MEMORY FOR AUDITORY ICONS AND EARCONS WITH LOCALIZATION CUES

Terri L. Bonebright

Psychology Department
DePauw University
Greencastle, IN USA 46135
tbone@depauw.edu

Michael A. Nees

Sonification Lab, School of Psychology
Georgia Institute of Technology
Atlanta, GA, USA 30332
mnees@gatech.edu

ABSTRACT

This study was designed to test whether associations between visual icons on a computer screen and auditory icons (environmental sounds that have a direct association with an object) or earcons (synthetic sounds that have no direct association with an object) are easier to learn. In addition, localization of sound presentation relative to the position of the icons on the screen was tested. Results revealed that participants made faster and more correct matches between visual icons and auditory icons than between visual icons and earcons. The results also suggested that localization may be a useful cue for learning the associations between icons and their auditory counterparts; however, more research is needed to provide conclusive evidence.

1. INTRODUCTION

Computer software manufacturers have been adding sound to programs to enhance the users' experience since the development of hardware capable of reliably presenting sounds to users. In many cases, sounds have been added to computer software packages simply as an additional feature without testing to determine if these sounds add to the usability of the programs. Recently researchers have started to systematically test how users interact with such sounds to examine which types of sounds are most effective for the user navigating the computer environment. Of particular interest for the present study is the work that has been done investigating different types of sounds that are used as auditory adjuncts to the visual icons that are a ubiquitous part of computer desktops. There are two main types of these sounds recognized by researchers: auditory icons and earcons [1]. Auditory icons represent visual desktop icons with ecological sounds that are directly associated with a sound the object would make in the environment [2]. For example, the sound of a pair of scissors cutting paper might be used to represent the "cut" menu function that has a scissor visual icon. In contrast, earcons are synthetic sounds that are not directly related to the object they represent [3]. An example of an earcon would be three ascending tones used to represent the "save" function that has a diskette visual icon.

Researchers have proposed design principles for auditory icons and earcons [3,4,5], investigated the overall usefulness of earcons in navigating computer software [5,6] and examined the utility of such sounds for the visually impaired [7]. The present study was designed to contribute to the basic understanding of how users learn the associations between icons and their auditory counterparts. Specifically, we believed that participants would be able to learn associations for icons with auditory icons more easily than with earcons due to the non-arbitrary mapping for auditory icon sounds. For a matching memory task, we hypothesized that there would be more correct matches for visual icons paired with auditory icons versus earcons. We also predicted that matches would be made more quickly between icons and auditory icons than between icons and earcons. In addition, localization of sound was tested to determine if this would impact memory for the sound associations with the icons. The auditory icons or earcons were presented to participants either through both ears (non-localized) or in the right ear for icons in the right column on the screen, the left ear for those in the left column, and to both ears for the icons in the middle column (localized). The hypothesis was that localized sound presentation should lead to more correct matches with icons for both auditory icons and earcons than non-localized sounds.

2. METHOD

2.1 Participants

Participants were 100 undergraduate students (70 females and 30 males) with a range of 18 to 22 years of age (M = 19.5), who received course credit in a psychology class. The racial background of the students showed that the majority (98%) were caucasian.

2.2 Apparatus

Four PowerMacs were used for the stimulus preparation and data collection procedure. Sony MDR-CD850 stereo headphones were used to present the sounds.

The visual stimuli for the experiment were two sets of 15 icons selected from a set that is available for PowerMacs. The visual stimuli appeared in two sets of 15 icons arranged in 3 columns and 5 rows (see Figures 1 & 2). For each set of 15 sounds, there were 3 different layouts for the positions of the icons, and participants were randomly assigned to one of these layouts by the computer program.

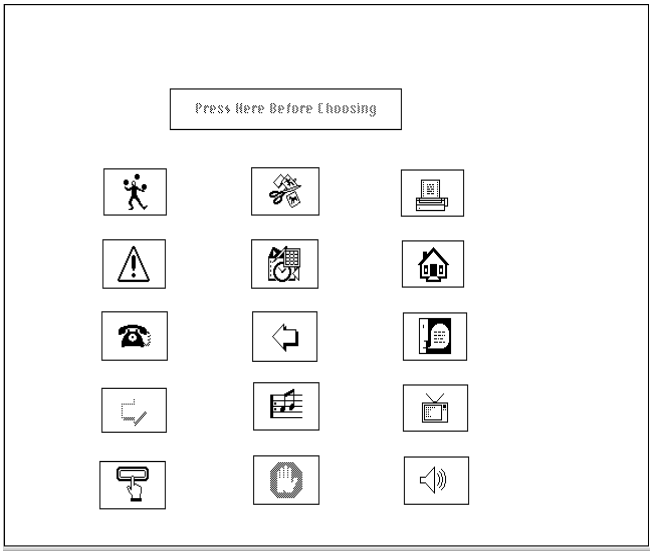

Figure 1. *Set 1 of the icons.*

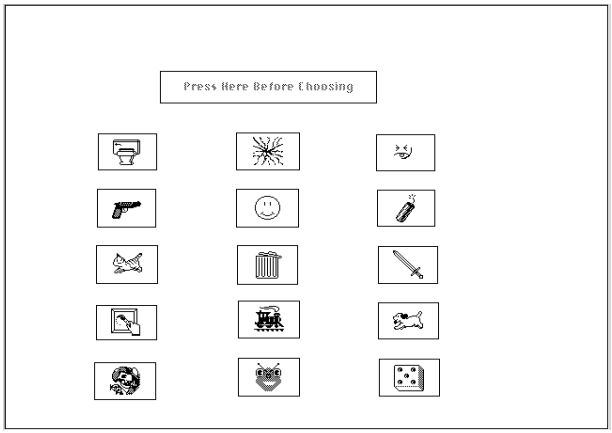

Figure 2. *Set 2 of the icons*

The auditory icons and earcons were selected for each of the visual icons in the two sets using synthesized sound files from sound effects collections and from sound files available at various websites. For example, an icon of a telephone had an auditory icon of a telephone ring and an earcon with 4 notes alternating in high and low frequency. In addition, the auditory icons and earcons were recorded for localization. For non-localized sounds, the recordings were mono-sounds presented binaurally so that both channels would be used during the data collection procedure. The localized sounds corresponded to the placement of the icons in columns such that the icons in the left column were played in the left channel, those in the right column were played in the right channel, and the icons in the middle were presented in both channels. In addition, the arrow icons had sounds that panned from one channel to the other in the direction of the arrow.

2.3. Procedure

Participants were tested individually for the 1-hour sessions that consisted of two tasks. All instructions and data collection were administered via computer except for the follow-up questionnaire. Participants first completed a set of demographic questions, including age, sex, and computer experience. Next, they were given the following instructions for the learning task of the study:

> You are being asked to do a learning task for associating sounds with icons (pictures) on a computer screen. There will be two parts to the computerized portion of the procedure.
>
> **Part 1: Learning Associations**
> For the first part of the procedure, you will be asked to learn the sound paired with each icon. As shown below, you will be presented with a set of 15 icons on the computer screen. Before you can listen to the sound for an icon, you will need to click on the bar on the top of the screen that says "press here before choosing". Please make sure that you wait to click until the letters have changed from gray to black. Click on the bar, which will activate the icons below (they will change from a gray color to black), and then click on one of the icons. This will play the associated sound. You should listen to each sound twice before going to the next icon. You should also begin with the icon in the upper left corner and move down the column before you go to the next column of icons. You will repeat the process of clicking on the bar and then the icon for each time you listen to a sound. Please make sure you are patient and wait for the bar or the icons to change to black to indicate that they are active before you click on either. Once you have listened to all the sounds twice, the program will take a moment and then it will present you with a second set of icons automatically. After you go through the second set, you will be given a memory task.

Each participant worked with both sets of icons, and the computer randomly determined which set each participant received first. However, participants were exposed to only one of the sets of sounds for each of the sets of icons. Thus, if a participant had auditory icons for the first set of icons, he or she would have earcons for the second set. The auditory icons and earcons were balanced across participants so that there would be a complete set of data for both types of sounds for both sets of icons. Finally, the computer program randomly determined whether the participant had localized or non-localized auditory icons and earcons. Participants were not told about the localization of the sound source since we were interested in whether they would notice and use these cues without prompting.

Once participants finished the learning task, they were given the following instructions for the second task, which was the memory task:

> This task will test your ability to remember which sound was associated with an icon. You will see a screen like the one presented above with the addition of a bar that says "repeat sound" that will be positioned below the "press here before choosing". When you click on the "press here before choosing" bar, you will hear a sound. Then you should look

> for the icon you think was associated with the sound and click on that icon. If you wish to hear the sound more than once, simply click on the "repeat sound" bar. Remember to wait until the bar and the icons are active before clicking on either. Do the best you can with the memory task. Even if you aren't sure which icon the sound belongs too, make the best choice you can! You must have an icon associated with each of the sounds to complete the task.

To finish the procedure, participants filled out a follow-up questionnaire and were fully debriefed by the experimenter.

3. RESULTS

Data collected included number of correct matches between the icons and the sounds, the reaction time for each match (measured in "ticks" : 1 tick = 1/60 of a second) and the number of times each sound was played during the memory task. A mixed factor 2 X 2 ANOVA with localization (BG factor) and number of correct matches (WG factor) revealed a non-significant interaction between the two factors. The main effect for localization was marginally significant, $F(1,98)=3.28$, $p=.075$, and showed a trend that the localized sounds ($M=10.92$, $SD=3.94$) were easier to correctly match with the icons than the non-localized sounds ($M=10.25$, $SD=4.55$). The main effect for number of times the sounds were correctly matched with the icons was significant, $F(1,98)=6.78$, $p=.01$. Inspection of the means revealed that icons were correctly matched with the auditory icons ($M=11.53$, $SD=4.02$) more often than with the earcons ($M=9.58$, $SD=4.64$).

A mixed factor 2 X 2 X 2 ANOVA was performed with localization (BG factor) and reaction time and number of times each sound was played (WG factors). The results showed a non-significant 3-way interaction as well as a non-significant 2-way interaction between number of times played and localization. The main effects for both of these variables failed to reach significance as well. However, there was a marginally significant interaction between localization and reaction time, $F(1,98)=3.48$, $p=.06$) and a main effect for reaction time, $F(1,98)=1364.31$, $p<.001$. The interpretation of the marginally significant interaction shows that localization had no effect on the reaction time for the earcons (localized mean = 56.86, $SD=27.6$; non-localized mean = 55.74, $SD=21.16$), but that the auditory icons showed a faster reaction time for the localized ($M=45.37$, $SD=21.84$) than the non-localized sounds ($M=57.11$, $SD=27.89$). The main effect for reaction time showed that participants were significantly faster at matching the icons with auditory icons ($M=51.71$, $SD=25.84$) than with the earcons ($M=56.26$, $SD=24.22$)

The follow-up questionnaire asked participants to rate the difficulty of learning the associations between the icons and the auditory icons and earcons using a 7-point rating scale (1=very easy to 7=very difficult). An ANOVA revealed that participants felt that it was easier to learn the associations between the icons and the auditory icons ($M=1.48$, $SD=.97$) than the earcons ($M=5.32$, $SD=1.27$), $F(1,99)=481.09$, $p<.001$. Participants were also asked what strategies they used to learn the associations between the icons and the two types of sounds. For the auditory icons, the majority (95%) responded that they used previous knowledge of the types of sounds such objects make. However, for the earcons, participants tried a variety of strategies: 30% simply tried memorization; 35% tried to form a relationship; 27% tried to think of a story or images that formed associations; and 8% used the localized sounds as cues. Finally participants were asked if any of the sounds were annoying to them: 33% responded that none of them were annoying while 55% said that the high pitched tones were bothersome and 15% responded that the loud sounds were difficult to listen to.

4. DISCUSSION

The results support the hypothesis that associations between icons and auditory icons are easier for users to "learn" than those with earcons. In fact, it is apparent that the participants were not actually learning these relationships, but rather they were using pre-stored semantic connections between the objects and the sounds. However, support for the second hypothesis, that localization would be helpful for learning the associations, was not as conclusive since the results with this factor were marginally significant. A follow-up study could investigate whether telling participants explicitly about the localization for the sounds relative to the icons would lead to better use of this cue. Comments made by participants in the final questionnaire revealed that some of them did notice and make use of the localization cue; however, these remarks were only made by a small number of the participants (8 out of 100). Future research should concentrate on longitudinal designs to determine how long it takes users to learn earcons, especially since these sounds can be shorter in duration and use less computer memory than most auditory icons [6]. In addition, further investigation of the usefulness of such auditory cues for visually impaired users could lead to more effective tools for their use with computers.

5. REFERENCES

[1] S.A.Brewster, P.C.Wright, and A.D.N.Edwards, "A detailed investigation into the effectiveness of earcons,", In *Auditory Display: Sonification, Audification and Auditory Interfaces. Proc. of the first Int. Conf. on Auditory Display (ICAD). Santa Fe Institute Studies in the Sciences of Complexity,* vol.18, pp. 471-498, 1994.

[2] W.Gaver, "The SonicFinder: An interface that uses auditory icons," *Human-Computer Interaction*, vol. 4, pp. 67-94, 1989.

[3] M.M.Blattner, D.A.Sumikawa, and R.M. Greenberg, "Earcons and icons: Their structure and common design principles," *Human-Computer Interaction*, vol. 4, 11-44, 1989.

[4] J. Hereford and W.Winn, "Non-speech sound in human-computer interaction: A review and design guidelines." *Journal of Educational Computing Research*, vol. 11, pp. 211-233, 1995.

[5] G.Kramer, "Some organizing principles for representing data with sound." In *Auditory Display: Sonification, Audification and Auditory Interfaces. Proc. of the first Int. Conf. on Auditory Display (ICAD). Santa Fe Institute Studies in the Sciences of Complexity*, vol. 18, pp. 83-221, 1994.

[6] S.A.Brewster, P.C.Wright, and A.D.N.Edwards, "Parallel earcons: Reducing the length of auditory messages." *International Journal of Human Computer Studies*, vol. 43, pp. 152-175, 1995.

[7] R.D.Stevens, A.D.N.Edwards, and P.A.Harling, "Access to mathematics for visually disabled student through multimodal interaction," *Human-Computer Interaction, v*ol. 12, pp. 47-92, 1997.