EYES-FREE INTERACTIONS WITH AURAL USER INTERFACES

Romisa Rohani Ghahari

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

August 2015

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

Davide Bolchini, Ph.D., Chair

Michael Justiss, Ph.D.

Doctoral Committee

Sarah Koskie, Ph.D.

August 11, 2015

Xi Niu, Ph.D.

Stephen Voida, Ph.D.

DEDICATION

To my parents.

Acknowledgements

It is a pleasure to thank those who have made this dissertation possible. First and foremost, I owe sincere and earnest thankfulness to my advisor, Dr. Davide Bolchini. Dr. Bolchini was the person who had the most influence on my academic profession. It has been an incredible experience working with him, and I always feel excited when thinking about how much I have grown during these years under his continuous support and inspiring instructions. I also want to thank Dr. Bolchini for his patient in helping me correct my mistakes. Without his help, I might not have been able to overcome the challenge so smoothly.

Many thanks to my research committee members: Dr. Michael Justiss, Dr. Sarah Koskie, Dr. Xi Niu, and Dr. Stephen Voida. Thanks go out to Dr. Justiss and Dr. Koskie for letting me use the driving simulation lab facility and helping me to be familiar with the logistics of conducting a study in a simulation lab. Their guidance helped me enhance the experimental design found in Chapter 6 of this dissertation. I want to thank Dr. Xi Niu for her feedback on this dissertation although she served as my committee member for the last one year. Dr. Xi Niu's feedback allowed me to enhance the data analysis section. Thanks also need to go out to Dr. Stephen Voida for his continuous feedback on the whole dissertation. He had provided me with many constructive comments to strengthen the reasoning on the core part of my dissertation. I would also like to thank Dr. Mark Pfaff, who served as my research committee member for a year. Dr. Mark Pfaff helped me a lot shaping the experimental design and statistical tests I needed to perform throughout this dissertation.

My life in the US would not have been as wonderful as it has been without the company of my colleagues and friends in the USER Lab. Thanks go out to Dr. Mexhid Ferati, Dr.

Tao Yang, and Dr. Afarin Pirzadeh for teaching me both the experimental design and statistical analysis skills. I always admire how knowledgeable they were in those fields. Thanks also need to go out to Dr. Jennifer George-Palilonis for helping me with the initial stages of this dissertation. Her background in journalism enabled me with discovering the news domain much better than discovering it on my own. I would also like to thank Dr. Arthur Liu, Siavash Mortazavi, Dr. Hossain Gahangir, Lindsay Kaser, Kevin Pritchard, Callie Archibald, Dr. Sung Pil Moon, Shiva Ladan, Prathik Gadde, Debaleena Chattopadhyay, and Yuan Jia for their help and ideas, all of which helped strengthen my dissertation. I wish you all good luck in your future endeavors.

I would also like to acknowledge the support of the National Science Foundation and the Funding Opportunities for Research Commercialization and Economic Success at IUPUI Office of the Vice-Chancellor for Research. Thanks to all participants for their contributions to this research.

Last, but not least, I would like to thank my parents and my sister for encouraging me to continue with my higher education, and for supporting me both emotionally and financially whenever I needed. I understand allowing me not to be by their sides for so many years was a tough decision, but thanks for letting me experience and accomplish this degree.

Abstract

Romisa Rohani Ghahari

EYES-FREE INTERACTIONS WITH AURAL USER INTERFACES

Existing web applications force users to focus their visual attentions on mobile devices, while browsing content and services on the go (e.g., while walking or driving). To support mobile, eyes-free web browsing and minimize interaction with devices, designers can leverage the auditory channel. Whereas acoustic interfaces have proven to be effective in regard to reducing visual attention, a perplexing challenge exists in designing aural information architectures for the web because of its non-linear structure.

To address this problem, we introduce and evaluate techniques to remodel existing information architectures as "playlists" of web content − *aural flows*. The use of aural flows in mobile web browsing can be seen in ANFORA News, a semi-aural mobile site designed to facilitate browsing large collections of news stories. An exploratory study involving frequent news readers (n=20) investigated the usability and navigation experiences with ANFORA News in a mobile setting. The initial evidence suggests that aural flows are a promising paradigm for supporting eyes-free mobile navigation while on the go. Interacting with aural flows, however, requires users to select interface buttons, tethering visual attention to the mobile device even when it is unsafe.

To reduce visual interaction with the screen, we also explore the use of simulated voice commands to control aural flows. In a study, 20 participants browsed aural flows either through a visual interface or with a visual interface augmented by voice commands. The results suggest that using voice commands decreases by half the time spent looking at the device, but yields similar walking speeds, system usability and cognitive effort ratings as using buttons.

To test the potential of using aural flows in a context featuring higher cognitive load and distraction, a study (n=60) was conducted in a driving simulation lab. Each participant drove through three driving scenario complexities: low, moderate and high. Within each driving scenario, participants went through an alternative aural application exposure: no device, voice-controlled aural flows (ANFORA*Drive*) or an alternative, commercially available solution (Umano). Results suggest that voice-controlled aural flows do not affect distraction, overall safety, cognitive effort, driving performance or driving behavior when compared to the no-device condition.

Davide Bolchini, Ph.D., Chair

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

Accessing the mobile web while on the go in a variety of contexts (e.g., walking, standing, jogging or driving) is becoming increasingly pervasive (Kane, Wobbrock, & Smith, 2008; Schildbach & Rukzio, 2010; Zhou, Rau, Zhang, & Zhuang, 2012). Mobile users are often engaged in another activity while looking at their mobile screens, making such actions inconvenient, distracting and, sometimes, dangerous (Anhalt et al., 2001; Christian, Kules, Shneiderman, & Youssef, 2000; Garlan, Siewiorek, Smailagic, & Steenkiste, 2002; Yang et al., 2011). Although existing visual user interfaces are efficient in regard to supporting the quick scanning of a page, they typically require highly focused attention and may not work well while walking on a busy street, crossing the road or driving a car. In order to combat this challenge, this research seeks to explore novel ways by which to enable users to effectively access the mobile web while on the go.

In our preliminary work, we introduced the ANFORA (Aural Navigation Flows on Rich Architectures) framework – a set of techniques aimed at remodeling existing web information architectures as linear, aural flows that can be listened to with minimal interaction via a device using touch or gesture (Ghahari & Bolchini, 2011; Rohani Ghahari, George-Palilonis, & Bolchini, 2013). Aural flows are concatenated sequences of pages extracted in real-time from web sources and played to users on their mobile devices, much like playlists for listening to music. They enable a new class of aural and semi-aural (i.e., a combination of visual and aural interfaces) applications and are anticipated to minimize the visual attention required for the use of mobile devices, while, at the same time, maximize consumption of relevant content without compromising safety during multitasking.

In order to investigate the potential of ANFORA, we applied our concept to the news domain because news websites are content-intensive and employ complex navigational structures. News consumption on mobile devices is also increasing, making news content an interesting test bed for aural browsing. Our approach was exemplified by ANFORA News, a set of semi-aural mobile application prototypes that generate real-time aural flows from web sources and enable users to listen to collections of news stories while on the go.

Interacting with aural flows using existing mechanisms, such as touch or gesture (Ghahari & Bolchini, 2011; Rohani Ghahari et al., 2013), forces users to pay attention to displays. To relieve users from this potential distraction and unleash a more complete eyes-free experience, we investigated voice commands. For example, Apple Siri™ has been marketed as the solution for eyes-free experiences for users on the go by enabling them to have more natural interactions using voice commands (Lager, 2012). iPhone users can check the weather, send tweets, post to Facebook, schedule meetings, find contacts, get directions and send texts using Siri. However, if users want to access the latest news stories on their iPhones, Siri will direct them to a Google page containing a list of news stories from different sources. This example demonstrates how today's most advanced consumer products for mobile voice browsing fail to provide fluid access to arbitrary web content unless the capabilities for interacting with that content are explicitly pre-programmed into the interaction agent.

To solve the above-mentioned challenges, the two main questions addressed by this research are: How do semi-aural mobile applications support users engaged in web navigation, while also carrying out a parallel primary task with lower (e.g., walking) and higher cognitive loads (e.g., driving)? How do different input modalities affect the user experience, while one is interacting with semi-aural mobile applications? Accordingly, this dissertation presents four interconnected research projects that explore these questions. The first project presents the different types of aural flows that underlie the ANFORA framework as well as the application of this framework on the National Public Radio (NPR) news website. Content-rich websites can adopt ANFORA to automatically convert their content to playlists that can then be listened to on the go.

To evaluate the ANFORA framework, the second project explores how well ANFORA supports an eyes-free browsing experience while walking. This project also explores the usability, enjoyment, strengths and weaknesses of the ANFORA framework. The results of this exploratory study suggest that the ANFORA framework minimizes visual engagement with the mobile device screen.

However, this framework still requires that the users interact with buttons and gestures, which requires visual attention. As such, in order to reduce the necessary visual attention to the screen, the third project establishes novel navigation vocabularies to aurally interact with the content playlist using voice commands. To support a more fluid and natural control of the aural flows, this project iteratively creates, deploys and experimentally evaluates the usability of a set of voice commands for aural web browsing on mobile devices. This project also enables us to understand the users' preferences for different voice commands that can be used to control the aural flows. We manifest the design ideas and vocabulary for the commands in a prototype named *Linkless* ANFORA.

Finally, to enhance the effectiveness and efficiency of aural web navigation, we explore the potential and limits of the voice-controlled aural flows on the user experience by performing a set of evaluation studies involving participants using mobile devices while walking. To understand how *Linkless* ANFORA will apply to the driving scenario, the fourth project evaluates the impact of voice-controlled aural flows on drivers in a driving simulation lab. This project is a significant next step because it evaluates the idea of voice-controlled aural flows in a context featuring higher cognitive load and distraction compared to the context of walking. Therefore, fourth project presents how the paradigm of aural flows for the news domain could impact the user experience, especially in regard to distractions, overall safety, cognitive efforts, driving performance and driving behavior.

The rest of the chapters in this dissertation are organized as follows. Chapter 2 reviews the theoretical background of this dissertation, which includes a discussion of aural user interfaces, voice user interfaces and their application in driving, and various distractions while using mobile devices. Chapter 3 introduces the ANFORA framework as it is related to remodeling existing web information architectures into aural flows and presents the resulting design issues raised by the framework and the ANFORA News prototype. Chapter 4 presents the preliminary evaluation of the ANFORA framework showing how aural flows support an eyes-free browsing experience while walking and listening to web content. Chapter 5 introduces *Linkless* ANFORA along with the voice command vocabulary and presents the findings of a second, controlled evaluation study. Chapter 6 evaluates the voice-enabled aural flows in the driving context with 60 participants. Chapter 7 summarizes the contributions of this dissertation, and Chapter 8 discusses possible future research directions.

Chapter 2. Review of Theoretical Background

Zhang and Lai (2011) noted that a number of studies has been conducted that suggests guidelines for modifying desktop-based websites to be usable on mobile devices for *visual consumption*. However, little research exists in regard to modifying desktop-based websites to be usable on mobile devices for *aural consumption*. This dissertation is rooted in five areas: (1) the theoretical background behind visual and auditory channels, (2) aural user interfaces, (3) solutions for automated browsing concepts, (4) voice user interfaces and their application in regard to use while driving and (5) research on distractions while using mobile devices.

2.1. Theoretical Background Behind Visual and Auditory Channels

The Multiple Resource Theory (MRT) (Wickens, 1980; Wickens, 2002) explains the importance of decoupling visual and auditory channels (Figure 1). This theory originated from an examination of how people time-share two or more number of activities. The examination showed that visual-auditory task (cross-modal) combinations could be time-shared more efficiently (in terms of performance and parallel processing) than either visual-visual or auditory-auditory (intramodal) task combinations. For example, the tasks of walking, monitoring the environment and listening to the content of a website simultaneously were performed more efficiently compared to the tasks of walking and browsing website content visually. The reason why cross-modal combinations are more efficient is because two different resources (i.e., visual and auditory resources) are used at the same time, while, in the intra-modal combinations, the same resource is used simultaneously.

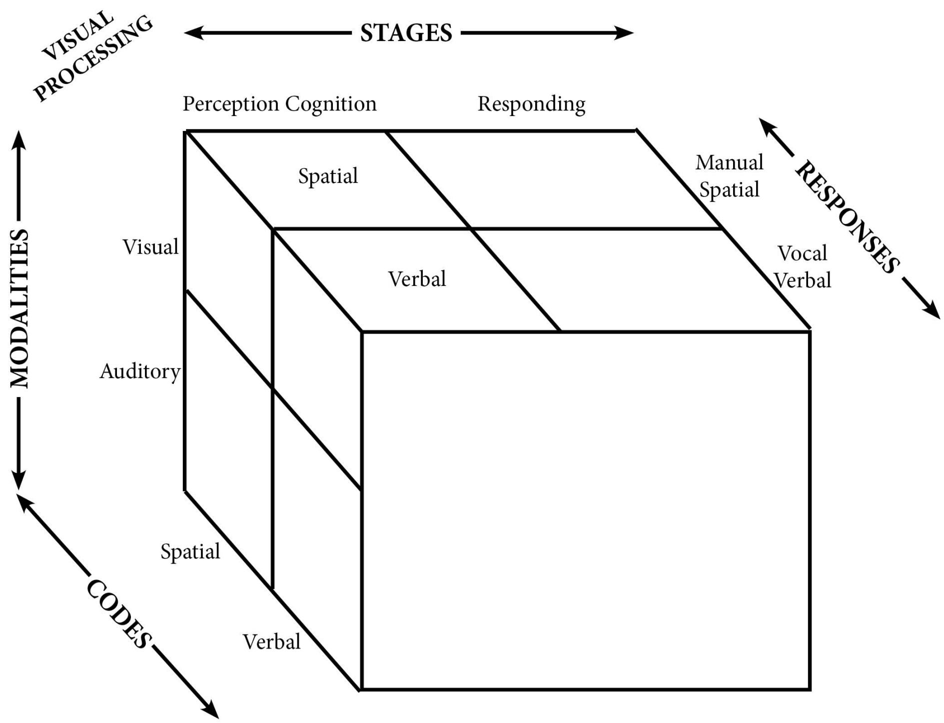
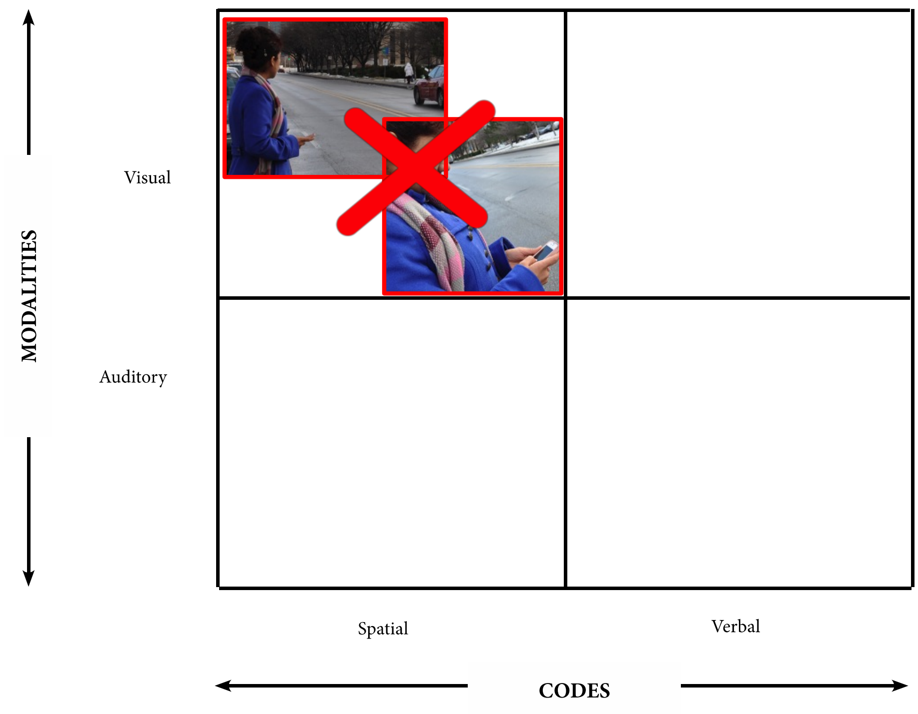
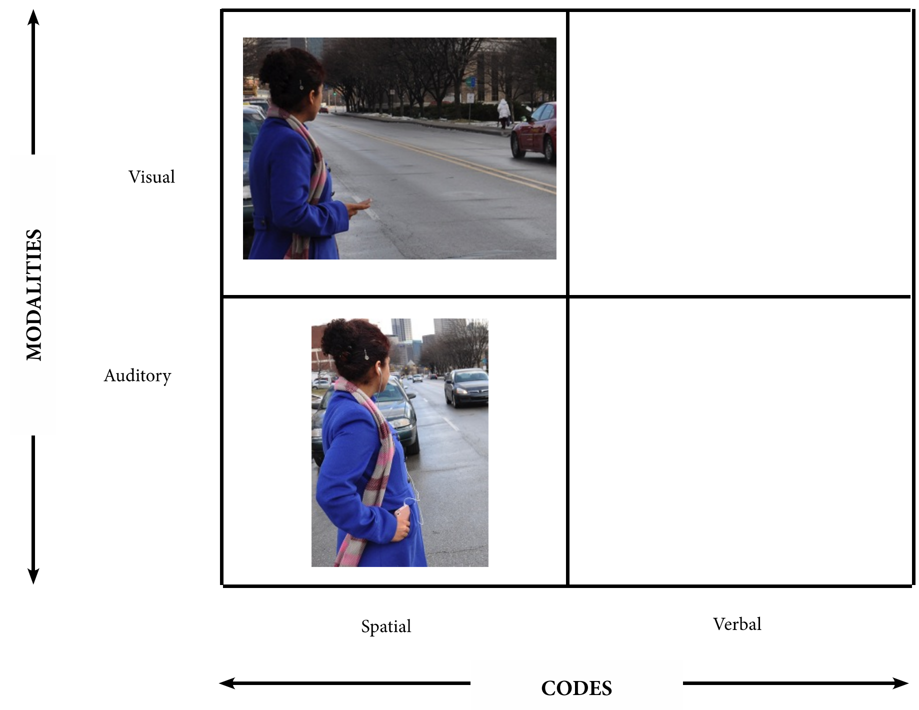


Figure 1. Multiple Resource Theory (MRT) (Wickens, 1980; Wickens, 2002).

Imagine a scenario in which users are involved in the tasks of walking and monitoring the environment, which uses the users' visual resources. When users time-share a task of browsing a website with the task of walking and monitoring the environment, they, again, use their visual resources (Figure 2a). Therefore, the performances related to browsing a website and walking are reduced, as outlined by the Multiple Resource Theory. In order to address this problem, users could use their auditory resources, instead of their visual resources, to browse the website (Figure 2b).



(a)



(b)

Figure 2. Application of the MRT as related to the tasks of walking, monitoring the environment and browsing web content at the same time. (a) Using visual resources for both of the tasks (monitoring the environment and browsing web content) simultaneously is less efficient in terms of performance than (b) using visual resources to monitor the environment and auditory resources to browse web content.

In addition, the auditory channel is omni-directional, meaning that information can be perceived from any direction. For example, listeners do not need to focus on a specific direction to hear sound. Therefore, users can direct their visual attentions to other tasks with the benefit of being able to focus on different things while listening to the information (Baldwin, 2012). Overall, this theory explains why using the auditory channel in addition to the visual channel can create opportunities for improving dual-task performances in a variety of contexts.

2.2. The Value of Aurally Navigated User Interfaces

A number of studies have emphasized using audio interfaces over visual interfaces to consume content, as well as provided reasons why audio interfaces may be preferred. Recent studies have shown that audio interfaces in cars are less distracting compared to traditional visual interfaces (Brumby, Davies, Janssen, & Grace, 2011). Users, however, select the modality according to their performance objectives. For example, Li, Baudisch, and Hinckley (2008) introduced the blindSight prototype, which helps users access calendars and contact lists via audio feedback, instead of looking at a screen. This study showed how audio interfaces could allow users to access quickly and interact with systems, while engaged in other primary tasks. In another study, Zhao, Dragicevic, Chignell, Balakrishnan, and Baudisch (2007) discussed five reasons why visual feedback might not be feasible: “competition for visual attention, absence of a visual display, user disability, inconvenience and reduction of battery life” (p. 1395).

A number of domains make use of audio navigation strategies, including audio museum navigation guides, audio books and audio playlists. Audio museum navigation guides allow users to carry a PDA in a museum to listen to linear information related to the artwork. Some of the examples of audio museum navigation guides include Ec(h)o (Wakkary & Hatala, 2007) and the Multimedia Museum Guide (Zancanaro, Stock, & Alfaro, 2003), both of which allow the user to pause, fast forward, rewind or stop the presentation by tapping on the PDA display.

Digital Talking Books (DTBs) are another form of text-based content that can be aurally navigated. DTBs give access to the full text of books, allowing users to interact with it using a keyboard (Morley, 1998). Likewise, the Mobile Rich Book Player prototype is a type of DTB that uses the Windows Mobile platform. However, for this platform, vast amounts of information cannot be displayed at once, since the screen size is too small. In order to overcome this drawback, its developers have implemented tabs and a variety of pages that can be navigated using a minimal set of physical buttons (Duarte & Carriço, 2009).

In addition, Jain and Gupta (2007) presented a system called VoxBoox, which generates automatic interactive talking books. This system converts digital books to audio books and makes them accessible to visually-impaired users using voice commands for navigation. Recently, commercial services, such as audible (Audible, 2015), also offer audio books via an iPhone application and users can download the audio books to listen to them on the go.

Capti narrator (Borodin et al., 2014) and Voice Dream (Voice Dream, 2015) are two types of audio playlists that allow users to add content or web pages to their playlists. For example, users can select a pdf or a Word document via their Dropbox or Google drive accounts and add them to their playlists to listen to later. Once users have populated their playlists with their favorite content, they can listening to the content.

2.3. Automated Browsing

Since the ANFORA framework is based on the notion that the aural flow allows the user to automate browsing tasks, it is worth acknowledging some similar technologies that exist to implement automated browsing. Automating repetitive browsing tasks, such as checking email and paying bills, can reduce user interactions with an application. Some of transactions might need the user's visual attention and feedback, while others can happen automatically (Borodin, 2008). For example, WebVCR allows users to record and replay their browsing steps (e.g., filling out a series of forms to access data on travel websites) in smart bookmarks as shortcuts to web content. This feature exists so that users do not have to repeatedly and manually enter the information each time they interact with the application. The pages involved in these browsing steps are hard-to-reach and, as such, are good candidates for this shortcut strategy (Anupam, Freire, Kumar, & Lieuwen, 2000).

Similarly, Chickenfoot, a Mozilla Firefox extension, allows users to automate and customize their web experiences without changing the source code of the website. Chickenfoot provides a programming environment in the sidebar of a web browser that allows users to write scripts to manipulate and automate web pages. This automation helps reduce tedious repetition of tasks (Bolin, Webber, Rha, Wilson, & Miller, 2005). Hence, the notion of automated browsing is not new in the field of Human-Computer Interaction (HCI), but its application in regard to aural navigation is new and will, hopefully, create new opportunities for browsing content-rich websites on the go.

2.4. Voice User Interfaces and Their Application in Regard to Driving

2.4.1. Voice Input User Interfaces

Recently, several studies have investigated the importance of voice commands as an interaction medium. For example, Aural Language for VoiceXML Interpretation and

Navigation (ALVIN) is a voice-based scripting language that allows users to define navigation strategies. It is completely voice/audio-based and intended to be used with voice/audio-only devices, such as telephones (Nichols, Gupta, & Wang, 2005).

Along the same line, the Dynamic Aural Web Navigation (DAWN) system translates HTML pages into VoiceXML pages (Gupta, Raman, Nichols, Reddy, & Annamalai, 2005). DAWN presents a small set of global voice commands for moving across documents, such as “skip” and “back.” It also allows users to create and attach voice anchors or labels to any part of a document in order to return to those points later simply by saying the name of the label.

Another example of a system that uses voice commands is the Web-based Interactive Radio Environment (WIRE), an in-car voice browser designed to be used safely by a driver while in transit. WIRE supports interactions from drivers via physical buttons and a simple vocabulary of speech commands (Goose & Djennane, 2002). Along the same line, Commute UX is a voice-enabled infotainment system used in the car. This system enables drivers to access their music players, respond to messages and search car manuals via voice commands (Tashev, Seltzer, Ju, Wang, & Acero, 2009).

Similarly, VoxBoox automatically translates HTML books into VoiceXML (Jain & Gupta, 2007), which creates pages enhanced with additional control facilities (i.e., voice commands) in order to provide an enhanced browsing experience and additional navigation controls. Voice commands, such as “skip,” “back,” “start,” “end,” “repeat” and “pause,” are available to users. In addition, users can place voice bookmarks (or voice anchors) on various paragraphs and return to them later by saying the name of the voice anchor.

Likewise, Nomadic Radio is a wearable device that delivers information, such as emails, voicemails, news broadcasts and personal calendar events in the form of audio data (Sawhney & Schmandt, 2000). It is designed as a neckset (Neckset, 2015) with two directional speakers and one directional microphone to be used in indoor and outdoor environments. Users can navigate and interact with Nomadic Radio using voice commands (e.g., go to my email, move forward, move back and play audio). They can also use a push-to-talk strategy to activate voice commands while in noisy environments or use a continuous monitoring strategy (i.e., always in a listening mode) when in quiet environments. Nomadic Radio notifies users about incoming information using different scaled auditory cues based on the priority of the information, usage level and user context, which will help reduce annoyance on the part of the user related to constant auditory notifications. Apple's Siri (Apple Siri, 2015) uses Nuance Dragon (Nuance, 2015), which enables people to use voice commands and ask their "personal assistant" to do things for them, such as check the weather, schedule a meeting or set an alarm. Siri allows users to have natural, conversational interactions with their device (Hearst, 2015) by selectively retrieving information and services from the phone or web.

2.4.2. Disadvantages of Voice Interaction

In the previous section, we discussed several interfaces that use voice inputs as their modalities of interaction. Although voice inputs are beneficial in hands- and eyes-free interactions, several disadvantages exist to using the voice to interact. The first problem is that speech is slow due to its sequential and transient nature (Sawhney & Schmandt, 2000). The second problem is that users need to recall the voice commands unlike on-screen buttons (Sawhney & Schmandt, 2000). The third problem is the effect of the environment on the success of the voice command recognition program (Sawhney & Schmandt, 2000). For example, noisy environments can reduce the system's voice recognition success and, eventually, frustrate the user. However, the addition of a noise-canceling microphone tends to resolve this issue. The fourth problem is that users do not feel comfortable talking to themselves (Sawhney & Schmandt, 2000) or a device (Patel et al., 2009) when in social environments. Users also feel that they might lose their privacy if they have to say confidential information, such as passwords, when in public (Sawhney & Schmandt, 2000). The fifth problem is the effect of motion on recognition error rates. Recent research (Price et al., 2006) demonstrated that motion causes higher recognition error rates, but it may be possible to lessen the effects of motion through a system adaptation. The final problem is the difficulty that exists in regard to recovering from system recognition errors (Patel et al., 2009) or errors in speech (Patel et al., 2009; Tang, Wang, Bai, Zhu, & Li, 2013). Some of the abovementioned issues with speech commands will be resolved as technology advances.

2.4.3. A Design Method for Voice Commands

Several studies have introduced the Wizard-of-Oz approach to designing voice commands. This method (Dahlbäck, Jönsson, & Ahrenberg, 1993; Green & Wei-Haas, 1985) means that subjects are told that they are interacting with a computer system, when, in fact, they are not. Instead a human operator, the wizard, mediates the interaction. For example, SUEDE (Klemmer et al., 2000; Sinha, Klemmer & Landay, 2002) is an informal prototyping tool used to map natural language interactions quickly and then test those interactions using the Wizard-of-Oz approach. SUEDE consists of two modes: design and test. The design mode allows designers to map interaction flows and record voices to act as both the computer and user. The test mode converts the dialogue sequences to a browser-based interface for the ‘wizard’ to use while performing the test.

Along the same line, Salber and Coutaz (1993) demonstrated how the Wizard-of-Oz approach could be extended to analyzing the multimodal interfaces. In addition, Fong and Frank (1992) designed a rapid, semi-automatic simulation method to compare pen and voice as interaction modalities. Another study used the Wizard-of-Oz approach to test how users use a system in order to build a multimodal interface, using speech and pen as an input (Vo & Wood, 1996). Similarly, the Wizard-of-Oz approach was found to be beneficial in regard to simulating speech recognition systems and is recommended for similar experiments in the future (Tsimhoni, Smith, & Green, 2004). These studies support the notion that the Wizard-of-Oz approach is a possible method for the rapid design of voice command vocabulary as another interaction modality.

2.4.4. Guidelines for Effective Voice Commands

Researchers have introduced guidelines by which to design the vocabularies of voice commands. One experiment demonstrated that participants made significantly more memorization errors when using speech versus a mouse for command activation (Karl, Pettey, & Shneiderman, 1998). Other studies have focused on improving voice commands in order to enable users’ memorizations and recall of the commands. For example, one study suggested that designers should only use a few short and aurally distinct words. Moreover, speech recognition software could be configured to respond similarly to a lowered tone of voice. This configuration would permit a user to carry on a normal conversation without inadvertently activating a link (Christian et al., 2000).

Another study mentioned that applications using small vocabularies and predefined commands can significantly reduce error rates and improve recognition accuracy (Feng & Sears, 2009). It is important to avoid multiple commands that sound alike, as such choices will lead to errors and confusion. In addition, the dialogue should effectively leverage the user's vocabulary, making the interaction with the system natural. In this way, many vocabulary problems can be reduced, and commands easier to learn, remember and retrieve. Another study suggested that a short command vocabulary remains easier to discern and understand in short-term memory (Bradford, 1995). Hence, these guidelines informed our design of the high-level vocabulary of the voice commands in the mobile setting used to control the aural flows.

2.4.5. Voice Interaction in Driving Context

Figure 3. Ecosystem of devices in the car (Google Images, 2015).

In the driving scenario, the primary task is defined as the actual driving task and is often performed out of habit, grounded in people's prior driving experience. However, secondary tasks (e.g., selecting music from a hand-held or hands-free music player, receiving and accepting a call, entering data into a navigation system) are not part of the natural driving response. As such, these secondary tasks have the capability to divert the driver's attention away from the driving task (Peissner, Doebler, & Metze, 2011). Considering the evolution of modern, in-vehicle technologies (Figure 3), several studies have focused on the impact of distractions due to driver interactions with information systems (Harbluk, & Lalande, 2005; Lee, Caven, Haake, & Brown, 2001; Peissner et al., 2011; Tchankue, Wesson, & Vogts, 2012; Tsimhoni et al., 2004; Yang, Reimer, Mehler, Wong, & McDonald, 2012). "Driver distraction can be defined as the diversion of attention away from activities critical for safe driving toward a competing activity" (Young, Lee, & Regan, 2008, p. 34).

The findings of a 100-car study conducted by Neale, Dingus, Klauer, Sudweeks, and Goodman (2005) shows that "lapses in selective attention either through inattention or distraction, cause many crashes" (Trick & Enns, 2009, p. 64). Therefore, two types of distractions (i.e., cognitive and visual distractions) can occur due to interactions with car systems. Complexity of the interactions plays a role in causing the cognitive distractions, while the interaction modes and nature of the secondary tasks affect the visual distraction.

Strayer, Watson, and Drews (2011) introduced a third type of distraction called the manual distraction, which occurs when "drivers take their hands off the steering wheel to manipulate a device" (p. 31). Figure 4 illustrates three types (i.e., visual, cognitive and manual distractions) and levels (i.e., low, moderate and high) of the distractions. For example, a low level of distraction occurs when a driver listens to the radio while driving. In this situation, a low level of demand occurs on the driver's visual, manual and cognitive resources. An example of a high level of distraction occurs when a driver uses a touchscreen device while driving, which places a high level of demand on the driver's visual, manual and cognitive resources.

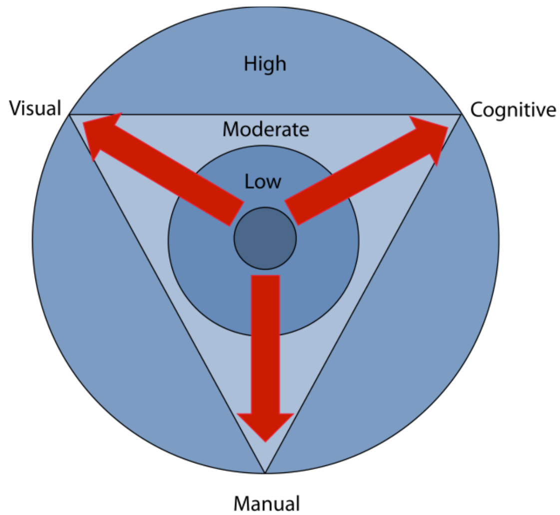


Figure 4. Driver distraction framework (Strayer et al., 2011).

Another factor that plays an important role in distracting drivers is the duration of the secondary tasks with which they are engaged. For example, when the secondary task involves interacting with a visual interface in a car, the length of time that the driver spends interacting with the interface (e.g., five seconds vs. 150 seconds) plays a strong role in how distracted the driver becomes.

A number of strategies have been attempted to address various types of distractions. For example, some studies provide evidence that speech-based interactions can reduce visual (Barón & Green, 2006; Peissner et al., 2011; Ranney, Harbluk, & Noy, 2005) and manual distractions (Harbluk, Eisenman, & Noy, 2002; Peissner et al., 2011; Ranney et al., 2005), which can also improve driving performance (Barón & Green, 2006; Maciej &

Vollrath, 2009; Tsimhoni et al., 2004; Yang et al., 2012) and reduce accidents (Peissner et al., 2011). However, it is important for the speech recognition system to be accurate and easy-to-use in order to enable better and safer interactions in the car (Peissner et al., 2011).

Contradictory research results also exist on the use of audio/voice interaction systems in the car. For example, several studies have mentioned that audio/voice-based interactions introduce significant cognitive overload (Harbluk & Lalande, 2005; Harbluk et al., 2002; Hua & Ng, 2010; Lee et al., 2001; Strayer et al., 2013; Winter, Grost, & Tsimhoni, 2010) when compared to baseline tasks, such as driving only or driving and listening to the radio. However, one study demonstrates that audio/voice-based interactions introduce less cognitive overload when compared to visual/manual-based interactions (Barón & Green, 2006). Another study stated that in-car systems with advanced auditory cues can decrease cognitive overload when compared to visual systems (Gable, Walker, Moses, & Chitloor, 2013).

Two sources of cognitive distractions exist when using voice-based interfaces: (1) listening to audio interfaces and (2) using voice commands to interact with interfaces.

- *Listening to audio interfaces:* Some studies (Harbluk, & Lalande, 2005; Lee et al., 2001) have shown that listening to audio interfaces, paying attention to what is being said and acting upon it consumes cognitive resources. The more cognitive resources are being used, the higher the potential for distraction.
- *Using voice commands to interact with interfaces:* One study mentioned that the voice commands used to interact with audio-based systems can also increase

cognitive load (Winter et al., 2010) due to the need to memorize the commands. With the increasing number of domains in which speech applications are applied, drivers must memorize a number of command words to control traditional speech interfaces. This study showed the commands that are more dialogic in nature can be easily memorized (Winter et al., 2010). In order to address this memorization issue, we introduced several guidelines (e.g., using short and distinct words, small vocabularies, predefined commands) in the previous section.

A body of work also exists in regard to predicting driving performance measurements, while using any user interface in a car simulator (Liu & Salvucci, 2001; Salvucci, 2001; Salvucci, 2002; Salvucci, 2005; Salvucci, 2006; Salvucci, 2013; Salvucci & Taatgen, 2008). Through this body of work, researchers developed a novel simulation software called Distract R, which provides a way for researchers to design an interface, set an interaction with the interface, set the cognitive level in the simulator and run a simulation to receive a few of the predicted measurements for the driving performance (Salvucci, 2009; Salvucci, Zuber, Beregovaia, & Markley, 2005). Three limitations exist in regard to using Distract R. First, it only supports comparative evaluations among identical prototypes with different interaction modalities. Second, it only predicts some of the simulation measurements (e.g., brake response time, longitudinal speed deviation). Third, it only predicts these measurements at the time of the interaction.

Hence, conducting a study in a driving simulation lab is a more promising way because it allows for the prediction of all possible measurements (e.g., peak longitudinal acceleration, lane keeping/displacement, number of lane departures, mean and SD of following distance, number of accidents, mean of glance time) and looks at the user experience as a whole. In addition, the framework for distraction (Strayer et al., 2011) is fundamental in regard to understanding the limitations and benefits of ANFORA as it can impact a number of different distraction dimensions while driving.

2.5. Measuring Distraction Due to Interactions with Mobile Devices

Interacting with mobile devices while walking requires both visual (Bragdon, Nelson, Li, & Hinckley, 2011; Lemmelä, Vetek, Mäkelä, & Trendafilov, 2008) and cognitive attention (Lemmelä et al., 2008), which can be distracting. The complexity of the interactions play a role in causing cognitive distractions (Young, Regan, & Hammer, 2007), while the interaction modes and nature of the secondary tasks affect the visual distractions (Young et al., 2007). Visual distractions are measured by the number and duration of glances towards the mobile device (Metz & Krueger, 2010), while cognitive distraction is measured through cognitive load.

As shown in Table 1, cognitive load can be measured directly using the NASA Task Load Index (NASA-TLX) questionnaire (Hart & Staveland, 1988) or indirectly using the cognitive load theory (CLT) (Sweller, 1988). Sweller introduced different types of cognitive loads, such as Intrinsic Cognitive Load (ICL), Extraneous Cognitive Load (ECL) and Germane Cognitive Load (GCL). ICL (Sweller & Chandler, 1994) is the integral level of difficulty related to the task. ECL (Chandler & Sweller, 1991) is engendered by the approach through which information is presented to the subject as a part of the system design. GCL (Sweller, Van Merrienboer, & Paas, 1998) is the load devoted to the processing, construction and automation of the system operations related to the subject’s prior experiences. Measuring these three types of cognitive loads are important in regard to understanding how interaction modalities while navigating aural flows can effect cognitive efforts. In addition, understanding and measuring different types of distractions that may occur while walking and interacting with mobile devices facilitate a better experimental setup in terms of adopting the right questionnaires and data collection methods.

Table 1. Direct and indirect measurement of cognitive workload.

Direct Measurement	*Indirect Measurement*
NASA-TLX Questionnaire	Cognitive Load Theory (CLT) = Intrinsic Cognitive Load (ICL), Extraneous Cognitive Load (ECL) and Germane Cognitive Load (GCL)

In summary, this literature review has examined the role of user interfaces in dual-task scenarios. The following chapter introduces novel ideas that enable users to listen to content-rich websites, while engaged in another primary activity, such as walking, jogging or driving.

Chapter 3. Introducing the ANFORA Framework

In Chapter 2, we presented different studies in the area of multitasking, while using mobile devices and discussed different types of distractions that may occur. This chapter introduces the ANFORA framework, which contains a set of techniques to be used to remodel existing web information architectures as linear, aural flows.

3.1. ANFORA Framework

ANFORA is a conceptual framework built on top of existing, content-rich, information architectures (Ghahari & Bolchini, 2011). ANFORA framework provides a method to remodel existing websites into a set of aural flows. An aural flow is a concatenated, design-driven sequence of content pages with self-activating links; thus, an aural flow can be listened to with minimal interaction required. ANFORA provides a vocabulary and simple set of design principles by which to define flows of aural content on top of the existing web navigation structures. Such vocabulary is extended from the tradition of hypermedia design models (Bolchini & Paolini, 2006), which aim to describe information and navigation structures at the conceptual level independently of the implementation mechanisms.

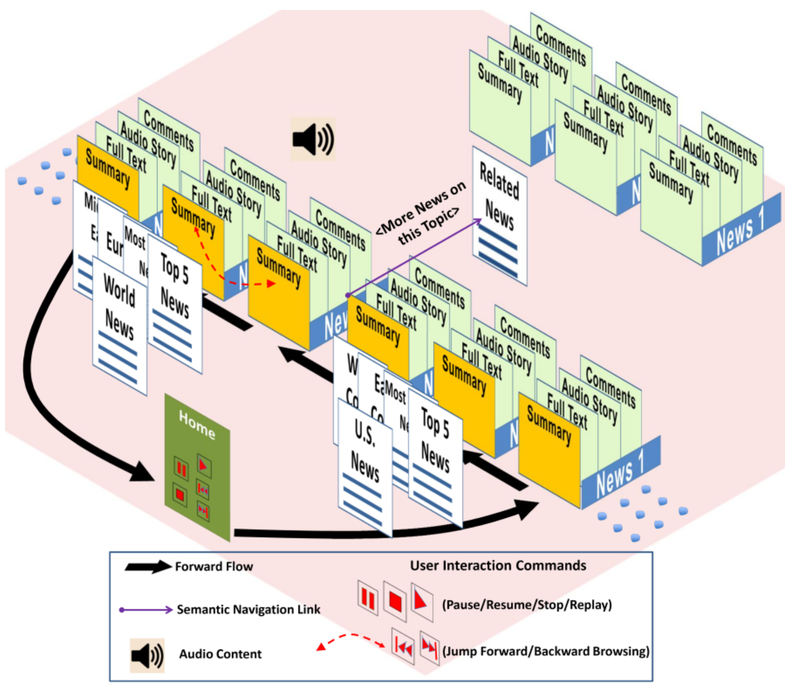
ANFORA could be applied to websites in a number of domains, such as museum, travel, tourism and news sites. By making use of an aural navigation system, ANFORA presents a number of design alternatives that have the potential to enhance quick scanning through content-rich pages when time, contextual and physical constraints are at play. When using ANFORA, users can choose content from the news categories in which they are most interested. Then, users can choose how in-depth they want to delve into those categories based on how much time they have. Finally, ANFORA transforms text on information-rich web pages into a Text-to-Speech (TTS) presentation that users can listen to instead of read. These strategies are an evolution of the guided tour concept, which is a common pattern in media modeling. In a guided tour navigation, users are “led around” by the application (e.g., selecting “next” or “previous” commands), according to the appropriate sequences of content conceived by the designers (Paolini, Garzotto, Bolchini, & Valenti, 1999). Through ANFORA, we investigate new ways by which different types of aural flows can be effectively applied to conventional web information architectures. In an effort to further describe the ANFORA experience, we have identified two main types of aural flows (group flow and full flow) that will be used to describe the interaction patterns outlined below.

3.1.1. Full Flow for Prolonged Aural Experiences

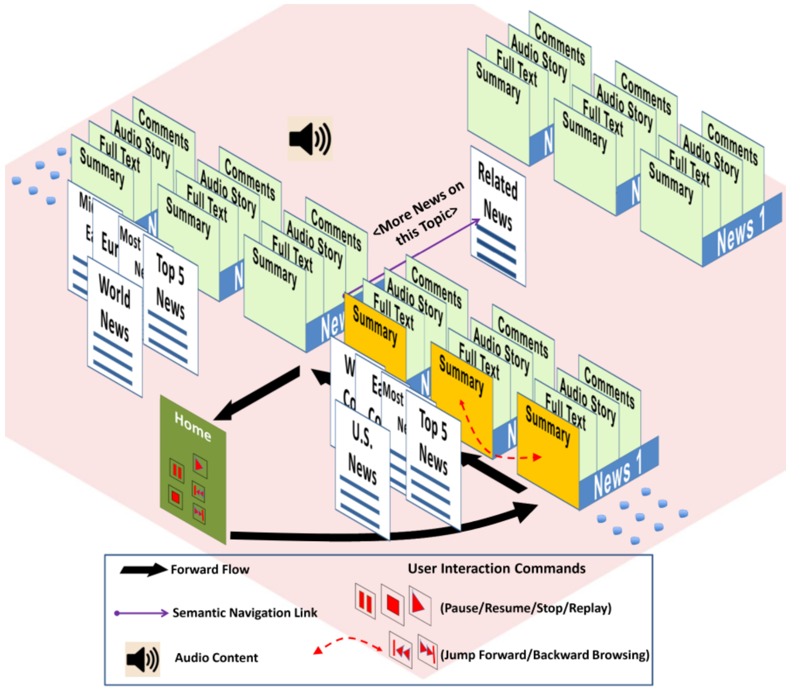
Full flow is the concatenation of some or all of the categories of content (e.g., u.s. news, local news and world news). Full flow allows users to experience all of the main content available (Figure 5a). The length of the flow is determined by the number of items (e.g., news stories) in each group as well as by the number of groups. One advantage of full flow is that it caters to situations in which users have relatively long periods of time to listen to content while on the go. Some of the disadvantages, however, are that users might not perceive changes from one category to another and may have difficulty building mental models of the content structure being played. In addition, some content types can become rather lengthy and, in these situations, the computer-generated voice may cause users to lose interest or become bored.

3.1.2. Short Aural Explorations with Group Flow

Group flow provides users with aural access to a selected category of content (e.g., u.s. news) and plays all of the individual items (e.g., news stories) within the selected group (Figure 5b). The flow stops when all of the items in the category have been read. At that time, the user is led back to the homepage. Obvious advantages of this flow are that users can decide from the outset which category of content they would like to listen to and they have this choice every time a category ends. They can also avoid categories of content in which they are not interested. A favorite group flow can also be bookmarked. A disadvantage, however, is that users need to interact with the interface every time they wish to select a new category.



(a)



(b)

Figure 5. Aural flows in a typical web architecture: (a) Full flow through all categories and (b) Group flow through one category at a time.

3.2. Reifying ANFORA in the News Domain: ANFORA News

In order to reify the ANFORA concept, we have applied it to the news domain because traditional news sites require active navigation and constant visual engagement.

3.2.1. The Four Different News Consumption Experiences of ANFORA News

In order to give ANFORA News' listeners a number of listening options based on how much time they have and how in-depth they want to 'read' into a story, ANFORA News offers several types of listening experiences, each based on the length of the story (Table 2). This design strategy is based on a number of well-defined news consumption experiences: scanning, sampling and comprehensive reading/listening as modeled in *Eyetracking the News*, a widely cited study on print and online news consumption conducted by the Poynter Institute for Media Studies (Quinn, Stark, Edmonds, Moos, & Van Wagener, 2007).

Scanning is defined as the quick perusal of headlines, other display type, hyperlinks and visual elements. Scanning readers rarely read full-text versions of stories, opting instead for a cursory glance at the news through top-level headlines and links (Quinn et al., 2007). Sampling occurs when news consumers go one step further than scanning by also engaging with brief summaries (one to five sentences) of the text-based stories. If summaries aren't available, samplers sometimes read the first one or two paragraphs of a story, but rarely go further (Quinn et al., 2007). Comprehensive reading/listening occurs when news consumers read full stories. Comprehensive readers/listeners tend to engage with news products (i.e., newspapers, magazines and websites) more entirely than scanners and samplers (Quinn et al., 2007). Supplemental reading/listening is a fourth category that has been added to identify an interaction pattern that is more specific to the web information architecture that includes hyperlinks and the ability to comment on web content. This category occurs when news consumers choose to read deeper into a topic for which they have acquired an interest. To do so, they may click on hyperlinks to related stories. They may also choose to comment on a story they have read as a means for interaction with the news source and/or other readers.

Table 2. Aural flow navigation patterns.

<table>
<tr><th rowspan="2">Flow</th><th rowspan="2">Advantages</th><th rowspan="2">Disadvantages</th><th colspan="4">Listening Experiences</th></tr>
<tr><th>Scanning</th><th>Sampling</th><th>Comprehensive Reading/Listening</th><th>Supplemental Reading/Listening</th></tr>
<tr><td>Group</td><td>Decide the Category From the Outset</td><td>Interact Every Time to Select a Different Category</td><td rowspan="2">Headline</td><td rowspan="2">Headline
+
Summary</td><td rowspan="2">Headline
+
Full Story</td><td rowspan="2">Headline
+
Full Story
+
Related Story
+
Readers' Comments</td></tr>
<tr><td>Full</td><td>Less Interaction</td><td>Difficulty in Building a Mental Model</td></tr>
</table>

3.2.2. ANFORA News' User Profiles

ANFORA News targets a broad audience of news consumers characterized largely by individual differences in news consumption habits. As previously mentioned, some news consumers are more likely to scan and sample stories of interest. Others tend to immerse completely, spending more time reading a wider range of news stories from several categories. Still others engage in both types of activities, depending on how much time they have to devote to the news at that moment. Thus, we envision that ANFORA News' users can be broken down into three key categories: light, heavy and combination users. These titles are based on the premise that different users exhibit varying levels of information motivation, technical savvy and expectations regarding the time commitment related to news consumption at a given time.

Light users most often choose to scan headlines or sample story summaries, rather than listen to full stories. They do so because they are motivated by both time constraints and a less intense desire to spend time listening to news. Major news stories – regardless of category – are generally of interest. Examples of such stories include the death of Osama bin Laden, presidential elections or breaking news stories. Light users are also often motivated to engage with only the stories in which they are personally interested. Heavy users are generally more likely to regularly spend more time with news than light users. They are more likely to listen to stories on a variety of topics, regardless of personal interest, and are more willing to listen to full stories than light users. Finally, combination users may exhibit behaviors common of both light and heavy users based on how much time they have and/or how motivated they are at a given time to engage with the news.

These user profiles were used to inform the design of ANFORA News' user experience. By providing users with a number of levels of listening – scan headlines, sample news stories, listen to full stories and supplement with related headlines and/or reader comments – ANFORA News allows them to listen to the news in whatever format fits their current time constraints, interests and desired levels of detail when it comes to story length.

3.3. The ANFORA News Prototype

The ANFORA News design capitalizes on common news consumption habits by allowing users to choose which level of listening (i.e., scanning, sampling, comprehensive listening or supplemental listening) they wish to engage (Figure 6). Thus, we have designed a mobile version of this audio-based news website that looks like an application and implements different aural flow types in one prototype. After users access the website, an introductory page is displayed for few seconds before they are redirected to the home page where they can decide how deeply they want to listen to the news. Users can select “scan headlines,” “sample story summaries” or “listen to full stories.” They can also add “related stories” or “readers’ comments.” Next, users are redirected to a page where they can select the main categories of news; the subcategories are decided based on the main category choices. Once all of these choices are made, the news is automatically read via TTS. Users can also follow along if they wish by looking at the screen (See Appendix A for detailed screenshots of the ANFORA News prototype).

Figure 6. ANFORA lets users choose how much time they want to spend with the application and then creates a custom aural flow of news stories.

ANFORA News is designed to minimize visual and physical interaction with the screen, using self-activating links that concatenate pages in the flow (Figure 6). However, if they want to, users can interact by using either tap button commands and/or touch-based gesture commands. These commands allow users to pause, resume, replay and stop the flow. Commands also allow users to fast forward to go to the next segment of a single news story (e.g., related stories or readers' comments) or rewind to the previous segment of a single news story. Finally, users can skip to the next news story or go back to the previous one at any time by using the "jump forward/backward" commands. Figure 7 shows the gesture commands that correspond with these interaction patterns. Figure 8 shows the appearance of the button commands.

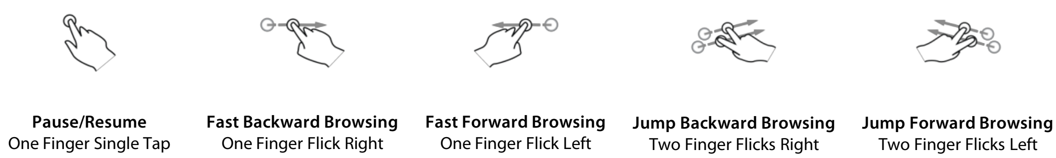


Figure 7. Touch-based gesture commands can be used at any time during the flow experience.

Figure 8. Button commands can be used at any time during the flow experience.

Consider, for example, a scenario in which a user decides to listen to ANFORA News during his 30-minute walk to work (as shown in Figure 9). He chooses to listen to the summaries for the "top 5" and "most recent" stories in the "world news" category as well as the "most recent" story summaries in the "national news" category and "indiana" stories in the "local news" category. Between stories and categories, the user hears sound effects (i.e., earcons) to indicate when a new story or category begins. Earcons are “non-verbal audio messages used in the user-computer interface to provide information to the user about some computer object, operation or interaction” (Blattner, Sumikawa, & Greenberg, 1989, p. 13). In this scenario, the user employs gesture commands to skip to the next story summary or replay a summary.

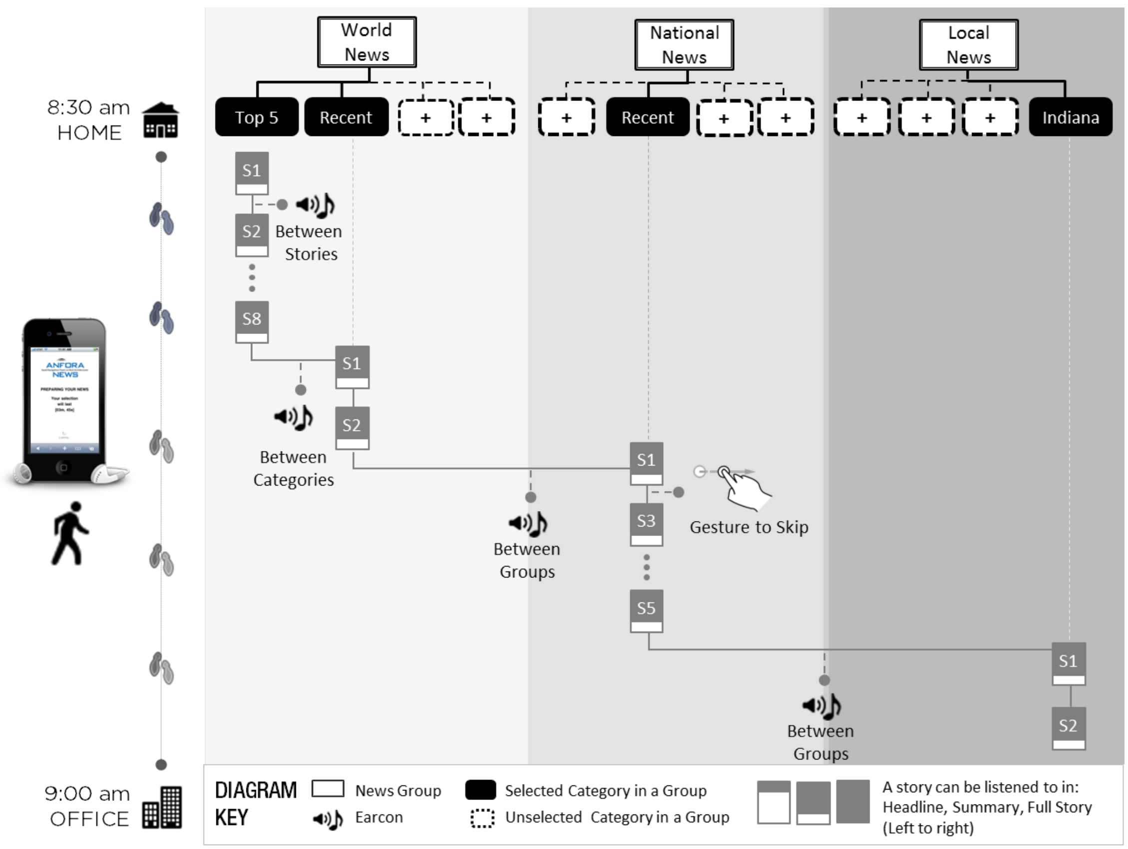


Figure 9. Visualization of an ANFORA News experience scenario.

3.3.1. Content, Styles and Formats

The ANFORA News prototype contains news stories pulled from the NPR news website (www.npr.com). NPR was chosen for its comprehensive coverage of u.s. and world news, as well as its regional focus on several local markets, including the market in which this study was conducted. Some of the stories used for the ANFORA News prototype were downloaded audio files from NPR programs. Others were text-based stories converted to TTS. ANFORA News could allow news organizations to offer a mix of broadcast quality reports along with TTS news stories.

The news stories were divided into three main categories: “local”, “national” or “World.” Stories that would remain interesting to a general audience for several months were chosen so that the prototype wouldn’t have to be updated every day with new stories. Stories were then assigned to four sub-categories within each main news category (e.g., “top 5” stories, “most recent” stories, etc.). The number of stories in each main category varied, just as it would on a news website. Some stories could fall into multiple news categories or sub-categories. ANFORA News stories are tagged in such a way that when such redundancies occur, they appear in only one group/category, namely the first category encountered according to the order of the groups and categories selected by the user.

3.3.2. Design Challenges for the Aural Experience

Blending two distinct modalities, such as a TTS technology and news, is not without its challenges. In fact, a number of characteristics exist that are rather unique to the way news organizations operate and present content that poses notable roadblocks to the implementation of ANFORA News. These challenges are certainly not insurmountable. However, they are worth noting here, along with some of the ways in which the current iteration of ANFORA News responds to them.

A few key issues arose in the early development of ANFORA: time, orientation and TTS voice quality. *Time* refers to the time it takes users to complete a full news listening experience. Of course, different users will intend to spend varying amounts of time with the ANFORA News application, depending on the time of day and how much time they have. Therefore, ANFORA News was designed to accommodate a number of different interaction lengths, from five to 10 minutes up to 45 to 60 minutes. Since ANFORA News was built with a number of time and engagement options, users can quickly become disoriented when engaging with multiple news stories from different news groups (e.g., “local news”, “national news” or “world news”). Thus, a number of strategies for maintaining user orientation were designed.

3.3.2.1. Ensuring User Awareness of Time Commitments

News stories vary in length, depending on the importance of the story and the amount of space and resources available for its coverage. For example, many news organizations repurpose stories originally written for print, a medium that is very space dependent, for online news sites. Significant stories are often written in greater depth and length than stories deemed less newsworthy. This concept is significant for a TTS application because it results in variations in regard to the time that it takes for each story to play. As ANFORA News is designed to be used primarily when news consumers are engaged in other tasks and since news consumption itself has been defined as a “snacking” activity when executed on a mobile device (Meijer, 2007), it is imperative that users are always aware of how much time they are investing in ANFORA News.

ANFORA News employs a few key strategies to address the issue of time. First, each news sub-category (e.g., within the “local news” group, users may choose to listen to "top stories," "most recent" news stories or stories focused on "indiana" or "indianapolis") is labeled with the amount of time it will take to listen to the news sub-category in its entirety. Second, each news story is displayed on the device screen as it is being read and users can scroll through it to see how long it is. Third, each story segment (e.g., summary, full story, related stories and reader comments) is labeled with its length in minutes and seconds. Finally, as each segment plays, a label indicates how much time is left in the article. Together, these strategies ensure that users are always aware of how much time their choices will take and how much longer a particular listening experience will last.

3.3.2.2. Ensuring User Orientation

ANFORA News can provide news headlines, summaries, full stories of varying lengths, reader comments and related summaries in a TTS format. Users may also choose to listen to several stories from a number of different news categories (e.g., "local", "national" and "world"). As ANFORA News transitions from a story in one category to another story in the same category, it is necessary to include clear labeling to ensure that users can quickly assess which category of news they are listening to at any point in time. Finally, as users' attention is often divided between ANFORA News and other tasks (e.g., cooking, walking and jogging), it is easy for users to quickly become disoriented. For this reason, it is important that users can easily reorient themselves.

Two levels of orientation exist in the program as shown in Figure 10. The first level of orientation, *flow-level orientation*, provides users with an indication of how many news stories they have listened to or how many news stories are left to listen to in the flow. For example, a user might listen to the first of 12 news stories across "top local" news, "most recent local" news and "indianapolis" news. This information enables users to plan ahead by providing them with a sense of how long the complete experience will last.

The second level, *group-level orientation*, provides users with an indication of how many news stories exist in each category. For example, a user might listen to the first of five news stories in the "top local" news category and then might listen to the second of four news stories in the "most recent local" news category. In this case, the user would not know how long the complete experience will take and cannot plan ahead.

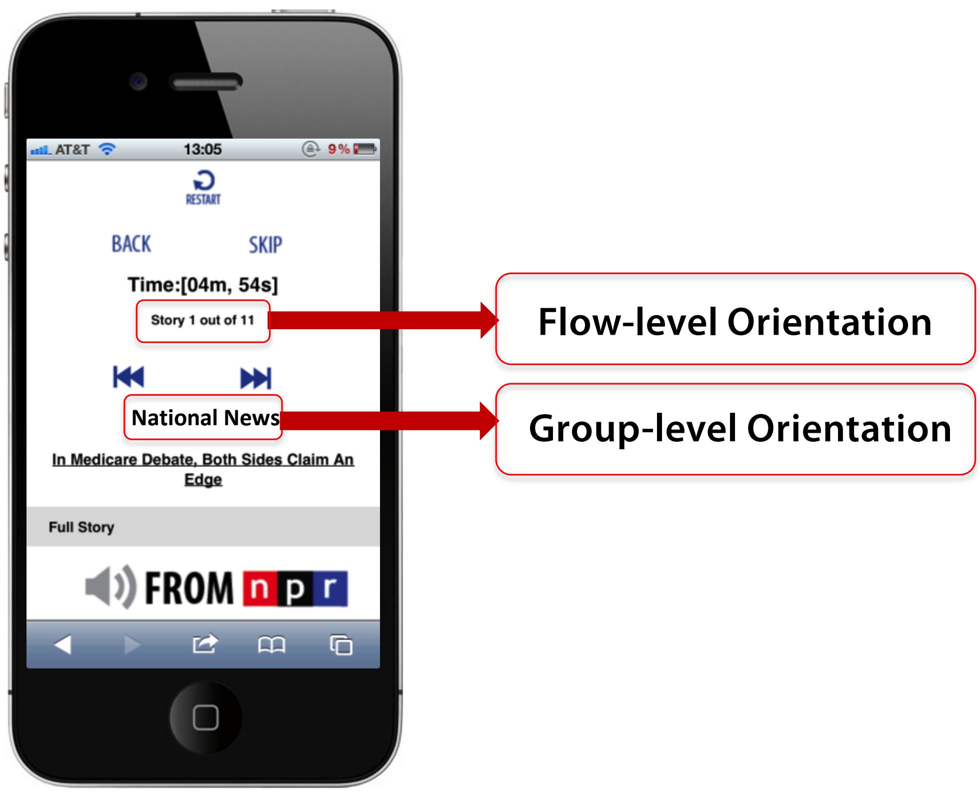


Figure 10. We combine group- and flow-level orientation by allowing users to *see and hear* which category of news they are listening to and *hear* which story they are listening to.

The challenge that exists for designers relates to which of these experiences – the flow- or group-level orientation – to offer. If designers show the orientation at both levels, they will likely overload the users aurally, increasing cognitive load. For example, imagine that you are listening to the first of 12 news stories in a flow made up of more than one category. For example, the first five stories might be “top local” news stories, the next four stories might be under the “most recent local” news category and the remaining three stories might be from the “indianapolis” news category. In a strictly visual interface, such as a common news websites, it is easy to illustrate these categorizations, while still allowing the user to view all 12 stories in a row. These divisions can be distinguished through the use of navigation labels, hierarchical menus and other visual cues. However, these strategies are not available in the aural experience. Users would have to listen to a large amount of orientation information (e.g., “reading story one of 12 total stories; story one of five in the “top local” news category”), which would disrupt the flow experience. Listening to large chunks of sequential information can be improved by having sounds that mark the breaks or movements between one story and the next and one category and the next.

In order to achieve this goal (Figure 10), we decided that it is more important for the user to know how many stories (12) make up the complete listening experience. As such, we opted for flow-level orientation, which provides the user with an overall sense of flow. At any point during the flow, however, a user could glance at the screen to see a label explaining to which category of news (e.g., “top local”, “most recent” and “indianapolis”) the story he is listening to belongs. This strategy enables the user to regain a sense of group-level orientation. Thus, although the primary function of ANFORA News is to provide a hand- and eyes-free TTS news experience, a visual interface exists to ensure that users clearly understand their time commitments and orientation at any given moment.

In this chapter, we have presented the ANFORA framework and its application to the news domain. The next chapter will report on an exploratory study conducted on ANFORA News that investigates how well aural flows support an eyes-free browsing experience that takes place while walking and listening to web content.

Chapter 4. Preliminary Evaluation of the ANFORA Framework

This chapter will present an evaluation of the ANFORA framework via an exploratory study and its results. The exploratory study had four goals:

1. To explore how well the initial ANFORA News design supported an eyes-free browsing experience;
2. To learn how well ANFORA can coexist with the physical and cognitive tasks inherent to the mobile experience (e.g., walking and paying attention to surroundings);
3. To explore the ecological validity of the ANFORA concept by testing the usability, enjoyment and information value of the aural flows and the semi-aural experience; and
4. To explore the strengths and weaknesses of ANFORA in regard to the user experience of listening to news.

4.1. Study Design

4.1.1. Physical Setup

The evaluation study was conducted in a controlled environment that consisted of a predetermined path that users had to walk while listening to ANFORA News. The path was established through the hallways in a highly populated building and included six sharp turns to simulate a real-world scenario in which people are required to avoid other people and objects. The users' interactions with ANFORA News was video recorded to capture their walking behavior along the path. The participants were encouraged to walk on the path as naturally as possible while listening to ANFORA News.

4.1.2. Participants

Twenty participants (10 male and 10 female; all graduate students at a large Midwestern university) were recruited for this study. All of the participants spoke English fluently and none of the participants had hearing or walking impairments. The participants were all daily users of a touchscreen mobile phone and regular news consumers. They received a $15 Amazon gift certificate for approximately 60 minutes of participation.

4.1.3. Procedures and Tasks

Five tasks were identified in order to ensure that the participants would engage in all of the interaction patterns available in the ANFORA News prototype. As each task yielded a listening experience that ranged from three to 15 minutes, the tasks were divided into two groups. This division of participants ensured that each research session would last no more than one hour in order to minimize the participants' fatigue. Group one contained three tasks and group two contained two tasks. The participants were assigned to one of the two task groups. Although the tasks were slightly different, depending on which type of listening interaction pattern (i.e., scanning, sampling or comprehensive listening) the users were asked to perform, the nature of the tasks was the same. Thus, although the users made different initial selections, their general experiences were the same. Once a listening session began, the only difference present was in the length and subject matter of the stories. Thus, we can consider the two groups to be a single sample consisting of 20 participants because the aspects of the interactions and listening experiences central to the study were the same.

Prior to commencing the study, the participants were given a brief explanation of ANFORA News. The researchers gave each participant a short demo of the interface and allowed each person to practice using it to get a feel for how ANFORA works. The first task group was asked to complete a series of three tasks focused on the *scanning, sampling* and *comprehensive listening* interaction patterns. In the comprehensive listening task, the participants also engaged in the *supplementing* interaction pattern by adding reader comments and related stories to the initial selection. The second task group was asked to complete two tasks focused on the *sampling* and *comprehensive listening* interaction patterns. The participants were told that they were not required to interact with the screen after creating the initial news playlist. However, they were also told that, if they wanted, they could use either/both control buttons on the screen and gesture commands to interact with the screen. The length of each task depended upon the interaction pattern. The tasks ranged in length from four minutes (scan headlines) to 15 minutes (listen to full stories).

The researchers accompanied the participants during the walking aural experience and video recorded the sessions. Three main types of data were recorded during this portion of the study. First, the researchers recorded whether the participants completed each task with or without assistance and whether they chose to stop the aural flow before the end of the flow. This data is referred to as the *aural flow completion rate*. Second, the researchers recorded the number of errors that occurred during each task and then categorized those errors according to their main causes. This data is referred to as *the* *occurrence of error percentage during the total number of listening sessions*. Third, the researchers recorded the amount of time that users visually or physically engaged with the screen. These data are referred to as *the percentage of time spent engaged with the screen*.

After completing the tasks, the participants completed two brief five-point Likert item surveys about their experience based on which group they were in. Both task groups engaged in a “sample story summaries” task, while task group one also engaged in “scan headlines” and “listen to full stories” tasks. Thus, task group one responded to 16 questions, while task group two responded to 14 questions. After completing the survey, the participants engaged in a 15-minute interview with the researchers (See Appendix B for the full list of tasks as well as the surveys and interview questions). In the interview, participants were asked to report whether they became distracted by their surroundings and, if so, whether the distractions prohibited them from paying attention to the news. Likewise, they were asked to report whether listening to the news or any interaction with ANFORA News interfered with their abilities to effectively navigate their surroundings.

4.2. Analysis

For the task performance data analysis, the aural flow completion rate, rate of occurrence in regard to different types of errors during the tasks and the amount of time that the users engaged with the screen during the tasks were recorded. These measures helped form an understanding of how easy or difficult it was for the users to use the ANFORA News interface while walking and to what extent they engaged in an eyes-free aural news consumption experience. The surveys were used to measure ease-of-use, willingness to use ANFORA News again, quality of TTS, perceptions of orientation and opinions about the value of the specific levels of reading (i.e., scanning, sampling, listening in full and supplementing) in which they engaged. The results for the surveys were averaged across participants across tasks. For the qualitative analysis of the post-task interviews, recurrent themes were extracted and comments were grouped by theme. The emerging issues highlighted user satisfaction with the ANFORA News listening experience, reflection on levels of distraction encountered during the listening experience, and positive and negative opinions about the interface.

4.3. Results

4.3.1. Task Performance Data Analysis

4.3.1.1. Aural Flow Completion Rate

Of the participants, 90% (18) completed the flow from start to finish with or without assistance (Figure 11). Only 10% of the participants (2) stopped the flow early (one during Task 4 and one during Task 5). The tasks varied in length, depending on the reading level (i.e., scan headlines, sample story summaries, listen to full stories) and number of stories in the particular selection. In cases in which the session was long (sometimes as long as 60 minutes) due to a large number of long stories, the participants were asked to stop after 15 minutes in order to reduce fatigue. The aural flow completion rate was defined by whether a user stopped the task before all of the stories in a selection were read or before the 15 minutes had been completed (See Appendix C for the tabulated data).

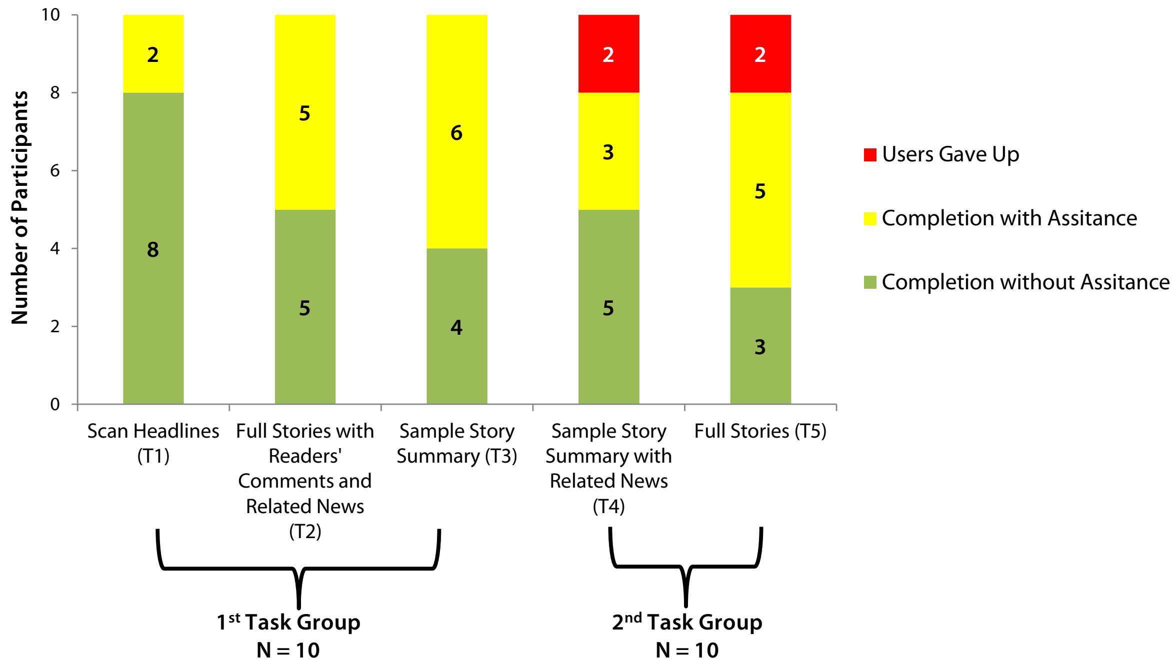


Figure 11. Aural flow completion rate across all five tasks.

Of the participants, 80% (8) completed Task 1 without assistance. This percentage is greater than the percentage of the participants who completed the other tasks. As Task 1 encompasses scanning headlines, the task is shorter than the other tasks. The longer the task, the more likely the user needed assistance, mainly due to technical errors (explained in the next section) and not the design or orientation. In addition, users were more likely to become disinterested during longer browsing tasks, such as listening to full stories.

4.3.1.2. Percentage of Error Occurrences During Total Number of Listening Sessions

Figure 12 shows the different types of errors that occurred during the total number of task sessions (n=50). These errors often caused the participants to engage with the screen either by looking at it or physically interacting with it through button or gesture commands. Overall, the reasons why the users engaged with the screen can be summarized as "confused by long pauses," "encountered technical problems," "poor recall of the gesture commands," "misunderstood button labeling" and "misunderstood TTS" (See Appendix C for the tabulated data).

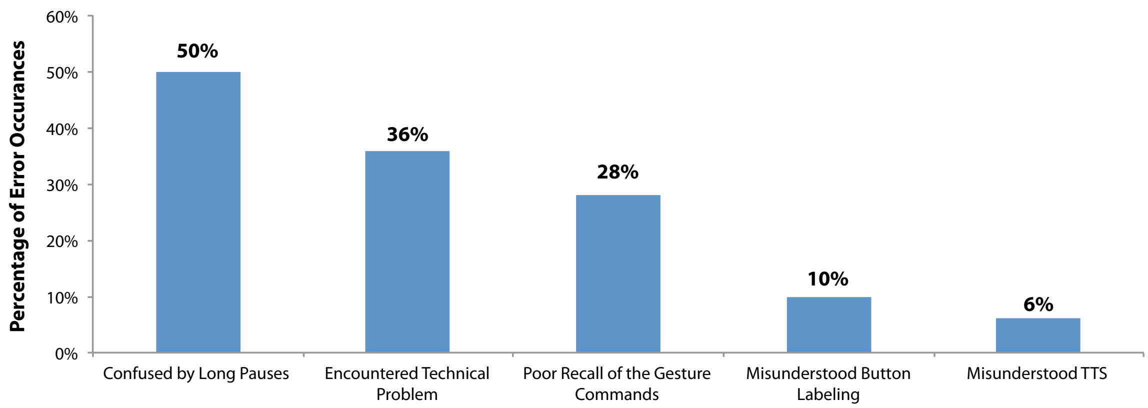


Figure 12. Percentage of error occurrences during total number of listening sessions (n=50).

Confused by long pauses: Confusion caused by long pauses between stories occurred in 50% of the total sessions and was the most frequent type of error that the participants encountered. Although the pauses between the stories were designed to be about one-half second, a slow network connection sometimes caused them to be as much as three seconds. These long pauses often caused the participants to look at the screen because they thought something was wrong.

Encountered technical problem: Technical problems accounted for 36% of the errors experienced. Sometimes, the application timed out due to network malfunctions. This error often caused the users to look at the screen in an attempt to determine why the flow had suddenly stopped.

Poor recall of the gesture commands: Twenty-eight percent of the error occurrences were due to a poor recall of the gesture commands. The participants also had trouble remembering the different gesture commands. Therefore, they sometimes incorrectly used one- or two-finger swipe commands.

Misunderstood button labeling: Ten percent of the errors occurred when the users didn't fully understand the functions of particular button commands. Although they understood that “next” and “back” would take them to the next or previous stories, they did not always know what the double arrow/line button (i.e., jump forward/jump backward) meant.

Misunderstood TTS: Across all of the error occurrences, misunderstood TTS accounted for 6% of the errors encountered. The participants often looked at the screen when they had trouble understanding the TTS. As the written stories appear on the screen as they are being read, users have the opportunity to clarify what they are hearing by visually following along with what they see on the screen.

4.3.1.3. Percentage of Time Engaged in the Aural Flow

Overall, the users spent more than two-thirds of the time on task engaged in the aural flow. The amount of time spent listening to the news without engaging with the screen increased from Task 1 to Task 3 for the first group of participants and from Task 4 to Task 5 for the second group of participants (Figure 13) (See Appendix C for the tabulated data).

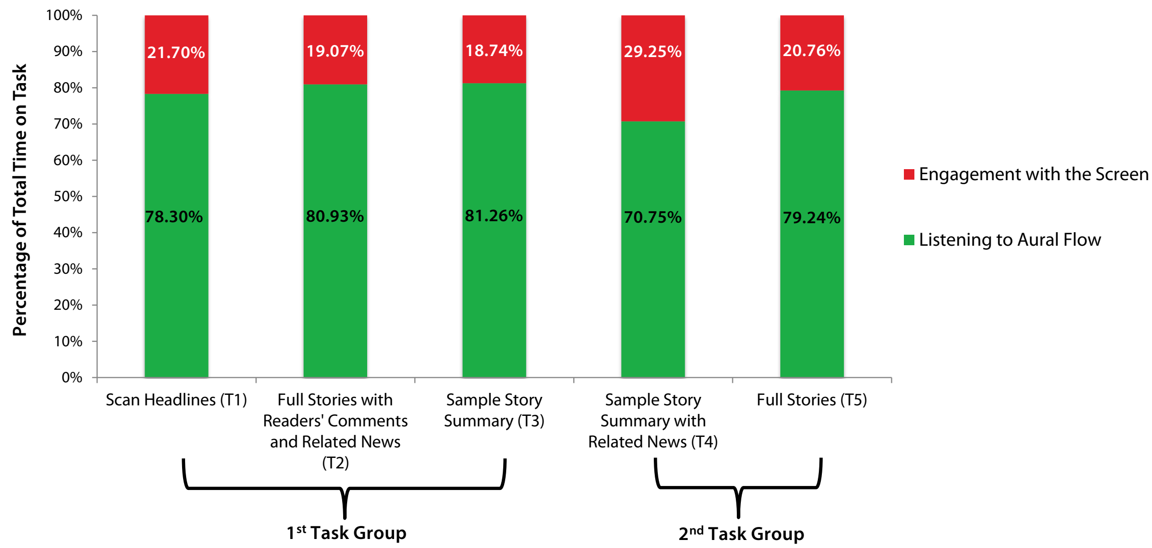


Figure 13. Users spent two-thirds of the task time listening to the aural flows without engaging with the screen.

4.3.2. Post-task Survey

The overall response to ANFORA News was positive. On average, the users found ANFORA News to be easy-to-use (average response: 4/5), enjoyable (average response: 3.95/5) and easy-to-navigate (average response: 3.7/5). Most users also reported that they would use ANFORA News again (average response: 3.85/5) and that the TTS voice was easy-to-understand (3.9/5). Figure 14 shows the average responses across all 20 participants to each of the 14 questions. As previously noted, although the two groups completed slightly different tasks, the nature of the tasks was the same, making the overall user experience the same among all 20 participants (See Appendix C for the tabulated data).

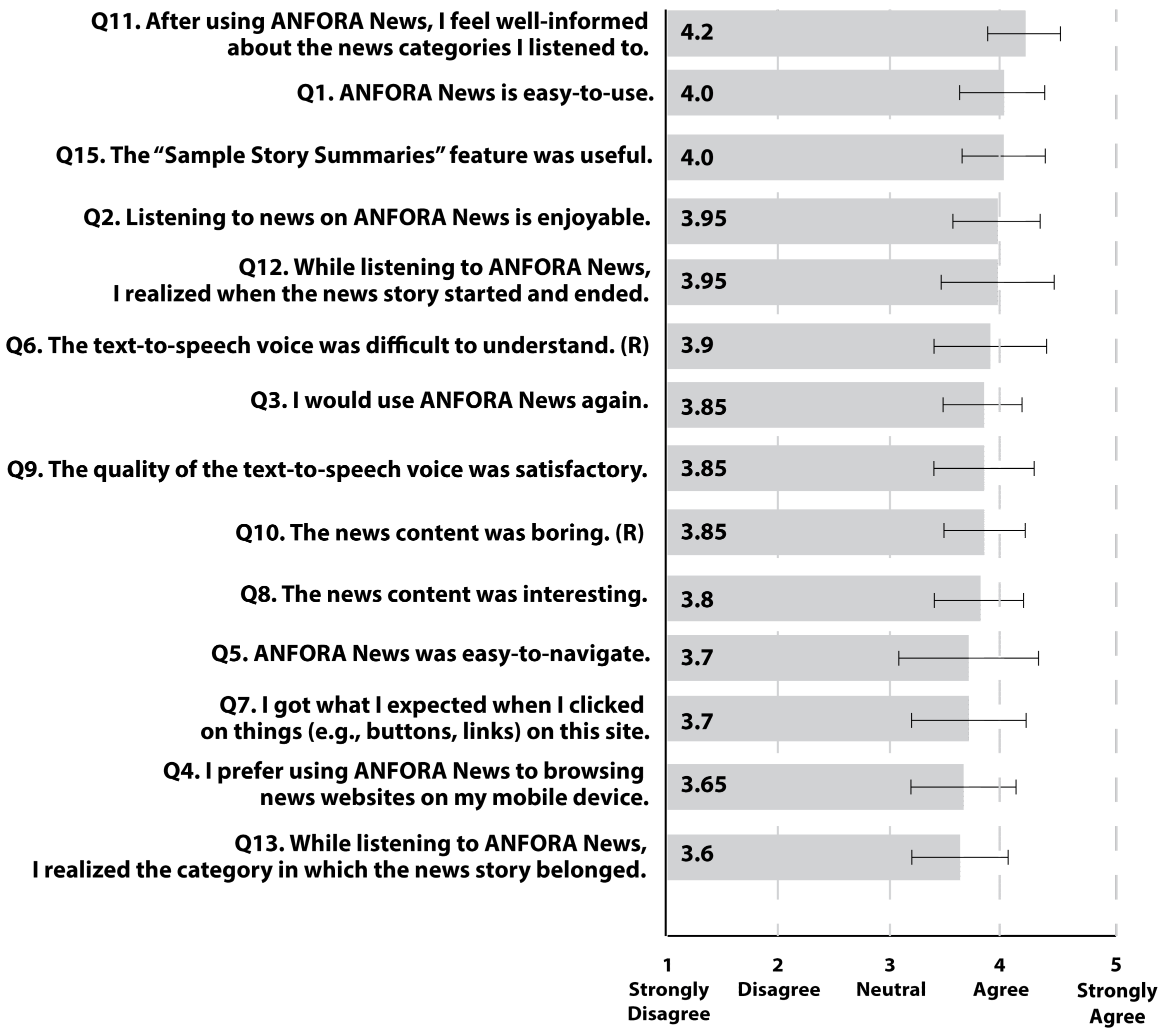


Figure 14. Average responses to the survey questions (N = 20).

It is worth noting that for six of the survey items, the deviation from the mean dropped below three. Responses to items five, six, 11, 12, 13 and 14 were more widely distributed. Items five (ANFORA News was easy-to-navigate) and six (The TTS voice was difficult to understand) can be directly correlated with the results of the error occurrences summarized above.

Finally, in order to determine the relationships between the questions, we examined the factorability of the 14 survey questions. Nine of the 14 questions correlated with each other, suggesting reasonable factorability (Table 3). The Keiser-Meyer-Olkin test of sampling adequacy was .54 and Bartlett's test of sphericity was significant, χ^2 (36) = 51.80, $p < 0.05$. The Cronbach's Alpha is .751.

Table 3. Extracted factors from post-task survey questions.

Factor	*Questions*
Factor 1. Enjoyability of ANFORA News	**Q11.** After using ANFORA News, I feel well-informed about the news categories I listened to. **Q3.** I would use ANFORA News again. **Q15.** The "sample story summaries" feature was useful. **Q2.** Listening to news on ANFORA News is enjoyable.
Factor 2. Content of ANFORA News	**Q10.** The news content was boring. (R) **Q8.** The news content was interesting. **Q9.** The quality of the TTS voice was satisfactory.
Factor 3. Navigation Structure and Orientation	**Q13.** While listening to ANFORA News, I realized the category in which the news story belonged. **Q12.** While listening to ANFORA News, I realized when the news story started and ended.

Table 4. Questions loading for each factor.

Rotated Component Matrix[a]			
Questions	*Component*		
	1	*2*	*3*
Q11. After using ANFORA News, I feel well-informed about the news categories I listened to.	.783		
Q3. I would use ANFORA News again.	.769		
Q15. The “sample story summaries” feature was useful.	.758		
Q2. Listening to news on ANFORA News is enjoyable.	.637		
Q10. The news content was boring. (R)		.879	
Q8. The news content was interesting.		.799	
Q9. The quality of the TTS voice was satisfactory.		.768	
Q13. While listening to ANFORA News, I realized the category in which the news story belonged.			.900
Q12. While listening to ANFORA News, I realized when the news story started and ended.			.827
Extraction Method: Principal Component Analysis Rotation Method: Varimax with Kaiser Normalization			
a. Rotation converged in four iterations.			

Three factors were extracted (see Table 4 for the questions loading on each factor). The first factor was the enjoyability of ANFORA News, which explained 34.71% of the total variation. The second factor was the content of ANFORA News, which explained 19.59% of the total variation. The third factor was the navigation and structure (i.e., orientation) of ANFORA News, which explained 14.21% of the total variation.

4.3.3. Post-task Interviews

These semi-structured interviews included 15 questions and focused on three main themes: *user satisfaction with the ANFORA News listening experience, reflection on the levels of distraction encountered during the listening experience* and *positive and negative opinions about the interface.* We will discuss these three themes in the remainder of this section.

4.3.3.1. User Satisfaction with the ANFORA News Listening Experience

The interviews confirmed the users' general satisfaction with ANFORA news. In addition, all of the participants reported that it was easy-to-use and convenient. The interviews allowed the participants to elaborate on their survey responses and they cited *ease-of-use* and *convenience* as the most appealing aspects of the application. In particular, six users stated that they liked that they only had to listen to the categories of news in which they were interested. One likened the experience to reading only one section of a newspaper.

Nineteen users reported that they would use ANFORA News if it were available today and noted that there were other contexts (besides walking) in which they would find it useful, such as while cooking or driving. One user said: "It's quick and easy-to-use and you spend a lot of your time in motion, in commute to somewhere; you don't have a lot of time to sit still and focus on a reading, or news articles or news online." Five users also noted that ANFORA is a good alternative to other news consumption activities, such as listening to the radio or podcasts or surfing the web.

For the most part, the users were satisfied with the quality of the TTS. Three even suggested that it should be faster in order to keep their attention. However, one user reported that the TTS was not pleasing, while two of the users stated that, at first, the TTS voice was confusing and hard to understand.

4.3.3.2. Reflection on the Levels of Distraction Encountered During the Listening Experience

When the participants were asked about whether they became distracted during their listening sessions, it became clear that distraction was a relative term in regard to aurally navigating the web while engaged in another task. In fact, distraction seemed to be measured on two ends of a continuum. At one end, the participants sometimes stopped carefully listening to the news in order to adequately monitor their surroundings. When they did, they often failed to fully process some of the content. At the other end, the participants were sometimes so engrossed in the story that they lost a sense of their surroundings. In these cases, continuing to listen may be dangerous. One participant noted that situational awareness fluctuated between the news story he was listening to and his surroundings.

4.3.3.3. Positive and Negative Opinions about the Interface

The participants were almost evenly split when it came to preferences regarding button or gesture commands for interacting with the screen. Ten users preferred gesture commands, while eight preferred button commands. Among those users who preferred gesture commands, the primary concern was efficiency. Several of the users noted that gesture commands allow them to quickly skip to the next story without having to look at the screen. On the other hand, those users who liked the button commands better noted that the buttons were more intuitive. Several of the participants said that the button commands made more sense because they were easier to understand than the gesture commands.

Although the users were generally happy with the ANFORA News experience, a few key recommendations surfaced repeatedly. Half of the users noted that they wanted even more choices in regard to selecting the news in which they were most interested. Ten of the participants specifically recommended that we give a list of headlines in each category so that users can choose individual stories for the playlist. Likewise, nine of the users reported that they would like more content options (e.g., sports, business, technology and entertainment).

Only two complaints consistently surfaced about the ANFORA News interface. One complaint was related to the button and gesture command functionality, while the other complaint was related to the length of the pauses between the news stories. Five of the users said that the button commands were confusing and eight users said that the gesture commands were confusing. Seven of the users said that the pauses between the stories were too long, while five noted that the long pauses between the stories often caused the participants to look at the screen because they thought something was wrong.

4.4. Discussion and Future Work

Through this study, we unearthed initial evidence suggesting that aural flows represent a promising paradigm through which to support eyes-free browsing of mobile devices while on the go. However, we acknowledge a number of limitations that still need to be addressed. For example, a few of the participants required initial assistance to make sense of the mechanics of ANFORA News. In addition, as this study was preliminary, the number of participants (n=20) is relatively small, making it difficult to generalize the results. In spite of these limitations, this study provided some key insights into the benefits of using aural flows to minimize the amount of visual attention necessary for users who wish to browse content-rich websites while on the go. Specifically, this study helped us address our research aims in the following ways:

Regarding the first aim – to explore how well the initial ANFORA News design helps support an eyes-free browsing experience – the fact that the participants spent more than two-thirds of the time on task engaged in aural flows suggests that ANFORA News achieves what it was designed to do (i.e., minimize visual interaction with the mobile device screen). In addition, with an aural flow completion rate of 90%, it is clear that, for these participants, ANFORA News was easy-to-use while on the go.

These results also lead us to believe that promise exists regarding the basis for the second aim, which was to explore how well ANFORA can coexist with the physical and cognitive tasks inherent to the mobile experience. As ANFORA News minimizes the amount of time users must engage with the screen during a rich news consumption experience, users are better able to monitor their surroundings while walking or engaged in other primary tasks. Unlike the experience of browsing news websites on a mobile device, ANFORA News promotes consumption of large amounts of information by listening to rather than looking at content.

ANFORA News also differs from other methods of listening to news, such as radio broadcasts and news podcasts, as shown in Figure 15. These differences are based on a few key principles, including flexibility of access; broader content selection at a high level of abstraction; a multimodal experience, which provides different output and input modalities; and various levels of reading/listening (e.g., scan headlines, sample story summaries and listen to full stories). A radio news broadcast, on the other hand, is synchronous in that users tune in to a complete newscast edited linearly by a producer for a predetermined time slot and a mass audience. Thus, listening to a particular program that contains multiple news stories requires that users do so at a predetermined time for stories presented in a pre-edited format. The news podcast provides a more asynchronous experience by allowing users to download programs and listen to them wherever and whenever they want. However, these programs are still edited by producers with a mass audience in mind. Thus, neither the radio newscast nor podcast can take into consideration any single individual's time constraints and/or personal interests. ANFORA News, however, lets users decide the length of time they will spend with the news, how in-depth they will delve into individual stories (e.g., scan headlines, sample story summaries or listen to full stories), what categories of news they will listen to.

Admittedly, the cost for this flexibility lies in the fact that a user must initially take the time to select the category of news in which he is interested as well as the amount of time he wants to spend listening to the news. However, by spending just a few minutes making initial selections, users can create an automated playlist, avoiding the need to visually engage with a complex news website to browse and read stories of interest one at a time. Most users spent about one minute making initial selections. Once these selections have been made, subsequent visual interactions with the screen are minimal, as users spend the rest of the time listening to the stories they selected. On the other hand, visual interaction with a mobile device is exponentially higher when a user must visually browse a news website and then read stories while on the go. Thus, the cost of initial interaction is mitigated by the fact that all subsequent interactions are eyes-free.

Finally, like a podcast, ANFORA News offers an asynchronous experience by allowing users to listen whenever they want to a concatenated linear broadcast entirely based on their individual choices. In this sense, users become producers/editors by creating their own, personalized news listening experiences.

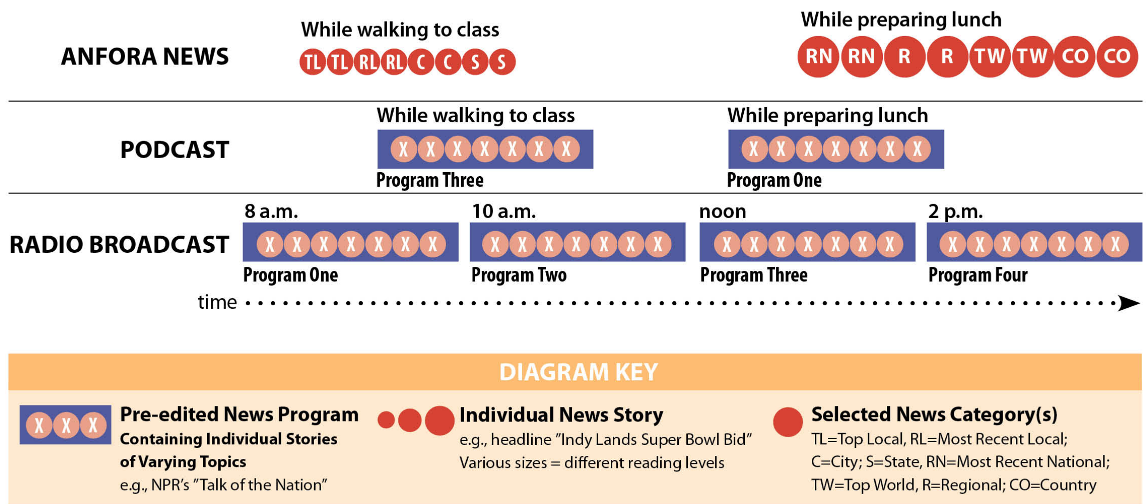


Figure 15. Comparing ANFORA News to podcasts and radio broadcasts. Aural flows provide different reading levels and flexible access by content categories.

The fact that most users found ANFORA News to be easy-to-use and preferred it to browsing news websites on their mobile devices lends additional promise to the third aim: To explore the ecological validity of the ANFORA concept by testing usability, enjoyment and information value of the aural flows and semi-aural experiences. This positive response was encouraging and even the more critical users provided great feedback as to how to improve ANFORA News for the future.

This feedback helped address the forth research aim: To explore the strengths and weaknesses of ANFORA in regard to the user experience of listening to news. The results of both the post-task survey and semi-formal interviews yielded a few narrowly-focused recommendations for improvement. For example, users preferered to have more categories of content (e.g., sports, business and entertainment) and a list of headlines in each section from which they could choose for their master playlists. We also learned that we need to redesign the button and gesture commands to make them more intuitive and utilize shorter pauses between stories. Thus, Chapter 5 will focus on improving the modes of interaction through the addition of vocal commands for infrequent interaction as a means for navigating the flow. The results from a 2005 study on the use of vocal commands showed that participants evaluated speech modality as more satisfying, entertaining and natural to use than using the mobile keypad to interact with the mobile device (Lee & Lai, 2005). As our participants were not fully satisfied with the gesture and control commands, we will implement vocal commands to determine whether this control modality is preferred. We are currently exploring a logical vocabulary for a vocal library as well as planning additional user studies to inform that process.

An important evolution of ANFORA is the seamless and automatic extraction of fresh content from existing websites. For example, in Chapter 5, we present a software engine that connects to the NPR Application Program Interfaces (API) in order to automatically extract daily news to be used to populate the ANFORA News database. This evolution would enable people to use ANFORA News as a stand-alone service.

4.4.1. Validity of the Study

4.4.1.1. Internal Validity

Several strategies were used to maximize internal validity. First, consistent training (i.e., a demo of the interface and practice using the interface) was conducted with the participants before the experiment commenced so that the participants could reach a common threshold of experience with ANFORA News. Second, in order to reduce fatigue, the tasks were divided into two groups and no participant walked for more than 30 minutes. Based on our observations, it was clear that, although the walking tasks were potentially tiresome, the users were not overworked during these tasks. Third, only those users who reported that they are regular news consumers were chosen to participate in the study. This decision was important because those users who have little to no interest in news would likely not find ANFORA News to be relevant to their lives. At a minimum, the participants must have had a general interest in news and a propensity to regularly browse news websites for their feedback about ANFORA News to be useful. Fourth, the survey and interview questionnaires were brief and provided the information needed to accomplish the research aims.

4.4.1.2. External Validity

As previously noted, in order to maximize internal validity, two groups were established. We acknowledge, however, that this decision poses a threat to external validity (or the generalizability of these results) because the sample size is low. However, given that the nature of the tasks is the same for both groups (as explained in section 4.1.3), we can view the total sample size as 20, which is a suitable sample size given the preliminary nature of this study.

As a further indication of the ongoing work on ANFORA News, we (Bolchini & Ghahari, 2013) filed a U.S. non-provisional patent application (No.: 14/024,612 on September 11, 2013) titled “Aural navigation of information rich visual interfaces (Appendix D).” It is our hope that, after additional research and revision, ANFORA News will be ready for public use.

Additional limitations include that the study was conducted in the hallways of a busy academic building, not on a city street. This decision was due to inclement weather and a desire to avoid fatigue and discomfort on the part of the participants. In addition, an experimenter effect may have existed on the users' general opinions about ANFORA News, in that they may have been more inclined to respond favorably in order to please the researchers who conducted the experiment.

4.5. Conclusion

Through this preliminary exploratory study, we learned that aural flows can support eyes-free browsing. Although the participants needed some support to initially make sense of the new/novel framework, they were able to quickly grasp the ANFORA News concept and begin listening to news stories while walking with minimal interaction with the screen. The aural flows allowed them to engage with the web-based news content without having to visually browse the screen. Likewise, the participants reported that they generally enjoyed the experience and found the ANFORA News concept to be easy-to-use. Granted, this study was not a comparative study with a controlled condition. However, when the participants were asked to consider ANFORA News in relation to its alternative (i.e., visually navigating news websites while on the go), they reported that they believed ANFORA News would be safer and easier-to-use. These results provide initial evidence that aural flows support eyes-free browsing and can, therefore, mitigate the physical and cognitive tasks inherent to the mobile experience.

Admittedly, ANFORA News needs additional improvement and development. Thus, Chapter 5 will include enhanced prototypes that address the less intuitive aspects of the existing ANFORA News design. Specifically, we will improve upon the selection and navigation controls and introduce voice commands in order to further minimize the amount of visual interaction required of the users. Chapter 5 will also present the findings from a controlled study used to examine the time taken to visually interact with the device, users' cognitive effort and usability of the button- versus voice-controlled aural flows in the context of walking.

Chapter 5. *Linkless* ANFORA and Evaluation

As we have seen in Chapter 4, touch and gesture still force users to have a visual interaction with aural flows. In this chapter, we introduce voice as another modality of interaction to control and navigate aural flows. We also compare voice- and button-controlled aural flows and examine the potential of voice commands to reduce visual interactions with the device.

5.1. Linkless Navigation Over Aural Flows

The ability to control aural flows using voice commands unleashes a ‘linkless’ interaction paradigm, in which users need not select interface link labels on specific pages and, instead, can activate a limited set of dialogic commands at any time.

5.1.1. Design Methodology

In order to manifest the concept of linkless navigation, we first established full flow as the default setting for the user experience. Full flow enables users to listen to the summaries and full versions of each news story (Figure 16). Full flow also allows users to skip a story or go back and re-listen to a story. In addition, users have the option to listen to related news stories for any given story.

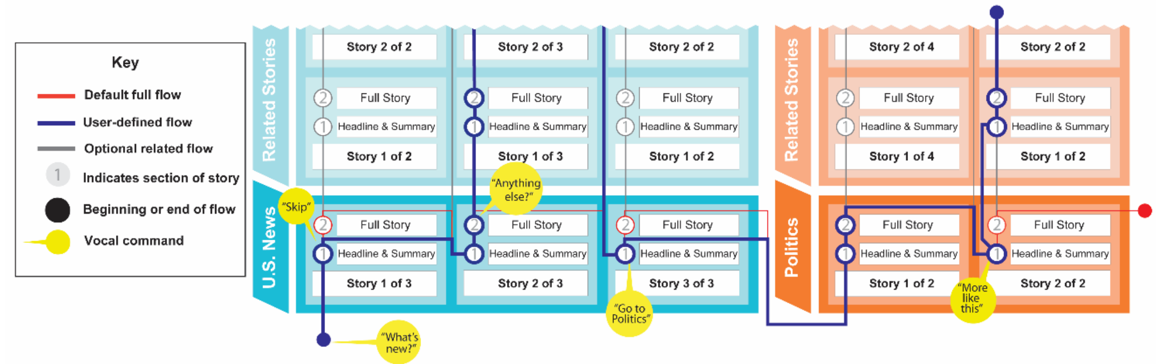


Figure 16. Semi-aural, linkless navigation strategy on ANFORA News: Architecture of aural flow types augmented by voice commands. Patent Pending (Bolchini & Ghahari, 2013).

Second, we defined the aural 'navigation vocabulary' to be used when moving within complex information architectures and interacting with aural flows (Figure 16). This small and simple vocabulary of commands was inspired by common primitives identified in conceptual navigation models (Bolchini & Paolini, 2006; Bradford, 1995; Feng & Sears, 2009; Garzotto, Paolini, & Schwabe, 1993). An aural navigation vocabulary was developed by matching new aural commands with each of the possible navigation strategies for the website. For example, a user could navigate from one news story to the next by saying "next." The design process for developing the final set of commands involved a team of seven designers who explored the commands and simulated the user experience through two rounds of Wizard-of-Oz approach. During these two rounds of Wizard-of-Oz approach, one team member said the voice command, the other team member played the related piece of audio, and all other team members provided their feedback on the voice commands and the piece of audio they heard. Although the Wizard-of-Oz approach (Fong & Frank, 1992; Klemmer et al., 2000) was used, the voice commands were kept short and simple because we wanted users to exert less cognitive effort to enact the commands (Bradford, 1995). Table 5 lists the voice commands (and the corresponding semantics) that were iteratively developed using this Wizard-of-Oz approach. For some of the semantics, we provided a few options in regard to the voice commands in order to determine which commands would be used the most.

Our set of voice commands belong to the following sources:

- The voice commands were partially inspired by the elements used to control a music player (e.g., next, skip, back, previous, pause, stop and play).
- Other commands were borrowed from traditional mechanisms used to control linear media (e.g., rewind, forward, restart and start).
- Another set of commands that we introduced was specific to the nature of aural flows (e.g., category name, what’s new, recent news, home, more, tell me more, like this and anything else).

Table 5. The vocabulary of the voice commands to control the aural flows.

Voice Commands	*System Action on Aural Flows*
U.S., World, Politics, Sports, Health, Science, Economy, or Technology	Select U.S., World, Politics, Sports, Health, Science, Economy, or Technology News Category
Start, What's New, Recent News	Starts Playlist of News
Restart	Restart Playlist of News
Rewind	Previous Section in News Story
Forward	Next Section in News Story
Back, Previous	Previous News Story
Skip, Next	Next News Story
More, Tell Me More, Anything Else, Related, Like This	Related News Stories
Home	Return to Home Page
Pause, Stop, Play	Click on the Button to Pause, Resume or Play

5.1.2. Manifesting Designs in *Linkless* ANFORA

In order to explore and evaluate the implications of the proposed navigation vocabulary for users browsing complex information architectures, we leveraged and improved on ANFORA News with *Linkless* ANFORA, which supports voice control over aural flows. In *Linkless* ANFORA, the aural flows were generated in real-time from existing news source (i.e., NPR website) and read aloud to users using a TTS service (www.ispeech.org). In order to demonstrate the navigation vocabularies used for dissemination and testing, two versions of *Linkless* ANFORA were instantiated in this study: one with button commands and one with both voice and button commands. Although the aural flows were fully implemented, the Wizard-of-Oz approach was used to control the participants' device when they used any of the voice commands (See Appendix E for the *Linkless* ANFORA prototypes). Hence, one researcher manually activated the commands voiced by the user through a control console.

The Wizard-of-Oz approach is a very common testing strategy for early designs of complex interfaces that need quick iterations of features that would normally require lengthy implementation processes (Dahlbäck et al., 1993). In the evaluation study, however, the researchers did not use the Wizard-of-Oz approach to do a complete exploratory evaluation of the voice commands. This decision was made because it would have been difficult for the researchers to execute a random command and guess what the participants meant in a controlled evaluation study.

5.2. Evaluation Hypotheses

Based on the principles of linkless navigation as applied to an aural website scenario, our research question (RQ) and hypotheses are as follows:

RQ: When navigating aural flows while on the go, does a set of voice commands reduce a user's visual interaction with the device and improve the user experience compared to clicking buttons in order to navigate through content?

- H1: Using voice commands, instead of button commands, requires less visual interaction with the device. (Although, by definition, using voice commands is expected to reduce the visual interaction, there are other factors that could come into play. For example, users might look at the screen while using voice commands because they are not yet familiar with the interaction modality or to check to see if the system did what they asked it to do.)

- H2: Users will find voice commands easier to use than button commands. (Although the voice commands are expected to be a more natural form of input, both voice and button commands could cause cognitive distractions.)
- H3: Users will find voice commands more enjoyable than button commands.

5.3. Study Design

In order to test the hypotheses, this paper conducted a controlled evaluation study with 20 users and adopted a within-subjects design in order to maximize internal validity.

5.3.1. Physical Setup

The evaluation study was conducted in an indoor navigation environment that included one large room connected to the main entrance corridor via another hallway (Figure 17). This study established a 54.4-meter long area that users walked while executing the aural browsing tasks. The path was marked on the floor using tape and included four sharp turns, two slight turns and two U-turns. Different static objects, such as tables and chairs, were placed along the route to simulate a real-world scenario in which an individual must safely recognize and navigate around obstacles. The participants were led through the path before they started with their tasks. The researchers limited the distractions to the available artifacts on the wall.

In order to effectively compare the experience of using voice commands to button commands, this study controlled for the condition of a noisy environment by conducting the study in an indoor environment. The researchers did not expect that the potential degradation of performance that might occur in a noisy setting would affect any particular problem; rather, they expected a reduction in accuracy, which would improve as the voice recognition system advanced. Additionally, the lists of voice commands were printed on an A4 size paper and placed on all the walls around the path (Figure 17). The lists of voice commands were comfortably readable from a distance of 190 cm. Therefore, the users could refer to these lists at any time in order to isolate the 'command learnability' factor of the study.

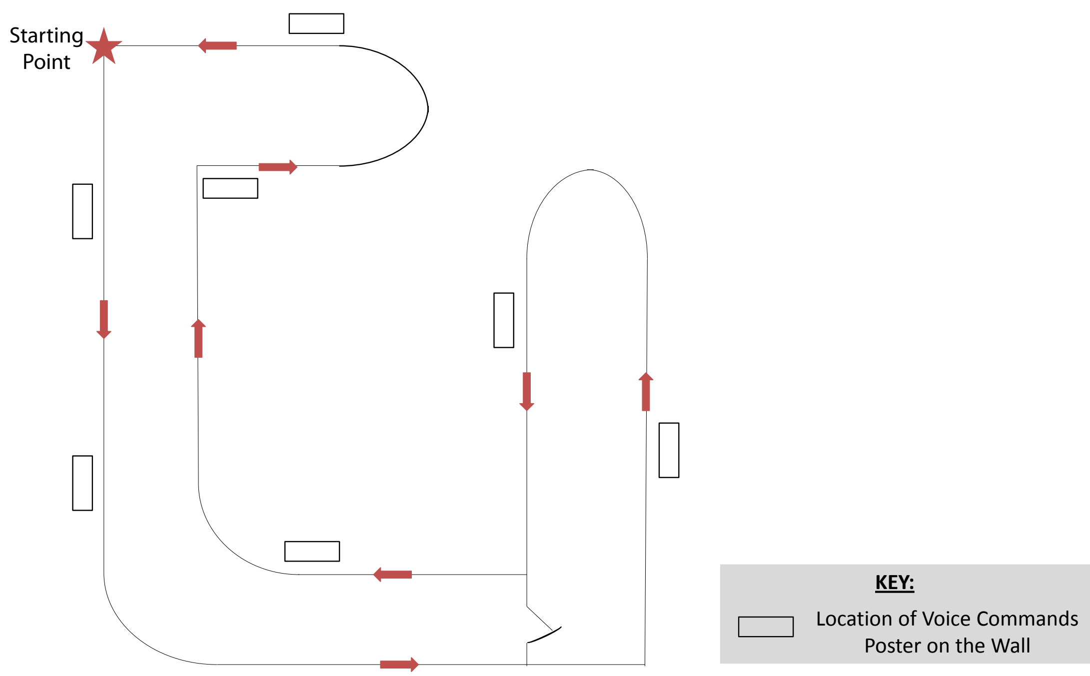


Figure 17. The path layout used in the experiment was 54.4-meters long with four sharp turns, two slight turns and two U-turns.

A distant side observer used a video camera to record the users' sessions and visual engagements with the application (Figure 18). A video recorder was used for two reasons. First, the researchers did not want to add new distractions to the experiment by making people walk around with a head-mounted eye-tracking devices (HEDs). Moreover, the condition of using an HED while walking is not externally valid. Second, the recorded video allowed the researchers to conduct post-task analyses and capture all other user activities (e.g., looking at the posters or the list of voice commands on the wall) during each task.

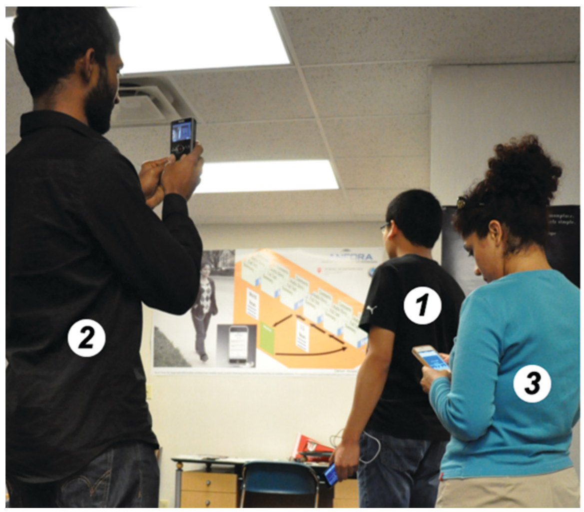


Figure 18. Experimental setup: 1. Participant listens to aural flows on *Linkless* ANFORA. 2. Researcher video records the session. 3. Researcher controls the flow and interaction.

The participants were encouraged to listen to the TTS content using Apple headphones and interact with the application using buttons or voice commands. They were instructed to hold the phone in one of their hands with their arms down while listening to the TTS content and hold the phone up when they used the button commands to interact with it (Figure 19). When the participants used a voice command, they had to click the button on the Apple Headphones Remote Button to simulate the real-world voice command activation. As the researcher had to walk behind the participants to hear their voice commands, the participants were made aware that the researcher was manually activating the voice commands through a control console.

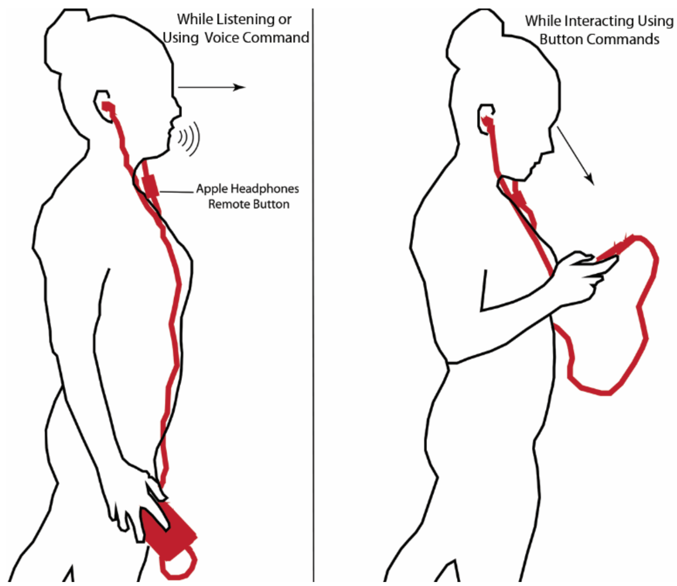


Figure 19. (Left) Participant is holding the phone in her hand with her arms down while listening to the aural flows. (Right) Participant is holding the phone up when she uses the button commands to interact with the aural flows.

5.3.2. Experimental Conditions and Study Variables

The independent variable was the style of navigation over the aural flows, which varied on two levels: (1) button- or (2) voice-plus-button commands. The researchers did not include a voice-only condition on the basis that current interfaces, such as Apple's Siri and Android's Google Voice (Android, 2015), typically provide voice commands as only one of the possible modalities, and almost never employ only one interaction modality to interact. Having multiple modalities for interaction is likely to accommodate a range of individual user preferences.

The dependent variables were as follows:

- *Interaction time (IT):* The overall time that the users were interacting with the interface *regardless of the modality* (voice or button).
- *Visual interaction time (VIT):* The time that the users spent listening to the aural flows while looking at or touching the interface.
- *Speed of walking*: The speed at which the participants walked while listening to the aural flows calculated by the total distance walked during a 15 minute task.
- *Frequency of using voice commands:* The number of times each voice command was used.
- *Instructed activities:* The number of activities performed by the users as instructed in the task, such as interacting via button/voice commands.
- *Non-instructed activities:* The number of activities performed by the users in addition to what was instructed in the task, such as looking at and/or reading text on the interface.
- *System usability:* The usability of the system as measured by the System Usability Scale (SUS) (Brooke, 1996) on the scale of 0 to 100 with cronbach alpha above .90 (Bangor, Kortum, & Miller, 2008; Lewis & Sauro, 2009).
- *Cognitive load:* The perceived mental demand of the task, as measured by the NASA-TLX (Hart & Staveland, 1988) (Cronbach Alpha above .70 (Hoonakker et

al., 2011)) on the scale of 0 to 100. Another strategy used to measure cognitive load is adding up the ICL, ECL and GCL scores. These scores are calculated indirectly through some of the questions in the SUS (Brooke, 1996).

The main purpose of using voice commands was to provide the users with a more eyes-free navigation experience. Thus, the researchers measured the visual interaction time in order to understand whether using voice commands required the users to look at the interface less than when they used only the button commands. In addition, visual interaction time and cognitive load were selected in order to measure visual and cognitive distraction, respectively.

5.3.3. Participants

Twenty participants from a large Midwestern University (10 male, 10 female) were recruited for this study. The participants ranged in age from 19 to 49 (M = 27; SD = 8.14) and were native English speakers and frequent news consumers. All of the participants had experience with touchscreen mobile devices and none had hearing impairments. None of the participants had prior experience with *Linkless* ANFORA or ANFORA News prortoype. The participants each received a $20 Amazon gift card for their 90 minutes of participation.

5.3.4. Procedure

Each participant engaged in a session that consisted of three parts executed in this order: (1) training; (2) two-stage task session, including the use of *Linkless* ANFORA in one of the two conditions, followed by usability and cognitive load surveys; and (3) a post-task interview.

5.3.4.1. Training

The participants attended a 30 minute training session, during which they were introduced to *Linkless* ANFORA and briefed about the voice and button commands. In order to make sure that all of the participants could reach a common threshold of familiarity with *Linkless* ANFORA, each participant executed simple navigation tasks using different versions of *Linkless* ANFORA.

5.3.4.2. Task Sessions and Post-task Surveys

The participants engaged in two stages of tests. The first stage used the button commands (B) as the control condition. The second stage used voice-plus-button commands (VB) as an experimental condition (hereafter to be referred to as "voice" condition). The order in which participants engaged in each style of navigation was systematically counterbalanced across all of the participants in order to minimize the learning effect. Overall, each participant executed two tasks (Figure 20):

a) One task (15 minutes) for the button condition and

b) One task (15 minutes) for the voice condition.

The structure of each task was the same across the different conditions. The only difference was the category of news stories covered. For example, the voice task was as follows:

"In this version, you may navigate using either the voice or button commands. You have 15 minutes to use *Linkless* ANFORA. Please browse at least eight news stories during this time period and change the category to any other category at least once. Try not to listen to the category of news to which you have already listened."

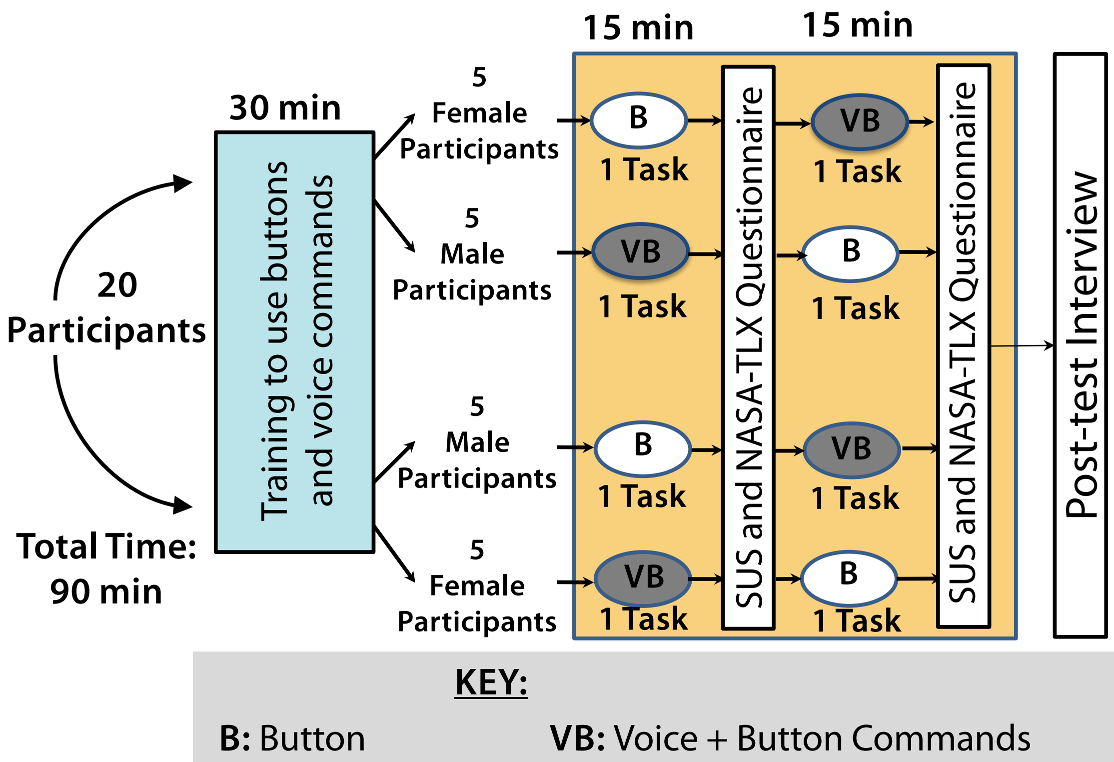


Figure 20. Within-subject design for the comparative evaluation of the different interaction modes.

The task for each condition was designed to be 15 minutes long because it was a good compromise between the depth and breadth of aural flows exploration and the fatigue caused by walking and listening to content. Overall, the researchers controlled for the task time (15 minutes), modality of interaction and continuous interaction. Within the constraint of time and modality of interaction, the researchers let the participants browse the aural flows freely in order to explore the content.

In a natural setting, users would be likely to employ several modalities at once. The combination of interaction techniques in one condition – voice and button – was used to preserve external validity. Moreover, the researchers' intentions were not to completely replace the existing button interaction techniques. Rather, they sought to provide users with more flexibility and additional options for navigating a semi-aural interface with natural and efficient aural navigation flows.

Finally, after each task, the participants rated the system's usability as well as their cognitive load using the SUS questionnaire (Brooke, 1996) and NASA-TLX questionnaire (Hart & Staveland, 1988), respectively.

5.3.4.3. Post-task Interview

After the two-stage task sessions and usability and cognitive load questionnaires, the participants answered interview questions related to both conditions. The purpose of the interview was to understand how the participants described their experience using *Linkless* ANFORA with different modalities; which modality of interaction they preferred to use in the voice condition and why; what they liked best or least about *Linkless* ANFORA; whether they listened to the news while walking and adequately monitored their surroundings; whether the orientation cues were clear to the participants; and in what other context would the participants prefer to use *Linkless* ANFORA (See Appendix F for the introductory script, training, tasks, surveys and interview questions).

5.4. Analysis

For the quantitative data, *repeated measure t-tests* were used in order to analyze the efficiency and effectiveness of the linkless navigation strategy as well as the effect of the interaction style. We used the interaction style (i.e., button vs. voice commands) as the within-subject factor. Several outcome variables (i.e., IT, VIT, walking speed, frequency of using voice commands, instructed activities, non-instructed activities, system usability and cognitive load) were compared.

Two researchers watched the recorded videos in order to measure both the IT and VIT in order to maximize the reliability of our measurements (inter-rater reliability metrics). Walking speed, instructed vs. non-instructed activities, and frequency of using voice commands were also measured by watching the recorded videos. System usability was reported using the SUS questionnaire and perceived cognitive load was calculated using the NASA-TLX on the scale of 0 to 100.

During our analysis, however, we connected the questions from SUS to specific types of cognitive load (see Table 1) that we wanted to capture. We choice to utilize the SUS in this manner because cognitive load is an important variable. Hence, in order to increase the reliability of our results, we measured cognitive load both directly and indirectly. Table 6 shows an example of how the SUS questions were mapped to different types of cognitive load. For the qualitative analysis of the interviews, we transcribed each of the interviews, extracted the recurrent themes and grouped the comments by type. The emerging issues highlighted user preference for the interaction paradigms and the difficulties faced while using the voice and button commands.

Table 6. Example of how the questions from the SUS were mapped to specific types of cognitive load.

Different Types of Cognitive Load	*Questions Selected from the SUS*
Intrinsic Cognitive Load (ICL)	**Q2.** I found this application unnecessarily complex. **Q3.** I thought this application was easy-to-use.
Extraneous Cognitive Load (ECL)	**Q5.** I found the various functions in this application well-integrated. **Q6.** I thought that too much inconsistency existed in this application.
Germane Cognitive Load (GCL)	**Q4.** I think that I would need assistance to be able to use this application. **Q10.** I needed to learn a lot of things before I could get going with this application.

5.5. Results

5.5.1. Interaction Times with Aural Flows

Figure 21a shows that the IT with the interface in the voice condition (M = 84.50 sec., SE = 9.93) was lower than the button condition (M = 114.35 sec., SE = 15.66) ($t(19) = 1.835$, p = .082). However, this difference was not found to be statistically significant. In the voice condition, on average, participants spent 55.1 second out of 84.5 seconds interacting with the device using the buttons (Figure 21a) and 29.4 second out of 84.5 seconds interacting with the device using the voice commands. On average, the participants spent 18 seconds looking at the voice commands posters on the wall. This activity was essential in regard to the users being able to interact with the voice commands, but the amount of time for this activity would decrease as users learn the voice commands. Hence, the time taken for this activity was not included in our interaction time measurement.

Two researchers measured the VIT. Based on the first researcher's measurements (Figure 21b), the users spent 51.11% less time visually interacting with the interface in the voice condition (M = 104.20 sec., SE = 20.32) than they did in the button condition (M = 213.15 sec., SE = 20.73) ($t(19) = 4.289$, $p < .01$), which resulted in a statistically significant difference. Based on the second researcher's data, the users spent 40.20% less time visually interacting with the interface in the voice condition (M = 121.00 sec., SE = 22.65) than they did in the button condition (M = 202.35 sec., SE = 19.36) ($t(19)$ = 3.693, $p < .01$), which is also a statistically significant difference. The inter-rater reliability correlations for the VIT by the two researchers were $r(19) = .057$, $p < .01$.

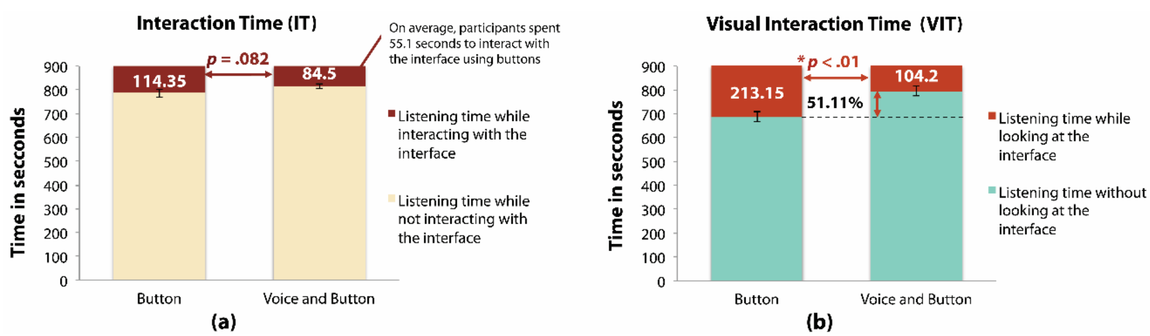


Figure 21. The voice commands (a) reduced the IT with respect to using buttons (with no statistical significance present), while the voice commands (b) also reduced the VIT with respect to using buttons (with statistical significance present).

5.5.2. Walking Speed, System Usability and Cognitive Effort

The participants' walking speeds while listening to the aural flows appears to be similar in the button (*M* = 58.22 cm/s, *SE* = 7.03) and voice conditions (*M* = 59.79 cm/s, *SE* = 6.94) ($t(19) = .536$, $p = .59$) (Figure 22a). Based on the SUS questionnaire, the system's usability appears to be similar in the button (*M* = 80.33%, *SE* = 2.75) and voice conditions (*M* = 77.50%, *SE* = 2.91) ($t(19) = .921$, $p = .37$) (Figure 22b) as well.

Based on additional user experience questions, in general, the participants reported that controlling the aural flows was slightly more comfortable, enjoyable, satisfactory, pleasing, simple and easy to understand in the button condition than in the voice condition (Figure 23). However, the participants found that their experience of using the voice commands to be more *engaging* than using the button commands. *Engaging* was presented to the participants and measured as a polar opposite in the semantic differential scale to *boring*.

The users' cognitive efforts – as based on the NASA-TLX questionnaire – in the two interaction conditions are compared in Figure 22c. The button condition (*M* = 23.57%, *SE* = 2.82) yielded a similar cognitive effort as the voice condition (*M* = 24.64%, *SE* = 2.74) ($t(19) = .550$, $p = .59$). The users' cognitive efforts were also calculated indirectly using some of the questions in the SUS (Table 6). The results showed that indirectly calculated cognitive load (using SUS) was significantly correlated with directly calculated cognitive load (using the NASA-TLX) in both the button ($r(19) = .491$, $p < .05$) and voice conditions ($r(19) = .632$, $p < .01$).

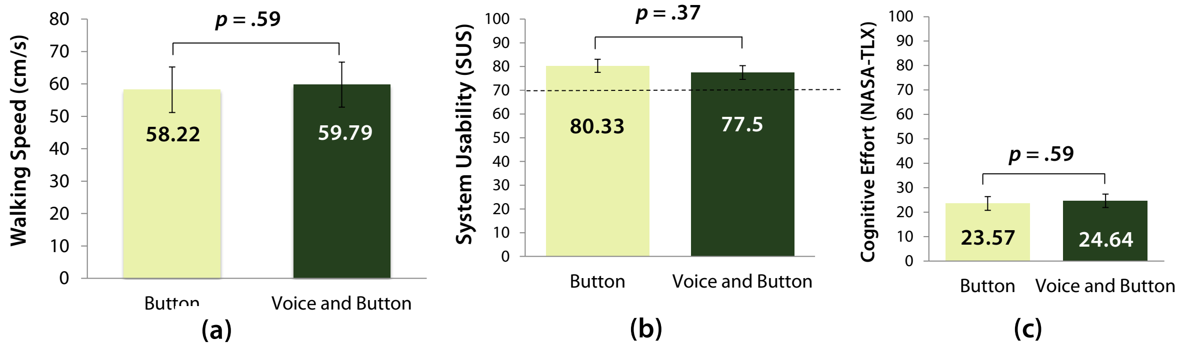


Figure 22. From left to right: No significant difference was found between the conditions for (a) the speed of walking, (b) system usability and (c) cognitive effort.

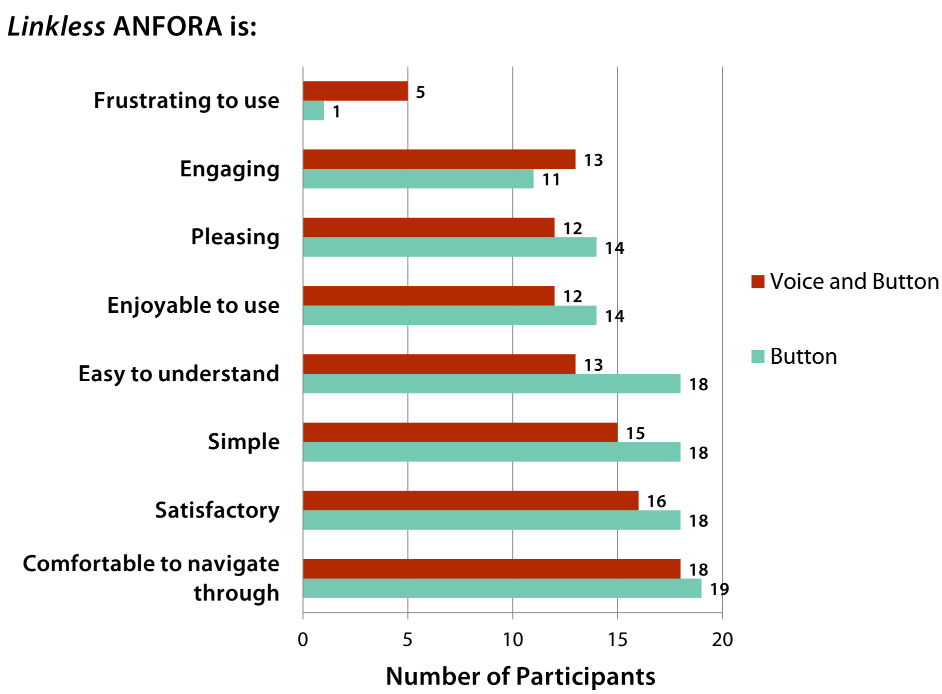


Figure 23. The participants who responded strongly agree/agree on every aspect of *Linkless* ANFORA experience.

5.5.3. Voice Command Usage

In the voice condition, the frequency of using the voice commands (M = 15.05, SE = 1.28) was significantly higher than the frequency of using the button commands (M = 4.85, SE = .97) ($t(19) = 5.293$, $p < .01$) (Figure 24). The average amount of time spent using the voice commands was 14.7 seconds. The three sets of commands used most often were as follows: (1) the "next/skip" command was used significantly more than all of the other commands (used 155 times; an average of eight times per participant; SD = 4.46); (2) the category selection commands, such as "technology," "world" and "health," were used the next most often (used 45 times; an average of two times per participant; SD = 1.92); and (3) the "forward" command was used to move from a story summary to a full version of the same story (used 41 times; an average of two times per participants; SD = 1.85). The "anything else" and "like this" commands were never used.

The results show that the participants used "next" (124 times) more than the "skip" command (19 times) to go to the next story and "back" (four times) more than the "previous" command (two times) to go back to the previous story. The participants used "related" (nine times) more often than "more" (five times) and "tell me more" (two times) to go to a related story. They also used "recent news" (five times) more than "what's new" (two times) and "start" (once) to begin listening to the aural flows playlist.

Additionally, the results show that one participant said, "reverse" instead of "back" or "previous" and "skip next" instead of "skip" or "next." Another participant used "related link" instead of "related" and 11 participants said "summary" for "rewind" and "full story" for "forward."

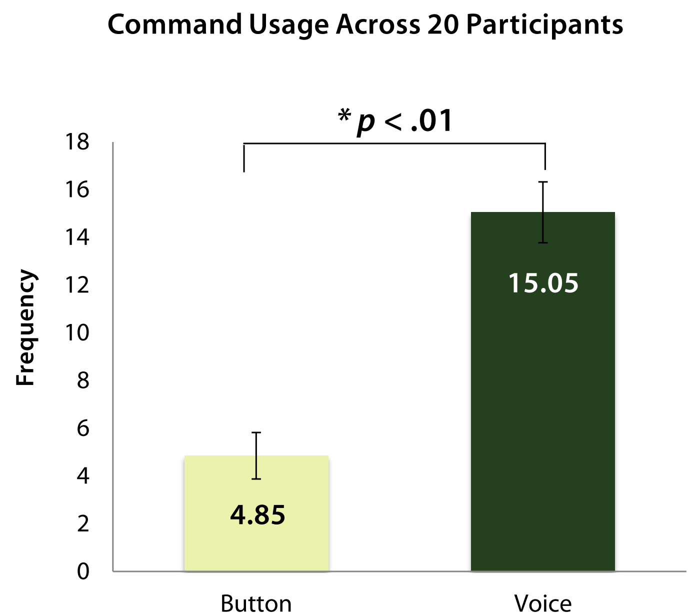


Figure 24. The participants used significantly more voice commands than button commands.

5.5.4. Instructed vs. Non-instructed Activities

In the voice condition, the participants performed significantly more non-instructed (*M* = 26.65, *SE* = 3.18) than instructed activities (*M* = 19.90, *SE* = 1.20) ($t(19) = 2.281$, $p < .05$) (Figure 25). Examples of instructed activities were the use of voice or button commands to interact with the interface. We also observed that the users looked at the list of voice commands or other artifacts available on the walls and glanced/read the news on the mobile interface, all of which are considered to be non-instructed activities. The participants either stopped to read the list of voice commands on the wall or glanced at it by turning their heads without stopping.

Similarly, in the button condition, the participants executed significantly more non-instructed (*M* = 23.40, *SE* = 3.07) than instructed activities (*M* = 10.95, *SE* = 1.42) (*t*(19)

= 3.701, $p < .01$). Taken together, these sets of results show that the participants performed more non-instructed than instructed activities regardless of the modality condition.

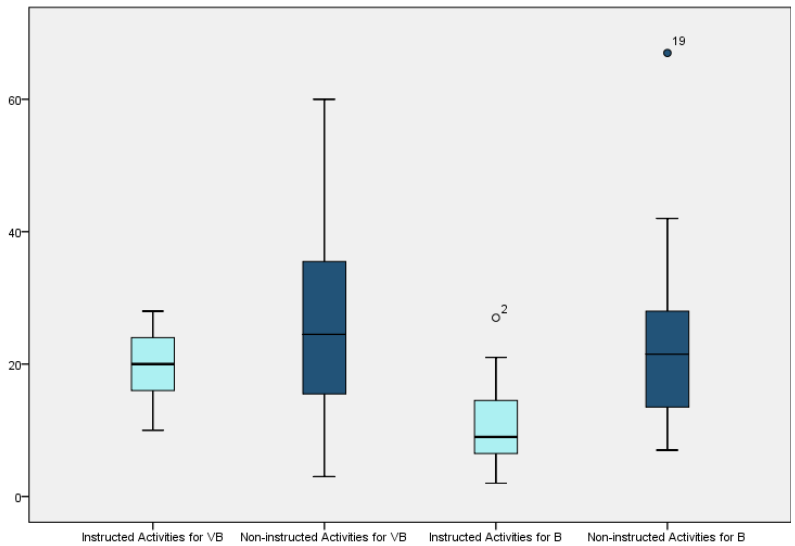


Figure 25. The participants performed significantly more non-instructed than instructed activities in both the voice and button conditions.

5.5.5. Interview Results

5.5.5.1. Self-reported Experiences

The interviews confirmed the users' general satisfaction with *Linkless* ANFORA as all 20 participants reported that it was easy-to-use and convenient. In particular, three users said that they liked the wide range of categories and content taken from NPR. For example, one participant (P18) noted,

> I liked that you guys used NPR. I liked that there was lots of different news categories. It wasn't just world news. I usually like the special interest, health and science, so I liked that it had those categories available.

Flexibility

Four of the participants reported that they liked the flexibility associated with not having to look at the screen. Furthermore, two participants reported that they liked moving from one category to another by using the voice commands. One user (P6) noted, "I was able to walk and not get distracted. I did not have to stop walking in order to press button commands on the screen and I felt safer because I was aware of my surroundings." Another user (P13) said, "I enjoyed the flexibility of not looking at the screen and being able to control the news category you liked to listen to."

Orientation

Fifteen users reported that they did not feel lost (in terms of where they were in the news content) while listening to the news story and felt that the orientation of information was good. Likewise, all of the participants recognized when a news story started or ended. One user (P12) noted, "I did not get lost, but if I did, I could have looked at the phone to know where I was." Another user (P18) said, "I did not get lost in what category I was in or what story I was listening to."

Competitive Uniqueness

Most participants reported that they had previously used other news applications, such as NBC news, CNN, BBC News, NPR, USA Today, Technews, and Stitcher. They all said that none of the applications they used previously are similar to *Linkless* ANFORA. In particular, two participants noted that they perceived *Linkless* ANFORA to be a new idea that is more akin to consuming news from radio and television broadcasts than from the web.

For example, one of the participants (P2) noted, "*Linkless* ANFORA differs from radio because with the radio you cannot skip over stuff that you do not want to listen to, and you have to wait to get to the next one. But *Linkless* ANFORA is broad, as far as the topics. If you like a certain topic, you can go back to that instead of going through everything." The other participant (P10) said, "It was very up to date and up to the point unlike the radio or TV news, there are commercials in between. Sometimes, I just do not like hearing them again and again. *Linkless* ANFORA was just very short, you could listen to the summary and if you are interested, you could listen to the full story." Finally, a third user (P18) commented on carrying *Linkless* ANFORA everywhere:

> I can use *Linkless* ANFORA in the morning with my headphones when I cannot turn on the radio while my roommates are sleeping. Even if I can turn on the radio, when I leave my room, I cannot hear the radio any more but with my phone, I can just walk wherever I want and I do not miss anything when I walk from my apartment to my car.

5.5.5.2. Multitasking

Eighteen of the 20 participants said that they could adequately monitor their surroundings while listening to the news. However, one participant (P10) had to stop walking while using the button commands and was not able to monitor his surroundings. He said, "I wonder how different [my experience will be] when I am walking in a crowded area." Three participants mentioned that the walking path was the same in both conditions and that there were not many obstacles, making it easy to monitor their surroundings.

5.5.5.3. Combining the Visual and Voice Commands

The participants were asked whether they preferred to use the voice commands, button commands or a combination of both types in order to interact with *Linkless* ANFORA. All of the participants used the voice commands, but three noted that they would prefer the button commands. They did not like the voice commands for four reasons. First, it was odd to speak aloud while alone in a public setting. Second, they had had prior negative experiences with the use of voice commands, particularly when it came to voice recognition interfaces. For example, they had to speak the voice commands several times until the system recognized it. Third, the participants had to learn and memorize commands that were named differently than they were on the interface, which could be time-consuming. For example, the voice command to move to a full story while in the summary is “forward” instead of “full story” and the command to go back to a story summary is “rewind” instead of “summary.” Forth, the difference between the “forward” and “next” commands was also confusing because “next” would go to the next story, while “forward” would go to the full story within the same story.

Other participants, however, reported that they liked using the voice commands. Five of the participants noted that they did not have to stop walking to look down at the screen. Instead, they could do other things while using the voice commands, such as monitor their surroundings and look at posters on the walls. According to one participant (P6), “I felt safer because I was aware of my surroundings.” Another participant (P14) said, “The voice commands were quicker compared to the button commands.” One user (P9) noted, “It was easy to go from category to category just by speaking into it without going back to the home screen, so it was convenient. It was just all on the fly.”

Seventeen of the 20 participants mentioned that they preferred to use a combination of the voice and button commands, but they had a variety of reasons. For example, one participant (P14) said, “If voice does not work, I can still benefit from the button commands.” In other words, the button commands can be used as a backup navigation method if the voice commands are not working properly. Having button commands as a backup navigation method is a significant concept, as tone and tenor of voice, as well as voice quality and accents vary among individuals, making voice commands potentially less precise than button commands.

The other main reason that the participants cited for preferring a combination of the voice and button commands relates to the contexts in which *Linkless* ANFORA might be used. For example, one user (P3) noted, “I would use the voice, but, if I’m leaving class, I would click on a story and go walking from there and then use the voice.” Another user (P8) said, “If I am at a noisy place, like a subway, I would use the button commands. If I am walking in a quiet place, I would use the voice. I think it depends on the environment.” A third participant (P15) reported, “If you come to talk to somebody, you would want to pause it with your finger, but if you are just walking around, you could just tell it what to do and do it.” Another participant (P19) noted, “Like, if I were crossing a busy street or riding my bike, I would definitely prefer to use the voice than the button commands.” Finally, another participant (P3) said, “If I were sitting somewhere, like a coffee shop or something, I might use the button commands because I’m not moving, but, if I’m walking, then I would use the voice.”

5.5.5.4. Other Contexts for Voice-controlled Aural Flows

The participants suggested other contexts in which *Linkless* ANFORA could be useful. Three participants noted they would use *Linkless* ANFORA while driving, when their eyes and hands are busy. One participant (P5) noted, “This app is more appropriate for a driving context than only a walking context because, while walking or sitting down, I prefer to read it, which is faster than just listening to the content.” Another participant (P18) said, “If I was driving, probably, I would use the voice commands because I did not have to look at my phone screen.” Several other potential contexts of use included: while on the way to work/class, outside a classroom, while sitting in a coffee shop, on the bus, while exercising, while riding a bike and while working around the house.

5.5.5.5. Limitations and Improvements Suggested by the Users

The users also provided suggestions on how to optimize the usability of *Linkless* ANFORA.

Repetition of the Orientation Information

Seven of the participants were frustrated with the repetition of the orientation information. For example, each time a new story began, *Linkless* ANFORA included audio that reported the story number, category and news headline. Two of the users said that the story number was of little interest. One participant (P8) added, “If I was listening to a research paper, maybe it would be necessary, but not for a news story.”

Confusing Category Transition

Additionally, four participants said that the transition between two categories of news was not clear. One participant (P4) said, “I guess I didn’t understand when it switched from one category to another and I was like, oh wait, I'm not in Science anymore. I'm in Economy or whatever it was." Two users wanted some indication of when a story was finished, such as audio stating 'end of story.'

5.6. Discussion

5.6.1. Voice Commands and Eyes-free Browsing

Our study provides some empirical support to H1: Using voice commands, instead of button commands, requires less visual interaction with the device. On average, compared to the button condition, the voice condition saved about 40.20% to 51.11% of the time in visual interaction with the device. Therefore, combining voice commands with aural flows and button commands reduced visual interaction with the screen when compared to using button commands with aural flows. Likewise, this result validates the primary value of extending the interaction with aural flows through voice commands.

In the voice condition, we also observed that the participants looked at the screen not only when they used the button commands, but, also, when they used voice commands for different reasons. For example, users were not yet familiar with the interaction modality or they checked to see if the system did what they asked it to do. We hypothesize that this visual interaction while using voice commands could decrease as users become familiar with and trust the application.

Our study also confirms the findings from another recent study (Brumby et al., 2011) on the use of mobile devices during secondary tasks. This study indicated that, although audio-based interfaces are slower to use, they are less distracting than visual interfaces. However, an important question is still unanswered: To what extent do combinations of aural flows with voice commands support eyes-free browsing while driving a car? Some of our participants noted that they would prefer to use *Linkless* ANFORA while driving. Furthermore, a recent study (Strayer et al., 2013) reported that using TTS systems for sending and receiving text or email messages in the car is risky because too many and continuous voice interactions can also cause higher levels of cognitive distraction. We hypothesize that, by using a small vocabulary of voice commands (Feng & Sears, 2009), which are short and easy to remember (Bradford, 1995) as discussed in the suggested design guidelines, the cognitive effort required for the use of *Linkless* ANFORA is still minimal and will not distract users from effectively monitoring their environments. This hypothesize is because users will not continuously have voice interactions with *Linkless* ANFORA and will only use a few short commands that will not tax their attention. Thus, future research needs to focus on using aural flows with voice commands while driving a car.

5.6.2. Similar System Usability, Users' Cognitive Efforts and Walking Speed

Both the button and voice conditions yielded a similar system usability and cognitive effort. Therefore, H2 was not confirmed. This similarity in the two conditions is, most probably, because aural flows already improve system usability and reduce cognitive effort so significantly – with respect to visually interacting with content-intensive websites on a mobile device – that merely changing the interaction style has no additional effect. Figure 22b shows that the system usability for the button and voice conditions were reported as 80.33% and 77.5%, respectively, which is close to an excellent rating (Bangor, Kortum, & Miller, 2009). Cognitive effort for both the button and voice conditions is 23.57% and 24.64%, respectively, which is a low cognitive effort score

(Figure 22c). Overall, our results show that aural flows yield a very good user experience in both the button and voice conditions.

Additionally, the low cognitive effort engendered by aural flows regardless of the interaction modality allowed the participants to do more non-instructed than instructed activities. This finding is because the users spent 12.71% and 9.39% of the time interacting with the aural flows (i.e., instructed activities) in the button and voice conditions, respectively (Figure 21a) and engaged in non-instructed activities during the remaining time. For example, the participants looked at the posters on the wall or glanced at the mobile visual interface, which were not instructed to them as part of the task. This result is mainly relevant for multitasking experiences while on the go because attention to the mobile device and the risk of having an accident are minimized.

Similarly, the participants' walking speeds were similar in both the button and voice conditions. This result shows that the interaction modality did not have measurable effect on their walking speeds. As we discussed previously, the voice commands significantly reduced the amount of time required to interact visually with the device. However, the participants' walking speeds show that not focusing on the device does not necessary make the users walk faster. This finding could be because the participants had to walk the same path in an indoor environment repeatedly. Figure 8a shows that the walking speeds for the button and voice conditions were 58.22 and 59.79 cm/s, respectively, which is far below the average walking speed for adults (140 cm/s) in the 20- to 30-years-old age range (Bohannon, 1997). This finding could be because the participants had 15 minutes for the task and were not in a rush to finish the path or reach a particular destination. We realize that the participants walked in an environment where there were no dynamic obstacles and the static obstacles were always present in the same position. Therefore, it is difficult to reach an ultimate conclusion about the real effects of distracted walking because of the nature of our environment.

5.6.3. Experience with Voice Commands

The analysis of the recorded videos revealed that the participants used the voice commands significantly more than the button commands to interact with the aural flows. However, the participants' answers to the interview questions revealed that 85% of them chose a combination of both the voice and button commands by which to interact with the aural flows for different reasons. One of the reason was because some of the users reported poor previous experiences with voice commands. The main reason for their criticism was related to their perception that the tone and tenor of their voices, as well as voice quality and individual accents, affects systems' abilities to understand them.

5.6.3.1. Contradictory User Experiences with Navigation Modalities

A few possible reasons exist as to why the user experience was slightly less favorably in the voice condition than in the button condition (Figure 23). The Wizard-of-Oz approach introduced a longer pause between actions for when a voice command is used compared to when a button command is clicked. Additionally, it may be difficult for users to quickly learn the voice commands and differentiate them from one another (e.g., next and forward). For example, in response to the statement, "I found this application [voice condition] very cumbersome/awkward to use," a participant rated the application as a five on a scale of one to seven (one = strongly disagree, seven = strongly agree). This same participant also rated "I needed to learn a lot of things before I could get going with this application [voice condition]" with a 7.

One participant reported that using the button commands was less satisfactory and less enjoyable, but also simple, easy to understand and engaging. This discrepancy between user experience attributes could exist because, although the button interface is easy-to-use, the user had to stop walking to click the button commands. Three of the participants reported that using the voice commands was more frustrating than the button commands, but that the voice commands were simple, pleasing and enjoyable. The reason for this apparent contradiction is that, although the interface is easy-to-use, the user was frustrated with the repetition of orientation information (reported in our Interview Results, Section 5.5.5.5).

Our participants rated their user experiences slightly less favorably for the voice condition than for the button condition. However, they enjoyed using the voice commands slightly more than the button commands. One possible reason for this finding is that users do not have to look at the screen to interact with the device and can, instead, enjoy listening to the news while navigating with the voice commands.

5.6.4. Consistency between the Aural and Visual Interfaces

Our study reinforces the importance of the principle of 'consistency' between the voice commands and the written labels on the button commands. For example, the *Linkless* ANFORA interface includes two button commands, "summary" and "full story," but users must say "rewind" and "forward" to move between summaries and full stories. Our design included very simple playlist-like commands (e.g., forward and rewind), which were applicable to the playlist metaphor. On the other hand, to control the visual condition, we used a tab structure that includes "summary" and "full story," which represents different sections of the news (i.e. world news vs. local news). At times, users said "summary" or "full story" instead of "rewind" or "forward." Users reported that the labels on the button commands were not consistent with the voice commands, which caused confusion. While the common principle of consistency (Nielsen & Molich, 1990) usually applies to visual interfaces, studying semi-aural interfaces suggests the importance of examining issues related to *cross-modal consistency* (Evans & Treisman, 2010; Spence, 2011). For example, how consistent do aural and visual interfaces need to be? Does the consistency contribute to having natural interactions with the semi-aural interfaces?

5.6.5. Limitations of the Study

One limitation of our experimental design is that the users had to walk in a controlled lab environment in order to avoid putting them in danger. Additionally, not having natural distractors in our environment could have affected the cognitive load measurements. The interview findings suggest that additional studies in which participants are put in new scenarios might be valuable in the future. The second limitation is that the users had to walk the same path with the presence of static obstacles for both conditions. Familiarization to the path, however, is partially lessened by the counterbalancing of two conditions.

The third limitation is that the participants had to learn the voice commands and the *Linkless* ANFORA interface in a short period of time. Therefore, they were provided with lists of voice commands on all of the walls surrounding the path in the event that they could not remember them. Thus, learnability was factored out of the cognitive load measurement.

The fourth limitation is that the voice commands were not fully implemented in the system. Instead, we used the Wizard-of-Oz approach in order to simulate voice interaction. The decision to use the Wizard-of-Oz approach was made in order to minimize the chances that many different speech patterns and/or accents would result in a high number of system errors, which would interfere with our ability to effectively measure the linkless user experience. Additionally, the Wizard-of-Oz approach led to a faster response time than might be expected in a real system.

The fifth limitation is that we did not accurately capture whether the participants preferred button commands for certain types of interactions, although we did observe patterns of preferences while recording the participants' videos. For example, to go to the next or previous news story, sometimes the participants preferred the button commands. However, in order to change the news category, the participants preferred the voice commands instead of going through the menu selection using the button commands. The sixth limitation is that the participants were not restricted to listening to a certain number of news stories, but were simply told to listen to a minimum of eight news stories. Therefore, all participants did not have the equal number of interactions with aural flows, which might have affected on some of the outcome variables.

5.7. Conclusion and Future Work

This study is the first study to demonstrate the properties of aural flows in the context of how to interact with them. Aural and semi-aural interfaces have the potential to amplify users' abilities to navigate the mobile web more safely and with fewer visual distractions from their surroundings. This work compared navigating aural flows with two different interaction modalities (i.e., voice and button). The results suggest that voice commands in combination with aural flows and button commands reduce visual interaction time with the device up to one-half compared to using button commands in combination with aural flows while walking. The results of the two conditions were also similar in terms of walking speed, system usability and cognitive effort. Overall, the low cognitive effort engendered by aural flows (regardless of the interaction modality) allowed the participants to do more non-instructed than instructed activities. We must consider that a noiseless environment and no errors in voice recognition were included as assumptions to reach the above conclusion. Hence, the ecological validity of the study is limited. In future studies, we will add errors in the Wizard-of-Oz approach (Fong & Frank, 1992; Klemmer et al., 2000) to better simulate a more realistic scenario.

Several of our participants suggested that they would like to use *Linkless* ANFORA while driving a car. A recent study (Strayer et al., 2013) suggested that using *speech-to-text* systems in the car is risky because too many voice interactions still tax our attention bandwidth. We are interested in studying whether the user's ability to listen to aural flows as he unfolds minimizes interaction and mitigates this problem. In the next chapter, I will present ways by which to use aural flows to mitigate the distraction by reducing both the visual and vocal interactions in a driving scenario.

Chapter 6. ANFORA*Drive* and Evaluation

As shown in the previous chapter, our participants were highly interested in using *Linkless* ANFORA as a form of infotainment technology while driving. Infotainment technologies provide a combination of information and entertainment contents, such as are available via a radio, CD player or smartphone (Demers, 2005). These infotainment technologies are widely used by young drivers (Alt et al., 2010), but studies have shown that they can distract them from safe driving (Lee, 2007). In this chapter, I assess the impact of *Linkless* ANFORA on drivers in order to gain a better understanding of a potential infotainment technology that provides content, while being less distracting than traditional infotainment technology. Moreover, driving was selected versus other contexts (e.g., jogging, exercising, biking or cooking) suggested by our participants because the cognitive load in the context of driving is higher than the other contexts. Therefore, I could test *Linkless* ANFORA in two extreme environments (walking and driving) in terms of cognitive load requirement. For simplicity, in this context, we will refer to *Linkless* ANFORA as ANFORA*Drive*.

6.1. Aural Flows in the Context of Driving

The web survey conducted by Alt et al. (2010) showed that more than 90% of the respondents used a fixed or mobile display for navigation or entertainment purposes while in the car. This web survey (Alt et al., 2010) had two important findings relevant to our research. First, more than 70.3% of the respondents preferred audio to text, images, emails and videos as a form of entertainment. Second, 83.6% of the respondents preferred general news as the type of content to listen to while driving. The reason for this preference could be as a result of an adaptation to the use of radio in cars (Alt et al., 2010). Therefore, ANFORA*Drive* could be a perfect example of embodying these elements (i.e., audio plus news) in a new way (e.g., aural flows).

Additionally, the rapid evolution of infotainment technology has become a distraction for young drivers more than other driving populations. This distraction occurs because young drivers have less experience in being able to anticipate and manage hazards while driving . Hence, an important question emerges: Could ANFORA*Drive* be an example of suitable alternative in driving scenarios to enhance content-rich, non-distracting infotainment technologies?

6.1.1. Comparing Competing Aural Browsing Solutions

Based on the user evaluation study conducted in Chapter 5, we discovered that a number of applications similar to ANFORA*Drive* exist that contain one or more of the following items: pre-built playlist, default semi-aural or aural access, and voice-based category access. Some of these news applications are the BBC, CNN, NPR, Stitcher, Umano and USA Today. In a preliminary activity, we scanned the environments of these competing applications, investigated how users could interact with and consume the news through them and, finally, decided which application to pick for our controlled experiment.

After careful consideration, we selected Umano (Umano, 2015), which provides news stories in audio format with an easy-to-use interface to compare with ANFORA*Drive*. Umano, however, has some differences with respect to ANFORA*Drive* (Table 7), including:

- **Pre-built playlist of all of the categories concatenated vs. one category:** ANFORA*Drive* enables users to listen to a pre-defined or pre-built playlist of news. This pre-built playlist covers all of the available categories in a list (i.e., full flow), but users can decide on the category of news from which they are interested in starting. While listening to the playlist, users can change the category by clicking a button to activate the device microphone and then say the voice command (Figure 26). The concept of a pre-built playlist also exists for Umano, but it only covers one category at a time (the equivalent to group flow in ANFORA*Drive*). For example, users can listen to only U.S. News. Therefore, in Umano (Figure 27), to listen to a different category, users have to return to the list of selected channels, select the channel they are interested in and choose a news story. This entire process consists of four clicks (Figure 26). Therefore, in order to prepare the news playlist, Umano requires to user to visually interact with the device more than would be required of ANFORA*Drive*.
- **All-to-All vs. index category access:** ANFORA*Drive* provides users with an all-to-all navigation pattern among new stories across all categories (Figure 26), which means that users can begin to listen to the news from any category and can move to any other category without having to return to the index page. Umano, however, provides users with a separate index navigation pattern for each category (Figure 26), which means that the users are required to return to the index page every time they want to change the category.

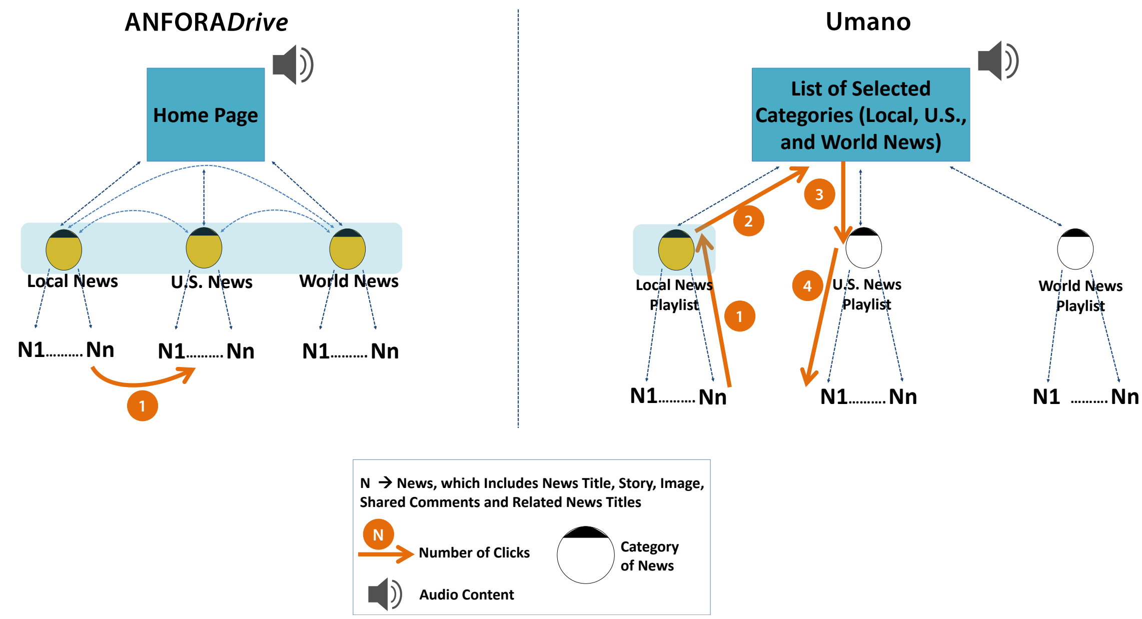


Figure 26. (Left) ANFORA*Drive* provides all-to-all access and needs only one click of the steering wheel button to change the category via voice commands. (Right) Umano provides index access and needs four clicks on its interface to change the category.

- **Voice- vs. visually-based category access:** In order to interact with the abovementioned navigation patterns, ANFORA*Drive* enables users to say the category name (i.e., eyes-free modality of interaction) and the playlist jumps to that category. In Umano, users have to return to the index page by clicking on the back buttons (visual interaction), select another category, and then select the news stories they are interested in.
- **Multiple reading levels vs. one reading level:** ANFORA*Drive* introduces different types of content categorizations especially suited for aural navigation. For example, users can choose to listen to a segment of news stories (i.e., title, summary or full story) based on their time constraints and degree of interest in

the content by selecting related stories. Umano only provides the full story and does not provide access to summaries or related stories.

- **Default semi-aural vs. aural access:** ANFORA*Drive* provides the news in both audio and text formats simultaneously. In Umano, users have to make two clicks to see the text of a news story, if interested.

Table 7. ANFORA*Drive* and Umano Comparison.

Aural Flows (Manifest in ANFORADrive)	*Alternative Solutions on the Market (Manifest in Umano)*
Pre-built Playlist of All Categories Concatenated (i.e., Full Flow)	Pre-built Playlist of One Category (i.e., Group Flow)
All-to-All Category Access	Index Category Access
Voice-based Category Access	Visually-based Category Access
Multiple Reading Levels	One Reading Level
Default Semi-aural Access	Default Aural Access

In summary, we can characterize ANFORA*Drive* and Umano as follows:

- **ANFORA*Drive*** is a voice-controlled full flow with all-to-all access to news categories that supports different reading levels, including a summary, full story and related stories.
- **Umano** is a button-controlled group flow with index access to news categories and access to the full story only (Figure 27).

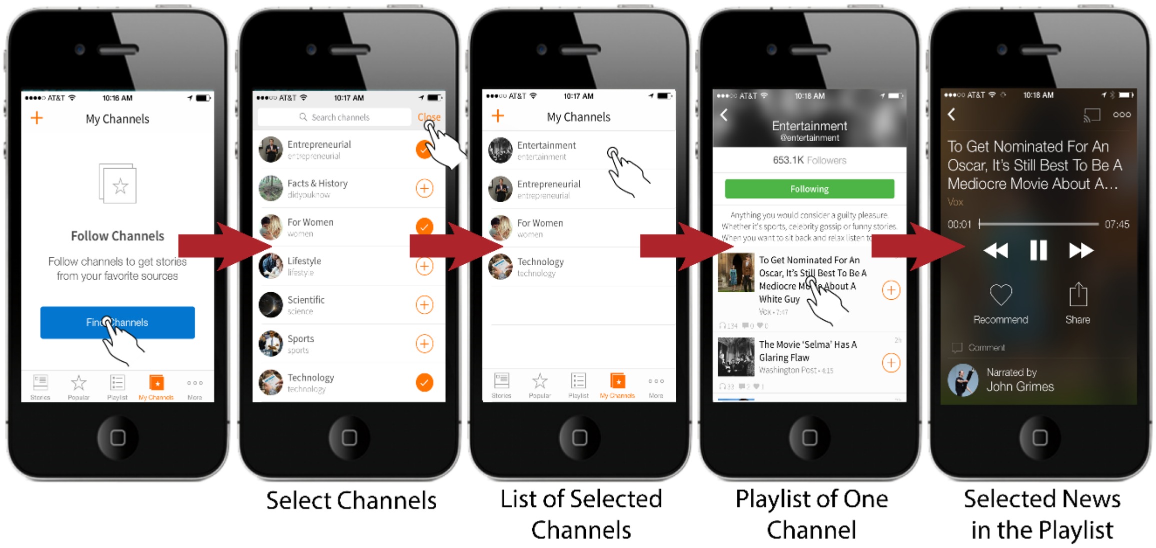


Figure 27. The Umano application interface displaying the step-by-step process of adding channels to a list and selecting which story to listen to.

6.2. Evaluation Hypotheses

Based on the abovementioned comparisons, we have defined our research question and hypotheses as follows:

RQ: How the use of a *voice-controlled aural flow* (e.g., ANFORA*Drive*) provide less of a distraction and improve driving performance than an *alternative solution on the market* (e.g., Umano) or a situation in which no flow or solution is utilized?

Compared to the driving only condition (i.e., No Device),

- H1.1: ANFORA*Drive* does not increase the driver's cognitive effort.
- H1.2: ANFORA*Drive* does not increase driver distraction.
- H1.3: ANFORA*Drive* does not reduce overall safety.

- H1.4: ANFORA*Drive* does not reduce driving performance.
- H1.5: ANFORA*Drive* does not increase the driver's visual interaction time with the device.

Compared to the driving only condition (i.e., No Device),

- H2.1: Umano increases the driver's cognitive effort.
- H2.2: Umano increases driver distraction.
- H2.3: Umano reduces overall safety.
- H2.4: Umano reduces driving performance.
- H2.5: Umano increases the driver's visual interaction time with the device.

Compared to Umano,

- H3.1: ANFORA*Drive* reduces the driver's cognitive effort.
- H3.2: ANFORA*Drive* reduces driver distraction.
- H3.3: ANFORA*Drive* increases overall safety.
- H3.4: ANFORA*Drive* increases driving performance.
- H3.5: ANFORA*Drive* reduces the driver's visual interaction time with the device.
- H3.6: ANFORA*Drive* increases user satisfaction while using the device.

6.3. Study Design

In order to test our hypotheses, we conducted a controlled evaluation study with 60 users and adopted a within-subjects design (from the participants' perspectives) to maximize internal validity.

6.3.1. Preliminary Pilot Study

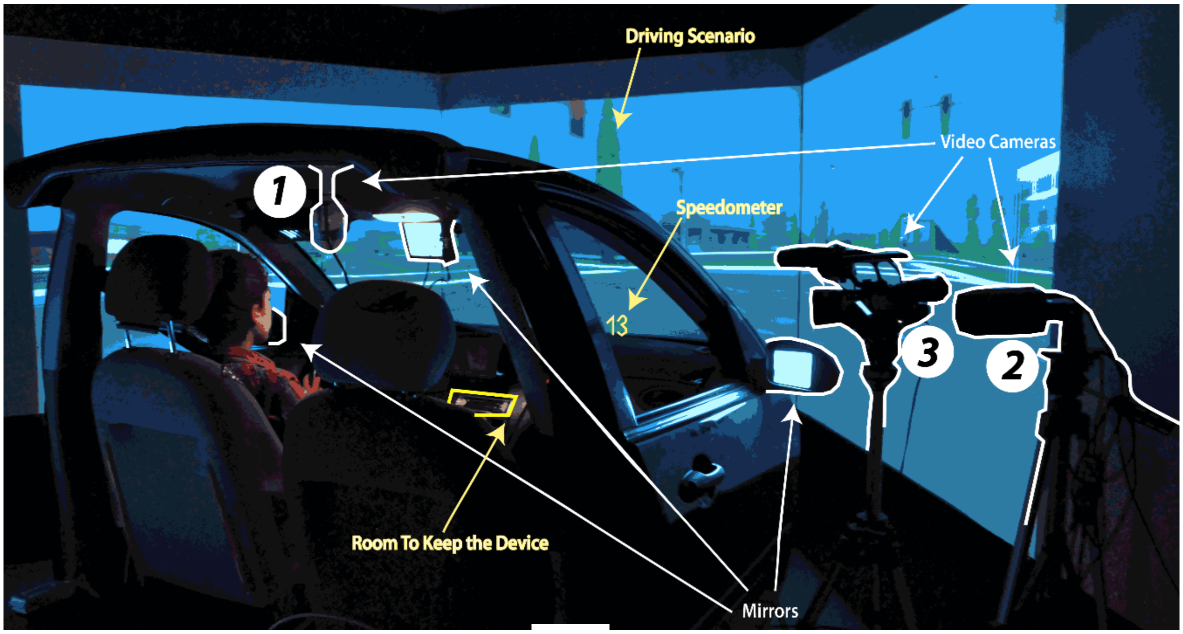
Before we conducted the controlled study, we ran several iterations of the pilot study in the Transportation Active Safety Institute (TASI) lab (TASI, 2015) at the School of Engineering and Technology at IUPUI. The preliminary study was undertaken with five participants who tested the experiment procedures, accuracy and appropriateness of the driving scenario, relevance of the tasks and length of the training. The pilot study also enabled us to improve different aspects of the controlled study. Based on the results of the pilot study, we conducted a controlled evaluation study from November 2014 to April 2015.

6.3.2. Physical Setup

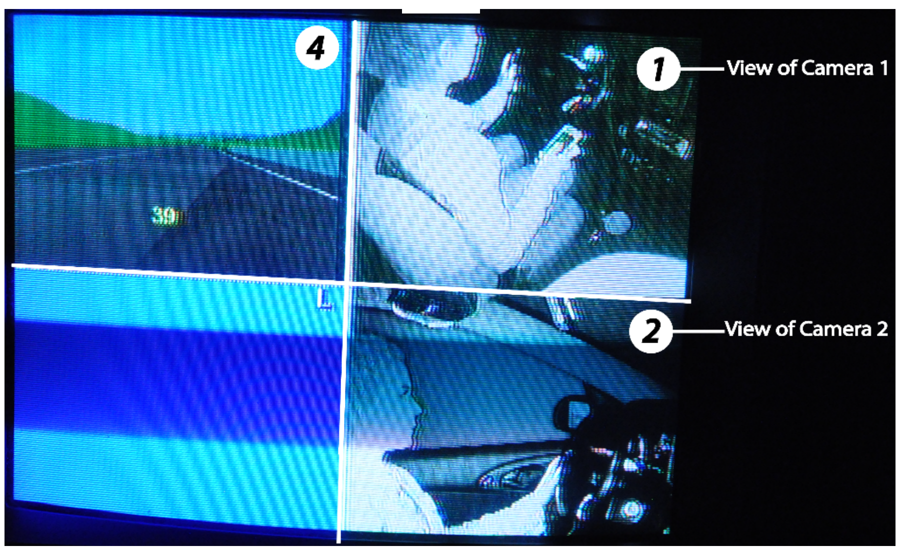
The evaluation study was conducted in the TASI facility at IUPUI, which is a controlled driving simulation environment. The driving simulation used at the TASI lab is called DriveSafety DS-600c.

The DriveSafety DS-600c provides a flexible and realistic environment for testing. The Drive Safety DS-600c projects roadway images onto three large screens positioned in front of the cab of a Ford Focus to provide an immersive driving experience.

This driving simulation also utilizes three mirrors: a center mirror and two side mirrors to account for blind spots. In the TASI lab, the users drove in the car simulation while executing aural browsing tasks. We recorded the user sessions as well as the users' visual engagements with the applications using three cameras mounted inside and outside of the car (Figure 28a). The participants were encouraged to use both ANFORA*Drive* and Umano during the study.



(a)



(b)

Figure 28. (a) Physical setup – Three video cameras record the user's visual interactions with the device while driving and the speedometer is displayed on the screen in front of the driver. The feed of cameras one and two were displayed on the control monitors and the feed of camera three was recorded separately. (b) Controlled monitor with three feeds: (1) view of camera one, (2) view of camera two and (4) the driving scenario.

6.3.3. Experimental Conditions and Study Variables

For this evaluation study (Figure 30), the two major *independent variables* were the *aural application* and *driving scenario complexity*. The *aural application* varied on three levels: (1) no aural applications or devices (i.e., driving only task / control condition), (2) ANFORA*Drive* and (3) alternative solution on the market (Umano). In order to gain a better understanding of the impact of aural applications on driving performance, system usability and distraction in various conditions, we also modelled three *driving scenario complexities*: (1) low, (2) moderate and (3) high. The low complexity scenario consisted of a single-lane environment with low traffic volume and a low-speed limit, such as would be found in a residential neighborhood. The moderate complexity scenario consisted of two lanes with a higher traffic volume and higher speed limit, such as would be found in the suburbs. The high complexity scenario consisted of a multiple lane environment with left or right turns, a much higher density of traffic volume and much higher speed limits, such as would be found in highway and city driving. The design of the various complexity levels was consistent with the guidelines indicated in previous studies (Horberry, 1998; Horberry, Anderson, Regan, Triggs, & Brown, 2006; Horberry & Edquist, 2008; Justiss, Mann, Stav, & Velozo, 2006).

The major *dependent variables* were as follows.

- *Perceived distraction:* Self–reported distractions measured using two questions (See Appendix F for questionnaire).
- *Overall safety:* The driver’s safety was measured by one question (See Appendix F for questionnaire).

- *User satisfaction:* User satisfaction with the aural application was measured using one question, but the participants also rated how pleasing vs. annoying, enjoyable vs. unenjoyable, simple vs. difficult, engaging vs. boring, and easy to understand vs. confusing they found the aural applications.
- *System usability:* The usability of the system was measured using the SUS score (Brooke, 1996) on the scale of 0 to 100 with cronbach alpha above .90 (Bangor, Kortum, & Miller, 2008; Lewis & Sauro, 2009).
- *Cognitive workload:* The perceived mental demand of the task as measured by the NASA-TLX (Hart & Staveland, 1988) (Cronbach Alpha above .70 (Hoonakker et al., 2011)) on the scale of 0 to 100.
- *Aural flow and voice command usage:* The average number of times the participants changed the categories and used voice commands in each of the aural applications.

The *driving performance variables* were as follows.

- *Number of lane departures:* The number of times the participants went out of the lane without using the right or left turn signal.
- *Response time:* The amount of time the participants took to hit the break or use the left or right turn signal before taking an exit or turning left or right.
- *Number of accidents:* The number of times the participant crashed into another car, pedestrians or bicyclist.
- *Lateral lane position (SD):* The standard deviation of the lane position angle.
- *Steering wheel angle (SD):* The standard deviation of the steering wheel angle.

- *Longitudinal speed (mean and SD):* The average time the participants went five miles or more per hour over the speed limit.

The study utilized one *driver behavior variables*. This variable was the Times Eyes off the Road [TEOR], which was the average amount of time the participants were visually interacting with the mobile interface instead of focusing on driving. Visual interaction time and cognitive load were selected in order to measure the visual and cognitive distractions, respectively.

6.3.4. Participants

Seventy participants were recruited for this study, but only 60 (26 male, 34 female) participants completed the entire study. The remaining 10 individuals could not complete the complete study as they experienced motion sickness caused by the driving simulation. The participants ranged in age from 18 to 57 (M = 27; SD = 7.52) (figure 29), were native English speakers and were frequent news consumers. All of the participants had experience with touchscreen mobile devices and none of the participants had hearing or cognitive impairments. The participants were tested for cognitive impairments using the Mini-Mental State Exam (Folstein, Folstein, & McHugh, 1975) and they all scored above four out of the total score of six (Anthony, LeResche, Niaz, Von Korff, & Folstein, 1982). Twenty-four participants did not have any visual impairments, while 22 wore glasses and 14 wore contact lenses at the time of the study. All of the participants had a minimum of two years of driving experience in the U.S. and 45 of the participants drove on a daily basis. None of the participants had a history of motion sickness and they did not have a prior experience of using ANFORA*Drive* or Umano. For 120 minutes of participation, each participant received a $20 Amazon gift card.

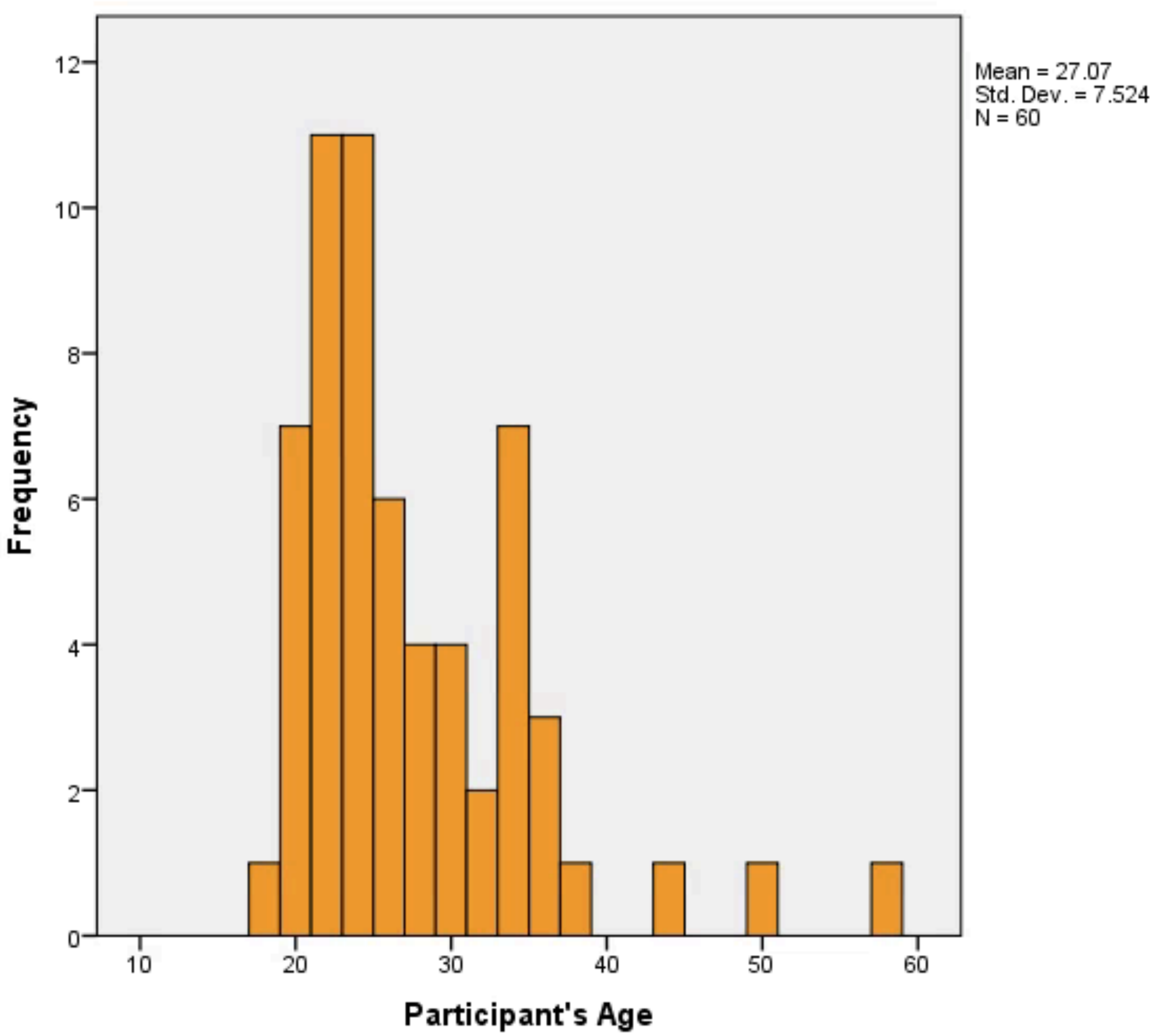


Figure 29. Histogram of participants' age range.

The participants were asked to answer a questionnaire related to their engagement in distracting behaviors while driving (Feng, Marulanda, & Donmez, 2014) before the study was conducted. All of the participants reported that they engaged in some distracting behaviors while driving, such as holding phone conversations, manually interacting with a phone, continually checking roadside accident scenes, daydreaming, reading roadside advertisements, chatting with passengers in their cars and adjusting the settings of the in-vehicle technology (Figure 29).

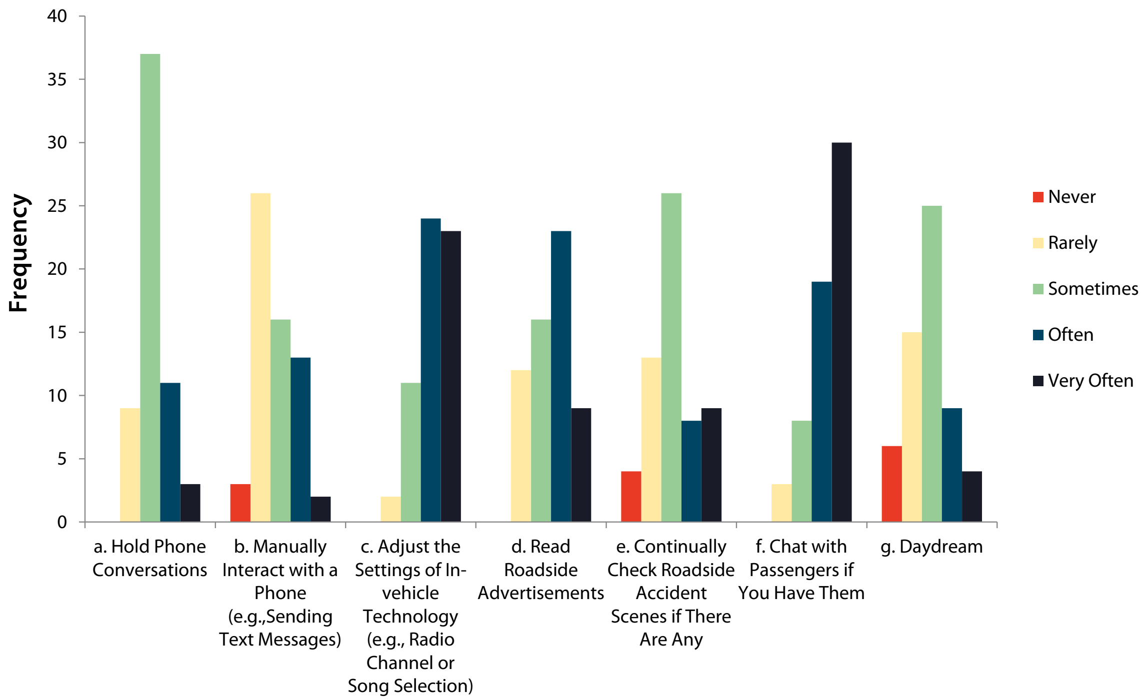


Figure 29. The results of a distraction engagement questionnaire taken by the participants prior to the study.

6.4. Procedure

The participants engaged in sessions that consisted of four parts (two hours): (1) warm up; (2) training; (3) a three-stage task session, which consisted of a session in which they used ANFORA*Drive,* a session in which they used Umano and a session in which they did not use either of the applications (i.e., No Device); and (4) completion of the simulator sickness, usability and cognitive load questionnaires as well as a post-task interview (Figure 30).

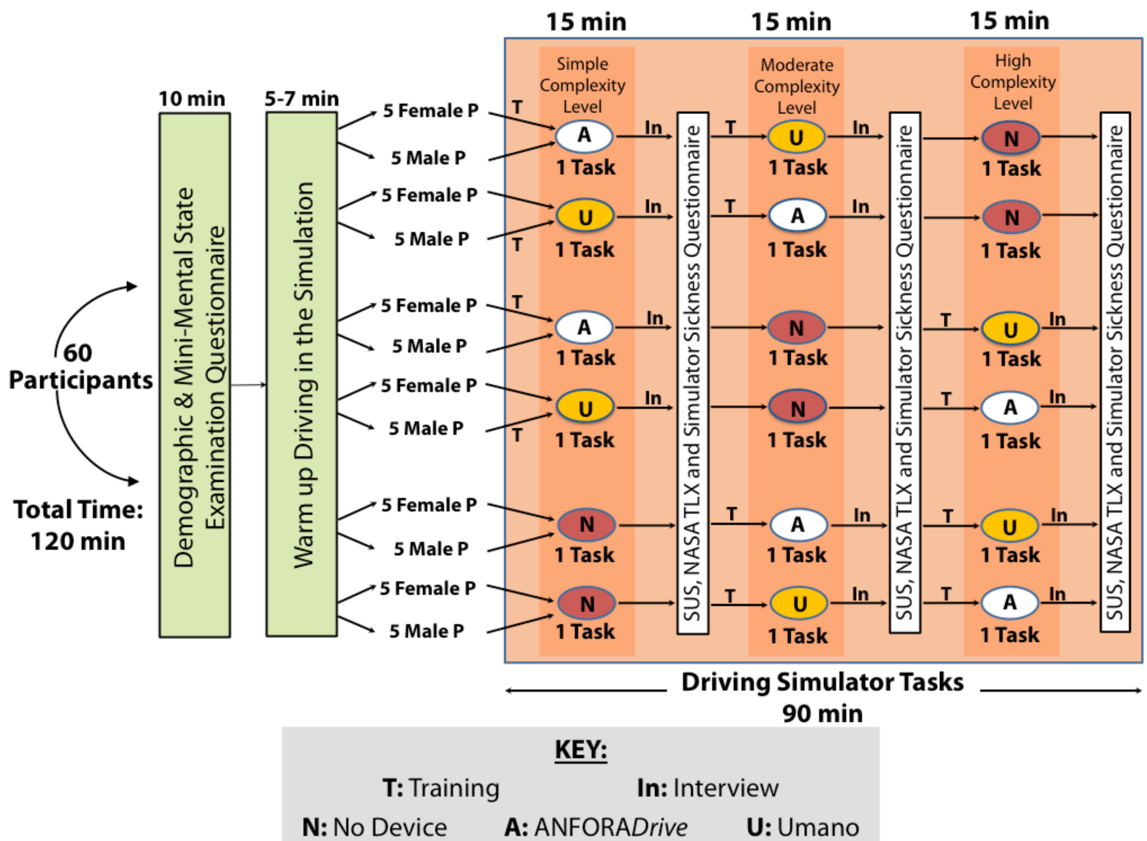


Figure 30. The within-subject design for the comparative evaluation of the different aural applications (N = 60).

6.4.1. Warm up

Each participant drove one warm-up scenario for 5-7 minutes to get familiarized with the driving simulator. The warm-up scenario took place in a residential neighborhood, similar to the low driving complexity scenario. The researcher pointed out the speedometer on the screen (Figure 28a) and the general controls in the car before starting the warm up session.

6.4.2. Training

In order to mitigate the *learning effort* in regard to remembering the navigation buttons or voice commands, I performed 15-minute training sessions with the participants prior to having them use *ANFORADrive* and Umano. The purpose of the training was to allow all of the participants to try out both of the applications and gain a common threshold of familiarity in regard to how to work them. For example, in ANFORA*Drive*, the participants were trained on how to use the application using the button and voice commands. They were told to click on the steering wheel button before using any of the voice commands in order to initiate that feature. Then, they practiced for five minutes with the ANFORA*Drive* voice commands by clicking on the steering wheel button. Finally, they were asked to memorize the eight categories available and repeat them for the researcher before starting the actual task. In the Umano training, the participants were trained to use only the part of the application that was relevant to the purpose of this study (i.e., the channel section). They were trained on how to choose their channels, add channels to their playlists, select stories and skip to the next story.

6.4.3. Task Sessions and Post-task Surveys

The participants engaged in three driving scenario complexity stages: low, moderate and high. The order of the stages remained the same and the participants always began with the low traffic, neighborhood streets before progressing the higher traffic stages, such as in the city and on the freeway (Odenheimer et al., 1994). Within each driving complexity stage, the participants went through alternative aural application exposure: no aural application/no device (N), ANFORA*Drive* (A) and Umano (U). The order of the aural application exposure was systematically counterbalanced across all of the participants in order to minimize the learning effect. Overall, the participants executed three tasks (Figure 30):

a) One task (approx. 15 minutes) for the No Device condition in which the participants drove in the low, moderate or high complexity stage without using any applications; and

b) One task (approx. 15 minutes) for the ANFORA*Drive* condition in the low, moderate or high complexity stage; and

c) One task (approx. 15 minutes) for the Umano condition in the low, moderate or high complexity stage.

The structure of each task was the same across the ANFORA*Drive* and Umano conditions (Figure 31). Each participant initially had to drive for two minutes without using the application. Then, the researcher would instructed him to begin listening to the news. These two minutes of driving were designed to help the participant become familiar with that particular driving scenario complexity. Once the participant began to listen to the news, he could listen to the news, but was only to interact with the application when prompted (e.g., change the story or category) during the next eight minutes. During these eight minutes, the participant would be prompted four times to change the story or category. One (for the low and moderate complexity) or two (for the high complexity) of the navigation prompts would be followed by maneuver prompts five seconds later. This design would enable the researcher to measure the participant's driving performance by calculating the response time to the instructed maneuver. At the end of the ten minutes, the participant would hear a prompt that he could interact with the application whenever he wanted until the end of the task (i.e., exploration time). During these five minutes, the researcher was able to examine how the participant used the aural flows in ANFORA*Drive* and whether he preferred to listen to the summary, full story or both. The researchers could also see whether the participant was moving between the categories within the application.

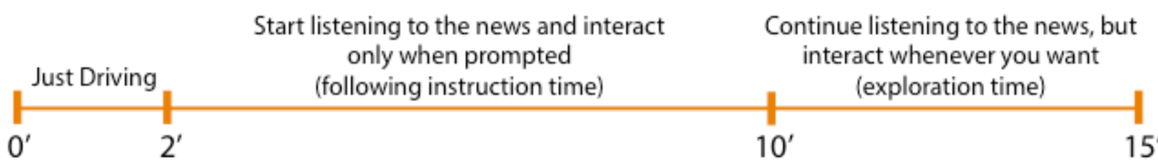


Figure 31. ANFORA*Drive* and Umano task designs during the 15 minutes of driving.

For example, the ANFORA*Drive* task was as follows:

> You have 15-minutes to drive as you would normally do and listen to the playlist of news stories using the ANFORA*Drive* app. In the first two minutes, you are to drive without using the app. Once I prompt you, you can begin listening to ANFORA*Drive* by selecting any category of interest. Once you begin listening to the news, for the next eight minutes, please don't do anything until I prompt you to change the news story or category [played the prompts for the participants to become familiar with them]. After eight minutes of listening to the playlist, I will prompt you to listen and interact as you would normally would for the remaining five minutes. I will stop you at the end of 15 minutes. You may start driving now for two minutes.

While using ANFORA*Drive* or Umano, the participants were instructed to keep the phone below the radio when not interacting visually with it (Figure 32a). The participants could interact with Umano using only button commands, but they could interact with

ANFORA*Drive* using either button or voice commands. If the participants used the voice commands, they had to click on the steering wheel button and then say the voice command. Once the participants used the voice command, the researcher repeated it as a way of giving them a feedback and controlling the participants' devices using the Wizard-of-Oz approach (explained in Chapter 5). We randomly generated voice recognition errors for participants so that they would have a natural experience (as described below).

(a) **(b)**

Figure 32. (a) A participant is using ANFORA*Drive* and clicking on the steering wheel button to initiate the voice command. (b) A participant is visually interacting with Umano.

After each of the three stages in which the participants used ANFORA*Drive* or Umano, they rated their motion sickness, system usability and their cognitive efforts using the simulator sickness questionnaire (SSQ) (Kennedy, Lane, Berbaum, & Lilienthal, 1993), SUS (Brooke, 1996) and NASA-TLX (Hart & Staveland, 1988), respectively. The participants also answered interview questions related to each of the applications. After the no device condition, the participants rated only their motion sickness and cognitive efforts. They did not rate system usability. The participants also did not have any interview questions after the no device condition (See Appendix G for the introductory script, training, tasks, surveys and interview questions).

6.4.3.1. Generating random errors during the voice interaction with ANFORADrive

Modern voice recognition systems (such as Apple Siri™) are far from perfect and errors are common. Therefore, in order to improve the external validity of the study on the Wizard–of–Oz ANFORA*Drive* prototype, we devised a strategy to include a random recognition error when a voice command is used. We leveraged the information in (Fong & Frank, 1992) that used a 3% voice recognition error in the context of testing a new pen/voice system as a future portable device. Based on this prior work, we included a 3% random voice recognition error for the ANFORA*Drive* prototype across all instances of system activation expected in the study. For example, there were a total of 60 participants in this study and each participant used a minimum of 10 voice commands to interact with the system. Therefore, there was a minimum of 600 voice commands across all of the participants. As such, for 18 of the 600 voice commands (3%), a recognition error was randomly triggered.

We introduced two types of recognition errors. The first type was *inaccurate* recognition, which was caused when the system did not recognize the actual command voiced by the user and, as such, provided an incorrect response. For example, the user would say "next" and the system recognizes it as "sports." The second type was *missed* recognition, which occurred when the system missed the command and provided the user with a missed recognition. For example, the user would say "technology" and the system would respond, "I am not sure what you just said" (Similar to Apple Siri™ response).

For the purpose of this study, we randomly generated 18 numbers between 1 and 600 and randomly assigned the type of recognition error for each of these numbers. The instrument generated for the research was a simple table that indicated when and what type of error (i.e., inaccurate or missed recognition) must be triggered. This table allowed the researcher to keep track of the number of voice commands said and, at which point, the error must be activated. For more information, please refer to Appendix H.

6.5. Analysis

We analyzed the collected data for each of the three driving complexity scenarios (i.e., low, moderate and high) separately using SPSS. We used the aural application (no device vs. ANFORA*Drive* and Umano) as the between-subject factor from the analysis perspective. The outcome variables were compared: perceived distraction, overall safety, user satisfaction, system usability, cognitive workload, driving performance and driving behavior. For the quantitative data, an independent t-test was used to analyze perceived distraction, overall safety, user satisfaction and system usability of the two aural applications (i.e., ANFORA*Drive* and Umano). A Univariate Analysis of Variance (ANOVA) was used to analyze the cognitive workload, driving performance and driver behavior variables of the two aural applications vs. the no device condition.

We did use two-way ANOVA to look into the interaction of participans' gender, age range and number of times they drove in a week with the aural application they used on the driving performance under each different driving scenario complexities. We did not use repeated-measure ANOVA because each participant did not go through nine different conditions (3 aural applications * 3 driving complexity scenarios). This decision in the experimental design was because having nine different conditions with each of them lasting for 15-minute driving in a simulation would cause fatigue. We also did not use mixed ANOVA because each participant used a different aural application under a different driving complexity. That means both aural applications and driving scenario complexities were within subject factors for each participant.

Three researchers watched the videos recorded by the three cameras in order to measure the TEOR, voice command and flow usage. System usability was reported using the SUS questionnaire and perceived cognitive workload was calculated using the NASA-TLX on the scale of 0 to 100. For the qualitative analysis of the interviews, we extracted the recurrent themes and grouped the comments by type. The emerging issues highlighted user preference for the interaction paradigms as well as the difficulties faced while using ANFORA*Drive* or Umano.

6.6. Results

6.6.1. Self-reported Cognitive Workload

6.6.1.1. Low Complexity

The users' cognitive efforts, based on the NASA-TLX questionnaire, in the three conditions are compared in Figure 33. A significant effect of the aural applications existed on the self-reported cognitive efforts for the three conditions ($F_{(2,57)} = 17.075$, $p < .001$, $\eta^2 = .375$). Games-Howell was used for the post-hoc comparisons because the assumption of homogeneity of variances was violated. Compared to the no device condition (*M* = 19.71%, *SE* = 3.285), we observed significantly higher cognitive effort for Umano (*M* = 45.04%, *SE* = 3.285) ($p < .001$), but not for ANFORA*Drive* (*M* = 23.92%, *SE* = 3.285) (p = .545). Compared to ANFORA*Drive*, we observed significantly ($p < .01$) higher cognitive efforts for Umano.

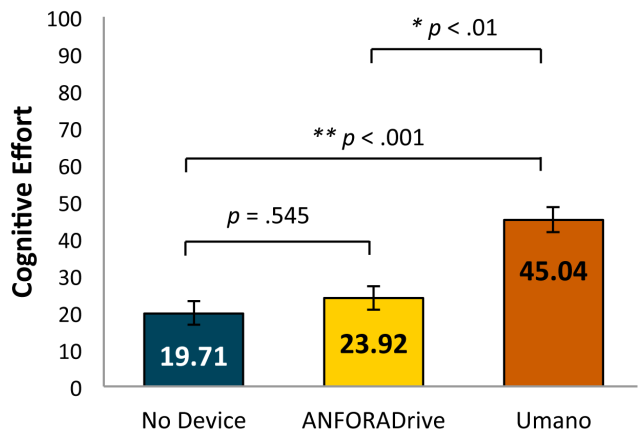


Figure 33. Compared to the no device condition, Umano significantly increased cognitive effort, but ANFORA*Drive* did not add additional cognitive effort in the *low* driving complexity scenario.

6.6.1.2. Moderate Complexity

Figure 34 shows that a significant effect of aural application existed on cognitive effort for the three conditions ($F_{(2,57)} = 6.608$, $p < .01$, $\eta^2 = .188$). The post-hoc comparisons using the Tukey test indicated that the cognitive effort for Umano (*M* = 49.54%, *SE* = 4.274) was significantly higher than for ANFORA*Drive* (*M* = 28.92%, *SE* = 4.274) ($p < .05$) and the no device condition (*M* = 32.67%, *SE* = 4.274) ($p < .05$). However, ANFORA*Drive* did not significantly differ from the no device condition. Tukey test was used for the post-hoc comparisons because the assumption of homogenety of variances was not violated and the sample sizes were equal.

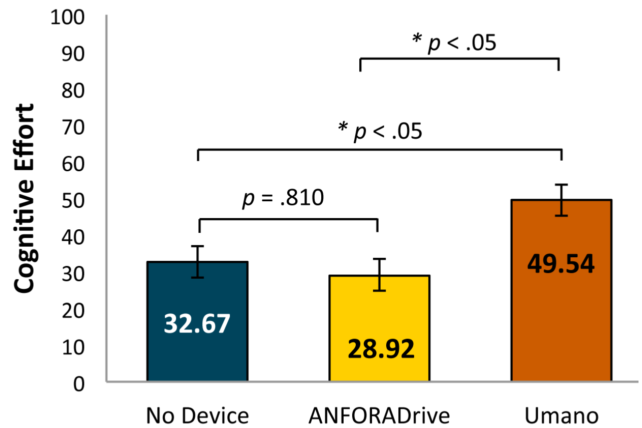


Figure 34. Compared to the no device condition, Umano significantly increased cognitive effort, but ANFORA*Drive* did not add additional cognitive effort in the *moderate* driving complexity scenario.

6.6.1.3. High Complexity

A significant effect of aural application existed on cognitive effort for the three conditions ($F_{(2,57)}$ = 6.539, $p < .01$, $\eta^2 = .187$) (Figure 35). Games-Howell was used for the post-hoc comparisons. Compared to ANFORA*Drive* (*M* = 25.29%, *SE* = 4.319), we observed significantly higher cognitive effort for Umano (*M* = 46.83%, *SE* = 4.319) ($p < .05$). The no device condition (*M* = 31.83%, *SE* = 4.319) did not significantly differ from both ANFORA*Drive* or Umano.

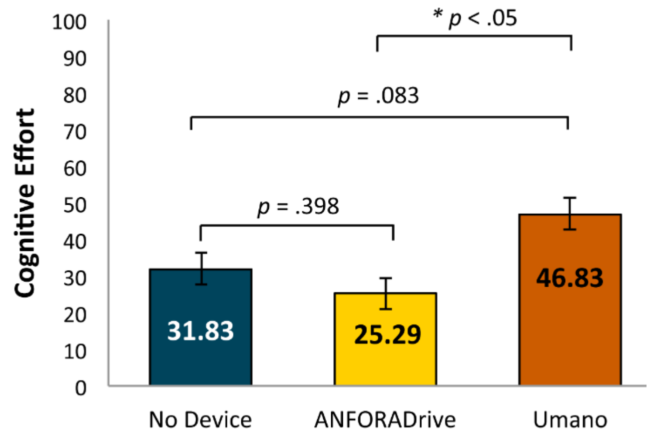


Figure 35. Compared to the no device condition, Umano increased cognitive effort, while ANFORA*Drive* decreases it, but not significantly, in the *high* driving complexity scenario.

6.6.2. Self-reported System Usability, Distraction, Overall Safety and User Satisfaction

6.6.2.1. Low Complexity

Based on the SUS questionnaire, the system usability of ANFORA*Drive* (*M* = 81.00%, *SE* = 3.23) was significantly better than the system usability of Umano (*M* = 64.13%, *SE* = 4.18) ($t(38) = 3.198$, $p < .01$) (Figure 36a). The participants were asked to rate how distracted they were while driving and using ANFORA*Drive* or Umano on a scale of 1 to 100 (1 = *very low* and 100 = *very high*). We found that the participants reported being significantly more distracted when they used Umano (*M* = 71.00, *SE* = 4.51) than when they used ANFORA*Drive* (*M* = 32.75, *SE* = 4.10) ($t(38) = 6.279$, $p < .001$) (Figure 36b). The participants were asked to rate their level of distraction on an additional semantic-differential scale of 1 to 7 (1 = *not distracted* and 7 = *very distracted*). Again, the participants felt significantly more distracted when they used Umano (M = 4.95, SE = .31) than when they used ANFORA*Drive* (M = 3.15, SE = .31) ($t(38) = 4.093$, $p < .001$) (Figure 37).

The participants were not asked to rate their distraction for the no device condition. However, the self-reported cognitive workload and distraction for both ANFORA*Drive* ($r(19) = .682$, $p < .01$) and Umano ($r(19) = .599$, $p < .01$) were correlated. Therefore, the self-reported distraction for the no device context could be correlated with self-reported cognitive workload. The self-reported distraction for Umano was significantly higher than the no device condition, but ANFORA*Drive* was not significantly higher.

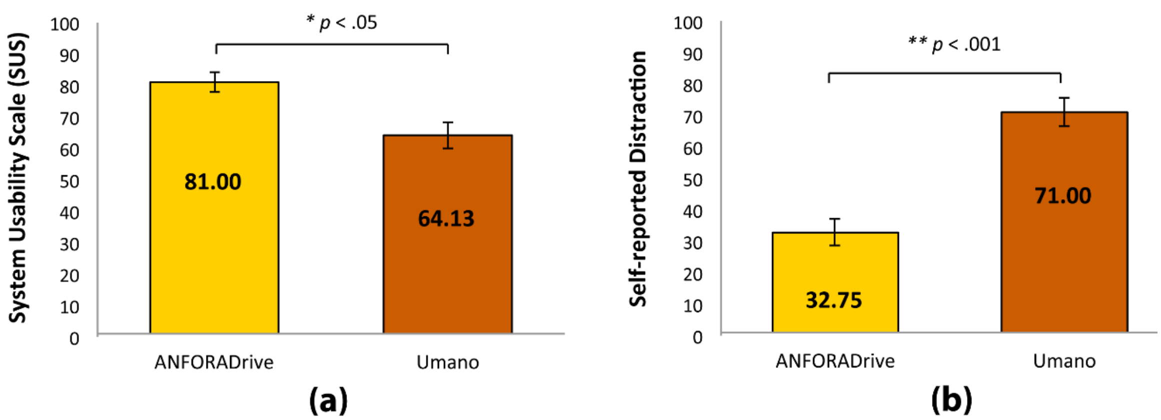


Figure 36. ANFORA*Drive* significantly (a) has a better system usability than Umano and (b) reduces self-reported distraction by 38.25% when compared to Umano in the *low* driving complexity scenario.

The participants rated their experiences (e.g., overall safety, satisfaction, difficulty, pleasant, engagement and enjoyment) with both applications using an additional semantic-differential questionnaire on a scale of 1 to 7 (for example, 1 = *difficult* and 7 = *simple*). Figure 37 shows that the participants felt that using ANFORA*Drive* (M = 5.45, SE = .27) was significantly safer than using Umano (M = 3.00, SE = .24) ($t(38) = 6.826$,

$p < .001$) while driving. They were not asked to rate their feeling of safety for the no device condition, but the self-reported distraction and safety for both ANFORA*Drive* ($r(19) = .482$, $p < .05$) and Umano ($r(19) = .849$, $p < .01$) were correlated. Therefore, the self-reported distraction for the no device context could be correlated with self-reported safety. Self-reported safety for Umano was significantly higher compared to the no device condition, while ANFORA*Drive* was not significantly higher.

Additionally, the participants rated their experiences using ANFORA*Drive* ($M = 5.45$, $SE = .34$) as more satisfactory than using Umano ($M = 4.35$, $SE = .28$) ($t(38) = 2.468$, $p < .05$). They also found ANFORA*Drive* ($M = 6.05$, $SE = .18$) to be simpler to use than Umano ($M = 4.80$, $SE = .34$) ($t(38) = 3.195$, $p < .05$). All of these reported differences were statistically significant. In general, the participants reported that ANFORA*Drive* was slightly more enjoyable, more pleasing and easier to understand than Umano. However, the participants found Umano to be slightly more engaging than ANFORA*Drive* (Figure 37).

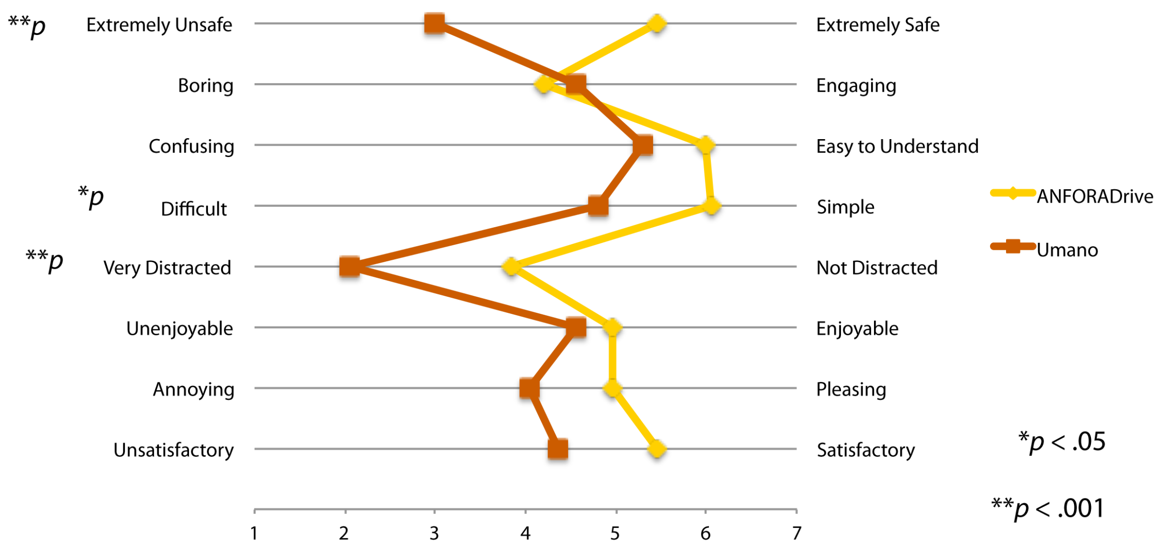


Figure 37. In the *low* driving complexity scenario, ANFORA*Drive* rated significantly safer, simpler to use and more satisfactory than Umano.

6.6.2.2. Moderate Complexity

Based on the SUS questionnaire, the system usability of ANFORA*Drive* (M = 78.63%, SE = 2.89) was significantly better than the system usability of Umano (M = 58.88%, SE = 3.97) ($t(38)$ = 4.019, p < .001) (Figure 38a). We found that the participants were significantly more distracted when they used Umano (M = 62.25, SE = 5.82) than when they used ANFORA*Drive* (M = 29.50, SE = 5.52) ($t(38)$ = 4.081, p < .001) (Figure 38b). When using the semantic-differential scale, the participants again felt significantly more distracted when they used Umano (M = 5.40, SE = .28) than when they used ANFORA*Drive* (M = 3.10, SE = .40) ($t(38)$ = 4.71, p < .001) (Figure 39).

Self-reported cognitive workload and distraction for both ANFORA*Drive* ($r(19)$ = .597, p < .01) and Umano ($r(19)$ = .799, p < .01) were correlated. Therefore, self-reported distraction for no device could be correlated with self-reported cognitive workload. Self-reported distraction for Umano was significantly higher than the no device condition, while it was not significantly higher for the ANFORA*Drive* condition.

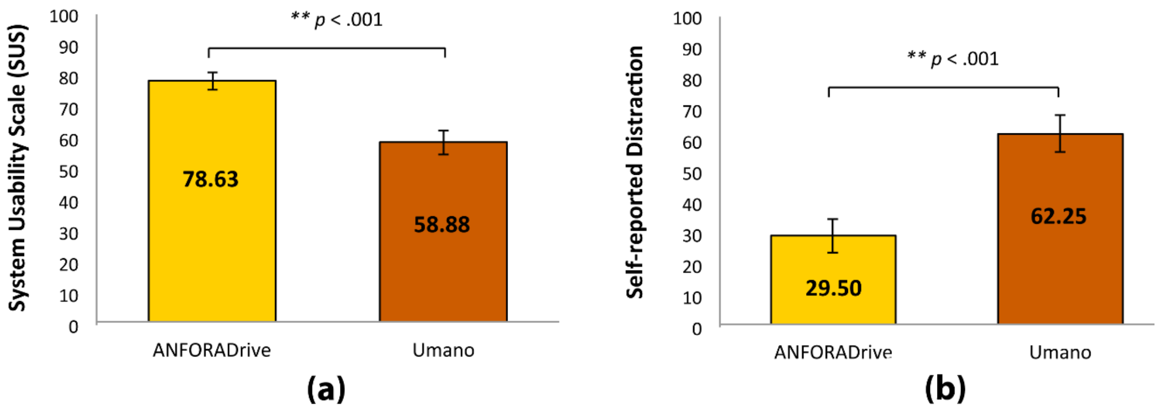


Figure 38. ANFORA*Drive* significantly (a) had better system usability than Umano and (b) reduced self-reported distraction by 32.75% when compared to Umano in the *moderate* driving complexity scenario.

Figure 39 shows that the participants felt that using ANFORA*Drive* (M = 4.80, SE = .32) was significantly safer than using Umano (M = 2.70, SE = .30) ($t(38) = 4.778$, $p < .001$) while driving. They were not asked to rate their feeling of safety for the no device condition; however, self-reported cognitive workload and safety for both ANFORA*Drive* ($r(19) = .520$, $p < .05$) and Umano ($r(19) = .696$, $p < .01$) were correlated. As such, self-reported cognitive workload for the no device condition could be correlated with self-reported safety. Therefore, self-reported safety for Umano was significantly higher than the no device condition, while it was not significantly higher for the ANFORA*Drive* condition.

Additionally, the participants rated their experiences using ANFORA*Drive* (*M* = 5.55, *SE* = .21) as being more satisfactory than using Umano (*M* = 3.80, *SE* = .39) (*t*(38) = 3.909, *p* < .001). They found ANFORA*Drive* (*M* = 5.60, *SE* = .35) simpler to use than Umano (*M* = 4.15, *SE* = .36) (*t*(38) = 2.865, *p* < .05) and easier to understand than Umano ((*M* = 5.80, *SE* = .35) vs. (*M* = 4.75, *SE* = .30) (*t*(38) = 2.275, *p* < .05)). All of these reported differences were statistically significant. In general, the participants reported that ANFORA*Drive* was slightly more enjoyable, pleasing and engaging than Umano (Figure 39).

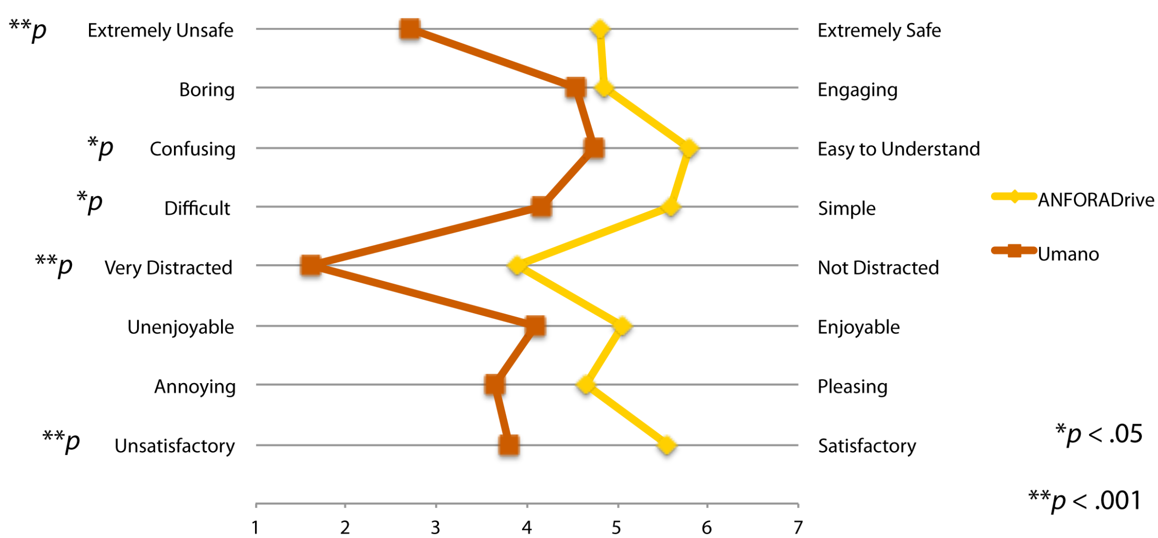


Figure 39. In the *moderate* driving complexity scenario, ANFORA*Drive* was rated significantly safer, simpler to use, easier to understand and more satisfactory than Umano.

6.6.2.3. High Complexity

Based on the SUS questionnaire, the system usability of ANFORA*Drive* (*M* = 77.88%, *SE* = 3.13) was significantly better than the system usability of Umano (*M* = 62.13%, *SE* = 5.99) (*t*(38) = 2.330, *p* < .05) (Figure 40a). We found that the participants were significantly more distracted when they used Umano (*M* = 69.25, *SE* = 5.17) than when they used ANFORA*Drive* (*M* = 18.85, *SE* = 2.68) ($t(38) = 8.655$, $p < .001$) (Figure 40b). Using the semantic-differential questionnaire, the participants indicated that they felt significantly more distracted when they used Umano (*M* = 5.30, *SE* = .34) than when they used ANFORA*Drive* (*M* = 2.25, *SE* = .25) ($t(38) = 7.213$, $p < .001$) (Figure 41).

The participants were not asked to rate their distraction for the no device condition; however, self-reported cognitive workload and distraction for both ANFORA*Drive* ($r(19) = .528$, $p < .05$) and Umano ($r(19) = .682$, $p < .01$) were correlated. As such, self-reported distraction for the no device condition could be correlated with self-reported cognitive workload. Therefore, compared to the no device condition, self-reported distraction for Umano and ANFORA*Drive* were higher, but not significantly.

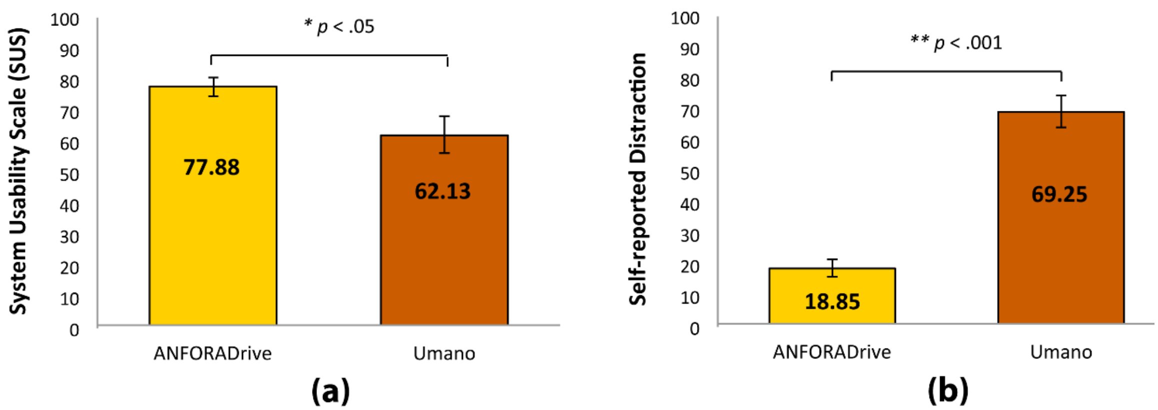


Figure 40. ANFORA*Drive* significantly (a) had a better system usability than Umano and (b) reduced self-reported distraction by 50.40% compared to Umano in the *high* driving complexity scenario.

Figure 41 shows that the participants felt that using ANFORA*Drive* (*M* = 5.60, *SE* = .23) was significantly safer than using Umano (*M* = 2.65, *SE* = .28) ($t(38) = 8.025$, $p < .001$)

while driving. They were not asked to rate their feelings of safety for the no device condition; however, self-reported distraction and safety for both ANFORA*Drive* and Umano were correlated ($r(39) = .747$, $p < .01$). As such, self-reported distraction for the no device condition could be correlated with self-reported safety. Therefore, compared to the no device condition, self-reported safety for Umano and ANFORA*Drive* were higher, but not significantly.

The participants rated their experiences using ANFORA*Drive* ($M = 5.15$, $SE = .31$) as more satisfactory than using Umano ($M = 4.10$, $SE = .40$) ($t(38) = 2.064$, $p < .05$). They also found ANFORA*Drive* ($M = 6.10$, $SE = .18$) simpler than Umano ($M = 4.65$, $SE = .39$) ($t(38) = 3.370$, $p < .05$), but they found Umano ($M = 5.55$, $SE = .30$) more engaging than ANFORA*Drive* ($M = 4.50$, $SE = .30$) ($t(38) = 2.447$, $p < .05$). All of these reported differences were statistically significant. In general, the participants reported that ANFORA*Drive* was slightly more enjoyable, more pleasing and easier to understand than Umano (Figure 41).

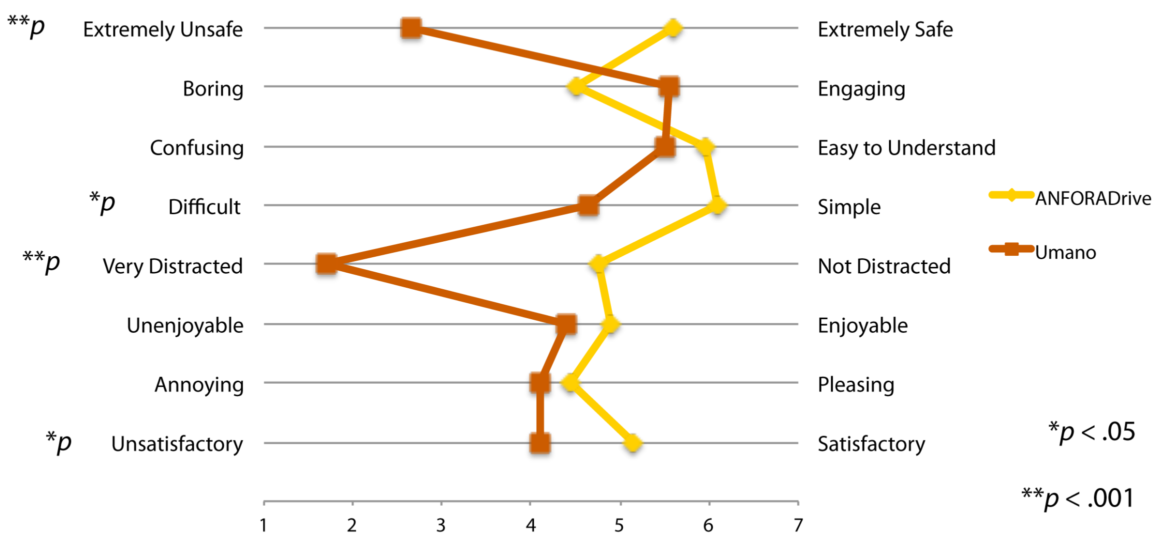


Figure 41. In the *high* driving complexity scenario, ANFORA*Drive* was rated as significantly safer, simpler to use and more satisfactory than Umano, but Umano was rated as significantly more engaging than ANFORA*Drive*.

6.6.3. Driving Performance

6.6.3.1. Low Complexity

A significant effect of the aural application did not exist on the number of collisions ($F_{(2,57)}$ = 1.000, p = .374, η^2 = .034), lane departures ($F_{(2,57)}$ = 1.440, p = .245, η^2 = .048), lane positions ($F_{(2,57)}$ = .006, p = .994, η^2 = .0), steering wheel angles ($F_{(2,57)}$ = .879, p = .421, η^2 = .03), longitudinal speeds ($F_{(2,57)}$ = .720, p = .491, η^2 = .025) or response times ($F_{(2,57)}$ = .107, p = .899, η^2 = .004) for the three conditions. Additionally, participants' gender, age and number of times they drive in a week did not have a significant main effect on the number of lane departures and response time.

6.6.3.2. Moderate Complexity

A significant effect of the aural application did not exist on the number of collisions ($F_{(2,57)}$ = .199, p = .820, η^2 = .007), lane departures ($F_{(2,57)}$ = .035, p = .966, η^2 = .001), lane positions ($F_{(2,57)}$ = 2.236, p = .116, η^2 = .069), steering wheel angles ($F_{(2,57)}$ = 1.907, p = .158, η^2 = .064), longitudinal speeds ($F_{(2,57)}$ = 2.191, p = .121, η^2 = .073) or response times ($F_{(2,57)}$ = 3.147, p = .051, η^2 = .099) for the three conditions. Moreover, participants' gender and number of times they drive in a week did not have a significant main effect on the number of lane departures and response time. Participant's age also did not have a significant main effect on the number of lane departure. Participants' age, however, did have a significant main effect on the response time ($F_{(14,45)}$ = 2.875, $p < .05$, η^2 = .242).

6.6.3.3. High Complexity

A significant effect of the aural application did not exist on the number of collisions ($F_{(2,57)}$ = .924, p = .403, η^2 = .031), lane positions ($F_{(2,57)}$ = .663, p = .519, η^2 = .023), steering wheel angles ($F_{(2,57)}$ = 1.258, p = .292, η^2 = .034), longitudinal speeds ($F_{(2,57)}$ = 2.682, p = .077, η^2 = .086) or response times ($F_{(2,57)}$ = 2.977, p = .059, η^2 = .095) for the three conditions. However, a significant effect of the aural application exist on lane departures ($F_{(2,57)}$ = 3.707, $p < .05$, η^2 = .115) (Figure 42).

The post-hoc comparisons using the Games-Howell test indicated that the number of times the participants went out of their lanes when they used Umano (*M* = 4.40, *SE* = .722) was not significantly more than when they used ANFORA*Drive* (*M* = 1.80, *SE* = .722) (*p* = .098) or when they did not use a device (*M* = 2.25, *SE* = .722) (*p* = .198). In addition, the number of lane departures in ANFORA*Drive* did not significantly differ from the no device condition. Although the F-test (overall) was significant, the post-hoc comparison (pairwise) was not significant because the overall and the pairwise tests ask different questions and they get different answers. Moreover, this different could be due to sensitivity of ANOVA which is greater than pairwise test sensitivity. Additionally, participants' gender, age and number of times they drive in a week did not have a significant main effect on the number of lane departures and response time.

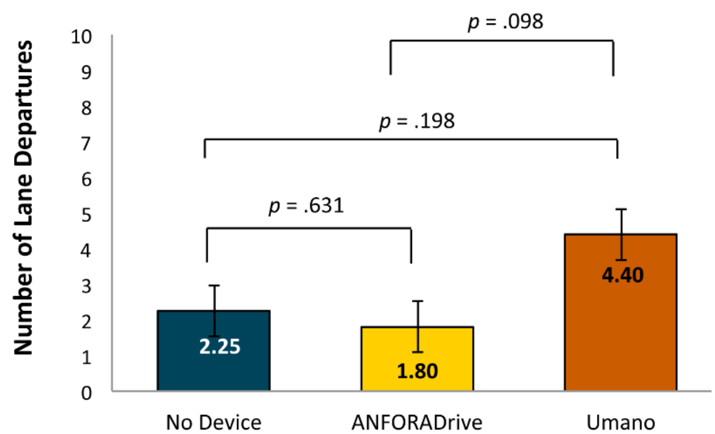


Figure 42. Although not significant, the number of lane departures increased when the participants used Umano than when the used ANFORA*Drive* or did not use any device in the *high* driving complexity scenario.

6.6.4. Driving Behaviors (TEOR)

6.6.4.1. Low Complexity

The aural application significantly affected the amount of time that the participants took their eyes off the road for the three conditions ($F_{(2,57)}$ = 196.268, p < .001, η^2 = .831). The post-hoc comparisons using the Tukey test indicated that the participants took their eyes off the road (TEOR) for a significantly longer time when they used Umano (M = 99.25

sec., *SE* = 3.963) than when using ANFORA*Drive* (*M* = 6.50 sec., *SE* = 3.963) ($p < .001$) or in the no device condition (*M* = .00 sec., *SE* = 3.963) ($p < .001$). However, ANFORA*Drive* did not significantly differ from the no device condition (Figure 43).

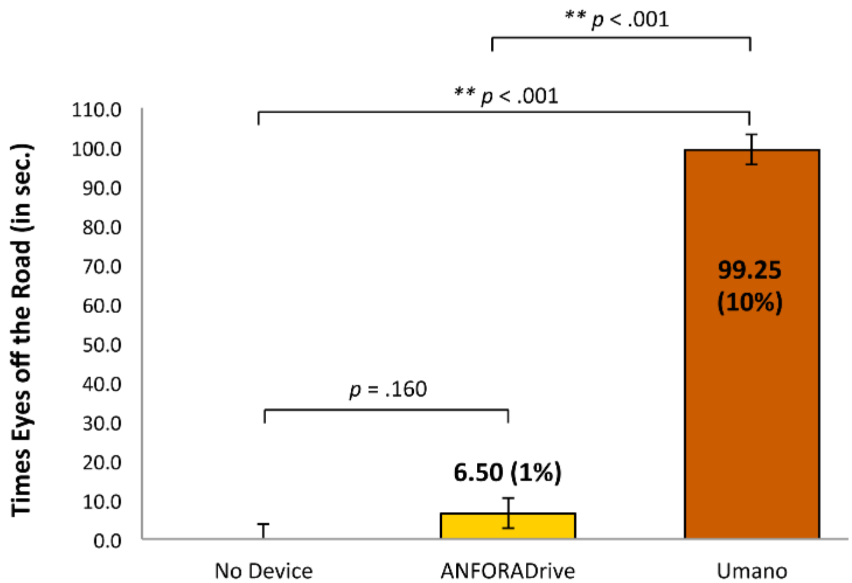


Figure 43. The TEOR for Umano was significantly higher than for ANFORA*Drive* and the no device condition in the *low* driving complexity scenario. The percentage value is the percentage of the total task time (15 minutes = 900 seconds).

6.6.4.2. Moderate Complexity

Figure 44 shows that a significant effect of the aural application on TEOR existed for the three conditions ($F_{(2,57)} = 140.322$, $p < .001$, $\eta^2 = .099$) in the moderate complexity scenario. Tukey was used for the post-hoc comparisons. The TEOR for Umano (*M* = 84.15 sec., *SE* = 3.877) was significantly higher than for ANFORA*Drive* (*M* = 10.20 sec., *SE* = 3.877) ($p < .001$) and the no device condition (*M* = .00 sec., *SE* = 3.877) ($p < .001$). However, ANFORA*Drive* did not significantly differ from the no device condition.

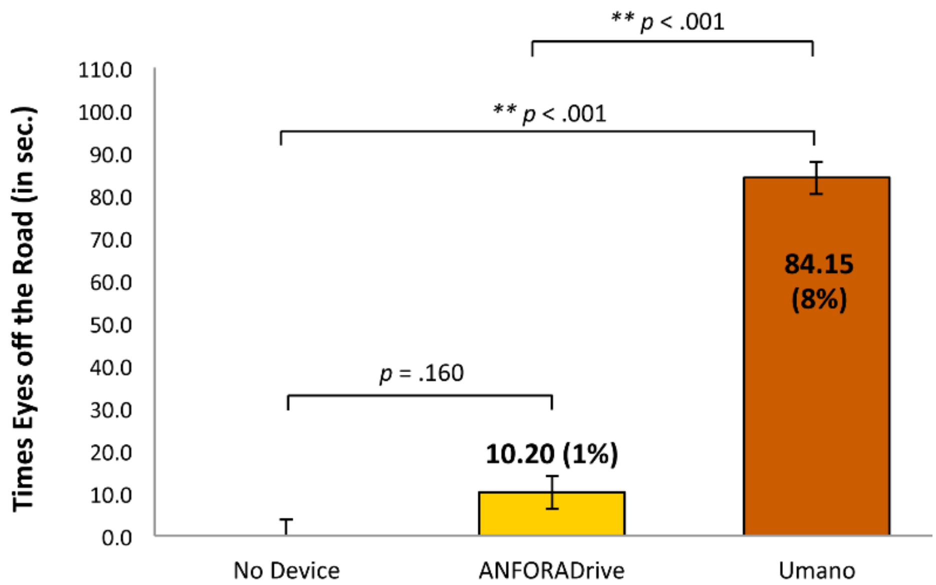


Figure 44. The TEOR for Umano was significantly higher than ANFORA*Drive* and the no device condition in the *moderate* driving complexity scenario. The percentage value is the percentage of total task time (15 minutes = 900 seconds).

6.6.4.3. High Complexity

The aural application significantly affected TEOR for the three conditions ($F_{(2,57)}$ = 105.712, p < .001, η^2 = .788). The post-hoc comparisons using the Tukey test indicated that the participants had their eyes off of the road (TEOR) when they used Umano (M = 74.90 sec., SE = 4.043) for a significantly longer time than when they used ANFORA*Drive* (M = 6.20 sec., SE = 4.043) (p < .001) and when they did not use a device (M = .00 sec., SE = 4.043) (p < .001). However, ANFORA*Drive* did not significantly differ from the no device condition (Figure 45).

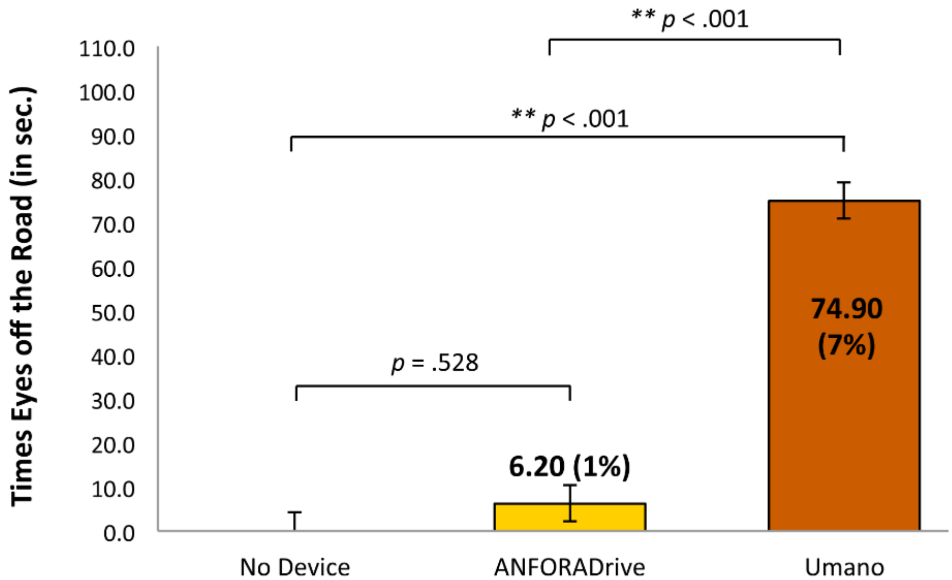


Figure 45. The TEOR for Umano was significantly higher than for ANFORA*Drive* and the no device condition in the *high* driving complexity scenario. The percentage value is the percentage of total task time (15 minutes = 900 seconds).

6.6.5. Voice Command and Aural Flow Usage

In this section, we report the data corresponding to the last five minutes of the tasks (exploration time) in which the participants used the aural application (ANFORA*Drive* or Umano). We conducted the following comparisons: how often the participants changed the news category; whether they listened to the summary, full story or both; how often they used the button vs. voice commands; and which voice commands were primarily used. We will also report on voice command usage while beginning the playlist in ANFORA*Drive*, which is considered exploratory in nature since the participants had the freedom to use any of the voice commands.

6.6.5.1. Navigating the Aural Flows

The results show that, on average, the participants changed their news categories twice when they used ANFORA*Drive* with the voice or button commands and once when they used Umano with the button commands. Overall, on average, the participants listened to three categories using ANFORA*Drive* and two categories using Umano.

Fifty-two of the 60 participants let the aural flow move through the entire summary and full story. However, six of the participants preferred the full story only. Every time they heard the title or a bit of a summary, they immediately used the "full story" command to listen to the entire story. Two of the participants preferred to only listen to the summary. Every time they finished listening to a summary, they changed either the category or moved to the next news story. Additionally, nine participants used the "related," "tell me more," "more" or "like this" commands to listen to related stories.

6.6.5.2. Input Modalities: Voice Commands vs. Button Commands

Overall, the 60 participants used 309 voice commands in ANFORA*Drive*. On average, each participant used five voice commands (M = 5.22, SD = 2.64) and zero button commands (M = .08, SD = .38) to interact with ANFORA*Drive*. The three sets of commands used most are as follows: (1) the "next/skip" command was used significantly more than all of the other commands (used 146 times; an average of three times per participant; SD = 1.48); (2) the category selection commands, such as "technology," "world" and "health," were used next most often (used 107 times; an average of two times per participant; SD = .94); and (3) the "full story" command was used to move from a story summary to a full version of the same story (used 34 times; an average of one time per participant; *SD* = .73). The percentage of voice command usage is displayed in Table 8.

Table 8. The percentage of voice commands used in decreasing order.

Voice Commands	*Percentage of Usage*
Skip, Next	47%
Full Story	11%
Health	8%
Technology	6%
U.S.	5%
Related	4%
World	4%
Sports	4%
Politics	3%
Science	3%
Economy	3%
Restart	1%
Summary	1%
Previous	1%
More	0%
Tell Me More	0%
Anything Else	0%
Like This	0%

The participants could use “start,” “what’s new?” and “recent news” to begin the default playlist (i.e., U.S. news) or they could use the name of the category (e.g., “world,” “technology,” “health”) they were interested in listening to. The results showed that 22 participants used “start” or “what’s new?” to start the default playlists, while 34 participants used one of the eight categories in which they were interested. For example, 10 participants said “U.S.”, while five used “world.” Two of the participants used the voice command “play” to begin, even though it was not an approved command. Finally, two of the participants used the button commands instead of the voice commands to begin their playlists.

6.7. Interview Results

6.7.1. ANFORA*Drive* vs. Umano in the Context of Driving

Fifty-two of the 60 participants stated that they would prefer to use ANFORA*Drive* while driving, while only 14 participants said that they would use Umano while driving in particular circumstances. For example, five participants said that they would use Umano on long trips. One participant (P46) noted, “When there is not a lot of traffic around or one stretch of road, it would be more useful than in the city switching lanes.” Other participants said that they would create a playlist beforehand or would listen to only one channel, so that they would not have to manually interact with Umano while driving. For example, one participant (P25) noted, “I just wouldn’t hold it [Umano]. I would just play it, put it in my car, drive while listening and let it go automatically.” A few other participants stated that they would manually interact with Umano only when stopped or when they were familiar with the road. For example, one participant (P37) noted, “I would wait until I was going to a stop sign or traffic light to change the channel." Another participant (P50) said, "Yes. I would use it only on drives that I know exactly where I am going since I have to look at the screen and take my eyes off the road."

Thirty-six of the 60 participants said that they would not use Umano while driving for several reasons. First, they did not like the way in which they had to interact with the application in order to change the story or channel as it did not have voice controls and the button commands were small, close to each other (i.e., pause and next button) and not sensitive enough (i.e., back button). Second, they did not like that they had to visual select the news story by reading the headlines. Third, they did not like that only the full story played and there was no way to listen to a summary or related stories. These participants suggested other contexts in which they would want to use Umano, such as while washing dishes at home, sitting and having coffee, sitting at their desks or computers, getting ready in the morning, commuting on the bus, walking or waiting for a class.

6.7.2. Voice Commands as a Preferred Interaction Modality with ANFORADrive

Forty-one of the 60 participants said that they preferred to use the voice commands to interact with ANFORA*Drive* because it was easier, safer and less distracting. For example, one of the participants (P42) noted, "Voice commands. It's just easier when you are driving just to speak than look for the button command. Less distracting."

Sixteen of the participants said that they would use a combination of the voice and button commands to interact with ANFORA*Drive* while driving for the following reasons. First, the users thought that the voice recognition system had not yet reached the point where it could operate without any errors. As such, they wanted a backup method in case of errors. Second, they wanted to use both the button and voice commands until they got used to the commands. Then, they felt that they might only use the voice commands. Third, the users preferred to use the button commands when they were at a red light, stop sign or driving on a long road; however, they preferred to use voice commands while driving on a busy street with a lot of traffic. For example, one participant (P56) noted,

> Probably combination of the two. If it were a long road, it would be ok to use button commands and take it out few times and hit the button commands, but, if it were a busy road, voice commands would be nice.

Three of the participants said that they preferred the button commands for security reasons or because they were not being able to adopt the new technology. For example, one participant (P6) noted, “I will use the button commands because of security issues in that I don’t know where my voice is being saved to.” Another participant (P31) said, “I will use button commands because I am used to it.” Another participant (P7) noted,

> I prefer the button commands, if there was a lot of traffic and a lot of stops where I could easily take it out and play with it or with Indiana’s law about texting and not driving, in the area where I was sure there is no police, in the familiar area.

6.7.3. Self-reported User Experiences with ANFORA*Drive*

Five of the 60 participants did not have anything negative to say about ANFORA*Drive.* For example, one of the participants (P15) noted, “I can’t think of anything negative.”

The positive aspects of ANFORA*Drive*, as mentioned by the participants, were categorized into different themes and are discussed below.

6.7.3.1. Easy to Learn, Use and Navigate

Thirty of the 60 participants found ANFORA*Drive* easy to learn, simple to use, and easy to navigate. For example, some of the participants noted that it was easy to learn and use the voice commands. One participant (P7) said,

> I liked how easy it was. We had a five minute little training session and, at no point, was I confused and the voice commands were simple enough. I didn't have to say a special word or memorize it. It was all natural and I could recall it.

Another user (P21) noted, "I think it added a lot of solid voice commands that were easy to know without trying very hard." Other participants said that they liked how easy it was to click on the steering wheel button before using any of the voice commands. For example, one participant (P10) noted, "ANFORA*Drive* was much easier to use while driving, especially clicking on the steering wheel, that helped a lot." Another participant (P41) said, "I liked that you could use the steering wheel button, just press the button and speak. It was a lot less distracting than looking at it on the screen."

Three of the 60 users noted that they liked that they could select what to listen to. For example, one participant (P22) said, "I liked that I could pick what I wanted to listen to. That is the only part of the radio that I don't like that, just waiting and waiting for something interesting to come up." Two of the 60 users said that although ANFORA*Drive* was easy to learn and navigate but it was not fluid and seamless going from one news to another because it was a prototype.

6.7.3.2. Hands- and Eyes-free

Fourteen of the 60 participants said that they liked that ANFORA*Drive* was hands-free, eyes-free and safe to use. For example, one of the participants (P24) noted, "I definitely like the hands-free interface with ANFORA*Drive* and the usability." Another participant (P43) said, "The good thing was that you did not need to pay attention to the screen and could focus on the road."

6.7.3.3. Educational and Informative

Four of the participants found ANFORA*Drive* both informative and educational. One participant (P19) noted, "Being a person who loves news, it [gives] me a burst of what is going on." Another participant (P52) said, "I just thought it was very educational."

6.7.4. Combining Best Features of ANFORA*Drive* and Umano into One Application to Use It While Driving

When the participants were asked to name the features that they would select from ANFORA*Drive* or Umano to combine into one application that could be used while driving, they suggested the following features of ANFORA*Drive* (listed based on the highest to the lowest number of times suggested): voice commands (43 participants), story summary (20 participants), full story (10 participants), related story (8 participants), hands-free (4 participants), navigation (2 participants), ease and simplicity to use (1 participant) and freedom of flexibility (1 participant).

The participants also suggested the following features of Umano (listed based on the highest to the lowest number of times suggested): variety of news categories, sub-categories, news sources and content (31 participants), narrators and the human voice instead of TTS audio (25 participants), interface design with colors and pictures (11 participants), transition between stories with a little music (4 participants), swiping to the next story (2 participants), smoothness and continuous flow (2 participants), car mode with bigger interface and button commands (2 participants), setting up the list of my channels (2 participants), going back 15 seconds within a news story (1 participant) and playing the story from where it was paused (1 participant).

6.7.5. Preferences for ANFORA*Drive* Features and Improvements Suggested by Users

6.7.5.1. Reading Level (Summary vs. Full Story)

Six of the 60 participants said that they liked the option of having both a summary of the story as well as the full story. Four of the participants noted that they liked the option to be able to get related news. While six of the participants said that they preferred listening to the summary by default, another six participants did not want to listen to the summary. Instead, they wanted to listen to the full story only. For example, one participant (P54) said,

> I don't know if I liked how they did the summary. I felt like the title could be a summary because the summary seemed a little long. By the time I was done with the summary, I was like, I guess it goes to the story now.

Another participant (P60) noted, "I didn't like the fact that there was a summary and then a full story. I would have liked to listen to the full story and, if I didn't like it, I could just move on."

Finally, one participant (P28) thought that the summary and full story be confusing for users to differentiate between, as such, he noted, "I can understand getting lost in summary and the full story just because it sounds the same. If there was ambient background music behind summary, but not behind full story to differentiate them [that would be useful]."

6.7.5.2. Orientation Information

Two of the participant liked the orientation information, such as the story number, category name and summary vs. full story. However, 11 of the users did not like that the story number and the category name were repeated every time they listened to a new news story. One of the participants (P7) suggested that "maybe, when you start the app, it could say the total number of stories, even that we could just cut it." Another participant (P4) noted,

> Before every story, it will tell you the category. It was kind of monotonous. I knew I was in U.S. News, so I didn't necessarily hear it [category name] after every story. I would want to know only when the category changes.

6.7.5.3. Variety of News

Six of the participants said that they liked the variety of selection provided by ANFORA*Drive*. However, 12 of the participants said that not enough news categories, sub categories and news sources were provided when compared to Umano. For example, one participant (P30) noted, "I had like if ANFORA*Drive* had more sources, such as CNN and science daily, available since these news were all from NPR." Another participant (P16) commented on not having sub-categories, "The topics were broad, like it didn't have basketball or football."

6.7.5.4. TTS Audio

Thirty-three of the 60 participants said that they did not like the TTS voice because it was robotic and monotone. The participants noted that they wanted a human voice that they could also adjust the speed of and for which they could choice different options (e.g., male and female). For example, one of the participants (P7) said, "Without a lot of inflection in the voice (i.e., monotone), you kind of zone out." Another user (P19) noted, "It would be good if there was an option to hear another type of voice." On the other hand, two participants liked the TTS voice and its pace of reading.

6.7.5.5. Voice Commands

Six of the 60 participants noted that it was easy to learn and use the voice commands. However, two of the participants said that it took a bit longer for them to learn the voice commands and remember them than learning to interact using button commands. For example, one participant (P55) noted that "I think it was easy to use the voice commands while driving." However, another participant (P54) said, "voice command was okay, but it seemed like it took longer to learn how to use it because I had to learn the voice commands and it was also just not as obvious as the Umano app."

Although one participant (P29) wanted to have more alternatives for each of the voice commands, another participant (P5) thought that we had enough alternatives for each of the voice commands. Three of the participants wanted a voice command that would take them to the beginning of the story. Two of these participants found the "restart" command confusing and thought it was to be used to go back to the beginning of a story rather than to the beginning of the playlist. For example, one of the participants (P2) noted, "I expected the "restart" command to restart the article instead of restarting the playlist. Maybe having a separate command to do both would be nice." Similarly, one of the participants (P56) commented that he could not rewind within a story. He said, "I was not able to go back a little bit in the story by like five or 10 seconds or go back to the beginning of the story."

For the related stories in ANFORA*Drive*, the participants could just listen to the list of related stories. They could not select a particular news story in the list other than selecting them manually. Two of the participants suggested having a numbered list of related stories so that they use the number as the voice command to select a specific related story.

One participant (P5) also commented that the feedback form the voice commands made his experience go smoothly. He noted, “I liked the feedback part of the system, which repeats the voice command. They have that for lot of things to make sure that it understood what I said.”

6.8. Discussion

The research question for this study was focused on discovering the impact of *voice-controlled aural flows* (i.e., ANFORA*Drive*) and an *alternative solution on the market* (i.e., Umano) on distraction and driving performance with respect to not using any device in the context of driving. To answer this question, the study was conducted in a driving simulation lab. Overall, the findings suggested that *voice-controlled aural flows* do not significantly distract drivers or worsen driving performance with respect to not using any devices. This study showed that voice-controlled aural flows belong to a low level on the distraction framework (see Chapter 2, Section 2.4.5). In addition, the aural flow usage patterns confirmed the initial design of ANFORA, which allows participants to customize content (see Chapter 3, Section 3.2.1). These findings are discussed in details in the following sections.

6.8.1. Hypotheses Revisited

6.8.1.1. Cognitive Workload

This study confirms H1.1 and H2.1: Compared to the driving only condition, ANFORA*Drive* does not increase the driver's cognitive effort, but Umano increases the driver's cognitive effort in all the three driving complexities. This study also confirms H3.1: Compared to Umano, ANFORA*Drive* reduces the driver's cognitive effort. The cognitive effort ratings of the no device condition for the low, moderate and high complexity scenarios were 19.71%, 32.67% and 31.83%, respectively (Figure 46). This result shows that the cognitive effort for the moderate driving complexity scenario was slightly higher than for the high driving complexity scenario. However, this result could have occurred due to experiencing many curvy roads in the design of the driving scenario.

The ANFORA*Drive* cognitive effort ratings for the low, moderate and high complexity scenarios were 23.92%, 28.92% and 25.29%, respectively (Figure 46), which were below 30%. However, the Umano cognitive effort ratings for the low, moderate and high complexity scenarios were 45.04%, 49.54% and 46.83%, respectively (Figure 46), which were between 45% and 50% cognitive effort. The cognitive effort for ANFORA*Drive*, Umano and the no device condition increased from the low to moderate complexity scenarios, but decreased from the moderate to high complexity scenarios. Based on previous studies (Horberry, 1998; Horberry et al., 2006; Horberry & Edquist, 2008; Justiss et al., 2006), highway and city driving belongs to the higher complexity scenario. However, using many curvy roads in the scenario design could also belong to the high complexity because it adds additional overhead to cognitive effort.

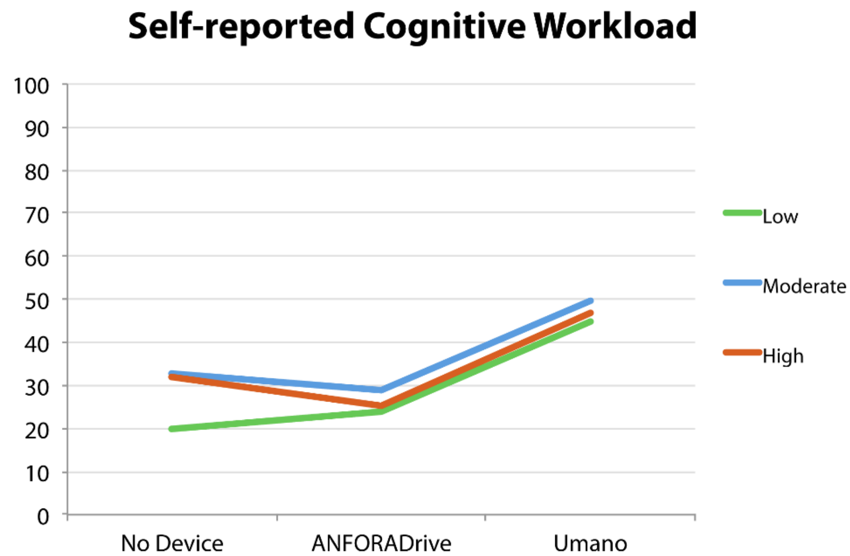


Figure 46. The ANFORA*Drive* cognitive workload was below 30% in low, moderate and high driving complexity scenarios.

6.8.1.2. Distraction and Overall Safety

Our study confirms H3.2: Compared to Umano, ANFORA*Drive* reduces driver's distraction in all the three driving complexities (low, moderate, and high). On a scale of 1 to 100, ANFORA*Drive*'s distraction ratings for the low, moderate and high complexity scenarios were 32.75, 29.50 and 18.85, respectively (Figure 47), which were below the 40% distraction level. However, the Umano distraction ratings for the low, moderate and high complexity scenarios were 71.00, 62.25 and 69.25, respectively (Figure 47), which were between the 60% to 75% distraction levels. Although driving complexity increased from the low to high complexity scenarios while using ANFORA*Drive*, the distraction level decreased from 32.75 to 18.85 (Figure 47). This result suggests that, as driving difficulty increases, ANFORA*Drive* does not add additional distraction, but reduces self-reported distraction.

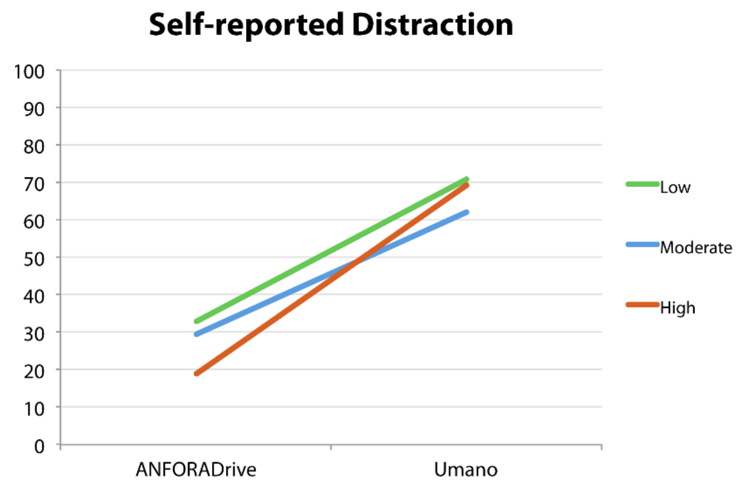


Figure 47. Self-reported distraction decreases as the driving complexity scenario increases for ANFORA*Drive*.

Our study also confirms H3.3: Compared to Umano, ANFORA*Drive* increases the overall safety in all the three conditions. On a scale of 1 to 7, ANFORA*Drive*'s overall safety ratings for the low, moderate and high complexity scenarios were 5.45, 4.80 and 5.60, respectively. However, Umano's overall safety ratings for the low, moderate and high complexity scenarios were 3.00, 2.70 and 2.65, respectively. As driving complexity increases, Umano's overall safety decreased slightly. However, ANFORA*Drive*'s overall safety increased slightly. These results were confirmed by our qualitative results in which the participants noted that using voice commands to interact with ANFORA*Drive* was safer and less distracting. Interacting with Umano was perceived as being more distracting, which is one of the reasons why the participants did not prefer to use Umano while driving (reported in our Interview Findings, Section 6.7.1).

6.8.1.3. System Usability and User Experiences

Our study shows that, compared to Umano, ANFORA*Drive* has a better system usability in all the three driving complexities. ANFORA*Drive*'s system usability ratings for the low, moderate and high complexity scenarios were 81.00%, 78.63% and 77.88%, respectively (Figure 48), which indicated an acceptable interface with a rating close to excellent (Bangor et al., 2009). However, the Umano system's usability ratings for the low, moderate and high complexity scenarios were 64.13%, 58.88% and 62.13%, respectively (Figure 48), which indicated marginal acceptability of interface (Bangor et al., 2009).

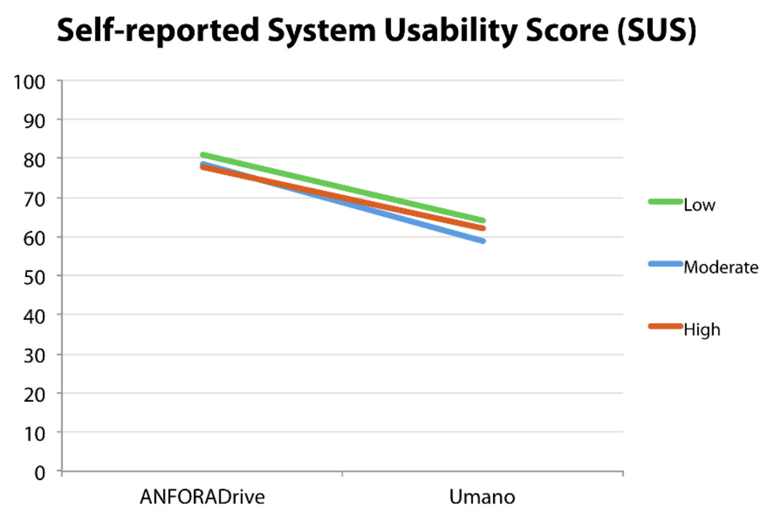


Figure 48. ANFORA*Drive*'s usability score was above 75% in the low, moderate and high driving complexity scenarios, which was close to an excellent rating.

This study also showed that compared to Umano, ANFORA*Drive* is simpler to use and provides the user with a satisfactory experience while using it (confirming H3.6). For both the low and high complexity scenarios, ANFORA*Drive* is less engaging than

Umano. However, this result is only significant in the high complexity scenario because of two reasons. First, the participants could listen to only a small set (eight) of news categories and the TTS audio using ANFORA*Drive*. Second, the high complexity scenario was always the last task and the participants were already fatigued. Therefore, other elements (e.g., news varieties and TTS) could affect them more than in the other scenarios.

Overall, the participants had a better user experience when they used ANFORA*Drive* than when they used Umano in all three of the driving complexities. Our interview findings also confirmed this result because our participants clearly noted that they found ANFORA*Drive* easy to learn, use and navigate.

6.8.1.4. Driving Performance

The two main outcome measurements for driving performance in our study are the number of lane departures and response time. Participants using ANFORA*Drive*, Umano and no device yielded a similar response time in all three of the driving complexities (Figure 49). However, an increasing trend occurred from no device to ANFORA*Drive* to Umano for the moderate and high complexity scenarios (Figure 50). In addition, as the cognitive load increases, the response time also increases. Testing response time with additional participants could confirm whether the aural application has a significant effect on response time.

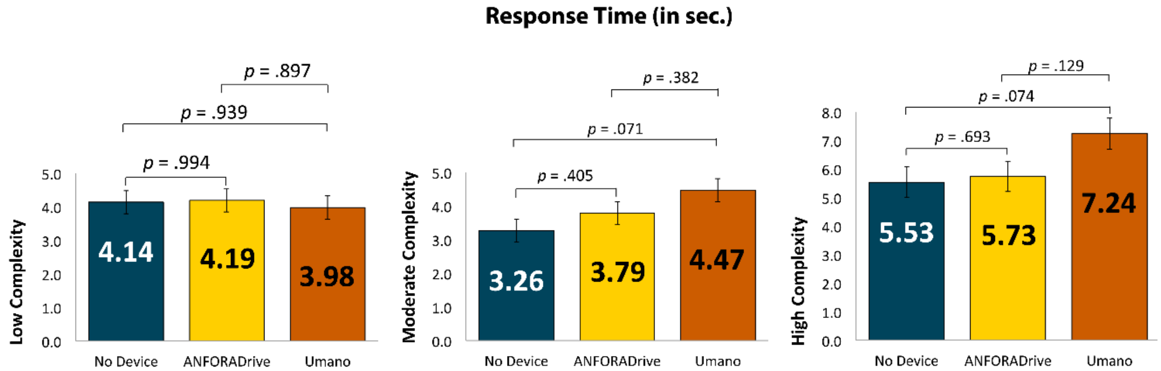


Figure 49. ANFORA*Drive*, Umano and no device yielded a similar response time in all three of the driving complexity scenarios.

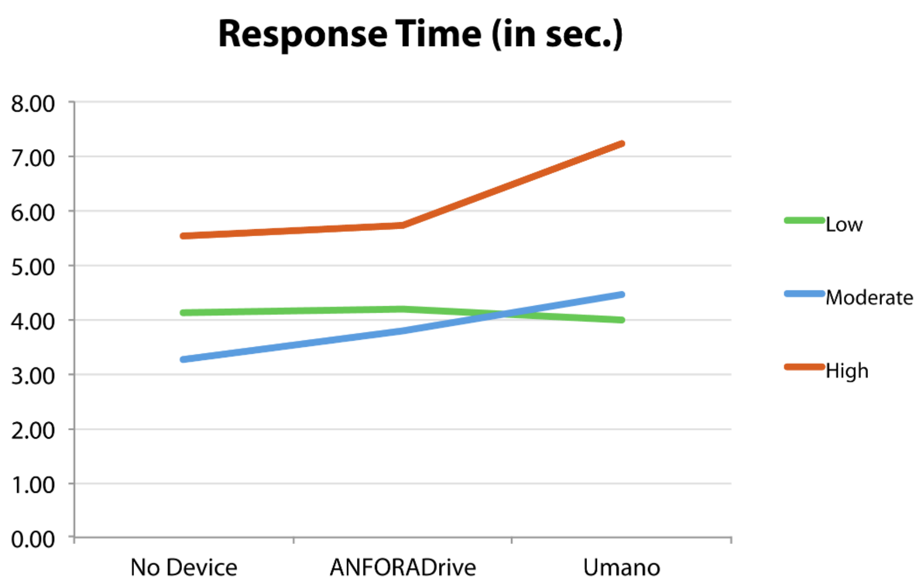


Figure 50. Response time has an increasing trend from no device to ANFORA*Drive* to Umano for both the moderate and high complexity scenarios (with no statistical significance present).

The participants using ANFORA*Drive*, Umano and no device yielded similar number of lane departures in the low and moderate complexity scenarios (Figure 51). However, the number of lane departures were significantly different for the high complexity scenarios

(Figure 51). In addition, an increasing trend exists from no device to ANFORA*Drive* to Umano for the low and moderate complexity scenarios (Figure 52). Overall, this study confirms H1.4: Compared to the driving only condition, ANFORA*Drive* does not reduce driving performance. However, this study does not fully confirm H2.4 and H3.4. It is likely that having more participants in each condition would help to confirm those two hypotheses as well since we did discover an increasing trend.

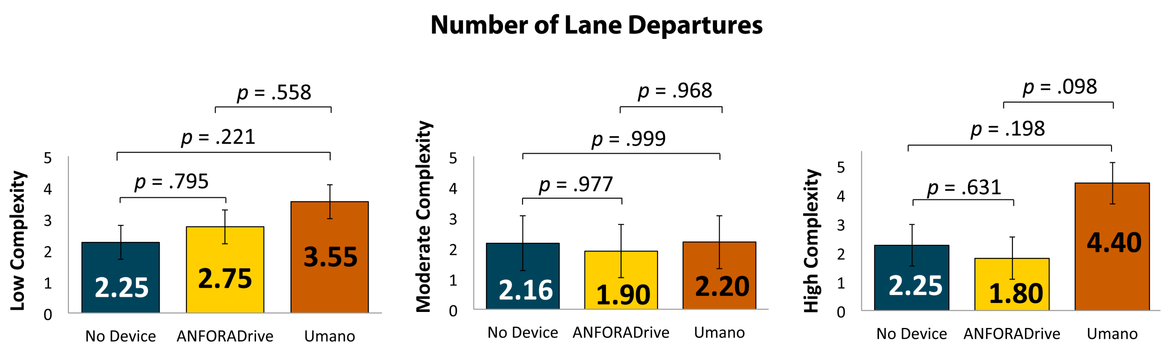


Figure 51. The number of lane departures was significantly different for the high complexity scenario.

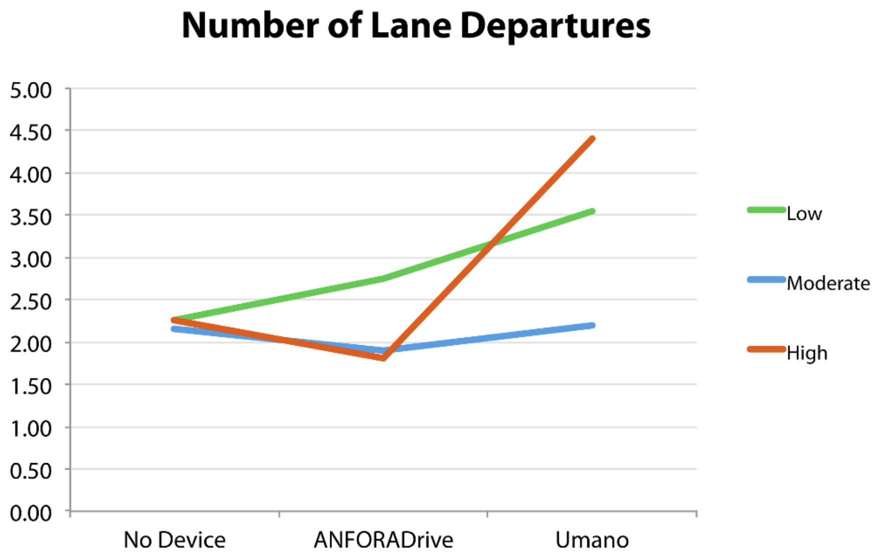


Figure 52. The number of lane departures has an increasing trend from no device to ANFORA*Drive* to Umano for the low and moderate complexity scenarios.

6.8.1.5. Driving Behavior

Our study confirms H1.5 and H2.5: Compared to the driving only condition, ANFORA*Drive* does not increase driver's visual interaction time with the device, but Umano increases driver's visual interaction time with the device in all the three driving complexities. This study also confirms H3.5: Compared to Umano, ANFORA*Drive* reduces driver's visual interaction time with the device. The TEOR of ANFORA*Drive* for the low, moderate and high complexity scenarios was 6.50 sec., 19.20 sec. and 6.20 sec., respectively (Figure 53), which was 1% of the total task time. However, the TEOR of Umano for the low, moderate and high complexity scenarios was 99.25 sec., 84.15 sec. and 74.90 sec., respectively (Figure 53), which was between 7% and 10% of the total task time. As the driving complexity increased, the TEOR while using Umano decreased because its participants needed to pay closer attention to the road. In addition, the visual interaction time with Umano was longer due to the manual interaction to change the news stories and categories.

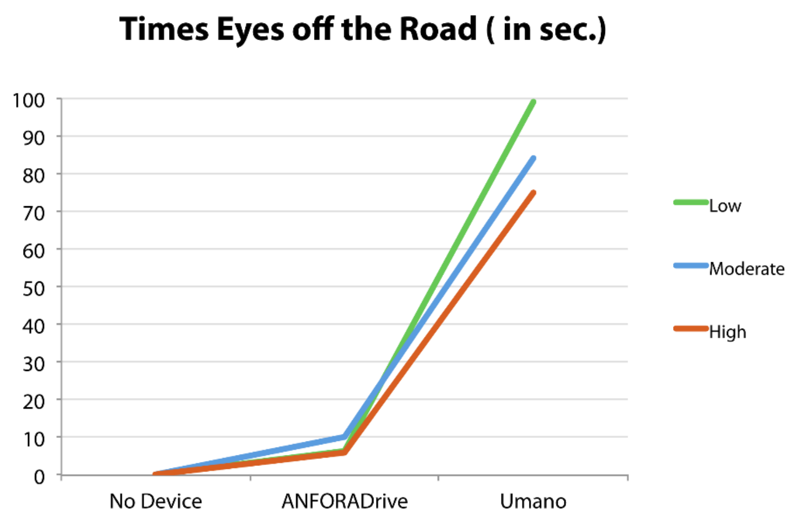


Figure 53. The visual interaction time with ANFORA*Drive* was 1% of the total task time (15 minutes).

All of our hypotheses, other than H2.4 and H3.4, were confirmed (Table 9). In summary, these results suggest that ANFORA*Drive* is similar to the no device condition in terms of driving performance, driving behavior, cognitive effort, distraction and overall safety. These findings suggest that using ANFORA*Drive* does not add any additional overhead or distraction when compared to not using any device.

Table 9. Hypotheses Revisited.

Hypotheses	*Confirmed/Rejected*
H1.1: Compared to the driving only condition, ANFORA*Drive* does not increase driver's cognitive effort.	Confirmed
H2.1: Compared to the driving only condition, Umano increases driver's cognitive effort.	Confirmed
H3.1: Compared to Umano, ANFORA*Drive* reduces driver's cognitive effort.	Confirmed
H1.2: Compared to the driving only condition, ANFORA*Drive* does not increase driver distraction.	Partially Confirmed
H2.2: Compared to the driving only condition, Umano increases driver distraction.	Partially Confirmed
H3.2: Compared to Umano, ANFORA*Drive* reduces driver distraction.	Confirmed
H1.3: Compared to the driving only condition, ANFORA*Drive* does not reduce overall safety.	Partially Confirmed
H2.3: Compared to the driving only condition, Umano reduces overall safety.	Partially Confirmed
H3.3: Compared to Umano, ANFORA*Drive* increases overall safety.	Confirmed
H1.4: Compared to the driving only condition, ANFORA*Drive* does not reduce driving performance.	Confirmed
H2.4: Compared to the driving only condition, Umano reduces driving performance.	Not Confirmed
H3.4: Compared to Umano, ANFORA*Drive* increases driving performance.	Not Confirmed
H1.5: Compared to the driving only condition, ANFORA*Drive* does not increase the driver's visual interaction time with the device.	Confirmed
H2.5: Compared to the driving only condition, Umano increases the driver's visual interaction time with the device.	Confirmed

H3.5: Compared to Umano, ANFORA*Drive* reduces the driver's visual interaction time with the device.	Confirmed
H3.6: Compared to Umano, ANFORA*Drive* increases user satisfaction while using the device.	Confirmed

6.8.2. The Role of Aural Flows While Driving

In this section, we discuss where aural flows belong in the driver distraction framework introduced by Strayer et al. (2011) as shown in Figure 4. Understanding the level of distraction generated by aural flows is important because prior studies (Barón & Green, 2006; Gable et al., 2013; Harbluk et al., 2002; Hua & Ng, 2010; Lee et al., 2001; Strayer et al., 2013; Winter et al., 2010) have provided contradictory findings on the role of audio/voice-based in-car systems on the cognitive overload. In our study, we used TEOR and NASA-TLX to measure visual and cognitive distraction, respectively. Manual distraction was very similar to visual distraction because, whenever the participants took their eyes off of the road, they manually interacted with their phones or the steering wheel button.

In order to understand the role of aural flows in the driver distraction framework, we first need to know where the driving only condition belongs in the driver distraction framework. The driving only condition has a *low* visual, manual and cognitive distraction on drivers because they simply drive and do not engage in any secondary tasks. As demonstrated in our results through TEOR and NASA-TLX, visual and cognitive distraction scored low for ANFORA*Drive* compared to the driving only condition (no device) (as reported in our Results, Sections 6.6.2 and 6.6.4). Therefore, aural flows also belong to *low* level for visual, manual and cognitive distractions. These low levels of distraction are evident when users *listen to* and *interact with* aural flows (Figure 54).

Similarly, *listening* to Umano also belongs to low level of distractions because it is comparable to listening to the radio, which belongs to low level distractions (Strayer et al., 2011). However, *interacting* with Umano could belong to the moderate or high level distractions (Figure 54) because both the TEOR and NASA-TLX scores increased significantly for Umano when compared to the no device condition (reported in our Results, Sections 6.6.2 and 6.6.4). In summary, our findings suggest that ANFORA*Drive* could be used as a non-distracting infotainment technology while driving.

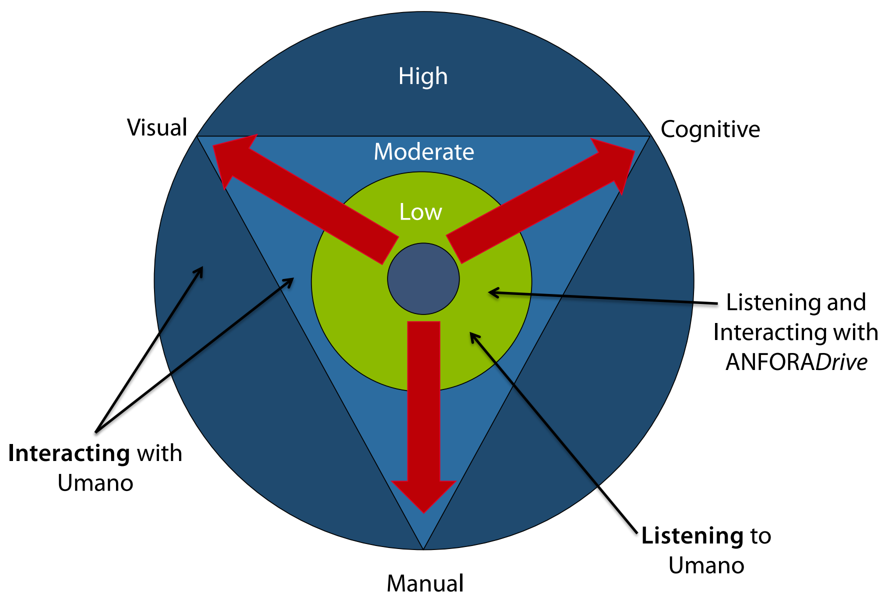


Figure 54. Listening to and interacting with aural flows belong to the low level condition for visual, manual and cognitive distractions. Listening to Umano also belongs to low level distractions. However, interacting with Umano could belong to the moderate or high level of distractions.

6.8.3. Usage of Aural Flows

6.8.3.1. Navigation Model: Full Flow with All-to-All Access vs. Group Flow with Index Access

In the results section, we showed that the participants changed the news category only once while using Umano, possibly because it takes four clicks to change a category (Figure 26). As reported in the results, we observed a radical difference between voice command usage in ANFORA*Drive* and button command usage in Umano. This difference could be because using button commands to interact visually with Umano needed more time and the participants preferred not to use such a time consuming method. This result was supported by our interview results, which showed that one of the main reasons why the participants did not want to use Umano was that it required visual interactions in order to change channels. For example, one of the participants (P2) explicitly noted the difference between ANFORA*Drive* and Umano,

> I liked how easy it was to switch between articles or between categories [in ANFORA*Drive*]. I didn't have multiple steps to go through. With the other app, I had to first go to the category and then I had to say it to start playing one of the articles. That was multiple steps to just do something as simple as starting a playlist.

These results show that full flow, along with all-to-all access and voice commands, could reduce visual interactions with the device and improve on the user's experience compared to group flow along with index access and button commands. Hence, full flow with all-to-all access and voice commands could better suit the driving context than group flow with index access and button commands.

6.8.3.2. Structural Navigation and Listening Experience within a News Story

While Umano is designed to provide its participants with only full news stories, ANFORA*Drive* provides both summary and full news stories. We observed, in our results, that four patterns of usage exist when the aural flows are used in ANFORA*Drive* (Figure 55). In the first pattern, the participants let the aural flow run through both the summary and full story. This pattern is the default flow provided to the users in ANFORA*Drive*. For example, the user will start the flow by accessing the first news in the U.S. category. He listens to the title of the story, summary and full story. Toward the end of the full story, he decides to move to the next story and says "next." He now listens to the second story in the U.S. category by listening to the title, summary and part of the full story. He then says "world" and the flow moves to the world news. He listens to the title, summary and full story of the first story in world news. In the second pattern, the participant prefers to listen to the summary of the news only.

The second pattern is called *sampling* and is based on the initial design ideas introduced for ANFORA News in Chapter 3 (Table 2). For example, the user listens to the title and summary of the first news story in the U.S. category, then he says "next" before the flow moves to the full story. He now listens to the summary of the second news story. Then, he says "world" and listens to the summary of the first news story in the world category. In the third pattern, the *comprehensive* pattern, the participant prefers to listen to only the full news story (Table 2 in Chapter 3). In this pattern, the user listens to either the title or a bit of summary and then says "full story."

In the fourth pattern, the *supplemental* pattern, the participants listen to related stories (Table 2 in Chapter 3). For example, the user listens to the summary and full story of the first news story in the U.S. category and then says “next.” He now listens to the summary and full story of the second news story in the U.S. category. He realizes that he is interested in listening to similar news on this topic and says “related.” Once he listens to the related story, the flow moves to the third story in the U.S. category. He, again, likes the third story and says “tell me more.” The flow takes him to a related story.

These four patterns confirmed the initial intention of the design of ANFORA, which was specified in Chapter 3. As reported in our results, we observed that the majority of our participants (87%) adhere to the first pattern (i.e., the default function of the aural flow). Fewer than 10% of our participants adopted both the second and third usage patterns. Finally, 15% of our participants utilized the fourth usage pattern. These results were also supported by our interview results, where 23 of our participants commented they were not going to use Umano because they could only listen to the full story, but not the summary and related stories.

These results show that providing both the summary and full story as default could be a good option in the context of driving, but designers need to give the users ability to set their preferences beforehand or while listening to the flow. For example, the users might want to listen to summaries of breaking news, but full stories in the science news category. They might be able to do it by saying “breaking news summaries” or “science news full stories.”

Additionally, a two-way ANOVA was conducted to find out if the driving complexity scenarios and aural flows usage did affect the driving performance measurements. The results showed that both driving complexity scenarios and aural flows usage did not have a significant main effect on both number of lane departures and response time.

Hence, enabling the users to select the aural flows based on their preference will not affect their driving performance.

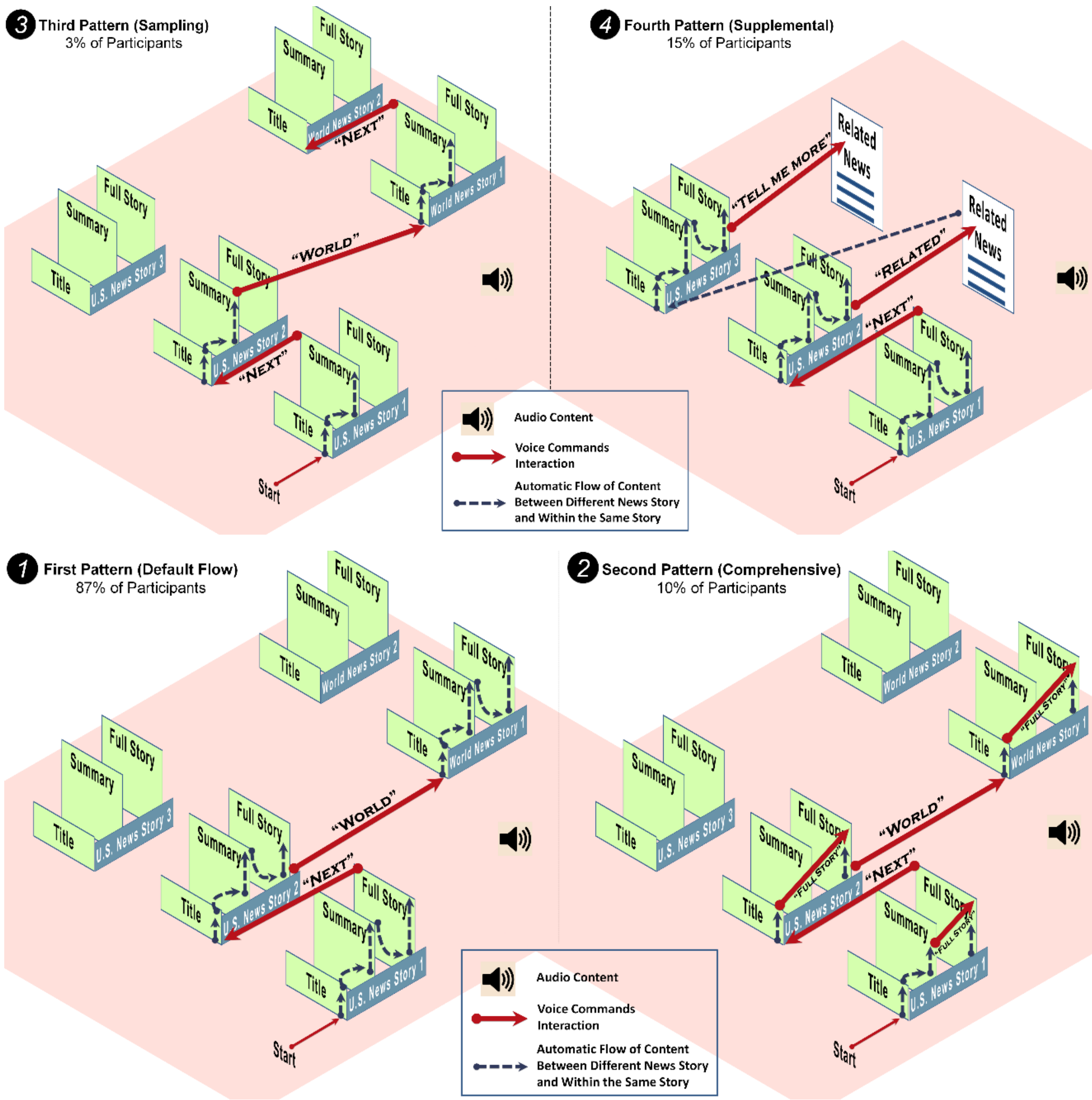


Figure 55. Four different patterns of aural flow usage: 1) 87% of the participants let the aural flow goes through both the summary and full story, 2) 10% of the participants preferred to listen to the full news story only, 3) 3% of the participants preferred to listen to the summary of the news only and 4) 15% of the participants listened to related stories in addition to the summary or full story.

6.8.4. Limitations of the Study

One limitation of our experimental design is that we conducted the study with 60 participants due to time and resource constraints. Conducting the study with over 100 participants would give better power for our statistical data analysis. The second limitation is that distraction and overall safety questions were not asked for the "no device" condition. As such, we had to look into correlation of distraction and safety with NASA-TLX measurement to make a judgment about the distraction and safety ratings for the "no device" condition.

The third limitation of the study is that we had a between-subject design (20 participants for each condition) for the aural applications in each driving complexity scenario (i.e., low, moderate and high). However, having a within-subject design with 20 participants for each condition would give us a more accurate result, since the same participant would use different aural applications. The only limitation with this experimental design would be that the same user would go through only one driving complexity scenario. Therefore, the participants would become familiar with the path of the driving scenario. In order to reduce the learning effect of the path, we would have to create three versions for each driving complexity scenario. For example, the low complexity scenario would have three versions. Similarly, each of the moderate and high complexity scenarios would have three versions. Therefore, we would have nine versions for each driving complexity scenario. We did not pursue this path due to resource constraints.

6.9. Conclusions

Through this controlled evaluation study, we learned that the ANFORA*Drive* condition was similar to the no device condition in terms of driving performance, driving behavior, cognitive workload, distraction and overall safety. These findings are positive and show that ANFORA*Drive* does not add any additional cognitive overhead for drivers even though they are aurally listening to and interacting with their mobile devices. These findings are contradictory to the recent study by Strayer et al. (2013), which suggested that using speech-to-text systems to text message in the car is risky because too many voice interactions still tax our attention bandwidth. This contradiction could occur because our participants were using voice commands to interact with aural flows compared to the larger number of voice commands required for sending a text message while driving.

Overall, this study showed that aural flows allow participants to engage with web-based news content without having to visually browse the screen while driving. Admittedly, ANFORA*Drive* needs further improvements and developments based on the findings gathered during this study. In the next chapter, I will discuss the main contributions of this dissertation to the HCI research community, news industry and automobile industry.

Chapter 7. Summary of Contributions

7.1. HCI Research Community

This research contributes novel HCI knowledge that informs the design of a new class of aural and semi-aural user interfaces for the mobile experience (i.e., systems that transform existing web information architecture into linear, aural flows to be comfortably listened to, thus off-loading the eyes from continuous attention to mobile devices). Our approach is exemplified in ANFORA, a set of semi-aural mobile application prototypes optimized to generate real-time aural flows from web sources and allow the user to listen to large collections of news stories on the go. This research also investigated eyes-free input modalities used to interact in the context of walking and driving with semi-aural user interfaces and control aural flows created from the web. This dissertation provides five main significant contributions to consuming content-rich websites while on the go:

- **Continuous Flows of Content:** ANFORA eliminates the need for intermittent navigation by providing aural flows. A flow is governed by aural design rules that determine which pages of the information architecture to concatenate automatically as well as how users can control these flows. Aural flows act as playlists of content. The application provides the following types of aural flows based on the breadth and the time length of the content covered: group flow and full flow. These flow types are associated with different aspects of the information architecture of a content-intensive website.

- **Enhancing the Mobile Experience:** Users can employ the proposed application on modern smart phones (i.e., iPhone and Android devices). Hence, they do not need to sit in front of their personal computers to use it.
- **Making Complex Websites Simpler:** The structure of content rich websites (such as news, education or tourism websites) is not only hierarchical, but also hypertextual. As an example of a hypertextual feature, while browsing a news website, a user could quickly reference related news stories or news stories within the same subcategory. The aural browsing experience can become difficult when users have to navigate non-hierarchical websites. In order to address this challenge, ANFORA provides aural flows that cover the hypertextual relationship among the content.
- **Topical Access to Content:** ANFORA introduces different types of content categorizations specifically suited for aural navigation. For example, users can choose to listen to segments of news stories based on time constraint (e.g., five or 10 minute aural flows) or the degree of the coverage of the content (e.g., only a summary of the news or the full story).
- **All-to-All Access to Categories** ANFORA enables users to begin listening to any content and move to any other content without returning to an index or home page to re-select options. For example, users listening to a technology news story can simply select “World” in the menu options to listen to the world news instead of returning to an index page.

This research investigated the role of aural flows in two different contexts, such as walking and driving. These two contexts were selected as an example of contexts featuring both low and high cognitive load and distraction. In the walking scenario, where lower cognitive demand existed, using button vs. voice commands did not strongly effect the system's usability and cognitive workload. However, in the driving scenario, where a higher cognitive demand existed, using voice vs. button commands increased system usability and reduced cognitive workload. Additionally a significant contribution of this research is identifying that voice-controlled aural flows belong to low level visual, manual and cognitive distraction on driving distraction framework (Figure 56).

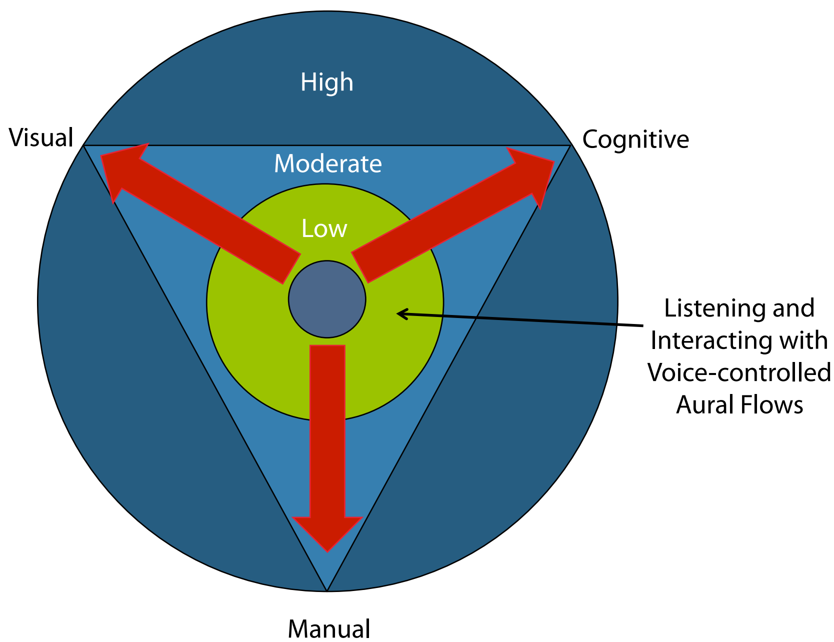


Figure 56. Listening to and interacting with voice-controlled aural flows belong to the low level for visual, manual and cognitive distractions.

7.2. Potential Contributions to the News Industry

ANFORA News differs from other methods of listening to the news, such as radio broadcasts and news podcasts, due to differences in a few key principles, including flexibility of access and the level of abstraction of the content selection. As such, ANFORA provides a multimodal experience that provides different output and input modalities as well as various levels of reading/listening (e.g., scan headlines, sample story summaries and listen to full stories). A radio news broadcast, on the other hand, is synchronous in that users tune into a complete newscast edited linearly by a producer for a predetermined time slot and mass audience. The news podcast provides a more asynchronous experience by allowing users to download programs and listen to them wherever and whenever they want. However, these programs are edited by producers who have the mass audience in mind. Thus, neither the radio news broadcast nor news podcast can take into consideration any single individual's time constraints and/or personal interests. ANFORA, however, lets users decide the length of time they want to spend with the news and how in-depth they want to delve into individual stories. Therefore, ANFORA provides an unmatched user experience opportunity in the midst of a dramatic transition as the news industry struggles to keep up with the rapidly evolving media landscape.

In today's society, the news industry is searching for methods by which to reach young audiences using their phones and tablets. ANFORA represents a potential paradigm shift in an industry that is struggling to reinvent itself and more effectively reach audiences by leveraging paradigms with which younger users are already familiar (e.g., listening to playlists on the go). Finally, it is important to note that the innovations introduced by ANFORA apply to a variety of content-intensive domains, for which new casting is a prominent example.

7.3. Automobile Industry

According to Richard Robinson, the director of the Automotive Multimedia and Communications Service (AMCS), “in five years, nearly 25% of the cars will be connected to the Internet (Car and Driver, 2015).” He also noted, in his article about the future of in-car technology, that “your dashboard may soon become as versatile as your laptop (Car and Driver, 2015).” The same article stated that, in the near future, customers would be able to visit an automaker’s app store in order to install software in their cars instead of buying a new device. For example, MyFord Touch enables car drivers and passenger to configure and listen to their own Internet music “station” via Pandora (Car and Driver, 2015). Another article noted that Android will soon be integrated into cars (Digital Afro, 2015). Similarly, in the near future, car touchscreen dashboards will enable drivers and passengers to listen to their personalized news playlists.

Hence, this research contributes to novel HCI techniques used to design applications that could be installed in car touchscreen dashboards. This application could transform existing web information architectures (e.g., news, education or government websites) into playlists of content to be comfortably listened to and interacted with via voice. In addition, if users installed this application on their phone or laptop and they were just listening to the content playlist at their home, once they go to their car, they could continue listening to the same playlist via their car dashboard. In the autonomous cars of the future, this transition could be done seamlessly as the system in the car recognizes that the users were just listening to the news vs. music before they go to their vehicle.

The study conducted in the driving simulation lab showed that the current design of aural flows is suitable while driving since it does not add any significant cognitive workload, distract users or change the users' driving performance. Moreover, the results of this study showed that listening to and interacting with aural flows belong to low level visual, manual and cognitive distraction framework (Figure 56). This research will enable car drivers to keep their eyes on the road and their hands on the steering wheel to avoid future accidents. Ultimately, this research enables us to understand the possibility of cooperating aural flows in autonomous cars.

Chapter 8. Future Research Directions

8.1. Controlling Aural Flows Using Touch

There are four directions in which this research could be expanded. One possibility is to further investigate controlling aural flows with touch/gesture. A few studies, such as those studies that investigated earPod (Zhao et al., 2007) and Bezel-Tap (Serrano, Lecolinet, & Guiard, 2013), have shown that touch/gesture can decrease the visual interaction with an interface. In addition, in our study, two of our participants liked how they could swipe to go to the next or previous story using Umano while driving (See Section 6.7.4). Hence, we could explore a vocabulary of gesture interactions for controlling aural flows via the interface or a car's steering wheel (Döring et al., 2011).

Previous research has examined using gestures to interact with infotainment systems in the car (Ohn-Bar, Tran, & Trivedi, 2012). For example, a single-finger swipe right or left might enable movement between the stories and up and down swipes might move between categories. Single taps could go to the full story, and double taps could stand in for "pause" or "play" commands. These gestures could apply to use on both the interface and on the steering wheel. It is also important to investigate which part of the interface should be used for gestural interaction so that users do not hit the wrong button commands by mistake.

8.2. Investigating Additional Voice Commands for Other Interactions

Another possibility is to further investigate controlling aural flows with the additional voice commands. Right now, the *Linkless* ANFORA prototype has a limited number of categories, all of which are taken from NPR. However, in the marketable application, we would want to have a broader variety of categories as well as subcategories and news stories from different sources, such as CNN, BBC and *The New York Times*. Therefore, it is worth investigating whether our users need to remember the voice commands for all of the categories or only for those categories they access regularly. We could also explore how to provide users with personalized flows after repeated usage. For example, if a user accesses only health and technology news the first 10 times that he accesses the app, then the next time the user accesses the app, the aural flow would begin by default showing only health and technology news.

In our final experiment, some of the participants preferred listening only to summaries, while other participants preferred listening to full stories depending on the category in which the stories were being listened. Moreover, users might want to listen to summaries of breaking news, but full stories for science news. As such, the voice commands could be “breaking news summaries” or “science news full story.” Another interesting pattern of aural flows navigation was that our users liked to listen to some of the related stories. Therefore, it is necessary to investigate how these additional voice commands could be used in ways that would provide users with the freedom to interact with the device in more meaningful ways.

8.3. Applying Aural Flows to Other Domains

The third possibility is to explore aural flows within other content-rich websites, such as social networking, education or government websites. For example, on Facebook, the user could listen to the 10 most recent posts to his feed or listen to the feeds of a select group of friends. He could also listen to the comments for each of the feeds. This idea could be transferred to Twitter, where a user could to listen to the tweets of users whom he follows.

We could also expand on the main idea of aural flows and explore a generic framework. This framework could be built on top of any content-rich website, allowing the user to access the website's API and content and convert the sites to aural flows. Once the content is obtained, the main challenge would be to identify the category into which the content falls. For example, the framework must be able to distiguish between a feed's content and the comments. In addition, each social networking website tags its content differently from its peers, so the program would need to be able to distinguish between the types. Another challenge would be to identify whether the website was a news, social networking or government website.

8.4. Exploring Aural Flows for Visually-impaired Users

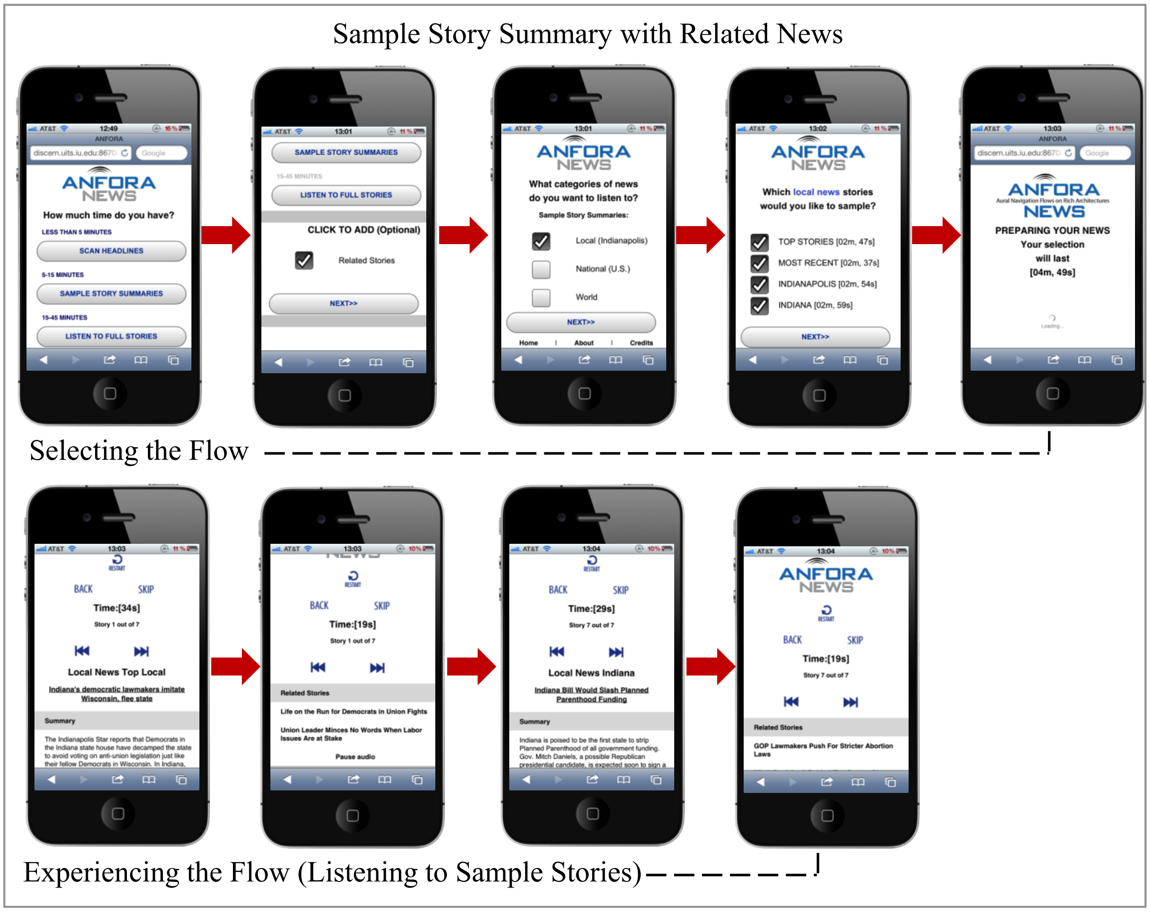
The fourth possibility is to investigate how to use *Linkless* ANFORA for visually-impaired users. Since visually-impaired individuals consume web content by listening to it using screen readers, it is worth exploring how to use aural flows for the visually-impaired users, especially since accessing aural flows using voice commands has proven to be useful for eyes-free scenarios, such as driving. For example, we could conduct a geussability study with visually-impaired users. First, we could train them on how to use the *Linkless* ANFORA and interact with it using voice commands. Then, we could ask them to provide us with other voice commands or gestures that would be helpful within the program. This geussability study could inform us about what voice commands or gestures are more natural for visually-impaired users when interacting with aural flows.

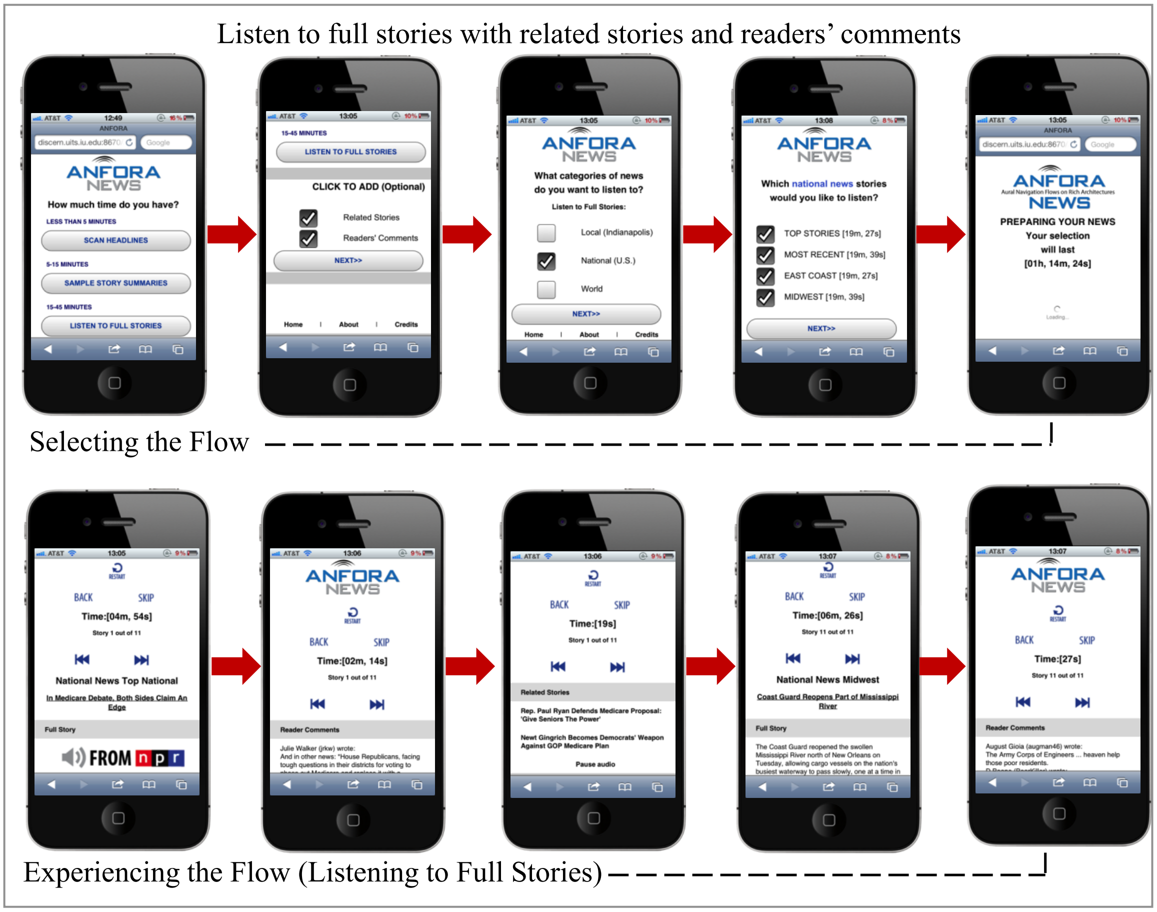
Appendices

Appendix A: Detailed Screenshots of ANFORA News Prototype

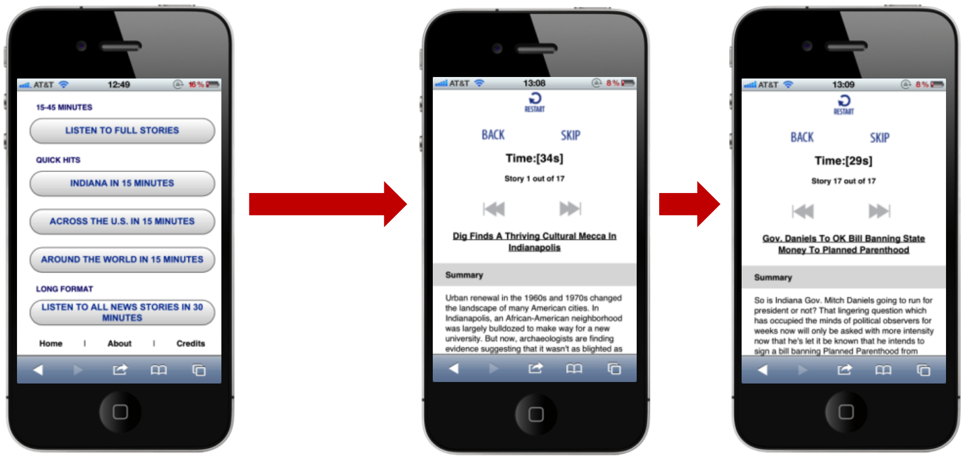
The ANFORA News prototype is available at: http://discern.uits.iu.edu:8670/ANFORA/

Full Source code and database are available at:
http://discern.uits.iu.edu:8670/downloads/ANFORA(Feb15_2012).zip


Sample Story Summary with Related News
Selecting the Flow
Experiencing the Flow (Listening to Sample Stories)


Listen to full stories with related stories and readers' comments
ANFORA NEWS
How much time do you have?
SCAN HEADLINES
SAMPLE STORY SUMMARIES
LISTEN TO FULL STORIES
CLICK TO ADD (Optional)
Related Stories
Readers' Comments
NEXT>>
What categories of news do you want to listen to?
Local (Indianapolis)
National (U.S.)
World
Which national news stories would you like to listen?
TOP STORIES [19m, 27s]
MOST RECENT [19m, 39s]
EAST COAST [19m, 27s]
MIDWEST [19m, 39s]
PREPARING YOUR NEWS
Your selection will last [01h, 14m, 24s]
Selecting the Flow
BACK
SKIP
Time:[04m, 54s]
National News Top National
Full Story
FROM npr
Time:[02m, 14s]
Reader Comments
Time:[19s]
Related Stories
Time:[06m, 26s]
National News Midwest
Coast Guard Reopens Part of Mississippi River
Time:[27s]
Experiencing the Flow (Listening to Full Stories)

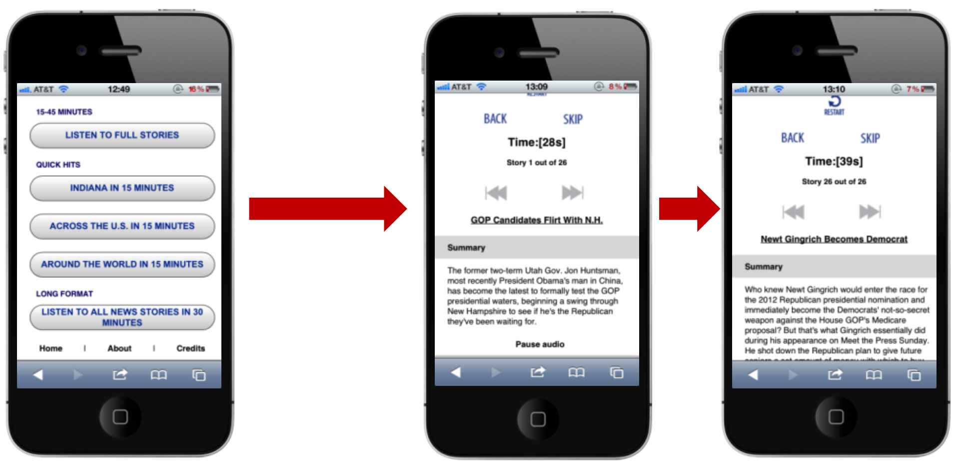
Quick hits “Indiana in 15 min”



Selecting the Flow

Experiencing the Flow (Listening to Sample Story)

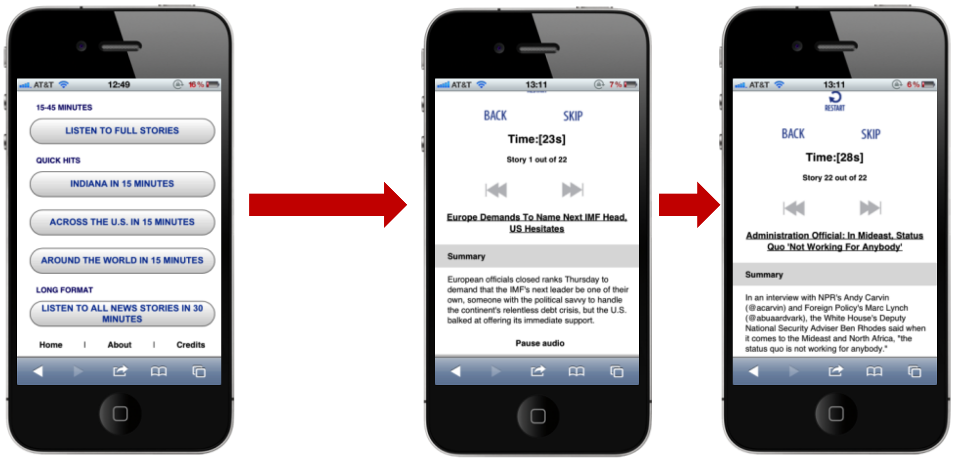
Quick hits “Across the U.S. in 15 min”



Selecting the Flow

Experiencing the Flow (Listening to Sample Story)

Quick hits “Across the world in 15 min”



Selecting the Flow

Experiencing the Flow (Listening to Sample Story)

Listen to all news stories in 30 minutes

Selecting the Flow

Experiencing the Flow (Listening to Full Stories)

Interrupting the Flow to Go Back to Homepage

Appendix B: Evaluation Study Instruments and Scripts

Introductory Script

Thank you for agreeing to participate in our research project. For this project, we have developed a mobile news application called ANFORA News that will allow you to listen to news stories in the form of text-to-speech while on-the-go. ANFORA News is designed to allow you to customize your news experience, by first choosing the types of news stories you would like to listen to based on how much time you have.

Imagine that you want to browse a news website using your mobile phone. But you also must walk to work or class, which may make it difficult to visually browse a news site and read stories while you walk. ANFORA News allows you to select the categories of news you want to listen to before beginning your walk. Once your selections have been made, ANFORA News creates a playlist of those stories and allows you to listen to them, one after another, without further visual interaction with the screen. In other words, ANFORA News provides a customized, eye-free news listening experience.

In general, we aim to test ANFORA News' usability, collect your opinions regarding its strengths and weaknesses and determine whether you find the ANFORA News experience to be enjoyable. Therefore, you will be asked to complete up to three simple tasks focused on interaction with the ANFORA News interface while walking through a busy hallway. I will join you on your walk to observe your interactions with the interface, video record your session and help you if any technical problems should arise. During this experience, please let me know if you become distracted by your surroundings and/or obstacles encountered while walking. When we return to the lab, I will ask you a series of questions regarding your experience. The entire session should last about one hour.

You do not have to interact with the screen after making an initial news playlist. However, if you want to, there are both control buttons on the screen and gesture commands you can use to do so. The buttons should be self-explanatory. Gesture commands are as follows: One-finger swipe left allows you to go to the next section within a story; one-finger swipe right allows you to go to the previous section within a news story; two-finger swipe left allows you to go to the next news story; and two-finger swipe right allows you to go to the previous news story. You can also scroll to the top of the page and use the button control commands if you like.

You can skip to the next story or stop the flow at any time. However, ANFORA News is designed to minimize interaction.

First Task Set

1. From the home screen, select “Scan Headlines.” Then, select all three categories “local, national and world.” For each category, select at least two sub-categories. ANFORA News is designed to minimize interaction and allow you to listen to a playlist of news stories on-the-go. However, you can skip to the next story or stop the flow any time you like by using gesture or control commands.
2. From the home screen, select “Listen to Full Stories” and add “Related Stories” and “Comments.” Then, select one category, “local, national or world.” Finally, select all four sub-categories. Remember that you can skip to the next story or stop the flow any time you like by using gesture or control commands. However, you are not required to interact with the screen after making these initial selections.
3. From the home screen, select one of the three “Quick Hits” options. Remember that ANFORA News is designed to minimize interaction. But you can skip to the next story or stop the flow any time you like.

Second Task Set

1. From the home screen, select “Sample Story Summaries” and add “Related Stories.” Then, select two of three categories, “local, national and world.” For each of the two categories you selected, choose at least two sub-categories. Remember that you can skip to the next story or stop the flow any time you like. ANFORA News is designed to minimize interaction and allow you to listen to a playlist of news stories on-the-go. However, you can skip to the next story or stop the flow any time you like by using gesture or control commands.
2. From the home screen, select “Listen to All News Stories” under the “Long Format” option. Remember that you can skip to the next story or stop the flow any time you like by using gesture or control commands. However, you are not required to interact with the screen after making these initial selections.

Survey for First Task Set

On a scale of 1-5 (1 = strongly disagree, 5 = strongly agree) rate your level of agreement with the following statements:

1. ANFORA News is easy to use.

2. Listening to news on ANFORA News is enjoyable.

3. I would use ANFORA News again.

4. I prefer using ANFORA News to browsing news websites on my mobile device.

5. ANFORA News was easy to navigate.

6. The text-to-speech voice was difficult to understand.

7. I got what I expected when I clicked on things (buttons, links, etc.) on this site.

8. The news content was interesting.

9. The quality of the text-to-speech voice was satisfactory.

10. The news content was boring.

11. After using ANFORA News, I feel well informed about the news categories I listened to.

12. While listening to ANFORA News, I realized when the news story started and ended.

13. While listening to ANFORA News, I realized the category in which the news story belonged to.

14. The “Scan Headlines” feature was useful.

15. The “Sample Story Summaries” feature was useful.

16. The “Listen to Full Stories” feature was useful.

Survey for Second Task Set

1. ANFORA News is easy to use.

2. Listening to news on ANFORA News is enjoyable.

3. I would use ANFORA News again.

4. I prefer using ANFORA News to browsing news websites on my mobile device.

5. ANFORA News was easy to navigate.

6. The text-to-speech voice was difficult to understand.

7. I got what I expected when I clicked on things (buttons, links, etc.) on this site.

8. The news content was interesting.

9. The quality of the text-to-speech voice was satisfactory.

10. The news content was boring.

11. After using ANFORA News, I feel well informed about the news categories I listened to.

12. While listening to ANFORA News, I realized when the news story started and ended.

13. While listening to ANFORA News, I realized the category in which the news story belonged to.

14. The “Sample Story Summaries” feature was useful.

Interview Questions

1. Overall, how would you describe your experience with ANFORA News?

2. How convenient was it for you to set up your news playlist? In other words, how easy was it for you to choose the categories of news you wanted to listen to?

3. Were you able to adequately monitor your surroundings while walking? If no, why not?

4. Was it clear when a new news story started/ended?

5. At any point, did you feel confused by the interface? If so, can you recall when?

6. At any point, did you feel lost in the while listening to the news? If so, can you recall when?

7. Did you notice any sound effects such as music or bells in between stories? If yes, what did they mean to you?

8. At any point, did you stop ANFORA News before your playlist ended? If yes, why?

9. Did you use gesture commands? Control commands? Both? Why or why not?

10. How did you feel about the way ANFORA allowed you to make initial choices about what types of stories you wanted to listen to and then automatically played stories in order after those choices were made?

11. If ANFORA News were available today, when would you use it? How? Why or why not?

12. What did you like best about ANFORA News?

13. What did you like least about ANFORA News?

14. How many news stories did you listen to today?

15. Briefly tell me about a news story that you remember.

Appendix C: Tabulated Data

Task Performance

Aural Flow Completion Rate	*Scan Headlines (T1)*	*Full Stories with Readers' comments & Related News (T2)*	*Sample story Summary (T3)*	*Sample Story Summary with Related News (T4)*	*Full Stories (T5)*
Completion without assitance	8	5	4	5	3
Completion with assitance	2	5	6	3	5
Users Gave up	0	0	0	2	2

	Confused by Long Pauses	*Encountered Technical Problem*	*Poor Recall of Gesture Commands*	*Misunderstood Button Labeling*	*Misunderstood TTS*
Percentage Occurrence of Error During Total Number of Listening Sessions (50)	50%	36%	28%	10%	6%

	Scan Headlines (T1)	*Full Stories with Readers' comments & Related News (T2)*	*Sample story Summary (T3)*	*Sample Story Summary with Related News (T4)*	*Full Stories (T5)*
Engagement with the Screen	21.70	19.07	18.74	29.25	20.76
Listening to Aural flow	78.30	80.93	81.26	70.75	79.24

Survey Questionnaire

	Age	*Gender*	*Kind of Phone*	*News Web*	*Mobile News*	*Radio News*	*TV News*
P1	30	M	iPhone	6+ hrs	5-30 min	no time	no time
P2	27	F	Epic 4G	5-30 min	5-30 min	no time	5-30 min
P3	27	M	iPhone	3-6 hrs	30-60 min	5-30 min	6+ hrs
P4	26	F	Samsung	5-30 min	5-30 min	5-30 min	1-3 hrs
P5	23	F	iPhone 4S	1-3 hrs	5-30 min	no time	no time
P6	24	F	Basic Model	1-3 hrs	no time	no time	no time
P7	25	M	Blackberry Torch	30-60 min	5-30 min	3-6 hrs	3-6 hrs
P8	24	F	iPhone	5-30 min	1-3 hrs	no time	no time
P9	27	M	Nokia- M73	6+ hrs	no time	1-3 hrs	1-3 hrs
P10	55	F	LG Optimus	30-60 min	30-60 min	1-3 hrs	6+ hrs
P11	26	M	Android	5-30 min	no time	5-30 min	no time
P12	50	M	Blackberry	5-30 min	5-30 min	6+ hrs	no time
P13	29	F	iPhone 4	1-3 hrs	5-30 min	no time	1-3 hrs
P14	37	F	Android Samsung Fascinate	1-3 hrs	5-30 min	no time	1-3 hrs
P15	23	F		30-60 min	5-30 min	5-30 min	5-30 min
P16	27	M	LG CU500 (java)	3-6 hrs	no time	5-30 min	5-30 min
P17	37	M	iPhone 4	1-3 hrs	1-3 hrs	1-3 hrs	no time
P18	30	F	iPhone	30-60 min	5-30 min	3-6 hrs	3-6 hrs
P19	24	M	Android-SGH T959	1-3 hrs	30-60 min	5-30 min	5-30 min
P20	34	M	Regular P.O.S	30-60 min	no time	1-3 hrs	no time

	Q1. ANFORA News is easy to use.	*Q2. Listening to news on ANFORA News is enjoyable.*	*Q3. I would use ANFORA News again.*	*Q4. I prefer using ANFORA News to browsing news websites on my mobile device.*	*Q5. ANFORA News was easy to navigate.*	*Q6. The text-to-speech voice was difficult to understand.*	*Q7. I got what I expected when I clicked on things (buttons, links, etc.) on this site.*	*Q8. The news content was interesting.*	*Q9. The quality of the text-to-speech voice was satisfactory.*	*Q10. The news content was boring.*	*Q11. After using ANFORA News, I feel well-informed about the news categories I listened to.*	*Q12. While listening to ANFORA News, I realized when the news story started and ended.*	*Q13. While listening to ANFORA News, I realized the category in which the news story belonged to.*	*Q14. The "Scan Headlines" feature was useful./ The "Sample Story Summaries" feature was useful.*	*Q15. The "Sample Story Summaries" feature was useful.*	*Q16. The "Listen to Full Stories" feature was useful.*
P1	2	4	3	2	1	1	2	4	5	2	4	4	4	3	3	4
P2	4	4	3	3	4	2	3	4	4	2	4	3	3	4		
P3	4	5	5	5	4	3	2	3	3	3	5	5	5	4	4	3
P4	5	4	4	5	5	1	3	3	4	2	4	4	4	5		
P5	5	5	4	5	1	5	4	5	3	3	5	5	5	5	5	5
P6	3	3	4	3	4	2	3	5	5	2	4	3	4	4	5	3
P7	4	4	4	3	4	2	4	4	4	2	5	5	4	3		
P8	3	3	2	2	4	1	4	4	4	2	4	4	2	3		
P9	3	2	3	3	3	4	4	4	2	2	4	3	3	4	4	4
P10	4	5	5	3	5	1	4	5	5	1	5	4	4	4	4	4
P11	5	5	4	5	4	1	4	4	4	2	3	4	3	4		
P12	5	3	4	4	5	2	4	3	2	3	4	5	4	4	4	4
P13	4	4	4	4	5	1	5	4	5	1	5	5	3	5		
P14	4	4	4	3	5	2	5	4	4	2	4	3	4	4	4	4
P15	4	3	3	3	4	2	4	3	4	3	3	4	4	3	3	3
P16	4	4	4	4	4	2	4	4	4	2	3	5	4	3		
P17	4	4	4	4	2	2	1	3	3	2	4	1	2	2	3	4
P18	4	5	5	5	2	3	5	5	5	1	5	5	4	5		
P19	4	4	4	3	3	2	4	2	4	4	5	3	2	5		
P20	5	4	4	4	5	3	5	3	3	2	4	4	4	4		

Appendix D: ANFORA News Patent

US 20140282006A1

(19) **United States**
(12) **Patent Application Publication** (10) **Pub. No.: US 2014/0282006 A1**
Bolchini (43) **Pub. Date: Sep. 18, 2014**

(54) **AURAL NAVIGATION OF INFORMATION RICH VISUAL INTERFACES**

(71) Applicant: **Davide Bolchini**, Indianapolis, IN (US)

(72) Inventor: **Davide Bolchini**, Indianapolis, IN (US)

(21) Appl. No.: **14/024,612**

(22) Filed: **Sep. 11, 2013**

Related U.S. Application Data

(60) Provisional application No. 61/799,748, filed on Mar. 15, 2013.

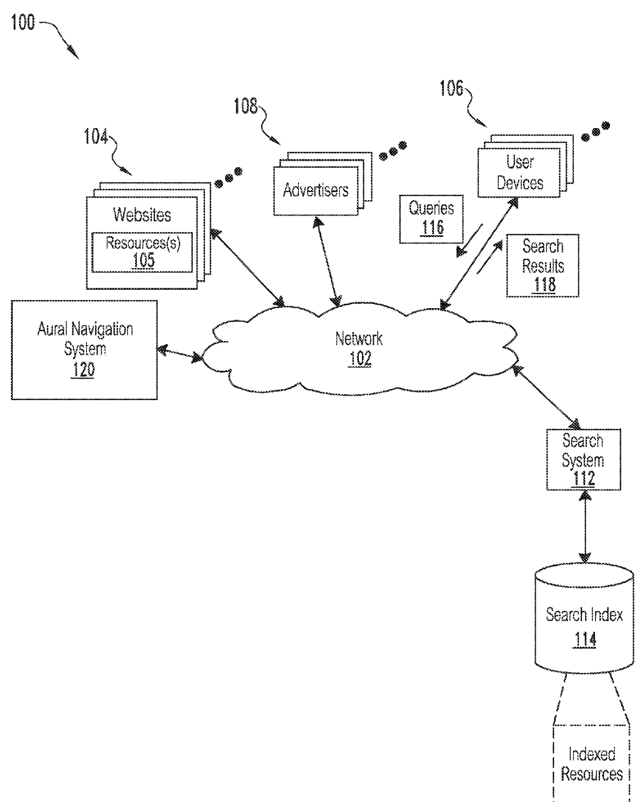
Publication Classification

(51) **Int. Cl.**
G06F 3/16 (2006.01)

(52) **U.S. Cl.**
CPC ***G06F 3/167*** (2013.01)
USPC **715/728**

(57) **ABSTRACT**

A method comprising generating, by a computer, a model of a website using user interaction primitives to represent hierarchical and hypertextual structures of the website; generating, by the computer, a linear aural flow of content of the website based upon the model and a set of user constraints; audibly presenting, by the computer, the linear aural flow of the content such that the linear aural flow of content is controlled through the use of user supplied primitives, wherein, the linear aural flow can be turned into a dynamic aural flow based upon the user supplied primitives.



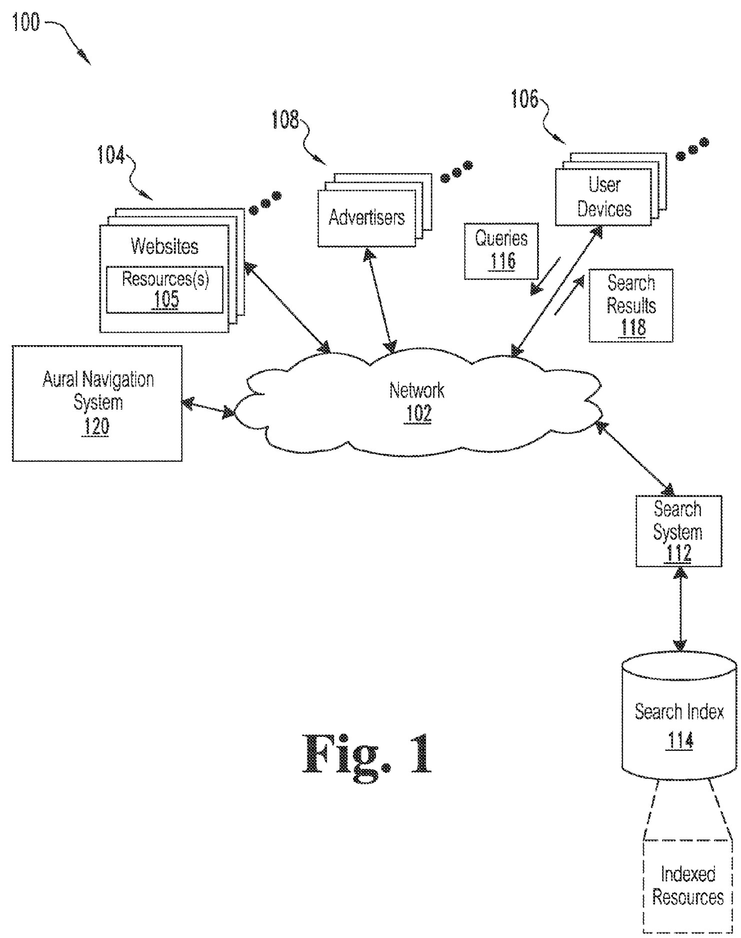

100
104
108
106
Websites
Resources(s)
105
Advertisers
User
Devices
Queries
116
Search
Results
118
Aural Navigation
System
120
Network
102
Search
System
112
Search Index
114
Indexed
Resources

Fig. 1

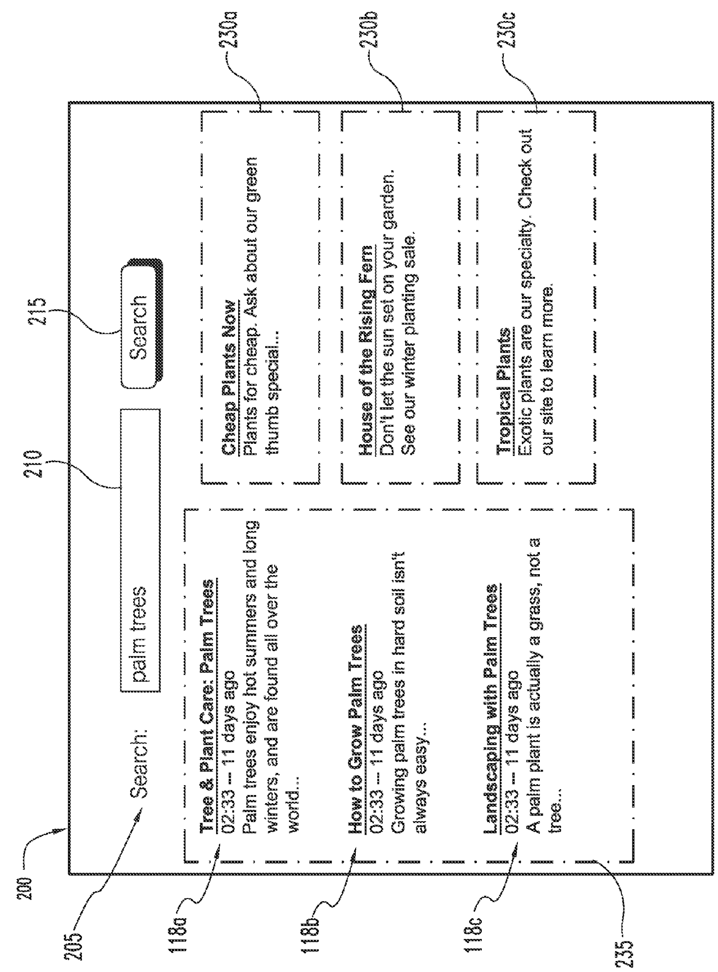

200
205
Search:
palm trees
210
Search
215
118a
Tree & Plant Care: Palm Trees
02:33 -- 11 days ago
Palm trees enjoy hot summers and long winters, and are found all over the world...
118b
How to Grow Palm Trees
02:33 -- 11 days ago
Growing palm trees in hard soil isn't always easy...
118c
Landscaping with Palm Trees
02:33 -- 11 days ago
A palm plant is actually a grass, not a tree...
235
230a
Cheap Plants Now
Plants for cheap. Ask about our green thumb special...
230b
House of the Rising Fern
Don't let the sun set on your garden. See our winter planting sale.
230c
Tropical Plants
Exotic plants are our specialty. Check out our site to learn more.

Fig. 2

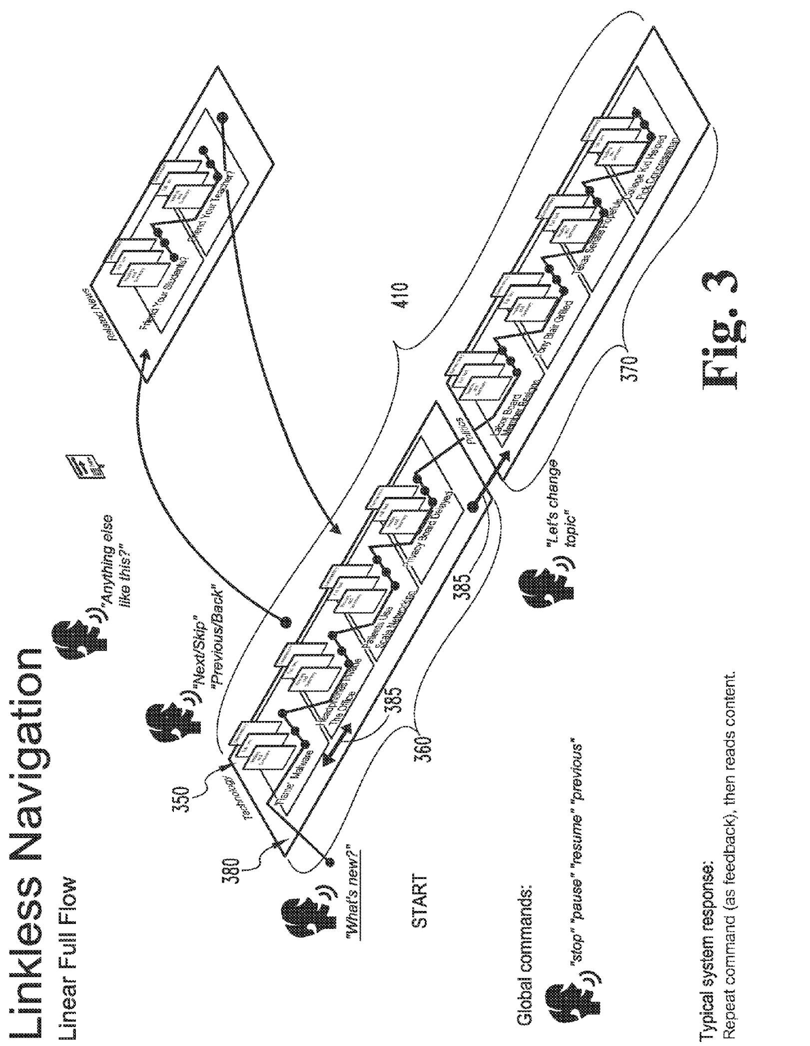
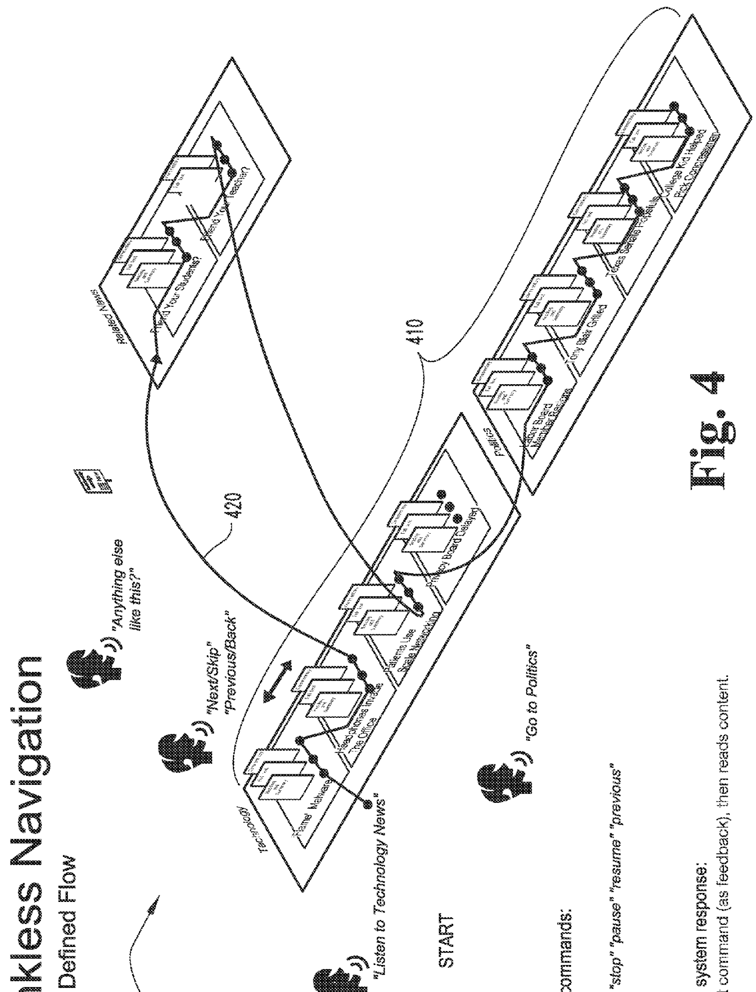

Linkless Navigation
Linear Full Flow
"What's new?"
START
Global commands:
"stop" "pause" "resume" "previous"
Typical system response:
Repeat command (as feedback), then reads content.
380
350
360
385
385
"Next/Skip"
"Previous/Back"
"Anything else like this?"
"Let's change topic"
410
370

Fig. 3



Linkless Navigation
User Defined Flow
400
"Listen to Technology News"
"Next/Skip"
"Previous/Back"
"Anything else like this?"
420
410
Technology
Politics
Related Items
START
"Go to Politics"
Global commands:
"stop" "pause" "resume" "previous"
Typical system response:
Repeat command (as feedback), then reads content.

Fig. 4

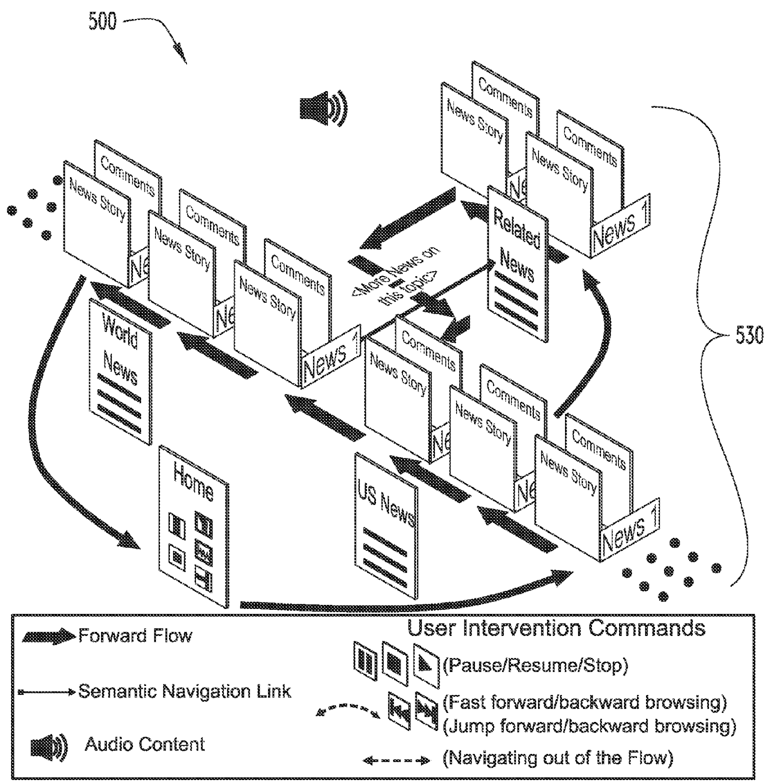

500
530
Comments
News Story
News 1
Related
News
World
News
Home
US News
<More News on this topic>
Forward Flow
Semantic Navigation Link
Audio Content
User Intervention Commands
(Pause/Resume/Stop)
(Fast forward/backward browsing)
(Jump forward/backward browsing)
(Navigating out of the Flow)

Fig. 5

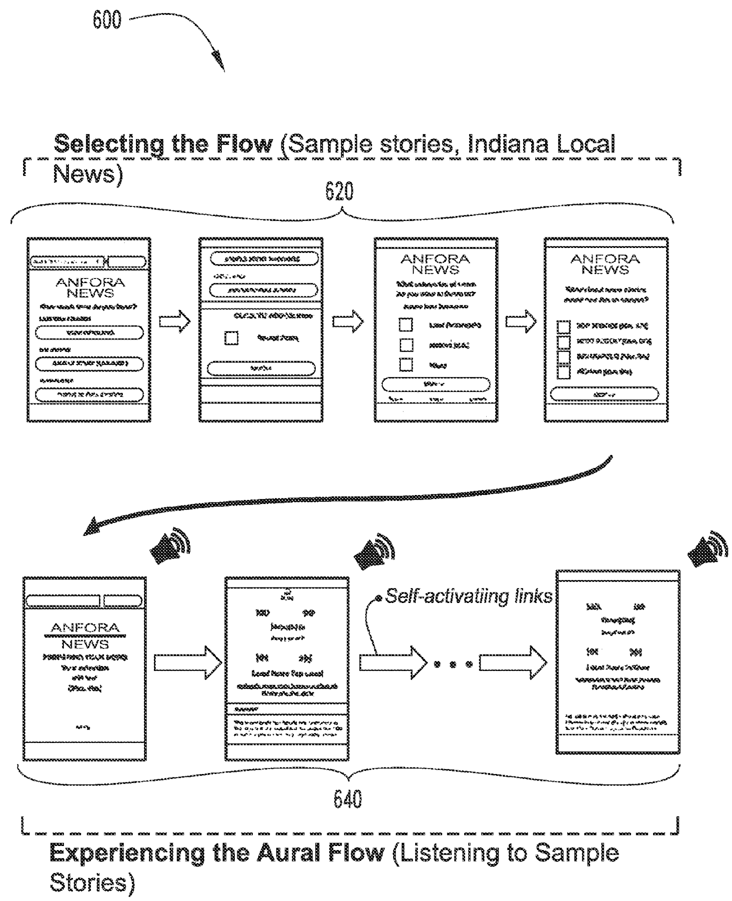
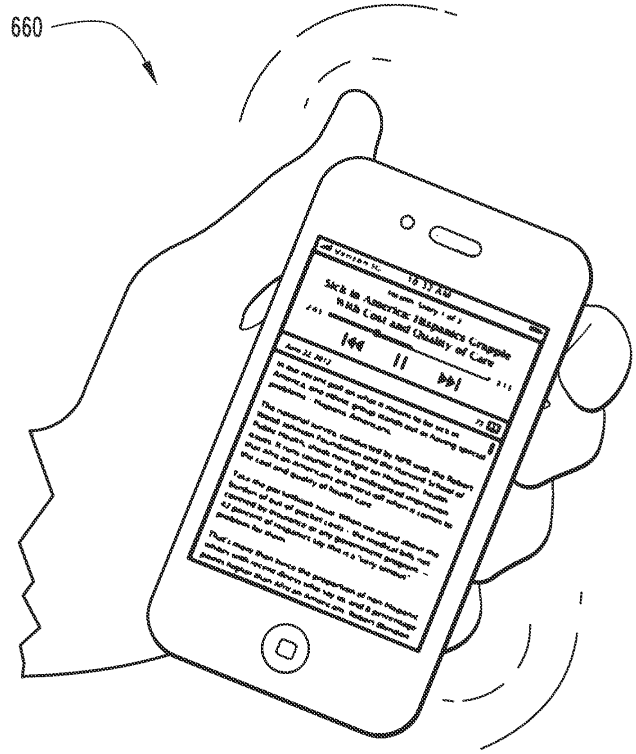

600
Selecting the Flow (Sample stories, Indiana Local News)
620
ANFORA NEWS
ANFORA NEWS
ANFORA NEWS
ANFORA NEWS
Self-activating links
640
Experiencing the Aural Flow (Listening to Sample Stories)

Fig. 6



Accelerometer-based shake gesture to barge-in duing aural flow

Fig. 7

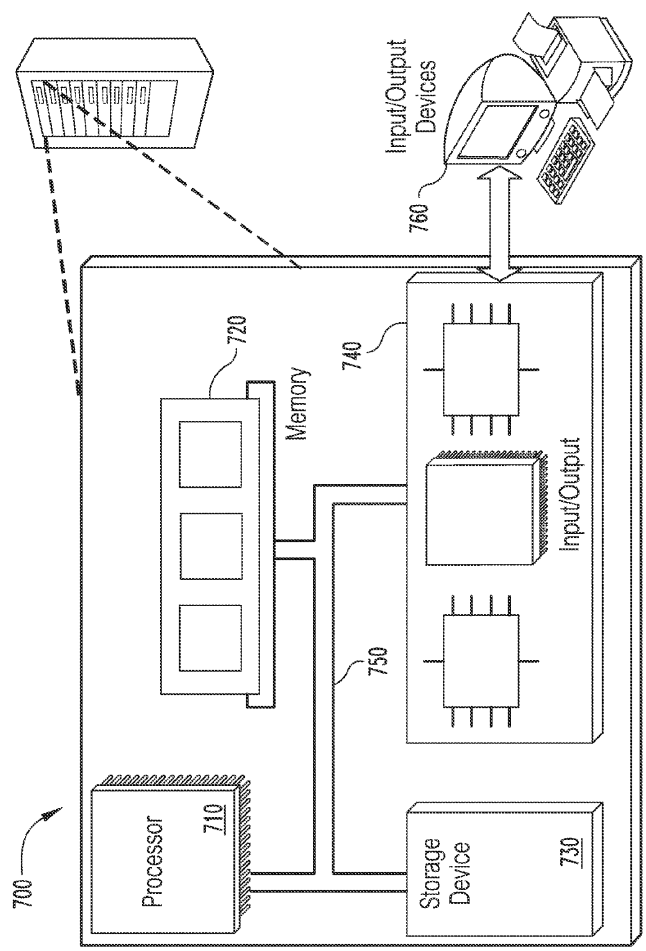

Input/Output
Devices
760
720
Memory
740
Input/Output
750
700
Processor
710
Storage
Device
730

Fig. 8

AURAL NAVIGATION OF INFORMATION RICH VISUAL INTERFACES

[0001] This patent application claims priority to copending U.S. provisional application No. 61/699,748, filed on Sep. 11, 2012 and incorporates the same herein by reference.

BACKGROUND

[0002] This specification relates to navigation of information and content rich interfaces and applications and specifically the navigation of web based interfaces and applications. Accessing the mobile web on-the-go and in a variety of contexts (e.g., walking, standing, jogging, or driving) is becoming more and more pervasive. Mobile users are often engaged in another activity when it is inconvenient, distracting or even dangerous to continuously look at the web display device at all times. Although existing visual user interfaces can be efficient to support quick scanning of a page, they typically require highly focused attention and may not work well or require a dangerous level of attention in certain situations. It is known that the use of audio-based interfaces of mobile and non-mobile devices during secondary tasks are less distracting and demanding when compared to visual interfaces.

[0003] Another concern is the degree of required or desired interactivity with the web application. Continuous or visually detailed interaction with a conventional web interface requires the user to expend visual attention to the web interface. For example, a user is walking on a city street and would like to catch up with the weekly local news during his 10-minute walk to work. Continuous interaction with a conventional news site on your smart phone would force the user to scan the homepage, ascertain the latest news, selecting a category, potentially followed by selecting a subcategory, and then finally select a news story to read. Once read, the user may want to know more about it or select another news story in the same category, etc. Much of this interactivity is in conflict with the current task of the user's walk to work. Furthermore, the effort expended to both walk and visually interact with the web interface likely amounts to an undesirable user experience. Thus, there is a need for an audio-based system of interaction with data rich interfaces. The present invention addresses this need.

SUMMARY

[0004] This specification describes technologies relating to audio based web navigation and audio web content presentation.

[0005] In general, one innovative aspect of the subject matter described in this specification can be embodied in methods that include the actions of generating a model derived from the analysis of user interactions that represents the hierarchical and hypertextual structures of a website and using that model and user supplied constraints to generate a linear aural flow of content from the said website. An audible presentation based on the linear aural flow is then presented to the user with options for the user to dynamically direct and alter the content of the audio presentation.

[0006] Other embodiments of this aspect include corresponding systems, apparatus, and computer programs, configured to perform the actions of the methods, encoded on computer storage devices.

[0007] The details of one or more embodiments of the subject matter described in this specification are set forth in the accompanying drawings and the description below. Other features, aspects, and advantages of the subject matter will become apparent from the description, the drawings, and the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] FIG. **1** is a block diagram of an example environment in which a paradigm for implementing aural navigation flows on rich architectures manages content delivery services.

[0009] FIG. **2** is an example web page such as might be navigated by an aural navigation system.

[0010] FIG. **3** is a block diagram of an aural navigation system's linear full flow of a collection of web pages.

[0011] FIG. **4** is a block diagram of an aural navigation system's user defined flow of a collection of web pages.

[0012] FIG. **5** is a sample block diagram of a group aural flow in a simplified example web architecture.

[0013] FIG. **6** is a representation of a sample user interface for a mobile device that supports aural navigation flows.

[0014] FIG. **7** is a representation of accelerometer-based shake gesture to interact with an aural flow.

[0015] FIG. **8** is a block diagram of a personal computing device capable of implementing a portion or all of the described technology.

[0016] Like reference numbers and designations in the various drawings indicate like elements.

DETAILED DESCRIPTION

[0017] Before the present methods, implementations and systems are disclosed and described, it is to be understood that this invention is not limited to specific synthetic methods, specific components, implementation, or to particular compositions, and as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular implementations only and is not intended to be limiting.

[0018] As used in the specification and the claims, the singular forms "a," "an" and "the" include plural referents unless the context clearly dictates otherwise. Ranges may be expressed in ways including from "about" one particular value, and/or to "about" another particular value. When such a range is expressed, another implementation may include from the one particular value and/or to the other particular value. Similarly, when values are expressed as approximations, for example by use of the antecedent "about," it will be understood that the particular value forms another implementation. It will be further understood that the endpoints of each of the ranges are significant both in relation to the other endpoint, and independently of the other endpoint.

[0019] "Optional" or "optionally" means that the subsequently described event or circumstance may or may not occur, and that the description includes instances where said event or circumstance occurs and instances where it does not. Similarly, "typical" or "typically" means that the subsequently described event or circumstance often though may not occur, and that the description includes instances where said event or circumstance occurs and instances where it does not.

[0020] This application describes a novel, semiinteractive aural paradigm for implementing aural navigation flows on rich architectures enabling users to listen to information-rich interfaces, such as web pages, utilizing complex, hypertextual structures while interacting with the interfaces infrequently. Further, this technology provides for the "aural flow" and investigates of new ways in which different types of aural flow can be applied to conventional information rich architectures such as web pages. An aural flow is a design-driven, concatenated sequence of pages that can be listened to with minimal interaction required. A flow is governed by aural design rules that determine which pages of the information architecture to automatically concatenate and at which point of the flow the user can interact.

[0021] This technology additionally provides the ability to quickly scanning through content-rich data interfaces, such as web pages, allowing effective but time and/or contextual and/or physical constrained scanning. Finally, the described technology provides a generic design framework applicable to any non-linear, content-rich architecture, such as that which underlies modern web systems. For example, the described technology is appropriate for any large website that features hierarchical and hypertextual structures, such as a commerce, travel planning, or tourism site, and the like.

[0022] FIG. **1** is a block diagram of an example environment **100** in which a paradigm for implementing aural navigation flows on rich architectures manages content delivery services. The example environment **100** includes a network **102**, such as a local area network (LAN), a wide area network (WAN), the Internet, or a combination thereof. The network **102** connects websites **104**, user devices **106** (also known as personal computing device), content sponsors (e.g., advertisers **108**), and an aural navigation system advertisement management system **120**. The example environment **100** may include many thousands of websites **104**, user devices **106** and advertisers **108**.

[0023] A website **104** is one or more resources **105** associated with a domain name and hosted by one or more servers. An example website is a collection of web pages formatted in the hypertext markup language (HTML) that can contain text, images, multimedia content and programming elements, such as scripts. Each website **104** is maintained by a publisher/sponsor, which is an entity that controls, manages and/or owns the website **104**.

[0024] A resource **105** is any data that can be provided over the network **102**. A resource **105** is identified by a resource address that is associated with the resource **105**. Resources include HTML pages, word processing documents, and portable document format (PDF) documents, images, video, and feed sources, to name a few. The resources can include content, such as words, phrases, images and sounds, that may include embedded information (such as meta-information in hyperlinks) and/or embedded instructions (such as JavaScript scripts). Units of content that are presented in (or with) resources are referred to as content items.

[0025] A user device **106** is an electronic device that is under control of a user and is capable of requesting and receiving resources over the network **102**. Example user devices **106** include personal computers, mobile communication devices, and other devices that can send and receive data over the network **102**. A user device **106** typically includes a user application, such as a web browser, to facilitate the sending and receiving of data over the network **102**.

[0026] A user device **106** can request resources **105** from a website **104**. In turn, data representing the resource **105** can be provided to the user device **106** for presentation by the user device **106**. The data representing the resource **105** can also include data specifying a portion of the resource or a portion of a user display (e.g., a presentation location of a pop-up window or in a slot of a web page) in which advertisements can be presented. These specified portions of the resource or user display are referred to as slots or advertisement slots.

[0027] To facilitate searching of these resources **105**, the environment **100** can include a search system **112** that identifies the resources **105** by crawling and indexing the resources **105** provided by the publishers on the websites **104**. Data about the resources can be indexed based on the resource **105** to which the data corresponds. The indexed and, optionally, cached copies of the resources **105** are stored in a search index **114**.

[0028] User devices **106** can submit search queries **116** to the search system **112** over the network **102**. In response, the search system **112** accesses the search index **114** to identify resources that are relevant to the search query **116**. The search system **112** identifies the resources in the form of search results **118** and returns the search results **118** to the user devices **106** in search results pages. A search result **118** is data generated by the search system **112** that identifies a resource that is responsive to a particular search query, and includes a link to the resource. An example search result **118** can include a web page title, a snippet of text or a portion of an image extracted from the web page, and the URL of the web page. Search results pages can also include one or more slots in which other content or advertisements can be presented.

[0029] When a resource **105** or search results **118** are requested by a user device **106**, the advertisement management system **110** receives a request for advertisements to be provided with the resource **105** or search results **118**. The request for advertisements can include characteristics of the slots that are defined for the requested resource or search results page, and can be provided to the advertisement management system **110**.

[0030] For example, a reference (e.g., URL) to the resource for which the slot is defined, a size of the slot, and/or media types that are eligible for presentation in the slot can be provided to the advertisement management system **110**. Similarly, keywords associated with a requested resource ("resource keywords") or a search query **116** for which search results are requested can also be provided to the advertisement management system **110** to facilitate identification of advertisements that are relevant to the resource or search query **116**.

[0031] Based on data included in a given request, the advertisement management system **110** selects advertisements or other content that is eligible to be provided in response to the request (e.g., eligible advertisements). For example, eligible advertisements can include advertisements having characteristics matching those of slots and that are identified as relevant to specified resource keywords or search queries **116**. In some implementations, advertisements that have target keywords that match the resource keywords or the search query **116** are selected as eligible advertisements by the advertisement management system **110**.

[0032] A targeting keyword can match a resource keyword or a search query **116** by having the same textual content ("text") as the resource keyword or search query **116**. The relevance can be based, for example, on root stemming, semantic matching, and topic matching. For instance, an advertisement associated with the targeting keyword "hockey" can be an eligible advertisement for an advertisement request including the resource keyword "hockey." Similarly, the advertisement can be selected as an eligible advertisement for an advertisement request including the search query "hockey."

[0033] A targeting keyword can also match a resource keyword or a search query **116** by having text that is identified as being relevant to a targeting keyword or search query **116** despite having different text than the targeting keyword. For example, an advertisement having the targeting keyword "hockey" may also be selected as an eligible advertisement for an advertisement request including a resource keyword or search query for "sports" because hockey is a type of sport, and therefore, is likely to be relevant to the term "sports."

[0034] The Aural navigation system **120** in some implementations provides a generic design framework applicable to any non-linear, content-rich architecture that is depicted in this example environment **100**. The Aural navigation system **120** provides for aural flows that are modeled on top of existing web information and navigation architectures and can co-exist with the traditional navigation and search mechanisms such as depicted in this example environment **100**. In some implementations, the aural navigation system **120** takes the existing structures and linearizes them appropriately for the aural experience, eliminating the need for changes to the existing websites. For example, the aural navigation system **120** can analyze an existing website **104** such as a news website, and linearize the website for audio presentation such that only simple commands are needed by the user to navigate the audio presentation of the content of the news website.

[0035] In some implementations, the aural navigation system **120** can also utilize user directives, past user browsing and audio browsing history, user stated preferences, and other user information such as user location, user online socio presence, and user schedule when linearizing a website for audio presentation. User directives can be thought of as user supplied defaults. For example, for sites that employ popularity ordering of the article, the user can add defaults to instruct the aural navigation system **120** to ignore articles below a certain ranking. As another example, the aural navigation system **120** can analyze a user's past browsing history to determine that the user typically doesn't review sports articles. Using such information, the aural navigation system **120** could neglect the sports content of a news website **104** when linearizing its content for audio presentation to that user. However, the aural navigation system **120** could override the user's past browsing habits upon encountering sports content that has a significant socio connection with the user. One example of a significant socio connection with the user is the sports content referencing a friend of the user.

[0036] In some implementations, the Aural navigation system **120** is able to perceive and respond to user input (oral or otherwise) and such user input is interpreted within the context of the user's session and user's history. Example commands can include "Change to", "Switch to", "Back", or "Previous" which are sensitive to the users' flow history, not a default flow. Most implementations include various forms of bookmarking enabling the continuing of a story or a topic from a previous session. In some implementations, multiple bookmarks can be maintained enabling the user to go back and continue any of several paused stories. In some implementations, the aural navigation system **120** implements a time-based relevance decay enabling past bookmarked articles to eventually lose their bookmark if not referenced after a period of time.

[0037] Other sample commands include but are not limited to: "What's new?", "Anything else (like this)?", "Next" or "Skip", "Stop" or "Pause", "Resume" "Continue" or "Play" "Listen to" "Go to" "Switch to" or "Change to", "More" or "Tell me more", and "Restart" or "Start over". Note that in some implementations, the aural navigation system **120** is implemented by a user device **106**.

[0038] FIG. **2** is an example web page **200** such as might be navigated by an aural navigation system **120**. The example web page **200** is the resource **105**. The example web page **200** includes a title **205**, a search text slot **210**, a search button **215**, a search results container **235** and advertisement slots **230***a*-**230***c*. The search results container **235** contains the search results **118** of a search performed on this resource. In some implementations, the aural navigation system **120** would provide a "linear flow" of the content of web page **200**, contemplating pre-designated page exits while other implementations provide a "user defined flow" enabling user designated exits and/or content expansion.

[0039] FIG. **3** is a block diagram **300** of an aural navigation system's **120** linear full flow **310** of a collection of web pages **104**. In some implementations of a linear full flow **310**, the flow of information in is strictly linear. Users are able to leave the flow **320** for related stories **330**; upon finishing related stories, they are returned **340** to the original flow. They are only able to jump forward and backward. The flow begins with the first story **350** in the first group of topics **360**. Headline, summary and full story are read in that order. Upon finishing the first story, the system will move on to the next story in that group of topics **360**. Upon finishing the last story in a group of topics **360**, the system will move on to the next group of topics **370**.

[0040] In the block diagram **300**, the lines **380** show the default flow. The lines **320**, **340**, and **385** represent where users can interrupt the flow and move to different parts. The system begins with an orientation cue letting the user know which group of topics they are listening to and the position of the current story in the flow (e.g. "World News, Story 1 of 3). As shown, each story contains a headline, summary, full story and optional related stories.

[0041] In some implementations, the aural navigation system **120** can review the user's browsing history, audio browsing history, location, device **106** usage, socio presence, and calendar when generating a linear full flow of a collection of web pages. For example, a user's browsing history may demonstrate a preference for only the top ranked stories from a particular website **104**. As such, the aural navigation system **120** can anticipate the user's continued browsing pattern by generating a linear flow of the webpages corresponding to the user's anticipated preferences. As another example, a user's browsing pattern could be based upon his location. For example, the content that the user wishes to review in the car can be vastly different than the content that the user wants to review when at work.

[0042] FIG. **4** is a block diagram **400** of an aural navigation system's **120** user defined flow **410** of a collection of web pages. In some implementations, the aural navigation system **120** pauses after reading, audibly disclosing, each dialogue (e.g. summary, full story, reader comments) allowing a user to speak a command. Users can interrupt this flow at any time with any command from the vocabulary. In some implementations, users are able to speak the name of a group of topics (e.g. Politics, U.S., World) and begin the flow in that group. As such, in some implementations each category of content from the website being accessed is available as a command. Categories act as keywords to allowing users the freedom to define their own navigation strategy.

[0043] In this example **400**, line **420** indicates a scenario in which a user leaves the flow to listen to related stories and then changes categories during the flow. Each story contains a headline, summary, full story, reader comments, and two related stories. Users are free to navigate the topics as they please.

[0044] FIG. **5** is a sample block diagram **500** of a group aural flow **530** in a simplified example web architecture. Even in this simplified example, the non-linearity typical nature of such information sources is clearly visible. For example, the example contains different organizational structures (e.g., hierarchical and hypertextual).

[0045] In some implementations, the features of the architecture along with the hypertextual connections are modeled through a collection of primitives and notions known in the art as Interactive Dialogue Module (IDM). IDM provides basic concepts to describe and model hypertextual non-linear architectures. IDM is based on the notion that user interaction can be considered a dialogue between the user and the system. In a nutshell, core content entities (e.g., the news) are multiple topics. A multiple topic can be structured in dialogue acts (news story, commentary on the news story) corresponding to different pages or interaction units composing the topic. Multiple topics are typically organized in groups of topics (e.g., U.S. news or world news) at different hierarchical levels. Hypertextual or semantic associations are typed and can be characterized as structural relationships between multiple topics.

[0046] Using IDM, one or more aural flows are modeled on top of existing web information and navigation architectures as represented by IDM. Thus, the aural flows can co-exist with the traditional navigation and interaction paradigm. As a more complete explanation, an aural flow can be thought of as a design-driven, concatenated sequence of web pages that can be listened to with minimal interaction. The flow is governed by aural design rules that determine which pages of the information architecture to automatically concatenate and at which point of the flow the user can interact. Such design rules can be proposed and refined through various machine learning statistical techniques. For example, concatenation rules can be derived from topic popularity as determined by related topic page hits. Or as another example, statistical models can be derived from topic popularity measures and web activity measures. Similar to predicting conversion for a sales event, the popularity measures and activity measures can be used to derive a "conversion-like" predictor capable of providing a predictive expectation value for topic popularity.

[0047] In some implementations, the user is presented with two flow patterns. The user may either follow the Default Full Flow with little to no interaction, or they may navigate where they please within the flow, creating their own User-Defined Flow. The Default Full Flow, unless interrupted by the user, follows a linear, concatenated flow of information. Typical implementations provide the headline and summary of articles and then provide a portion of the content based upon the aural flow rules. For example, the aural flow rules could provide the full content along with the commentary. The flow continues for each content or story deemed to be above a certain threshold of interest or automatically included by a default behavior. Upon finishing a content, the system will move on to the next content in that group of topics. Upon finishing the last story in a group of topics, the system will move on to the next group of topics. The next group of topics can be based upon the underlying web architecture, derived interest rules (where a topic may have a perceived higher interest than another), or from user derived interest rules (such rules can be derived from previous user actions or directly obtained through a user initiation where the user acts provide rules to govern topic interest).

[0048] However, a user can interrupt the default flow at any time with command from the vocabulary (e.g. "stop" or "change to"). They may navigate wherever they please, at any time they want. This freedom of control creates a User-Defined Flow. An important feature of this flow type is that the system will keep track of a user's history and context during each session. For example, saying a command like "Previous" will take them to the last story they heard, not the previous story in the default flow. In some implementations, the User-Defined Flow still follows the order of the Default Full Flow until a user utters a command. A table of example commands and their respective actions are presented below. In most implementations, users have at least four basic categories of interaction. The four categories are a) Pause, resume, replay and stop: The user can pause and resume the flow. The same dialogue act can be replayed from the beginning. The user can also stop the flow to go back to the home page, b) Fast forward/backward browsing: The user can fast forward to go to the next dialogue act of the same topic or fast backward to go back to the previous dialogue act of the same topic, c) Jump forward/backward browsing: The user can jump forward to the next topic or jump backward to the previous one at any time, d) Navigating out of the flow: The user might want to listen to the related topic by clicking on its link. This action breaks the current flow and moves outside the flow to the desired content (e.g., Related News).

[0049] Note that some implementations provide for a preliminary input or presentation guiding input from a user. Table 2 provides an example of the different characteristics of aural flow types. This preliminary input enables the system to tailor the aural flow to the user's current expectations and/or limitations. For example, in some implementations the user can tell the system an amount of time that the user has with which to listen. For example, the user can tell the system that he has 20 minutes, in which the aural flow through possible content will be streamlined. As another example, a user after choosing a main group of topic, such as U.S. news, could listen to all of the headlines or story summaries in that category. Users would be able to navigate through all the news stories in one category and continue the flow with the next category of news or related stories.

[0050] It has been observed that two sources of error account for a sizable portion of the errors between user and aural flow interaction. The two sources of errors are speech recognition errors and navigation errors. Recognition errors occur when the system either does not understand the uttered command, or the uttered command is not within the scope of the command vocabulary. In some implementations, recognition errors are handled through the notification of the user by the system emitting an earcon, a distinct and noticeable. After being such notified, the user can then reissue the command or issue a different command.

[0051] It is worth mentioning that some implementations provide for a hybrid interface consisting of both the audio presentation along with a visual interface dynamically cued to the content of the current flow. Such hybrid interfaces enable the "At-a-glance" visual confirmation of content. Additionally, such implementations provide for a more extensive visual coverage of the current topic. Such implementations typically provide an interactive mechanism, for example, a swipe on the personal computing device's touch sensitive screen, to visually provide the full coverage of the topic currently being disclosed.

[0052] Navigation errors occur when a user utters a command that is not applicable in the current part of the flow (e.g. saying "Forward" while in the commentary"). These errors should be handled with audio orientation cues provided by the system (e.g. "There is no more content for this story"). In some implementations, the system responds by reverting back to a default flow. Alternatively, some implementations respond by audibly providing a shortened menu of possible actions based upon the user's current location in the flow.

[0053] FIG. **6** is a representation of a sample user interface **600** for a mobile device that supports aural navigation flows. The aural flow experience consists of two main components, as highlighted in the figure: Selecting the flow **620** and experiencing the flow **640**. In Selecting the flow **620**, the system provides to the user several options to choose and customize the coverage of the available content, based on time constraints, types of aural flow and user's interest. A simple sequence of user interface screens is shown supporting this selection task. Once the user has selected values for such simple parameters, the system immediately generates and makes available the aural flow corresponding to the user's selection. At this point, the user enters the Experiencing the flow **640** part, in which the system plays the aural flow, which concatenates the web pages through self-activating links.

[0054] FIG. **7** is a representation **660** of accelerometer-based shake gesture to interact with an aural flow. In some implementations, the aural flow can be interacted and altered through vocal and/or tactile user actions and/or a locational value of the personal computing device. In such implementations, activating a microphone will temporarily stop the system output and activate the "listening mode." During this pause, the system will wait for a command. If the button is released with no command having been uttered, the system will simply resume its output. If a command was uttered and understood by the system, the system will react accordingly. Shaking the personal computing device and utilizing the accelerometer to activate the listening mode works similarly to curing the microphone.

[0055] The locational input utilizes a geographical positional system component that is typical of many personal computing devices. However, the locational input functionality differs from the input of direct user actions. Locational input is typically configured by the user to respond in certain ways to locational values. For example, a user could configure the system such that the content, as presented upon arriving at the user's place of employment, consist of the latest topics on the company's intranet. As another example, a user could request that the content be based upon the user's geographic position.

[0056] FIG. **8** is a block diagram of a personal computing device **700** capable of implementing a portion or all of the described technology. The example of one such type of personal computing device **700** shows a block diagram of a programmable processing system (system) **700** suitable for implementing apparatus or performing methods of various aspects of the subject matter described in this specification. The system **700** includes a processor **710**, a random access memory (RAM) **721**, a program memory **730** (for example, a writable read-only memory (ROM) such as a flash ROM) and an input/output (I/O) controller **740** (typically endowed with GPS capability) coupled with a bus **750**. The system **700** can be preprogrammed, in ROM, for example, or it can be programmed (and reprogrammed) by loading a program from another source (for example, downloaded from an application site, or another personal computing device).

[0057] The I/O controller **740** is operationally connected to I/O interfaces **760**. The I/O interface receives and transmits data (e.g., stills, pictures, movies, and animations for importing into a composition) in analog or digital form over communication links such as a serial link, local area network, wireless link, and parallel link, cellular, touch and shake inputs, geographic locational input, and the like.

TABLE 1

Sample system navigation and commands (primitives)

Command	System	System Action
What's new?"	"Recent stories in (topic) news"	Begin default full flow
"Anything else (like this)?" "More like this"	"Related stories"	Go to related stories
"Next" or "Skip"	"Next story"	Go to next story
"Previous" or "Back"	"Previous story"	Go to previous story in user history
"Stop" or "Pause"	Earcon	Pauses story
"Resume" "Continue" or "Play"	"Resuming (headline)"	Resumes story
"Listen to" "Go to" "Switch to" or "Change to"	"Switching to (topic) news"	Switch to selected topic
"Forward" or "Rewind"	Title of next section is read	Move between sections within a story
"Restart" or "Start over"	"Restarting (reads headline)"	Restarts story

TABLE 2

example of the different characteristics of aural flow types

Flow	Characteristics	Time	Advantages	Disadvantages
Group	A selected group of topics	5 min	Decide the category from the outset	Interact every time to select a different category
Full	All groups of topics	Longer period of time - 30 min.	Less interaction	Difficulty building mental model
Deep	All groups of topics + semantic associations	Longer period of time - 1 hr.	In-depth coverage of content	Difficulty building mental model
Light	Agile overview of each topic (default dialogue act)	Shorter period of time	More stories in less time (agile overview)	Details of each topic will not be played
Rich	Extensive coverage of each topic (all dialogue acts)	Longer period of time	Extensive coverage	Time-consuming and constraining

[0058] Embodiments of the subject matter and the operations described in this specification can be implemented as a method, in digital electronic circuitry, or in computer software, firmware, or hardware, including the structures disclosed in this specification and their structural equivalents, or in combinations of one or more of them. Embodiments of the subject matter described in this specification can be implemented as one or more computer programs, i.e., one or more modules of computer program instructions, encoded on computer storage medium for execution by, or to control the operation of, data processing apparatus. Alternatively or in addition, the program instructions can be encoded on an artificially-generated propagated signal, e.g., a machine-generated electrical, optical, or electromagnetic signal that is generated to encode information for transmission to suitable receiver apparatus for execution by a data processing apparatus. A computer storage medium can be, or be included in, a computer-readable storage device, a computer-readable storage substrate, a random or serial access memory array or device, or a combination of one or more of them. Moreover, while a computer storage medium is not a propagated signal, a computer storage medium can be a source or destination of computer program instructions encoded in an artificially-generated propagated signal. The computer storage medium can also be, or be included in, one or more separate physical components or media (e.g., multiple CDs, disks, or other storage devices).

[0059] The operations described in this specification can be implemented as operations performed by a data processing apparatus on data stored on one or more computer-readable storage devices or received from other sources.

[0060] The term "data processing apparatus" encompasses all kinds of apparatus, devices, and machines for processing data, including by way of example a programmable processor, a computer, a system on a chip, or multiple ones, or combinations, of the foregoing The apparatus can include special purpose logic circuitry, e.g., an FPGA (field programmable gate array) or an ASIC (application-specific integrated circuit). The apparatus can also include, in addition to hardware, code that creates an execution environment for the computer program in question, e.g., code that constitutes processor firmware, a protocol stack, a database management system, an operating system, a cross-platform runtime environment, a virtual machine, or a combination of one or more of them. The apparatus and execution environment can realize various different computing model infrastructures, such as web services, distributed computing and grid computing infrastructures.

[0061] A computer program (also known as a program, software, software application, script, or code) can be written in any form of programming language, including compiled or interpreted languages, declarative or procedural languages, and it can be deployed in any form, including as a stand-alone program or as a module, component, subroutine, object, or other unit suitable for use in a computing environment. A computer program may, but need not, correspond to a file in a file system. A program can be stored in a portion of a file that holds other programs or data (e.g., one or more scripts stored in a markup language document), in a single file dedicated to the program in question, or in multiple coordinated files (e.g., files that store one or more modules, sub-programs, or portions of code). A computer program can be deployed to be executed on one computer or on multiple computers that are located at one site or distributed across multiple sites and interconnected by a communication network.

[0062] The processes and logic flows described in this specification can be performed by one or more programmable processors executing one or more computer programs to perform actions by operating on input data and generating output. The processes and logic flows can also be performed by, and apparatus can also be implemented as, special purpose logic circuitry, e.g., an FPGA (field programmable gate array) or an ASIC (application-specific integrated circuit).

[0063] Processors suitable for the execution of a computer program include, by way of example, both general and special purpose microprocessors, and any one or more processors of any kind of digital computer. Generally, a processor will receive instructions and data from a read-only memory or a random access memory or both. The essential elements of a computer are a processor for performing actions in accordance with instructions and one or more memory devices for storing instructions and data. Generally, a computer will also include, or be operatively coupled to receive data from or transfer data to, or both, one or more mass storage devices for storing data, e.g., magnetic, magneto-optical disks, or optical disks. However, a computer need not have such devices. Moreover, a computer can be embedded in another device, e.g., a mobile telephone, a personal digital assistant (PDA), a mobile audio or video player, a game console, a Global Positioning System (GPS) receiver, or a portable storage device (e.g., a universal serial bus (USB) flash drive), to name just a few. Devices suitable for storing computer program instructions and data include all forms of non-volatile memory, media and memory devices, including by way of example semiconductor memory devices, e.g., EPROM, EEPROM, and flash memory devices; magnetic disks, e.g., internal hard disks or removable disks; magneto-optical disks; and CD-ROM and DVD-ROM disks. The processor and the memory can be supplemented by, or incorporated in, special purpose logic circuitry.

[0064] To provide for interaction with a user, embodiments of the subject matter described in this specification can be implemented on a computer having a display device, e.g., a CRT (cathode ray tube) or LCD (liquid crystal display) monitor, for displaying information to the user and a keyboard and a pointing device, e.g., a mouse or a trackball, by which the user can provide input to the computer. Other kinds of devices can be used to provide for interaction with a user as well; for example, feedback provided to the user can be any form of sensory feedback, e.g., visual feedback, auditory feedback, or tactile feedback; and input from the user can be received in any form, including acoustic, speech, or tactile input. In addition, a computer can interact with a user by sending documents to and receiving documents from a device that is used by the user; for example, by sending web pages to a web browser on a user's client device in response to requests received from the web browser.

[0065] Embodiments of the subject matter described in this specification can be implemented in a computing system that includes a back-end component, e.g., as a data server, or that includes a middleware component, e.g., an application server, or that includes a front-end component, e.g., a client computer having a graphical user interface or a Web browser through which a user can interact with an implementation of the subject matter described in this specification, or any combination of one or more such back-end, middleware, or front-end components. The components of the system can be interconnected by any form or medium of digital data communication, e.g., a communication network. Examples of communication networks include a local area network

("LAN") and a wide area network ("WAN"), an inter-network (e.g., the Internet), and peer-to-peer networks (e.g., ad hoc peer-to-peer networks).

[0066] The computing system can include clients and servers. A client and server are generally remote from each other and typically interact through a communication network. The relationship of client and server arises by virtue of computer programs running on the respective computers and having a client-server relationship to each other. In some embodiments, a server transmits data (e.g., an HTML page) to a client device (e.g., for purposes of displaying data to and receiving user input from a user interacting with the client device). Data generated at the client device (e.g., a result of the user interaction) can be received from the client device at the server.

[0067] While this specification contains many specific implementation details, these should not be construed as limitations on the scope of any inventions or of what may be claimed, but rather as descriptions of features specific to particular embodiments of particular inventions. Certain features that are described in this specification in the context of separate embodiments can also be implemented in combination in a single embodiment. Conversely, various features that are described in the context of a single embodiment can also be implemented in multiple embodiments separately or in any suitable subcombination. Moreover, although features may be described above as acting in certain combinations and even initially claimed as such, one or more features from a claimed combination can in some cases be excised from the combination, and the claimed combination may be directed to a subcombination or variation of a subcombination.

[0068] Similarly, while operations are depicted in the drawings in a particular order, this should not be understood as requiring that such operations be performed in the particular order shown or in sequential order, or that all illustrated operations be performed, to achieve desirable results. In certain circumstances, multitasking and parallel processing may be advantageous. Moreover, the separation of various system components in the embodiments described above should not be understood as requiring such separation in all embodiments, and it should be understood that the described program components and systems can generally be integrated together in a single software product or packaged into multiple software products.

[0069] Thus, particular embodiments of the subject matter have been described. Other embodiments are within the scope of the following claims. In some cases, the actions recited in the claims can be performed in a different order and still achieve desirable results. In addition, the processes depicted in the accompanying figures do not necessarily require the particular order shown, or sequential order, to achieve desirable results. In certain implementations, multitasking and parallel processing may be advantageous.

What is claimed is:

1. A method comprising:

generating, by a computer, a model of a website using user interaction primitives to represent hierarchical and hypertextual structures of the website;

generating, by the computer, a linear aural flow of content of the website based upon the model and a set of user constraints;

audibly presenting, by the computer, the linear aural flow of the content such that the linear aural flow of content is controlled through the use of user supplied primitives, wherein, the linear aural flow can be turned into a dynamic aural flow based upon the user supplied primitives.

2. The method of claim **1** wherein user supplied primitives comprises a spoken command.

3. The method of claim **1** wherein the linear aural flow is further based on a ranking of current topics based upon each topic's page hits the website has received.

4. The method of claim **2** wherein interrupted audibly presented content is bookmarked such that the bookmark ages over a user stated period and is eliminated upon an ending of a user stated period.

5. The method of claim **1** wherein the set of user constraints is derived from a user's past audio browsing history in conjunction with the device used to perform the past audio browsing.

6. The method of claim **1** wherein the user supplied primitives are interpreted in context of a user's session.

7. The method of claim **1** wherein the linear aural flow sequences individual articles into dialogues for audio presentation including a dialog for an article's headline, a dialog for the article's summary, and a dialog for the article's content.

8. The method of claim **2** wherein a spoken command is a name of a category of content available on the website.

9. The method of claim **1** wherein the set of user constraints is derived from popularity measures of articles present on the website.

10. A computer storage medium encoded with a computer program, the program comprising instructions that when executed by a user device cause the user device to perform operations comprising:

receiving a model of a website, the model representing hierarchical and hypertextual structures of the website, wherein the model uses user interaction primitives to represent the hierarchical and the hypertextual structures of the website;

receiving a set of user derived constraints;

generating a linear aural flow of content of the website based upon the model and a set of user derived constraints;

audibly presenting the linear aural flow of the content;

determining whether a user command indicates a desire for a dynamic aural flow;

upon determining that a user command indicates a desire for a dynamic aural flow, audibly presenting a dynamic aural flow.

11. The method of claim **10**, wherein interrupted audibly presented content is bookmarked such that the bookmark ages over a user stated period and is eliminated upon an ending of a user stated period.

12. The method of claim **10** wherein the set of user constraints is derived in part from a user's past audio browsing history in conjunction with the device used to perform the past audio browsing.

13. A system comprising:

a user device;

one or more computers operable to interact with the device;

instructions stored on a machine readable storage device for execution by the one or more computers, wherein upon execution the instructions cause the one or more computers to perform the operations of:

generate a model of a website, the model representing hierarchical and hypertextual structures of the website through usage of user interaction primitives;

generate a linear aural flow of content of the website based upon the model and a set of user constraints;

provide instructions to the user device causing the user device to audibly present the linear aural flow of the content;

upon receiving input from a user, provide instructions to the user device causing the user device to audibly present a dynamic aural flow of the content.

14. The system of claim **13**, wherein the one or more computers comprise the user device.

15. The system of claim **13**, wherein the linear aural flow is further based on a ranking of current topics based upon each topic's page hits the website has received.

16. The system of claim **13**, wherein the one or more computers comprise a server operable to interact with the device through a data communication network, and the user device is operable to interact with the server as a client.

17. The system of claim **13**, wherein the one or more computers consist of one computer, the user device is a user interface device, and the one computer comprises the user interface device.

* * * * *

Appendix E: *Linkless* ANFORA Prototypes

Button condition: http://discern.uits.iu.edu:8670/ANFORA_B/

Voice + Button condition: http://discern.uits.iu.edu:8670/ANFORA_VB/

Control console to manually activate voice commands:
http://discern.uits.iu.edu:8670/ANFORA_VB/admins/home

Appendix F: *Linkless* ANFORA Evaluation Study Instruments and Scripts

Introductory Script

Thank you for agreeing to participate in our research project. For this project, we have developed a working prototype mobile news application called ANFORA News that will allow you to listen to news stories while on the go. ANFORA News is designed to allow you to customize your news experience, by choosing the categories of news stories you would like to listen to based on your interests.

Imagine that you want to browse a news website using your mobile phone. But you also must walk to work or class, which may make it difficult to visually browse a news site and read stories while you walk. ANFORA News allows you to select the category of news you want to listen to before beginning your walk. Once your selections have been made, ANFORA News starts playing the requested playlist of those stories and allows you to listen to them, one after another, without further visual interaction with the screen. In other words, ANFORA News provides a customized, eye-free news listening experience.

We are studying different ways to interact with ANFORA News, and would like to collect your feedback regarding your experience with this application. You will be asked to complete two simple tasks focused on interaction with the ANFORA News interface while walking in the halls of the USER LAB at Walker Plaza building at IUPUI. We will join you on your walk to observe your interactions with the interface, video record your session and help you if any technical problems should arise. I will ask you a series of questions regarding your experience after each interaction with the ANFORA prototype. The entire session should last about 90 minutes.

You can use either button controls on the screen or voice commands to interact with the interface. The button control commands should be self-explanatory. We will spend 30 minutes training you on how to use both button and voice control commands.

Training Tasks

Step 1: This is the interface of ANFORA News prototype. For the purposes of this study, we have chosen a number of news categories from the NPR website from which playlist of News will be created. You can start listening to this playlist by selecting any category that you are interested in, such as U.S., world, technology, sports, health, science, economy, and politics. If you don't select any category, the order of the news stories will be U.S. News, World, politics, sports, technology, health, science, and economy. The volume level has been predetermined, so you will not need to make changes to it. In this part of training you can use both buttons and voice commands to interact with the application and listen to the news. When you want to issue a voice command, touch the cord from your earphones, and say the command you wish to use. For e.g., Touch the cord and say "What's New today?" Then, I will simulate the response to your voice command using my controls. Each news story has two sections, a story summary and the full story. In this application, by default, you can listen to both sections. But if you are interested in listening to one or the other, you can use a voice command to indicate that.

Here is the list of voice commands; you can review it for a few minutes (2-3 minutes) (Hands the list to the participant)

Now we are on the home page.

- Now if you want to start with your playlist, say "**What's New?", or "Start" or "Recent."**
- To pause the playlist, please click on the pause button in the interface.
- If you want to go back to the home page, say **"Home".**
- We have 8 different categories of news in our application: U.S., World, Politics, Sports, Technology, Health, Science and Economy. You can select any of them by saying the name of the section. If you want to listen to a specific category of news just say the name of the category you want to listen to. For example, if you want to listen to sports news, simply say, "Sports."
- Please pause.

- Once you are listening to a playlist of news stories, you can go to the next story just by saying "**Next" or "Skip"**.
- Imagine that you have listened to a few news stories and realize there was a story that you want to listen to again. You can go back to that story by saying **"Previous"** or **"Back".**
- At any time during your listening experience, you can change the current category of news to another one that you are more interested in by saying the category name. For example, if you are currently listening to U.S. and you want to switch to Sports News, simply say Sports.
- If you decide you are not interested in a particular news story – meaning both a story summary and a full story – you can say **"Forward"** to go to the next section of a story.
- You can say **"Rewind"** to go to the previous section.
- At any time, if you decide to go back to the beginning of the playlist, say **"Restart".**
- Imagine you are listening to one news story and you decide you are interested in listening to more stories on that topic if they are available on the NPR website. To do that, you can say any of the following: **"Anything Else?", "More", "Tell Me More", "Like This"** or **"Related".** Try saying, **"More"** and see what happens.
- We have now explored all of the voice commands you can use to interact with ANFORA News.
- For your convenience the list of voice commands and the news categories are available on both the walls of the hallway**.**

Step 2: If you want, you can review the list of voice commands for a few more minutes. (2-3 minutes)

Step 3: We will no longer tell you what to say. Feel free to use any of the commands you have learned to interact with the application. Now, please start listening to U.S. News. If at any time you face any difficulty, I am here to assist you (5 minutes)

Step 4: In this part of the training, you may only use buttons to interact with the application.

- To begin your playlist, click on the button. Click the checkbox next to U.S. News and then click submit. You can use this button to switch to another category of news at any time.
- In the middle of the story, you may decide you want to stop listening or listen to the news later. In this case, click .
- To continue listening to a story, click .
- Once you are in the flow, you can go to the next story by clicking on the next button .
- Imagine that you have listened to a few news stories and realize there was a story that you want to listen to again. You can go back to that story by clicking on the previous button .
- Imagine you are listening to one news story and you decide you are interested in listening to more stories on that topic if they are available on the NPR website. To do this, you can click on Related.
- If you decide you are not interested in a section of news, you can click Summary or Full Story to go to one of those sections.
- At any time, if you decide to go back to the beginning of the playlist, click RESTART.

Step 5: Please start listening to U.S. news, and use the button control commands to interact with it. (5 minutes)

Task List

NOW you are going to walk and use ANFORA News on the go. Let me first show you the path and the list of voice commands on the wall. Please follow me.

Task List – Button Control Commands

1. In this version, you may navigate using **button control commands**. You have 15 minutes to use ANFORA. Please browse at least 8 news stories during this

time period and change the category once. Try not to listen to the category of news you already listened to. From the home screen, start listening to any news category you like. We will stop you after 15 min.

Task List – Button Control Commands + Voice Commands

1. In this version, you may navigate using either **voice or button control commands**. You have 15 minutes to use ANFORA. Please browse at least 8 news stories during this time period and change the category once. Try not to listen to the category of news you already listened to. From the home screen, start listening to any news category you like. We will stop you after 15 min.

ANFORA News Questionnaire

Thank you for your participation in today's study. Now, please take a few moments to answer the following demographic questions. Circle your answers.

Demographic Questions

Age: _____ **Gender:** M F

What kind of mobile phone do you own? _______________

How much time to you spend on news websites each week? (circle the answer that most closely applies to you.)

No time
5-30 minutes
30-60 minutes
1-3 hours
3-6 hours
More than 6 hours

How much of this time, do you use your mobile to access the news?

No time
5-30 minutes
30-60 minutes
1-3 hours
3-6 hours
More than 6 hours

How much time do you spend listening to news on the radio each week?

No time
5-30 minutes
30-60 minutes
1-3 hours
3-6 hours
More than 6 hours

How much time do you spend watching news on TV each week?

No time
5-30 minutes
30-60 minutes
1-3 hours
3-6 hours
More than 6 hours

ANFORA News Questionnaire	B

On a scale of 1-5 (1 = strongly disagree, 5 = strongly agree) rate your level of agreement with the following statements:

	strongly disagree			neutral			strongly agree
1. I think that I would like to use this application (ANFORA News) frequently.	1	2	3	4	5	6	7
2. I found this application unnecessarily complex.	1	2	3	4	5	6	7
3. I thought this application was easy to use.	1	2	3	4	5	6	7
4. I think that I would need assistance to be able to use this application.	1	2	3	4	5	6	7
5. I found the various functions in this application were well integrated.	1	2	3	4	5	6	7
6. I thought there was too much inconsistency in this application.	1	2	3	4	5	6	7
7. I would imagine that most people would learn to use this application very quickly.	1	2	3	4	5	6	7
8. I found this application very cumbersome/awkward to use.	1	2	3	4	5	6	7
9. I felt very confident using this application.	1	2	3	4	5	6	7
10. I needed to learn a lot of things before I could get going with this application.	1	2	3	4	5	6	7
11. I felt comfortable navigating through this application.	1	2	3	4	5	6	7
12. I felt frustrated using this application.	1	2	3	4	5	6	7
13. I felt this application is enjoyable to use.	1	2	3	4	5	6	7

	Unsatisfactory						Satisfactory
14. Rate your experience using button control commands.	1	2	3	4	5	6	7
	Annoying						Pleasing
15. Rate your experience using button control commands.	1	2	3	4	5	6	7
	Unenjoyable						Enjoyable
16. Rate your experience using button control commands.	1	2	3	4	5	6	7
	Difficult						Simple
17. Rate your experience using button control commands.	1	2	3	4	5	6	7
	Confusing						Easy to understand
18. Rate your experience using button control commands.	1	2	3	4	5	6	7
	Boring						Engaging
19. Rate your experience using button control commands.	1	2	3	4	5	6	7

ANFORA News Questionnaire B

	Very Low — Very High
20. How mentally demanding was the task?	
21. How physically demanding was the task?	
22. How hurried or rushed was the pace of the task?	
23. How successful were you in accomplishing what you were asked to do?	
24. How hard did you have to work to accomplish your level of performance?	
25. How insecure, discouraged, irritated, stressed, and annoyed were you?	

ANFORA News Questionnaire V & B

On a scale of 1-5 (1 = strongly disagree, 5 = strongly agree) rate your level of agreement with the following statements:

	strongly disagree			neutral			strongly agree
1. I think that I would like to use this application (ANFORA News) frequently.	1	2	3	4	5	6	7
2. I found this application unnecessarily complex.	1	2	3	4	5	6	7
3. I thought this application was easy to use.	1	2	3	4	5	6	7
4. I think that I would need assistance to be able to use this application.	1	2	3	4	5	6	7
5. I found the various functions in this application were well integrated.	1	2	3	4	5	6	7
6. I thought there was too much inconsistency in this application.	1	2	3	4	5	6	7
7. I would imagine that most people would learn to use this application very quickly.	1	2	3	4	5	6	7
8. I found this application very cumbersome/awkward to use.	1	2	3	4	5	6	7
9. I felt very confident using this application.	1	2	3	4	5	6	7
10. I needed to learn a lot of things before I could get going with this application.	1	2	3	4	5	6	7
11. I felt comfortable navigating through this application.	1	2	3	4	5	6	7
12. I felt frustrated using this application.	1	2	3	4	5	6	7
13. I felt this application is enjoyable to use.	1	2	3	4	5	6	7

	Unsatisfactory						Satisfactory
14. Rate your experience using both voice and button control commands.	1	2	3	4	5	6	7
	Annoying						**Pleasing**
15. Rate your experience using both voice and button control commands.	1	2	3	4	5	6	7
	Unenjoyable						**Enjoyable**
16. Rate your experience using both voice and button control commands.	1	2	3	4	5	6	7
	Difficult						**Simple**
17. Rate your experience using both voice and button control commands.	1	2	3	4	5	6	7
	Confusing						**Easy to understand**
18. Rate your experience using both voice and button control commands.	1	2	3	4	5	6	7
	Boring						**Engaging**
19. Rate your experience using both voice and button control commands.	1	2	3	4	5	6	7

ANFORA News Questionnaire V & B

	Very Low — Very High
20. How mentally demanding was the task?	
21. How physically demanding was the task?	
22. How hurried or rushed was the pace of the task?	
23. How successful were you in accomplishing what you were asked to do?	
24. How hard did you have to work to accomplish your level of performance?	
25. How insecure, discouraged, irritated, stressed, and annoyed were you?	

ANFORA News Interview Questions

Based on your experience today with ANFORA News,

- Overall, how would you describe your experience with ANFORA News?
- Did you listen to the news while walking?

    If yes, were you still able to adequately monitor your surroundings?

    If no, why not?
- Was it clear when a new news story started?
- Was it clear when a new news story ended?
- At any point, did you feel confused by the interface? If so, can you recall when?
- At any point, did you feel lost in the while listening to the news? If so, can you recall when?
- At any point, did you stop ANFORA News before your playlist ended?

    If yes, why?
- When you used the version in which both voice control and buttons were available, did you use voice commands?

    Why or why not?
- Do you prefer to use voice commands, label clicking or a combination of the two? Why?
- Were the homepage instructions clear about how to get started with the application? If not, what part was confusing/unclear to you?
- If ANFORA News were available today, when would you use it? How? Why or why not?
- What did you like best about ANFORA News?
- What did you like least about ANFORA News?
- How many news stories did you listen to today?
- Tell me briefly about one news story that struck you.
- Do you have anything else you would like to share about your experience with ANFORA News?

Appendix G: ANFORA*Drive* Evaluation Study Instruments and Scripts

MINI-MENTAL STATE EXAMINATION (MMSE) – 6-Item Screener

I have some preliminary set of questions before we get started.

*Instructions:

1. Say to the participant the words "apple", "table" and "penny".
2. Have them repeat the words and tell them to remember the words.
3. Go through the first three questions.
4. Have the patient recall the three words.
5. Record Total Score.

Questions:	**Response**	**Score** * Score one point for each correct answer
What day of the week is today?		
What month is it?		
What year is it?		
*Recall the first word (apple)		
*Recall the second word (table)		
*Recall the third word (penny)		

Introductory Script

Thank you for agreeing to participate in our research project. For this project, we have developed a working prototype mobile news application called ANFORADrive that will allow you to listen to news stories while driving. ANFORADrive is designed to allow you to customize your news experience, by choosing the categories of news stories you would like to listen to based on your interests.

Imagine that you want to browse a news website using your mobile phone. But you also must drive to work or class, which may make it difficult to visually browse a news site and read stories while you drive. ANFORADrive allows you to select the category of news you want to listen to before beginning your drive. Once your selections have been made, ANFORADrive starts playing the requested playlist of those stories and allows you to listen to them, one after another, without further visual interaction with the screen. In other words, ANFORADrive provides a customized, eye-free news listening experience.

We are studying different ways to access and interact with ANFORADrive, and compare it with another existing application called Umano. We would like to collect your feedback regarding your experience with these two applications. You will be asked to complete two simple tasks focused on interaction with the ANFORADrive interface and Umano while driving in the driving simulator at TASI (Transportation Active Safety Institute) lab. In addition to these two tasks, you will be also asked to complete one task where you will not use any of the applications. I will spend 5-10 minutes training you on how to use each of the applications before you use them. I will join you on your drive to observe your interactions with the interface, and help you if any technical problems should arise. I will also ask you a series of questions regarding your experience after using each of the interfaces. The entire session should last about 2 hours. Any question before we start? Great, let's start now with the warm up session, so you can get familiar with the simulator.

Training Tasks

Task List – Warm up

1. For you to get familiar with the simulator environment, we will have approximately 5-minute warm up session. Drive as you would normally do following the rules of the road. Remember you will hear the instruction to take turns or change lanes, so if you don't hear any instruction, please keep going straight. Your speedometer is shown on the screen. Feel free to try and stop and start the car several times to get a feeling of how it is and get used to the break. Start driving, as you would normally do following the rules of the road now.

[You have reached end of the warm up session, you can go ahead and put the vehicle in the park. Let's now move to the next part.]

Remember you will hear the instruction to take turns or change lanes, so if you don't hear any instruction, please keep going straight.

Task List – No device

1. You have approximately 15-minute. Drive as you would normally do following the rules of the road.

[You have reached end of the scenario, you can go ahead and put the vehicle in the park and fill up the questionnaire.]

[Let participants do all the steps themselves/Hand in the list of Voice commands]

ANFORA*Drive* Application Training

Take 2-3 minutes to look at the list of voice commands, then I will start training you on how to use the application and the voice commands.

Linkless Navigation - Voice Commands

Voice Commands	System Action
U.S.:	Select U.S. news category
WORLD, POLITICS, SPORTS, HEALTH, SCIENCE, ECONOMY, or TECHNOLOGY:	Select world, politics, sports, health, science, economy or technology news category
START, WHAT'S NEW, RECENT NEWS:	Starts playlist of news
RESTART:	Restart playlist of news
BACK, PREVIOUS:	Previous news story
SKIP, NEXT:	Next news story
SUMMARY:	Summary of the news story
FULL STORY:	Full story of the news story
MORE, TELL ME MORE, ANY-THING ELSE, RELATED, LIKE THIS:	Related news stories
PAUSE, STOP, PLAY	**Click on the button to pause, resume or play**

Step 1: Let's walk you through ANFORA*Drive* application and train you on how to use its interface.

- This application is basically a big, giant playlist of news **from the web**. This playlist of news is chunked in different sub playlist, each of a given category (for example, US news, technology news, politics news). If you let the playlist play, the playlist will go automatically from one category to another and will read one story after the other until all stories will be exhausted. In each story, the playlist will play the summary and then the full story. In this case, we have fresh news story coming from **NPR** into this application.
- We are now on the homepage of ANFORA*Drive*, please browse up and down for few seconds.
- Now if you want to start with the playlist, you have two options. You can click either click on Begin Flow or . Begin Flow starts the US news by default and it will go automatically from US to technology to politics and other categories available until all the news are exhausted.
- But if you are interested to start from a different category such as politics, you can click on , select the category ("Politics") and then click submit. Try clicking on Begin Flow

Now you are in the playlist of news. For example, this playlist contains X news stories across all available categories. As I mentioned before, if you let the playlist play, the playlist will read one story after the other until all X stories will be exhausted. In each story, the playlist will play the summary and then the full story. If at any point, you want to interact with the playlist, you can use button or voice control commands. For example with button, you can do the following:

- You can play and pause by clicking
- You can go to the next story by clicking , please click.
- You can go back to the previous story by clicking .
- You can move between different sections of news by clicking Summary or Full Story.
- You may decide that you are interested in listening to the news related to this story and you can just click on "**related news.**"
- At any time during your listening experience, you can change to another category of news by clicking the and then select another category and click Submit.
- At any time, if you decide to go back to the first category of news you started from, click RESTART.
- And you can click on the logo to go back to them home page.

Step 2: Other than button, you can also use voice control commands to start and interact with the playlist. To mention: this is a simulated prototype for voice interaction, which means you voice out your command and I will operate your command through my device.

- Now if you want to start with the playlist, press and hold the ***simulated button on the steering wheel*** and say one of the followings: "**What's New?" or "Start" or "Recent."** Then, release the button. Try pressing the button, saying, **"What's New"** and then release to see what happens.
- Or if you want to listen to a specific category of news like U.S., say "**U.S.**" Try saying, **"U.S."** and see what happens.
- Now you are in the playlist of news, you can go to the next story just by saying "**Next" or "Skip"**. Try saying, **"Next"** and see what happens.
- You can also go back to the previous story by saying **"Previous"** or **"Back".** Try saying, **"Back"** and see what happens.
- You can also move back and forth between different section of news by saying **"Full Story"** or **"Summary"**. Try saying, **"Full Story"** and **"Summary"** to see what happens.
- And if you decide you are interested in listening to the news related to this story. You can just say any of the following: **"Anything Else?", "More", "Tell Me More", "Like This"** or **"Related".** Try saying, **"More"** and see what happens.
- At any time during your listening experience, you can change the category of news story you are listening to another one that you are more interested in by saying "**Switch To Technology"** or **"Change To Technology."** Try saying, **"Change To Technology"** and see what happens.
- At any time, if you decide to listen to the news story from the beginning of the playlist, say **"Restart".** Try saying, **"Restart"** and see what happens.
- **Play and Pause** works using button and not voice.

Step 3: Review the list of voice commands for a few minutes. (2-3 minutes)

Step 4: Now, please start using ANFORA*Drive*, and use both button and voice control commands to interact with it. (5 minutes)

[Make sure to remove the audio files from the server]

Remember you will hear the instruction to take turns or change lanes, so if you don't hear any instruction, please keep going straight.

You can place your phone under the radio.

Remember you can activate and control the news either by voice or button. If you use voice, remember to click on the steering wheel and say your voice command. Since this is a prototype, sometimes there will be a long pause to load the content and audio whether you use voice or button, just bare with the system and don't take it out to look at the device, it will eventually play the story.

Task List – ANFORADrive

1. You have 15-minute to drive as you would normally do and listen to the playlist of news stories using ANFORA*Drive* app. In the first 2 minutes, you just drive without using the app. Once I prompt you, you can start listening to ANFORA*Drive* by selecting **any category of your interest**. Once you start listening to the news, for the rest of 8 minutes please don't do anything until I prompt you to change the news story or the news category [play the prompts for the users so they are familiar]. After 8 minutes of listening to the playlist, I will prompt you to listen and interact as you would normally wish to do for the remaining 5 minutes. I will stop you in 15 min. You can start driving now for 2 minutes.

[You have reached end of the scenario, you can go ahead and put the vehicle in the park and now come out to fill up the questionnaire.]

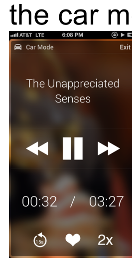
[Let participants to do all the steps themselves]

Umano Application Training umano

Step 6: Before we start, I would like to explain that each news story in umano app has only the full story and you can only use button control commands to interact with this application.

Step 7: Once you open umano app, there are 5 buttons in the bottom of the app, which are stories, popular, playlist, my channels and more.

- Stories Stories have the most recent news stories.
- Popular Popular has the most popular news stories.
- Playlist Playlist has the news stories you have added to your own playlist.
- My channels My Channels also show the channels you have selected based on your interest.

- When you are in Stories or Popular, you can either click on the story to listen to it or add the story to your playlist to listen to it later by clicking on .
- Once the news is added to your playlist, you can see next to the story. To remove the news from the playlist, you can click on wherever you are.
- While listening to the news stories, you can go to the next story by clicking .
- Imagine that you have listened to a few news stories and realize there was a story that you want to listen to again. You can go back to that story by clicking .
- Or imagine that while listening to a news story, you missed a part; you can click on which takes you back 15 seconds.
- Now if you want to pause while listening to news story, you can click
- And if you want to start playing the news story again, you can click on
- While listening to a news story, to change to the car mode, click on and select "Car Mode" and the page changes to 

and you can go to the next or previous news by clicking on or . You can also pause or play by clicking on or
- You can also exit the car mode, by clicking on the Exit on the top left corner.
- If you are interested to read the text or glance the news story text, you can click on View
- To add channels to "my channels", click on the top left corner and click on the

- Once the channel is added to my channels, you can see  next to the channel. To remove the channels from “my channel”, you can click on 

Step 8: Please start listening using umano app, and use the button control commands to interact with it. (5 minutes)

[Make sure to remove the stories from the playlist and also the channels from my channel before starting the real tasks.]

Remember you will hear the instruction to take turns or change lanes, so if you don’t hear any instruction, please keep going straight.

Please lower the volume for Umano Application. You can place your phone under the radio. Make sure to use the CarMode.

Task List – Umano

1. Please prepare your playlist in the channel section of Umano app selecting minimum 3 category you are interested in. [Once the playlist preparation is done] You have 15-minute to drive as you would normally do and listen to the playlist of news stories using Umano app. In the first 2 minutes, you just drive without using the app. Once I prompt you, you can start listening to Umano by picking **one of the categories you already selected**. Once you start listening to the news, for the rest of 8 minutes please don’t do anything until I prompt you to change the news story or the news category [play the prompts for the users so they are familiar]. After 8 minutes of listening to the playlist, I will prompt you to listen and interact as you would normally wish to do for the remaining 5 minutes. I will stop you in 15 min. You can start driving now for 2 minutes.

[You have reached end of the scenario, you can go ahead and put the vehicle in the park and now come out to fill up the questionnaire.]

ANFORA*Drive* Questionnaire

Thank you for your participation in today's study. Now, please take a few moments to answer the following questions about your experience with ANFORA*Drive*. Circle your answers.

Demographic Questions

Age: _____ **Gender:** M F

What kind of mobile phone do you own? _______________

How much time do you spend on news websites each week? (cirlce the answer that most closely applies to you.)

No time
5-30 minutes
30-60 minutes
1-3 hours
3-6 hours
More than 6 hours

How much of this time do you use your mobile to access the news?

No time
5-30 minutes
30-60 minutes
1-3 hours
3-6 hours
More than 6 hours

How much time do you spend listening to news on the radio each week?

No time
5-30 minutes
30-60 minutes
1-3 hours
3-6 hours
More than 6 hours

How much time do you spend watching news on TV each week?

No time
5-30 minutes
30-60 minutes
1-3 hours
3-6 hours
More than 6 hours

How often in a week do you drive?

Daily
4 - 6 days in a week
1 - 3 days in a week
1 day in a week
less than once in a week

Distraction Engagement Questions

When driving, you:	Never	Rarely	Sometimes	Often	Very Often
a. hold phone conversations					
b. manually interact with a phone (e.g.,sending text messages).					
c. adjust the settings of in-vehicle technology (e.g., radio channel or song selection).					
d. read roadside advertisements.					
e. continually check roadside accident scenes if there are any.					
f. chat with passengers if you have them.					
g. daydream					

No Device Questionnaire N

Circle how much each symptom below is affecting you right now.

Simulator Sickness Questionnaire

1. General Discomfort	None	Slight	Moderate	Severe
2. Fatigue	None	Slight	Moderate	Severe
3. Headache	None	Slight	Moderate	Severe
4. Eye strain	None	Slight	Moderate	Severe
5. Difficulty focusing	None	Slight	Moderate	Severe
6. Salivation increasing	None	Slight	Moderate	Severe
7. Sweating	None	Slight	Moderate	Severe
8. Nausea	None	Slight	Moderate	Severe
9. Difficulty concentrating	None	Slight	Moderate	Severe
10. Fullness of the head *	None	Slight	Moderate	Severe
11. Blurred vision	None	Slight	Moderate	Severe
12. Dizziness with eyes open	None	Slight	Moderate	Severe
13. Dizziness with eyes closed	None	Slight	Moderate	Severe
14. Vertigo **	None	Slight	Moderate	Severe
15. Stomache awareness ***	None	Slight	Moderate	Severe
16. Burping	None	Slight	Moderate	Severe

* Fullness of the head is sinus like fullness

** Vertigo is experienced as loss of orientation with respect to vertical upright

*** Stomache awareness is usually used to indicate a feeling of discomfort which is just short of nausea.

No Device Questionnaire N

Please respond to the following questions based on **the task of driving.** Please mark between the lines.

NASA TLX Questionnaire

	Very Low	Very High
1. How mentally demanding was the task?		
2. How physically demanding was the task?		
3. How hurried or rushed was the pace of the task?		
4. How successful were you in accomplishing what you were asked to do?		
5. How hard did you have to work to accomplish your level of performance?		
6. How insecure, discouraged, irritated, stressed, and annoyed were you?		

Distraction Question

	Very Low	Very High
1. How distracted were you while driving?		

ANFORA*Drive* Questionnaire	A

Circle how much each symptom below is affecting you right now.

Simulator Sickness Questionnaire

1. General Discomfort	None	Slight	Moderate	Severe
2. Fatigue	None	Slight	Moderate	Severe
3. Headache	None	Slight	Moderate	Severe
4. Eye strain	None	Slight	Moderate	Severe
5. Difficulty focusing	None	Slight	Moderate	Severe
6. Salivation increasing	None	Slight	Moderate	Severe
7. Sweating	None	Slight	Moderate	Severe
8. Nausea	None	Slight	Moderate	Severe
9. Difficulty concentrating	None	Slight	Moderate	Severe
10. Fullness of the head *	None	Slight	Moderate	Severe
11. Blurred vision	None	Slight	Moderate	Severe
12. Dizziness with eyes open	None	Slight	Moderate	Severe
13. Dizziness with eyes closed	None	Slight	Moderate	Severe
14. Vertigo **	None	Slight	Moderate	Severe
15. Stomache awareness ***	None	Slight	Moderate	Severe
16. Burping	None	Slight	Moderate	Severe

* Fullness of the head is sinus like fullness

** Vertigo is experienced as loss of orientation with respect to vertical upright

*** Stomache awareness is usually used to indicate a feeling of discomfort which is just short of nausea.

ANFORA*Drive* Questionnaire	A

Please respond to the following questions based on the task of **using ANFORA*Drive* and driving.** Please mark between the lines.

NASA TLX Questionnaire

	Very Low … Very High
1. How mentally demanding was the task?	
2. How physically demanding was the task?	
3. How hurried or rushed was the pace of the task?	
4. How successful were you in accomplishing what you were asked to do?	
5. How hard did you have to work to accomplish your level of performance?	
6. How insecure, discouraged, irritated, stressed, and annoyed were you?	

Distraction Question

	Very Low … Very High
1. How distracted were you while driving and using ANFORA*Drive*?	

ANFORA*Drive* Questionnaire A

Rate your level of agreement based on using **ANFORA*Drive* app while driving.**

System Usability Survey (SUS) Questionnaire	strongly disagree		neutral		strongly agree
1. I think that I would like to use this system or app frequently.	1	2	3	4	5
2. I found the system unnecessarily complex.	1	2	3	4	5
3. I thought the system was easy to use.	1	2	3	4	5
4. I think that I would need assistance to be able to use this system.	1	2	3	4	5
5. I found the various functions in this system were well integrated.	1	2	3	4	5
6. I thought there was too much inconsistency in this system.	1	2	3	4	5
7. I would imagine that most people would learn to use this system very quickly.	1	2	3	4	5
8. I found the system very cumbersome/awkward to use.	1	2	3	4	5
9. I felt very confident using the system.	1	2	3	4	5
10. I needed to learn a lot of things before I could get going with this system.	1	2	3	4	5

General ANFORA*Drive* Experience

Rate your experience using **ANFORA*Drive* app while driving.**

Unsatisfactory	1	2	3	4	5	6	7	**Satisfactory**
Annoying	1	2	3	4	5	6	7	**Pleasing**
Unenjoyable	1	2	3	4	5	6	7	**Enjoyable**
Not distracted	1	2	3	4	5	6	7	**Very distracted**
Difficult	1	2	3	4	5	6	7	**Simple**
Confusing	1	2	3	4	5	6	7	**Easy to understand**
Boring	1	2	3	4	5	6	7	**Engaging**
Extremely unsafe	1	2	3	4	5	6	7	**Extremely safe**

ANFORA*Drive* Interview Questions A

Based on your experience today with ANFORA*Drive*,

• Overall, how would you describe your experience with ANFORA*Drive* app while driving today?

• Were you able to pay attention to the news while driving?

If yes, were you still able to adequately attend to the road?

If no, why not?

• Was it clear when a new news story started? What was the indication?

• Was it clear when a news story ended?

• At any point, did you feel confused by the interface or the voice commands? If so, can you recall when?

• At any point, did you feel lost in the while listening to the news? If so, can you recall when?

• At any point, did you stop listening before your playlist ended?

If yes, why?

• Do you prefer to use voice commands, button or a combination of the two? Why?

• If ANFORA*Drive* was available today, would you use it? when would you use it? Why or why not?

• What did you like best about ANFORA*Drive* application?

• What did you like least?

• How many news stories did you listen to today using ANFORA*Drive*?

• Tell me briefly about one news story that struck you.

• Do you have anything else you would like to share about your experience with ANFOR-A*Drive* app?

Umano Questionnaire U

Circle how much each symptom below is affecting you right now.

Simulator Sickness Questionnaire

1. General Discomfort	None	Slight	Moderate	Severe
2. Fatigue	None	Slight	Moderate	Severe
3. Headache	None	Slight	Moderate	Severe
4. Eye strain	None	Slight	Moderate	Severe
5. Difficulty focusing	None	Slight	Moderate	Severe
6. Salivation increasing	None	Slight	Moderate	Severe
7. Sweating	None	Slight	Moderate	Severe
8. Nausea	None	Slight	Moderate	Severe
9. Difficulty concentrating	None	Slight	Moderate	Severe
10. Fullness of the head *	None	Slight	Moderate	Severe
11. Blurred vision	None	Slight	Moderate	Severe
12. Dizziness with eyes open	None	Slight	Moderate	Severe
13. Dizziness with eyes closed	None	Slight	Moderate	Severe
14. Vertigo **	None	Slight	Moderate	Severe
15. Stomache awareness ***	None	Slight	Moderate	Severe
16. Burping	None	Slight	Moderate	Severe

* Fullness of the head is sinus like fullness

** Vertigo is experienced as loss of orientation with respect to vertical upright

*** Stomache awareness is usually used to indicate a feeling of discomfort which is just short of nausea.

Umano Questionnaire U

Please respond to the following questions based on the task of using **Umano and driving.** Please mark between the lines.

NASA TLX Questionnaire

	Very Low … Very High
1. How mentally demanding was the task?	
2. How physically demanding was the task?	
3. How hurried or rushed was the pace of the task?	
4. How successful were you in accomplishing what you were asked to do?	
5. How hard did you have to work to accomplish your level of performance?	
6. How insecure, discouraged, irritated, stressed, and annoyed were you?	

Distraction Question

	Very Low … Very High
1. How distracted were you while driving and using Uamno?	

Umano Questionnaire	U

Rate your level of agreement based on using **Umano app while driving.**

SUS Questionnaire	strongly disagree		neutral		strongly agree
1. I think that I would like to use this system or app frequently.	1	2	3	4	5
2. I found the system unnecessarily complex.	1	2	3	4	5
3. I thought the system was easy to use.	1	2	3	4	5
4. I think that I would need assistance to be able to use this system.	1	2	3	4	5
5. I found the various functions in this system were well integrated.	1	2	3	4	5
6. I thought there was too much inconsistency in this system.	1	2	3	4	5
7. I would imagine that most people would learn to use this system very quickly.	1	2	3	4	5
8. I found the system very cumbersome/awkward to use.	1	2	3	4	5
9. I felt very confident using the system.	1	2	3	4	5
10. I needed to learn a lot of things before I could get going with this system.	1	2	3	4	5

General Umano Experience

Rate your experience using **Umano app while driving.**

	1	2	3	4	5	6	7	
Unsatisfactory	1	2	3	4	5	6	7	Satisfactory
Annoying	1	2	3	4	5	6	7	Pleasing
Unenjoyable	1	2	3	4	5	6	7	Enjoyable
Not distracted	1	2	3	4	5	6	7	Very distracted
Difficult	1	2	3	4	5	6	7	Simple
Confusing	1	2	3	4	5	6	7	Easy to understand
Boring	1	2	3	4	5	6	7	Engaging
Extremely unsafe	1	2	3	4	5	6	7	Extremely safe

Umano Interview Questions	U

Based on your experience today with Umano,

• Overall, how would you describe your experience with Umano application today?

• Were you able to pay attention to the news while driving?

If yes, were you still able to adequately attend to the road?

If no, why not?

• Was it clear when a new news story started? What was the indication?

• Was it clear when a news story ended?

• At any point, did you feel confused by the interface? If so, can you recall when?

• At any point, did you feel lost in the while listening to the news? If so, can you recall when?

• At any point, did you stop listening before your playlist ended?

If yes, why?

• How did you find interacting with Umano using buttons while driving?

• Since Umano app is available today, would use you it? when would you use it? Why or why not?

• What did you like best about Umano app?

• What did you like least?

• How many news stories did you listen to today using Umano app?

• Tell me briefly about one news story that struck you.

• Do you have anything else you would like to share about your experience with Umano app?

Comparison Interview Questions

Based on your experience today with both ANFORA*Drive* and Umano,

• Which one of these two applications do you like to use while driving? why?

• Did you not engage with these two applications (device) because of traffic condition?

Appendix H: Random Recognition Errors Generated

Voice Command Number	Error Recognition Type if any
1 2 3 ….	N/A
29	Inaccurate Error
30 ….	N/A
73	Inaccurate Error
74 ….	N/A
95	Inaccurate Error
96 ….	N/A
123	Inaccurate Error
124 ….	N/A
134	Inaccurate Error
135 ….	N/A
163	Missing Error
164 ….	N/A
184	Inaccurate Error
185 ….	N/A
261	Missing Error
262 ….	N/A

273	Inaccurate Error
274 …. 299 300	N/A

References

Alt, F., Kern, D., Schulte, F., Pfleging, B., Shirazi, A. S., & Schmidt, A. (2010, November). Enabling micro-entertainment in vehicles based on context information. In *Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 117-124). ACM.

Android Google Voice. https://play.google.com/store/apps/details?id=com.google.android.apps.googlevoice&hl=en (accessed August 27, 2015)

Android Will Be Integrated Into Cars As Early As 2015. http://www.digitalafro.com/android-will-be-integrated-into-cars-as-early-as-2015/ (accessed August 27, 2015)

Anhalt, J., Smailagic, A., Siewiorek, D. P., Gemperle, F., Salber, D., Weber, S., . . . Jennings, J. (2001). Toward Context-Aware Computing: Experiences and Lessons. *IEE Intelligent Systems*, *16*(3), 38-46. http://doi.ieeecomputersociety.org/10.1109/5254.940025

Anthony, J. C., LeResche, L., Niaz, U., Von Korff, M. R., & Folstein, M. F. (1982). Limits of the 'Mini-Mental State' as a screening test for dementia and delirium among hospital patients. *Psychological medicine*, *12*(02), 397-408. http://dx.doi.org/10.1017/S0033291700046730

Anupam, V., Freire, J., Kumar, B., & Lieuwen, D. (2000). Automating Web navigation with the WebVCR. *Computer Networks*, *33*(1-6), 503-517. doi:10.1016/S1389-1286(00)00073-6

Apple Siri. http://www.apple.com/iphone/features/siri.html (accessed July 1, 2015)

Audible. http://www.audible.com (accessed August 10, 2015)

Baldwin, C. L. (2012). *Auditory cognition and human performance: Research and applications*. Boca Raton, FL: Taylor & Francis.

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction*, *24*(6), 574-594. doi:10.1080/10447310802205776

Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies*, *4*(3), 114-123. doi:10.1.1.177.1240.

Barón, A., & Green, P. (2006). *Safety and usability of speech interfaces for in-vehicle tasks while driving: A brief literature review* (No. UMTRI-2006-5). University of Michigan, Transportation Research Institute.

Blattner, M. M., Sumikawa, D. A., & Greenberg, R. M. (1989). Earcons and Icons: Their Structure and Common Design Principles. *Human–Computer Interaction*, *4*(1), 11-44. doi: 10.1207/s15327051hci0401_1

Bohannon, R. W. (1997). Comfortable and maximum walking speed of adults aged 20—79 years: reference values and determinants. *Age and ageing*, *26*(1), 15-19. doi: 10.1093/ageing/26.1.15

Bolchini, D. & Ghahari, R. R. (2013). *U.S. Patent Application 14/024,612.*

Bolchini, D., & Paolini, P. (2006). Interactive dialogue model: A design technique for multichannel applications. *Multimedia, IEEE Transactions on*, *8*(3), 529-541. doi: 10.1109/TMM.2006.870733

Bolin, M., Webber, M., Rha, P., Wilson, T., & Miller, R. C. (2005). Automation and customization of rendered web pages. *Proceedings of the 18th annual ACM symposium on User interface software and technology*, 163-172. doi: 10.1145/1095034.1095062

Borodin, Y. (2008). Automation of repetitive web browsing tasks with voice-enabled macros. *Proceedings of the 10th international ACM SIGACCESS conf. on Computers and Accessibility*, 307-308. doi: 10.1145/1414471.1414552

Borodin, Y., Puzis, Y., Soviak, A., Bouker, J., Feng, B., Sicoli, R., . . . Ramakrishnan, I. V. (2014, April). Listen to everything you want to read with Capti narrator. In *Proceedings of the 11th Web for All Conference* (p. 33). ACM. doi: 10.1145/2596695.2596728

Bradford, J. H. (1995). The human factors of speech-based interfaces: a research agenda. *SIGCHI Bulletin*, *27*(2), 61-67. doi: 10.1145/202511.202527

Bragdon, A., Nelson, E., Li, Y., & Hinckley, K. (2011, May). Experimental analysis of touch-screen gesture designs in mobile environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 403-412). ACM. doi: 10.1145/1978942.1979000

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, *189*(194). 4-7.

Brumby, D. P., Davies, S. C., Janssen, C. P., & Grace, J. J. (2011). Fast or safe?: how performance objectives determine modality output choices while interacting on the move. *Proc. CHI* (473-482). ACM. doi: 10.1145/1978942.1979009

Chandler, P., & Sweller, J. (1991). Cognitive Load Theory and the format of Instruction. *Cognition and instruction*, *8*(4), 293-332. doi: 10.1207/s1532690xci0804_2

Christian, K., Kules, B., Shneiderman, B., & Youssef, A. (2000). A comparison of voice controlled and mouse controlled web browsing. *Proc. ASSETS* (72-79). ACM. doi: 10.1145/354324.354345

Dahlbäck, N., Jönsson, A., & Ahrenberg, L. (1993, February). Wizard of Oz studies: why and how. In *Proceedings of the 1st international conference on Intelligent user interfaces* (pp. 193-200). ACM. doi: 10.1145/169891.169968

Demers, D. P. (2005). *Dictionary of mass communication & media research: a guide for students, scholars and professionals*. Marquette Books. p.143.

Döring, T., Kern, D., Marshall, P., Pfeiffer, M., Schöning, J., Gruhn, V., & Schmidt, A. (2011, May). Gestural interaction on the steering wheel: reducing the visual demand. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 483-492). ACM. doi: 10.1145/1978942.1979010

Duarte, C., & Carriço, L. (2009). When You Can't Read It, Listen to It! An Audio-Visual Interface for Book Reading. *Universal Access in Human-Computer Interaction. Applications and Services*, *5616,* 24-33. doi: 10.1007/978-3-642-02713-0_3

Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of vision*, *10*(1), 6, doi:10.1167/10.1.6.

Feng, J., & Sears, A. (2009). Speech Input to Support Universal Access. The Universal Access Handbook, (pp. 1-12): CRC Press.

Feng, J., Marulanda, S., & Donmez, B. (2014). Susceptibility to Driver Distraction Questionnaire (SDDQ): Development and Relation to Relevant Self-Reported

Measures. In *Proceedings of the Transportation Research Board 93rd Annual Meeting (14-3009)*. http://dx.doi.org/10.3141/2434-04

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*, *12*(3), 189-198. doi:10.1016/0022-3956(75)90026-6

Fong, M., & Frank, M. P. (1992, October). A rapid semi-automatic simulation technique for investigating interactive speech and handwriting. In *Proc. Int'l Conf. on Spoken Language Processing* (Vol. 2, pp. 1351-1354). University of Alberta Press, Edmonton, Canada.

Gable, T. M., Walker, B. N., Moses, H. R., & Chitloor, R. D. (2013, October). Advanced auditory cues on mobile phones help keep drivers' eyes on the road. In *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 66-73). ACM. doi: 10.1145/2516540.2516541

Garlan, D., Siewiorek, D. P., Smailagic, A., & Steenkiste, P. (2002). Project aura: Toward distraction-free pervasive computing. *Pervasive Computing, IEEE*, *1*(2), 22-31. doi: 10.1109/MPRV.2002.1012334

Garzotto, F., Paolini, P., & Schwabe, D. (1993). HDM—a model-based approach to hypertext application design. *TOIS*, *11*(1), 1-26. doi: 10.1145/151480.151483

Ghahari, R. R., & Bolchini, D. (2011, September). ANFORA: Investigating aural navigation flows on rich architectures. In *Web Systems Evolution (WSE), 2011 13th IEEE International Symposium on* (pp. 27-32). IEEE. doi: 10.1109/WSE.2011.6081816

Google Images. https://images.google.com (accessed August 10, 2015)

Goose, S., & Djennane, S. (2002). WIRE[3]: Driving Around the Information Super-Highway. *Personal and Ubiquitous Computing*, *6*(3), 164-175. doi: 10.1007/s007790200017

Green, P., & Wei-Haas, L. (1985, October). The Rapid Development of User Interfaces: Experience with the Wizard of Oz Method. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 29, No. 5, pp. 470-474). SAGE Publications. doi: 10.1177/154193128502900515

Gupta, G., Raman, S. S., Nichols, M., Reddy, H., & Annamalai, N. (2005). DAWN: Dynamic Aural Web Navigation. *Proc. HCI*, Las Vegas.

Harbluk, J. L., & Lalande, S. (2005). Performing e-mail tasks while driving: The impact of speech-based tasks on visual detection. In *Proceedings of the Third International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Rockport, ME* (pp. 304-310).

Harbluk, J. L., Eisenman, M., & Noy, Y. I. (2002). *The impact of cognitive distraction on driver visual behaviour and vehicle control* (No. TP# 13889 E). Transport Canada.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human mental workload*, *1*(3), 139-183. doi:10.1016/S0166-4115(08)62386-9

Hearst, M. A. *Towards "Natural" Interactions in Search User Interfaces*. Retrieved from http://people.ischool.berkeley.edu/~hearst/papers/cacm11.pdf (accessed August 27, 2015)

Hoonakker, P., Carayon, P., Gurses, A. P., Brown, R., Khunlertkit, A., McGuire, K., & Walker, J. M. (2011). Measuring workload of ICU nurses with a questionnaire survey: the NASA task load index (TLX). *IIE transactions on healthcare systems engineering*, *1*(2), 131-143. doi:10.1080/19488300.2011.609524

Horberry, T. (1998). Bridge strike reduction: the design and evaluation of visual warnings (Doctoral dissertation, University of Derby).

Horberry, T., & Edquist, J. (2008). 13 Distractions outside the Vehicle. *Driver Distraction: Theory, Effects, and Mitigation*, 215.

Horberry, T., Anderson, J., Regan, M. A., Triggs, T. J., & Brown, J. (2006). Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. *Accident Analysis & Prevention, 38*(1), 185-191. doi:10.1016/j.aap.2005.09.007

Hua, Z., & Ng, W. L. (2010, November). Speech recognition interface design for in-vehicle system. In *Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 29-33). ACM. doi: 10.1145/1969773.1969780

Jain, A., & Gupta, G. (2007). VoxBoox: A System for Automatic Generation of Interactive Talking Books. *Proc. UAHCI* (329-338). Springer Berlin Heidelberg. doi: 10.1007/978-3-540-73283-9_37

Justiss, M. D., Mann, W. C., Stav, W., & Velozo, C. (2006). Development of a Behind‐the‐Wheel Driving Performance Assessment for Older Adults. *Topics in Geriatric Rehabilitation, 22*(2), 121-128.

Kane, S. K., Wobbrock, J. O., & Smith, I. E. (2008). Getting off the treadmill: Evaluating walking user interfaces for mobile devices in public spaces. In *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services* (pp. 109-118). ACM. doi: 10.1145/1409240.1409253

Karl, L. R., Pettey, M., & Shneiderman, B. (1998). Speech Versus Mouse Commands for Word Processing: an Empirical Evaluation.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, *3*(3), 203-220. doi: 10.1207/s15327108ijap0303_3

Klemmer, S. R., Sinha, A. K., Chen, J., Landay, J. A., Aboobaker, N., & Wang, A. (2000). Suede: a Wizard of Oz prototyping tool for speech user interfaces. In *Proceedings of the 13th annual ACM symposium on User interface software and technology* (pp. 1-10). ACM. doi: 10.1145/354401.354406

Lager, T. (2012). The Spoken Web: a survey from a Web Science perspective. *Proceedings of Fonetik 2012*.

Lee, J. D. (2007). Technology and teen drivers. *Journal of safety research*, *38*(2), 203-213. doi:10.1016/j.jsr.2007.02.008

Lee, J. D., Caven, B., Haake, S., & Brown, T. L. (2001). Speech-based interaction with in-vehicle computers: The effect of speech-based e-mail on drivers' attention to the roadway. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *43*(4), 631-640. doi: 10.1518/001872001775870340

Lee, K. M., & Lai, J. (2005). Speech versus touch: A comparative study of the use of speech and DTMF keypad for navigation. *International Journal of Human-Computer Interaction*, *19*(3), 343-360. doi: 10.1207/s15327590ijhc1903_4

Lemmelä, S., Vetek, A., Mäkelä, K., & Trendafilov, D. (2008, October). Designing and evaluating multimodal interaction for mobile contexts. In *Proceedings of the 10th international conference on Multimodal interfaces* (pp. 265-272). ACM. doi: 10.1145/1452392.1452447

Lewis, J. R., & Sauro, J. (2009). The factor structure of the system usability scale. In *Human Centered Design* (pp. 94-103). Springer Berlin Heidelberg. doi: 10.1007/978-3-642-02806-9_12

Li, K. A., Baudisch, P., & Hinckley, K. (2008). BlindSight: Eyes-Free Access to Mobile Phones. *Proceeding of the 26th annual SIGCHI conference on human factors in computing systems*, 1389-1398. doi: 10.1145/1357054.1357273

Liu, A., & Salvucci, D. (2001, August). Modeling and prediction of human driver behavior. In *Proceedings of the 9th International Conference on Human-Computer Interaction* (pp. 1479-1483).

Maciej, J., & Vollrath, M. (2009). Comparison of manual vs. speech-based interaction with in-vehicle information systems. *Accident Analysis & Prevention*, *41*(5), 924-930. doi:10.1016/j.aap.2009.05.007

Meijer, I. C. (2007). Checking, snacking and bodysnatching. How young people use the news and implications for public service media journalism. *From Public Service Broadcasting to Public Service Media. Gothenburg: Nordicom*.

Metz, B., & Krueger, H. P. (2010). Measuring visual distraction in driving: The potential of head movement analysis. *Intelligent Transport Systems, IET*, *4*(4), 289-297. doi: 10.1049/iet-its.2009.0116

Morley, S. (1998). Digital talking books on a PC: A usability evaluation of the prototype DAISY playback software. *Proceedings of the third international ACM conference on assistive technologies*, 157-164. doi: 10.1145/274497.274527

Narayanan, S., & Potamianos, A. (2002). Creating conversational interfaces for children. *Speech and Audio Processing, IEEE Transactions on*, *10*(2), 65-78. doi: 10.1109/89.985544

Neale, V. L., Dingus, T. A., Klauer, S. G., Sudweeks, J., & Goodman, M. (2005). An overview of the 100-car naturalistic study and findings. *National Highway Traffic Safety Administration, Paper*, (05-0400).

Neckset. http://neckset.com (accessed August 10, 2015)

Nichols, M., Gupta, G., & Wang, Q. (2005). Voice-commanded Scripting Language for Programming Navigation Strategies on-the-fly. *Paper presented at the Proceedings of the HCI International 2005*, Las Vegas.

Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. *Proc. CHI* (249-256). ACM. doi: 10.1145/97243.97281

Nuance – Dragon Dictation: iPhone – Dragon Dictation for iPadTM, iPhoneTM and iPod touchTM is an easy-to-use voice recognition application. http://www.nuance.com/for-business/by-product/dragon-dictation-iphone/index.htm (accessed August 27, 2015)

Odenheimer, G. L., Beaudet, M., Jette, A. M., Albert, M. S., Grande, L., & Minaker, K. L. (1994). Performance-based driving evaluation of the elderly driver: safety, reliability, and validity. *Journal of Gerontology, 49*(4), M153-M159. doi: 10.1093/geronj/49.4.M153

Ohn-Bar, E., Tran, C., & Trivedi, M. (2012, October). Hand gesture-based visual user interface for infotainment. In *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 111-115). ACM. doi: 10.1145/2390256.2390274

Paolini, P., Garzotto, F., Bolchini, D., Valenti, S. (1999). 'Modelling by Pattern' of Web Applications, in Peter P.S. Chen (Eds). *Proceedings of Advances in Conceptual Modeling (ER'99) Workshops on Evolution and Change in Data Management, Reverse Engineering in Information Systems, and the World Wide Web and Conceptual Modeling*, 293-306.

Patel, N., Agarwal, S., Rajput, N., Nanavati, A., Dave, P., & Parikh, T. S. (2009, April). A comparative study of speech and dialed input voice interfaces in rural India. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 51-54). ACM. doi: 10.1145/1518701.1518709

Peissner, M., Doebler, V., & Metze, F. (2011). Can voice interaction help reducing the level of distraction and prevent accidents? Meta-study on driver distraction and voice interaction. *Nuance Communications*, Burlington, MA, USA, May 19, 2011, White Paper.

Price, K. J., Lin, M., Feng, J., Goldman, R., Sears, A., & Jacko, J. A. (2006). Motion does matter: an examination of speech-based text entry on the move. *Universal*

Access in the Information Society, *4*(3), 246-257. doi: 10.1007/s10209-005-0006-8

Quinn, S., Stark, P., Edmonds, R., Moos, J. and Van Wagener, A. (2007). Eyetracking the News. *St. Petersburg: Poynter Institute for Media Studies.*

Ranney, T. A., Harbluk, J. L., & Noy, Y. I. (2005). Effects of Voice Technology on Test Track Driving Performance: Implications for Driver Distraction. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *47*(2), 439-454. doi: 10.1518/0018720054679515

Rohani Ghahari, R., George-Palilonis, J., & Bolchini, D. (2013). Mobile Web Browsing with Aural Flows: An Exploratory Study. *International Journal of Human-Computer Interaction*, *29*(11), 717-742. doi: 10.1080/10447318.2013.773875

Salber, D., & Coutaz, J. (1993). Applying the Wizard of Oz Technique to the Study of Multimodal Systems. In *Human-Computer Interaction* (pp. 219-230). Springer Berlin Heidelberg. doi: 10.1007/3-540-57433-6_51

Salvucci, D. D. (2001). Predicting the effects of in-car interfaces on driver behavior using a cognitive architecture. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 120-127). ACM. doi: 10.1145/365024.365064

Salvucci, D. D. (2002). Modeling Driver Distraction from Cognitive Tasks. In *Proceedings of the 24th Annual Conference of the Cognitive Science Society* (pp. 792-797).

Salvucci, D. D. (2005). Modeling tools for predicting driver distraction. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 49, No. 12, pp. 1149-1152). SAGE Publications. doi: 10.1177/154193120504901211

Salvucci, D. D. (2006). Modeling driver behavior in a cognitive architecture. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *48*(2), 362-380. doi: 10.1518/001872006777724417

Salvucci, D. D. (2009). Rapid prototyping and evaluation of in-vehicle interfaces. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *16*(2), 9. doi: 10.1145/1534903.1534906

Salvucci, D. D. (2013, April). Distraction beyond the driver: Predicting the effects of in-vehicle interaction on surrounding traffic. In *Proceedings of the 2013 ACM annual conference on Human factors in computing systems* (pp. 3131-3134). ACM. doi: 10.1145/2470654.2466427

Salvucci, D. D., & Taatgen, N. A. (2008). Threaded cognition: An integrated theory of concurrent multitasking. *Psychological review*, *115*(1), 101-130. http://dx.doi.org/10.1037/0033-295X.115.1.101

Salvucci, D. D., Zuber, M., Beregovaia, E., & Markley, D. (2005). Distract-R: Rapid prototyping and evaluation of in-vehicle interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 581-589). ACM. doi: 10.1145/1054972.1055052

Sawhney, N., & Schmandt, C. (2000). Nomadic radio: speech and audio interaction for contextual messaging in nomadic environments. *(TOCHI)*, *7*(3), 353-383. doi: 10.1145/355324.355327

Schildbach, B., & Rukzio, E. (2010). Investigating selection and reading performance on a mobile phone while walking. In *Proceedings of the 12th international* *conference on Human computer interaction with mobile devices and services* (pp. 93-102). ACM. doi: 10.1145/2470654.2481421

Serrano, M., Lecolinet, E., & Guiard, Y. (2013). Bezel-Tap gestures: quick activation of commands from sleep mode on tablets. *Proc. CHI* (3027-3036). ACM. doi: 10.1145/2470654.2481421

Sinha, A. K., Klemmer, S. R., & Landay, J. A. (2002). Embarking on spoken-language NL interface design. *IJST*, *5*(2), 159-169, doi: 10.1023/A:1015424215015.

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, *73*(4), 971-995, doi:10.3758/s13414-010-0073-7.

Strayer, D. L., Cooper, J. M., Turrill, J., Coleman, J., Medeiros-Ward, N., & Biondi, F. (2013). Measuring cognitive distraction in the automobile.

Strayer, D. L., Watson, J. M., & Drews, F. A. (2011). Cognitive Distraction While Multitasking in the Automobile. *Psychology of Learning and Motivation-Advances in Research and Theory*, *54*, 29.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science*, *12*(2), 257-285, doi:10.1016/0364-0213(88)90023-7.

Sweller, J., & Chandler, P. (1994). Why Some Material Is Difficult to Learn. *Cognition and instruction*, *12*(3), 185-233. doi: 10.1207/s1532690xci1203_1

Sweller, J., Van Merrienboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational psychology review*, *10*(3), 251-296, doi:10.1023/A:1015424215015.

Tang, Y., Wang, D., Bai, J., Zhu, X., & Li, M. (2013). Information distance between what I said and what it heard. *CACM*, *56*(7), 70-77, doi:10.1145/2483852.2483869.

Tashev, I., Seltzer, M., Ju, Y. C., Wang, Y. Y., & Acero, A. (2009). Commute UX: Voice enabled in-car infotainment system. In *Mobile HCI* (Vol. 9).

TASI. http://www.engr.iupui.edu/tasi/ (accessed August 10, 2015)

Tchankue, P., Wesson, J., & Vogts, D. (2012). Are mobile in-car communication systems feasible?: A usability study. *Proceedings of the South African Institute for Computer Scientists and Information Technologists Conference*, 262-269. doi: 10.1145/2389836.2389867

The Future of In-Car Technology. http://www.caranddriver.com/features/the-future-of-in-car-technology (accessed August 27, 2015)

Trick, L. M., & Enns, J. T. (2009). A Two-Dimensional Framework for Understanding the Role of Attentional Selection in Driving. *Human factors of visual and cognitive performance in driving*, 63-73.

Tsimhoni, O., Smith, D., & Green, P. (2004). Address entry while driving: Speech recognition versus a touch-screen keyboard. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(4), 600-610, doi: 10.1518/hfes.46.4.600.56813.

Umano. http://umano.me/ (accessed March 10, 2015)

Vo, M. T., & Wood, C. (1996, May). Building an application framework for speech and pen input integration in multimodal learning interfaces. In *Acoustics, Speech, and Signal Processing, 1996. ICASSP-96. Conference Proceedings., 1996 IEEE International Conference on* (Vol. 6, pp. 3545-3548). IEEE, doi: 10.1109/ICASSP.1996.550794.

VoiceDream. http://www.voicedream.com/ (accessed July 1, 2015)

Wakkary, R., & Hatala, M. (2007). Situated play in a tangible interface and adaptive audio museum guide. *Personal Ubiquitous Comput., 11*(3), 171-191. doi: 10.1007/s00779-006-0101-8

Wickens, C. D. (1980). The structure of attentional resources. *Attention and performance VIII*, *8*.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical issues in ergonomics science*, *3*(2), 159-177. doi: 10.1080/14639220210123806

Winter, U., Grost, T. J., & Tsimhoni, O. (2010). Language Pattern Analysis for Automotive Natural Language Speech Applications. In *Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 34-41). ACM. doi: 10.1145/1969773.1969781

Yang, J., Sidhom, S., Chandrasekaran, G., Vu, T., Liu, H., Cecan, N., ... & Martin, R. P. (2011). Detecting driver phone use leveraging car speakers. In *Proceedings of the 17th annual international conference on Mobile computing and networking* (pp. 97-108). ACM. doi: 10.1145/2030613.2030625

Yang, Y., Reimer, B., Mehler, B., Wong, A., & McDonald, M. (2012). Exploring differences in the impact of auditory and visual demands on driver behavior. *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 173-177. doi: 10.1145/2390256.2390285

Young, K., Lee, J. D., & Regan, M. A. (Eds.). (2008). *Driver distraction: Theory, effects, and mitigation*. CRC Press.

Young, K., Regan, M., & Hammer, M. (2007). Driver distraction: A review of the literature. *Distracted driving. Sydney, NSW: Australasian College of Road Safety*, 379-405.

Zancanaro, M., Stock, O., & Alfaro, I. (2003). Using Cinematic Techniques in a Multimedia Museum Guide. *Proceedings of Museums and the Web 2003, 12*.

Zhang, D., & Lai, J. (2011). Can Convenience and Effectiveness Converge in Mobile Web? A Critique of the State-of-the-Art Adaptation Techniques for Web Navigation on Mobile Handheld Devices. *International Journal of Human-Computer Interaction, 27*(12), 1133-1160. doi: 10.1080/10447318.2011.559876

Zhao, S., Dragicevic, P., Chignell, M., Balakrishnan, R., & Baudisch, P. (2007). Earpod: eyes-free menu selection using touch input and reactive audio feedback. *Proc. CHI* (1395-1404). doi: 10.1145/1240624.1240836

Zhou, R., Rau, P. L. P., Zhang, W., & Zhuang, D. (2012). Mobile phone use while driving: Predicting drivers' answering intentions and compensatory decisions. *Safety Science*, *50*(1), 138-149. doi:10.1016/j.ssci.2011.07.013

CURRICULUM VITAE

Romisa Rohani Ghahari

Email: rrohanig@iupui.edu
Web: http://mypage.iu.edu/~rrohanig/

EDUCATION

Indiana University, School of Informatics, Indianapolis, IN — August 2015
PhD in Human-Computer Interaction (GPA 3.94)

Syracuse University, School of Information Studies, Syracuse, NY — August 2010
MS in Information Management (GPA 3.65)

Visveswaraiah Technological University, B.M.S College of Engineering, Bangalore, India — June 2006
BE in Information Science and Engineering

COMPUTING SKILLS

Database Management: Oracle 10g, MS SQL Server 2005, MS-Access

Operating Systems: Linux, UNIX, Windows XP, Windows Vista

Programming Languages: C, C++, C#, Java, UNIX Programming, Visual BASIC, XML, HTML & CSS, PHP, CakePHP, Javascript, SQL, PL/SQL

Software Packages: SPSS, R, VB 6.0, MS Visual Studio

HUMAN-COMPUTER INTERACTION (HCI) SKILLS

HCI Skills: User Experimental Design, Contextual Inquiry, Wireframes, Rapid Prototyping, User Interface Design, Mobile Web User Interfaces Design, Interaction Design, Usability Testing, Wizard-of-Oz Testing, User Studies, Think-aloud Sessions, Focus Group, Diary Study, Heuristics Evaluation, Cognitive Walkthrough, Paper Prototyping, Low-Hi Fidelity Prototyping, User Interview

HCI tools: Morae Recorder and Manager, Camtasia, Balsamiq, Axure, Adobe Creative Suite

PATENT

Aural Navigation of Information Rich Visual Interfaces
Date filed: 09/11/2012. Co-inventors: Davide Bolchini and Romisa Rohani Ghahari. **Submitted as Regular US Patent (pending) on September 12, 2013.**
https://www.google.com/patents/US20140282006?dq=bolchini&hl=en&sa=X&ei=gJutVMaEEoL3yQT114KgDQ&ved=0CB0Q6AEwAA

EMPLOYMENT HISTORY

Indiana University, School of Informatics, Indianapolis, IN — August 2010 – Present
Research Assistant
Assisting Dr. Davide Bolchini (PI) in the NSF-funded research project “Navigating the Aural Web”

- Investigating advanced navigation strategies in complex information architectures to support the "aural" consumption of web content
- Exploring multimodal interaction design, running user studies with sighted users on the go or blind users
- Generated publications: Back navigation shortcuts for screen reader users, Aural Browsing On-The-Go: Listening-based Back Navigation in Large Web Architectures, ANFORA: Investigating aural navigation flows on rich architectures

Syracuse University, School of Information Studies, Syracuse, NY September 2009 – May 2010
Research Assistant
Assisted Dr. Ping Zhang with a research project on students' use of social media for various purposes

- Arranged focus groups with students to collect data until saturation
- Analyzed and coded data using Atlas.ti software
- Reviewed literature for the project

Syracuse University, Maxwell School, Syracuse, NY May 2009 – July 2010
Research Assistant
Assisted Dr. Mehrzad Boroujerdi with a research project on background of Iranian politicians

- Read the profiles of Iranian politicians to select useful information, then entered information into a database
- Entered data into the database in English, though the profiles were in Persian

RESEARCH PROJECTS EXPERIENCE

Indiana University School of Informatics, Indianapolis, IN
NSF-NIH funded research project "From Critique to Collaboration: A Fundamental Rethinking of Computerized Clinical Alerts" September 2014 – August 2015
PI: Dr. Jon Duke and Dr. Davide Bolchini

- Conducted three contextual inquiries (totaling 255 minutes, 22 participants) at Wishard-Eskenazi Health
- Found eight themes driving trusted advice and seven design directions for trust-based alerts
- Prototyped novel drug-drug interaction alert user interfaces to convey drug safety information to providers

NIH-funded project "Anti-Smoking Video Game for Pre-adolescent Girls"
January 2011 – February 2011

- Reviewed recorded usability testing of anti-smoking video games using Morae Manager software
- Analyzed the collected data
- Consolidated the findings and coded key requirements for improvements
- Created usability evaluation and analysis report

NSF-funded research project "Aural Navigation Flows on Rich Architecture (ANFORA)"
January 2011 – August 2015
http://www.youtube.com/watch?v=rSuGcvDUh70&feature=plcp, PI: Dr. Davide Bolchini

- Designed different types of aural flows for news domain information web architectures
- Implemented different types of aural flows in ANFORA News mobile web application prototype
- Evaluated the aural flows designs by conducting user studies with 20 participants

- Explored and designed other modes of interaction with aural flows such as voice commands
- Evaluated voice-controlled aural flows in the context of walking and driving

NSF-funded research project “Navigating the Aural Web” September 2010 – February 2014
PI: Dr. Davide Bolchini

- Designed two advanced back navigation strategies in complex information architectures to support the “aural” consumption of web content for screen reader users
- Implemented both the advanced back navigations in Webtime prototype
- Evaluated the efficiency, effectiveness, navigation experience and cognitive load with 10 blind high-school students at Indiana School for the Blind and Visually Impaired

Priceline.com – Usability and Evaluative Methods I543 August 2010 – October 2010
Graduate Student in HCI

- Conducted scenario-based and heuristics evaluation inspection on the usability of priceline.com website
- Reported the usability problems and recommended strategies for improvements

Priceline.com – Usability and Evaluative Methods I543 October 2010 – December 2010
Graduate Student in HCI

- Conducted user testing with 10 participants across 5 tasks
- Reported success rate, time on task and other findings tasks by tasks
- Recommended strategies for improvements based on the user study findings

Healthcare Inventory Tracking Procurement & Scheduling (H.I.T.P.S) – HCI1 I541
August 2010 – October 2010
Graduate Student in HCI

- Designed a best-in-class system that is ultra-portable, intuitive, reliable, accurate and insightful, which will engage nurses who need a more efficient and effective way of managing the localized supply chain and logistics.
- Conducted the design evaluation with interviewing the users and the internal walkthrough
- Revised the interactive prototype based on the usability study findings

iSee (Advanced Visual File Management) – HCI1 I541 October 2010 – December 2010
Graduate Student in HCI

- Created a tool that takes the files attributes and plots them in a graphical way, such as color and scale
- Sketched the paper prototype of iSee system
- Conducted the design evaluation with interviewing the users and the cognitive walkthrough
- Reported the usability issues and revised the interactive prototype based on these findings

Healthy Living Kids – HCI2 I561 January 2011 – March 2011
Graduate Student in HCI

- Designed an application that could be both educational and enjoyable for children while promoting a healthy living lifestyle during a shopping experience at the grocery store
- Conducted contextual inquiry on 3 participating parents and their children
- Consolidated all the information from the previous work to build several work and activity models
- Sketched the low-fidelity prototype based on the synthesized data
- Conducted internal walkthrough and incorporated the feedback into a high-fidelity prototype

ChaCha Powered by People – HCI2 I561 March 2011 – May 2011
Graduate Student in HCI
- Conducted a two-week diary study with 8 participants on existing ChaCha iPhone application
- Developed a high-fidelity prototype based on the collected users' feedbacks

USPS Self - Service Kiosk January 2010 – May 2010
- Designed a user friendly kiosk prototype followed by user need testing, context analysis, user and task analysis
- Applied formative evaluation using Morae Recorder and Manager software, also applied heuristic evaluation

Orchid Resort and Villas Database September 2009 – December 2009
- Identified business problem, established business rules, implemented logical and physical design for the database objects
- Created profiles, users, roles, tables, views, functions, triggers, cursors, procedures and packages using PL/SQL
- Performed backup, recovery and performance tuning
- Designed and implemented Oracle forms and reports for user interface

NAYEB Restaurant Table Reservation and Food Ordering Database January 2009 – May 2009
- Designed a database using SQL Server 2005 as a back end and MS Access 2007 as a front end
- Created an ordering system for taking and managing restaurant customer's orders
- Created reporting system for the manager to generate daily sales reports to monitor sales and help in decision making

BMS College of Engineering, Bangalore, India February 2006 – May 2006
SARE'*e*'SHOPPING A Web Application
- Developed a website using asp.net (using C#) for online shopping of Indian Traditional Dress and Accessories

DEVELOPED INTERACTIVE SYSTEMS

Aural Navigation Flows on Rich Architectures (ANFORA News) May 2011 – December 2011
http://discern.uits.iu.edu:8670/ANFORA/
ANFORA is a novel design framework that remodels existing information architectures as linear, aural flows appropriate to listening experience.
- Modeled and implemented the database
- Designed and implemented the text-to-speech enabled interface
- Optimized the interface for Safari, iOS 4.0

Webtime September 2010- December 2010
http://discern.uits.iu.edu:8670/NSF_WEB/
Webtime introduces two novel navigation strategies implemented on its different versions. These two novel strategies provide back browsing shortcuts by leveraging the conceptual structure of content-rich websites for screen-reader users.
- Modeled and implemented the database
- Designed and implemented the three different versions of Webtime interface
- Optimized the interface for Internet Explorer V8.0 and Windows Eyes V7.5

MENTORING AND TEACHING EXPERIENCE

Mentored 2 Master students to analyze the video recorded data on a recent study conducted at the driving simulation lab, April 2015

Instructor of an Online Graduate Class, Interaction Design Methods (INFO-H543), 8 students, Fall 2014

Teaching Assistant, Interaction Design Practice (INFO-H541), Fall 2014

Mentored undergraduate and graduate students in the various projects in the USER Lab (PI: Davide Bolchini), May 2012 – August 2012

ANFORA: Investigating Aural Navigation Flows on Rich Architectures, Seminar (I) in Health Informatics, (INFO I667), November 8, 2012

Multimodal Research in HCI Examples from the User Simulation and Experience Research Lab (U.S.E.R. Lab), Seminar (I) in Human-Computer Interaction, (INFO I624), October 30, 2012

ANFORA: Investigating Aural Navigation Flows on Rich Architectures, Usability Principles in New Media Interfaces Class (NEWM450), April 19, 2012

REFEREED JOURNAL ARTICLES

Chattopadhyay, D., **Rohani Ghahari, R.**, Duke, J., D., & Bolchini, D. Understanding Advice Sharing among Physicians: Towards Trust-Based Clinical Alerts. Interacting with Computers. 2015. (*Accepted and to appear*)

Rohani Ghahari, R., George-Palilonis, J., & Bolchini, D. (2013). Mobile Web Browsing with Aural Flows: An Exploratory Study. International Journal of Human-Computer Interaction, 29(11), 717-742.

REFEREED PUBLICATIONS

Rohani Ghahari, R., Ferati, M., Yang, T., & Bolchini, D., Back navigation shortcuts for screen reader users, Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility (ASSETS '12), pp. 1-8. 22-24 Oct. 2012. doi: 10.1145/2384916.2384918

Yang, T., Ferati, M., Liu, Y., **Rohani Ghahari, R.**, Bolchini, D., Aural Browsing On-The-Go: Listening-based Back Navigation in Large Web Architectures, Proceedings of the thirtieth annual ACM SIGCHI Conference on Human Factors in Computing Systems (CHI '12), 2012.

Ghahari, Romisa Rohani; Bolchini, Davide; ANFORA: Investigating aural navigation flows on rich architectures, Proceedings of the 13th IEEE International Symposium on Web Systems Evolution (WSE), vol., no., pp.27-32, 30-30 Sept. 2011. doi: 10.1109/WSE.2011.6081816

MANUSCRIPTS IN REVIEW

Rohani Ghahari, R., George-Palilonis, J., Gahangir, H., Kaser L., & Bolchini, D. Semi-aural Interfaces: Investigating Voice-controlled Aural Flows. Interacting with Computers. 2015. (*In review*)

DOCTORAL CONSORTIUM PAPERS

Rohani Ghahari, R., Bolchini, D., Eyes-Free, Natural Interaction with Aural User Interfaces, Accepted abstract at Richard Tapia Celebration of Diversity in Computing Conference, February 7, 2013.

NON-REFEREED PUBLICATIONS, POSTERS AND INTERACTIVE DEMOS

Bolchini, D., **Rohani Ghahari, R.,** George-Palilonis, J., Moon, S., Archibald, C., Kaser, L., Eyes-free Web Browsing with Linkless Navigation, IUPUI Innovation to Enterprise Showcase & Forum, IUPUI Campus Center, Indianapolis (IN), November 28, 2012.

Rohani Ghahari, R., George-Palilonis, J., Bolchini, D., ANFORA – Aural Navigation Flows on Rich Architectures, poster presented at the 2012 Grace Hopper Celebration of Women in Computing (GHC), Baltimore Convention Center, Baltimore (MD), October 3-6, 2012.

Rohani Ghahari, R., George-Palilonis, J., Bolchini, D., ANFORA – Aural Navigation Flows on Rich Architectures, poster presented at the 2012 IUPUI Research Day, Student Showcase, IUPUI Campus Center, Indianapolis (IN), April 13, 2012.

Bolchini, D., Ferati, M., **Rohani Ghahari, R.**, Yang, T., Navigating the Aural Web, poster presented at the 2012 IUPUI Research Day, Faculty and Community Showcase, IUPUI Campus Center, Indianapolis (IN), April 13, 2012.

Rohani Ghahari, R., George-Palilonis, J., Bolchini, D., ANFORA – Aural Navigation Flows on Rich Architectures, poster presented at the 2012 Women in Technology, IUPUI Campus Center, Indianapolis, April 11, 2012.

Bolchini, D., Ferati, M., Liu, Y., Luebke, J., **Rohani Ghahari, R.**, Yang, T., Navigating the Aural Web, poster presented at the 2011 World Usability Day, organized by the Indiana Chapter of the Usability Professionals' Association (UPA) Indianapolis (IN), November 10, 2011.

Bolchini, D., **Rohani Ghahari, R.**, George-Palinonis, J., ANFORA – Aural Navigation Flows on Rich Architectures, poster presented at the 2011 World Usability Day, organized by the Indiana Chapter of the Usability Professionals' Association (UPA) Indianapolis (IN), November 10, 2011.

Bolchini, D., Ferati, M., Liu, Y., Luebke, J., **Rohani Ghahari, R.**, Yang, T., Navigating the Aural Web, invited poster presented at the 2011 IUPUI TRIP (Translating Research Into Practice) Showcase, Indianapolis, IUPUI Campus Center, September 12, 2011.

Bolchini, D., **Rohani Ghahari, R.**, George-Palinonis, J., ANFORA – Aural Navigation Flows on Rich Architectures, invited poster presented at the 2011 IUPUI TRIP (Translating Research Into Practice) Showcase, Indianapolis, IUPUI Campus Center, September 12, 2011.

Bolchini, D., Ferati, M., **Rohani Ghahari, R.**, Liu, Y., Luebke, J., Yang, T., Navigating the Aural Web, poster presented at the World Usability Day in Indianapolis, IUPUI Campus Center, November 11, 2010.

Bolchini, D., Ferati, M., **Rohani Ghahari, R.**, Liu, Y., Luebke, J., Yang, T., Navigating the Aural Web, poster presented at the World Usability Day in Indianapolis, IUPUI Campus Center, November 11, 2010.

Bolchini, D., Ferati, M., **Rohani Ghahari, R.**, Liu, Y., Luebke, J., Yang, T., Navigating the Aural Web, poster presented at the Indiana TechPoint Innovation Summit, Indiana Convention Center, October 27, 2010.

INVITED PRESENTATIONS AND TALKS

Back Navigation Shortcuts for Screen Reader Users, 14th international ACM SIGACCESS conference on Computers and accessibility (ASSETS '12), Boulder (CO), October 22, 2012.

Navigating the Aural Web, invited presentation at the Indiana Chapter of the Usability Professionals' Association (UPA), Indesign LCC, Indianapolis (IN), October 15, 2012.

ANFORA: Investigating Aural Navigation Flows on Rich Architectures, 13th IEEE Symposium on Web Systems Evolution, Williamsburg (VA), September 30, 2011.

INTERNSHIP EXPERIENCE

SAP, Newtown Square, PA — February 2014 – August 2014

User Experience Researcher Intern, Knowledge & Enablement Solution (K&ES) Team

- Evaluated the information architecture and overall experience of WordPress newsletter templates by interviewing end-users, identified usability issues and provided recommendations for improving the usability of all internal newsletters
- Validated the usability and values of preliminary web concept to disseminate knowledge management best practices by conducting user interviews, resulted in a successful launch of a portal page with best practices content
- Evaluated the usability of proposed redesign for a Tutorial & Help portal page of internal Human Resource System, resulted in a successful launch of the new HR portal
- Analyzed SAP Services LoB portal and presented the findings to enable stakeholders identify out dated content and improve the usage of the portal with relevant and up-to-date content
- Evaluated the usability of media sharing site through expert reviews, analytics, and user interviews, provided functional and interface recommendations for short-term improvements
- Validated knowledge management survey instrument by multivariate analysis of 2 pilot implementations of data from multiple departments and regions, conducted data analysis in SPSS and compared it to R with a publication report in progress
- Analyzed the correlation between system usability survey and net-promoter scores (2012-2013) using SPSS to understand the relationship between the KPIs, and overviewed the whole internal tool landscape for the next phase of the project
- Designed and conducted exploratory research around key information behaviors (e.g. bookmarking resources, tools, links, etc.) by interviewing employees on their current and desired practices
- Designed a prototype to support integrated bookmarking for employees used as part of the exploratory interviews as well as to illustrate parts of the overall experience strategy

SocialYell, New York — June 2009 – August 2009

Consultant

- Consulted the founder of SocialYell to plan the product development of the website
- Prioritized new features to implement website based on priority
- Made recommendations about including how to translate business requirements into technical requirements for development

HDFC Bank, Bangalore, India March 2008 – May 2008
Intern, Direct Sales Department
Market Research Analysis of TASC (Trusts, Associations, Societies and Clubs) Segment in Bangalore

- Contacted customers to find out with which bank they have their association account and what are the benefits they are getting from that specific bank
- Analyzed collected data from customers and determined banking requirements of this targeted segment
- Reported data to HDFC Bank Manager

HONORS AND AWARD

Richard Tapia Travel Scholarship to attend and present a poster February 2014 & 2015
At Richard Tapia Celebration of Diversity in Computing Conference

School of Informatics & Computing Travel Scholarship to attend and present a poster
At Grace Hopper Celebration of Women in Computing October 2014

Salesforce Scholarship to attend and present a poster October 2013
At Grace Hopper Celebration of Women in Computing

Richard Tapia Travel Scholarship to attend and present at doctoral consortium February 2013
At Richard Tapia Celebration of Diversity in Computing Conference

FactSet Scholarship to attend and present a poster October 2012
At Grace Hopper Celebration of Women in Computing

Graduate Research Assistantship August 2010 – Present
Indiana University Purdue University Indianapolis

Member of Honor Society for International Scholars (Phi Beta Delta) March 2010 – March 2011
Syracuse University

SERVICE ACTIVITIES

Reviewer for Journal and Conference Papers
August 2010 – Present

- Journal of the Institute for Ergonomics and Human Factors – 2014
- AVI'14 International Working Conference on Advanced Visual Interfaces – 2014
- Interacting with Computers – 2013

LANGUAGES

English, Persian (Farsi)

VOLUNTEER WORK

HCI International Conference, Orlando, FL July 2011
Student Volunteer

Syracuse University, Syracuse, NY
June 2009 – August 2009

Lillian and Emanuel Slutzker Center for International Services

- Helped the staff upon international students' arrival
- Gave introduction seminars to group of 5 to 8 new students at a time

REFERENCES

Dr. Davide Bolchini, dbolchin@iupui.edu

Dr. Mark Pfaff, mpfaff@iupui.edu