Local Danger Warnings for Drivers: The Effect of Modality and Level of Assistance on Driver Reaction

Yujia Cao
Human Media Interaction
University of Twente
P.O. Box 217, 7500 AE,
Enschede, The Netherlands
y.cao@utwente.nl

Angela Mahr
German Research Center
for Artificial Intelligence
Building D3$_2$, Campus
Saarbrücken, Germany
angela.mahr@dfki.de

Sandro Castronovo
German Research Center
for Artificial Intelligence
Building D3$_2$, Campus
Saarbrücken, Germany
sandro.castronovo@dfki.de

Mariët Theune
Human Media Interaction
University of Twente
P.O. Box 217, 7500AE,
Enschede, The Netherlands
m.theune@utwente.nl

Christoph Stahl
Cluster of Excellence -
Multimodal Computing &
Interaction
Saarland University, Germ.
stahl@cs.uni-sb.de

Christian Müller
German Research Center
for Artificial Intelligence
Building D3$_2$, Campus
Saarbrücken, Germany
christian.mueller@dfki.de

ABSTRACT

Local danger warning is an important function of Advanced Driver Assistance Systems (ADAS) to improve the safety of driving. The user interface (the warning presentation) is particularly crucial to a successful danger avoidance. We present a user study investigating various warning presentations using a scenario of emergent road obstacles. Two presentation factors were selected: *modality* and *level of assistance*. The modality factor had 4 variants: speech warning, visual and speech warning, visual warning with blinking cue, and visual warning with sound cue. The level of assistance varied between with or without action suggestions (AS). In accordance with the ISO usability model, a total of 6 measurements were derived to assess the effectiveness and efficiency of the warnings and the drivers' satisfaction. Results indicate that the combination of speech and visual modality leads to the best performance as well as the highest satisfaction. In contrast, purely auditory and purely visual modalities were both insufficient for presenting high-priority warnings. AS generally improved the usability of the warnings especially when they were accompanied by supporting information so that drivers could validate the suggestions.

Author Keywords

automotive, multimodal interfaces, car2car communication

ACM Classification Keywords

H.5.2 Information interfaces and presentation: User Interfaces, User-centered design

General Terms

Design, Human Factors

IUI 2010, February 7 10, 2010, Hong Kong, China.

INTRODUCTION

Local danger warning is an important function of Advanced Driver Assistance Systems (ADAS) to improve the safety of driving. Besides directly sensing the environment to detect danger [8], recent advances in inter-vehicle communication technology (e.g.wireless ad-hoc networks car-2-car communication) further allow the exchange of information between cars [2] [13]. This enables a much wider application of local danger warnings, as drivers can be alerted to approaching danger that is not yet visible. A crucial part of successful danger avoidance is the user interface – the presentation of the warning to the driver, which is investigated in this study. We focus on a scenario where drivers are warned about road obstacles that are a short distance ahead but not yet visible (e.g. due to a bend in the road or a leading vehicle), therefore requiring an immediate reaction.

According to the situation awareness (SA) theory from Endsley [6], driving can be considered as a dynamic decision making task based on real-time maintenance of SA. SA has three hierarchical phases which all contribute to 'knowing what is going on' ([6], p.36) in a dynamic environment. The first step, perception, is to perceive the dynamics of relevant elements in the environment. The second step, comprehension, is to obtain an understanding of the perceived elements, including their significance to the task. The third step, projection, is to predict the future states of the environment. Finally, based on this updated SA, a decision can be made on how to react. For example, a driver D perceives a newly-presented warning message in the car. D understands from the warning that there is a stationary vehicle on the road-side shortly ahead. Then D predicts that there might be people moving around that vehicle. Finally D decides on a significant decrease in driving speed in order to be able to pass safely. From this perspective, the presentation of local danger warnings should aim at assisting a timely update of the D's SA, as well as helping him or her to make proper decisions. Accordingly, we chose to investigate two presentation factors, namely *modality* and *level of assistance*. A user experiment was conducted to evaluate various presentation modes for obstacle warnings, manipulated by these two factors.

PRESENTATION FACTORS: MODALITY AND LEVEL OF ASSISTANCE

Modality is a factor that is known to influence the quality and efficiency of information perception, the very first phase of SA. In the context of in-vehicle warnings, the study in [21] showed that drivers appeared more vigilant for hazards when the warnings were delivered aurally (speech) rather than visually (text). It was also suggested in [12] that auditory modalities, sound in particular, are the most suitable for presenting warnings with a high priority, whereas visual modalities are not adequate. This can be explained from several aspects. Regarding the sensory resources, driving mainly consumes visual perception resources. Thus, auditory messages interfere less with the driving task and are less likely to cause mental overload [22]. Auditory modalities are also omnidirectional, so that the information can be picked up while the eyes are kept on the road [20]. In addition, auditory modalities have a very high salience and this means they can trigger an immediate attention shift to the warning messages [3]. However, sound warnings, if over-used or not well-designed, can cause annoyance [17]. They can produce faster reactions than visual modalities, but suffer from more inappropriate responses [9]. Speech might be too time-consuming in a situation where a real quick reaction is needed. For example, it was demonstrated in [4] that the duration of speech warning messages (less than 10 words, about 5 seconds) is longer than the time needed to obtain the same information through visual modalities (1.8 seconds for icons and 3.6 seconds for text). Regarding the visual presentation of warnings, the major weak point of visual modalities is a lack of salience so that a timely perception cannot be guaranteed. However, once attended to, well-designed visual messages can be more efficient time-wise in conveying information. Therefore, in this study, additional cues (a blinking bar in the peripheral visual field or a beep sound) were provided to enhance the salience of visual warning messages, which were then compared to pure speech warnings. Furthermore, we also investigated the combination of speech and visual warnings to see whether their somewhat complementary characteristics could bring together “the best of both worlds”.

While the modality factor influences SA, the factor *level of assistance* plays a role in decision making. It manipulates whether a warning contains only the information of the approaching danger or also an action suggestion (e.g. brake, change lanes). Action suggestions are expected to shorten the time needed to infer a proper action from the latest status of SA. However, a few previous studies showed that a high level of assistance was not always favored. Using a headway maintaining scenario, [11] presented that the lowest level of assistance (providing information about leading cars) allowed the smoothest driving in terms of speed variance and was the best accepted by drivers. In contrast, the brake command to the drivers and the automated brake by the vehicle hampered the driving safety and were less acceptable. In their study, the usage of modality differed in all conditions. Therefore, the findings might be confounded with a modality influence, as the authors also discussed. Another study [14] compared a command message style (e.g. “Reduce speed”) with a notification style (e.g. “Curve ahead”). Results revealed that the command style increased the drivers’ compliance to the ADAS. However, the trust in the ADAS and the self-confidence both declined when drivers were not provided with sufficient information and were forced to rely on the command messages. Based on the SA theory, the decision making process relies on an up-to-date SA. Therefore, when drivers are aware of an information lack, they cannot be confident in their decisions. In our study, the action suggestions were always presented in combination with warning messages. When receiving the action suggestion, drivers also had the opportunity to know why they were suggested so and to decide whether they should comply. The higher level of assistance was expected to enhance the effectiveness of local danger warnings.

EXPERIMENT

We conducted a user study in which obstacle warnings were presented in various modes. For evaluation, we followed the usability guidance from ISO (ISO 9241-11, [1]). In this ISO standard, the usability of a user interface is assessed in terms of effectiveness, efficiency and satisfaction. **Effectiveness** is the accuracy and completeness with which users achieve specified goals. In our driving context, it can be reflected by the performance of danger avoidance and the level of situation awareness. SA was assessed on-line by placing recall tasks along the driving course [5]. **Efficiency** is the (temporal and cognitive) resources expended to achieve the goals, which can be measured by the reaction time and the subjective evaluation of driving load. Finally, **Satisfaction** is the users’ positive attitude towards the use of the interface, and can be obtained in terms of subjective rating.

Subjects. 32 drivers participated in this experiment, 16 men and 16 women. Their age varied between 20 years and 62 years (mean = 32.6, SD = 10.8). All participants had been in possession of a valid driver’s licence for at least two years. They were all native German speakers. Each driver was paid 15 Euros for approximately 2 hours of experiment time.

Apparatus. To promote a realistic sensation of driving, the experiment was carried out inside a real car (Mercedes-Benz R Klasse). The simulator software was hosted by a normal PC. The driving scene was a one-way highway with two lanes. No extra traffic was involved. The scene was projected onto the windshield and was updated at approximately 25Hz. The vehicle stood indoors enclosed by extra shields to reduce ambient light. This ensured a good visibility of the projection at all times of the day. Visual warning messages were presented on a 10.6-inch head-down display mounted next to the steering wheel on the right hand side. The display was also a touch screen through which the recall task was performed. Auditory signals (speech and sound) were delivered through a PC speaker located in the center of the vehicle. A web camera was mounted on the dashboard to record the frontal facial view of the driver throughout the experiment. Figure 1 shows a subject driving.

Figure 1. A subject driving in the experiment.

Messages and Presentations. A warning message described an obstacle in terms of its type (what), location (where) and distance (how far). Four types of obstacles were included: broken vehicle, fallen tree, rock and lost cargo. The location could be on the left lane, on the right lane or on the right roadside. The distance varied between 150m and 180m. In this experiment, these three information units were not equally relevant to the drivers' reaction (explained in the next section). In order to let drivers carefully perceive all three information units before reacting to the obstacles, we also presented irrelevant messages with one of the three units being 'out of range'. The irrelevant type, location and distance were air traffic, on the left roadside and more than 10km, respectively. Note that the irrelevant messages were included for experimental purposes only; it should never be a function of real ADAS systems to give false alarms.

The modality factor had 4 variants: visual + blinking cue, visual + sound cue, visual + speech, and speech only. For visual warnings, icons were designed for the four types of obstacles (Figure 2); the location of an icon on a road image indicated the obstacle location; and the distance was presented in text (Figure 3). This design was found to require shorter perception time in comparison with three other visual presentations [4]. The blinking cue was provided by a flashing red bar on top of the warning display area (Figure 3). The sound cue was a beep tone that lasted for about 350ms. All cues were delivered at the onset of a visual warning. Speech warnings narrated the obstacle type, distance and location sequentially, such as "Broken vehicle in 180 meters on the right roadside" (translated from German). This sequence was determined in an informal survey with 8 German native speakers.

Figure 2. Icon design of the four types of obstacles used in the study: broken vehicle, fallen tree, rock and lost cargo.

The level of assistance varied between with or without action suggestions (AS). AS were always given in speech, such

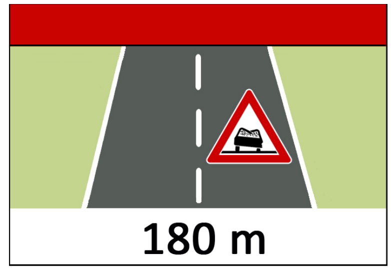

Figure 3. A visual warning displayed with a (blinking) top bar.

as "Change lanes" or "Brake" for warnings and "Attention" for irrelevant messages. A full factorial design of the two presentation factors resulted in 8 (4 × 2) presentation conditions. They are summarized in Table 1. In the visual channel, visual message and blinking bar started simultaneously, when both present. In the auditory channel, if more than one component was included, they were presented sequentially. The order from first to last was beep sound, AS and speech message.

Table 1. Presentation conditions used in the experiment.

Condition Index	1	2	3	4	5	6	7	8
Action Suggestion					×	×	×	×
Visual Message	×	×	×		×	×	×	
Speech Message	×			×	×			×
Beep Sound		×				×		
Blinking Bar			×				×	

Driving Task. The driving speed could be controlled at two levels (120km/h and 60km/h), using the gas pedal and the brake pedal. For example, when the speed was 120km/h, pressing the brake pedal would change it to 60km/h immediately. The basic driving requirement was to drive 120km/h in the nearside lane.

Obstacle warnings were presented at random intervals between 800m and 1300m. Drivers were required to change to the offside lane if the obstacle was on the nearside lane and brake if the obstacle was on the offside lane or on the right roadside. After passing the obstacle, they should change back to the nearside lane or accelerate to the higher speed again. When a warning started, the obstacle was not yet visible. It only appeared in the scene when it was 40 meters ahead of the vehicle, in the form of a colored box (Figure 4). Drivers were instructed not to wait for the obstacles to appear. Instead, they should react to the warnings immediately after they had decided on the right action to perform, because given the distance settings in the experiment there was no risk of acting unnecessarily early. Regarding the AS, they were given total freedoms from complying to it immediately to ignoring it and relying only on the warnings. Irrelevant messages did not require any reaction and were to be ignored in all conditions. In addition, brake or lane change actions should be completed at a safe distance from the obstacles. We defined a distance of more than 20m to be safe, which means the speed should have been decreased to 60km/m or the vehicle should be on the offside lane when the car reached 20m in front of the obstacle. A low pitch error sound was delivered in case of a late or missed reaction.

Speech messages all finished before the obstacles appeared and visual messages were taken off the display when the obstacles were passed. During the interval between two obstacles, drivers were asked to recall the situation of the most recent obstacle. A question was asked via speech, regarding either the type, the distance mentioned in the warning message, or the color of the obstacle. Drivers answered by pressing one of the four options displayed on the touch screen. No questions were asked after irrelevant messages.

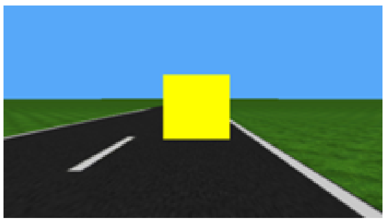

Figure 4. An obstacle on the nearside lane. The color can be yellow, green, blue or red.

Procedure. The procedure of the experiment was as follows. When entering the car, the drivers first adjusted the seat to their comfort. Then the experiment was introduced, including the driving scenario, messages, presentations and task. Then the drivers drove two practice tracks of about 15 minutes in total. In the first track, obstacles appeared in the scene without warnings and the drivers had to react as soon as they saw them. The second track included all 8 types of warning presentations, as well as the irrelevant messages. The recall task was also practiced. The main part of the experiment started after the practicing section. First, a baseline drive was performed for 5 minutes. The drivers encountered 16 obstacles without warnings. Afterwards, based on a within subject design, the drivers drove 8 tracks with different presentation conditions. Each track lasted for about 5 minutes, containing 8 warning messages and 3 irrelevant messages. The message order was randomized. During a short break after each track, the drivers filled in a questionnaire, rating driving activity load and the satisfaction with the warning presentation. A size-8 Latin square was used to counterbalance the order of the 8 tracks. At the end of the experiment, an open questionnaire was provided to obtain more comments and feedback.

Measurements. Based on the ISO usability model [1], a total of 6 measurements were derived to assess effectiveness, efficiency and satisfaction, as summarized in Table 2.

An effective warning is supposed to enhance the driving safety and the driver's SA. Therefore, to assess effectiveness, we used one measurement for the driving safety and two measurements for drivers' SA. Regarding unsafe behaviors, 3 error types were distinguished: 1) incorrect reaction, such as lane change instead of braking; 2) late reaction, which was performed less than 20m in front of the obstacle; and 3) no reaction to the obstacle warning. The *unsafe behaviors* measured the total amount of these three types of behaviors in

Table 2. Summary of dependent measurements.

Parameter	Dependent Measurement
Effectiveness	Unsafe behaviors Correct recalls Reaction to irrelevant messages
Efficiency	Reaction Time Driving activity load (Effort of attention, Visual demand, Auditory demand, Temporal demand, Interference and Situational stress)
Satisfaction	Expected satisfaction with the warnings in real-life driving

each condition. The *correct recalls* assessed the amount of recall questions that were correctly answered in each condition as a measure of SA. The *reaction to irrelevant messages* refers to the unnecessary brake and lane change actions after irrelevant messages. It also reflects how well drivers were aware of the situation conveyed by the warnings.

The efficiency of warnings was evaluated by the reaction time and the subjective ratings of driving load. The reaction time was defined as the time interval between the moment when a warning presentation started and the moment when an action was performed. A brake action was identified when the speed changed from 120km/h to 60km/h. A lane change action was identified when the lateral displacement of the car reached 10% of the maximum lateral displacement during the course of a lane change (Figure 5). To obtain subjective evaluation of the driving load, we used the Driving Activity Load Index (DALI), which is a revised version of the NASA Task Load Index adapted to the driving task [18, 19]. It contains 6 workload factors as described in Table 3. Each factor can be rated on a 6-level scale from 0 (low) to 5 (high).

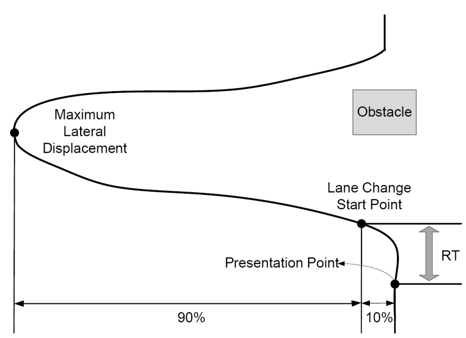

Figure 5. The calculation of the lane change starting point.

Finally, the subjective satisfaction with the usability of the warnings was obtained by a 9-level rating scale, from -4 (very unsatisfied) to 4 (very satisfied). Drivers were asked to base their judgements on the expected usability in a real-life driving situation. They always had access to the ratings they had already given, so that they could make adjustments to maintain the relative rankings between conditions.

Table 3. Description of DALI factors.

Factor	Description
Effort of attention	To evaluate the attention required by the driving activity – to think about, to decide, to choose, to look for etc.
Visual demand	To evaluate the visual demand necessary for the driving activity.
Auditory demand	To evaluate the auditory demand necessary for the driving activity.
Temporal demand	To evaluate the specific constraints due to timing demands when driving.
Interference	To evaluate the possible disturbance when driving simultaneously with any other supplementary task, such as perceiving the warning messages.
Situational stress	To evaluate the level of constraints / stress while driving – fatigue, insecure feeling, irritation, discouragement etc.

RESULTS

Unsafe Behaviors. During the baseline drive without obstacle warnings, all drivers except one showed unsafe behaviors. On average, 19.1% of the obstacles were not passed safely, because drivers reacted either incorrectly or too late. There were more unsafe lane change reactions than unsafe braking reactions. This baseline performance indicates that it is indeed a challenge for drivers to react to obstacles at short notice.

With the assistance of warnings, 18 drivers (56.3%) safely avoided all obstacles in all presentation conditions. The percentage of unsafe behaviors was reduced to 1.4% on average. The percentage within each condition is shown in Figure 6. When AS were provided, unsafe reactions only occurred in the speech condition (1.2%), because subjects did not immediately comply to the AS and did not act fast enough after the speech messages were over. When AS were not available, the number of unsafe behaviors increased in all modality conditions, especially when speech was used alone (9.0%). Figure 7 further shows the distribution of all unsafe behaviors over the two action types and three error types. Most unsafe situations were caused by incorrect actions, counting up to 58.6% from both action types. In another 31.1% of the cases, brake or lane change actions were performed correctly but too late. In the remaining 10.3% of the cases, subjects did not brake and passed the obstacles with the higher speed. Late actions and no actions all occurred in the pure speech conditions.

A two-way repeated-measure ANOVA further revealed a significant assistance effect ($F(1, 31) = 15.8$, $p < .001$) and a significant modality effect ($F(3, 29) = 5.1$, $p < .01$). The higher level of assistance (with AS) led to safer driving than the lower level (without AS). Among the four modality variants, the performance was equally good for the three variants that included visual modalities. However, driving safety decreased significantly when speech was used alone, compared to the other three variants ($F(1, 31) = 16.2$, $p < .001$ by Helmert contrast). There was also a significant interaction between modality and the assistance level ($F(3, 29) = 5.5$, $p < .01$), meaning that when drivers received assistance in the form of AS, presenting the warnings using speech alone was much less harmful to the driving safety.

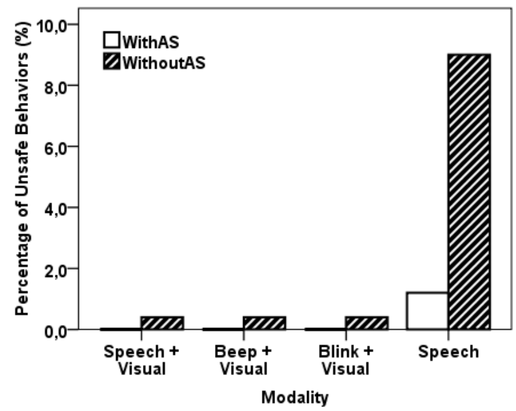

Figure 6. Percentage of unsafe behaviors in each condition.

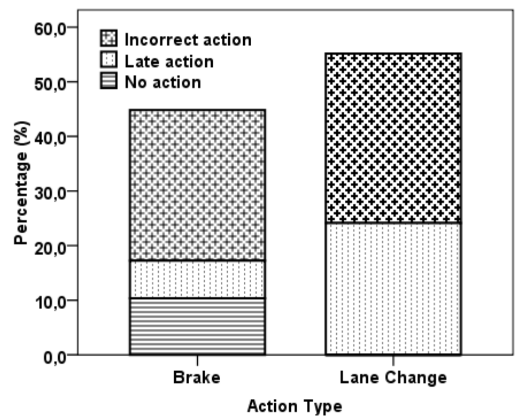

Figure 7. Distribution of unsafe behaviors over action and error types.

Correct Recalls. The recall task was generally performed well. On average, 91.7% of the questions were answered correctly. The performance was the best when AS, speech and visual information were all provided (96.1% correct), and was the worst in the two pure speech conditions (87.9% correct). Repeated ANOVA was conducted and confirmed a significant modality effect on the quality of recall ($F(3, 29) = 3.5$, $p < .05$). Combining speech and visual messages led to significantly more correct recalls than the other three modality variants ($F(1, 31) = 10.2$, $p < .01$, by Helmert contrasts). This result suggests that the maintenance of SA could be assisted by presenting information through more than one sensory channel. Pure speech resulted in the worst recall performance ($F(1, 31) = 6.6$, $p < .05$). This might be due to the transience of speech, which does not allow repeated perception. No difference in recall was found between the two visual conditions with additional cues. The level of assistance did not influence the recall performance either ($F(1, 31) = 0.03$, n.s.). This is not surprising, because the action suggestions did not contain relevant information to the questions.

Besides modality, the topic of the questions also had an effect on the quality of recall ($F(2, 30) = 8.0$, $p < .01$). As shown in Figure 8, the color of obstacles was recalled the worst, compared to the type and the distance ($F(1, 31) = 8.9$, $p < .01$). A possible explanation is that of the three types of information, color was not mentioned by the warnings and was the least relevant to the driving task, so the users may have paid less attention to it. In addition, the obstacle type was recalled better than the distance ($F(1, 31) = 9.8$, $p < .01$). The reason might be that the icon presentations of obstacles were more vivid than the text presentations of distances, because shapes and colors have great salience to human information processors due to the sharp contrasts they are able to create [16].

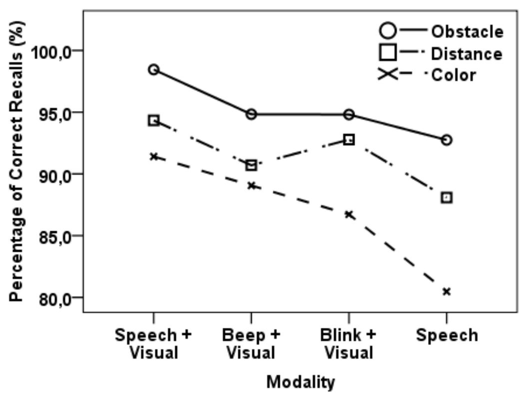

Figure 8. Percentage of correct recalls (averaged over the assistance levels for each modality variant).

Reactions to Irrelevant Messages. To obtain more insight into the level of SA created by the warning messages, we looked into drivers' reactions to irrelevant messages. On average, drivers reacted to 12.1% of irrelevant messages with unnecessary braking or lane changes. In most cases, drivers self-corrected their actions very soon. The number of unnecessary reactions was influenced by the modality factor ($F(3, 29) = 5.2$, $p < .01$) but not by the level of assistance ($F(1, 31) = 0.1$, n.s.). The latter finding is not surprising, since the irrelevant action suggestion ("Attention") did not refer to a physical driving action.

It can been seen in Figure 9 that unnecessary reactions occurred the most when messages were presented visually with the blinking cue ($F(1, 31) = 9.8$, $p < .01$, by Helmert contrast: blink + visual vs. other three). Since this is the only modality variant that is purely visual, this result might suggest that the lack of auditory modalities could make the drivers less vigilant (cf. [21]) and less careful. Interestingly, unnecessary reactions occurred the least when speech was used alone, which seems to contradict its negative impact shown by the measure of unsafe behaviors. However, a possible explanation is obtained when analyzing the three types of irrelevant information separately. As Figure 9 shows, when visual messages were provided (except speech only), drivers reacted more to irrelevant distances than to irrelevant obstacles and locations ($F(1, 31) = 29.5$, $p < .001$ by Helmert contrast: distance vs. other two). Irrelevant obstacles again resulted in more reactions than irrelevant locations. This pattern suggests that drivers may have used a common sequence to scan the visual messages. Since the location of an obstacle was the most relevant to the driver's reaction, it was probably inspected first. Then the type of an obstacle would have been observed before the distance, because it was more spatially integrated with the location. Since early reactions could be performed at any time before all three information units were analyzed, the later the irrelevant unit was perceived, the more unnecessary reactions were performed. In the pure speech conditions, however, the sequence of perception was fixed, with the location information being presented last. This guaranteed the detection of irrelevant obstacles and distances before any action could be performed. However, this advantage does not mean that the speech modality is the most able to enhance drivers' SA. In contrast, when looking at those cases where the location was irrelevant, the speech condition still produced the most unnecessary reactions, which is consistent with the fact that 82.4% of the incorrect reactions to regular warnings occurred in the pure speech conditions. Finally, these results suggest that it takes more effort to interpret spatial information (e.g. locations) when it is presented only orally than when it is also presented visually. This is consistent with literature (e.g. [7]) stating that spatial information can be effectively presented by visual non-verbal modalities, such as image.

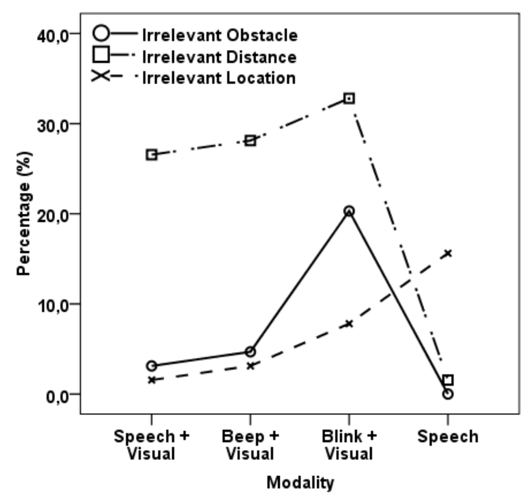

Figure 9. Percentage of irrelevant messages that were unnecessarily reacted to (averaged over the assistance levels for each modality variant).

Reaction Time. First, we looked at whether reaction time differed between the brake and the lane change actions. A three-way repeated measure ANOVA was conducted and revealed that the action type (brake/lane change) did significantly influence the reaction time. On average, the lane change actions were performed 0.24 seconds faster than the brake actions ($F(1, 31) = 15.4$, $p < .001$). This result falls in line with previous findings stating that steering is 0.15 to 0.3 seconds faster than braking, because of a lower response complexity [10]. Due to the difference, further analyses were conducted separately on the two types of actions.

When speech was used alone without AS, the reaction time was particularly long, because no action could be performed before the end of the speech. (Remember that the location information, which was essential for deciding which action should be performed, was given at the end of the message.) As Figure 10 shows, the average reaction time was 5.1s (170m) for braking and 4.8s (160m) for lane change. When drivers were assisted with AS, the reaction time in the pure speech condition was reduced to 2.3s for braking and 2.1s for lane change. However, these reaction times are still the longest among the four conditions where AS were provided. It seems that the drivers hesitated to comply to the action suggestions when no visual information was available. Playing back the video recordings, we saw that when visual messages were provided, most drivers checked the display during the course of the action suggestion or shortly afterwards. This might explain their hesitation in the pure speech condition. Generally, drivers still chose to comply to the AS without the complete information of an obstacle. However the reaction was more or less delayed.

Besides the two pure speech conditions, the difference in reaction time among the other six conditions was relatively minor. However, ANOVA still revealed significant modality effects (braking: $F(2, 62) = 25.0$, $p < .001$; lane change: $F(2, 30) = 22.0$, $p < .001$) as well as significant assistance effects (braking: $F(1, 31) = 27.9$, $p < .001$; lane change: $F(1, 31) = 24.7$, $p < .001$). Regarding the level of assistance, both braking and lane change were performed faster with AS than without AS, as shown in Figure 10. On average, AS accelerated braking by 470ms (15.7m) and lane change by 130ms (4.3m). This finding suggests that although most drivers spent time on validating the action suggestions with the visual messages, it was still less time/effort consuming compared to making decisions without any suggestion. In this driving simulation, the structure of visual displays and the content of warning messages were rather simple, thus the benefit of AS in terms of reaction time was only below a half second. However, this benefit could be larger in a more dynamic situation where the warning messages are more diverse and less expected.

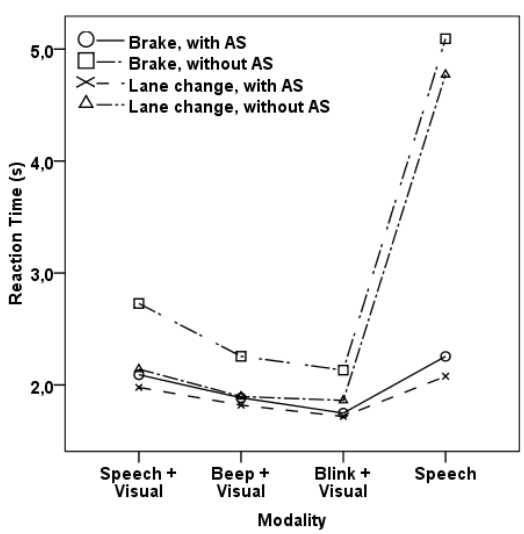

Figure 10. Average reaction time in each condition for brake and lane change respectively.

Among the three modality variants with the visual modality, the combined use of speech and visual messages led to the longest reaction time for both braking reactions ($F(1, 31) = 44.8$, $p < .001$) and lane change reactions ($F(1, 31) = 44.6$, $p < .001$). Comparing the two visual conditions with attentional cues, both types of reactions were faster with the blinking cue than with the beep cue. However, the difference was only significant for the lane change action ($F(1, 31) = 6.9$, $p < .05$).

A somewhat reversed pattern could be found when comparing the reaction time with the drivers' reactions to irrelevant messages (Figure 9). It seems that the longer the reaction time was, the fewer unnecessary reactions there were. This observation suggests a variation in the level of vigilance. When speech was presented with the visual information, drivers seemed the most willing to carefully inspect the visual messages before reacting to them. On the other hand, in the purely visual conditions, drivers reacted the fastest but most often overlooked the irrelevant items. As a couple of drivers commented, they found the purely visual condition boring. This again suggests that the lack of auditory signals could decrease the drivers' vigilance, causing them to be less attentive and behave less carefully.

Driving Activity Load. Using the DALI questionnaire, the load of driving activities (to avoid obstacles during driving) was rated on 6 dimensions: effort of attention, situational stress, visual demand, auditory demand, temporal demand, and interference. Each dimension could be rated from 0 (low) to 5 (high). First, we averaged the rating scores over all 6 dimensions as a global assessment of driving load. Repeated ANOVA revealed that the global driving load was influenced by both the modality ($F(3, 93) = 5.0$, $p < .01$) and the assistance level ($F(1, 31) = 23.4$, $p < .001$). As Figure 11 shows, the driving task was generally less demanding with AS than without AS. Regarding modality, the driving task was rated the least demanding when speech and visual modalities were combined ($F(1, 31) = 5.8$, $p < .05$, by Helmert contrast), and the most demanding when speech was used alone ($F(1, 31) = 5.3$, $p < .05$). No significant difference in the driving load was found between the visual conditions with attentional cues.

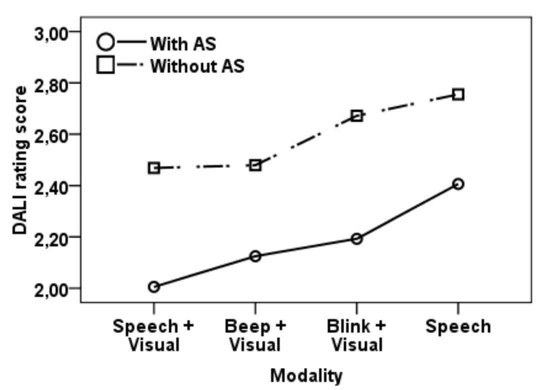

Figure 11. Average DALI rating scores over 6 dimensions.

When zooming into the DALI dimensions, we found that besides the two perceptional (visual and auditory) demands, the remaining 4 showed a common pattern over the 8 conditions. To validate this observation, we conducted the reliability analysis on these four variables in each of the 8 conditions. Results showed that the Cronbach's Alpha (coefficient of internal consistency) values were all greater than 0.8, indicating that ratings on these 4 DALI dimensions are indeed highly consistent. The common pattern among them obviously corresponds with the global driving load (Figure 11). In contrast, the variance in the perceptional demands was closely related to the manipulation of modalities (Figure 12). As expected, the visual demand was rated the lowest in the speech conditions. Additional demand was induced by visual warning messages, but the increment was at a minimum when the visual warnings were combined with speech. AS also generally reduced the visual demand, probably because less effort was needed to analyze the information on the display. Regarding auditory demand, the scores were the lowest in the purely visual condition and the highest in the pure speech conditions. Interestingly, the auditory demand induced by speech warnings was less when visual messages were provided at the same time. The reason might be that drivers spent less effort on listening to the speech because they could also get some information visually.

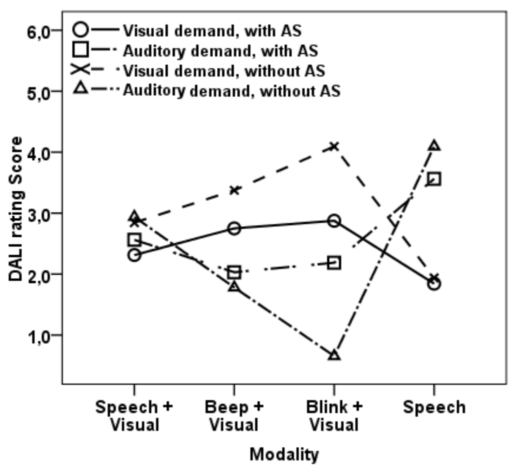

Figure 12. DALI ratings on visual and auditory demand.

Subjective Satisfaction. Drivers rated how satisfied they would be to use each warning presentation in real-life driving. On a 9-level rating scale from -4 (very unsatisfied) to 4 (very satisfied), the judgments were generally positive on average, except for the purely visual warnings without AS (Figure 13). The condition with speech, visual messages and AS was rated the highest. ANOVA confirmed that the subjective satisfaction was affected by both the modality ($F(3, 29) = 22.1$, $p < .001$) and the level of assistance ($F(1, 31) = 43.2$, $p < .001$). Subjects were more satisfied with the higher level of assistance than the lower one. Among the four modality variants, combing speech and visual messages was the most satisfying ($F(1, 31) = 33.0$, $p < .001$). The visual condition with beep cues was the second best ($F(1, 31) = 8.7$, $p < .01$). No significant difference was found between the purely auditory (speech) and the purely visual (with blinking cues) variants. This is due to an interaction between the modality and the level of assistance. When drivers were given suggestions on what to do, they preferred the purely visual messages to the pure speech messages, because they could quickly validate the action suggestions and react with confidence. Without AS, on the other hand, the purely visual messages were rated lower than speech. Several drivers explained that the messages were less salient without any auditory signals. Although the blinking cue worked well in this experiment, it might be less effective when there are more distractions in the car, such as radio and conversations. It was also mentioned that when there was no action suggestion, it was more important to keep one's eyes on the road.

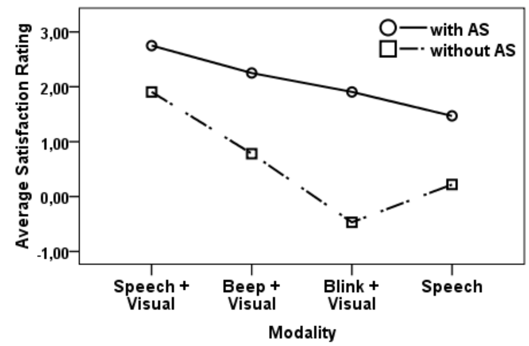

Figure 13. Subjective ratings on the satisfaction with warnings.

DISCUSSION

In this experiment, obstacle warnings were presented to drivers, in order to assist them to safely avoid emergent local dangers. It is shown that warnings presented in various modes all resulted in less unsafe behaviors, compared with the baseline condition where drivers were not warned in advance. This indicates that the safety of driving could indeed benefit from providing local danger warnings to drivers. In addition, this experiment also revealed that the way the warnings were presented, with regard to modality and level of assistance, could influence their effectiveness and efficiency, and drivers' satisfaction with them.

Level of Assistance. The level of assistance varied between with and without AS. The AS were supposed to directly assist drivers to decide on how to react to the obstacles. In this experiment, the advantage of using AS was consistently shown by various measurements. Less unsafe behaviors occurred with than without AS. AS accelerated both braking and lane change actions. Although the benefit was below a half second, it could be more pronounced when the warning messages are more dynamic and less expected and the situational decisions are more complex. Regarding the driving load ratings, the attentional effort, stress, visual demand, temporal demand and interference were all rated lower when AS were provided. Moreover, drivers expected the warnings to be more satisfying in real-life driving if AS would be available. In summary, based on the ISO usability model, adding AS made the warnings more effective, efficient and satisfying in this experiment.

This finding seems inconsistent with some previous studies which favored a lower level of assistance for drivers. In the headway assistant presented in [11], users either received braking commands that were presented in speech, or information about a leading car that was presented visually. Information was always presented much earlier than the commands so that there would be sufficient time to analyze it. Subjects in this study accepted the information style more than the command style. It is not surprising that drivers preferred to be notified as early as possible and decide on what to do themselves. However, in the situation of an emergent danger, action commands could be more effective because time might not always be sufficient for analyzing the information and deciding the next steps to take. In our study, the drivers were notified of the obstacles less than 10 seconds ahead, which might explain their strong preference for action suggestions. In addition, AS in our experiment always pointed to correct actions. Although aware of this fact, most drivers still validated the AS with the visual messages when they were available. This finding suggests that when integrating AS into real ADAS systems, it is important to provide relevant information at the same time, because the AS might not always be perfect and drivers should still have the chance to decide for themselves. In summary, we could say that AS are probably more beneficial when the situation is more urgent and they should be provided together with relevant information about the situation.

Modality. Regarding the four modality variants used in this experiment, our usability assessment generally suggests that the speech and visual modality combined variant is the most usable, speech alone is the least usable and the two visual variants with attentional cues lie in between. Speech has three major drawbacks when used in our scenario. First, the duration of the speech messages was too long, leaving the drivers only about 2 seconds to react. Second, speech does not allow multiple perception without a repeat function. Third, it is not suitable to convey spatial information such as object locations, which were the most relevant to this task. As a result, using only speech led to the most unsafe behaviors, the worst recall, the longest reaction times and the highest ratings on driving load. In line with previous findings [12], our result suggests that speech alone is not adequate for presenting high priority warnings.

Comparing the two kinds of attentional cues, although no significant difference was found in terms of effectiveness and efficiency, the beep cue was clearly preferred by the drivers. The satisfaction ratings showed that drivers expected the warnings with beep cues to be more satisfying in real-life driving than the warnings with blinking cues. In the general questionnaire at the very end of the experiment, drivers were asked to rate the usefulness of the two kinds of cues, on a 6-level scale from 0 (not useful) to 5 (very useful). The beep cue received significantly higher scores (mean=3.8, sd=1.3) than the blinking cue (mean=2.1, sd=1.8), shown by a t-test ($p < .001$). In addition, 15.6% of the drivers commented that they had not noticed the blinking top bar at all. This result confirms that visual modalities have a lack of salience, thus are less suitable for warning presentations than auditory modalities. In this experiment, this lack of salience did not harm the driving safety, because the warnings were always well expected, so no warning was missed even if the blinking cues were not detected. However, the visual cues might be less useful in a real-life driving situation, because warnings are normally unexpected and thus are more likely to be missed.

Besides the lack of salience, our results also suggest that purely visual presentations could reduce the drivers' vigilance for hazards. In the condition with visual message, blinking cue and no AS, drivers reacted the fastest but overlooked the most irrelevant messages. It seems that they were less willing to check the warnings properly and were less careful with their reaction. Regarding the satisfaction rating, the purely visual condition was the only one that received a minus score on average. Several drivers described this condition as boring. These findings stand in line with the study in [21] where drivers appeared more vigilant when warnings were delivered aurally than visually. Therefore, it could be suggested that the presentation of high-priority warnings needs to include auditory signals.

Although pure speech and purely visual messages both have major drawbacks, the combination of the two significantly improved the usability of the warnings. Their complementary characteristics provided both salience and a freedom of perception. Consequentially, this modality variant had the best recall performance, the lowest driving load score and the highest satisfaction score. Although the reaction was not the fastest, drivers reacted the least to irrelevant messages (when presented visually), indicating a better awareness of the situation conveyed by the warnings. As several drivers explained, receiving information through multiple channels would let it be picked up more quickly, because there are more choices for perception and drivers could choose one that is compatible with the driving activities in that specific situation. It seems that a redundant use of multiple modalities did bring "the best of both world" in this study. This redundancy benefit was also previously obtained using an in-car navigation task [15], where drivers made the fewest errors and showed the most proper vehicle control when the navigation messages were delivered both visually and aurally (relative to the single modalities).

CONCLUSIONS AND FUTURE WORK

In this study, we investigated the presentation of local danger warnings to drivers. A user experiment was conducted in which obstacle warnings were presented in 8 modes, manipulated by four modality variants and two levels of assistance. In accordance with the ISO usability model, 6 measurements were derived to assess effectiveness and efficiency of the warnings and drivers' satisfaction. Based on the results, several suggestions are made for the design of local danger warnings in ADAS systems. First, regarding modality, purely auditory or purely visual modalities are both insufficient for presenting high-priority warnings. The combination of both can be beneficial, especially the combination of speech and well-designed visual messages. In speech messages, the critical information needs to be presented first. The speech should be kept short and a 'repeat' function might be worth consideration. If spatial information such as locations needs to be presented, it is better to present it visually and graphically. Second, if it is necessary to use additional cues to attract drivers' attention to the warnings, auditory signals (e.g. sounds) are more suitable than visual signals (e.g. blinking objects), due to a higher level of salience. Finally, it could be beneficial to provide action suggestions in urgent situations. However, AS should always be accompanied by supporting information, because they might not always be perfect and thus drivers always need the opportunity to validate the suggested action.

Future work can be considered in the following directions: using a head-down display, we found that providing attentional cues in the peripheral visual field was not very effective. However, this finding needs to be reexamined using head-up displays, such as direct projection on the windshield. This is expected to enhance the effectiveness of visual cues. Moreover, the warning messages were relatively simplex in this experiment and there was no extra traffic involved. It would be useful to extend the current study with more diverse warnings and more dynamic driving scenarios.

ACKNOWLEDGEMENT

This work was funded by the Dutch Ministry of Economic Affairs (project Interactive Collaborative Information Systems, grant nr: BSIK03024) and the German Ministry of Education and Research (project Car-Oriented Multimodal Interface Architecture, grant number 01IW08004).

REFERENCES

1. ISO 9241: Ergonomic requirements for office work with visual display terminals (vdts) - part 11: Guidance on usability, 1998.
2. Identify intelligent vehicle safety applications enabled by DSRC. Technical Report DOT-HS-809-859, U.S. Department of Transportation, National Highway Traffic Safety Administration, Springfield, Virginia, USA, March 2005.
3. N. Bernsen and L. Dybkjaer. Exploring natural interaction in the car. In *CLASS Workshop on Natural Interactivity and Intelligent Interactive Information Representation*, pages 75–79, 2001.
4. Y. Cao, S. Castronovo, A. Mahr, and C. Müller. On timing and modality choice with local danger warnings for drivers. In *1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2009)*, pages 75–78, 2009.
5. M. R. Endsley. Measurement of situation awareness in dynamic systems: Situation awareness. *Human Factors*, 37(1):65–84, 1995.
6. M. R. Endsley. Towards a theory of situation awareness in dynamic systems. *Human Factors*, 37(1):32–64, 1995.
7. S. K. Feiner and K. R. McKeown. Automating the generation of coordinated multimedia explanations. *IEEE Computer*, 24:33–41, 1991.
8. D. M. Gavrila. Sensor-based pedestrian protection. *IEEE Intelligent Systems*, 16(6):77–81, 2001.
9. R. Graham. Use of auditory icons as emergency warnings: evaluation within a vehicle collision avoidance application. *Ergonomics*, 42(9):1233–1248, 1999.
10. M. Green. "How long does it take to stop?" Methodological analysis of driver perception-brake times. *Transportation Human Factors*, 2(3):195–216, 2000.
11. A. Kassner. Meet the driver needs by matching assistance functions and task demands. In *European Conf. on Human Centred Design for Intelligent Transport Systems*, pages 327–334, 2008.
12. C. Kaufmann, R. Risser, A. Geven, and R. Sefelin. Effects of simultaneous multi-modal warnings and traffic information on driver behaviour. In *European Conf. on Human Centered Design for Intelligent Transport Systems*, pages 33–42, 2008.
13. T. Kosch. Local danger warning based on vehicle ad-hoc networks: Prototype and simulation. In *1st International Workshop on Intelligent Transportation (WIT 2004)*, 2004.
14. J. D. Lee, B. F. Gore, and J. L. Campbell. Display alternatives for in-vehicle warning and sign information: message style, location and modality. *Transportation Human Factors*, 1(4):347–375, 1999.
15. Y. C. Liu. Comparative study of the effects of auditory, visual and multimodality displays on drivers performance in advanced traveller information systems. *Ergonomics*, 44(4):425–442, 2001.
16. N. H. Lurie and C. H. Mason. Visual representation: Implications for decision making. *Journal of Marketing*, 71:160–177, 2007.
17. D. C. Marshall, J. D. Lee, and P. A. Austria. Alerts for in-vehicle information systems: Annoyance, urgency, and appropriateness. *Human Factors*, 49(1):145–157, 2007.
18. A. Pauzié. Evaluating driver mental workload using the driving activity load index (DALI). In *Proc. of European Conference on Human Interface Design for Intelligent Transport Systems*, pages 67–77, 2008.
19. A. Pauzié and G. Pachiaudi. Subjective evaluation of the mental workload in the driving context. *Traffic and Transport Psychology*, pages 173–182, 1997.
20. N. Sarter. Multimodal information presentation: Design guidance and research challenges. *International journal of industrial ergonomics*, 36(5):439–445, 2006.
21. B. Seppelt and C. Wickens. In-vehicle tasks: Effects of modality, driving relevance, and redundancy. Technical Report AHFD-03-16/GM-03-2, GM Corp, 2003.
22. C. D. Wickens. Multiple resources and mental workload. *Human Factors*, 50(3):449, 2008.