AUDITORY SIGNS TO SUPPORT TRAFFIC AWARENESS

Johan Fagerlönn¹, Håkan Alm²
1. Interactive Institute - Sonic Studio, SE-94128, Piteå, Sweden, +46(0)70-3689810, johanf@tii.se.
2. Department of Human Science, Luleå university of technology, SE-94187, Luleå, Sweden, +46(0)920 491270, hakan.alm@ltu.se.

ABSTRACT
Informative systems might contribute to sensory and cognitive driver distraction, which in turn can lead to a more dangerous driving behavior. In this study we evaluated auditory signs to support drivers traffic awareness during simulated driving. 18 truck drivers identified traffic situations based on information conveyed by brief sounds. Aspects of learning, interpretation and pleasantness of sounds were monitored and rated by the drivers. Sounds which were arbitrary mapped to traffic situations required longer learning times, resulted in degraded choice reaction performance, and were rated as less pleasant compared to sounds with a high level in context specific meaning.

KEYWORDS
Auditory display, warning signals, distraction, cognitive load, acceptance.

INTRODUCTION
Technical development of Intelligent Transport Systems (ITS) is often associated with expectations of their potential to increase traffic safety [1]. But systems designed to inform drivers (IVIS) can potentially contribute to both sensory and cognitive driver distraction. This in turn may lead to a more dangerous driving behavior [2,3,4] and increase the risk of traffic accidents [5]. Driver distraction might be especially problematic in urgent, unusual and otherwise challenging situations that already put high demands on the drivers’ resources. In such situations, the effectiveness of a system relies not only on the interplay between the driver and the system, but also on the drivers’ capability to take in and process signals from the system while simultaneously performing the driving task.

Visually based solutions might not be appropriate in these situations. Previous studies have shown how increased visual load during driving can affect detection performance [6] and lane keeping [3]. Systems that allow the driver to keep visual focus on the road, such as auditory solutions, can be better from a safety point of view. On the other hand, research has demonstrated that involvement in auditory tasks can affect driving performance. For instance, Engström et al [3] showed that an Auditory Continuous Memory Task (ACMT) resulted in increased gaze concentration towards the road centre. Alm et al [2] found that a mobile telephone task increased
choice reaction time - that was not compensated enough in headway. Strayer et al [4] showed how communicating on a hands-free mobile phone affected drivers ability to remember roadside objects. Thus, finding sounds that are easy to perceive and process while driving might be especially important when designing for safe IVIS communication.

In a collaboration project between Scania CV AB, Interactive Institute and Luleå university of technology research addresses how audio design can meet the requirements of safe IVIS communication within heavy vehicles. The work presented in this article focus on the broad range of potential systems aimed at supporting drivers’ traffic awareness.

**Auditory cues to support traffic awareness**

Many Intelligent Transport Systems monitors the surroundings automatically. This means that more information than ever before can now be delivered to drivers to enhance their awareness of the traffic environment. Traffic awareness can be seen as one component of the more general construct of situation awareness. While no established definition of traffic awareness seems to exist, authors have defined the concept of situation awareness (SA). According to Endsley [7] SA is “the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future”. Information that help drivers maintain their awareness about the traffic situation may be regarding other road users and various types of dangers, their position in relation to the own vehicle, or how the current situation is evolving over time.

One strategy to support traffic awareness can be to direct drivers’ visual attention to road dangers [8,9,10]. Fung et al [8] demonstrated that an effective auditory cue for a Forward Collision Warning System was a simple tone of 2 kHz. The sound quickly made drivers pay attention to the road ahead and break fast. In a similar way, attention grabbing visual or tactile cues can be used to make the driver shift visual attention towards crucial information. In a recent study Ho et al [10] found that tactile cues can even be more effective that auditory cues for this purpose.

But even though attention-grabbing auditory signals are effective for some IVIS, there are reasons why to further examine the use of more informative auditory signals. First, many types of systems such as night vision, blind spot detection, and various types of road user protection systems, are introduced in heavy vehicles to assist drivers in conditions and areas with low or no visibility. Other systems are implemented to make drivers aware of accident-prone areas, such as intersections, bus stops, and school areas. Information through sound can be an attractive means of communication in absence of fundamental visual information. Further, all systems that rely on visual information processing can potentially impact the drivers ability to process other important information from the traffic scene. Informative auditory signals can reduce the risk of visual overload in visually demanding traffic situations.

**Auditory signs**

Previous auditory display research has shown how verbal and non-verbal auditory signs can be used to convey information in various types of user environments [11,12,13,14]. Research on
non-verbal signs has largely been focusing on advantages and disadvantages of sound types referred to as Earcons and Auditory Icons. The concept of Earcons was first introduced by Blattner [15]. He defined them as “non-verbal audio messages used in the user-computer interface to provide information to the user about some computer object, operation, or interaction”. Blattner suggested that Earcons, like icons, could be divided into the classes representational, abstract, and semi-abstract. Gaver [16] investigated representational earcons, although he called them auditory icons. Gaver defined auditory icons as “everyday sounds mapped to computer events by analogy with everyday sound producing events”. Since first introduced, the definition of Earcons seems to have changed. Brewster et al [17] defined Earcons as “abstract, synthetic tones that can be used in structured combinations to represent parts of an interface”. In contrast to Earcons, Auditory Icons are often associated with iconic representations of objects and events.

At the International Community for Auditory Display (ICAD) conference 2008 Mustonen [18] stated that the current definitions of non-speech sounds are not compatible with the sign descriptions of e.g. semiotic science. “Most signs we encounter are neither purely abstract nor iconic but combine both iconic and symbolic dimensions to make sense”. Further, he pointed out that the same sound could be listened to with different outcomes in different situations and orientations. When listening to interface elements we intuitively recognize familiar parts from the sound and construct the meaning from their relation to the situation. In line with this re-established approach to the design of non-verbal sounds, we do not divide the sounds into defined types. Instead, focus lies on the meaning that the sounds can have in a context and situation where it is perceived.

Auditory signs and IVIS

A body of research has addressed the use of auditory signals to convey traffic related information in vehicles. A number of experiments have been focusing on warning signals for forward collision warning systems [7,19,20]. A few studies have been investigating sounds for traffic related information that may require other types of driver actions. Baldwin [21] examined in two simulator studies the effectiveness of verbal auditory signs with various types of collision warning systems. She found that the verbal warnings reduced crash rate, especially for older drivers. Chen et al [22] evaluated the use of a 3D sound reproduction technique and various sound types to improve traffic awareness in a number of traffic situations. The study focused primarily on system acceptance and conclusions were based on subjective driver responses. Chen et al recommended the use of auditory icons due to their intuitiveness. Mc Keown [11] evaluated four sound types (auditory icons, environmental sounds, earcons and speech) to convey various types of in-vehicle information, including some traffic related information. Response time, accuracy of response, perceived urgency and scores of pleasantness were evaluated. It should be pointed out that the study was not conducted in a driving context. Mc Keown favored the use of auditory icons and speech to convey information in vehicles. In other experiments, verbal signs and non-verbal signs referred to as auditory icons have proven more effective than other sound types both in terms of learning [23,24], response time [25, 26] and accuracy of response [26].

From a semiotic perspective this is not too surprising. Mustonen [18] stated “the important difference of the earcon paradigm is that the design in auditory icon paradigm has been more
focused on how the sound itself, through similarities and metaphors motivates the meaning creation process”. The importance of meaning within the driving context was demonstrated in the study by McKeown [11]. The environmental sounds consisted of real-world sounds that were likely to familiar, but did not have specific meanings within the vehicle interface. The sounds were mapped to scenarios on the basis of their perceived urgency, but they did not specifically represent any of the driving scenarios. These sounds resulted in both longer response times and more errors compared to verbal messages and the real world sounds defined as Auditory Icons.

Objectives

On the basis of previous research it is reasonable to believe that sounds that have a particular meaning within the driving context have a less distracting effect compared to sounds arbitrary mapped to traffic situations. However, we still know relatively little about such effects when designing for traffic awareness. The aim of this study was to evaluate differences in learnability and interpretation of auditory signs assumed to have a high level of context specific meaning, and auditory signs that are arbitrary matched to traffic situations. A further objective was to investigate perceived pleasantness of the auditory signals. The experiment did not focus on a particular traffic event or system but covered a range of traffic events handled by current and developing IVIS.

METHOD

Subjects

18 truck drivers (17 males and 1 female) participated in the study. Ages ranged between 22 and 61 (mean 39). Their truck driving experience ranged between 2 and 32 years (mean 16.8). Self-reported annual driving ranged between 2000 and 150000 km (mean 75410). All drivers had self-reported normal hearing.

Apparatus

The experiment was conducted in a Scania R truck cab. A 10.4” touch monitor (Lilliput electronics, CA, USA) showing videos of four hazardous traffic situations simultaneously during the driving sessions was positioned approximately on one arm length 30° to the right of the driver. The video clips were brief .gif animations (1.26 s in length, continuously repeated), showing traffic situations from above. A Lane Change Test represented the driving task. The visual driving scene was projected with an Optoma EP 755 XGA DLP projector (Optoma Technology Inc, CA, USA) 3.34 meters in front of the cab. Presentation of auditory stimuli, traffic situations and monitoring of driver responses were handled by a stand-alone Java application running on a Lenovo Thinkpad T60 (Lenovo, NC, USA). Sound files were processed using a Kontakt 3 sampler (Native Instruments, Berlin, Germany) and an E-MU 1616m digital sound card (E-MU Systems Inc, CA, USA). The sounds were played to participants at a comfortable listening level through an Anthony Gallo Nucleus Micro 5.1 system (Anthony Gallo Acoustics Inc, CA, USA).
Dependent and independent variables

Two sets of non-verbal auditory signs and one set of verbal signs were designed prior to the experiment. Each set contained five sounds mapped to road users (car, truck, pedestrian, children and bicycle). The first set of non-verbal auditory signs (arbitrary) consisted of five short musical motives. Brewster et al [17] has suggested how musical timbre and rhythm can make sounds distinguishable from each other. Different rhythms and timbres were selected to make the sounds easy to tell apart. The second set of non-speech auditory signs (meaningful) consisted of sounds assumed to have a specific meaning when driving. Road users make noise and they have their natural ways of catching the attention and of other road users. A number of signs were designed on the basis of these naturally occurring sounds. A panel of three people not involved in the study judged their comprehension. On the basis of panel feedback the sounds were changed again into the final set of signs used in the study. The final non-verbal signs are presented in table 1. A soft-spoken male voice was used for the verbal signs (verbal). Each speech message consisted of a keyword presented in Swedish language describing the road danger. The different auditory signs were not exactly equal in length but ranged between one and two seconds. Spatial positions of the road users were represented by presenting the sounds in the corresponding direction (front, rear, front-left, front-right, rear-left, rear-right). Combinations of road users and spatial positions resulted in a total number of 30 different traffic scenarios used in the study.

<table>
<thead>
<tr>
<th>Meaningful</th>
<th>Arbitrary (musical timbre / tones)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bicycle</td>
<td>Bicycle bell</td>
</tr>
<tr>
<td>Car</td>
<td>Car horn</td>
</tr>
<tr>
<td>Truck</td>
<td>Truck horn</td>
</tr>
<tr>
<td>Pedestrian</td>
<td>Man shout</td>
</tr>
<tr>
<td>Children</td>
<td>Children laugh</td>
</tr>
</tbody>
</table>

Table 1 - Description of the non-verbal auditory signs.

Learning time and trials, response time in judging traffic situations, accuracy of response (identity and position of road dangers), and subjective ratings of pleasantness and interpretation defined the dependent variables. A loosely structured interview was conducted in the end of each trial in order to get complementary driver judgments. In this interview the drivers were allowed to talk freely about any issues experienced during the sessions.

Procedure

The experiment was conducted using a within-subjects design in which all 18 drivers listened to all sounds. The participants were introduced to the simulator and the Lane Change Test in a 5-10 min test drive. The experimenter also provided a short demonstration session explaining the auditory sign judgment task. Subjects were told they were evaluated on the basis of driving performance. They were required to judge the auditory signals as fast as possible so that it would not affect the driving more than absolutely necessary. However, no parameters related to driving performance were actually monitored during the trials.
The experimental session consisted of three blocks (*verbal, arbitrary* and *meaningful*) that lasted for about 30 minutes each. Each block started with a learning session (no driving task) in which the participants were required to learn the intended mappings between sounds and road dangers. Images of the road dangers were presented on the touch screen and the driver were required to play the sounds until he/she felt comfortable with their meaning. Both learning time and trials were monitored.

The learning session was followed by the driving task. The auditory signs were presented in random order to the drivers with 20-60 seconds intervals. When a sound was played the drivers were required to select one out of four traffic situations presented on the touch screen. After each judgment the driver received feedback (text and color code) regarding the response. A green light indicated a correct response. A red light with a text presenting the correct answer indicated an incorrect response. This feedback was built-in to allow drivers to learn from mistakes during the sessions. After each driving session the drivers rated interpretation difficulties and pleasantness using on-screen rating scales ranging from 1 (not at all difficult/annoying) to 10 (very difficult/annoying).

**RESULTS**

The statistical analysis of the test data was performed using the computer package Minitab (Minitab Data Analysis Software, Philadelphia, PA, USA).

**Learning**

Average learning times and number of trials for the sign sets is presented in table 2. As predicted the drivers had serious problems learning the intended mappings between the five musical motives and the traffic events. In mean the drivers listened to each sound 5.9 times compared to 1.93 times for the set of meaningful non-verbal signs and 1.52 times for the verbal signs. A one way ANOVA revealed significant differences in learning time $F(2,34)=21.03$, $p<0.01$ and trials $F(2,34)=28.22$, $p<0.01$. Post-hoc analyses were conducted using Tukey’s HSD test. Both learning time and trials were significantly higher in the non-verbal random condition than the two other conditions ($p=0.01$).

<table>
<thead>
<tr>
<th></th>
<th>Time in seconds (SD)</th>
<th>Trials (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>meaningful</td>
<td>24,7 (8,7)</td>
<td>9,6 (2,9)</td>
</tr>
<tr>
<td>arbitrary</td>
<td>90,2 (59,6)</td>
<td>29,5 (16,4)</td>
</tr>
<tr>
<td>verbal</td>
<td>23,3 (12,4)</td>
<td>7,6 (3,3)</td>
</tr>
</tbody>
</table>

**Table 2 - Learning time and trials.**
Interpretation

Table 3 shows response times and average number of judgment errors for the different sign sets, and table 4 presents mean response times, accuracy of responses in terms of identity of road dangers and position of road dangers. As predicted, the arbitrary sounds resulted in longer response times and reduced response accuracy compared to the other two sound sets. A one way ANOVA revealed significant differences in response time F(2,34)=27.56, p<0.01 and accuracy F(2,34)=56.19, p<0.01. Tukey’s HSD showed a significant difference between the non-verbal random condition and the two other conditions (p=0.01). Significant effects was also found both in terms of identification of road danger F(2,34)=69.99, p<0.01 and position of road dangers (F(2,34)=16.77, p<0.05. Tukey’s HSD test indicated that the arbitrary condition resulted in more errors both in terms of identity and position of road dangers (p=0.01).

<table>
<thead>
<tr>
<th>Time in seconds (SD)</th>
<th>Errors (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>meaningful</td>
<td>3.00 (0.67)</td>
</tr>
<tr>
<td>arbitrary</td>
<td>4.12 (1.30)</td>
</tr>
<tr>
<td>verbal</td>
<td>3.11 (0.62)</td>
</tr>
</tbody>
</table>

Table 3 - Mean response time and accuracy of response.

<table>
<thead>
<tr>
<th>Identity (SD)</th>
<th>Position (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>meaningful</td>
<td>0.72 (1.67)</td>
</tr>
<tr>
<td>arbitrary</td>
<td>8.00 (3.66)</td>
</tr>
<tr>
<td>verbal</td>
<td>0.78 (1.73)</td>
</tr>
</tbody>
</table>

Table 4 - Accuracy of response in terms of identity and position of road users.

Subjective ratings

The subjective ratings of pleasantness and interpretation are presented in table 5. The arbitrary sounds were rated more annoying and more challenging to interpret while simultaneously performing the driving task compared to the other two conditions. A significant effect was found for interpretation F(2,34)=15.32, p<0.01 and annoyance F(2,34)=7.47. The post-hoc analysis reported significant effects between the arbitrary non-verbal sounds and the other two conditions (p=0.01).

<table>
<thead>
<tr>
<th>Pleasaness (SD)</th>
<th>Interpretation (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>meaningful</td>
<td>3.45 (2.09)</td>
</tr>
<tr>
<td>arbitrary</td>
<td>5.39 (2.83)</td>
</tr>
<tr>
<td>verbal</td>
<td>3.56 (1.76)</td>
</tr>
</tbody>
</table>

Table 5 - Subjective ratings of pleasantness and interpretation.
DISCUSSION

The aim of this study was to evaluate learnability, interpretation and pleasantness of auditory signs used to support drivers’ traffic awareness. Sounds assumed to have a high level of context specific meaning were compared to arbitrary assigned sounds and verbal signs.

In the interview carried out in the end of each trial, a majority of the drivers reported that they tried to make the arbitrary sounds meaningful by finding intuitive similarities and forming associations between the signs and the road users. For instance, one subject stated, “the musical motive representing the big truck was the easiest one to recognize since the cello is playing in a low register”. Low register can signify something large moving. Another driver said that he tried hard to establish an association between the sound of a trumpet and a child playing the trumpet. However, the drivers also reported that they failed to establish durable and useful associations between the sounds and the traffic events. Even though they spent considerable more time and trials trying to learn the meaning of the arbitrary sounds, their judgment performance was degraded during this condition compared to the other two conditions. Significant effects were found both in terms of response time and accuracy of response. Also, the drivers rated the sounds as more challenging to interpret while simultaneously performing the driving task. It should be pointed out that none of the 18 drivers performed faster or more accurate during this condition compared to the other two conditions. Taken together, the results support a significant effect on interpretation during the simulated driving task, which was not compensated by the longer learning time. It seems like more indirect sounds has the potential to increase cognitive load considerably more than sounds with a natural meaning within the driving context. As indicated in previous research on auditory tasks and cognitive distraction, this may have negative consequences in demanding situations.

During the trials the drivers were required to respond to the sounds as fast as possible so that it would not affect driving performance (Lane Change Test) more than absolutely necessary. The drivers were told they were judged primarily on the basis of their performance on the LCT track. In mean the driver judged the arbitrary sounds in 4.12 seconds. But sometimes the arbitrary sounds resulted in exceptionally long response times (10-15 seconds). This happened for 3 subjects (17%) and represented 1.5 percent of the measurements. In these situations the driver was not able to find a solution within a reasonable time frame. Despite the time pressure, the drivers remained motionless staring at the touch screen for long periods, losing focus on the road completely. When a perceived sound did not seem to match any of the situations presented on the screen this may have contributed to an increased cognitive load, which in turn resulted in a higher level of perceptual narrowing. In the study, this behavior was very inappropriate since the drivers were not able to focus visually on the screen and on the road scene simultaneously. In a real driving situation, perceptual narrowing might have both positive and negative consequences depending on the driving situation. One potential negative consequence can be a decreased level of situation awareness in eventful driving situations.

The results indicated that verbal signs might be as effective as the meaningful non-verbal auditory signs when driving related information is to be conveyed. These results partially replicates the results of a previous study conducted by McKeown [11]. However, it should be pointed out that the use of verbal messages might have drawbacks in that they are language dependent and may interfere with other verbal communication in the vehicle-cab. Also, there
have been some indications of how advantage of speech messages may be reduced when mental workload is high [26].

An issue for auditory display designers is how to design sounds that are not annoying. Previous experiments have been focusing on perceived annoyance of individual sounds and the acoustic properties that affect pleasantness [11, 27]. However, perceived annoyance of an individual auditory signal may be a desirable quality in very urgent situations. In other, more everyday situations, it might be crucial to find signals that are pleasant. In the present study, the participants rated perceived annoyance in the end of each experimental condition. In mean, the driver rated the arbitrary assigned sounds more annoying then the more meaningful sounds. The results give us some indication of how arbitrary sounds used by IVIS may affect the general satisfaction of the interface.

One potential criticism of the study is that the laboratory set-up is not really comparable to realistic driving. The Lane Chang Test track requires the driver to focus on the forward road scene. However, it does not fully represent the dynamic task of driving. Also, in the study the drivers listened to 90 auditory signs is about one hour of driving. This high frequency of warnings can have promoted a higher attentiveness than is typical in routine driving.

CONCLUSIONS AND FUTURE WORK

The present study investigated brief auditory signs as a mean to support drivers’ traffic awareness in different types of traffic situations. Aspects of learning, interpretation and pleasantness were monitored and rated by drivers. Sounds that were arbitrary mapped to traffic events resulted in significantly longer learning times and more trials compared to sounds with a higher level of context specific meaning. Still, the arbitrary sounds resulted in degraded performance when identifying traffic situations, both in terms of response time and accuracy of response. The pleasantness scores indicated how the level of meaning in individual sounds might impact the general acceptance of an auditory display. The results may have implications for a broad range of systems implemented to support the driver in absence of fundamental visual information or in visually demanding situations.

The concept of using an auditory display as a primary channel for traffic related information need to be further evaluated. Future experiments should examine driver behavior and effects on driving which can be more directly related to performance and safety in different types of traffic situations. We know relatively little about the potential in auditory signals to fully support driver behavior and actions in critical situations. Also, the level of mental workload may significantly affect the processing of auditory signals [26]. Future studies should investigate the usefulness of sound in traffic situations with different levels of workload.

In application, the appropriate type and speed of a driver response depend on the urgency of the situation. However, we know little about how intense signals with a high level of “sonic urgency” can distract the driver in demanding situations. Future studies should examine how perceived urgency in brief sounds can impact performance and interplay with information processing in critical traffic situations.
ACKNOWLEDGEMENTS

Thanks to Stefan Lindberg at the Interactive Institute Sonic Studio for helping us with the sound design. Robert Friberg and Josefin Nilsson at Scania CV AB for support and inspiration.

REFERENCES


[22] Chen, F and Jarlengrip, J, ”Listen! There are other road users close to you - improve the traffic awareness of truck drivers”, In Proceedings of HCI 2007, Beijing, China, 2007.


