

User Preferences for Auditory Device-driven Menu Navigation

Elsa Eiriksdottir, Micheal Nees, Jeff Lindsay and Raymond Stanley
School of Psychology, Georgia Institute of Technology
Atlanta, Georgia

A device-driven auditory menu interface that “pushes” the menu options to the user and requires minimal user input can be beneficial to users with both limited vision and input range. Presentation rate is important for auditory menus; they need to be slow enough for comprehension and responding, but fast enough to prevent frustration. User preferences for inter-item gap on an auditory device-driven cell phone menu were investigated. Participants ($n=11$) navigated a two-level auditory menu to target a randomly chosen item. Before each trial participants set the inter-item gap by manipulating a slider. Results showed a mean preference for .853 second inter-item gap in experimental trials, and a significant decrease in preferences between practice (15) and experimental (30) trials [$t(10) = 4.28, p = .002$]. Accuracy was 91.5% over all the trials. The findings indicate that it is feasible to design a limited-input cell phone with an auditory device-driven menu.

INTRODUCTION

The successful use of a menu system requires two distinct subtasks of the user: 1) She/he must traverse the menu options, and 2) she/he must provide input to indicate the selection of a target. Most traditional menu systems require the user to perform the navigation to the target (by “pulling” or “scrolling” through the menu system to arrive at the desired selection). An alternative approach is a device-driven navigation system, whereby the interface facilitates navigation by “pushing” the menu to the user (e.g., telephone-based interfaces, see Brewster, 1997).

The primary advantage of a device-driven menu navigation interface is that it decreases the amount of input required of the user. In fact, when the device “pushes” the menu options to the user, menu navigation can be accomplished with just one form of input used to select the desired target (e.g., a single button, voice command, etc.). This advantage can be particularly important for some classes of users. Some users’ capabilities do not permit a wide range of input options; Mankoff, Dey, Moore, & Batra (2002) dubbed this a “low bandwidth input” limitation. One example of a user class with this type of limitation is physically impaired users, who may not possess a full range of input responses. Also, elderly users of menu-oriented devices can experience difficulty with some input options due to dexterity losses associated with aging. Finally, children and cognitively impaired users may struggle to remember a large number of input options. In addition to user limitations, certain environments (e.g., ones in which gloves are required) and tasks (e.g., when the eyes are busy) may limit the available capacity for input from users.

Research has also indicated that more complex menus and navigation functions can inhibit performance with devices (e.g., Ziefle, 2002). A well-designed, device-driven menu structure with simple and limited input functions could potentially benefit the entire user population for some devices or specific tasks.

The auditory modality seems particularly suitable for presentation of device-driven menus. Hearing is not affected by the aforementioned “eyes busy” task restrictions. Furthermore, for devices such as cell phones, the small display size may present problems for blind or visually impaired users (e.g., the elderly, see Omori, Watanabe, Takai, Takada, & Miyao, 2002). As Marsden, Thimbleby, Jones, and Gillary (2002) point out, most cell phones can only concurrently display a very limited number of menu options (i.e., as few as one) at any given time.

When designing a device-driven auditory menu the rate of item presentation is important. A presentation rate that is either too hurried or too slow can cause frustration for the user. A presentation rate that is too hurried can lead to difficulties in comprehension and responding. Conversely, a rate of presentation that is too slow can cause frustration by making the user wait for information. It is therefore necessary to determine the rate of item presentation that is comfortable for the user. Auditory presentation rate can be divided into speech rate and inter-item gap. This study gathered information regarding participants’ preferences for inter-item gaps (while keeping speech rate constant) in an auditory, device-driven menu navigation task. The researchers expected to find that users preferred a small inter-item gap.

METHOD

Participants

Participants ($n=11$) were recruited from undergraduate psychology courses at the Georgia Institute of Technology. Participants were compensated with course credit.

Apparatus

All visual presentations were made on a 17 in (43.2 cm) Apple studio display. All auditory presentations were

delivered via an Apple G4 system speaker. All stimulus presentations and data collection were conducted using MatLab.

Stimuli

The stimulus set employed in the current study featured the spoken menu entries from a Nokia cell phone menu structure. Stimuli were created as .WAV files with AT&T’s text-to-speech conversion (www.research.att.com/projects/tts/demo.html). All stimuli were created using the same female voice and compressed to 60% of their original length using the Audacity sound editing program. The compression of the speech stimuli was controlled so that it did not result in pitch shifting.

Menu structure

The menu structure in the current study featured 8 top-level or first-level menu items (see Figure 1 below). A varied number of sub-items or second-level items were nested under each top-level menu. Most of these menu items were derived from the menu items on a Nokia cell phone. For each trial, the top-level menu items were randomly ordered and the items in the sub-menus were randomly ordered, such that a given sub-menu item always fell under the same top-level menu from trial to trial. This random ordering was implemented in order to prevent anticipation of a given target (and therefore to discourage participants from responding before a target item was presented). Furthermore, participants were provided with a visual aid (Figure 1) that indicated the first-level menu heading to which the sub-menu targets belonged. This was provided to ensure that participants would have the necessary information to locate each target on each trial. It is important to note that the current study was not concerned with the arrangement of the menus per se; we merely sought to determine an optimal inter-item time gap.

Messages Text messages Multimedia messages Voice messages Message settings Service command	Call Log Missed calls Received calls Dialed numbers Delete recent call Call timers GPRS data counter GPRS connection timer Message counter	Profiles Normal Silent Meeting Outdoor My profile 1 My profile 2	Settings Tone settings Time settings Call settings Phone settings Display settings Network Services Security settings Restore factory settings
Organizer Alarm clock Calendar To-do list	Games Parachute Solitaire Combat Rad Racer	Extras Calculator Stopwatch Voice commands	Contacts Find entry Add entry Speed dial numbers My phone number

Figure 1. Menu structure. Top-level menu headers are in bold.

Procedure and Task

Informed consent was obtained from participants, and they were seated in front of the computer. Participants received

instructions that included a detailed description of the task. Before a trial, participants were presented with a screen that featured a slider. The starting slider position was in the center of its possible range. Participants were instructed to use the slider to adjust the inter-stimulus (i.e., inter-item) gap for auditory menus in the upcoming trial. Participants were requested to select a between-item length that was both comfortable and efficient. The available inter-item gap ranged from 250 ms to 2250 ms. It is of note that the minimum inter-item gap created a considerable perceptual delay between items, but 250 ms represents a reasonable lower bound to allow for an input reaction after any given item.

After choosing a slider setting, participants were given a text presentation on the screen of a target item to locate within the menu structure. Target items were always second-level menu entities. Thus, the task was to select the higher level menu item containing the target with a key press and then select the target itself with another key press.

Participants then pressed a key to begin the trial. Once a trial began they had no control over the rate of presentation of the stimuli. Participants would hear the speech presentation of the randomly ordered, first-level menu items, that were presented with the participant selected inter-stimulus gap. When participants pressed the key, they would hear the second-level sub-menu that corresponded to their top-level selection. A trial ended when participants pressed a key to select a sub-menu target.

Participants were given 15 practice trials and 30 test trials, for a total of 45 identical trials of the task. The primary measures in the current study were participants’ preferred inter-item slider settings, as well as their accuracy in responses.

RESULTS

Over all 45 trials, the mean slider setting was .998 seconds (*SD* = .65). Participants accurately found the target in 91.5% of trials. There was no significant difference in slider setting for correct versus incorrect trials. (.9969 versus 1.011 seconds, respectively). For practice trials (trials 1-15), the mean slider setting was 1.288 seconds (*SD* = .68), and participants responded with 84% accuracy.

For experimental trials (trials 16-45), the mean slider setting was .853 seconds (*SD* = .58) seconds, and participants responded with 95% accuracy. The mean slider setting difference between practice and experimental trials was significant [$t(10) = 4.28, p = .002$] such that participants decreased the time between menu items for test trials compared to practice trials. The accuracy mean difference between practice and experimental trials was not significant [$t(10) = -2.15, p = .057$].

Figure 2 shows the mean gap chosen for each trial. In the practice trials there was a considerably decrease in the selected length of the inter-item gap, from approximately 2 seconds to 1 second. In the experimental trials the selected

length of the inter-item gap continued to shorten gradually, and was 0.73 minutes on average in the last five trials.

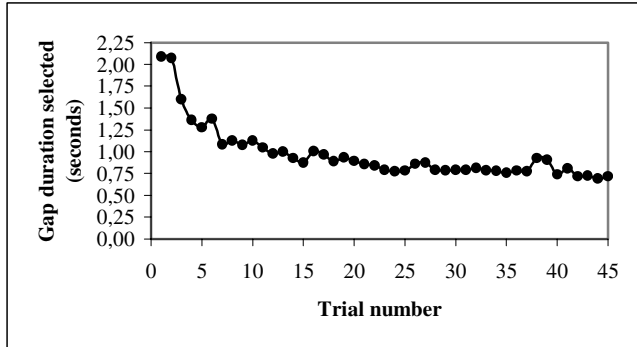


Figure 2. Mean gap duration selected on each trial.

CONCLUSIONS

This study was performed to describe the user preferences of the gap between menu items on an auditory, device-driven menu. The preferred gap was about .85 seconds.

Interestingly, there was no floor effect. People did not default to the fastest available setting. The results indicated that participants decreased the slider setting over time. Slider settings did not have a significant impact on accuracy, and accuracy did not improve statistically over time. However, a difference in 85% accuracy for practice test trials versus 95% accuracy for test trials likely represents an improvement over time with practical relevance. It is of note that participants in this study were not allowed to “go back” to correct errors in first-level menu selection. A “go back” option would most likely increase time to target as well as accuracy in instances where the wrong upper-level menu was selected.

Based on the results, it would seem feasible to design a limited-input cell phone with a device-driven menu such as the one used on the study. There are many obvious users for whom such a phone could potentially be very beneficial, as well as for certain tasks or in certain environments. A device-driven, single input cell phone interface could be implemented without a stigmatizing detriment in aesthetics, and visual appeal has been shown to be an important indicator of device satisfaction in cell phones (e.g., Moody & Burtner, 2005; Yun, Han, Hong, & Kim, 2003).

Although critics of a device-driven auditory menu navigation system for cell phones might suggest that the simplification of the input and menu structure will result in a loss of functionality, a study involving interviews with blind and disabled participants suggested that this population prefers “low-tier or limited function phones” (Smith-Jackson, Nussbaum, & Mooney, 2003, p. 556).

Future research should examine the rate at which menu items are presented to find an optimal rate of speech presentation and also to look for any potential interaction between the rate of speech presentation and the preferred inter-item menu gap. Studies have indicated that blind users in particular can use speech rates much faster than normal

speech with text-to-speech conversion software (Asakawa, Takagi, Ino, & Ifukube, 2003).

REFERENCES

- Asakawa, C., Takagi, H., Ino, S., & Ifukube, T. (2003). Maximum listening speeds for the blind. *Proceedings of the 2003 International Conference on Auditory Display*, 276-279.
- Brewster, S. (1997). Navigating telephone-based interfaces with earcons. *Proceedings of BCS HCI '97*, 39-56.
- Marsden, G., Thimbleby, H., Jones, M., & Gillary, P. (2002). Data structures in the design of interfaces. *Personal and Ubiquitous Computing*, 6(2), 132-140.
- Mankoff, J., Dey, A., Moore, M., & Batra, U. (2002). Web accessibility for low bandwidth input. *Proceedings of the Fifth International ACM SIGCAPH Conference on Assistive Technologies (ASSETS 2002)*.
- Moody, L., & Burtner, J. (2005). Evaluating both usability and desirability in the evaluation of cell phones: Toward an integrated model. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*, 1758-1761.
- Omori, M., Watanabe, T., Takai, J., Takada, H., & Miyao, M. (2002). Visibility and characteristics of the mobile phones for elderly people. *Behaviour & Information Technology*, 21(5), 313-316.
- Smith-Jackson, T.L., Nussbaum, M.A., & Mooney, A.M. (2003). Accessible cell phone design: Development and application of a needs analysis framework. *Disability and Rehabilitation*, 25(10), 549-560.
- Yun, M.H., Han, S.H., Hong, S.W., & Kim, J. (2003). Incorporating user satisfaction into the look-and-feel of mobile phone design. *Ergonomic*, 46(13/14), 1423-1440.
- Ziefle, M. (2002). The influence of user expertise and phone complexity on performance, ease of use and learnability of different mobile phones. *Behavior and Information Technology*, 21(5), 303-311.