In-vehicle Auditory Interactions: Design and Application of Auditory Displays, Speech, Sonification, & Music

Myounghoon Jeon, Seyedeh M., FakhrHosseini, Steven Landry, & Jason Sterkenburg
Michigan Technological University
Houghton, USA
mjeon@mtu.edu

ABSTRACT
This tutorial presents an overview of the in-vehicle auditory interactions, including theoretical background, design considerations, and practical design procedure and guidelines based on the first author's 20 year experience in sound design and automotive user interface design, and the Mind Music Machine Lab's research on both topics. The tutorial will be composed of four sessions: auditory menu navigation, speech interactions, music & sonification, and sonic branding. Particularly, in each session hands-on activities will be conducted first and then, a short lecture on the topic will be followed. This tutorial aims to inspire practitioners as well as researchers to appropriately apply auditory displays and sonification to in-vehicle technologies, and thus, facilitate necessary multi-tasking processes in automotive user interfaces.

Author Keywords
Auditory displays; emotions; menu navigation; music; sonification; sonic branding; speech; warnings

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): User Interfaces – Auditory (non-speech) feedback, interaction styles (e.g., commands, menus, forms, direct manipulation), user-centered design, voice I/O; sound and music computing – methodologies and techniques.

INTRODUCTION
Multitasking is pervasive in our daily lives and so does in the driving context. In addition to the primary task (i.e., driving), drivers are frequently engaged in the secondary task (e.g., navigating the telematics system) or the tertiary task (e.g., searching music). Using alternative perceptual channels via auditory and multimodal presentation can facilitate performance with interfaces where there is visual overload [1, 2]. Further research suggested that auditory and multimodal interfaces may overcome some of the problems associated with visually taxing in-vehicle technologies [3].

Based on this background, in this tutorial, we would like to trace out the past, review the status quo, and glean some hints about the next stage of in-vehicle auditory interactions. To this end, we invite researchers, designers, engineers, and practitioners, regardless of their capability or experience in sound design or auditory display research. The distinction between traditional in-vehicle auditory display literature and this tutorial's content is twofold: First, we will not focus on speech recognition. We will slightly touch speech interactions, but will focus more on non-speech interactions. Second, we will not deal with collision warning sounds. We will include warning sounds in some of the sessions, but will explore more possibilities of the use of auditory displays beyond the traditional warning purpose.

THEORETICAL BACKGROUND FOR USE OF MULTIMODAL INTERFACES IN VEHICLES
Psychological models and theories explain how we can efficiently conduct multitasking through the distribution of our attentional resources [4] and/or different processing mechanisms in our working memory [5]. On the one hand, Multiple Resources Theory predicts that two tasks that demand *separate* modalities (e.g., one visual and one auditory tasks) will interfere with each other *less* than two tasks, both of which demand the same modality (e.g., two tasks demanding visual perception). On the other hand, Working Memory Model has a main controller (central executive) capable of attentional focus, storage, and decision making, and two subsystems: The visuo-spatial sketchpad and the phonological loop. Again, different subsystems process different information and thus, visual/spatial information and auditory information can be well time-shared. However, because both theories mainly deal with speech interactions, more research is required on non-speech interactions.

TIME SCHEDULE AND SESSIONS
The tutorial begins with introduction of the organizer and participants followed by a brief introduction to auditory displays and sonification. Then, we will move onto four sessions. Each session will consist of 1) hands-on experiences/experiments with sound and auditory display samples, 2) brief lecture on taxonomy and issues, 3) presentation on research that the author and collaborators conducted on the topic, and 4) wrapping-up with practical

Tentative Schedule	
09:00-09:10	Opening and introduction (organizer and participants)
09:10-09:20	Intro to Auditory Display & Sonification
09:20-09:50	Session 1: Activity “Sound Card Sorting”
09:50-10:10	Session 1: Tutorial “Auditory Menu Navigation”
10:10-10:40	Session 2: Activity “Dialogue Development”
10:40-11:00	Session 2: Tutorial “Speech Interactions”
11:00-11:10	Coffee Break
11:10-11:40	Session 3: Activity “Emotional Music Selection”
11:40-12:00	Session 3: Tutorial “Music & Sonification”
12:00-12:30	Session 4: Activity “Brand Image Mapping”
12:30-12:50	Session 4: Tutorial “Sonic Branding” for vehicle
12:50-13:00	Closing (*wrap-up, next steps)*

Table 1. Tentative schedule of the tutorial

guidelines. Of course, we do not have users design sounds per se, but we try to conduct a participatory design (involving users in each stage) throughout the sound design process. Thus, in this tutorial participants will also take part in hands-on activities which are used in the actual design procedure.

Session 1: Auditory Menu Navigation

Menu is a traditional HCI research topic. Users communicate with the system (mobile devices, home appliances, in-vehicle technologies, etc.) mainly via menus. However, most menu systems have been developed as a visual format. Depending on users, tasks, or contexts, auditory menus can supplement or even replace visual menus. This session includes auditory displays developed specifically for auditory menus and auditory menu navigation research for vehicle environments. Auditory researchers have developed a number of auditory displays we can apply (or have applied) for the vehicle context: auditory icons [6] (representative part of sounds of objects), earcons [7] (ear + icons, short musical motives as symbolic representations of objects), spearcons [8] (speech + icons, compressed speech) and spindex [9] (speech + index), lyricons [10] (lyrics + earcons), etc. in addition to speech. We will demonstrate these auditory displays and how to design each. The auditory menu research presentation will include research on smartphones, touch screens, and gesture interfaces with various auditory cues.

In this session, participants will experience mapping procedure between sounds and menu items/functions using the sound card sorting task and draw their own confusion matrix as an evaluation of the mapping [11].

Session 2: Speech Interactions

One of the main research interests of in-vehicle technologies includes the natural, intuitive interaction between a driver and a car [12]. An obviously natural way to communicate with an in-vehicle system is using speech, just as with a human co-driver. However, covering the entire speech recognition process is beyond the scope of the tutorial. In this section, we will particularly focus on emotional dialogue between an intelligent in-vehicle agent and a driver. Research has identified design considerations for speech-based in-vehicle systems. For example, a young adult voice made older drivers feel more confident while driving, need less time to complete the driving course, and have fewer accidents than using an old adult voice [13]. Also, a familiar voice (famous TV and radio presenters) yielded better performance (avoiding accidents, following traffic rules and lane keeping) for angry drivers than using an unfamiliar voice [14]. Nass et al. [15] showed that when the in-vehicle voice emotion matched the driver’s emotional state (e.g., energetic to happy and subdued to upset), drivers had fewer accidents and attended more to the road (actual and perceived), and even spoke more with the car. Subsequent research used emotion regulation speech cues, using “cognitive reappraisal” [16] and “attention deployment” [17].

Participants will develop their own specific dialogue structure for emotional drivers in the activity session. Then, outcomes will be discussed in terms of the psychological emotion regulation model [18].

Session 3: Music & Sonification

Even though speech is a clear communication means, it might interfere with concurrent conversation or create annoyance as a virtual backseat driver. Given that most drivers listen to music while driving, we can consider using drivers’ music or creating new sonification for different purposes. For example, we have developed a prototype fuel efficient driving interface using interactive sonification [19]. We extract all the driving performance data (speed, lane deviation, torque, steering wheel angle, pedal pressure, crash, etc.) from the simulator (NADS miniSim) in real-time. Then, these data can be mapped onto sound parameters, which can generate soundscape to facilitate, for example, eco-friendly driving behaviors consciously or subconsciously. Since music and sound are closely related to emotions, this type of interface can also be extended to emotional driving interfaces by collecting and integrating a driver’s affective state data (e.g., ECG, respiration, skin conductance response, facial expression, fEMG, eye-tracking, and EEG, etc.) with driving behavior data. If well integrated, this system can make a driver drive in a more fuel efficient way and also less aggressive (safer) way. Our recent research has shown that listening to music (either happy or sad) reduces angry drivers’ driving errors but does not increase driver workload compared to the control group (angry group with no music) [20]. However, details about musical (or sonification) parameters should be further identified.

In this session, participants will have an opportunity to select their own music (style, genre, etc.) and discuss how to effectively use it for emotional drivers and what are critical variables to be more effective.

Session 4: Sonic Branding

The preceding three sessions are related to improving performance, whereas sound can also be used for corporate/brand identity, which we call sonic brand, sonic logo, or jingle. Imagine Intel’s commercial sound. For this

type of project, the sound design process needs to integrate with other components, including marketing and product-planning, etc. Among many of the sounds in the vehicle, drivers are likely to remember the sound generated when they turn on or off the infotainment system. For example, people may be easily able to recall the Microsoft Windows' or Macintosh's opening or closing sound. Therefore, one of the strategies is to create a system opening (or booting) sound as a sonic brand and apply it to other areas, such as demos and advertisements on TVs, radios, and web sites. Even though people's vision is occupied with other tasks, the sonic brand can remind them of a unique brand image. Therefore, the sonic brand should be matched with the image of the corporate identity. We will demonstrate a variety of sonic brands and splash sounds and the procedure we take when we design a sonic brand.

Participants will practice to draw their car's brand map and find out some strategies (or some sound clips) to represent that brand image.

CONCLUSION

The session will be closed with short discussions on design considerations and processes. If time allows, we will also address issues in the potential application of auditory displays in electric and automated vehicles. We hope this tutorial can help participants understand the in-vehicle auditory interactions in terms of the components and guidelines. We also hope that we can provide reference points where the AutoUI community can consult whenever they need help or collaboration for in-vehicle auditory display design.

PREVIOUS WORKSHOPS & TUTORIALS ON THE RELATED TOPIC

The authors have (co)organized more than 12 workshops and tutorials on auditory displays and automotive user interfaces with collaborators in the international conferences (AutomotiveUI, UbiCom, PersuasiveTech, ICAD). Specifically, "research methods for assessing auditory displays" tutorial, "in-vehicle auditory interactions" w/s, "sonic information design" w/s at ICAD (International Conference on Auditory Display) are directly related to this tutorial.

MISCELLANEOUS

- Goal and topics of the workshop and how they match the conference goals and topics: To inspire practitioners as well as researchers on the application of auditory displays and sonification to in-vehicle technologies, and thus, facilitate the use of multimodal displays in automotive user interfaces.
- Outcome of the tutorial: 1) Participants understand the sound design elements and experience overall design procedure, 2) Hosting the follow/up workshop at AutoUI or other venues, and 3) forming a community (or special interest group), which is interested in in-vehicle auditory interactions.
- Expected attendance (12-20)
- Supported needed: power, projector, whiteboards (or poster) and markers.

BIOS

Myounghoon "Philart" Jeon. Dr. Jeon is an Associate Professor in the Department of Cognitive and Learning Sciences and the Department of Computer Science at Michigan Tech and director of the Mind Music Machine Lab. He also serves as a director of the Center for Human-Centered Computing at the Institute of Computing and Cybersystems (ICC) at Tech. His research areas include HCI (Human-Computer Interaction) and HRI (Human-Robot Interaction), with a focus on Auditory Displays, Affective Computing, Assistive Technologies, Automotive User Interfaces, and Aesthetic Computing. His research has yielded more than 150 publications across top peer reviewed journals and conference proceedings. His research is currently supported by NIH (National Institutes of Health), DOT (Department of Transportation), MTTI (Michigan Tech Transportation Institute) and automotive industry partners. He serves as an Associate Editor of MIT Press Journal, Presence: Teleoperators and Virtual Environments. He actively works in international conferences - chairing programs and sessions, organizing workshops, and serving as the program committee in AutomotiveUI, ICAD, HFES, CHI, MobileHCI, UbiComp, IEA, PersuasiveTech, etc. Currently, he serves as affective design technical committee of IEA (International Ergonomics Association). Dr. Jeon completed his MS and PhD degrees at Georgia Tech in Engineering Psychology and Human-Computer Interaction. He also earned his MS degree in Cognitive Sciences (Cognitive Engineering) at Yonsei University in Korea and studied Film Scoring at the same university. Before returning to academia, Dr. Jeon worked for LG Electronics and Daum Communications. He was responsible for all of the sound design for LG Electronics Corporate Design Center. He also designed all series of Hyundai Motors Telematics systems. Besides, he has led many industry projects in collaborations with General Electric, Samsung Electronics, Panasonic Automotive, Toyota, Equos Research, and Hyndai-Kia Motors Company, etc. mostly in sound design and automotive user interface domains. These works have been recognized by awards such as the IF Communication Design Award and the Korean Ergonomic Design Award and best papers in the flagship international conferences (ArtsIT, HCII, HFES). He is still playing guitar and drums in his church, and teaching her daughter how to compose music and play various instruments. He lives in UP (Upper Peninsula), Michigan with his wife and 9 year old daughter (EunYool) whose name means "fragrant melody".

Seyedeh Maryam FakhrHosseini. Maryam is a PhD student in the Applied Cognitive Science and Human Factors Program at Michigan Tech. Her research focuses on the effects of music on emotional drivers.

Steven Landry. Steven is a PhD student in the Applied Cognitive Science and Human Factors Program at Michigan Tech. His research focuses on auditory displays and sonification for the vehicle context and artistic performance.

Jason Sterkenburg. Jason is a PhD student in the Applied Cognitive Science and Human Factors Program at Michigan Tech. His research focuses on the sonically-enhanced in-vehicle gesture interactions.

ACKNOWLEDGEMENT

This tutorial is partly supported by Michigan Tech Transportation Institute (MTTI).

REFERENCES

1. Brewster, S. 1997. Using non-speech sound to overcome information overload. *Displays*, 17, 179-189.

2. Brown, M.L., Newsome, S.L., and Glinert, E.P. 1989. An experiment into the use of auditory cues to reduce visual workload'. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI 89), 339-346.

3. Nees, M.A., and Walker, B.N. 2011. Auditory displays for in-vehicle technologies. *Reviews of Human Factors and Ergonomics* 7, 1: 58-99.

4. Wickens, C.D. 2002. Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science* 3, 2: 159-177.

5. Baddeley, A. 1992. Working memory. *Science* 255, 5044: 556-559.

6. Gaver, W.W. 1986. Auditory icons: Using sound in computer interfaces. *Human-Computer Interaction* 2, 167-177.

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

8. Walker, B.N., Lindsay, J., Nance, A., Nakano, Y., Palladino, D.K., Dingler, T., and Jeon, M. 2013. Spearcons (speech-based earcons) improve navigation performance in advanced auditory menus. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 55, 1: 157-182.

9. Jeon, M., and Walker, B.N. 2011. Spindex (Speech Index) improves auditory menu acceptance and navigation performance. *ACM Transactions on Accessible Computing* 3, 3:10:11-26

10. Jeon, M. 2013. Lyricons (Lyrics+ Earcons): Designing a new auditory cue combining speech and sounds. In *Communications in Computer and Information Science (CCIS) 373. C. Stephanidis* (ed.). Springer Press, Heidelberg. pp. 342-346.

11. Jeon, M. 2015. Auditory user interface design: Practical evaluation methods and design process case studies. *The International Journal of Design in Society* 8, 2: 1-16.

12. Eyben, F., Wollmer, M., Poitschke, T., Schuller, B., Blaschke, C., Farber, B., and Nguyen-Thien, N. 2010. Emotion on the road: Necessity, acceptance, and feasibility of affective computing in the car. *Advances in Human-Computer Interaction* (ID 263593), pp. 1-17.

13. Jonsson, I.-M., Zajicek, M., Harris, H., and Nass, C. 2005. Thank you, I did not see that: In-car speech based information systems for older adults. In *Extended Abstracts on Human Factors in Computing Systems* (CHI'05), 1953-1956.

14. Jonsson, I.-M. 2009. *Social and emotional characteristics of speech-based in-vehicle information systems: Impact on attitude and behaviour*. PhD Dissertation. Linkoping University, Linkoping, Sweden.

15. Nass, C., Jonsson, I.-M., Harris, H., Reaves, B., Endo, J., Brave, S., and Takayama, L. 2005. Improving automotive safety by pairing driver emotion and car voice emotion, Improving automotive safety by pairing driver emotion and car voice emotion. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI'05), 1973-1976.

16. Harris, H., and Nass, K. 2011. Emotion regulation for frustrating driving contexts'. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI'11).

17. Jeon, M., Walker, B.N., and Gable, T.M. 2015. The effects of social interactions with in-vehicle agents on a driver's anger level, driving performance, situation awareness, and perceived workload. *Applied ergonomics*, 50, 185-199.

18. Gross, J.J. 2002. Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology*, 39, 281-291.

19. Landry, S., Tascarella, D., Jeon, M., and FakhrHosseini, S.M. 2016. Listen to your drive: Sonification arthitecture and strategies for driver state performance. In *Proceedings of the International Conference on Auditory Display* (ICAD'16).

20. Fakhrhosseini, S.M., Landry, S., Tan, Y.Y., Bhattarai, S., and Jeon, M. 2014. If you're angry, turn the music on: Music can mitigate anger effects on driving performance. In *Proceedings of the 6th International Conference on Automotive User Interfaces and Vehicular Applications* (AutomotiveUI'14), 1-7.