Auditory and Head-Up Displays in Vehicles

Christina Dicke[1], Grega Jakus[2], and Jaka Sodnik[2]

[1] Quality and Usability Lab, Telekom Innovation Laboratories, TU Berlin, Germany
christina.dicke@tu-berlin.de
[2] University of Ljubljana, Faculty of Electrical Engineering, Slovenia
{grega.jakus,jaka.sodnik}@fe.uni-lj.si

Abstract. The aim of the user study presented in this paper was to investigate the efficiency of single and multimodal user interfaces for in-vehicle control and information systems and their impact on driving safety. A windshield projection (HUD) of a hierarchical list-based visual menu was compared to an auditory representation of the same menu and to a combination of both representations. In the user study 30 participants were observed while operating a driving simulator and simultaneously solving tasks of different complexity with the three interfaces. The variables measured in the user study were task completion times, driving performance and the perceived workload. Our study shows that the single modality auditory interface is the least efficient representation of the menu; the multimodal audio-visual interface, however, shows a strong tendency to be superior to both the auditory and visual single modality interfaces with regards to driver distraction and efficiency.

Keywords: Human-computer interaction, auditory interface, head-up display, car simulator, driving performance.

1 Introduction

Head-up displays (HUDs) are the current state-of-the-art solution intended to reduce driver errors originating in distractive interfaces, such as onboard entertainment displays or navigation systems. HUDs reduce the frequency and duration of glances at Head-down displays (HDDs) by presenting information directly on the windshield in the driver's field of vision.

HUDs, when compared to HDDs, have been shown to reduce the response times to unanticipated road events [1], reduce navigational errors [2], and lead to smaller variances in lateral acceleration and steering wheel angle [3]. Charissis et al. [4] found that HUDs dramatically reduce the number of collisions and improve the maintenance of following distance in low visibility situations.

On the other hand, HUDs have also been shown to increase mental workload as indicated by longer response times in high workload situations [5, 6]. The so-called cognitive capture, i.e. when the drivers' attention is unconsciously shifted away from the road and is mainly focused on processing the information presented by the HUD [7, 8], has been identified as one of the disadvantages of HUDs. The resulting

M. Kurosu (Ed.): Human-Computer Interaction, Part II, HCII 2013, LNCS 8005, pp. 551–560, 2013.

perceptual tunneling may lead to a delayed reaction or a complete absence of response to situational changes in the driving task [9, 10].

The results of our previous study [11] give grounds for the assumption that by presenting information non-visually, negative influences such as visual distraction and cognitive capture can be reduced. The results also suggest that auditory information presentation can reduce the perceived workload and lead to fewer driving errors when compared to visual interfaces presented on a HDD. These findings are supported by the results of a recent study by Weinberg et al. [12], who established that navigating through aurally presented lists had a lower impact on mental workload when compared to HDDs and HUDs.

These findings support the assumption that a combination of auditory, especially speech-based information presentation and text-based information presentation through a HUD may in fact present the "best of both worlds": the attention and safety benefits of an eyes-free approach with the higher information processing rate of visual information presentation. By offering drivers a multimodal interface, they can choose to either listen to or read the required information depending on their situational cognitive capacities and/or their personal preferences.

The study presented in the following chapters was conducted to evaluate this assumption and to gain further insight into the subjectively and objectively measured impacts of using a single modality or a multimodal in-car information display while driving.

2 User Study

The aim of our user study was to evaluate the potential benefits of a multimodal user interface compared to two single modality interfaces for an in-vehicle information system through a series of representative tasks. The study used a within-subjects design (repeated measures) and compared a spoken auditory representation of a hierarchical menu structure to a visual representation (in the form of a HUD) of the same structure and a combined, multimodal representation. To imitate a realistic environment, the study was conducted in the driving simulator shown in Fig. 1, left.

The simulated on-board computer was operated through a custom-made interaction device attached to the steering wheel (see Fig. 1, right) enabling various tasks related to navigation, communication, and various on-board control systems. The primary interest of this study was to evaluate the efficiency of the interfaces and their impact on driving performance.

The efficiency was assessed by measuring the time required to complete a given set of tasks and the self-reported mental effort. The individual driving performance was analyzed by noting and rating anomalies and unsafe driving - such as swerving, sudden and unnecessary speed reductions, disobeying the traffic rules or even causing an accident - which occurred while performing the tasks.

2.1 Interaction with the Menu

Participants were asked to perform a set of tasks with the simulated on-board information system. The system imitated various navigational, communicational and other functionalities, while the interaction was based on a structured set of menu options. The top-most level of the structure was labeled the “Main menu”. It consisted of five sub-menus that were further structured into sub-sub-menus.

The menu was operated through a custom made device located at the backside of the steering wheel (Fig. 1, right). The device’s scrolling wheel and its two buttons supported the following three interactions: descending into a sub-menu, confirming an option and exiting a sub-menu or returning to the next higher level.

2.2 Conditions

Three experimental conditions were compared:

1. a visual representation of the menu (V),
2. an auditory representation of the menu (A), and
3. an audio-visual representation of the menu (AV).

Visual Interface (V). The visual interface was a HUD displaying the available items of the menu structure. The menu was projected to the right-central part of the windshield (Fig. 1, left). The approximate size of the projection was 20 x 20 cm. The text was displayed on transparent background in high contrast colors:

- the title of the active menu was displayed in green on the top of the HUD,
- the available but unselected items were displayed in yellow,
- the selected item was highlighted in red and a slightly bigger font size .

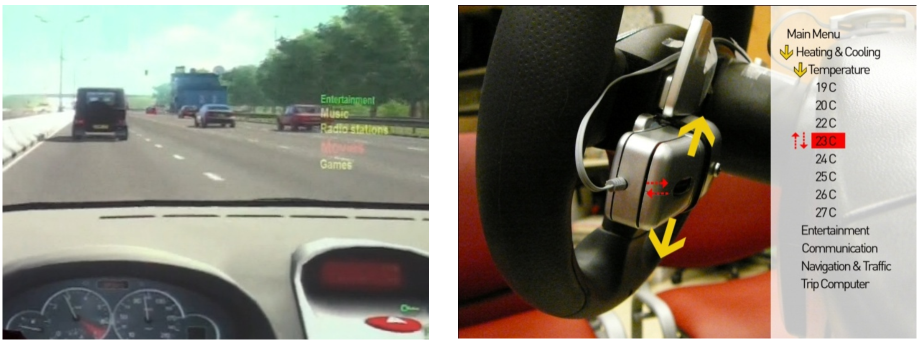

Fig. 1. The visual interface implemented as a HUD (left) and the interaction device (right). By clicking the lower mouse button (yellow highlight), the participants selected an item or descended in the hierarchy; the scrolling wheel (red highlight) was used to select an item; the upper mouse button (yellow highlight) was used for ascending in the hierarchy.

The HUD was implemented as a "sliding window" displaying at most five of the available items of the active menu. When available, the remaining items became accessable by scrolling up or down the menu.

Auditory Interface (A). In the auditory interface condition, pre-recorded readings of menu items were played through two computer speakers placed at the sides of the simulator imitating the in-vehicle speakers mounted in the side doors. The title of each sub-menu was announced upon entering the menu and menu items were read each time the virtual cursor's position/selection was changed. When participants reached the last (first) item of the menu, the end (beginning) of the list was indicated to them by a repeated reading of the item.

Audio-Visual Interface (AV). In the audio-visual condition, the visual and auditory interfaces described in the previous sections were presented simultaneously.

2.3 Tasks

Within each of the three experimental conditions, each participant was given three simple ("conventional") tasks followed by two difficult ("complex") tasks. All tasks in the conventional and all tasks in the complex group had the same minimum number of interactions (clicks, turns of the scrolling wheel) required for completion. The tasks in each group were chosen to be comparable in cognitive workload (e.g. scanning through text, listening to messages, considering options). A sample of a conventional task is: "Set the temperature to 24 degrees." A sample of a complex task is: "You are waiting for an e-mail from Denny Crane with the name of the restaurant where he would like to meet you. Please check your e-mail inbox and tell the name of the restaurant mentioned in the e-mail to the experimenter."

3 Methods

3.1 Test Subjects

A total of 30 test subjects (9 female and 21 male) with a valid driving license participated in the study. The participants were in average 29 years old and had in average 11 years of driving experience. They all reported to have normal sight and hearing.

3.2 Test Groups

To ensure comparable driving conditions, all participants were given loose navigation instructions while driving. To eliminate the influence of traffic density and street layouts, each of the three conditions (A, V, AV) was repeated with two different driving dynamics: a busy city center and a motorway. To avoid learning effects, six experimental groups were formed, each with a different combination of conditions and routes. For example, the participants from the group "A" started with the auditory interface on the "high speed motorway" route, then proceeded with the visual interface on the same route, and concluded with the AV interface on the "low speed city" route.

In the “high speed” scenario, participants drove on a motorway with a low traffic density with an average speed of 72 km/h (45 mph). In the “low speed” scenario, participants drove through a busy city center with an average speed of 33 km/h (21 mph).

In order to obtain a measure of the overall driving performance without the interference of a secondary task, a virtual control group was formed. The data of this control group was obtained by evaluating the driving performance of all drivers in the intervals between tasks.

3.3 Experiment Procedure

Before the experiment participants were informed about the nature and structure of the study. Participants were then asked to complete a pre-study questionnaire (age, gender, hearing and sight disabilities, driving experience, prior experience with simulators, and proneness to sea or simulator sickness). The participants were then thoroughly introduced to the driving simulator, the interaction device, the structure and content of the menu, and the three interfaces. Finally, they were given 20 minutes to familiarize themselves with the simulator while driving on a test route.

Participants then proceeded with a series of tasks.. Each task was read to participants loudly and clearly. Participants were asked to perform the given tasks as quickly as possible, but to also obey traffic rules and drive the car safely.

Task completion times, interaction activity and driving behavior were recorded automatically for each task. The measurements started when participants started to solve tasks and were stopped when the task was completed successfully. For the purpose of post-evaluation of the driving performance, the entire user study was recorded with a digital video camera.

After finishing each of the experimental conditions, participants were given an electronic version of the NASA TLX questionnaire [13] to evaluate their perceived workload for each particular interface.

After all three conditions were completed, participants were asked to fill in a short post-study questionnaire on their overall perception of the interfaces, their design, the design and realism of the driving simulator, and the complexity of given tasks.

3.4 Technical Setup

The experimental environment consisted of three basic elements: the driving simulator, the user interface application (UI) and the management and logging software (ML) suite.

The UI application controlled the interaction device and the output interfaces (A, V, AV). The application reported events associated with the interaction device (button clicks, scrolling wheel turns) to the ML software suite. The setup of the experimental environment was described in more detail in [14].

4 Results and Interpretations

4.1 Efficiency of Interfaces

The efficiency of the interfaces was measured through task completion times. The averaged results (of all participants in the study) are presented separately for each condition.

The average task completion times are presented in Fig. 2, left. Due to non-homogeneity of variances between data groups, the Kruskal-Wallis test was chosen to confirm the significance of differences between experimental conditions: H = 40.279, 2 d.f., p < .001. A post-hoc Games-Howell test with a .05 limit on family-wise error rate confirmed significant differences (p < .001) between the auditory interface (A) and other two interfaces (V and AV). The difference between the V and AV interfaces was not found to be significant.

The results confirm one part of our initial hypothesis: the auditory interface proved to be the least efficient among the three tested interfaces, as it was found to be the slowest and it required the most (physical) effort to complete tasks. However, a superiority of the multimodal interface over the visual interface could not be confirmed.

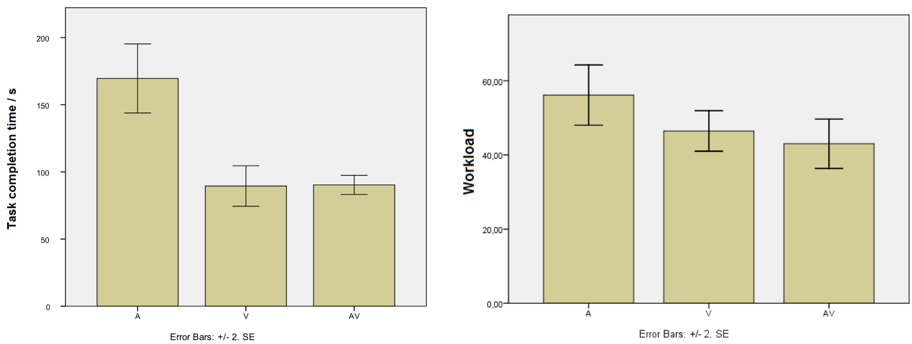

Fig. 2. Average task completion times (left) and subjectively perceived workload (unweighted) for all thee interfaces (right)

4.2 Workload

Participants' perceived workload was measured using the NASA-TLX assessment tool [13]. The procedure derives an overall workload score from the ratings on the following six subscales: mental demand, physical demand, temporal demand, performance, effort and frustration.

Fig. 2, right shows the overall workload for all three interface conditions. The results of the ANOVA show a significant difference between the three conditions: F(2.87) = 4.035; p = .021. The post-hoc Bonferroni test with a .05 limit on a family-wise error rate revealed that the reported workload in the audio condition (A) was significantly higher (p = .019) when compared to the workload in the AV condition.

The differences between the visual (V) and other two conditions (A and AV) were not found to be significant.

No significant difference between the visual and the multimodal interface was found, but again a trend towards a lower impact of the multimodal interface on the overall perceived workload is indicated.

4.3 Driving Performance

The driving performance was evaluated on the basis of the logged data in the simulator and then revised by a retrospective analysis of the video recordings. Four aspects were assessed: (a) lateral instability (swerving), (b) anomalies in driving speed (sudden and unnecessary speed reductions, inappropriate speed for a certain situation), (c) obeying common driving rules and signs (d) and causing accidents.

For each aspect, the performance of the participants was rated using the following system:

- 0 points – no anomalies were observed
- 1 point – one or very few anomalies were observed
- 2 points – anomalies were repeated several times
- 3 points – anomalies were repeated constantly resulting in very unsafe driving.

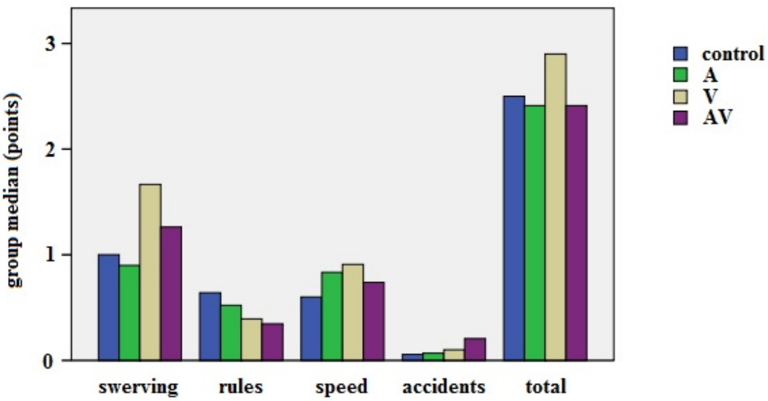

Fig. 3. Penalty points for driving errors per error category and total for all three interfaces and the control group

The scores of the individual aspects were then combined into an overall driving performance assessment. For the purpose of comparability, the performance of the control group was determined by measuring the performance of participants while driving without any distraction from secondary tasks. The results are shown in Fig. 3.

The results showed no significant differences between the four groups regarding the overall driving performance. However, significant differences were established for lateral control ($H = 9.294$, 3 d.f., $p = .026$.). A post-hoc Games-Howell test revealed a significantly lower number of errors for the control group when compared to the visual condition (p = .016) and for the auditory condition when compared to the visual condition (p = .047). This result corresponds with the findings of our initial study [11], in which the auditory interface caused the least lateral instability.

The results show no significant difference in the overall safety (total) of the individual interfaces. As expected, the analysis indicates that the driving simulator itself acts as a confounding variable, as it has a major influence on the overall driving performance. This influence is somewhat inevitable as the simulator allows for a semi-realistic scenario yet without causing any threats to safety. However, special care has been taken to create a high degree of realism of both the driving experience and the tasks.

5 Discussion

The analysis of task completion times shows a significant difference between the auditory interface and the other two interfaces. The auditory interface required more time and interaction activity from the participants to complete the given set of tasks. We assume this result is a consequence of the sequential nature (and hence lower information density) of sound, making a quick "scanning" of a menu for an item impossible. The transience of sound also implies that, in order to understand the meaning of the spoken words, one has to focus on the playback. In the other two conditions information was continuously projected on the windscreen until participant decided to make changes to the display. Information was presented just once in the auditory condition and the only way to re-access them was to deselect and then reselect the item again. These two factors may have contributed to task completion times, interaction activity and a higher perceived workload of the non-visual interface.

Participants' comments strengthen this assumption pointing out that for them eyes-free information presentation is inappropriate for lengthy messages, such as emails. For some it was too much of a strain to memorize information even over a limited amount of time while also attending to road events.

The visual interface had the strongest impact on lateral driving stability. Even though the display was projected onto the simulated windshield, a measurable influence could be noted. While participants in the audio-visual condition could attend to either of the modalities depending on the current situation, participants in the visual-only condition had to always focus on the visual representation of the menu which may have affected their stability within lanes.

When asked about their personal preference, the majority of participants preferred the multimodal interface (60%). The comments about this interface version were mostly positive. Only 16.7% preferred the visual interface and even fewer participants preferred the auditory interface (23.3%). Besides the benefits of the AV already mentioned, we assume that while the A and V interfaces are especially efficient with "auditory" or "visual" types of people, the multimodal AV interface, allows participants to exploit the part of the interface that best suits his or her preferences and needs.

Our data suggests that for an average, randomly selected user, the multimodal interface may not perform significantly better than a single modality interface, but it is likely to have the least disadvantages. In most aspects of the evaluation, the multimodal interface proves to have a slight (although not always significant) advantage, but it never has the strong disadvantages of the auditory interface (high task completion times) and of the visual interface (worse lateral control).

6 Conclusion

Visual interfaces based on HUDs proved to be very efficient and easy to learn. HUDs already represent a safer alternative to HDDs in many modern high-end vehicles. However, the majority of existing HUDs is used for displaying only a limited amount of information, such as navigation directions and speed or RPM. We believe their functionality could be extended to also enable advanced interaction with, for example, entertainment and communication systems.

The non-visual auditory interface proved to be safe but significantly less efficient than the HUD. Despite this fact, we see its great potential particularly in a combination with a HUD. Such a combined interface would take advantage of two non-competing human senses and enable the driver to rely on just one of them depending on the driving and traffic conditions. Although our experiment did not confirm any significant benefits of such a combination of interfaces, the majority of users responded positively to them.

The aim of our future research is to find new, efficient, safe, and enjoyable combinations of visual and auditory displays. We believe audio and visual output could be used simultaneously or interchangeably, depending on the driving conditions. We also believe that new and better alternative or supplement haptic interfaces to input devices attached to the steering wheel could be explored. Eye-gaze systems offer a big potential and could enable the on-board system to be aware of the user's current focus and eye activity. Similarly, tangible tracking has the potential to expand the driver's interaction beyond the limitations of physical buttons and knobs.

References

1. Sojourmer, R.J., Antin, J.F.: The effects of a simulated head-up display speedometer on perceptual task performance. Human Factors 32(3), 329–339 (1990)
2. Burnett, G.E.: A Road-Based Evaluation of a Head-Up Display for Presenting Navigation Information. In: Proc. of the Tenth International Conference on Human-Computer Interaction, pp. 180–184. Lawrence Erlbaum Associates (2003)
3. Yung-Ching, L.: Effects of using head-up display in automobile context on attention demand and driving performance. Displays 24(4-5), 157–165 (2003)
4. Charissis, V., Papanastasiou, S., Vlachos, G.: Comparative Study of Prototype Automotive HUD vs. HDD: Collision Avoidance Simulation and Results. In: Proc. of the Society of Automotive Engineers World Congress 2008, Detroit, Michigan, USA (2008)
5. Fisher, E., Haines, R.F., Price, T.A.: Cognitive issues in head-up displays. NASA Technical Paper 1711, NASA Ames Research Center (1980)
6. Wickens, C.D., Martin-Emerson, R., Larish, I.: Attentional tunnelling and the head-up display. In: Jensen, R.S. (ed.) Proc. of the Seventh International Symposium on Aviation Psychology, pp. 865–870. Ohio State University, Columbus (1993)
7. Tufano, D.: Automotive HUDs: The overlooked safety issues. Human Factors 39(2), 303–311 (1997)
8. Prinzel, L.J., Risser, M.: Head-up displays and attention capture. Tech. Rep. NASA/TM-2004-213000. NASA – Langley Research Center, Hampton, VA (2004)
9. Haines, R.F.: A breakdown in simultaneous information processing. In: Obrecht, G., Stark, L. (eds.) Presbyopia Research, pp. 171–175. Plenum, New York (1991)

10. Thomas, L.C., Wickens, C.D.: Visual Displays and Cognitive Tunneling: Frames of Reference Effects on Spatial Judgments and Change Detection. In: Proc. of the Human Factors and Ergonomics Society 45th Annual Meeting, pp. 336–340 (2001)
11. Sodnik, J., Dicke, C., Tomazic, S., Billinghurst, M.: A user study of auditory versus visual interfaces for use while driving. Int. J. Human-Computer Studies 66(5), 318–332 (2008)
12. Weinberg, G., Harsham, B., Medenica, Z.: Evaluating the Usability of a Head-Up Display for Selection from Choice Lists in Cars. In: Proc. of the International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2011), Salzburg, Germany (2011)
13. Hart, S.G., Wickens, C.: Workload assessment and prediction. In: Booher, H.R. (ed.) MANPRINT, an Approach to Systems Integration, pp. 257–296. Van Nostrand-Reinhold, New York (1990)
14. Dicke, C., Jakus, G., Tomazic, S., Sodnik, J.: On the Evaluation of Auditory and Head-up Displays While Driving. In: Proc. of the Fifth International Conference on Advances in Computer-Human Interactions (ACHI), Valencia, Spain, pp. 200–203 (2012)