Recognition of Car Warnings: An Analysis of Various Alert Types

Edin Sabic
Scott Mishler
Jing Chen
Bin Hu
New Mexico State University
Las Cruces, NM 88003, USA
sabic@nmsu.edu
smishler@nmsu.edu
jingchen@nmsu.edu
bhu2@alumni.nd.edu

CHI'17 Extended Abstracts, May 06-11, 2017, Denver, CO, USA
ACM 978-1-4503-4656-6/17/05.
http://dx.doi.org/10.1145/3027063.3053149

Abstract
Warnings are integral to ensuring the safe operation of a vehicle. The use of auditory alerts and warnings has the potential to alleviate drivers' workload, increase drivers' situation awareness, and facilitate efficient and safe driving. The present study assessed the ability of individuals to react to auditory warnings. The warnings took the form of text-to-speech (TTS), spearcons at two levels of linear compression, and auditory icons. To assess usability, the NASA Task Load Index and an annoyance question were used. Participants pressed a space bar when they recognized a warning, and then identified auditory warnings by selecting a picture corresponding to the meaning of the warning. The results showed that participants responded to the fastest spearcon warnings more quickly compared to TTS and auditory icons. Responses to auditory icons were slowest compared to all other auditory types. Importantly, responses to spearcon warnings were no less accurate in comparison to TTS warnings.

Author Keywords
Warning systems; auditory icons; text-to-speech; car alerts; spearcons; sonification.

ACM Classification Keywords
H.1.2 [Information Systems]: User/Machine Systems

Introduction

Car warnings can represent various issues and levels of urgency but all aim to communicate with the driver in an unobtrusive and informative manner. Recognition of the meaning behind a warning is important, as ambiguous warnings and symbols can be neglected or even completely ignored. Further, if the warning is hard to understand it may be misidentified, leading to an undesired response. As a result, the semantic mapping of a car warning to its referent is crucial in facilitating a response behavior [9]. An ambiguous and unclear relationship between the auditory warning and the phenomena it aims to represent can lead to less accurate and slow responses. If car warnings are either confusing or inefficient, they may impair drivers' performance in response to potential danger.

Auditory alerts or feedback can take shape in many forms, including text-to-speech (TTS), auditory icons, earcons, tones, and spearcons, to name just a few. Earcons were introduced as brief melodies that do not semantically map onto the construct which they represent [3]. In contrast, auditory icons are sounds that can accurately depict the construct intended through the use of semantically-congruent relationships. Auditory icons utilize pre-existing constructs, such as the connection between the shutter sound and a camera. While auditory icons have the benefit of being language-free, and therefore, are more generalizable cross-culturally, some constructs or messages are hard to represent without using language. Lastly, spearcons are sped-up versions of TTS that have increased tempo but are no different in pitch compared to TTS [15]. Spearcons are sped up to the point where they may not be recognizable as speech, although it is not strictly necessary. Studies assessing the learnability of sound cues have shown that auditory icons and spearcons both require less time to learn than earcons [1,4,11], and that spearcons are comparable to TTS in terms of learnability [1]. Further, spearcons and TTS outperform auditory icons and earcons in terms of learnability for longer lists [1].

Previous research investigating the efficacy of various auditory car warnings has supported the integral role of semantic congruency in auditory warnings [9], while others have found that the use of auditory warnings in a vehicle context was perceived as less effort-intensive than visual warnings [10]. Research analyzing gaze movements and interaction with an in-vehicle GPS interface also supported auditory warnings in vehicles by demonstrating that sonification of an in-vehicle device increased time spent looking at the primary (driving) task as compared with visual warnings [14]. Similarly, other research has demonstrated that supplementing a vehicle's infotainment system with auditory cues compared with no sound at all increased both driving and menu navigation performance [7]. Another study comparing the implementation of spearcons, earcons, or no sound at all to an interactive vehicle interface showed that the use of spearcons decreased the time spent looking at the visual display and increased subjective driving performance [8].

Previous studies [12,16] demonstrated that spearcons could be responded to more quickly than TTS, without sacrificing response accuracy. Other research has also demonstrated that manipulating properties of auditory warnings, such as speed, can convey urgency [2,6,13]. Urgency has been measured by rating and ranking sounds based on perceived urgency [2,6], or by having participants manipulate the properties of sounds to create their own urgent warnings [13]. With these findings in mind, the implementation of spearcons into a car warning system may allow drivers and passengers to respond more quickly to an aversive event. This facilitated response is important because in driving situations, time can be critical.

Experiment
So far, no previous studies have evaluated the efficacy of spearcons compared to other auditory alert types with potential usage for car warnings. In addition, further research in the usability aspect of auditory warnings is necessary as limitations, such as the potential for annoyance [9] and interference, are apparent. Thus, the purpose of the present study was to determine the efficacy of various forms of auditory alerts (TTS, auditory icons, and spearcons), while also measuring workload and annoyance associated with these forms of auditory alerts. Further, the present study assessed the learnability of the auditory alert types across the four experimental blocks in terms of both RT and accuracy. These auditory types were included due to their relevance in the field of auditory displays and alerts. The independent variable was the auditory alert type. Dependent variables consisted of reaction time (RT), accuracy, the NASA Task Load Index (NASA-TLX) [5] and an annoyance measure.

Methods
A between-subjects design was used consisting of one between-subjects factor, auditory alert type, with four levels: (a) spearcons at 40% of original audio length (40% spearcons), (b) spearcons at 60% of original audio length (60% spearcons), (c) TTS, and (d) auditory icons. A between-subjects design was chosen due to the semantic similarity across alert types. There were 22 participants in each of the groups.

Participants
Eighty-eight students (52 female; mean age = 19.38; *SD* = 3.12) from New Mexico State University participated in the experiment for course credit. All of them reported normal or corrected-to-normal hearing.

Apparatus
Participants completed the experiment on a computer, a Dell OptiPlex 7020, running E-Prime software. Participants wore a pair of headphones, the Audio-Technica ATH-M30x Professional Studio Monitor Headphones, to hear all stimuli during the experiment. Computer volume was kept constant at 30% of full volume. Practice served as a check to ensure participants could hear the stimuli at this volume, and all participants reported as having normal hearing.

Stimuli
TTS was created using NaturalReader 14.0 software, which reads entered text using a synthesized voice. Sound was captured using a playback option on Audacity, a free software used for audio editing. The WASAPI playback option allows for the recording of computer playback without the usage of a microphone. Spearcons were produced by taking each corresponding TTS alert or warning and increasing the tempo within Audacity. Tempo was increased through linear compression to create 40% and 60% spearcons. That is, if the original sound duration was 1 second, the 40% spearcon version was 400 ms. All auditory icons were gathered from the British Broadcasting Corporation (BBC) Sound Effects Library, from CDs: 1, 5, 12, 13, and 19. These auditory icons were used in previous

TTS	Auditory Icons
Headway closing fast	Collision sound
Drifting off road	Rumble strips
Speed limit exceeded	Car speeding past
Car in blind spot	Honk
Tire pressure low	Air release
Door is open	Door shutting
Hand break on	Creaking sound
Gas is low	Water pouring
Oil is low	Steam and water

Table 1: TTS phrases and descriptions of auditory icons.

research [9] to assess signal-referent relationships. All sound effects were edited through Audacity to ensure that sounds did not include trailing noise, and were only as long as needed to convey meaning. All warning phrases (see Table 1) and illustrations (see Figure 1) were borrowed from a previous study that compared TTS to other auditory feedback types [9].

Subjective measures
Participants responded to seven questions at the end of the experiment. NASA-TLX was used to assess mental workload [5]. This survey measures workload on six scales, which include: mental demand, physical demand, temporal demand, performance, effort, and frustration. Further, an annoyance question consisting of the prompt, "How annoying did you find these warnings?" was used to assess perceived annoyance. All questions utilized a 21-point scale.

Procedure
Participants first signed the consent form and provided demographics information. At the beginning of the experiment, participants were familiarized with the auditory type that would be used in the study. They were then presented with a slideshow, which paired the auditory warning with a picture that would be used to assess accuracy (see Figure 1). The practice session then began, which could last between 9 and 45 trials depending on the participant's performance. The practice session was completed when participants performed at above 85% accuracy. The practice trials mirrored the experimental blocks completely. Each trial began with the presentation of a sound, and participants were required to press the space bar as soon as they recognized the warning. RT was logged from stimulus onset until the space bar was pressed. After this bar press, they were taken to a screen with nine pictures (see Figure 1) corresponding to all auditory warnings, and asked to select which warning that they just heard. After practice was completed, participants performed 432 experimental trials split up into four blocks. At the end of each block, participants answered the NASA-TLX and annoyance questions based on their experience with the sound warnings. Participants were given the choice to take a break lasting up to five minutes between blocks.

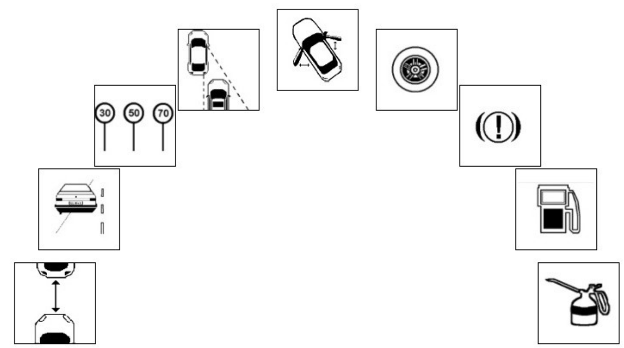

Figure 1: Illustrations for car warnings as seen by participants. Each illustration corresponded to a specific warning.

Results
A one-way analysis of variance (ANOVA) was conducted with auditory alert type as a between-subjects factor on RT. The main effect of alert type on RT was significant, $F(3, 84) = 42.11$, $p < .001$, $\eta_p^2 = .60$. Mean RTs tended to decrease across TTS, 60% spearcons, and 40% spearcons (see Figure 2). Pairwise comparisons (Sidak) showed that auditory icons were responded to more slowly compared to all other auditory types, *p*s < .001. 40% spearcons were responded to more quickly than

TTS, p = .002. Responses to 60% spearcons were not significantly different from 40% spearcons, p = .357, or from TTS, p = .319.

To determine if there were any practice effects for RT, a mixed-design ANOVA was conducted with block number being a within-subjects factor and auditory type a between-subjects factor. There was a significant effect of block number, $F(3, 252) = 98.40$, $p < .001$, $\eta_p^2 = .54$, and an interaction between block number and condition, $F(9, 252) = 2.21$, $p = .022$, $\eta_p^2 = .07$. Pairwise comparisons (Sidak) showed that there were significantly faster RTs moving from block 1 to 2, and from block 2 to 3, ps < .001. However, RTs were not significantly different in blocks 3 and 4, p = .369

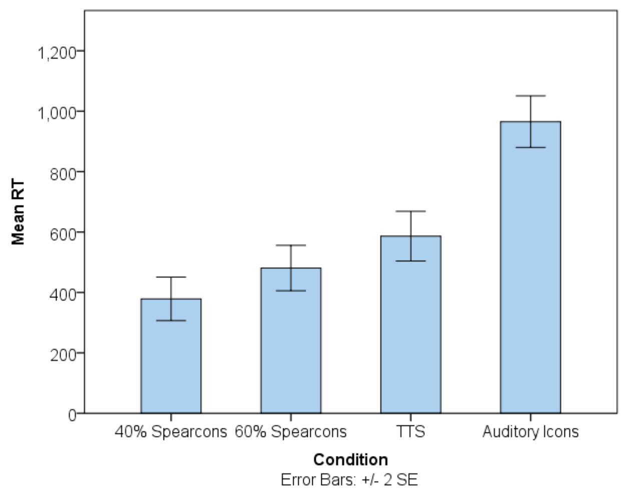

Figure 2: Reaction time (RT) across auditory type.

A one-way ANOVA was conducted with alert type as the between-subjects factor on accuracy. The overall analysis of accuracy was significant, $F(3, 84) = 3.81$, $p = .013$, $\eta_p^2 = .12$. Pairwise comparisons (Sidak) showed that the only significant difference in accuracy was between TTS and Auditory icons, p = .015, with auditory icons exhibiting a lower accuracy (M = 95.57%, SD = .03) than TTS (M = 97.81%, SD = .01). Across all auditory alert types, there were no significant learning effects for accuracy.

Subjective measures
To determine whether there was a difference of temporal demand across auditory types, a one-way ANOVA was conducted with auditory alert type as the between-subjects factor. Only the temporal demand data was used to determine whether the increased tempo of spearcons would create a sense of urgency. Responses to other questions from the NASA-TLX were not analyzed. The main effect of alert type was significant, $F(3, 84) = 3.71$, $p = .015$, $\eta_p^2 = .12$. However, pairwise comparisons (Sidak) showed that 40% spearcons were not perceived as being significantly more temporally demanding than 60% spearcons, TTS, or auditory icons, ps > .05. To determine whether there was a difference of perceived annoyance across auditory alert types, a one-way ANOVA was conducted with auditory alert type as the between-subjects factor. The main effect of alert type was not significant, $F(3, 84) = 2.54$, $p = .062$, $\eta_p^2 = .08$. Planned comparisons (Sidak) showed that 40% spearcons were not perceived as more annoying compared to the other alert types, ps > .05.

Discussion
The present study compared the efficacy and usability of four auditory alert types in a car warning recognition task. After only a brief training session, accuracy was not significantly lower for spearcons compared to TTS. Auditory icons had the lowest accuracy among all alert

types. Considering that the auditory icons contained lower semantic congruency than TTS and spearcons, this result is consistent with previous research investigating signal-to-referent relationships [9].

Although there were few differences in accuracy when comparing the auditory alert types, there were significant differences in terms of RTs. Analyses showed that auditory icons were responded to more slowly when compared to 40% spearcons, 60% spearcons, and TTS. This may be a result of longer durations of the warning, or a weaker signal-to-referent relationship for auditory icons [9]. Importantly, 40% spearcon warnings were responded to more quickly than TTS warnings. While responses to 60% spearcon warnings were not significantly different from TTS, RTs decreased numerically going from TTS, to 60% spearcons, and to 40% spearcons (see Figure 2), suggesting that further research should be conducted to analyze whether this level of linear compression (60% of original audio duration) can be consistently responded to more quickly when compared to TTS. Analyses assessing RTs across all alert types provided evidence for learning effects, indicating that training may be required to reach optimal RTs for all alert types.

The finding that 40% spearcons were responded to most quickly without any significant accuracy loss has important implications for auditory car warnings. Car warnings often need to communicate highly critical and time sensitive information. Facilitating a faster response by providing a more efficient warning may lead to safer driving and fewer accidents on the road. The finding that these spearcons were also not perceived as more annoying than the other alert types is important in assessing usability. Further research is needed to establish which types of auditory alerts should be designated to which warnings, especially when considering severity. Considering that annoyance and fatigue can set in with repetition of one type of auditory alert, designating low-risk and highly repetitive events with auditory icons may increase satisfaction. However, for high-risk situations that will demand quick and accurate responses, the present findings support that spearcons may be an appropriate design choice. Further research is also needed to establish whether the present findings can be replicated outside of a quiet environment while using a dual-task paradigm to assess the true validity of these auditory alert types.

Conclusion
The 40% spearcon warnings were responded to more quickly when compared to both TTS and auditory icons, and were not perceived as being more annoying compared to 60% spearcons, TTS, or auditory icons. Further, there was no significant difference in accuracy when comparing spearcons to TTS. The present research provides evidence that even brief spearcon alerts may be useful for providing information in time-critical situations which could potentially increase safety and driving performance. However, these findings need to be replicated in a driving task to ensure that spearcons can provide viable communication during the operation of a vehicle.

Acknowledgements
The authors thank Jaymison Miller, Steven Archuleta, Gabrielle Campbell, and Graham Strom for their help in data collection. Funding for this research was provided by the National Science Foundation under Grant 1314688.

References

1. Tilman Dingler, Jeffrey Lindsay, and Bruce N. Walker. 2008. Learnability of Sound Cues for Environmental Features: Auditory Icons, Earcons, Spearcons, and Speech. *14th International Conference on Auditory Display*, 1-6.
2. J. Edworthy and N. Stanton. 1995. A user-centered approach to the design and evaluation of auditory warning signals: 1. Methodology. *Ergonomics*, *38*, 11: 2262-2280.
3. William W. Gaver. 1986. Auditory icons: Using sound in computer interfaces. *Human-computer interaction 2*, 2: 167-177.
4. Stavros Garzonis, Simon Jones, Tim Jay, and Eamonn O'Neill. 2009. Auditory Icon and Earcon Mobile Service Notifications: Intuitiveness, Learnability, Memorability and Preference. *CHI 2009-Multimodal Mobile Interaction*: 1513-1522.
5. Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology* 52, C: 139-183.
6. E. J. Hellier and J. Edworthy. 1989. Quantifying the perceived urgency of auditory warnings. *Canadian Acoustics* 17, 3-11.
7. Myounghoon Jeon, Thomas M. Gable, Benjamin K. Davison, Michael A. Nees, Jeff Wilson, and Bruce Walker. 2015. Menu Navigation With In-Vehicle Technologies: Auditory Menu Cues Improve Dual Task Performance, Preference, and Workload. *International Journal of Human-Computer Interaction* 31, 1: 1-16.
8. Pontus Larsson and Mathias Niemand. 2015. Using Sound to Reduce Visual Distraction from In-Vehicle Human–Machine Interfaces. *Traffic Injury Prevention* 16, sup1: S25-S30.
9. Denis McKeown and Sarah Isherwood. 2007. Mapping candidate within-vehicle auditory displays to their referents. *Human Factors* 49, 3: 417-428.
10. Michael A. Nees, B. Helbein, and A. Porter. 2016. Speech auditory alerts promote memory for alerted events in a video-simulated self-driving car ride. *Human Factors* 58, 3: 416-426.
11. Dianne K. Palladino and Bruce N. Walker. 2007. Learning rates for auditory menus enhanced with spearcons versus earcons. *13th International Conference on Auditory Display*, 274-279.
12. Edin Sabic and Jing Chen. 2016. Threshold of Spearcon Recognition for Auditory Menus. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 60, 1539-1543.
13. Jeremiah Singer, Neil Lerner, Carryl Baldwin, and Eric Traube. 2015. Auditory Alerts in Vehicles: Effects of Alert Characteristics and Ambient Noise Conditions on Perceived Meaning and Detectability. *24th International Technical Conference on the Enhanced Safety of Vehicles (ESV)*(No. 15-0455).
14. Julien Tardieu, Nicolas Misdariis, Sabine Langlois, Pascal Gaillard, and Céline Lemercier. 2015. Sonification of in-vehicle interface reduces gaze movements under dual-task condition. *Applied Ergonomics*, 50, 41-49.
15. Bruce N. Walker, Amanda Nance, and Jeff Lindsay. 2006. Spearcons: Speech-based earcons improve navigation performance in auditory menus. *12th International Conference on Auditory Display,* 63-68.
16. Bruce N. Walker, Jeff Lindsay, Amanda Nance, Yoko Nakano, Dianne K. Palladino, Tilman Dingler, and Myounghoon Jeon. 2013. Spearcons (speech-based earcons) improve navigation performance in advanced auditory menus. *Human Factors*, 55, 157-182.