Advanced Auditory Menus for Universal Access to Electronic Devices

Myounghoon Jeon, Benjamin Davison, Jeff Wilson, Paramesh Raman, & Bruce N. Walker

mh.jeon@gatech.edu, bruce.walker@psych.gatech.edu

Sonification Lab, School of Psychology

Georgia Institute of Technology

654 Cherry Streeet, Atlanta, GA, 30332 USA

Introduction

Most electronic devices, ranging from desktop computers and mobile phones, to in-vehicle "infotainment" systems, can be navigated using a menu structure. In cases where the user cannot look at or cannot see the interface, an auditory menu can provide access to the device's functions. Traditional auditory menus involve text-to-speech (TTS) synthesis of the menu items. Our research is taking this type of interface to the next generation of interaction. To this end, we are studying advances to the TTS itself, as well as additional (non-speech) auditory cues before, during, or after the TTS phrases. Our advanced auditory menus use a range of novel techniques like spearcons, spindexes, and auditory scrollbars to provide fast, efficient, error-free, and enjoyable interaction with an auditory or multimodal menu system. We evaluate the effectiveness, efficiency, learnability, and subjective impressions of all these enhancements, with a range of enhancement designs, in a variety of use situations, on several devices, and with a range of user populations. This paper provides an overview of the recently obtained empirical results of a series of research on advanced auditory menus project.

Speech Sound Menus

The use of speech (often via text-to-speech synthesis) is the most obvious way to convey

menu items with sound, instead of visual text. Most auditory menus to date have adopted speech interfaces, and many of the implementations have focused on using speech alone. However, even with regard to the speech itself, there remains much research to be done to optimize the user experience. As one simple example, it remains unresolved whether, or in what circumstances, users prefer a male or female voice in the user interface. Further, many developers have depended on naïve or arbitrary mappings between text and speech when developing auditory menus.

One specific example is the representation of menu items that are not currently available. There have been many inelegant approaches in speech menus, all of them seeming to ignore how the visual interface simply uses a change in text color (i.e., they are the grayed-out items in a visual menu). In a recent study [1] we adopted the simple idea of using a whispered voice for unavailable items in auditory menus, and a regular voice for available items. We also studied the possible use of male voice for available, and female for unavailable items, or vice versa. Loudness levels were also investigated. Users preferred female voice over male voice; the same voice gender over different voice gender; and whispered voice over attenuated voice for unavailable items. Moreover, whispering unavailable items outperformed simply skipping unavailable items after a moderate amount of practice. All of these data suggest that designers should go beyond the naïve translation from text into speech when creating auditory systems.

Non-Speech Menu Sounds

Just as text is not the only feedback in a visual system, speech is not the entirety of an auditory menu. Recently, non-speech auditory enhancements have been studied as a way to compensate for the limitations of speech. An *Auditory Scrollbar* [2] is a set of sounds that help

the user know how long an auditory menu is, and also which item in that list is currently highlighted. This can improve navigation, and also help listeners develop a mental model of the menu structure. In an auditory scrollbar study with both sighted and vision impaired participants, we demonstrated the utility of adding a “proportional” auditory scroll bars.

Speech + Non-Speech Menu Sound

Just as modern GUI menus can include both text and little icons, our relatively new auditory analogies, *Spearcons* [3] and *Spindex* [4], were introduced to overcome the shortcomings of either text-only (speech) or icon-only (auditory icons or earcons) menus.

Spearcons can be produced by speeding up spoken phrases, even to the point where the resulting sound is no longer comprehensible as a particular word. These unique non-speech sounds blend the benefits of text and icons because of their acoustic relation between the spearcons and the original speech phrases. Accordingly, spearcons are easier to learn than other audio icons [5]. The use of spearcons has enhanced navigational efficiency on the spoken auditory menus of two dimensional interfaces [6] as well as one dimensional menus. Also, we are looking at the possibility that we may replace long TTS phrases with short spearcons in a certain context such as the push-menus in telephone interactive voice systems.

A spindex is created by associating an auditory cue with each menu item, in which the cue is based on the pronunciation of the first letter of each menu item. For instance, the spindex cue for “Apple” would be a sound based on the spoken sound “A”. The set of spindex cues in an alphabetical auditory menu is analogous to the visual index tabs in an encyclopedia. The benefit of the spindex cues is most evident in long menus with many items [4] because they lead to per-item speedups. Also, since spindex cues are part of the original word and are natural sounds, they

do not require training. These enhanced prefix sounds improved navigation efficiency and users' satisfaction even in dual task contexts such as menu navigation while driving [7]. In our most recent research with visually impaired users, the spindex + TTS condition also enhanced navigation time and was preferred over the TTS-only condition. Now, we plan further research using touch screen devices with spindex-enhanced menus.

Conclusion and Future Directions

As described above, we have been applying Universal Design and Human Factors principles, along with creative innovations, to create effective and well-liked advanced auditory menus. We are extending our research to understand where and when it is best to use these of the basic elements (e.g., spindex, spearcons) alone or together.

Another direction is to integrate auditory display with other modalities including visual and tactile display. We hope to tease out the best combinations of these modalities, and study their dynamic interaction (e.g., in what situation; how differently several modalities interact; whether they compete or compensate for each other; and when additional modality is redundant).

One of our directions of auditory display research is on aesthetics, which is directly related to users' acceptance of the display. Even though many researchers point out that aesthetic and annoyance issues are more important in auditory displays than in other displays, to date, research in the field has mainly focused on performance issues. Users might turn off an annoying sound even though the presence of that sound enhances performance with the system. Likewise, well-designed system sounds can improve the aesthetic experience of an interface without changing performance. Therefore, more research is needed to determine which parameters improve the aesthetic aspects of auditory display. Through these multi-faceted efforts, we should

be able to obtain subjective satisfaction as well as usability and universal accessibility of the devices and can ultimately provide essential information to various user populations efficiently and pleasantly.

References

[1] M. Jeon, S., Gupta, and B. N. Walker (submitted), "Auditory menus are not just spoken visual menus: A case study of "unavailable" menu items," *in the SIGCHI Conference on Human Factors in Computing Systems (CHI 2010)*, Atlanta, GA, April, 2010.

[2] P. Yalla, and B. N. Walker, "Advanced auditory menus: Design and evaluation of auditory scroll bars," *in the 10th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS2008)*, Halifax, Canada, 2008, pp. 105-112

[3] B. N. Walker, A. Nance, and J. Lindsay, "Spearcons: Speech-based earcons improve navigation performance in auditory menus," *in the 13th International Conference on Auditory Display (ICAD2006)*, London, UK, 2006, pp. 95-98.

[4] M. Jeon, and B. N. Walker, ""Spindex": Accelerated initial speech sounds improve navigation performance in auditory menus," *in the Annual Meeting of the Human Factors and Ergonomics Society (HFES 2009)*, San Antonio, TX, 2009, pp. TBD.

[5] T. Dingler, J. Lindsay, and B. N. Walker, "Learnability of sound cues for environmental features: Auditory icons, earcons, spearcons, and speech," *in the 15th International Conference on Auditory Display (ICAD 08)*, Paris, France, 2008.

[6] D. Palladino, and B. N. Walker, "Navigation efficiency of two dimensional auditory menus using spearcon enhancements," *in the Annual Meeting of the Human Factors and*

Ergonomics Society (HFES 2008), New York, NY (22-26 September), 2008, pp. 1262-1266.

[7] M. Jeon, B. K. Davison, M. A. Nees, J. Wilson, and B. N. Walker, "Enhanced auditory menu cues improve dual task performance and are preferred with in-vehicle technologies," *in the First International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2009)*, Essen, Germany, 2009, pp. TBD.