The Ecological AUI (Auditory User Interface) Design and Evaluation of User Acceptance for Various Tasks on Smartphones

Myounghoon Jeon[1] and Ju-Hwan Lee[2]

[1] Michigan Technological University, Houghton, MIchigan, USA
mjeon@mtu.edu
[2] Korean German Institute of Technology, Seoul, South Korea
jhlee@kgit.ac.kr

Abstract. With the rapid development of the touch screen technology, some usability issues of smartphones have been reported [1]. To tackle those user experience issues, there has been research on the use of non-speech sounds on the mobile devices [e.g., 2, 3-7]. However, most of them have focused on a single specific task of the device. Given the varying functions of the smartphone, the present study designed plausibly integrated auditory cues for diverse functions and evaluated user acceptance levels from the ecological interface design perspective. Results showed that sophisticated auditory design could change users' preference and acceptance of the interface and the extent depended on usage contexts. Overall, participants gave significantly higher scores on the functional satisfaction and the fun scales in the sonically-enhanced smartphones than in the no-sound condition. The balanced sound design may free users from auditory pollution and allow them to use their devices more pleasantly.

Keywords: Auditory user interface, ecological user interface design, smartphones, user acceptance.

1 Introduction

Smartphones have become one of the most important necessities in our daily lives. However, with the rapid development of the touch screen technology, some usability issues of smartphones have been reported [1]. Because smartphones and other touch screen devices usually have overlapped control and visual display areas and lack tactile feedback, the appropriate use of sounds could not only provide solutions to usability issues, but also enhance user experience [7]. Research [8] has supported this notion by showing that auditory feedback is the most effective modality in physical user interface satisfaction, followed by tactile and motion feedback.

Over the last two decades, there has been much research on the use of non-speech sounds to improve user interaction on the mobile devices [e.g., 2-7], but most of them have focused on a specific function or task of the device. It is not easy to find out literature that provides guidelines on how to lay out or integrate various auditory cues for complex functions depending on usage contexts. The present study attempts to

M. Kurosu (Ed.): Human-Computer Interaction, Part IV, HCII 2013, LNCS 8007, pp. 49–58, 2013.

design and evaluate a more integrated auditory user interface set to better represent the use of smartphones with sounds in varied situations from the perspective of the ecological interface design [9].

1.1 Types of Non-speech Auditory Cues on Mobile Devices

There are a number of types of non-speech auditory cues that have been applied to touch screen devices. Auditory icons [10] use representative part of sounds of objects, functions, and events which bear an analogic relationship with the object they represent. For example, Wilson and his colleagues employed auditory icons in their SWAN system [11], a system for wearable audio navigation for blind people. Earcons ("ear + icons") [12], on the other hand, use a short musical motive as symbolic representations of actions or objects. Earcons have shown superior performance compared to other less systematic sounds or no sound in a PDA [5] or a mobile phone [3]. Some researchers have also provided empirical guidelines for better aesthetics in creation of earcons [3, 13]. Auditory scrollbars also use musical sounds to represent a location of a user in a display (i.e., contextual information) as an analogy of visual scrollbars, which could be found on a computer application or a smartphone [14]. That previous study focused on a continuous scrollbar (the thumb could be located anywhere along the bar, whereas Yalla and Walker [15] examined the possibility of the use of discrete auditory scrollbars (the thumb can only be located at discrete, designated points) in mobile device menus. Yalla and Walker demonstrated the potential benefits of the proportionally mapped auditory scrollbars for visually impaired participants as well as sighted people. Brewster [14] once implemented a sonically enhanced widget set including buttons, menus, scrollbars, alert boxes, windows, and drag and drop on the desktop computer, but it focused on the use of only earcons and auditory scrollbars. Fairly recently, musicons ("music + icons") have been introduced to the HCI community [16]. Musicons are brief samples of well-known music used in the auditory interface design. Researchers provided some usage scenarios such as the use of musicons as a reminder at home and preliminary guidelines to create better musicons. However, a further validity test is needed for the relationship between musicons and their intended meanings in the interface, in addition to the discernability of musicons as original songs. As relatively new auditory cues that combines speech and non-speech sounds, spearcons ["speech earcons" 17] and spindex [6] were introduced into mobile devices to overcome the shortcomings of either purely non-speech sounds (auditory icons) or music-driven auditory cues (earcons, auditory scrollbars, and musicons). Spearcons use compressed speech which is produced by speeding up spoken phrases [17]. These unique sounds blend the benefits of speech and non-speech because of the acoustic relationship between the spearcons and the original speech phrases. The use of spearcons has enhanced navigational efficiency as well as subjective satisfaction on the various mobile devices [e.g., 4, 18]. A spindex ["speech index" 6] uses a short cue created based on the pronunciation of the first letter or phoneme of each spoken menu item. For instance, the spindex cue for "Super" would sound /es/ or even /s/. The set of spindex cues in an alphabetical auditory menu is analogous to the visual index tabs in a reference book (e.g., a large dictionary). The use of spindex cues has shown promising results in a mobile phone addressbook [6], on a touch screen device with various gesture styles [7], or in a dual task context [19].

To summarize, auditory display researchers have tried to use diverse non-speech auditory cues in mobile devices and have yielded successful outcomes as shown above. However, given the growing and varying functions and structures on the smartphone interface, more integrated research on the application of the sounds is needed for implementing sophisticated, but balanced auditory user interfaces.

1.2 Ecological User Interface Design

An ecological interface design (EID) approach focuses more on environments rather than on a specific task in complex and complicated systems [9]. This interface design framework aims to lessen mental workload and thus, enable users to more easily acquire advanced mental models about the system, by focusing on its entire architecture [20]. To this end, in the EID framework, various methods [e.g., abstraction hierarchy 21] have been used to determine what types of information should be displayed on the system interface and how the information should be arranged on different abstract levels. Given that environmental factors (e.g., noise, mask, interference) and an overall usage flow (e.g., harmony with one another) are as important in auditory displays as (or even more important than) in visual displays, the ecological interface design approach to auditory displays has been proposed and supported [e.g., 22, 23]. However, there has been rare empirical research on this topic.

From these backgrounds, the present study investigated users' acceptance of the plausibly integrated various auditory cues for specified purposes on a single touch screen smartphone. After identifying functions which auditory cues could be applied to through a preliminary function analysis of the target device, professional sound designers created several alternative sounds for each function. For all the task scenarios, we measured user acceptance levels and finally measured overall user experience of the presence of the sound. Based on this attempt, we expect that we could come up with a blueprint of the optimal layout of the auditory cues on the smartphone.

2 Method

2.1 Participants

Forty six (under) graduate students participated in this study (mean age = 23.4; female 27, male 19). All reported normal or corrected-to-normal vision and hearing, signed informed consent forms, and provided demographic details about age and gender.

2.2 Apparatus and Stimuli

Stimuli were presented using an LG LH 2300, smartphone with a 3 inch resistive wide full touch screen panel. The internal sound chip was used for sound rendering.

Participants listened to auditory stimuli using Sennheiser HD 202 headphones plugged into the phone's audio jack, and adjusted for fit and comfort.

Based on literature reviews on the guidelines of the sound applications to the electronic devices [24-26] and our function analysis of the target smartphone, we categorized generic functional groups to which sounds can be applied to facilitate user interaction as follows: (1) touch feedback, (2) sound widgets (3) hierarchical menu navigation, and (4) list menu navigation. A professional user interface designer, a sound designer, and an academic researcher iteratively discussed together and created all of the sound designs for this experiment. (1) Touch feedback: We focused on two types of touch tasks. The one was a flip, which is a pervasive gesture on the smartphone nowadays. The flip tones included a single tone, which provided a single feedback sound about users' flip action and revolving tones, which provided feedback sounds whenever each item passes by a specific zone until the rolling stops. These designs were made up of mechanical sounds (just like the wheeling sound on the Apple iPod), instead of an instrumental sound. The other touch feedback task was the dialing tone on the phone. The dialing tones included a single tone (the original sound of that smartphone), double tones (applying different sounds to "touch" and "release" actions each), and a voice (speaking out the touched number in addition to a single tone). (2) Sound widgets: For sound widgets, six different functions were included: an auditory scrollbar [e.g., 14, 15], auditory progress bar [e.g., 27] and auditory progress circulation, auditory check box, auditory toggle, and auditory pop-up. Auditory scrollbars and progress bars were studied before, but we extended the previous design and added alternatives. For the remaining functions, we devised new alternative ones. Two different auditory scrollbars were created. The one was the effect sound, which lasted by proportional length based on the distance that the scrollbar moved. Another was the location mapping sound, which was composed of four different notes. The first two notes presented the possible entire range of the scrollbar (e.g., if it goes down, top and bottom notes). The last two notes presented the current location of the scrollbar out of the range. (e.g., if it goes down and is in the middle of it, middle and bottom notes) The polarity of the sounds was changed depending on the moving direction of the scrollbar. The auditory progress bars involved three different designs. The first was the step sound, which got gradually faster as approaching the end. The second was the spatial sound, which was composed of one pitch, moving from left to right. The third was the pitch change, in which incremental tones with higher pitch were added as time goes by. For the auditory progress circulation, we created a single sound design, which was composed of a psychologically circulating sound called a 'Shepard tone'. The auditory check box sounds consisted of two alternatives. One was the single tone, which was heard whenever a user checks on a checkbox. Another was the selected number tone, in which the user could hear a total number of currently checked boxes whenever the user adds or subtracts a check mark (e.g., if a user checks the second checkbox, then he or she will hear two tones. If the user minuses a checkbox from the three marked checkboxes, he or she will also hear two tones because the currently marked boxes will be two). Two auditory toggle sounds were composed. In the first design, a single sound was generated whenever a user touches the toggle button. In the second design, if the toggle button had three different modes, it generated three different (but structurally similar) sounds (e.g., C-E-G, C-G-E, G-C-E). Four different pop-up sounds were devised. The first one was just a single tone that notified that 'there is a pop-up window'. The second was the continuous effect sound, which generated the sound intermittently until the user touches an OK button. The third was the simple repeated sound, in which a single tone was continuously repeated. The last one was the continuous melody, in which a melodious instrumental sound held out. (3) For the hierarchical menu navigation, sounds were tested in the two different functions, the menu depth and the menu content. The ambient menu depth cues were composed of the hierarchical background sound and unique background sound. The former included the same sound in the same depth, regardless of the functional category (e.g., if the menu structure has three different levels, only three distinct sounds are used in each depth). In contrast, the latter included a unique sound for each functional group, reflecting each depth. In both designs, the deeper level included incrementally more complex sounds (e.g., suppose the 1st depth includes a hi-hat sound. Then, the 2nd depth includes hi-hat and bass drum sounds. The 3rd depth includes hi-hat, bass drum, and snare drum sounds, which would provide navigation crumbs). The menu content auditory icons in each depth contained three different versions of sounds. The first one was a simple default sound. The second was the serial auditory icons that generated a couple of representative sounds of the functions in that category in a serial manner (e.g., generating a game sound and a camera shutter sound one by one when selecting the multimedia menu). The last one was the parallel auditory icons that generated a couple of representative sounds of the functions in that category in a parallel manner (e.g., generating a game sound and a camera shutter sound simultaneously). The serial auditory icons could clearly present functions, but they are relatively longer. In contrast, the parallel auditory icons could provide a quick auditory scan, but might be confusing. (4) List menu navigation: for the list menu navigation, two types of sounds were used. The first one was the single tone, which was the same as the single tone of the flip tone. Another sound condition used basic spindex cues. For instance, the spindex cue for "Michigan" would be a sound based on the spoken sound "M".

2.3 Design and Procedure

A within-subjects design was used in this experiment. Thus, one participant experienced all the sounds and compared them. Our experiment was designed in this way to focus more on the intra-participant's perception about overall integrated effects of various auditory cues on a single device, considering harmonization in between different tasks. After the consent form procedure, participants were seated in front of the desk on which they can play around with the smartphone. For each task, participants were instructed about the usage situation and how to control the smartphone and then, they conducted the task. The order of appearance of each task was counterbalanced across all participants. For every task, each condition was randomly presented to the participants. After completing each task, participants filled out the subjective questionnaire on the 'functional satisfaction', the 'fun', and the 'preference' using a seven-point Likert-type scale. Finally, participants provided comments on the study.

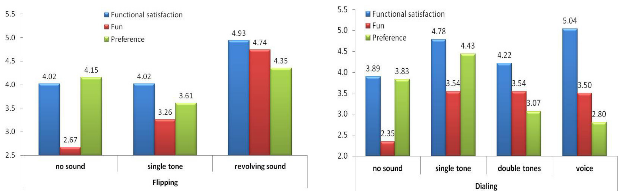
3 Results and Discussion

For all of the data analysis, we used repeated measures analysis of variance (ANOVA) first. Then, for the pairwise comparisons, paired-samples t-tests were used with a conservative alpha level (.01) instead of .05, across all the comparisons.

3.1 Touch Feedback

- Flipping tone: Participants showed higher functional satisfaction (M = 4.93) and fun (M = 4.74) scores in the revolving sound condition than in the no sound (M = 4.02; M = 2.67) or the single tone (M = 4.02; M = 3.26) conditions (ps < .01). Participants tended to prefer the revolving sound condition (M = 4.35) to the single tone condition (M = 3.61), but this difference did not lead to a traditionally significant level (p = .08). Based on the results, auditory user interface designers could consider employing revolving-type sounds to the flip function on smartphones.

- Dialing tone: For the functional satisfaction scale, the single tone (M = 4.78) and the voice (M = 5.04) conditions showed significantly higher scores (ps < .01) than the no-sound (M = 3.89) or the double tone (M = 4.22) conditions. Nevertheless, the voice (M = 2.80) condition was significantly less preferred than the single tone (M = 4.43) condition (p < .01), which means that even though the use of voice might functionally help users, they would choose a simple, brief single tone feedback for dialing, given that the dialing is a frequently used function. According to the results, the voice tone could be included in the setting option.



(a) (b)

Fig. 1. Rating results about touch feedback (flipping (a) and dialing (b))

3.2 Sound Widgets

- Auditory scrollbar: Participants showed significantly higher scores on the functional satisfaction, the fun, and the preference scales in the location mapping sound condition than in the no sound or the effect sound conditions (ps < .01). Note that in the preference rating scale, the effect sound condition was significantly lower than the no sound condition (p < .01). The use of inappropriate sounds may make users annoying.

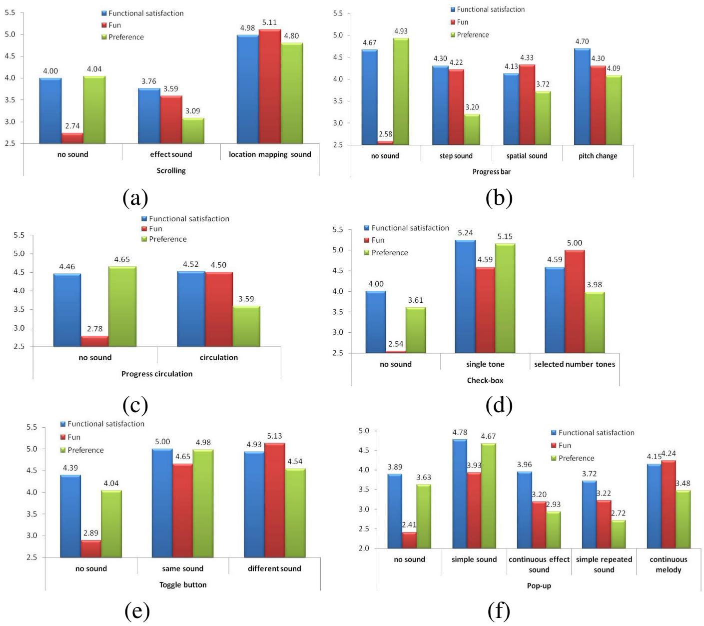


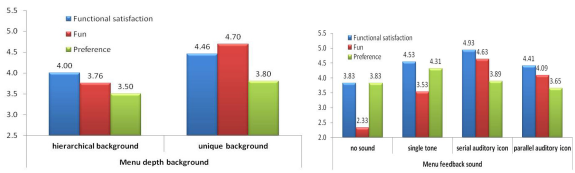
Fig. 2. Rating results about sound widgets (scrollbar (a) and progress bar (b), progress circulation (c), checkbox (d), toggle (e), and pop-up (f))

- Auditory progress bar and circulation: All the sound conditions showed significantly higher fun scores than the no sound condition ($ps < .01$), but showed lower preference scores. The pitch change condition showed significantly higher preference scores than the step sound or the spatial sound conditions ($ps < .01$), but it was still significantly lower than the no sound ($M = 4.93$) condition ($p < .01$). The results of the auditory progress circulations confirm this pattern. Whereas participants had fun in the circulation sound condition ($p < .01$) compared to the no sound condition, they preferred the no sound condition ($p < .01$) over the circulation sound. Previous research on auditory progress bars has shown that subjective workload can be lessened depending on the characteristics of the sounds [27]. However, in their experiment, there was no no-sound condition. Therefore, we may not over-generalize the results of auditory progress bars and circulations, but further research with more alternatives is needed.
- Auditory checkbox: Participants showed significantly higher functional satisfaction, fun, and preference in the single tone ($M = 5.24$; 4.59; 5.15) condition than in the no sound ($M = 4.0$; 2.54; 3.61) condition ($ps < .01$). For the fun scale, the selected number tone ($M = 5.0$) condition was also significantly higher than the no sound ($M = 2.54$) condition ($p < .01$). Moreover, given that all of the scores of the selected number tone condition were numerically higher than the no sound condition, sound designers could consider having this option in the setting menu even though they might want to have a single tone as a default.

- Auditory toggle: As to the auditory checkbox, participants favored the sound conditions over the no sound condition (*p*s < .01). The different sound condition showed the highest score ($M = 5.13$) on the fun scale, but the same sound condition showed the highest score on the functional satisfaction ($M = 5.0$) and the preference ($M = 4.98$) scales. There was no significant difference between the two sound conditions.
- Auditory pop-up: Participants preferred the simple notification pop-up sound to other conditions including the no sound condition (*p*s < .01). For the fun scale, the continuous melody condition was not significantly higher than the simple sound condition, but it showed numerically the highest score, which is promising for the next implementation. Improvements in terms of length, amplitude, or variations are expected to enhance functional satisfaction and preference of the continuous melody pop-up sound.

3.3 Hierarchical Menu Navigation

- Ambient menu depth cue: Participants showed significantly higher scores on the functional satisfaction and the fun scales in the unique menu depth background sound ($M = 4.46$; 4.70) condition than the hierarchically same depth background sound ($M = 4.0$; 3.76) condition (*p*s <. 01). The unique sound score ($M = 3.80$) was also numerically higher than the hierarchically same sound ($M = 3.50$). Overall, using a specific sound theme for each menu category could make more functionally helpful and fun user interfaces on smartphones.
- Menu content auditory icon: Participants showed significantly higher scores on the functional satisfaction and the fun scales in three different sound conditions than the no sound condition (*p*s < .01). For all of the measures, participants gave numerically higher scores to the serial auditory icon than the parallel auditory icon, which means that complexity matters. Even though the serial auditory icon should be intrinsically longer than the parallel auditory icon, participants favored hearing one at a time over hearing different sounds simultaneously.



(a) (b)

Fig. 3. Rating results about hierarchical menu navigation (ambient menu depth cue (a) and Menu content auditory icon (b))

3.4 List Menu Navigation

Participants showed significantly higher scores on the functional satisfaction and the fun scale in the spindex list-search condition than the no sound or the single tone conditions (*p*s < .01). However, the preference was not different. The subjective preference could be improved using alternative designs of the spindex (e.g., the attenuated version or decreased version) [6].

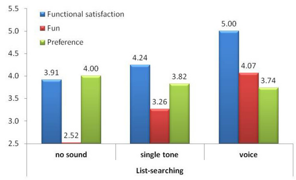


Fig. 4. Rating results about list menu navigation

4 Overall User Experience and Conclusion

For the overall user experience of the device, participants showed significantly higher functional satisfaction and fun scores in our sonically-enhanced smartphones than the no-sound condition (ps < .01). However, there was no preference difference. Based on overall results, sound designers could decide which sounds should be used as a default or as an option. Seven participants mentioned that the use of sound is going to be useful for the blind, the elderly, or children. More than half participants (26) stated that a variety of sounds are useful and functional, but simultaneously they mentioned that sounds should be turned off as an option. Thirteen participants said, "sounds are fun", but eight participants used a silent mode as a default on their phone. An appropriate application of sounds does not guarantee that users, who used to turn off their sound mode, would turn on. However, the balanced sound design might free users from auditory pollution and allow users to use smartphones more pleasantly.

References

1. Norman, D.A., Nielsen, J.: Gestural interfaces: A step backward in usability. Interactions 17(5), 46–49 (2010)
2. Brewster, S.A., Leplâtre, G., Crease, M.G.: Using non-speech sounds in mobile computing devices. In: Proceedings of the 1st Mobile HCI 1998, Glasgow, UK (1998)
3. Leplâtre, G., Brewster, S.A.: Designing non-speech sounds to support navigation in mobile phone menus. In: Proceedings of the 6th ICAD 2000, Atlanta, GA, USA (2000)
4. Palladino, D., Walker, B.N.: Efficiency of spearcon-enhanced navigation of one dimensional electronic menus. In: Proceedings of the ICAD 2008, Paris, France (2008)
5. Brewster, S.A., Cryer, P.G.: Maximising screen-space on mobile computing devices. In: Proceedings of the ACM CHI 1999, Pittsburgh, PA, USA (1999)
6. Jeon, M., Walker, N.B.: "Spindex" (Speech Index) improves acceptance and performance in auditory menu navigation for visually impaired and sighted users. ACM Transactions on Accessible Computing 3(3), 10:1–26:1 (2011)
7. Jeon, M., Walker, B.N., Srivastava, A.: "Spindex" (speech index) enhances menu navigation on touch screen devices with tapping, wheeling, and flicking gestures. ACM Transactions on Computer-Human Interaction 19(2), 14:1–27:1 (2012)

8. Oh, J.W., Park, J.H., Jo, J.H., Lee, C., Yun, M.H.: Development of a kansei analysis system on the physical user interface. In: Proceedings of the HCI, Kangwon, Korea (2007)
9. Rasmussen, J., Vicente, K.J.: Coping with human errors through system design: Implications for ecological interface design. International Journal of Man-Machine Studies 31, 517–534 (1989)
10. Gaver, W.W.: Auditory icons: Using sound in computer interfaces. Human-Computer Interaction 2, 167–177 (1986)
11. Wilson, J., Walker, B.N., Lindsay, J., Cambias, C., Dellaert, F.: SWAN: System for wearable audio navigation. In: Proceedings of the 11th ISWC 2007 (2007)
12. Blattner, M.M., Sumikawa, D.A., Greenberg, R.M.: Earcons and icons: Their structure and common design principles. Human-Computer Interaction 4, 11–44 (1989)
13. Helle, S., Leplatre, G., Marila, J., Laine, P.: Menu sonification in a mobile phone – a prototype study. In: Proceedings of the ICAD 2001, Espoo, Finland (2001)
14. Brewster, S.A.: The design of sonically-enhanced widgets. Interacting with Computers 11(2), 211–235 (1998)
15. Yalla, P., Walker, B.N.: Advanced auditory menus: Design and evaluation of auditory scroll bars. In: Proceedings of the ASSETS 2008 (2008)
16. McGee-Lennon, M., Wolters, M.K., McLachlan, R., Brewster, S., Hall, C.: Name that tune: Musicons as reminders in the home. In: Proceedings of the CHI 2011, BC, Canada (2011)
17. Walker, B. N., Lindsay, J., Nance, A., Nakano, Y., Palladino, D. K., Dingler, T., Jeon, M.: Spearcons (speech-based earcons) improve navigation performance in advanced auditory menus. Human Factors (2012) (Published online July 2, 2012 Print edition pending)
18. Walker, B.N., Kogan, A.: Spearcon performance and preference for auditory menus on a mobile phone. In: Stephanidis, C. (ed.) UAHCI 2009, Part II. LNCS, vol. 5615, pp. 445–454. Springer, Heidelberg (2009)
19. Jeon, M., Davison, B.K., Nees, M.A., Wilson, J., Walker, B.N.: Enhanced auditory menu cues improve dual task performance and are preferred with in-vehicle technologies. In: Proceedings of the AutomotiveUI 2009, Essen, Germany (2009)
20. Vicente, K.J.: Ecological interface design: Supporting operator adaptation, continuous learning, distributed collaborative work. In: Proceedings of the HCPC (1999)
21. Rasmussen, J.: The role of hierarchical knowledge representation in decision making and system management. IEEE Transactions on Systems, Man and Cybernetics 15, 234–243 (1985)
22. Sanderson, P.M., Anderson, J., Watson, M.: Extending ecological interface design to auditory displays. CSIRO (2000)
23. Walker, B.N., Kramer, G.: Ecological psychoacoustics and auditory displays: Hearing, grouping, and meaning making. In: Neuhoff, J. (ed.) Ecological psychoacoustics, pp. 150–175. Academic Press, New York (2004)
24. Jeon, M.: Two or three things you need to know about AUI design or designers. In: Proceedings of the ICAD (2010)
25. Kurakata, K., Mizunami, T., Yomogida, H.: Guidelines on the temporal patterns of auditory signals for electronics home appliances: Report of the association for electric home appliances. Acoust. Sci. & Tech. 29(2), 176–184 (2008)
26. Lee, J.-H., Jeon, M., Han, K.H.: Developing the design guideline of auditory user interface for domestic appliances. Korean Journal of the Science of Emotion & Sensibility 10(3), 307–320 (2007)
27. Kortum, P., Peres, S.C., Stallmann, K.: Mental workload measures of auditory stimuli heard during periods of waiting. In: Proceedings of the HFES (2010)