Applied Ergonomics 46 (2015) 184–192

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: www.elsevier.com/locate/apergo

A user study of auditory, head-up and multi-modal displays in vehicles

Grega Jakus [a, *], Christina Dicke [b], Jaka Sodnik [a]

[a] Faculty of Electrical Engineering, University of Ljubljana, Trzaska 25, Ljubljana 1000, Slovenia
[b] Quality and Usability Lab, Telekom Innovation Laboratories, TU Berlin, Ernst-Reuter-Platz 7, 10587 Berlin, Germany

ARTICLE INFO

Article history:
Received 13 September 2013
Accepted 4 August 2014
Available online 21 August 2014

Keywords:
Auditory display
Head-up display
Vehicle

ABSTRACT

This paper describes a user study on the interaction with an in-vehicle information system (IVIS). The motivation for conducting this research was to investigate the subjectively and objectively measured impact of using a single- or multi-modal IVIS while driving. A hierarchical, list-based menu was presented using a windshield projection (head-up display), auditory display and a combination of both interfaces. The users were asked to navigate a vehicle in a driving simulator and simultaneously perform a set of tasks of varying complexity. The experiment showed that the interaction with visual and audio-visual head-up displays is faster and more efficient than with the audio-only display. All the interfaces had a similar impact on the overall driving performance. There was no significant difference between the visual only and audio-visual displays in terms of their efficiency and safety; however, the majority of test subjects clearly preferred to use the multi-modal interface while driving.

1. Introduction

Head-up displays (HUDs) are the current state-of-the-art solution to reducing driver errors originating from distractive interfaces, such as on-board entertainment displays or in-vehicle information systems (IVIS). Compared to head-down displays (HDDs), which are integrated into the vehicle's control panel, the frequency and duration of glances towards the display are reduced by presenting information directly on the windshield in the driver's field of vision (Ablassmeier et al., 2007). Consequently, the response time to unanticipated road events is reduced when information is displayed on a HUD instead of a HDD (Horrey and Wickens, 2004; Liu and Wen, 2004; Sojourmer and Antin, 1990), and the number of collisions is significantly reduced (Charissis et al., 2008). Using HUDs also leads to a, generally, more consistent speed control and reduced mental workload (Liu and Wen, 2004), reduced navigational errors (Burnett, 2003), and smaller variances in lateral accelerations and steering wheel angle (Yung-Ching, 2003).

However, cognitively switching between two sources of information, i.e., the HUD and traffic, still poses a problem, especially in high-workload situations. The so-called cognitive or attention capture, i.e., when the driver's attention subconsciously shifts away from the road and becomes focused on processing the information presented by the HUD, has been identified as one of the disadvantages of HUDs (Gish and Staplin, 1995; Prinzel and Risser, 2004; Tufano, 1997). The resulting perceptual tunnelling may lead to a delayed reaction or a complete absence of response to situational changes in the environment (Haines, 1991; Thomas and Wickens, 2001).

Researchers have proposed to address this issue by utilising auditory interfaces (Lai et al., 2001; Sodnik et al., 2008). They report on the greater attention capturing properties of the auditory channel and the desire to keep attention focused longer on a complex auditory task to prevent a loss of information from the driver's working memory. Wickens et al. (2005) addressed this apparent contradiction in the context of a visual on-going task and an auditory interrupting task. They found that the auditory input improved the interrupting task performance when compared to a visual interrupting task, but degraded the on-going task performance because of an abrupt shift of attention at the onset of the auditory message (see, in addition, Biever, 2002; Horrey and Wickens, 2004).

A combination of auditory, especially speech-based, information presentations and text-based information presentations through a HUD may offer the attention and safety benefits of an eyes-free approach with the higher information-processing rate of a visual information presentation. The study presented in this paper was conducted to evaluate the assumption that allowing users to freely

* Corresponding author. Tel.: +386 1 4768 938.
E-mail addresses: grega.jakus@fe.uni-lj.si (G. Jakus), christina.dicke@tu-berlin.de (C. Dicke), jaka.sodnik@fe.uni-lj.si (J. Sodnik).

http://dx.doi.org/10.1016/j.apergo.2014.08.008
0003-6870/

switch their attention between visual and auditory information presentations may have less impact on the performance of the primary task than with a single-modal information presentation.

1.1. Related work

Liu (2001) conducted a simulator-based study comparing a visual display centrally mounted above the dashboard to an audio-only (digitised human female voice) and a multi-modal (audio and visual) display of traveller information with varying complexity (icons and text). Three to five information units were displayed in the low-complexity condition and between nine and fourteen information units in the high complexity condition. Liu measured the perceived workload, navigational performance, and information detection under both high- and low-load driving conditions. Both the auditory and multi-modality displays impacted the driving task but lead to a better performance in terms of response times, total number of correct turns, and subjective workload ratings when compared to the visual-only display. Liu summarised that the driving performance was the least affected by the multi-modal display.

Sodnik et al. (2008) compared three types of interfaces for in-vehicle information systems; an HDD was compared to two auditory interfaces: one with a monophonic output playing from different spatial positions and one with multiple, simultaneously playing sounds. The results of the user study showed significant differences between the visual and auditory interfaces: the interaction with the visual interface was faster and more efficient, while the auditory interface was significantly safer and preferred by the majority of users. A comparison of the two auditory interfaces showed no advantage of the spatial over non-spatial audio in the vehicles. On the contrary, the use of simultaneous spatial sounds increased the mental workload of drivers and resulted in an increased disturbance instead of an improved interaction. The total mental workload proved to be comparable in all three conditions.

Weinberg et al. (2011a) compared the impact of interacting with combinations of a HDD and text-to-speech (TTS) system, a HUD and a TTS system, and an audio-only TTS representation of lists of choices. A menu representing three domains of content (Navigation, Music and Contacts) with various subdomains and functions was used, and participants could traverse the hierarchical menu by using a steering-wheel-mounted jog dial device or speech input. Weinberg et al. (2011b) measured task completion times, perceived cognitive loads, and personal preferences while participants followed a leading vehicle in a driving simulator (the primary, ongoing task) and solved secondary tasks using either of the IVIS display types. They also reported on the driving performance as indicated by the lateral position, speed, steering and throttle and the following distance (Weinberg et al., 2011a). Generally, the HUD and audio-only interfaces had similar impacts on the driving performance and mental workload with a slight advantage found using the audio-only version. However, the task completion times were significantly longer when using the audio-only interface. Participants rated the auditory menu as the least distracting version, while they found both the audio-only and the HUD to be the easiest to use and the most desirable.

The results of previous studies (Sodnik et al., 2008; Weinberg et al., 2011a) provide a basis for the assumption that the display modality of an IVIS has a strong impact on both the primary and secondary tasks in a vehicle. Prior findings also support the assumption that a combination of auditory, especially speech-based information presentations and text-based information presentations through a HUD, may be the "best of both worlds" – the attention and safety benefits of an eyes-free approach with the higher information processing rate of a visual information presentation. In the following study, we compare two single-modality interfaces, an audio-only and a video-only interface, to a multi-modal HUD (Weinberg et al. (2011a) compared an audio-only display to two different multi-modal displays). We asked participants to perform a series of tasks while driving. In contrast with previous studies, our tasks varied in complexity between a simple "search for and find a piece of information" task and a more complex "identify and interpret" task.

The experimental design, including the technical setup, is described in the following section. This is followed by an analysis of the gathered data and an interpretation and discussion of the findings. The paper concludes with recommendations for the design of list-based head-up displays to access multi-modal in-vehicle information and entertainment systems.

2. Methods

2.1. User study design

Three different IVIS interfaces were evaluated in regards to conducting typical (secondary) tasks while driving a simulated vehicle in different traffic conditions ("high-speed" and "low-speed" traffic was randomly distributed among the experimental conditions). Three different experimental conditions were compared based on these three different display techniques:

- a visual display/HUD (V),
- an auditory display (A) and
- a multi-modal auditory and visual display/HUD (AV).

The research questions addressed in this study are as follows:

- Which interface is more efficient in terms of task completion time and interaction activity?
- Which interface is less detrimental to driving performance and why (measured by driving behaviour)?
- How do the different interfaces affect the user experience (measured by NASA TLX workload ratings, user experience (UEQ) and personal ratings collected in a post-hoc questionnaire and an informal interview)?

2.2. Design of the in-vehicle information system (IVIS)

All modern in-vehicle computer systems combine numerous functionalities related to navigation (e.g., traffic reports, navigation assistance), entertainment (e.g., audio, video, communication) and vehicle control (e.g., air conditioning, system information, cruise control). The IVIS used in the experiment simulated the majority of these common features and was accessible through a hierarchical menu structure. The top-most level of the structure was called the "Main menu". It consisted of five main categories that were further structured into sub-menus of various depths (cf. Fig. 1). Each level of the individual sub-menu consisted of up to eight options. At each level, the user could freely navigate between all the available options, select one of them and enter the corresponding sub-menu or exit and return to the previous menu. Navigation at a given level was "non-circular", with a virtual barrier at both ends of the menu. Therefore, when users reached the first or the last item of a menu, they could only continue by moving in the opposite direction.

2.3. The visual display

The visual display (V) used in the experiment is depicted in Fig. 2. The recommendations given in (Tretten et al., 2011; Wickens