explained that those who participated would be eligible for an incentive drawing at the conclusion of the study. The advertisement flyer contained a QR code that, when scanned by an iOS mobile device, would launch the Apple App Store and load the TactileTTS application for installation. The app is installed in the same manner as any other application, a procedure the participants are certainly familiar with. The time and location that the app was installed would be up to the participants, and not controllable. Neither was the time and place the study was undertaken. However, instructions prompted the participants to complete the study in a quiet environment away from distractions.

Figure 4 details the flow of the study as it traverses the various stages of the app. Phase one of the study was administered via the app but was controlled entirely within the Qualtrics® survey engine. Phase one consisted of informed consent, demographic questions, MSLQ questions, and comprehension pre-test questions. Once these, all of which are required, were answered, control was passed back to the app for phase two. Phase two was administered entirely within the app and consisted of written and verbal instructions for the study, a navigational gesture tutorial, and the audio text segment of the study. During phase three, participant and navigational data were passed back to Qualtrics®. In this phase, post-test comprehension questions and environmental questions were given. Upon completion, the participant was sent to a separate web page outside of the Qualtrics® survey engine to optionally enter the prize drawing.

Phase one

After installation and launch of the TactileTTS app, participants were randomly assigned to one of two treatments by the app; those who received enhanced navigational affordances (experimental group) and those who received basic navigational affordances (control group). Participants assigned to the control group were only provided basic pause and continue affordances. Participants assigned to the experimental group were given enhanced navigational affordances including pause and continue, go forward by sentence or paragraph, and go backward by sentence or paragraph. A unique participant identifier (i.e., GUID) was also generated during the group assignment to anonymously identify responses, although no information was collected that could uniquely identify the participant.

Upon loading the first page of the Qualtrics® survey, participant metadata were written to Qualtrics® from the app so that gesture data could be collected later. The app then loaded an informed consent statement derived from the current institutional review board template. The participant must have clicked the continue button to continue with the study. If the continue button was not clicked, the app would not load the rest of the survey or initialize the audio text portion of the study. This allowed participants to opt out of the study if they so chose and allowed the researcher to document affirmative informed consent.

Once the participant agreed to continue with the study, phase one of the survey (see Figure 4) was loaded. The survey questions provided by the Qualtrics® survey engine were presented in a web browser window from within the app that did not allow the participant to access browser controls, so that he or she must follow the navigation afforded in the survey to continue. Sections b, c, and d (cf. Figure 4) consisted of two demographic questions, questions comprising the scales of interest from the MSLQ, and ten pre-test comprehension questions. The two demographic questions asked the participant's age and academic classification and were collected to provide basic statistics that described the sample. The survey prompted for an additional question if the participant indicated they were younger than 18 years old. If they answered affirmatively, the app forwarded them to an informational web page which explained that they must be at least 18 years old to participate and could not give informed consent. The app is constructed in such a way that the study did not continue and the survey was terminated.

Next, MSLQ questions comprising 17 motivation scale questions (I, TA, ER), 12 monitoring and control questions (SR), and four control strategy questions (RE) were given using a seven-point Likert scale, which was specified in the instrument. All of the questions were required to be answered before the participant was allowed to proceed to the next set of questions. The order of the questions was randomized. Next, ten pre-test comprehension questions were given. These questions asked about knowledge contained in the same chapter as the audio text, but not over the same material. That is, the pre-test questions were on the same topic but are not covered by the audio text. These served as a covariate to control for prior knowledge. Both the order of the questions and the answers were randomized.

Phase two

Upon completing section one survey questions, both the control and experimental participants were given instructions about how to complete the study (see Figure 4). The text was given both visually and aurally. The instructions were context-specific for the participant's assigned group. Participants in both groups were taken to a training screen which explained the swipe gestures used to access the navigational affordances. The user was given an opportunity to test the navigational swipes prior to continuing the study (see Figure 4). The type and number of swipe gestures performed during training was recorded and saved in the result set.

After completing training, participants began the audio text portion of the study. A visual graphic of available navigational affordances was displayed and customized to the participant's group. Control group participants were given a visual navigational affordance graphic that included only the tap (pause and continue) gesture, while participants in the experimental group were additionally given back- and forward- by sentence and paragraph swipe gestures. A visual percent completion indicator was also displayed for both groups to give the participant a reference to how much of the text they had completed or had yet to complete.

During the audio text portion, all navigational gestures were recorded in the participant data set. For each gesture, the time and location of the gesture were recorded. Other types of data collected included whether or not the app was put in the background, and when it returned to the foreground (see Figure 5). These data indicated if another mobile app (e.g., the phone app) interrupted the study. While the participant was asked about interruptions at the end of the study, these data could confirm if a significant interruption occurred during the study, although it did not record the type of interruption.

Phase three

Upon completion of the audio text portion of the study, the app loaded the phase three survey which consisted of a set of comprehension questions about the text just heard (see Figure 4). The respondent data set from the audio text portion of the study was uploaded into Qualtrics® as soon as the phase three survey was loaded, thus conserving data even if the subject quit before answering any comprehension questions. The comprehension questions were taken from test questions related to the political science chapter just heard. Both the order of the questions and the answers were randomized.

After answering the phase three questions, the participant was asked four questions about the study environment related to potential distractions that may have occurred (see Figure 4). If no distractions occurred, there was a choice to indicate that as well. The question was phrased as types of distractions rather than number. Multiple types of distractions could be selected. The participant was also asked about whether they used a classmate's device to take the study prior to the current attempt, as well as whether or not they took any kind of notes. They were also asked if they encountered any problems with the app such as a freeze or crash. This condition was also measured by the app itself and sent with the response data.

Once environmental questions were completed, the participant was redirected to a separate form that explained how to enter the incentive drawing (see Figure 4). This form asked only for the participant's email. The form saved the incentive drawing data in a separate data set from the participant's response data, which could not be tied back to the participant's answers. This was done to insure anonymity of the response data. While the responses were anonymous, it is conceivable that a participant could be identified by their email address, thus the study cannot claim to be confidential.

CHAPTER IV

RESULTS

An initial examination of the data yielded an interesting observation related to a pattern of affordance use. There were a number of participants who "skipped" through the text by simply navigating to the end of the text quickly. These "skippers" were identified by a pattern of at least ten consecutive forward navigations (either by sentence or paragraph) in quick succession. This behavior indicated that the participant did not spend sufficient time to listen to the text. The criterion of ten consecutive forward navigations within five seconds was identified by comparing comprehension means of "skippers" versus "non-skippers" using a variety of time and frequency variables until the largest mean difference was found (Table 3). After identifying "skippers", it would follow that their comprehension scores would demonstrate their lack of comprehension of the material. Indeed, an ANOVA with a dependent variable of comprehension and an independent variable of "skipper" yields a difference, though not significant ($F = 1.81$, $\alpha = 0.181$). Means and standard deviations for the skippers and non-skippers were $\bar{X} = 8.95$, $SD = 2.62$ and $\bar{X} = 9.85$, $SD = 3.08$ respectively. The effect size as expressed by Glass' delta was $d = 0.38$. Levene's test of homogeneity of variances was not significant ($\alpha = 0.423$).

Item Analyses of Pre- and Post-Test Comprehension Items

An item analysis was conducted on both the pre- and post-test question sets. The "skippers" were excluded from these analyses. Since they did not utilize affordances in the intended manner, their answers to the pre- and post-test questions could also be suspect. For the pre-test questions, the reliability statistics were generated from 108 valid cases. Based upon these data (see Table 1), question Q10 was eliminated from the pre-test set, yielding an improved

Kuder-Richardson statistic of 0.339 (n = 9). The same analysis was performed on the post-test data set (see Table 2). For the post-test questions, the reliability statistics were generated from 88 valid cases. From this data set, questions Q4 and Q13 were eliminated for negative correlation. Question Q4 also demonstrated high difficulty. Additionally, questions Q2, Q7, Q8, and Q15 were eliminated to improve the overall reliability of the instrument. This yielded an improved KR20 statistic of 0.625 (n = 14). Even though the reliability of the test questions improved, the overall research question statistics did not significantly change. Therefore, the reduced test banks identified during the item analyses will be used.

Descriptive Statistics of the Dataset

A total of 174 participants responded to the study. Among these, 129 (92.1%) reported being between 18 and 24 years of age. Regarding academic classifications, 85 (48.9%) reported junior classification, while 29 (16.7%) reported senior classification. Table 4 contains a complete breakdown of demographic data collected.

Study mortality consisted of 64 cases out of the 174 total participants. For the 64 cases eliminated, all were a result of the participants terminating their participation in the study or experiencing a significant interruption which resulted in a delay in completing the study. Sixteen participants voluntarily terminated their participation during some portion of the study. Participants involuntarily terminated (i.e, by the app) amounted to 49. This could result from a variety of circumstances such as a phone call or other application interruption, or a significant pause in the protocol resulting in the device going to sleep. Recall that the TactileTTS app was designed to detect such interruptions. A breakdown of study mortality during the different phases of the study is contained in Table 5. Notably, 18 participants were eliminated during the audio protocol, nine each from the control experimental groups. This represented a 40.9% mortality for the control group while the mortality in the experimental group represented 8.3%.

Examining “skippers,” 21 participants were eliminated leaving 153 potential cases for study. After removing “skippers” and incomplete cases from the dataset, 88 cases were left, consisting of 77 who received enhanced navigational affordances (i.e., experimental group) and 11 who received basic navigational affordances (i.e., control group). A case was considered complete if both the pre-test and post-test data were available. Descriptive statistics for the remaining 88 cases are shown in Table 6. This included statistics for the comprehension measures, both before and after the item analyses, as well as the metacognitive and motivational factors.

RQ1-A: The Effect of the Presence of Navigational Affordances on Participants' Comprehension

Do enhanced navigational affordances improve participants' comprehension of audio text? To answer this question, the data were examined to see if there was a significant difference in comprehension (as measured by post-test questions) between the control and experimental groups when prior knowledge of the subject (as measured by the pre-test questions) is controlled for.

The first analysis performed was correlation without regard to covariates. A Levene's Test was performed and found no statistically significant difference in the error variances (α = 0.33). When comparing the means of the experimental group (those who had access to enhanced navigational controls) to the control group (those who only had basic navigational controls), an F = 0.25 and α = 0.62 (for the experimental group) indicated no significant difference between the groups with an effect size (estimated with Glass's delta) of d = 0.17.

Subsequent to the first analysis, the interaction between the group assignment and pre-test scores was examined. A Levene's Test failed to detect any statistically significant differences in variances, $\alpha = 0.08$, and a group assignment by pre-test interaction model with an $F = 2.83$ and $\alpha = 0.10$, thus indicating potential but non-significant interactions with the covariate. Subsequent to this analysis, the analysis of co-variance was performed with group assignment as the independent variable, post-test as the dependent variable, and pre-test as the covariate. This analysis produced a Levene's Test $\alpha = 0.19$, $F = 0.32$ and $\alpha = 0.57$ (see Table 7). Means and standard deviations for the control and experimental groups were $\bar{X} = 7.08$, $SD = 2.47$ and $\bar{X} = 7.50$, $SD = 2.82$ respectively. The effect size was $d = 0.17$. These data do not support a statistically significant difference in comprehension between the control and experimental groups.

RQ1-B: The Effect of the Use of Navigational Affordances on Participants' Comprehension

Does the use of navigational affordances affect comprehension? Recall that Credé and Phillips (2011) hypothesized that learning strategies may have a curvilinear relationship to academic performance because highly skilled learners don't need the strategy and low skilled learners don't realize they need the strategy. Assuming comprehension to be related to academic performance, this would suggest that the use of the learning strategy of navigational affordances may demonstrate a curvilinear relationship to measured comprehension. To determine if the use of enhanced navigational affordances affects comprehension, a curvilinear regression was performed by squaring the total navigations for each case and recoding the data into a new variable. For the purposes of this analysis, only the experimental group of participants, who had access to enhanced navigational affordances, was used. A hierarchical multiple regression was then performed with the post-test comprehension score (with item analysis applied) as the dependent variable, total navigations as the independent step-one variable, and total navigations squared as the independent step-two variable. The ANOVA was significant (F = 4.065, α = 0.022); therefore, these data support a relationship between comprehension and the use of navigational affordances (see Tables 8 and 9). Furthermore, the contribution of the curvilinear portion of the model (R^2 = 0.107) was much larger than the linear portion (R^2 = 0.024). This would indicate affordance use does predict comprehension, and the usage pattern may be quadratic across the comprehension continuum (see Figure 6). To further illustrate, total navigational use appears to be low in low- and high-comprehension participants, but moderate to high in average-comprehension participants for the dataset used.

RQ2: The Effects of Monitoring and Control Strategies on the Use of Navigational Affordances

What role do monitoring and control strategies play in the use of navigational affordances in audio text? Having been informed of the curvilinear nature of the relationship between affordance use and comprehension in the last research question, A multivariate ANOVA was performed using the MSLQ scales of metacognitive self-regulation (SR) and rehearsal (RE) as predictor variables acting upon the dependent variable of affordance use (A) as measured by total navigations as the independent step one variable, and total navigations squared as the independent step two variable. Interactions between the independent variables were also examined. For the purposes of this analysis, only the experimental group of participants, who had access to enhanced navigational affordances was used. The model demonstrated no significance for any of the factors, including interaction effects (see Table 10). Between subjects' effects were also non-significant (see Table 11). Surprisingly, self-regulation (SR) did not make a significant contribution to the model (β = 0.481, α = 0.62). As expected, rehearsal and self-regulation were moderately correlated ($\rho = 0.35$). As a check, a regression of the metacognitive factors with comprehension was also performed. The model did not demonstrate significant predictability with an $F = 1.66$ and $\alpha = 0.20$. Beta weights for the two factors were 0.15 for self-regulation, and -0.20 for rehearsal.

RQ3: The Effects of Motivational Factors on the Use of Navigational Affordances

What role do motivational factors play in the use of navigational affordances in audio text? Keeping in mind the potential curvilinear nature of affordance use, a multivariate ANOVA was performed using the MSLQ scales related to motivation including interest (I), test anxiety (TA) and effort regulation (ER) as predictor variables acting upon the dependent variable of affordance use (A) as measured by total navigations as the independent step one variable, and total navigations squared as the independent step two variable. Interactions between the independent variables were also examined. For the purposes of this analysis, only the experimental group of participants, who had access to enhanced navigational affordances was used. Cases who met the definition of “skipper” were excluded from the analysis since they could not be said to have utilized affordances in the intended manner. The model demonstrated no significance ($F = .124$, $\alpha = 0.88$) for any of the factors, including interaction affects (see Table 10). Between subjects’ effects were also non-significant (see Table 12).

To test the relationship between motivation and comprehension without affordance use as a moderating variable, a regression was performed using the three motivational measures as independent variables and post-test comprehension as the dependent variable. Not surprisingly, motivational factors did positively correlate with comprehension with an $F = 5.49$ and $\alpha = 0.002$. Beta weights for the three factors were 0.29 for interest, -0.35 for test anxiety, and 0.003 for motivation.

Distractions During the Study

Significant numbers of participants reported distractions during the course of the study. A total of 88 participants were not considered “skippers” and contributed both the pre- and post-test data. This includes participants from both the control and experimental groups. Referring to Table 13, 108 distraction types were reported among the 88 total participants, yielding an average of 1.24 distraction types per participant. Since types of distractions were queried rather than individual occurrences of distractions, this number does not reflect the actual number of distractions that occurred. This number is most likely higher. The distraction with the greatest proportion among the sample was a “person” interrupting. Nine participants encountered a delay of more than one minute, which is an interesting finding since the TactileTTS app is designed to eliminate participants when the app is put in the “background.” These nine participants must have kept their device from going to sleep manually, probably by touching the screen every 10 seconds or so.

From the list of 22 “other” distractions (see Table 14), one sees several patterns emerge including common distractions such as people, television-watching, and pets (41%), and boredom with the technology or the nature of audio material vs. print text (36%). One participant commented that the “reading was too fast” while another remarked that the “voice talked too slow and it was hard to pay attention.” In general, these data indicate a high degree of distractions encountered by participants in the study.

CHAPTER V

DISCUSSION AND CONCLUSIONS

Factors affecting the comprehension within the audio domain is an increasingly important topic to study because of the proliferation of natural speech audio texts as well as improvements in synthetic text-to-speech technology. While many studies have examined factors affecting comprehension of print texts, few have examined factors affecting comprehension of audio texts and fewer still the effects of specific moderating variables, such as affordance use. Affordances provide opportunities for learners to interact with instructional media. The literature has mainly described affordance effects on comprehension of written text, but little research has been done with audio text, and few studies have specifically described the effects of navigational affordances on comprehension of audio text. Further, intrinsic factors such as metacognitive skills and motivation also have been demonstrated to affect comprehension of written text. However, once again, few studies have tied these factors to the use of affordances with audio text and their relationship to comprehension. This study sought to explore the relationship of comprehension to the use of affordances, metacomprehension factors, and motivational factors within the audio domain.

The purpose of this study was to examine the use and effects of linear navigational affordances on comprehension within the audio text domain and the role of motivation and metacognitive strategies in the use of navigational affordances.

Presence of Navigational Affordances

When comparing the control and experimental groups of participants, there was no significant difference in comprehension based upon the *presence* of the affordances provided within the context of this study. It is notable that the sample size of the control group was smaller than desired, mainly because of study mortality. This decreased the statistical power of the result. Recall that the TactileTTS app randomly assigned participants at runtime to either of the two groups. A ratio of four to one, experimental to control, was selected to provide a sufficient number of cases for the second and third research questions. Overall, the number of control cases was very small, and may have affected the outcome of this question. Even so, the finding makes sense in light of the reviewed literature that mainly supports learner-centric factors affecting affordance use such as metacognitive skill and motivational factors rather than the design of the interface (e.g., Zabrucky & Moore, 1994). The author was unable to locate a finding in the literature in which the effects of the presence of an affordance on comprehension were studied. Thus, this finding may be a contribution to the literature. The assertion that the presence of an interface control (i.e., affordance) had no effect on learning is reminiscent of the Clark-Kozma debate (Clark, 1983; Clark, 1994; Kozma, 1991; Kozma, 1994) in which Richard Clark asserted that media do not influence learning. Richard Kozma countered that it was the interaction of the learner with the medium, content, and learning-environment that led to learning.

Use of Navigational Affordances

While the effect of the *presence* of navigational affordances on comprehension didn't yield a significant result, the *use* of navigational affordances did ($F = 4.065$, $\alpha = 0.022$). Moreover, the pattern of use of affordances supported a curvilinear, or more specifically a quadratic, usage pattern. The linear contribution to the regression accounted for 2.4% of the variance, while the quadratic portion accounted for 10.7% of the variance. Given Credé and Phillips' hypothesis (2011) that the use of learning strategies may have been undetected in previous studies because it is actually curvilinear, their hypothesis is supported in the present study. A similar inference can be drawn from Azevedo's findings (2005). Indeed, had the curvilinear portion of the regression not been added to the model, the result would not have been significant. Upon examination of Figure 6, an interesting observation is apparent. While those participants with moderate levels of affordance use achieved greater comprehension, the slope of the linear best-fit line is slightly positive across comprehension scores. This would indicate a slightly greater use of navigational affordances by participants with high comprehension compared to those with low comprehension.

Effects of Metacognitive Strategies on Affordance Use

Surprisingly, neither metacognitive self-regulation (SR) nor rehearsal (RE) was found to be a significant predictor of affordance use within the parameters of this study. Given the body of literature suggesting otherwise, one must ask why. Several explanations come to mind.

Uncontrolled variance of conducting the study "in the wild" could have attenuated the effects studied. Distractions in the environment could have confounded effects of metacognitive skills by preventing the recognition of a comprehension failure and subsequent use of navigational affordances to correct it. But the tradeoff between environmental control and the ecological validity of the methodology employed must be considered. Another possibility is that the MSLQ was not a good or appropriate measure for the construct under study. The MSLQ was selected as the instrument because it is widely used as a metacognitive skill and motivation measure for the population under study. However, the correlations reported in the MSLQ manual for the different scales were reported for the final class grade, not the construct of comprehension. And so, the reported correlations with academic performance may not necessarily correlate with comprehension. With regard to rehearsal, the four questions deal with repetition of material – repeating text material, re-reading class notes, memorizing key words, and making lists. None of these are directly related to using a navigational affordance to "re-hear" material, although the current study hypothesized that these behaviors would be related. One may also remark that the four questions comprising the rehearsal scale more broadly refer to study skills at a global level rather than during situational learning such as the act of reading or listening to a passage and comprehending the material. In this regard, the measure may not be appropriate.

Since metacognitive factors were not predictive of comprehension, this would not refute Hacker's model since affordance use could be a mediating factor. One cannot rule out unknown or unidentified factors that could also play a role. And so, the role that affordance use plays as a mediator between metacognitive skills and comprehension needs to be investigated further, controlling for distractions and refining the metacognitive measure.

Effects of Motivational Factors on Affordance Use

The motivational factors of interest (I), test anxiety (TA) and motivation (ER) as measured by the MSLQ were not found to be significant predictors of affordance use within the parameters of this study. Again, one may hypothesize that distractions may have played a role in confounding the use of affordances by preventing the learner from recognizing comprehension failures and thus having an opportunity to correct them. However, since motivational factors were predictive of comprehension, this would indicate that affordance use may not be a mediating factor between motivation and comprehension. Since motivation is not a factor of Hacker's model, the lack of predictability of affordance use would not refute the model. For the purposes of this study, motivational factors were investigated as a secondary concern. The findings would indicate that motivational factors should be left out of a monitoring and control model of audio comprehension, although these factors do seem to affect comprehension.

Distractions

Distractions were significant and mainly uncontrolled. The strongest control on distractions was the automatic termination of the study by the app for those participants who had significant delays in completing the study. This app feature eliminated data and in hindsight, was a poor design choice. A better strategy would have been to note the delay and continue to collect data. Even so, distractions may have had a significant effect on the outcomes of the study. The only participant control in place for distractions was a statement in the study instructions to "[be] sure you have a quiet place that is free from distractions to take the study." Yet the data support a high degree of distractions recorded by the participants (see Tables 13 and 14). Certainly, the environmental conditions contributed to the degree of distractions, but one may ask if the environment is more exemplary of real study environments for learners in our Internet-connected, "always on" world.

Distractions in the learning environment may have ecological validity for "real world" learning situations, but the tradeoff is lower internal validity offered by more controlled studies in less realistic environments. Nonetheless, the use of a data collection platform in the natural learning environment of the learner could prove fruitful for future studies if more specific data can be gathered about the distractions present in the environment. In contrast, much of the data collected on reading comprehension is based upon studies in a controlled environment. As Wood and her colleagues recently pointed out, "It is important to have randomized treatment-controlled studies to make causal inferences about the effect of text-to-speech/read-aloud tools in an educational context, thus increasing ecological validity by capturing the diversity of the current classroom. However, there are trade-offs between internal and external validity that are difficult to balance " (Wood et al., 2018, p. 80).

Even though some studies have introduced distractions into the experimental design, this is a different proposition from distractions as they naturally occur in the environment. It begs the question as to whether print text comprehension in the natural environment of a college-age learner is also affected by environmental distractions. How different is the *real* learning environment from the test environment?

In future studies, several available technologies could be used to collect better data on distractions. Mobile platform capabilities such as GPS, accelerometers, microphones and cameras could be leveraged to collect environmental data to describe distractions in the environment more clearly and completely. While this study did not go into a detailed analysis of distractions, the results would indicate that this could be a valuable path of research.

Mobile App Methodology

Considering the study app, while several commercial apps were researched, and one publisher was approached for collaboration, in the end, a custom app had to be produced. The author had significant difficulty in producing the app with the entire programming process requiring more than a year. Even with these difficulties, the final product is one that can be re-used for other purposes and with different experimental parameters. The author considered that the difficulty of producing the app would be outweighed by the innovative study methodology with its inherent potential to reach a far greater sample of participants than more conventional laboratory-based methodologies. Researcher-monitored testing environments are limited by location, whereas a mobile app can be distributed quickly over a geographic area beyond the laboratory. Since many participants may use a research app simultaneously, the methodology has the potential to reach a larger population more quickly than conventional methodologies.

This coupled with the increased ecological validity of using an app in a manner and environment similar to what college-age learners use today would be valuable. It is interesting to hypothesize if the mortality rate would've been as a high had the participants been in an observed, more controlled environment. One point of weakness was the production of the app for only the iOS platform. While Android has a significant installed base among the population of interest, there were simply neither the tools nor time to produce and support a second app.

An additional contribution to the design methodology of the TactileTTS app was the use of the Qualtrics® survey engine to take care of the data collection of the study from within the distributed app. This strategy proved to be efficient and useful utilizing the ecological value of an app with the power of a full-featured survey platform. This technique serves as a recommendation for future app studies.

Skippers

The participants previously identified as skippers presented a quandary in the data. On one hand, these participants used the navigational affordances in a manner which logically hindered and/or prevented them from accessing at least a portion of the audio text, bringing their results in question. On the other hand, the behavior in question was demonstrated on a continuum of two variables – frequency and rate – thus making identification problematic. Initially, several more traditional methods of outlier identification were considered, including the outlier labeling rule. But these methods assume a normal distribution, which the use of navigational affordances is not. Another method had to be found. Assuming a priori that the behavior of skippers would affect comprehension negatively, a series of variable sets with numbers of consecutive forward navigations ranging from six to ten within a time-frame of three to eight seconds was constructed. Interestingly, while the final ANOVA was not significant (F = 1.81, α = 0.181), an earlier analysis *had* yielded a significant result prior to adding the last group of participant data consisting of an additional 23 participants who completed the study in the spring of 2018 semester. Five of the 23 new participants met the definition of a skipper. Prior participants completed the study in the fall 2017 semester. Why these participants decreased the mean difference between the groups is unknown.

Application of Findings to the Hacker Model

Reviewing Hacker's model, the current study does not refute it, and in one way supports it. The lack of metacognitive skills to predict affordance use does not refute the model, although further investigation needs to be done in this area. The potential negative impacts of distractions on comprehension must be controlled and further investigated. Motivational factors were not included in the model, and the findings that motivation positively predicts comprehension but not affordance use would continue to suggest its exclusion from the model. The positive effect of affordance use on comprehension supports the model in its role as a strategy to control and affect comprehension.

But a clearer description of comprehension monitoring mechanisms in the audio domain must be developed. Why do some students use navigational affordances and others do not?

Threats to Validity – Internal and External

The most significant internal threat to validity was the unmonitored condition of the study environment. While variance could have been more controlled in a laboratory setting, the ecological validity of conducting the study in the field was balanced against a more controlled environment. Since the study solicitation was made to students using the text material in their course work, the results provided greater generalizability to the population studied.

Experimental mortality was also a significant internal threat. The high number of participants that either quit during the study or were eliminated by the app reduced the studied population significantly. As noted earlier, a much larger proportion of participants in the control group (40.8%) were eliminated when compared to the experimental group (8.3%). This five-fold difference greatly decreased the control group sample size and most certainly had a large impact on the power of the study. The larger mortality in the control group may have been a result of boredom with the lack of navigational affordances and the resulting lack of user control over the material, an idea which has some support based upon distraction responses, but this is speculative and would have to be tested. Changes in the study platform could help mitigate some of this, but if a natural learning environment continues to be a focus of this line research, one can only describe and control.

The identification of "skippers" utilized a procedure not based upon a statistical method of identifying outliers. As such, the appropriateness of the procedure is questionable. While it follows that a learner who doesn't listen to an audio text may not learn it, how are those who don't listen determined? Since skipping behavior seems to be demonstrated on a continuum, at what point does a learner move from "comprehending" to "not comprehending?" Indeed, the procedure used drew a definitive distinction between the two populations that conflicts with the continuous nature of the variable. In future studies, perhaps skippers could be identified at the time of testing and specific questions about why they skipped could be asked.

The reliability of the pre- and post-test measures posed a threat to the internal validity of the study conclusions. Although it was planned to use questions used previously by the textbook author that would provide reliability data, this did not come to fruition, mainly because many of the questions were taken from sections of the text that contained data presented in tabular format. The audio text selected avoided tabular data so as to be narrative. This resulted in most of the questions being created by the study author. Reliability data were calculated after the study was complete. A pilot study was not conducted, and in hindsight would have been instructive. Nonetheless, the post-hoc item analysis yielded a set of pre-test questions with a Kuder-Richardson 20 (KR20) reliability score of 0.304, which is low. Reliability for the post-test questions was much better with a KR20 reliability score of 0.576 which is moderate, but still somewhat low for this type of comprehension assessment. Since comprehension is the primary focus of most of the study, these statistics bring into question many of the conclusions which may be subsequently drawn.

Areas for Further Investigation

Many avenues of future study have already been suggested. Additionally, a detailed investigation of the type and location of navigations within the text could prove fruitful. Do certain types of audio navigational affordances (e.g., bookmarking, searching, key word skimming, etc.) correlate with greater comprehension? Do certain passages (i.e., hotspots) of a given audio text prompt greater use of certain types of navigations, and if so, why? Are usage patterns apparent? Only a small set of navigational affordances was examined in this study.

The curvilinear use of navigational affordances in relation to comprehension is intriguing, especially the slight difference in use between low and high comprehension learners. While this study supported the curvilinear hypothesis, the variance accounted for was modest ($R^2 = 0.11$). Further studies need to confirm this preliminary finding. If confirmed, the data could add support to the idea that, at least for low-comprehension learners, affordance use could offer greater gain in comprehension than for average or high comprehension learners. High-comprehension learners (i.e., those with high prior knowledge) may not see as much gain from affordance use. It may be beneficial to compare the use and effects of affordances between these groups. More beneficial would be studies that find ways to help low comprehension learners behave like higher comprehension learners in their use of affordances.

The appropriateness of the MSLQ as an instrument to measure the construct of metacognitive controlling strategies as they relate to affordance use should be further studied. It is possible that a better instrument may exist or could be created to measure this construct. An instrument that focuses on the specific monitoring and control strategies of audio comprehension could be developed. For example, a study with an audio text that specifically introduces ambiguous propositions, such as that conducted by Alessi, Anderson, and Goetz (1979) could clarify the process of monitoring and control within the audio domain.

The categorization and study of navigational affordance types is another area for study. For the purposes of this study, the total number of uses of navigational affordances, regardless of type, was used to compute the statistics even though the study platform collected detailed information about the type and location of navigational affordance use within the audio text. Were there differences in comprehension based upon type of affordance used? Were there certain areas of the audio text where navigational affordances were used more frequently?

Finally, further study of the potential mediating role of distractions is essential to develop the model. Modifying the study platform to collect more complete information about distractions should be investigated. Possible study designs that would collect better psychoeducational data on distractions could be used. Is it possible disentangle the effects of distractions from the role of navigational affordances to positively affect comprehension in the audio domain?

While the present study provides limited new knowledge about audio navigational affordances, further research on comprehension within the audio domain will become more important as more sources of audio learning are made available, and more applications of text-to-speech technology are developed. Synthetic text-to-speech technology is rapidly improving to the point of some systems being indistinguishable from real people. The combined use of artificial intelligence and text-to-speech is creating opportunities for new systems of communication and learning heretofore not possible. All these developments speak to the importance of understanding how people process audio communication and how interface design can affect comprehension.

REFERENCES

Abu Doush, I. T. (2010). *Non-visual navigation of Excel spreadsheet* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3391944)

Alessi, S. M., Anderson, T. H., & Goetz, E. T. (1979). An investigation of lookbacks during studying. *Discourse Processes, 2*(3), 197-212.

Alonso, M. B., & Keyson, D. V. (2006). MusicCube: A physical experience with digital music. *Personal & Ubiquitous Computing, 10*(2-3), 163-165.

Alvermann, D. E. (1988). Effects of spontaneous and induced lookbacks on self-perceived high- and low-ability comprehenders. *Journal of Educational Research, 81*(6), 325-331.

American Printing House for the Blind. (2015). Annual report 2014: Distribution of eligible students based on the federal quota census of January 7, 2013 (fiscal year 2014). Retrieved from http://www.aph.org/federal-quota/dist14.html

Andronico, P., Buzzi, M., Castillo, C., & Leporini, B. (2006). Improving search engine interfaces for blind users: A case study. *Universal Access in the Information Society, 5*(1), 23-40.

Andronico, P., Buzzi, M., & Leporini, B. (2005). (2005). Increasing usability of search engine interfaces for the blind. Paper presented at the *Proceedings of the 11th International Conference on Human-Computer Interaction,* Las Vegas, Nevada. 10pp.

Andronico, P., Castillo, C., & Leporini, B. (2006). (2006). A prototype of google interfaces modified for simplifying interaction for blind users. Paper presented at the *Proceedings of the 8th International ACM SIGACCESS Conference on Computers & Accessibility,* 267-268.

Archer, N. P., Head, M. M., Wollersheim, J. P., & Yuan, Y. (1996). Investigation of voice and text output modes with abstraction in a computer interface. *Interacting with Computers, 8*(4), 323-345.

Argyropoulos, V., Papadopoulos, K., Kouroupetroglou, G., Xydas, G., & Katsoulis, P. (2007). (2007). Discrimination and perception of the acoustic rendition of texts by blind people. Paper presented at the *Proceedings of the 12th International Conference on Human-Computer Interaction,* Beijing, P.R.China. 9pp.

Argyropoulos, V. S., Sideridis, G. D., Kouroupetroglou, G., & Xydas, G. (2009). Auditory discriminations of typographic attributes of documents by students with blindness. *British Journal of Visual Impairment, 27*(3), 183-203.

Arons, B. (1997). SpeechSkimmer: A system for interactively skimming recorded speech. *ACM Transactions on Computer-Human Interaction (TOCHI), 4*(1), 3-38.

Arons, B. (1993). (1993). SpeechSkimmer: Interactively skimming recorded speech. Paper presented at the *Proceedings of the 6th Annual ACM Symposium on User Interface Software and Technology,* Atlanta, Georgia, USA. 187-196.

Asakawa, C., & Takagi, H. (2001). One-dimensional user interface for retrieving information from the web for the blind. *Universal access in HCI: Towards an information society for all* (pp. 885-889).

Asakawa, C., Takagi, H., Ino, S., & Ifukube, T. (2003). A proposal for a dial-based interface for voice output based on blind users' cognitive listening abilities. *Universal access in HCI: Inclusive design in the information society* (pp. 1250-1254).

Asakawa, C., Takagi, H., & Itoh, T. (2001). *Document reader for the blind.* Retrieved from http://libezproxy.tamu.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=ega&AN=ega186739&site=ehost-live

Azevedo, R. (2005). Using hypermedia as a metacognitive tool for enhancing student learning? the role of self-regulated learning. *Educational Psychologist, 40*(4), 199-209.

Baker, L., & Anderson, R. I. (1982). Effects of inconsistent information on text processing: Evidence for comprehension monitoring. *Reading Research Quarterly, 17*(2), 281-294.

Balaji, T. S., & Roche, E. (2003). Whose turn is it? designing for overlapping grammar between applications and the voice browser. *Proceedings of the Human Factors & Ergonomics Society Annual Meeting,* 683-685.

Balaji, T. S., & Roche, E. (2004). Wizard of oz in a structured expert review (WiSER): An evaluative study. *Proceedings of the Human Factors & Ergonomics Society Annual Meeting,* , 2476-2479.

Bennett, R. W., Greenspan, S. L., Syrdal, A. K., Tschirgi, J. E., & Wisowaty, J. J. (1989). Speaking to, from, and through computers: Speech technologies and user-interface design. *AT&T Technical Journal, 68*(5), 17-30.

Bever, T. G., Garrett, M. F., & Hurtig, R. (1973). The interaction of perceptual processes and ambiguous sentences. *Memory & Cognition, 1*(3), 277-286.

Boulton, D. A. (1991). *U.S. Patent No. 4,985,697. Electronic book educational publishing method using buried reference materials and alternate learning levels.* Washington, DC: U.S. Patent and Trademark Office.

Boyle, E. A., Washburn, S. G., & Rosenberg, M. S. (2002). Reading's SLiCK with new audio texts and strategies. *Teaching Exceptional Children, 35*(2), 50-55.

Bretz, E. A. (2001). The car: Just a web browser with tires. *IEEE Spectrum, 38*(1), 92-94.

Burton, C., & Daneman, M. (2007). Compensating for a limited working memory capacity during reading: Evidence from eye movements. *Reading Psychology, 28*(2), 163-186.

Buzick, H., & Stone, E. (2014). A meta-analysis of research on the read aloud accommodation. *Educational Measurement: Issues & Practice, 33*(3), 17-30.

Chen, A. (2005). IBM's HPR reads more than web pages. *eWeek, 22*(6), 50.

Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research, 53*, 445-459.

Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research and Development, 42*(2), 21-29.

Commarford, P. M., Wilson, K. A., & Stanney, K. M. (2002). When should computers talk: Using multiple resource theory to determine whether to add synthetic speech to a user interface. *Proceedings of the Human Factors & Ergonomics Society Annual Meeting,* 612-616.

Connor, L. T., Dunlosky, J., & Hertzog, C. (1997). Age-related differences in absolute but not relative metamemory accuracy. *Psychology & Aging, 12*(1), 50.

Cornell University. (2015). Find U.S. disability statistics in 3 easy steps. Retrieved from http://www.disabilitystatistics.org/reports/acs.cfm?statistic=1

Craik, F.I. & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior,* (11), 671-684.

Credé, M., & Phillips, L. A. (2011). A meta-analytic review of the motivated strategies for learning questionnaire. *Learning & Individual Differences, 21*(4), 337-346.

Danks, J. (1980). Comprehension in listening and reading: Same or different? In F. B. Murray (Ed.), *Reading and understanding* (pp. 1-39). Newark: International Reading Association.

Davis, J. R., & Hirschberg, J. (1988). (1988). Assigning intonational features in synthesized spoken directions. Paper presented at the *Proceedings of the 26th Annual Meeting on Association for Computational Linguistics,* Buffalo, New York. 187-193.

Duffy, S. A., & Pisoni, D. B. (1992). Comprehension of synthetic speech produced by rule: A review and theoretical interpretation. *Language and Speech, 35 (Pt 4)*, 351-389.

Dunlosky, J., & Rawson, K. A. (2005). Why does rereading improve metacomprehension accuracy? evaluating the levels-of-disruption hypothesis for the rereading effect. *Discourse Processes A Multidisciplinary Journal, 40*(1), 37-55.

Ebina, T., Igi, S., Miyake, T., & Takahashi, H. (1999). GUI object search method using a tactile display. *Electronics & Communications in Japan, Part 3: Fundamental Electronic Science, 82*(8), 40-49.

Editorial. (1994). Geac intros OptiView and OptiVoice. *Information Today, 11*(11), 38.

Editorial. (1997, Voicemail on a PC. *Management Services, 41*, 31.

Editorial. (2007, October). Stranger than fiction. *Control & Automation, 18*, 28-33.

Edwards, A. D. N. (1991). Speech synthesis. technology for disabled people. *Paul Chapman Publishing, London*, 157.

Edwards, A. D. N. (1992). (1992). Graphical user interfaces and blind people. Paper presented at the *Proceedings of the 3rd International Conference on Computers for Handicapped Persons,* 114-119.

Elkind, J., Black, M. S., & Murray, C. (1996). Computer-based compensation of adult reading disabilities. *Annals of Dyslexia, 46*, 159-186.

Engwall, O., Balter, O., Oster, A. M., & Kjellstrom, H. (2006). Designing the user interface of the computer-based speech training system ARTUR based on early user tests. *Behaviour & Information Technology, 25*(4), 353-365.

Fellbaum, K., & Koroupetroglou, G. (2008). Principles of electronic speech processing with applications for people with disabilities. *Technology & Disability, 20*(2), 55-85.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. Resnick (Ed.), *The nature of intelligence* (pp. 231-236). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Foulke, E. (1962). Comprehension of rapid speech by the blind. *Exceptional Children, 29*, 134-141.

Foulke, E. (1966). Comparison of comprehension of two forms of compressed speech. *Exceptional Children, 33*(3), 169-173.

Foulke, E., & Louisville Univ., K. (1967). *The comprehension of rapid speech by the blind, part III.* Louisville Univ., KY.

Foulke, E. (1964). *The comprehension of rapid speech by the blind, part II*. Louisville Univ., KY.

Foulke, E., Amster, C. H., Nolan, C. Y., & Bixler, R. H. (1962). The comprehension of rapid speech by the blind. *Exceptional Children, 29*(3), 134-141.

Foulke, E., & Louisville Univ, KY Perceptual,Alternatives Lab. (1969). *The comprehension of rapid speech by the blind: Part III. final report*. Louisville Univ., KY.

Foulke, E., & Louisville Univ, KY Perceptual,Alternatives Lab. (1971). *Proceedings of the second louisville conference on rate and/or frequency-controlled speech (louisville, kentucky, october 22-24, 1969)*.

Foulke, E., Robinson, J., & Louisville Univ, KY Perceptual Alternatives Lab. (1970). *The development of accelerated speech as a useful communication tool in the education of blind and other handicapped children. progress report*. Louisville Univ., KY.

Franks, B. A., Therriault, D. J., Buhr, M. I., Chiang, E. S., Gonzalez, C. M., Kwon, H. K., . . . Wang, X. (2013). Looking back: Reasoning and metacognition with narrative texts. *Metacognition and Learning, 8*(2), 145-171.

Freitas, D., & Kouroupetroglou, G. (2008). Speech technologies for blind and low vision persons. *Technology & Disability, 20*(2), 135-156.

Fucci, D., Reynolds, M. E., Bettagere, R., & Gonzales, M. D. (1995). Synthetic speech intelligibility under several experimental conditions. *Augmentative and Alternative Communication,* (11), 113-117.

Garcia, L. G. (2004). Assessment of text reading comprehension by spanish-speaking blind persons. *British Journal of Visual Impairment, 22*(1), 4-12.

Garner, R., Hare, V. C., Alexander, P., Haynes, J., & Winograd, P. (1984). Inducing use of a text lookback strategy among unsuccessful readers. *American Educational Research Journal, 21*(4), 789-798.

Gaver, W. W. (1991). (1991). Technology affordances. Paper presented at the *Proceedings of the SIGCHI Conference,* 79-84.

Gibson, J. J. (1977). *The theory of affordances. perceiving, acting, and knowing: Toward an ecological psychology.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Gibson, J. J. (1979). *The ecological approach to visual perception*. New Jersey: Lawrence Erlbaum Associates.

Goudiras, D. B., Papadopoulos, K. S., Koutsoklenis, A. C., Papageorgiou, V. E., & Stergiou, M. S. (2009). Factors affecting the reading media used by visually impaired adults. *British Journal of Visual Impairment, 27*(2), 111-127.

Grammenos, D., Savidis, A., & Stephanidis, C. (2005). (2005). UA-chess: A universally accessible board game. Paper presented at the *Proceedings of the 11th International Conference on Human-Computer Interaction,* Las Vegas, NV. 10pp.

Grams, R. R., Smillov, A., & Li, B. (1992). Field-testing the new DECtalk PC system for medical applications. *Journal of Medical Systems, 16*(1), 31-38.

Grasso, M. A., Ebert, D. S., & Finin, T. W. (1998). The integrality of speech in multimodal interfaces. *ACM Transactions on Computer-Human Interaction (TOCHI), 5*(4), 303-325.

Gunderson, J. (1993). Intelligent task based model for the design of screen review systems for people with blindness. *Human-computer interaction: Software & hardware interfaces* (pp. 736-741)

Hacker, D. J. (1998). Self-regulated comprehension during normal reading. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 165-191). Mahwah, N.J.: L. Erlbaum Associates.

Hakulinen, J., Turunen, M., & Raiha, K. J. (2005). Evaluation of software tutoring for a speech interface. *International Journal of Speech Technology, 8*(3), 283-293.

Halstead-Nussloch, R. (1989). (1989). The design of phone-based interfaces for consumers. Paper presented at the *Conference on Human Factors in Computing Systems Proceedings,* Poughkeepsie, NY. 347-352.

Hansen, E. G., Lee, M. J., & Forer, D. C. (2002). A "self-voicing" test for individuals with visual impairments. *Journal of Visual Impairment & Blindness, 96*(4), 273-75.

Haskins Labs., New Haven, CT. (1972). *Status report on speech research, no. 29/30, January-June 1972.* (). Bethesda, MD: National Inst. of Child Health and Human Development (NIH).

Hertrich, I., Dietrich, S., Moos, A., Trouvain, J., & Ackermann, H. (2009). Enhanced speech perception capabilities in a blind listener are associated with activation of fusiform gyrus and primary visual cortex. *Neurocase, 15*(2), 163-170.

Hertrich, I., Dietrich, S., & Ackermann, H. (2013). How can audiovisual pathways enhance the temporal resolution of time-compressed speech in blind subjects? *Frontiers in Psychology, 4*, 530.

Hjelmquist, E., & Jansson, B. (1990). User performance and error analysis: Computer based newspaper reading by blind people. In L. Berlinguet, & D. Berthelette (Eds.), *Work with display units 89* (pp. 327-333). North Holland: Elsevier Science Publishers.

Hjelmquist, E., & Others, A. (1992). Visually impaired persons' comprehension of text presented with speech synthesis. *Journal of Visual Impairment and Blindness, 86*(10), 426-28.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journa, 2*(2), 127-160.

Hub, A., Diepstraten, J., & Ertl, T. (2004). (2004). Design and development of an indoor navigation and object identification system for the blind. Paper presented at the *ASSETS 2004.Proceedings of the 6th International ACM SIGACCESS Conference on Computers & Accessibility,* Atlanta, GA. 147-152.

Hunsucker, R. L. (2013). Making life easier for the visually impaired web searcher: It is now clearer how this should and can be done, but implementation lags. *Evidence Based Library & Information Practice, 8*(1), 90-93.

Jeon, M., Walker, B. N., & Srivastava, A. (2012). "Spindex" (speech index) enhances menus on touch screen devices with tapping, wheeling, and flicking. *ACM Transactions on Computer-Human Interaction (TOCHI), 19*(2), 14:27.

Kamm, C. (1995). User interfaces for voice applications. *Proceedings of the National Academy of Sciences of the United States of America, 92*, 10031-10037.

Karshmer, A. I. (1997). Interdisciplinary efforts to facilitate the production of tools to support the disabled and elderly in the information society. In G. Salvendy, & M. J. :. Smith R.J (Eds.), *Design of computing systems: Cognitive considerations* (pp. 409-412).

Kim, S. J., Kim, J. J., & Hahn, M. (2006). HMM-based korean speech synthesis system for hand-held devices. *IEEE Transactions on Consumer Electronics, 52*(4), 1384-1390.

Kinnunen, R., & Vauras, M. (1995). Comprehension monitoring and the level of comprehension in high- and low-achieving primary school children's reading. *Learning and Instruction, 5*(2), 143-165.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, England: Cambridge University Press.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363-394.

Klie, L. (2013). Over head under head. *Speech Technology Magazine, 18*(4), 19.

Knott, B. A., & Kortum, P. (2006). Personification of voice user interfaces: Impacts of user performance. *Proceedings of the Human Factors & Ergonomics Society Annual Meeting,* 599-603.

Kochanek, D. (1992). (1992). A hypertext system for the blind newspaper reader. Paper presented at the *Proceedings of the 3rd International Conference on Computers for Handicapped Persons,* Vienna, Austria. 285-293.

Kozma, R. B. (1991). Learning with media. *Review of Educational Research, 61*, 179-211.

Kozma, R. B. (1994). The influence of media on learning: The debate continues. *School Library Media Quarterly, 22*(4), 233.

Kuyk, T. K., & Story, S. M. (1988). Recognition of synthesized speech by experienced echo+ users: A follow-up. *Journal of Visual Impairment & Blindness, 82*, 382-383.

Lamell, L., Bennacef, S., Gauvain, J. L., Dartigues, H., & Temem, J. N. (2002). User evaluation of the MASK kiosk. *Speech Communication, 38*(1-2), 131-139.

Laws, C. K. (2005). (2005). Presentation-based web accessibility evaluation system: Design and implementation challenges and trade-offs. Paper presented at the *HCI International 2005.Proceedings of the 11th International Conference on Human-Computer Interaction,* Las Vegas, NV. 9pp.

Lazzaro, J. J. (1991). Windows of vulnerability. *Byte, 16*, 416.

Lee, K. M., Jung, Y., & Nass, C. (2011). Can user choice alter experimental findings in human-computer interaction?: Similarity attraction versus cognitive dissonance in social responses to synthetic speech. *International Journal of Human-Computer Interaction, 27*(4), 307-322.

Levitt, H. (1995). Processing of speech signals for physical and sensory disabilities. *Proceedings of the National Academy of Sciences of the United States of America, 92*(22), 9999-10006.

Li, H. (2014). The effects of read-aloud accommodations for students with and without disabilities: A meta-analysis. *Educational Measurement: Issues & Practice, 33*(3), 3-16. 10.

Loyd, D. B., Phalangas, A. C., & Barner, K. E. (1999). An audio- and speech-based interface for computer-controlled scientific instruments. *IEEE Transactions on Rehabilitation Engineering, 7*(2), 245-253.

Luxton, K. E. (1985). *Synthetic vs. natural speech and comprehension in blind and sighted adults* Available from PsycINFO. (617090832; 1985-57534-001).

Malyan, R. R., Sunthankar, Y., & Barnwell, P. (1987). (1987). An intelligent text reader for the blind. Paper presented at the *Proceedings of the European Conference on Speech Technology,* Edinburgh, Scotland. *1,* 344-347.

Marics, M. A., & Engelback, G. (1997). Designing voice menu applications for telephones. In M. G. Helander, T. K. Landauer & P. V. Prabhu (Eds.), *Handbook of human-computer interaction* (2nd ed., pp. 1085-1102). Amsterdam: Elsevier.

Markovits, H. (1993). The development of conditional reasoning: A Piagetian reformulation of mental models theory. *Merrill-Palmer Quarterly (1982-), 39*(1), 131-158.

Marsic, I., & Dorohonceanu, B. (2003). Flexible user interfaces for group collaboration. *International Journal of Human-Computer Interaction, 15*(3), 337-360.

Marston, J. R., & Church, R. L. (2005). A relative access measure to identify barriers to efficient transit use by persons with visual impairments. *Disability & Rehabilitation, 27*(13), 769-779.

Martial, O., & Dufresne, A. (1993). Audicon: Easy access to graphical user interfaces for blind persons - designing for and with people. *Human-computer interaction: Software & hardware interfaces* (pp. 808-813)

Maxwell, B. A., Meeden, L. A., Addo, N. S., Dickson, P., Fairfield, N., Johnson, N., . . . Silk, E. (2001). REAPER: A reflexive architecture for perceptive agents. *AI Magazine, 22*(1), 53.

McNulty, T., & TLT Group, Washington,,DC. (1994). Information technology and disabilities, 1994. *Equal Access to Software and Information (EASI), 1*(4).

Moller, S., Jekosch, U., Mersdorf, J., & Kraft, V. (2001). Auditory assessment of synthesized speech in application scenarios: Two case studies. *Speech Communication, 34*(3), 229-246.

Moller, S., Krebber, J., & Smeele, P. (2006). Evaluating the speech output component of a smart-home system. *Speech Communication, 48*(1), 1-27.

Morley, S., Petrie, H., O'Neill, A. M., & McNally, P. (1999). Auditory navigation in hyperspace: Design and evaluation of a non-visual hypermedia system for blind users. *Behaviour & Information Technology, 18*(1), 18-26.

Morrissey, W., & Zajicek, M. (2000). Can sound output enhance graphical computer interfaces? *Contemporary ergonomics 2000* (pp. 172-176).

Moshman, D. (1990). The development of metalogical understanding. In W. F. Overton (Ed.), *Reasoning, necessity and logic: Developmental perspectives* (pp. 205-225). Hillsdale: Erlbaum.

Murphy, E., Kuber, R., McAllister, G., Strain, P., & Yu, W. (2008). An empirical investigation into the difficulties experienced by visually impaired internet users. *Universal Access in the Information Society, 7*(1-2), 79-91.

Nass, C., Robles, E., Heenan, C., Bienstock, H., & Treinen, M. (2003). Speech-based disclosure systems: Effects of modality, gender of prompt, and gender of user. *International Journal of Speech Technology, 6*(2), 113-121.

Noot, H., & Ruttkay, Z. (2005). Variations in gesturing and speech by GESTYLE. *International Journal of Human-Computer Studies, 62*(2), 211-229.

Norén, N., Svensson, E., & Telford, J. (2013). Participants' dynamic orientation to folder navigation when using a VOCA with a touch screen in talk-in-interaction. *AAC: Augmentative & Alternative Communication, 29*(1), 20-36.

Norman, D. A. (1988). *The psychology of everyday things*. New York: Basic books.

Paap, K. R., & Cooke, N. J. (1997). Design of menus. In M. G. Helander, T. K. Landauer & P. V. Prabhu (Eds.), *Handbook of human-computer interaction* (2nd ed., pp. 533-572). Amsterdam: Elsevier.

Palladino, D. K., & Walker, B. N. (2008). Navigation efficiency of two dimensional auditory menus using spearcon enhancements. *Proceedings of the Human Factors & Ergonomics Society Annual Meeting,* 1262-1266.

Papadopoulos, K., Argyropoulos, V. S., & Kouroupetroglou, G. (2008). Discrimination and comprehension of synthetic speech by students with visual impairments: The case of similar acoustic patterns. *Journal of Visual Impairment & Blindness, 102*(7), 420-429.

Papadopoulos, K., Koutsoklenis, A., Katemidou, E., & Okalidou, A. (2009). Perception of synthetic and natural speech by adults with visual impairments. *Journal of Visual Impairment & Blindness, 103*(7), 403-414.

Parslow, A. (2005). An optical real-world text to speech reader. *International Congress Series, 1282*, 1056-1060.

Partala, T., & Surakka, V. (2004). The effects of affective interventions in human–computer interaction. *Interacting with Computers, 16*(2), 295.

Pavešic, N., Gros, J., Dobrišek, S., & Mihelic, F. (2003). Homer II—man–machine interface to internet for blind and visually impaired people. *Computer Communications, 26*(5), 438.

Perzanowski, D., Schultz, A. C., & Adams, W. (2001). Building a multimodal human-robot interface. *IEEE Intelligent Systems, 16*(1), 16-21.

Pintrich, P. R. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)* Ann Arbor, Mich. : University of Michigan ; Washington, DC] : U.S. Dept. of Education, Office of Educational Research and Improvement, Educational Resources Information Center, 1991].

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. Pintrick & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451-502). San Diego, CA: Academic.

Pintrich, P. R., & Smith, D. A. F. (1993). Reliability and predictive validity of the motivated strategies for learning questionnaire (MSLQ). *Educational & Psychological Measurement, 53*(3), 801-813.

Pontelli, E., Son, T. C., Kottapally, K., Ngo, C., Reddy, R., & Gillan, D. (2004). A system for automatic structure discovery and reasoning-based navigation of the web. *Interacting with Computers, 16*(3), 451-475.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading : The nature of constructively responsive reading. michael pressley, peter afflerbach* Hillsdale, N.J. : Lawrence Erlbaum Associates, 1995.

Qiu, L., & Benbasat, I. (2005). An investigation into the effects of text-to-speech voice and 3D avatars on the perception of presence and flow of live help in electronic commerce. *ACM Transactions on Computer-Human Interaction (TOCHI), 12*(4), 329-355.

Qiu, L., & Benbasat, I. (2009). Evaluating anthropomorphic product recommendation agents: A social relationship perspective to designing information systems. *Journal of Management Information Systems, 25*(4), 145-181.

Ralston, J. V., Pisoni, D. B., & Lively, S. E. (1991). Comprehension of synthetic speech produced by rule: Word monitoring and sentence-by-sentence listening times. *Human Factors, 33*, 471-491.

Raman, T. V. (1997). Emacspeak: A speech-enabling interface. *Dr.Dobb's Journal: The World of Software Development, 22*, 18.

Ramloll, R., Brewster, S., Yu, W., & Riedel, B. (2001). (2001). Using non-speech sounds to improve access to 2D tabular numerical information for visually impaired users. Paper presented at the *People & Computers XV - Interaction without Frontiers: Joint Proceedings of HCI 2001 & IHM 2001,* 515-529.

Reddy, H., Gupta, G., & Karshmer, A. I. (2005). (2005). Dynamic aural browsing of MathML documents via VoiceXML. Paper presented at the *HCI International 2005.Proceedings of the 11th International Conference on Human-Computer Interaction,* Las Vegas, NV. 10pp.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26*(4), 445-476.

Rezac, F., Kolenovsky, F., Mikulec, M., Voznak, M., Rozhon, J., Vychodil, J., . . . Safarik, J. (2012). IP telephony based danger alert communication system and its implementation. *Advances in Electrical & Electronic Engineering, 10*(4), 286-290.

Rhyne, J. M. (1982). Comprehension of synthetic speech by blind children. *Journal of Visual Impairment and Blindness, 76*(8), 313-16.

Rohani Ghahari, R., Ferati, M., Yang, T., & Bolchini, D. (2012). (2012). Back navigation shortcuts for screen reader users. Paper presented at the *Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility,* Boulder, Colorado, USA. 1-8.

Rosson, M. B. (1985). Using synthetic speech for remote access to information. *Behavior Research Methods, 17*(2), 250-252.

Sandweg, N., Hassenzahl, M., & Kuhn, K. (2000). Designing a telephone-based interface for a home automation system. *International Journal of Human-Computer Interaction, 12*(3-4), 401-414.

Schiefele, U. (1988). Motivational Factors of Text Comprehension. *Zeitschrift fur Padagogik, 34*(5), 687-708.

Schmandt, C., & Roy, D. (1996). Using acoustic structure in a hand-held audio playback device. *IBM Systems Journal, 35*(3), 453-472.

Shajahan, P., & Irani, P. (2007). One family, many voices: Can multiple synthetic voices be used as navigational cues in hierarchical interfaces? *International Journal of Speech Technology, 9*(1), 1-15.

Sharit, J., Czaja, S. J., Nair, S., & Lee, C. C. (2003). Effects of age, speech rate, and environmental support in using telephone voice menu systems. *Human Factors, 45*(2), 234-251.

Simpson, P. J. (1996). Defining the user model to enhance personal reader systems for the blind. *Contemporary ergonomics 1996: Proceedings of the annual conference of the ergonomics society* (pp. 458-463). London, England: Taylor & Francis.

Simpson, P. J. (1999). Personal reader systems for the blind: Some design issues. In W. S. Green, & P. W. Jordan (Eds.), *Human factors in product design: Current practice & future trends* (pp. 182-190). London: Taylor & Francis.

Smith, C. P. (1972). Conference record. conference on speech communication and processing, newton, massachusetts, 24-26 april 1972. (pp. 170-174). New York, NY: The Institute for Electronic and Electrical Engineers.

Sodnik, J., Jakus, G., & Tomažic, S. (2012). The use of spatialized speech in auditory interfaces for computer users who are visually impaired. *Journal of Visual Impairment & Blindness, 106*(10), 634-645.

Spiliotopoulos, D., Xydas, G., Kouroupetroglou, G., & Argyropoulos, V. (2005). (2005). Experimentation on spoken format of tables in auditory user interfaces. Paper presented at the *HCI International 2005.Proceedings of the 11th International Conference on Human-Computer Interaction,* Las Vegas, NV. 10pp.

Story, S. M., & Kuyk, T. K. (1988). First time recognition of synthesized speech: A comparison of three systems. *Journal of Visual Impairment & Blindness, 82*, 28-29.

Sweeney, D. (2015). *Mobile usage statistics at texas a&M university.* Unpublished manuscript.

Torihara, S., Nakamura, M., Ueda, N., Wada, T., & Ishizaki, S. (2006). (2006). English 'oblique' listening system - rapid listening system for the blind and visually impaired, and its evaluation. Paper presented at the *Computers Helping People Wth Special Needs.Proceedings of the 10th International Conference, ICCHP 2006,* Heidelberg, Germany. 53-60.

Tucker, H. J., & Luttbeg, N. R. (2015). *Comparing the states and communities* N.R.

Tunmer, W. E., & Chapman, J. W. (2012). The simple view of reading redux: Vocabulary knowledge and the independent components hypothesis. *Journal of Learning Disabilities, 45*(5), 453-466.

Uslan, M. M., Eghtesadi, K., & Burton, D. (2003). Accessibility of blood glucose monitoring systems for blind and visually impaired people. *Diabetes Technology & Therapeutics, 5*(3), 439-448.

Vertegaal, R., & Eaglestone, B. (1996). Comparison of input devices in an ISEE direct timbre manipulation task. *Interacting with Computers, 8*(1), 13-30.

Virgili, G., Acosta, R., Grover, L. L., Bentley, S. A., & Giacomelli, G. (2013). Reading aids for adults with low vision (review). *The Cochrane Library,* (10)10.

Walczyk, J. J., & Taylor, R. W. (1996). How do the efficiencies of reading subcomponents relate to looking back in text? *Journal of Educational Psychology, 88*(3), 537-545.

Walker, B. N., Lindsay, J., Nance, A., Nakano, Y., Palladino, D. K., Dingler, T., & Jeon, M. (2013). Spearcons (speech-based earcons) improve navigation performance in advanced auditory menus. *Human Factors, 55*(1), 157-182.

Wall, S., & Brewster, S. (2006). (2006). Feeling what you hear: Tactile feedback for navigation of audio graphs. Paper presented at the *Conference on Human Factors in Computing Systems Proceedings,* 1123-1132.

Weber, G. (1992). (1992). Adapting graphical interaction objects for blind users by integrating braille and speech. Paper presented at the *Proceedings of the 3rd International Conference on Computers for Handicapped Persons,* 546-555.

Wood, S. G., Moxley, J. H., Tighe, E. L., & Wagner, R. K. (2018). Does use of text-to-speech and related read-aloud tools improve reading comprehension for students with reading disabilities? A meta-analysis. *Journal of Learning Disabilities, 51*(1), 73-84.

World Wide Web Consortium. (2014). Web accessiblity initiative. Retrieved from http://www.w3.org/WAI/

Xydas, G., Argyropoulos, V., Karakosta, T., & Kouroupetroglou, G. (2005). (2005). An experimental approach to recognizing synthesized auditory components in a non-visual interaction with documents. Paper presented at the *HCI International 2005.Proceedings of the 11th International Conference on Human-Computer Interaction,* Las Vegas, NV. 10pp.

Yamada, K. (2009). Lexical patterns in L2 textual gist identification assessment. *Language Testing, 26*(1), 101-122.

Yuan, Y., Head, M., & Du, M. (2003). The effects of multimedia communication on web-based negotiation. *Group Decision & Negotiation, 12*(2), 89-109. 1023016804379

Zabrucky, K., & Moore, D. (1994). Contributions of working memory and evaluation and regulation of understanding to adults' recall.. *Journal of Gerontology, 49*(5), 201.

Zajicek, M., & Powell, C. (1997). (1997). Accessing the world wide web by telephone. Paper presented at the *Proceedings of the 16th International Symposium on Human Factors in Telecommunications,* Oslo, Norway. 177-184.

Zue, V. W. (1995). Navigating the information superhighway using spoken language interfaces. *IEEE Expert, 10*(5), 39-43.

APPENDIX A - TABLES

Table 1

Item Analysis of Pre-Test Comprehension Questions

	Item Difficulty	Point Biserial Correlation	KR20 if Item Deleted
Q1	.36	.068	.303
Q2	.20	.138	.271
Q3	.60	.141	.266
Q4	.55	.136	.269
Q5	.69	.113	.281
Q6	.51	.133	.270
Q7	.20	.123	.277
Q8	.64	.142	.266
Q9	.32	.125	.275
Q10	.19	-.038	.339

KR20 = 0.304
n = 10

Table 2

Item Analysis of Post-Test Comprehension Questions

	Item Difficulty	Point Biserial Correlation	KR20 if Item Deleted
Q1	.43	.140	.572
Q2	.49	.038	.588
Q3	.32	.194	.563
Q4	.06	-.115	.589
Q5	.43	.205	.561
Q6	.77	.395	.534
Q7	.30	.073	.581
Q8	.50	.156	.569
Q9	.43	.246	.554
Q10	.66	.155	.569
Q11	.36	.190	.563
Q12	.58	.178	.565
Q13	.22	-.016	.591
Q14	.66	.240	.555
Q15	.90	.049	.579
Q16	.90	.442	.540
Q17	.43	.188	.564
Q18	.60	.234	.556
Q19	.44	.373	.533
Q20	.41	.395	.529

KR20 = 0.576
n = 20

Table 3

Comparison of Means for Different Parameters of Skipper Navigational Behavior

		Skippers		Non-Skippers				
consecutive forward navigations	per unit time (secs.)	N	Mean Post-Test Comp.	N	Mean Post-Test Comp.	Difference in Means	α (2-tailed)	Power
6	3	23	9.17	73	9.90	-0.73	0.31	0.17
6	4	23	9.17	73	9.90	-0.73	0.31	0.17
6	5	25	9.12	71	9.94	-0.82	0.24	0.22
7	3	19	9.00	77	9.91	-0.91	0.24	0.22
7	4	23	9.17	73	9.90	-0.73	0.31	0.17
7	5	23	9.17	73	9.90	-0.73	0.31	0.17
7	6	24	9.33	72	9.86	-0.53	0.46	0.11
8	3	14	9.14	82	9.83	-0.69	0.43	0.12
8	4	22	9.00	74	9.95	-0.95	0.20	0.25
8	5	23	9.17	73	9.90	-0.73	0.31	0.17
8	6	24	9.33	72	9.86	-0.53	0.46	0.11
9	6	23	9.17	73	9.90	-0.73	0.31	0.17
9	7	23	9.17	73	9.90	-0.73	0.31	0.17
9	8	24	9.33	72	9.86	-0.53	0.46	0.11
10	5	21	8.95	75	9.95	-1.00	0.18	0.27
10	6	22	9.00	74	9.95	-0.95	0.20	0.25
10	7	22	9.00	74	9.95	-0.95	0.20	0.25
10	8	22	9.00	74	9.95	-0.95	0.20	0.25

Table 4

Descriptive Statistics of Population Demographics

		N	Percent
Age	18 to 24 years	129	74.1%
	25 to 34 years	1	0.6%
	35 to 44 years	4	2.3%
	45 to 54 years	4	2.3%
	55 to 64 years	1	0.6%
	65 years and over	1	0.6%
	Valid	140	80.5%
	Not Reported	34	19.5%
	Total	174	100.0%
Classification	Freshman	7	4.0%
	Sophomore	10	5.7%
	Junior	85	48.9%
	Senior	29	16.7%
	Doctoral Student	2	1.1%
	Other	7	4.0%
	Valid	140	80.5%
	Not Reported	34	19.5%
	Total	174	100.0%

Table 5

Participant Mortality

	Total	Control	Exp.	Percent
Began Protocol	174	29	145	100%
Discontinued procotol:				
before demographic questions	34	4	30	20%
before pre-test questions	9	3	6	5%
during the audio portion	18	9	9	10%
before post-test questions	4	2	2	2%
Elimated as Skippers	21	0	21	12%
Final Sample	88	11	77	51%

Table 6

Descriptive Statistics of Study Measures After Excluding Outliers

						Skewness		Kurtosis	
	N	Min	Max	Mean	S.D.	Statistic	S.E.	Statistic	S.E.
Pre-Test Score	108	0	10	4.27	1.71	0.14	0.23	0.19	0.46
Pre-Test Score*	108	0	9	4.08	1.68	0.13	0.23	-0.17	0.46
Post-Test Score	88	2	16	9.89	3.04	-0.39	0.26	-0.33	0.51
Post Test Score*	88	1	13	7.43	2.74	-0.34	0.26	-0.22	0.51
Total Navigations	78	0	39	10.5	8.07	1.57	0.27	2.63	0.54
Training Navigations	78	0	21	6.35	4.85	1.04	0.27	0.92	0.54
Self-Regulation (SR)	110	2.42	6.17	4.29	0.7	-0.05	0.23	0.31	0.46
Rehearsal (RE)	110	2.50	7.00	4.55	1.01	0.07	0.23	-0.36	0.46
Interest (I)	112	3.25	6.38	5.00	0.63	-0.26	0.23	-0.14	0.45
Test Anxiety (TA)	112	1.00	7.00	4.25	1.37	-0.22	0.23	-0.63	0.45
Effort Regulation (ER)	110	1.75	7.00	4.77	1.05	-0.10	0.23	0.10	0.46
Valid N (listwise)	74								

*after item analysis

Table 7

RQ1-A: Tests of Between-Subjects Effects with a Dependent Variable of Post-Test Comprehension

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Power
Corrected Model	11.2	2	5.60	0.73	0.48	0.02	1.47	0.17
Intercept	401.1	1	401.10	52.46	0.00	0.38	52.46	1.00
Pre-Test Comp.	9.2	1	9.22	1.21	0.28	0.01	1.21	0.19
Experimental	2.5	1	2.47	0.32	0.57	0.00	0.32	0.09
Error	642.2	84	7.65					
Total	5465.0	87						
Corrected Total	653.4	86						

Model R Squared = .017
Computed using alpha = .05

Table 8

RQ1-B - Model Summary of Curvilinear Regression of Affordance Use vs. Comprehension

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Err.of the Est.	R Square Change	F Change	df1	df2	Sig. F Change
1(a)	0.155	0.024	0.01	2.63	0.024	1.697	1	69	0.197
2(b)	0.327	0.107	0.081	2.535	0.083	6.302	1	68	0.014

a Predictors: (Constant), Total Navigations
b Predictors: (Constant), Total Navigations, Total Navigations Squared

Table 9

RQ1-B - ANOVA of Curvilinear Model of Affordance Use vs. Comprehension

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	11.75	1	11.75	1.697	.197b
	Residual	477.44	69	6.92		
	Total	489.18	70			
2	Regression	52.24	2	26.12	4.065	.022c
	Residual	436.94	68	6.43		
	Total	489.18	70			

a Dependent Variable: Post Test Score (IA)
b Predictors: (Constant), Total Navigations
c Predictors: (Constant), Total Navigations, Total Navigations Squared

Table 10

RQ2 and RQ3 - Multivariate ANOVA Effects of Metacognitive and Motivational Factors on Navigational Affordance Use

		Wilk's Lambda	F	Sig.	Power
Metacognitive Factors	Intercept	0.956	1.572	0.215	0.322
	SR	0.970	1.057	0.353	0.228
	RE	0.976	0.866	0.425	0.193
	SR * RE	0.968	1.128	0.330	0.241
Motivational Factors	Intercept	0.988	0.386	0.682	0.110
	I	0.991	0.301	0.741	0.096
	TA	0.993	0.235	0.791	0.085
	ER	0.989	0.374	0.690	0.108
	I * TA	0.993	0.23	0.795	0.085
	TA * ER	0.993	0.214	0.808	0.082
	I * ER	0.989	0.368	0.694	0.107
	I * TA * ER	0.992	0.25	0.780	0.088

Table 11

RQ2 Between-Subjects' Effects of Metacognitive Factors on Linear and Curvilinear Navigational Affordance Use

	Source	Dependent Variable	df	Mean Square	F	Sign.	Power
Metacognitive Factors	Corrected Model	Navigations	3	21	0.31	0.82	0.11
		Navigations Squared	3	1982	0.10	0.96	0.07
	Intercept	Navigations	1	71	1.07	0.31	0.18
		Navigations Squared	1	333	0.02	0.90	0.05
	SR	Navigations	1	27	0.41	0.52	0.10
		Navigations Squared	1	2349	0.12	0.73	0.06
	RE	Navigations	1	34	0.52	0.48	0.11
		Navigations Squared	1	477	0.02	0.88	0.05
	SR * RE	Navigations	1	40	0.60	0.44	0.12
		Navigations Squared	1	1022	0.05	0.82	0.06
	Error	Navigations	70	66			
		Navigations Squared	70	20076			
	Total	Navigations	74				
		Navigations Squared	74				
	Corrected Total	Navigations	73				
		Navigations Squared	73				

Table 12

RQ3 – Between-Subjects' Effects of Motivational Factors on Linear and Curvilinear Navigational Affordance Use

	Source	Dependent Variable	df	Mean Square	F	Sign.	Power
Motivational Factors	Corrected Model	Navigations	7	25	0.36	0.92	0.15
		Navigations Squared	7	6162	0.30	0.95	0.13
	Intercept	Navigations	1	0	0.00	1.00	0.05
		Navigations Squared	1	6696	0.32	0.57	0.09
	I	Navigations	1	0	0.00	0.98	0.05
		Navigations Squared	1	4851	0.23	0.63	0.08
	TA	Navigations	1	0	0.00	0.96	0.05
		Navigations Squared	1	4788	0.23	0.63	0.08
	ER	Navigations	1	0	0.01	0.94	0.05
		Navigations Squared	1	5196	0.25	0.62	0.08
	I * TA	Navigations	1	0	0.01	0.93	0.05
		Navigations Squared	1	2904	0.14	0.71	0.07
	TA * ER	Navigations	1	1	0.02	0.89	0.05
		Navigations Squared	1	2054	0.10	0.75	0.06
	I * ER	Navigations	1	3	0.04	0.84	0.06
		Navigations Squared	1	3098	0.15	0.70	0.07
	I * TA * ER	Navigations	1	5	0.08	0.78	0.06
		Navigations Squared	1	961	0.05	0.83	0.06
	Error	Navigations	66	69			
		Navigations Squared	66	20730			
	Total	Navigations	74				
		Navigations Squared	74				
	Corrected Total	Navigations	73				
		Navigations Squared	73				

Table 13

Percentages of Distractions Within the Valid Sample

Distraction Type	N	% of Valid Sample
No Distractions	22	25.3%
Person Interrupted	39	44.8%
Loud noise	19	21.8%
Phonecall or text	19	21.8%
Delay of more than 1 minute	9	10.3%
Other	22	25.3%
Total Reported Distractions	108	

Table 14

Comprehensive List of Other Types of Distractions

Other Distractions:
Audio books make it easy for me to tune out
Distraction by household pet
Doze off
Email
Friend watching TV mid-way through audio
got bored
Got lost mind drifted
Got sleepy
Had to stop a cat from going into a kitchen cabinet.
It was loud in the MSC
It's hard for me to listen. I would rather read on my own.
Just being bored of starring at nothing, lack of visuals
Multitasking
Roommates being loud
The reading had a weird pace and seemed to brush over necessary pauses in the reading.
The reading was very fast
The text never appeared
The voice talked too slow and it was hard to pay attention
Tiredness
TV
Vacuum cleaner noise
Went to restroom briefly

APPENDIX B - FIGURES

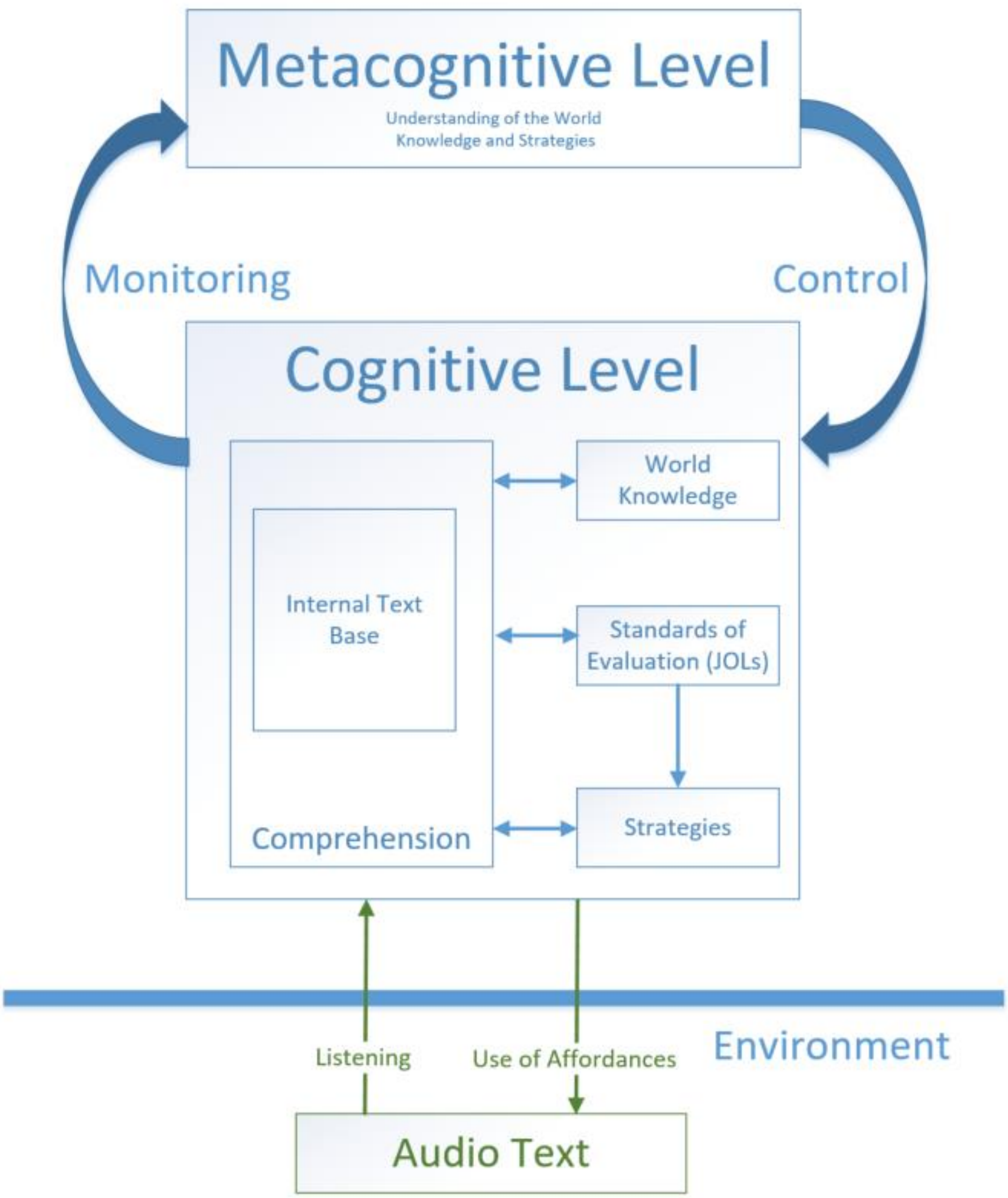

Figure 1. Model of comprehension monitoring and control with subsequent regulation of the environment through the use of affordances, adapted from Hacker (1998). Additions to the original model are colored in green.

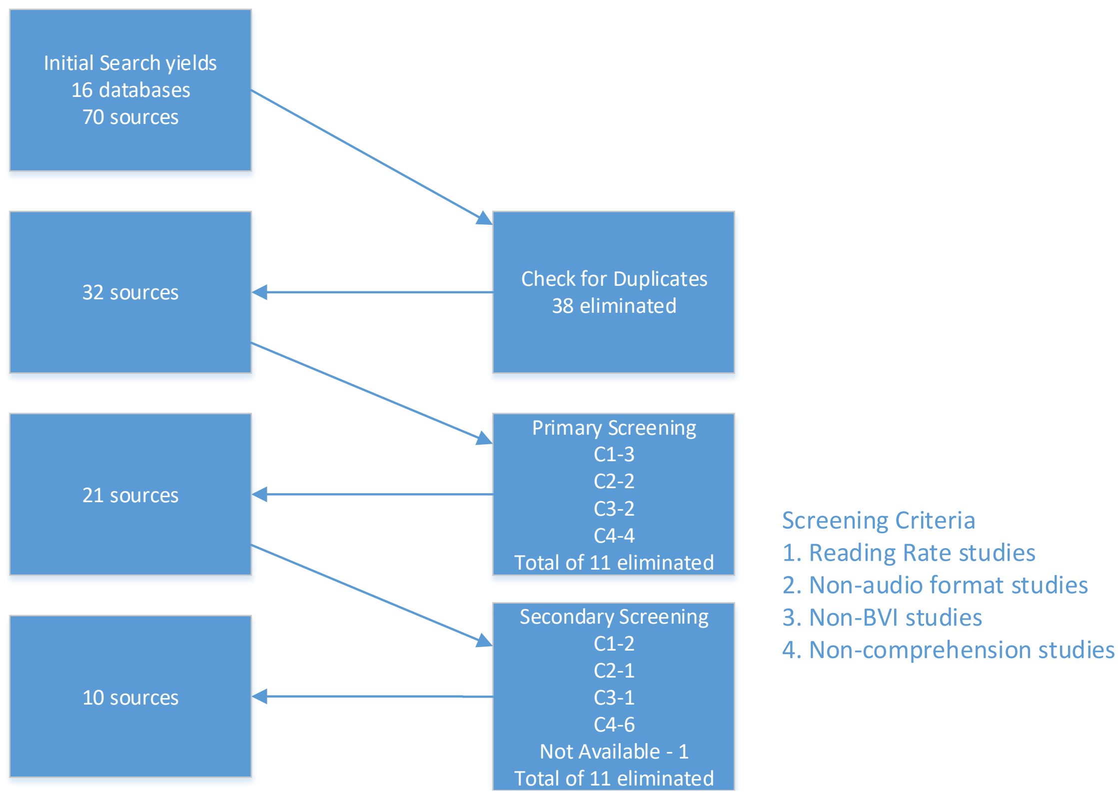

Figure 2. Systematic literature review process summary: Comprehension of audio text by blind and visually impaired learners.

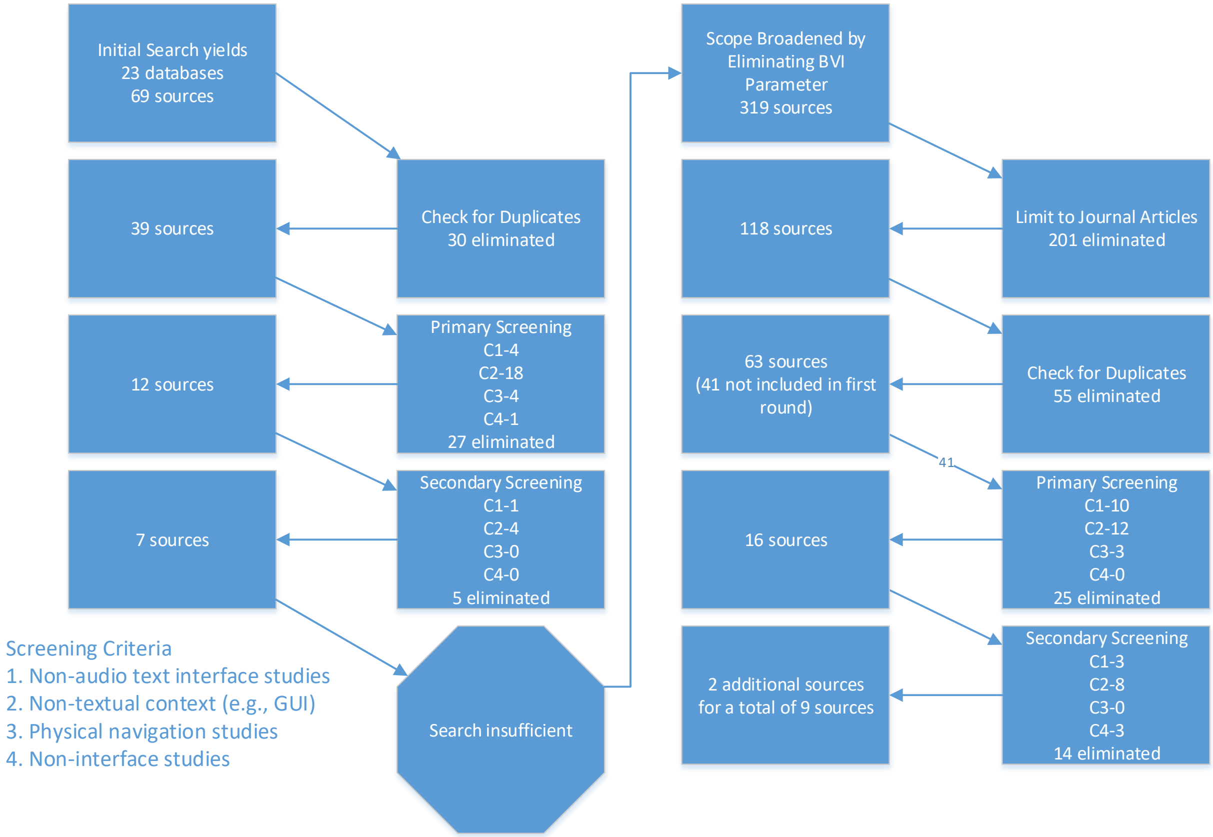

Figure 3. Systematic literature review process summary: Audio text interface design as it affects comprehension of audio text by blind and visually impaired learners.

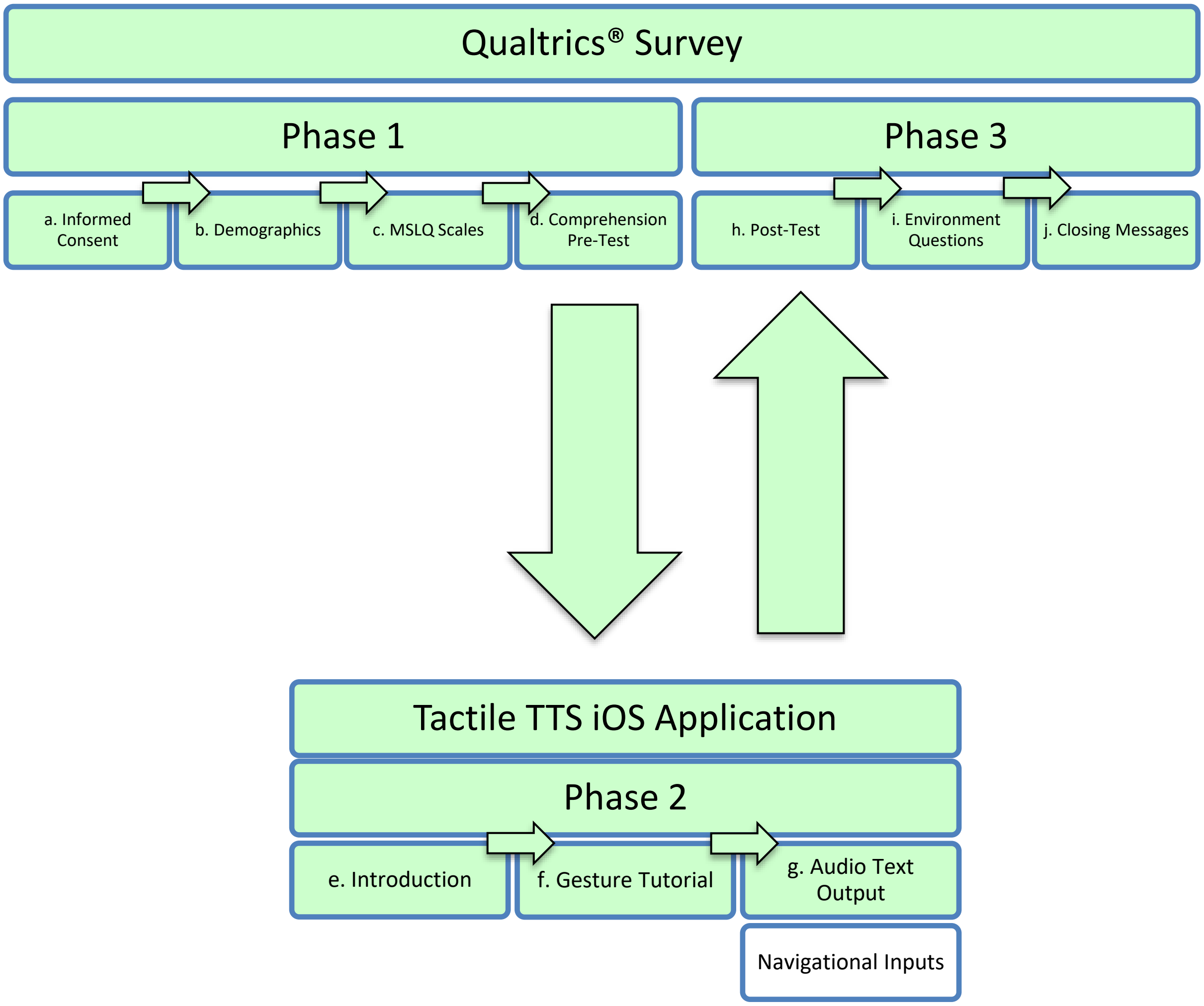

Figure 4. Study workflow.

```
participantGuid: 7CB3C369-35FD-4467-8DA7-112344B8387D
participantGroup: 2
participantTrial: 1
TPC,0,1454624439.67796
TPC,0,1454624441.79108
TPC,0,1454624443.57457
TFP,0,1454624445.49208
TBP,0,1454624447.67553
TF,0,1454624450.1252
TB,0,1454624452.02534
FP,32,1454624454.47778
FP,21,1454624454.84287
FP,356,1454624455.22423
P,979,1454624455.95996
C,979,1454624456.75934
P,994,1454624457.34338
C,994,1454624457.6428
F,1028,1454624459.00973
…
…
FP,4904,1454624486.72702
FP,5081,1454624486.97714
FP,5609,1454624487.26175
F,5603,1454624487.49294
F,5668,1454624487.74395
F,5713,1454624487.99287
F,5713,1454624488.24426
```

Figure 5. Example participant metadata log. The first three lines record the participant identifier, group and trial number. The remainder of the file records gestures during the study. Each line follows the format: Gesture type, location within the text, timestamp.

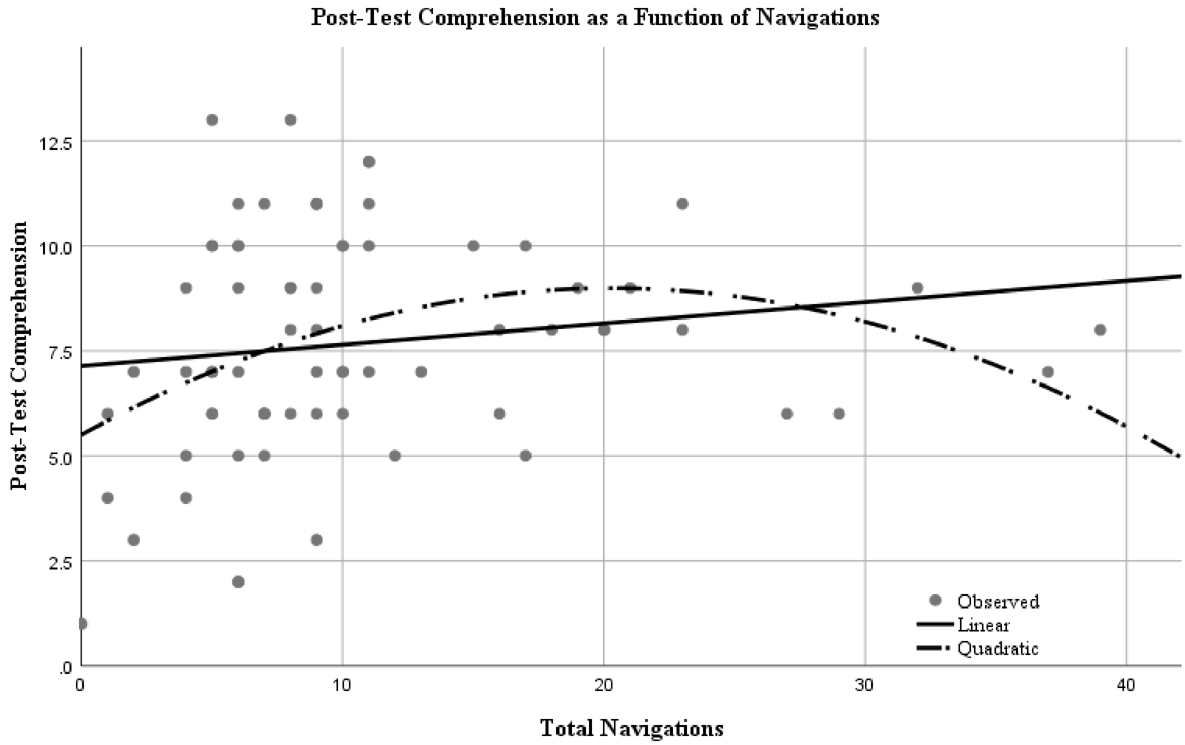

Figure 6. Post-test comprehension as a function of total navigations.

APPENDIX C – AUDIO TEXT

Heading. Minority Rights.

Decisions by courts to declare statutory laws and executive procedures as unconstitutional have been depicted as being counter to the majority rule principle fundamental to American democracy. Almost all government systems worldwide profess to follow the majority rule principle. Even hereditary monarchs and religious officials who hold power in theocracies claim the majority supports them.

We think the great genius of the American political system is that it is concerned with both majority rule and minority rights. The United States is among the small number of countries worldwide that make minority rights the more important, more fundamental concern.

As our constitutional system has evolved, we have come to recognize that, given the opportunity, majorities frequently choose to withhold from unpopular minority groups rights and privileges they enjoy themselves. The legislative and executive branches are structured to be more responsive to majority preferences. As a result, their actions have sometimes been detrimental to out-of-favor minorities.

Courts have been more active than the other two branches in championing the cause of minority rights under the law because that is the role assigned to them by federal and state constitutions. Although not all court decisions do so, judges have acted frequently to uphold and expand equal protection and equal application of criminal and civil laws.

Sub-Heading. Equal Protection Under Criminal Law.

In the United States, only a minority of citizens are accused or convicted of serious crimes. Those who are accused or convicted are an unpopular minority. Beginning in the 1950s, the you ess Supreme Court has, in a large number of its decisions, applied the federal minimum standards of justice in our national Constitution to state criminal procedures. Three principles stand out in their importance to state criminal procedures.

The first is the "exclusionary rule," which excludes illegally seized evidence from being presented in trials. The case is Mapp versus Ohio (1961). When miz Mapp refused to allow police to search her home, they ignored her and did so anyway. They found no evidence connecting her to a bombing for which she was suspected but rather found pornographic materials that were unlawful. She was arrested and convicted based on the gathered evidence. The Supreme Court overturned her conviction because without a search warrant specifying pornographic materials, the evidence was unlawfully obtained. Such evidence is to be excluded from the courts.

Although the initial principle was clearly drawn, the Supreme Court has weakened it with more recent decisions. The Court has established exceptions to the exclusionary rule to ensure the benefits of Fourth Amendment rights to privacy are not outweighed by the rule’s social costs. In

July 2016, it made a major modification in the exclusionary rule in Utah versus Strieff (2016). The court held that the exclusionary rule does not apply when an officer makes an illegal stop (that is, a stop without probable cause), runs a warrant check on the suspect's ID, learns that the suspect has an outstanding warrant, and then searches the suspect incident to arrest on the warrant. The court held that the evidence seized was admissible because the officer's discovery of the arrest warrant attenuated the connection between the unlawful stop and the evidence seized incident to arrest.

The second principle is the right to legal representation established in Gideon versus Wainwright (1963). Bari Gideon, unable to afford an attorney, unsuccessfully defended himself on a charge that he rifled a soda machine in a pool hall. While in prison, he studied law and successfully appealed his case to the Supreme Court. He maintained that it was unconstitutional for him to have been allowed to defend himself. Rather, he argued, the state should have provided an attorney. The Court agreed.

The third principle is the right to remain silent, first established Miranda versus Arizona (1966). Miranda confessed to a crime but was not told he could speak to an attorney and remain silent. The Supreme Court said the police had violated his rights to be informed. They had not told him that he had the right to remain silent, that anything he said might be used against him, and that he had a right to an attorney. This procedure of the police enumerating these rights has become called "Mirandizing" the accused.

One might think that the legal counsel one has a right to must be qualified or competent, but the you ess Supreme Court has not explicitly stated what level of competence is required to satisfy the right to counsel. In the absence of a national standard, states have determined their own standards. These standards vary widely.

States disagree on whether "qualified" requires previous experience or previous success in criminal law. In Texas, a defense attorney need only be qualified at the time he or she is appointed; previous or subsequent suspension of the attorney's license to practice law need not be considered. States disagree on whether or not the following are grounds for determining that defense counsel was inadequate: doing no work in preparation for trial, calling no witnesses, not challenging incompetent expert witnesses, permitting false testimony, sleeping during legal proceedings, having a debilitating medical condition, being disinterested in defending the accused because of racism, and having undisclosed business or romantic relationships with others involved in the case, including judges, prosecuting attorneys, and witnesses.

Sub-heading. Death Penalty.

The use of the death penalty has become an important criminal law issue in contemporary America. The number of states with death penalty punishments peaked in 1995 at 38. In 2016, 18 states have abolished the death penalty; the governors of five states have declared moratoriums on executions and four states have not had executions in the last ten years.

The you ess Supreme Court has not declared the death penalty unconstitutional, but it has acted to restrict its application. In 2002, the Court barred the execution of mentally retarded defendants. In 2005, it ruled that the Constitution bars the death penalty for crimes committed before the age of 18. The most recent restriction was in 2008. In Kennedy versus Louisiana (2008), the Court struck down as unconstitutional a Louisiana statute that allowed the death penalty for the rape of a child where the victim did not die.

The Kennedy versus Louisiana majority opinion stated that the 8th Amendment’s protection against excessive or cruel and unusual punishments "draws meaning from the evolving standards of decency that mark the progress of a maturing society." The court explicitly acknowledged "difficulties in administering the penalty to ensure against its arbitrary and capricious application."

Arbitrary and capricious application of the death penalty denies the equal protection guaranteed by the federal constitution and all 50 state constitutions. Throughout our history, African-Americans have been more likely to be convicted of capital crimes and more likely to be executed. The differences far exceed the standards for the statistical significance we use in this text. Even if capital punishment is constitutional in principle, the apparent inability or unwillingness of at least some state governments to administer it without racial bias may be arbitrary and capricious.

The procedure of execution by lethal injection has become an issue. In 2011, the European Union banned export of drugs used for lethal injection to the United States. In addition, a number of pharmaceutical companies in the United States stopped manufacturing drugs used for lethal-injection. When death-penalty states ran out of drugs used successfully in the past, they changed to different drugs. A series of botched executions by lethal injection in 2014 resulted in delays of executions and appeals that lethal-injection violated the you ess Constitution’s Eighth Amendment prohibition of cruel and unusual punishment. In Glossip versus Gross (2015) the Supreme Court ruled that Oklahoma’s use of lethal injection did not violate the ban on cruel and unusual punishment. The vote was 5 to 4, with the conservative justices in the majority.

Evidence of wrongful executions and racial bias notwithstanding, the American public favors the death penalty for persons convicted of murder. With the exception of the years 1957-1967, this has been true since polling began in 1937. According to the Gallup Poll, support for the death penalty peaked at 80 percent in 1994 and was 61 percent in 2015. So long as the decision is made by majority rule, the death penalty will continue in the United States in the foreseeable future. Only if the courts determine that the administration and implementation of the death penalty violates equal protection standards or the constitutional ban on cruel and unusual behavior is it likely to be abolished.

Sub-heading. Equal Protection Under Civil Law.

Perhaps the most important actions taken by courts to enhance equal protection under civil law have involved declaring criminal statutory laws unconstitutional and forbidding the states from criminalizing and punishing the behaviors involved. Voiding criminal laws and guaranteeing equal protection as a civil right under federal and state constitutions have advanced equal protection under civil law.

Consensus on which individuals and groups should receive equal protection has greatly changed since the Declaration of Independence and the Constitution of 17 89 were written. Originally, women and slaves had limited legal rights and no political rights. Even after slavery was changed from legal to illegal, state and local governments throughout the nation denied equal protection under law to African-Americans. In some states, inequality resulted from unequal application of laws. In other states, laws required separation of races. In every case, separation meant that African-Americans were denied opportunities available to others. It is worth noting that racial inequality occurred in more places after the abolition of slavery than before. Post-reconstruction racial discrimination was expanded from slaves to all African-Americans. Discrimination affected not only former slaves and their descendants but also individuals who had not been slaves--whether or not they had been in the United States or even alive when slavery was legal.

Today, the courts have established that equal protection cannot be arbitrarily withheld from individuals and groups on the basis of sex, race, national origin, religion, age, or disability. The contemporary legal standard is that government must have a legitimate reason for differential treatment. All three branches of the federal government have taken action to advance civil rights. Typically, action by some state governments to expand civil rights has come before comparable federal action. Nevertheless, it is the you ess Supreme Court that has led the way in declaring unconstitutional many occurrences of majority groups denying to unpopular minorities legal rights or privileges they grant to themselves.

The you ess Supreme Court’s landmark decision Brown versus Board of Education (1954) is one of the most important in the Court’s history. Prior to Brown, states had the option of operating dual, racially segregated public schools. Such segregation did not violate equal protection constitutional guarantees of due process or equal protection so long as segregated facilities and services were arguably equal. In Brown versus Board of Education, the Supreme Court declared that "separate but equal" public schools were inherently unequal and impermissible under the constitution. All public school systems were to be desegregated with "all deliberate speed." Southern states succeeded in delaying racial desegregation of public schools, but they were unable to prevent it.

Southern states’ efforts to prevent interaction between races included laws forbidding certain interracial activities. African-Americans had to use different restrooms and drinking fountains, wait for trains in separate areas, and ride in different sections of trains and buses. It was also a violation of criminal law for "whites" and "non-whites" to marry. Furthermore, couples legally married in other states committed criminal offenses by behaving as married couples in states where such marriages were forbidden. Illegal intermarriage was a felony punishable by many years in prison.

The you ess Supreme Court decision Loving versus Virginia (1967) declared that racial limitations on marriage violated the due process guarantees of the 14th Amendment. Not only did the Court decriminalize interracial marriage, it declared that marriage was one of the basic civil rights and the right to marry was a fundamental freedom. At one time or another, 41 states had laws making marriage or cohabitation between "whites" and "other races" illegal, although the definition of "other races" varied from state to state. All but 16 states had repealed these laws prior to the Loving versus Virginia decision.

In Lawrence versus Texas (2003) the you ess Supreme Court found that state laws could not make private sexual practices between consenting adults of the same sex criminal. The decision stated that, under the due process clause of the federal constitution, two adults have the right to engage in private con duct without government intervention. The majority opinion stated, "The Texas statute furthers no legitimate state interest which can justify its intrusion into the individual's personal and private life." The Lawrence versus Texas opinion expressly overruled the Court's decision in Bowers versus Hardwick (1986), which had come to an opposite conclusion. At the time, Texas was one of 13 states proscribing sodomy. Laws in four of those states, including Texas, made only homosexual sodomy illegal. Lawrence versus Texas also established privacy rights for both opposite sex and same sex couples.

The states remain free to define adulthood. As always, they reach different decisions. There are multiple definitions of adulthood across states and within states for different purposes.

You may be familiar with the concept "age of consent" but probably don't know that states can, and frequently do have different ages of consent for marriage without the approval of one's parents and ages of consent for adult private con duct. Most states have multiple exceptions to their nominal ages of consent. The minimum age is 16 in 30 states, 17 in 8 states and 18 in 12 states.

There is another limitation on private con duct between consenting adults related to marital status. Adultery is a criminal offense in about half the states, although there is no common definition of adultery. In some states, adultery is defined as relations between any two people not married to each other. In other states, at least one of the individuals involved must be married. In yet other states, adultery involves a married woman having relations with someone other than her husband. The criminal penalties for those found guilty range from a misdemeanor fine of $10 to a felony life sentence in prison. In the you ess military, adultery is a potential court-martial offense. The civilian statutes are rarely enforced, if ever. It is very difficult for the administration of any rarely enforced law to meet the equal protection standard.

APPENDIX D – INFORMED CONSENT

Tactile TTS Audio Learning Study

Please scroll down. This is an informed consent document which you must acknowledge before beginning the study.

You are invited to take part in a research study being conducted by David Sweeney, a researcher from Texas A&M University. This information is provided to help you decide whether or not to take part. If you decide to take part in the study, you will be asked to acknowledge your agreement to participate. If you decide you do not want to participate, there will be no penalty to you, and you will not lose any benefits you normally would have.

Why Is This Study Being Done?

The purpose of this study is to learn about audio comprehension and learner traits that can effect comprehension.

Why Am I Being Asked To Be In This Study?

You are being asked to be in this study because you are currently a student enrolled in Political Science 207 at Texas A&M University.

How Many People Will Be Asked To Be In This Study?

About 1500 people will be invited to participate in this study locally, although not all will participate.

Are there Alternatives to being in this study?

Yes, the alternative to being in the study is not to participate.

What Will I Be Asked To Do?

You will be asked to answer a series of questions about your learning preferences as they relate to the study of Political Science. When you complete the questions, you will be presented an audio text excerpt from the book Comparing the States and Communities: Politics, Government, & Policy in the United States by Harvey J. Tucker and Norman R. Luttbeg. Once the text is complete, you will be asked to answer some questions about the text you just heard. Your participation in this study will last up to 45 minutes. Once you answer the second set of questions, you are done with the study.

Will Photos, Video or Audio Recordings Be Made Of Me during the Study?

No, but your interaction with the application will be recorded and reported to the researcher. This includes how you interact with the audio text.

Are There Any Risks To Me?

The things that you will be doing are no more than risks you would come across in everyday life. Since the study asks questions about your learning strategies, you could feel discomfort at being asked these questions. Since the study asks questions about your understanding of the text, you could feel discouraged or anxious about your answers. Your responses will be coded with a random identifier, so there is no risk of identification from a breach of confidentiality of your answers. If you choose to signup for the prize drawing (detailed below), your email address will be collected, but not associated with your answers. This piece of identifying information will be separated from your answers when the study is complete. Otherwise, the researcher will not know your identity.

Are There Any Benefits To Me?

This study will present audio information from a common Political Science textbook used at Texas A&M University for Political Science 207 courses. This could be beneficial to you as a student.

Will There Be Any Costs To Me?

Aside from your time, there are no costs for taking part in the study.

Will I Be Paid To Be In This Study?

You will not be paid for being in this study. However, if you opt to enter the drawing, you will be asked to submit your email address. Two participants will be selected at random at the conclusion of the study to receive a $100 gift card.

Will Information From This Study Be Kept Private?

The individual records of this study will be kept private. No identifiers linking you to this study will be included in any sort of report that might be published. Research records will be stored securely and only the researcher will have access to the records.

Information about you will be stored in encrypted computer files protected with a password. This consent form will be filed securely in an official area. People who have access to your information include the Principal Investigator and research study personnel. Representatives of regulatory agencies such as the Office of Human Research Protections (OHRP) and entities such as the Texas A&M University Human Subjects Protection Program may access your records to make sure the study is being run correctly and that information is collected properly.

Information about you related to this study will be kept confidential to the extent permitted or required by law.

Who may I Contact for More Information?

You may contact the Principal Investigator, David Sweeney, to tell him about a concern or complaint about this research at 979-492-9598 or via email at sweeney@tamu.edu. You may also contact the Protocol Director, Susan Pedersen at 979-845-1128 or spedersen@tamu.edu.

For questions about your rights as a research participant, to provide input regarding research, or if you have questions, complaints, or concerns about the research, you may call the Texas A&M University Human Subjects Protection Program office by phone at 1-979-458-4067, toll free at 1-855-795-8636, or by email at irb@tamu.edu.

What if I Change My Mind About Participating?

This research is voluntary and you have the choice whether or not to be in this research study. You may decide to not begin, or to stop participating at any time. If you choose not to be in this study or stop being in the study, there will be no effect on your student status, medical care, employment, evaluation, relationship with Texas A&M University, etc.

STATEMENT OF CONSENT

I agree to be in this study and know that I am not giving up any legal rights by acknowledging this form. The procedures, risks, and benefits have been explained to me, and my questions have been answered. I know that new information about this research study will be provided to me as it becomes available and that the researcher will tell me if I must be removed from the study. I can ask more questions if I want by contacting the researcher as noted above, and I can still receive services if I stop participating in this study.

This informed consent document can be viewed any time by stopping and restarting the application.

To acknowledge your informed consent, click the arrow below. If you do not wish to participate, simply close the application.

APPENDIX E – METACOGNITIVE SKILLS LEARNING QUESTIONNAIRE:

QUESTIONS USED

NOTE: The MSLQ scale for each question is indicated in parentheses. These labels were not presented in the study.

Motivation

The following questions ask about your motivation for and attitudes about a college course. Think about a specific course you are currently taking or have taken to answer the questions. Remember there are no right or wrong answers, just answer as accurately as possible. Use the scale below to answer the questions. If you think the statement is very true of you, select 7; if a statement is not at all true of you, select 1. If the statement is more or less true of you, find the number between 1 and 7 that best describes you.

1. In a class like this, I prefer course material that really challenges me so I can learn new things. (I)
2. In a class like this, I prefer course material that arouses my curiosity, even if it is difficult to learn. (I)
3. The most satisfying thing for me in this course is trying to understand the content as thoroughly as possible. (I)
4. When I have the opportunity in this class, I choose course assignments that I can learn from even if they don't guarantee a good grade. (I)
5. Getting a good grade in this class is the most satisfying thing for me right now. (I)
6. The most important thing for me right now is improving my overall grade point average, so my main concern in this class is getting a good grade. (I)
7. If I can, I want to get better grades in this class than most of the other students. (I)
8. I want to do well in this class because it is important to show my ability to my family, friends, employer, or others. (I)

9. When I take a test I think about how poorly I am doing compared with other students. (TA)
10. When I take a test I think about items on other parts of the test I can't answer. (TA)
11. When I take tests I think of the consequences of failing. (TA)
12. I have an uneasy, upset feeling when I take an exam. (TA)
13. I feel my heart beating fast when I take an exam. (TA)

Learning Strategies

The following questions ask about your learning strategies and study skills for a college course. Think about a specific course you are currently taking or have taken to answer the questions. Again, there are no right or wrong answers. Answer the questions about how you study in this class as accurately as possible. Use the same scale to answer the remaining questions. If you think the statement is very true of you, select 7; if a statement is not at all true of you, select 1. If the statement is more or less true of you, find the number between 1 and 7 that best describes you.

1. When I study for this class, I practice saying the material to myself over and over. (RE)
2. When studying for this course, I read my class notes and the course readings over and over again. (RE)
3. I memorize key words to remind me of important concepts in this class. (RE)
4. I make lists of important items for this course and memorize the lists. (RE)
5. During class time, I often miss important points because I'm thinking of other things. (SR)
6. When reading for this course, I make up questions to help focus my reading. (SR)
7. When I become confused about something I'm reading for this class, I go back and try to figure it out. (SR)
8. If course readings are difficult to understand, I change the way I read the material. (SR)

9. Before I study new course material thoroughly, I often skim it to see how it is organized. (SR)
10. I ask myself questions to make sure I understand the material I have been studying in this class. (SR)
11. I try to change the way I study in order to fit the course requirements and the instructor's teaching style. (SR)
12. I often find that I have been reading for this class but don't know what it was all about. (SR)
13. I try to think through a topic and decide what I am supposed to learn from it rather than just reading it over when studying for this course. (SR)
14. When studying for this course I try to determine which concepts I don't understand well. (SR)
15. When I study for this class, I set goals for myself in order to direct my activities in each study period. (SR)
16. If I get confused taking notes in class, I make sure I sort it out afterwards. (SR)
17. I often feel so lazy or bored when I study for this class that I quit before I finish what I planned to do. (ER)
18. I work hard to do well in this class even if I don't like what we are doing. (ER)
19. When course work is difficult, I either give up or only study the easy parts. (ER)
20. Even when course materials are dull and uninteresting, I manage to keep working until I finish. (ER)

APPENDIX F – PRE-TEST QUESTIONS

Pre-Test Questions: Now you will be asked a brief set of multiple-choice questions related to the text you are about to hear. These questions are asked to gauge your prior knowledge of the topic.

Judicial activism has been defined to mean
- ❍ decisions that find government laws or behavior unconstitutional
- ❍ liberal decisions opposed by conservatives
- ❍ decisions that uphold and expand equal application of criminal law
- ❍ all of these

What has been the pattern of the number of violent and property crimes in the United States since 1994?
- ❍ both have increased
- ❍ both have decreased
- ❍ there has been little change
- ❍ numbers of crimes have increased but crime rates have decreased

Generally, Americans are affected more by
- ❍ the federal court system
- ❍ the state court system
- ❍ the federal courts in criminal cases
- ❍ the federal courts in civil cases

What is the typical name of a state's court of last resort?
- ❍ district court
- ❍ court of criminal or civil appeals
- ❍ state supreme court
- ❍ federal district court

Other than giving coverage to sensational trials, the mass media limit their coverage of the judicial branch largely to;
- ❍ decisions made by state courts
- ❍ decisions made by the U.S. Supreme Court
- ❍ decisions made by activist judges
- ❍ liberal decisions made by liberal activist judges

State and local courts;
- ❍ deal with more than 95 percent of criminal cases
- ❍ deal with more than 95 percent of civil cases
- ❍ employ more than 95 percent of judges and hearing officers
- ❍ all of the above

Our court system is best described as;
- a unitary system, federal law always takes precedence
- a dual system, one state and one federal
- a three part system—federal, state and local
- the judicial branch almost always deferring to the legislative and executive branches

The majority of court cases filed are;
- criminal cases
- civil cases
- resolved by trial
- referred to mediators employed by state governments

Justice of the peace courts;
- are being phased out by many states
- are increasing in number in all but five states
- are becoming more important venues for trying felony criminal cases
- none of the above

Which of the following methods of selecting judges is the newest?
- the partisan election plan
- the legislative appointment plan
- the nonpartisan election plan
- the Missouri plan or merit system

APPENDIX G – POST-TEST COMPREHENSION QUESTIONS

Comprehension

Now you will be asked a brief set of multiple-choice questions about the text you just heard. At the end of the survey, you will be able to enter the prize drawing by entering your email address. Winners will be contacted by the end of the semester.

Regarding rights, the American political system is most concerned with;

- ❍ minority rights
- ❍ majority rights
- ❍ majority and minority rights equally
- ❍ neither minority or majority rights

When compared with the judicial branch, the the legislative and executive branches are structured to be;

- ❍ more responsive to majority rights
- ❍ more responsive to white citizens
- ❍ more responsive to minority rights
- ❍ more responsive to the "silent majority"

When applying minimum standards of justice to state criminal procedures, which of the following was not cited as one of importance?

- ❍ the exclusionary rule
- ❍ the right to a trial by jury
- ❍ the right to legal representation
- ❍ the right to remain silent

The exclusionary rule forbidding the use of evidence unlawfully obtained;

- ❍ was established in Miranda v. Arizona
- ❍ always applies when police make searches without a search warrant or probable cause
- ❍ has exceptions that take into account the rule's social costs
- ❍ all of the above

The Supreme Court case that required those accused of crimes to be provided legal counsel at government expense if they cannot afford it was;

- ❍ Escobedo versus Illinois
- ❍ Miranda versus Arizona
- ❍ Baker versus Carr
- ❍ Gideon versus Wainwright

A lawyer may NOT be "Qualified Counsel," if he or she engages in which of the following?
- ❍ has a debilitating medical condition
- ❍ has an undisclosed romantic relationship with one of the involved parties
- ❍ sleeps during legal proceedings
- ❍ all of these may be disqualifying

In the United States, the death penalty is barred under all of the following conditions EXCEPT;
- ❍ the defendant is mentally retarded
- ❍ the defendant is under 18 years of age
- ❍ the defendant is not a U.S. citizen
- ❍ the defendant is guilty of rape and the victim survives

African-Americans are executed at a greater rate than other races. This is because;
- ❍ there is a greater proportion of African-Americans who commit crimes
- ❍ the courts may be acting in an arbitrary and capricious manner
- ❍ although the rate is greater, it is not statistically significant
- ❍ none of the above

In Kennedy versus Louisiana, the U.S. Supreme Court stated that the 8th Amendment's protection against excessive or cruel and unusual punishments "draws meaning from ________ that mark the progress of a maturing society.
- ❍ equal protection standards
- ❍ the evolving standards of decency
- ❍ constitutional interpretations
- ❍ a greater respect for human life

Equal protection under civil law;
- ❍ is guaranteed under federal law but not under state law
- ❍ is guaranteed under state law but not under federal law
- ❍ is guaranteed under both state and federal law
- ❍ is guaranteed by neither federal or state law; such guarantees apply only to criminal law

In Glossip versus Gross, the U.S. Supreme Court ruled that the use of lethal injection did not violate the ban on cruel and unusual punishment. Which state was cited in the case?
- ❍ Texas
- ❍ Alabama
- ❍ Louisiana
- ❍ Oklahoma

The death penalty;
- ❍ has been limited in application by the U. S. Supreme Court
- ❍ has been applied wrongfully and with racial bias
- ❍ is favored by a majority of the American public for persons convicted of murder
- ❍ all of the above

The death penalty in the U.S. could eventually be abolished if;
- it is found to violate equal protection standards
- popular opinion shifts in favor of a ban
- the Supreme Court rules that a specific method of execution (e.g., lethal injection) is a form of cruel and unusual punishment
- two-thirds of states independently abolish the death penalty

In the post-reconstruction U.S., discrimination;
- was lessened because of the Emancipation Proclamation
- was generalized to all African-Americans, regardless of whether or not they were slaves
- was greater for former slaves
- was more common in rural areas

The courts have established that equal protection cannot be withheld arbitrarily from individuals and groups on the basis of;
- religion
- disability
- age
- all of the above

The landmark case that ended school segregation was;
- Escobedo versus Illinois
- Brown versus Board of Education
- Baker versus Carr
- Gideon versus Wainwright

All of the following activities were cited in the text as legally prohibited for African-Americans EXCEPT;
- separate restaurants
- separate restrooms and drinking fountains
- separate accommodations on trains and buses
- interracial marriage

The U.S. Supreme Court has declared that marriage is a “basic civil right” and a “fundamental freedom.” Which case led to this decision?
- Loving versus Virginia (1967)
- Bowers versus Hardwick (1986)
- Lawrence versus Texas (2003)
- United States versus Windsor (2013)

The U.S. Supreme Court has declared that state laws could not make private sexual practices between consenting adults of the same sex criminal. Which case led to this decision?

- ❍ Loving versus Virginia (1967)
- ❍ Bowers versus Hardwick (1986)
- ❍ Lawrence versus Texas (2003)
- ❍ United States versus Windsor (2013)

Legally speaking, all of the following are true about adultery EXCEPT;

- ❍ adultery is a crime in about half the states in the U.S.
- ❍ the definition of what constitutes adultery varies from state to state
- ❍ penalties for adultery range from a $10 fine to a felony life sentence
- ❍ in the military, adultery is a serious offense, but rarely enforced

APPENDIX H – ENVIRONMENTAL QUESTIONS

Did you encounter any of the following distractions during the study? Check all that apply.

- ❑ I didn't encounter any distractions or interruptions
- ❑ Loud noise that made the audio text hard to hear
- ❑ Distracting Interruption such as a person
- ❑ Phone call or text that you answered
- ❑ Delay of more than 1 minute during the study
- ❑ Other

    If "other," please specify:

Did you encounter any problems with the app such as a freeze or crash?

- ❍ Yes
- ❍ No

Did you use a friend's mobile device to take the study prior to this attempt?

- ❍ Yes
- ❍ No

Did you take notes during the study?

- ❍ Yes
- ❍ No

APPENDIX I – STUDY INSTRUCTIONS

Control Group Instructions

This application will present a selection from the textbook, Comparing the States and Communities, by Harvey Tucker and Norman Luttbeg. The material will be presented as audio text and should take about twenty minutes to listen to. Be sure you have a quiet place that is free from distractions to take the study.

You will be able to pause and start the audio text by tapping the screen.

During the audio text portion of the study, a progress bar will be visible that will indicate how far in the text you have gone.

At the end of the audio, you will be directed to a survey that will ask you a brief set of multiple choice questions about the text you just heard. Please answer all of the questions.

At the end of the survey, you will be prompted to enter your email address if you want to enter the prize drawing. No other identifying information will be collected. Winners will be contacted by the end of the semester.

When you are ready to proceed, tap the Continue button in the upper right corner of this screen.

Experimental Group Instructions

This application will present a selection from the textbook, Comparing the States and Communities, by Harvey Tucker and Norman Luttbeg. The material will be presented as audio text and should take about twenty minutes to listen to. Be sure you have a quiet place that is free from distractions to take the study.

You will be able to navigate the audio text using swipe and tap gestures detailed in the next screen, including navigating the audio text by sentence and paragraph.

During the audio text portion of the study, a progress bar will be visible that will indicate how far in the text you have gone.

At the end of the audio, you will be directed to a survey that will ask you a brief set of multiple choice questions about the text you just heard. Please answer all of the questions.

At the end of the survey, you will be prompted to enter your email address if you want to enter the prize drawing. No other identifying information will be collected. Winners will be contacted by the end of the semester.

When you are ready to proceed, tap the Continue button in the upper right corner of this screen.

APPENDIX J – ORIENTATION INSTRUCTIONS

This screen provides an opportunity for you to practice navigation gestures. You may be able to access the training easier by turning your device sideways. There are five types of navigation. Tapping the screen alternaytes [misspelled to correct pronunciation] pausing and starting the speech. Swiping right with a single finger navigates back one sentence. Swiping left navigates forward one sentence. By swiping with two fingers, you will navigate by paragraph. When you are ready to start the study, tap the continue button at the top of the view. Try these gestures now.