DESIGN AND EVALUATION OF LYRICONS (LYRICS + EARCONS) FOR SEMANTIC AND AESTHETIC IMPROVEMENTS OF AUDITORY CUES

Myounghoon Jeon

Mind Music Machine Lab
Michigan Technological University
Houghton, MI 49931, United States
mjeon@mtu.edu

Yuanjing Sun

Mind Music Machine Lab
Michigan Technological University
Houghton, MI 49931, United States
Ysun4@mtu.edu

ABSTRACT

Auditory researchers have developed various non-speech cues in designing auditory user interfaces. A preliminary study of "Lyricons" (lyrics + earcons) has provided a novel approach to devising auditory cues in electronic products, by combining the concurrent two layers of musical speech and earcons (short musical motives). The purpose of the present study is to introduce iterative design processes and to validate the effectiveness of lyricons compared to earcons, whether people can more intuitively grasp functions that lyricons imply than those of earcons. Results favor lyricons over earcons. Future work and practical application directions are also discussed.

1. INTRODUCTION

Users expect instant feedback when they interact with electronic devices to make sure their operations are valid. Effective auditory feedback should convey a straightforward message and be distinguishable from other auditory feedback, but should not demand too much workload. To this end, auditory researchers have developed various non-speech and speech cues in designing auditory user interfaces. Auditory icons [1] use part of representative sounds of the object or item (e.g., shutter sound for the camera function). Earcons include short musical motives [2] (e.g., "Do Re Mi" sound for the "turn on" function). While auditory icons contain semantic closeness to the item, earcons are more aesthetic. Researchers have also tried to tweak speech (e.g., Spearcons or speech earcons [3]) or use part of speech (e.g., Spindex or speech index [4]), specifically for auditory menu navigation. On the one hand, non-speech cues (e.g., earcons) could be more aesthetic, but require learning. On the other hand, speech cues could be clearer, but might not be aesthetic or be more intrusive. From this background, a preliminary study of "Lyricons" (lyrics + earcons) [5] has provided a novel approach to devising auditory cues in electronic products, by combining the two layers of *musical* speech sounds (lyrics) and non-speech sounds (earcons) concurrently. This combination is expected to improve both semantics and aesthetics of auditory user interfaces (see Figure 1). The purpose of the present paper is to: 1) briefly present the results of focus groups conducted to obtain users' opinions about their awareness of auditory user interfaces in their everyday lives and comments on the initial design of lyricons; and 2) validate the effectiveness of lyricons compared to traditional earcons, whether people can more intuitively grasp the intended functions that lyricons imply than those of earcons.

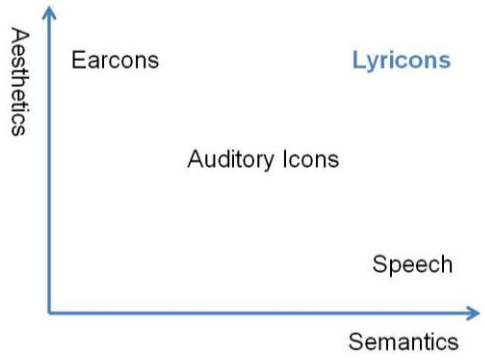

Figure 1. Relative position of auditory cue types along the two axes, which are important components in auditory user interface design.

2. INITIAL DESIGN

An experienced sound designer (>15 years) created 9 lyricons for 9 different functions (see Table 1). Earcon design follows literature and industry standards [6]. Lyrics came from previous research [5]. For more detailed design process and alternative designs, see [5].

Table1. Functions and corresponding lyricon designs

No.	Function Names and Descriptions	Music Notes	Speech in Lyricons
20	Power On (boot up system)	C4G4E4G4C5	- - Turn / It / On
16	Power Off (turn off system)	C5G4E4G4C4	- - Turn / It / Off
5	Function On (start laundry, start heat, start wind, etc.)	C4D4E4	Func / Tion / On
2	Function Off (turn off or pause heat, wind, etc.)	E4D4C4	Func / Tion / Off
11	Magnitudeup (speed up, increase values)	G4A4	- Up
12	Magnitude down (slow down, decrease values)	A4G4	- Down
1	Cancel (back, deselect)	G5G5	Back
21	Touch (press, entrance)	G5	Tink
22	Unavailable (disabled area, does not work)	G1	Bumb

3. FOCUS GROUP

Twelve undergraduate students (mean age = 23, female = 5)

participated in the focus group sessions. None of them has hearing impairments or professional music background. After a consent form procedure and introduction to the study, participants (3-5 in one session) discussed with a moderator their personal experience of the use of auditory user interfaces in electronic devices, and their advantages and disadvantages. Then, the moderator showed initial lyricon designs and participants provided comments on lyricons.

A majority of participants emphasized that auditory cues should convey a straightforward meaning, which is not necessarily the case in reality. They stated that they can easily fall into confusion when the meaning of the sound is uncertain, "Sometimes, I heard the sound but still don't know which part goes wrong, especially when I am driving. It's really annoying because neither can I stop the sound nor can I understand what the problem is." In addition to functional interpretation, some participants were likely to associate auditory cues with their memory or affect in their daily lives. For example, participant G mentioned, "I like the sound from vacuum when I just wake up. It links my memory with my mom." If the sound from products was used as a trigger of behavioral shift or attentional shift, participants allowed for an appropriate level of interference, "I like the prompt tone of SKYPE when someone is talking to me. I think it is OK for me if it's not too loud to be a noise." However, simultaneously, they want to have control over the auditory cue. Once they lose control of it, they tend to regard it as a noise. Some participants favored speech sounds, "I like natural voice to tell me what's wrong with my car," "It will be even better if the oven can talk to me. I mean I like to pretend all equipment at home is a human," which supports the application of lyricons. Participant L provided recommendations of the next lyricon designs, "To a new user, it will be better to have speech first and then, the sound, so he or she knows the specific function of the sound clearly. After a while, they can choose to skip the speech, but keep using the sound. If more instruments in different ranges were used, it would be easy to distinguish from each other."

4. EXPERIMENT IN PROGRESS

So far, thirteen undergraduate students (mean age = 23, f = 5) have attended the auditory cue-function mapping experiment. None of them has participated in the previous focus group sessions. They were randomly allocated to two groups: lyricon group or earcon group. After a consent form procedure, participants conducted a card sorting task. Nine function index cards were placed on the desk. Each card contains a definition and specific examples of the function so that all participants have the same understanding of the functions. The sound stimuli consisted of 9 lyricons and 9 earcons (exactly the same as those used in lyricons). Participants listened to sound stimuli generated from a SONY sr16 computer using a Sennheiser HD380 pro headphone. First, an experimenter explained the meaning of each function to participants. Before they start the sorting task, participants were told to listen to all of the sound recordings. Then, participants paired each sound stimulus with the function that the sound best represents. Participants were allowed to have as much time as they wanted to complete the sorting task. Before they confirm their response, they were also encouraged to finally listen to all the pairs one more time.

The average of accuracy rate of the lyricon group (70.38%) was higher than that of the earcon group (42.84%), which makes almost double (Table 2). On the other hand, the sorting time of the lyricon group (M = 4.8 mins) was shorter than that of the earcon group (M = 7 mins). We are recruiting more participants to get more ststitically reliable data. Moreover, we will analyze the confusion matrix to identify which functions make users confused about.

Table 2. Auditory cue mapping results (correct ones in green).

Lyricons	P1	P2	P3	P4	P5	P6		Earcons	P7	P8	P9	P10	P11	P12	P13	
Power-on	20	20	20	20	20	5	83.30%	Power-on	20	11	20	20	20	5	16	57.10%
Power-off	16	16	16	16	16	2	83.30%	Power-off	16	5	5	16	16	2	20	42.90%
Function-on	21	5	5	5	2	20	50.00%	Function-on	5	1	2	11	5	11	12	28.60%
Function-off	2	2	2	2	11	16	66.70%	Function-off	2	2	12	2	2	1	11	57.10%
Magnitude up	11	11	11	11	5	21	66.70%	Magnitude up	11	20	11	12	1	20	5	28.60%
Magnitude down	12	12	12	12	12	11	83.30%	Magnitude down	12	16	1	5	21	16	2	14.30%
Touch	1	1	1	21	1	1	83.30%	Touch	1	22	16	21	11	12	21	14.30%
Cancel	22	21	21	1	21	22	50.00%	Cancel	21	21	21	1	12	21	1	57.10%
Unavailable	5	22	22	22	22	12	66.70%	Unavailable	22	12	22	22	22	22	22	85.70%
Accuracy(%)	66.7	100	100	77.8	66.7	11.1	**70.38%**	Accuracy(%)	100	22.2	44.4	44.4	55.6	22.2	11.1	**42.84%**

5. CONCLUSION & FUTURE WORK

As expected, participants with lyricons showed higher accuracy rate and shorter mapping time than with earcons, which is promising in terms of lyricon applications in user interfaces. In the current study, however, we used only the piano sound for the experimental purpose. Based on the focus group results, we will iteratively redesign and enhance lyricons in terms of aesthetic quality as well. We also plan to analyze innate acoustic profiles of each functional speech [e.g., speech-to-song illusion, 7] and to reflect them on the earcon part design in order to enhance participants' perception and interpretation of the message conveyed by lyricons. In a practical application, once users get familiar with lyricons, they could use just the earcon part without the lyric part just as in the spearcon case. Based on this design and evaluation effort, researchers and practitioners could create more effective and efficient auditory interactions between a user and a system.

6. REFERENCES

[1] W. W. Gaver, "Auditory icons: Using sound in computer interfaces," *Human-Computer Interaction*, vol. 2, pp. 167–177. 1986.

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

[3] B. N. Walker, J. Lindsay, A. Nance, Y. Nakano, D. K. Palladino, T. Dingler, M. Jeon, "Spearcons (Speech-Based Earcons) Improve Navigation Performance in Advanced Auditory Menus." *Human Factors*, vol. 55, no.1, pp. 157-182, 2012.

[4] M. Jeon, B. N. Walker, "Spindex (Speech Index) Improves Auditory Menu Acceptance and Navigation Performance," *ACM Transactions on Accessible Computing*, vol.3, pp. 1-26, 2011.

[5] M. Jeon, "Lyricons (Lyrics + Earcons): Designing a new auditory cue combining speech and sounds," In C. Stephanidis (Ed.), Posters, Part I, HCII 2013, *CCIS* 373, Heidelberg: Springer. 2013. pp. 342-346.

[6] M. Jeon, "Tow or three things you need to know about AUI design or designers." *Proc. Int. Conf. on* Auditory Display, 2010.

[7] D. Deutsch, T. Henthorn, R. Lapidis, "Illusory transformation from speech to song," *J. Acoust. Soc. Am.*, vol. 129, no. 4, pp. 2245-2252, April 2011.