White Paper

Can voice interaction help reducing the level of distraction and prevent accidents?

Meta-Study on Driver Distraction and Voice Interaction

19 May 2011

Authors

Matthias Peissner, Vanessa Doebler
Fraunhofer-Institute for Industrial Engineering (IAO), Germany

Florian Metze
Carnegie Mellon University, USA

Study sponsored by

NUANCE COMMUNICATIONS Aachen GmbH, Germany

Abstract

This paper reviews the current state of knowledge in the field of driver distraction, its causes and correlates, its effects on driving performance and accident hazard, and the role of voice interaction for a better driving safety. It summarizes the results of an extensive literature review in the field of driver distraction and voice interaction and presents an overview of relevant findings from 35 selected publications between 2000 and 2011.

Driver distraction caused by interacting with technology while driving is a predominant and increasingly relevant accident hazard. The interaction with ICT products can be associated with cognitive and visual driver distraction. The actual level of driver distraction depends on characteristics of the secondary task and the used interaction mode. Cognitive distraction is mainly determined by interaction complexity whereas visual distraction can be avoided by using voice interaction instead of manually controlled systems with a graphical display. Compared to cognitive distraction, visual distraction has much stronger effects on driving performance and accident hazard. As a consequence, voice interaction is consistently shown to be superior in terms of driver distraction, driving performance and accident hazard when compared with manual control and graphical displays. Therefore, voice interaction is often referred to as an important and indispensable contribution to safer driving. However, the design and implementation of a speech-based system play an essential role for the effectiveness and safety of voice interaction in the car. An accurate speech recognizer and an easy-to-use voice user interface are the main prerequisites in order to make full use of the potentials of voice interaction.

In summary, well-designed speech-based in-vehicle systems will allow drivers to keep their eyes on the road and their hands on the wheel, making it possible to improve driving performance and reduce accident hazard.

Content

1 Introduction

In-vehicle-systems are becoming increasingly popular and extensively used. Functionalities such as navigation, air-conditioning, electric windows and seat adjustment, mobile telecommunication, DVD and multimedia content, etc. are becoming part of the daily driving experience. The increased complexity of control panels and menu structures of today's in-car information systems diverts the driver's attention away from his primary driving task and puts driving safety at risk.

Recent achievements in the field of automatic speech recognition seem to promise significant advances for driving safety. Together with acoustic system output, voice interaction might help to keep the driver's full visual attention on the road and manual attention at the steering wheel while using modern information and communication services.

The study at hand provides an overview of recent research results from scientific publications. The study presents an overview of relevant findings from 35 selected publications between 2000 and 2011. It covers the following main issues:

– How does interaction with interaction and communication systems (including in car systems and mobile devices) influence driver attention and performance?

– Which factors of a user interaction task cause driver distraction?

– Are there different types and levels of driver distraction depending on specific task characteristics?

– How do different types and levels of driver distraction influence driver performance and accident hazard?

– How does voice interaction compare to other interaction models in terms of driver distraction, driver performance and accident hazard?

– Is the design of ICT systems important for the level of driver distraction?

– Can voice interaction help to reduce the level of distraction and prevent accidents?

1.1 Statistics and facts

According to the National Highway Traffic Safety Administration (NHTSA, 2010) in the United States, 16 percent of driving fatalities were attributed to driver distractions in 2008 and 2009. In the year 2009 more than 5,000 deaths and almost half a million injuries occurred due distracted drivers. The NHTSA expects it to get worse. Just five years ago, in 2005, only 10 percent of driving deaths in the USA were attributed to distractions.

– During a phone conversation, drivers are four times as likely to get into crashes, serious enough to injure themselves. (Insurance Institute for Highway Safety, 2010)

- National Safety Council research indicates that cell phone use and texting while driving cause at least 28 percent of all traffic accidents in the USA - around 1.6 million accidents and 645, 000 injuries in 2008. (National Safety Council, 2010)
- Between 15%-20% of all distractions appear to involve driver interaction with technology. (Regan, Lee & Young, 2010)
- Texting has the highest risk to be involved in crash or near crashes of all cell phone related tasks. (Virginia Tech Transportation Institute, 2009)

These statistics from the United States indicate that phone conversation and technology interaction during driving have serious effects on distraction and accident hazard. However, as we will see in the following, not conversation and interaction per se are the main distractors but shifting manual and especially visual attention from the driving task to an interactive device.

1.2 What drives distraction?

Often drivers do not realize that they are distracted and that they switch their focus away from the road. Driver distraction can be interpreted differently. The following three definitions clarify the danger of being distracted while driving:

- »Driver Distraction can be defined as the diversion of attention away from activities critical for safe driving toward a competing activity.« (Regan, Lee & Young, 2009)
- The International Standards Organization defined distraction as »attention given to a non-driving-related activity, typically to the detriment of driving performance.« (ISO TC22/SC13/WG8 CD 16673)
- The AAA Foundation for Traffic Safety describes distraction as »when a driver is delayed in the recognition of information needed to safely accomplish the driving task because some event, activity, object, or person within or outside the vehicle compelled or tended to induce the driver's shifting attention away from the driving task.« (AAA Foundation for Traffic Safety, 2001)

Regan et al. (2009) state that the human brain is limited, and not able to do several tasks at the same time. Especially, when tasks are similar, highly demanding, and when they require continuous attention the performance will inevitably suffer. There are many potential sources of distraction in a vehicle: Talking to passengers, reaching for objects, listening to the radio, smoking, eating, daydreaming, applying cosmetics, and attending to potential sources of distraction outside the vehicle cockpit (e.g. advertising). Considering the increasing number of modern in-vehicle technologies, this paper will focus the impact of distraction due to driver interaction with entertainment, information and communication systems instead of everyday activities such as the above mentioned.

1.3 Primary and secondary task while driving

For a more comprehensive review of driver distraction, two types of driving tasks can be classified: the *primary task* and the *secondary task*.

– The *primary task* is defined as the actual driving task, keeping the vehicle on the road while obeying the traffic regulations and being thoughtful towards other traffic participants. The primary task includes physical actions such as braking, depressing the accelerator, operating the transmission, controlling the speed, and steering the vehicle. Primary tasks are often performed out of a habit due to the experience of driving.

– *Secondary tasks* are tasks, which are not part of the natural driving response, but function to please the comfort- and entertainment needs in a car just as selecting music from a hand-held or hands-free music player, receiving and indicating a call, entering some data to the navigation system, or regulating the air conditioning. Secondary tasks might divert the driver's attention away from the driving task.

Some researchers (Olson, Hanowski, Hickman & Bocanegra 2009; and Kern & Schmidt, 2009) define also *tertiary tasks* whereby the tertiary tasks are non-driving related tasks (equally to the above defined secondary tasks), and secondary task are defined as driving related tasks, but not required tasks for controlling the vehicle. In line with most published studies, we distinguish between primary and secondary tasks, without considering tertiary tasks.

1.4 Attention

Attention, especially attention bottlenecks play an important role for the concept of driver distraction, as we have seen in the definitions above. While driving the driver's attention will be split up between the primary task of driving and other secondary tasks, such as talking, listening to the radio, or interacting with entertainment or driver assistance systems.

Wickens and Hollands (2000) describe three categories of attention problems:

– *Selective attention*
Attention is focused towards one aspect in the environment whereas other aspects are not attended. (»cognitive tunnelling«).

– *Focused attention*
Tendency to be distracted by irrelevant information.

– *Divided attention*
Limited ability of attending to two or more concurrent tasks.

All these attention bottlenecks can apply to driving situations, e.g. »cognitive tunnelling« applies when a driver concentrates on entering an address into the navigation system and loses control of the primary task of driving. In general, whenever people are involved in more than one complex or cognitively demanding task in parallel, the human brain is in danger of missing critical information.

Gopher and Donchin's Multiple Resource Model (Gopher & Donchin, 1985) provides an explanation for different levels of interferences between tasks which are performed in parallel (see figure 1). The Multiple Resource Model describes three different pools of attention resources and information processing capacities:

1 Input and output modalities (visual vs. auditory),

2 stages of information processing (perception vs. working memory) and

3 codes of information processing (spatial/manual codes vs. verbal/vocal codes).

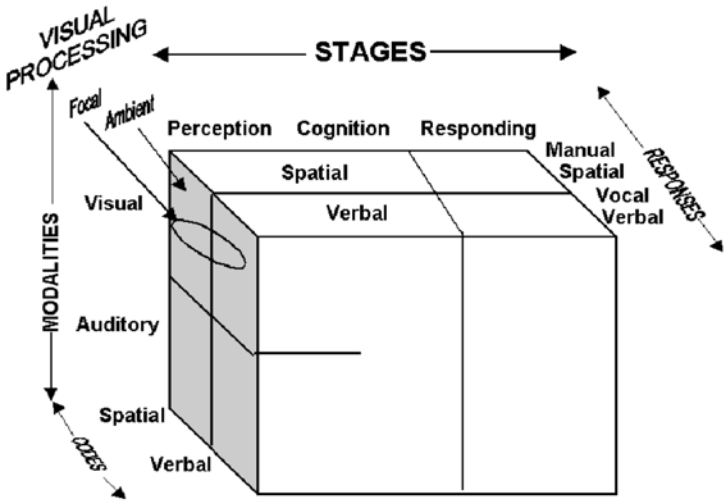

Figure 1 Multiple Resource Model (Gopher & Donchin, 1985)

Tasks which require resources from the same pool will interfere. For instance, operating a hand-held device will interfere with the primary task of driving because both tasks rely mainly on visual information and spatial/manual codes for perception and responses. According to the Multiple Resource Model, Maciej and Vollrath (2009) argue that speech-based interaction will be much safer while driving, since speech interaction requires different resources than driving.

2 Causes of driver distraction

Related to the use of ICT systems while driving, three questions frequently arise:

1 How does interaction with ICT systems (including in car systems and mobile devices) influence driver attention and performance?

2 Which factors of a user interaction task cause driver distraction?

3 Are there different types and levels of driver distraction depending upon specific task characteristics?

2.1 Influences on driver attention and performance

According to relevant research studies, we can conclude that interacting with ICT systems significantly decreases driver attention and performance.

In order to understand how and why interacting with modern ICT systems influences driver attention we first reconsider »cognitive tunnelling« as introduced in section 1.2. Using a cell phone leads to inattention and narrows the driver's scope – even if a hands-free device is used. This is also called »inattention blindness«. Drivers are looking out the windshield, but do not process everything in the roadway environment, they miss critical information on potential hazard in their surroundings and thus, they are not able to respond to unexpected situations. Their field of view narrows. In order to explore how the use of a cell phone may affect a driver's visual scanning, Harbluk, Noy and Eizenman (2002) tracked the eye movements of drivers while using hands-free phones, and while not using a phone. The figure below provides a comparison of the driver's fields of view in both conditions. The driver's field of view is significantly scaled down while being distracted by using a hands-free device.

Figure 2 Drivers' field of view while not using a hands-free device (left) and while using a hands-free device (right) (Image Source: National Safety Council, 2010)

The influences of using ICT systems on driving performance can clearly be demonstrated by measuring longer reaction and response times, and by detecting problems with lane keeping and variations in following distances. Often large variances in lane position are considered as the most serious sign of influences on driving performance when using an ICT system while driving. Therefore, Kun, Peak and Medenica (2007) recorded three measures of driving performance when investigating the effect of different accuracy levels of speech recognition: the lane position, the steering wheel angle, and the velocity of the participants when operating with different systems. They found significantly reduced driving performance for lower speech recognition accuracy levels. They used a high-fidelity driving simulator with a 180° field of view. 20 participants needed to follow a leading vehicle at a constant distance without departing from the lane.

A very well established experimental paradigm for assessing driver distraction is the Lane Change Task (LCT) as described by Mattes (Mattes, 2003). In a LCT simulation, a driver has to follow a straight three-lane road for about three minutes at a constant maximum speed of 60 km/h. During one trial, 18 signs along the track indicate that the driver has to change the lane as soon as possible. On the basis of measuring the longitudinal and lateral position, the speed, and the steering angle, the deviation of the actual driving performance from a normative, ideal model is calculated and serves as an indicator for the distraction caused by a secondary task.

2.2 The main factors of driver distraction

The main factors of driving distraction are *Head-up*, *Hand-off* and *Mind-off* operations, which can be mapped to visual, manual and cognitive attention bottlenecks:

– *Visual*
  visual tasks require eyes-off-the-road

– *Manual*
  manual tasks require hands-off-the-wheel

– *Cognitive*
  cognitive tasks require mind-off-the-road

Modern in-vehicle entertainment systems encompass diverse user interactions including map interactions, text entry, item selection, and reading while driving. While all these interactions can endanger the drivers' safety, Virginia Tech Transportation Institute found that texting is the most serious one because it involves all three above types of distraction. Firstly, the driver is looking away from the road and the glance is bestowed on the display of the cell phone. Secondly, the driver uses at least one hand to write and send the text message. And thirdly reading or writing is likely to involve emotional feelings or cognitive operations. Thus, it is not surprising that texting is associated with most frequent and longest lasting looking away from the road (cf. Olson, Hanowski, Hickman & Bocanegra, 2009 and Box, 2009).

2.2.1 Different types and levels of driver distraction

In order to analyse different types and levels of driver distraction, quite a number of observational studies have been conducted (e.g. Jensen, Skov, & Thiruravichandran, 2010; Barón & Green, 2006; Chang, Lien, Lathrop, & Hess, 2009).

Maciej and Vollrath (2009) show a comparison of manual and speech-based interfaces while driving and operating different in-vehicle information systems (IVIS). A driving simulator was used to investigate 30 drivers doing the Lane Change Task. Different secondary tasks such as selecting an audio artist, album or title, placing a phone call and entering points-of-interests (POI) or a specific address into the navigation system were performed either manually or by voice interaction. The address entry tasks was performed in two different conditions: stepwise affirmation of the entered address information (»multiple«) and collected affirmation of all entered address information at the end of the dialogue within one single step (»single«). The single entry navigation system was used with the speech dialogue only. Figure 3 shows the mean (bar) and standard deviation (whiskers) of the subjective distraction for the different IVIS comparing manual vs. speech control.

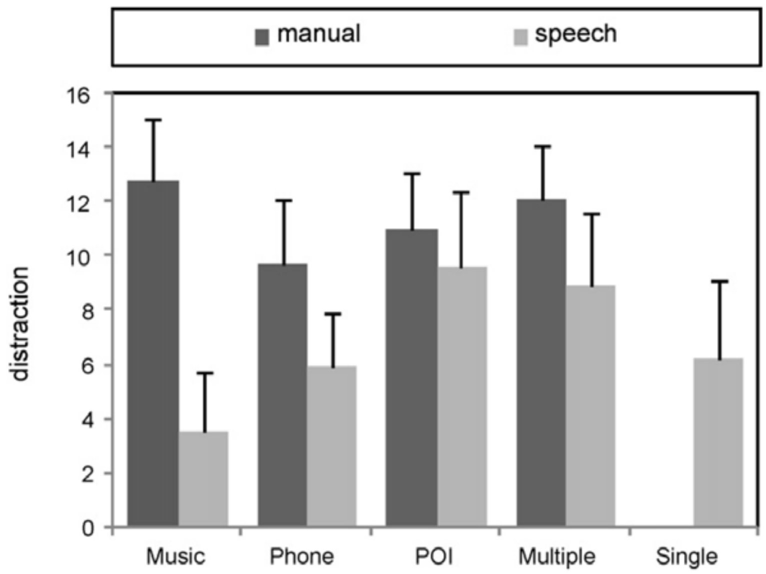

Figure 3 Subjective distraction (means – bars, standard deviations – whiskers) of different IVIS tasks comparing manual and speech control (Source: Maciej & Vollrath, 2009)

The results indicate two main findings: Firstly, manual control resulted in a substantial increase of perceived distraction in comparison with voice interaction during all tasks. Secondly, the level of decrease in subjective distraction differs a lot between the different tasks. With regards to the driving performance measures, there was even one task (POI entry) for which speech interaction did not yield a significant improvement compared to the manual control conditions. These results demonstrate that there are different types and levels of driver distraction depending upon specific characteristics of the secondary task and the used interaction mode.

2.2.2 Visual distraction

Frequency and duration of in-vehicle glances to infotainment systems are important influences with regards to visual distraction. The higher the frequency of glances off the road the bigger the threat to be involved in a danger situation while driving. Several authors point out that visual distraction of more than two seconds is considered to be a critical time of visual absence (cf. Zhang, Smith & Witt, 2006; Klauer, Dingus, Neale, Sudweeks & Ramsey, 2006; and Olson, Hanowski, Hickman & Bocanegra, 2009).

Hosking, Young and Regan (2009) examined the impact of text messaging on the mean frequency of in-vehicle glances as well as on the mean duration of in-vehicle glances. In this experiment, 20 young novice drivers were tested in an advanced driving simulator at the Monash University Accident Research Centre. Figure 4 shows that text messaging results in more and longer in-vehicle glances than driving in respective control conditions without text messaging. Both, retrieving and sending text messages negatively affect the driving performance. The driver’s eyes were focusing significantly less on the road during the activity of messaging compared to the control condition (no text messaging).

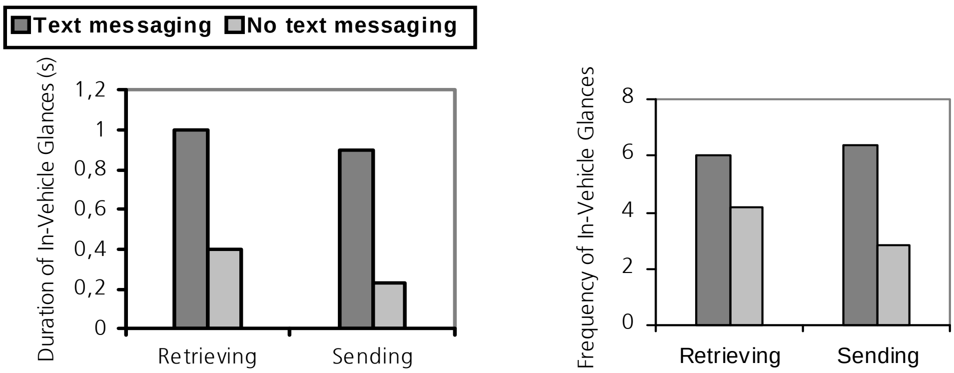

Figure 4 Mean durations and frequencies of in-vehicle glances while text messaging (Source: Hosking, Young & Regan, 2009)

In their literature review, Barón and Green (2006) state that speech recognition interfaces reduce driver distraction compared to traditional manual interfaces. The drivers do not need to attend to a graphical interface in order to select any functions. Rather, voice interaction allows the driver to visually concentrate on the driving task.

2.2.3 Cognitive distraction

»Cognition« refers to the mental functions of human understanding and information processing such as perception, learning, thinking, and remembering.

With regard to the cognitive workload, Green (2000) found out that also tasks that are not visually demanding, such as daydreaming or listening to a long complex auditory message from a phone call can increase the probability of crashes.

McCallum, Campbell, Richman, Brown and Wiese (2004) investigated the cognitive workload when using a personal digital assistant (PDA) while driving. Two different conditions (Speech PDA, and Manual PDA) and a control condition (No PDA) were compared. 24 participants were tested in a stationary vehicle including a display with the driving environment, simulation control and data collection modules. For subjective ratings a modified NASA TLX rating scale was used[1]. The results indicate that while driving, manual PDA operations produce a significantly higher cognitive workload than speech-based operations (cf. Figure 5).

[1] The NASA Task Load Index is a subjective workload assessment tool. It is used world-wide to evaluate workload in human-machine systems; cf. http://humansystems.arc.nasa.gov/groups/TLX/

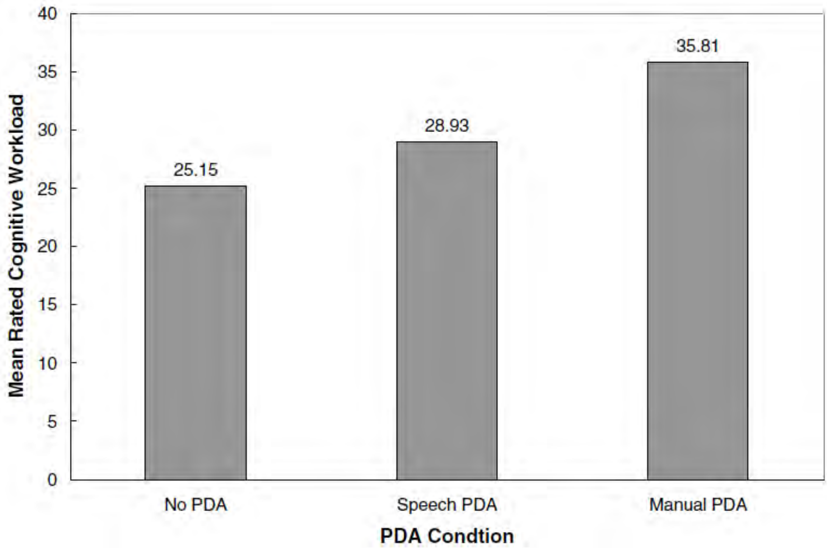

Figure 5 Mean rated subjective cognitive workload levels for different PDA interaction modes while driving (Source: McCallum, Campbell, Richman, Brown & Wiese, 2004)

When considering the advantages of voice interaction, one has to distinguish different types: voice interaction between a human and a computer and human-to-human voice interaction. In the latter, a further distinction between speaking on the phone and talking to a passenger in the car is essential. A passenger in a car is very likely to estimate how much talk in various traffic situations will be suitable. However, a person on the phone has no knowledge about the current situation, and thus, cannot respond appropriately to different traffic situations. Commonly, passengers avoid conversations about complex matters, which could lead to driver distraction whereas people on the phone do not mind (Green, 2000). Future in-car systems should provide similar intelligent mechanisms and adapt their interaction behaviour to the current traffic and the driver's workload in the primary task.

3 Effects on driver performance and accident hazard

Different types and levels of driver distraction will influence driver performance and accident hazard in different manners. The effect on driving performance depends on several interrelated factors: the nature of the activity, the ability and experience of the user, the complexity of the driving task and the design, location, and activities with in-vehicle technologies. Therefore, not only certain task characteristics but also the design of an in-vehicle information system can have a considerable effect on driver performance and crash risk. Driving performance is usually defined in terms of the following three measures:

– Longitudinal control:
  Speed and following distance

- Lateral control:
  Lane keeping and steering measures
- Reaction Time

The following sections will present and review findings on different effects on these driving performance indicators and the resulting risks of crashes or near-crashes.

3.1 Longitudinal and lateral control

Longitudinal control measures are commonly examined in distraction research. Two of the most commonly used longitudinal control measures are speed and following distance. Lateral control is usually measured by lane keeping and steering.

Hosking, Young & Regan (2009) investigated the impact of text messaging on driving performance by taking longitudinal control measures. The driver's ability to control the longitudinal vehicle position was significantly impaired during text messaging activities. The speed of the distracted driver did not differ from the driving speed in the control condition, however, the following distance increased. Hosking interpreted the increased following distance as the drivers attempting to compensate for being distracted during driving by reducing the speed.

A study on the effects of naturalistic cell phone conversations on driving performance published by Rakauskas, Gugerty and Ward (2003) yields similar results on distracted driver behaviour. In a driving simulator experiment, they found out that cell phone use causes drivers to have higher variation in accelerator pedal position, drive more slowly with more variation in speed, and report a higher level of workload.

Gärtner, König and Wittig (2002) investigated the influence of manual and speech input on different driving performance indicators.15 subjects were tested in a real traffic situation when performing different tasks using a Driver Information System (DIS). The driving errors were classified into eight classes from A to H. Figure 6 shows that »poor lane keeping« (lateral control) and »speed too low« (longitudinal control) are the major errors in distracted driving performance. Both types of driving errors, occurred significantly more frequently in the manual input condition than with speech input.

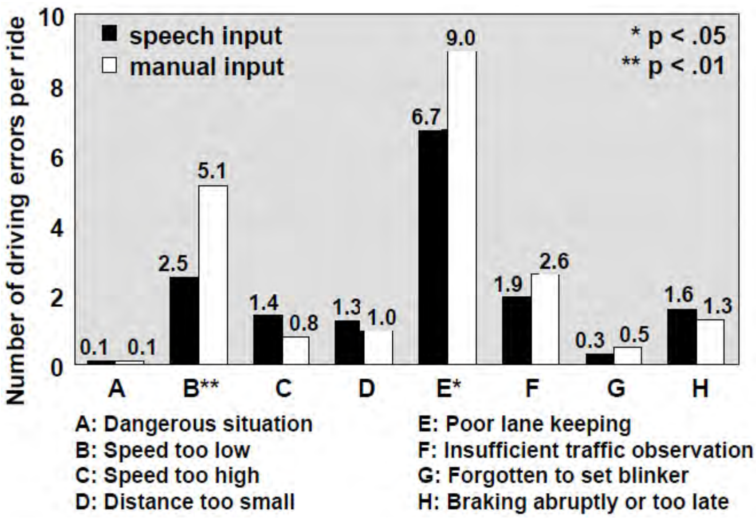

Figure 6 Driving errors for different error classes (Source: Gärtner, König & Wittig, 2002)

The effects on longitudinal and lateral control performance were also investigated by Jensen, Skov, and Thiruravichandran (2010). They inspected lateral control errors of three different output configurations (only audio, only visual and audio-visual) of a navigation system in real traffic driving. The measured variables include primary task performance (longitudinal control, lateral control and traffic violations) and secondary task performance (navigation errors and task completion time). The experiment was conducted in real traffic with 30 participants. The results showed that the participants had many more longitudinal control errors (i.e. speeding violations) in the visual and audio-visual condition than with the audio configuration. With regards to lateral control errors, the participants performed significantly better with the audio configuration than with visual or audio-visual output. The visual and audio-visual configurations collectively constituted 95% of all lateral control errors, whereas participants had only four out of a total of 77 lateral control errors.

In their 2003 literature review, Young, Regan and Hammer (2003) report that visually distracted drivers steer their car in a different way than attentive drivers do. Steering wheel movements are considered to be an indicator of secondary task load. Under normal conditions, when not performing secondary tasks, drivers usually do a number of small corrective steering wheel movements to keep the lateral position. When they perform a visually or manually demanding secondary task, drivers often make a number of large and abrupt steering wheel movements to correct heading errors (cf. Regan, Lee, & Young, 2009).

Liang and Lee (2010) compared the effects of different types of distraction on driving performance in a medium-fidelity simulator study: visual distraction, cognitive distraction and a combination of both. They found out that visual distraction interferes with driving performance more than cognitive distraction, and visual distraction dominates the performance decrements during combined distraction (see Figure 7). They conclude that minimizing visual demand is particularly important in the design of in-vehicle systems and in the development of distraction countermeasures.

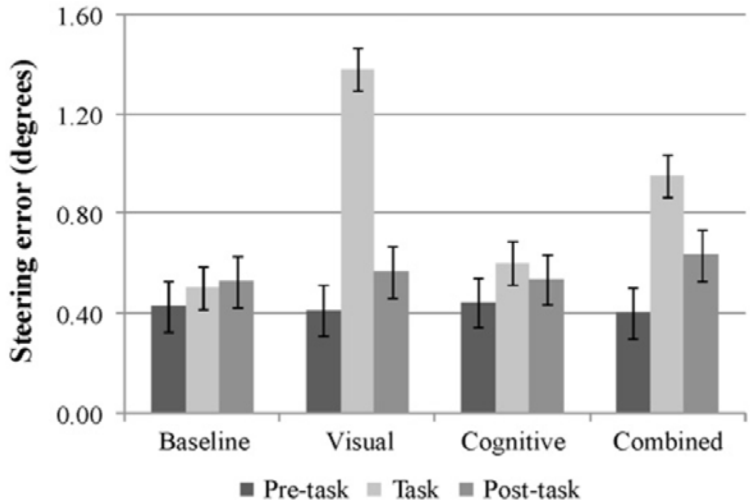

Figure 7 Lateral control performance under different driver distraction conditions
(Source: Liang & Lee, 2010)

3.2 Reaction time

Reaction time measures have become increasingly popular in in-vehicle system research, particularly when the used devices are complex (cf. Regan, Lee, & Young, 2009). Reaction time is regarded as an essential driving performance indicator as increases in reaction time can decisively raise the accident risk in unexpected and hazardous driving situations.

In a driving simulator study, Maciej and Vollrath (2009) found out that both manual and speech control in secondary IVIS interaction tasks led to significant increases in reaction times. Even in the least demanding experimental condition which resulted in the shortest response (music selection via speech control as a secondary task), the average reaction time was still significantly longer than in the baseline condition (driving without performing a secondary tasks). Figure 7 shows the means (bar) and standard deviations (whiskers) of the reaction times for the different in-vehicle-information-systems (IVIS) comparing manual vs. speech control. The baseline condition is shown as a black bar in the first column.

Maciej and Vollrath (2009) point to similar results from Chisholm, Caird and Lockhart (2008) who found that selecting music from an MP3 player significantly increased the reaction time to unexpected hazards in the traffic. Studies with speech controlled in-car systems indicate that also voice interaction can lead to increased reaction times (cf. Lee, Caven, Haake & Brown, 2001 or Jameson, Westerman, Hockey, & Carsten, 2004).

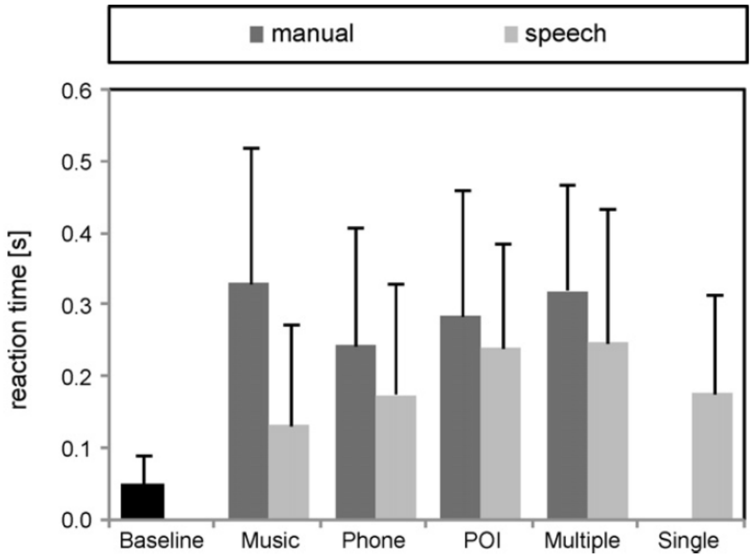

Figure 8 Reaction times in the Lane Change Test while performing different tasks on in-vehicle information systems with speech control and manual control (Source: Maciej & Vollrath, 2009)

3.3 Accident hazard

Not surprisingly, the above reported decreases in driving performance seem to have an effect also on the risk of being involved in an accident. Studies done by the Virginia Tech Transportation Institute (VTTI) which included light vehicle drivers and truck drivers indicate that using a cell phone can substantially increase the risk of safety-critical events such as crashes or near-crashes (cf. Box, 2009). An overview of different tasks and their effect on accident hazard is provided in figure 8 below.

CELL PHONE TASK	*Risk of Crash or Near Crash event*
Light Vehicle/Cars	
Dialing Cell Phone	2.8 times as high as non-distracted driving
Talking/Listening to Cell Phone	1.3 times as high as non-distracted driving
Reaching for object (i.e. electronic device and other)	1.4 times as high as non-distracted driving
Heavy Vehicles/Trucks	
Dialing Cell phone	5.9 times as high as non-distracted driving
Talking/Listening to Cell Phone	1.0 times as high as non-distracted driving
Use/Reach for electronic device	6.7 times as high as non-distracted driving
Text messaging	23.2 times as high as non-distracted driving

Figure 9 Cell phone use and risk of crash (Source: Box, 2009)

A main result of their studies is that manual manipulations of phones lead to massive risk increases whereas talking or listening seem to have only slight or no effects. Box (2009) explains this finding with eye glance analyses which showed that the tasks with the highest risks had the longest duration of eyes off road time. In summary, visual distraction seems to be a major accident risk factor which is by far more important than cognitive distraction.

3.4 Excursus: Validity of driving simulator studies

In the above mentioned paper, Box (2009) points out that the dangerous effects of talking and listening on the phone are often overestimated in driving simulator studies. Thus, the results of VTTI's naturalistic studies question the validity of driving simulator studies in general. The majority of research studies in the field of driver distraction, driving performance and accident hazard have been conducted in simulator settings. But how well do these findings apply to realistic driving situations? It seems that this question cannot be answered in a simple and general way. The two publications summarized in the following section suggest quite high levels of external validity for simulator studies.

Wang and colleagues (2010) found high consistencies in central distraction and performance indicators between studies in a driving simulation environment and in the field. They compared usability and safety implications of three manual address entry methods in an on-road study and in a medium fidelity fixed-base driving simulator. Measures of glance frequency, total glance duration, percentage of time eyes on the road, and mean task time were almost identical for simulation and field experiments. Rank ordering of the effects of the three input methods was consistent for each in both environments. The validity of central driving performance measures such as mean and deviation of forward velocity, standard deviation of lane position could not be assessed to the same extent because the differences between the three input conditions did not significantly differ from each other and because the distraction conditions had to be kept at a level which would not place too much accident risk on the driving participants.

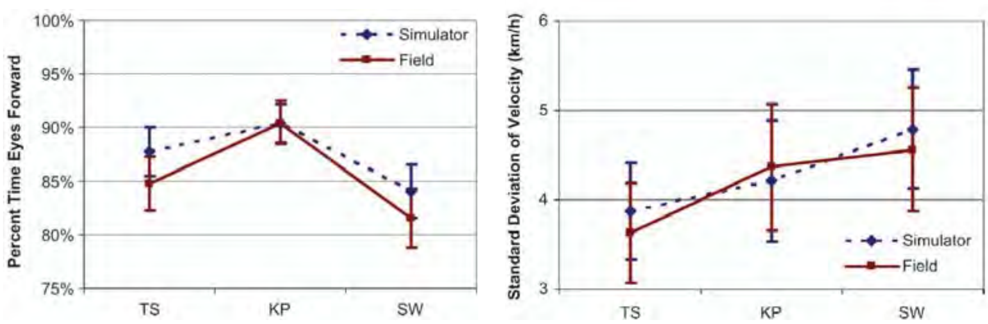

Figure 10 Comparison of visual attention (left: relative time drivers' eyes were directed forward toward roadway) and driving behaviour (right: standard deviation of forward velocity) for three device entry methods: TS – touch screen, KP – keypad, and SW – scroll wheel (Source: Wang, Mehler, Reimer, Lammers, D'Ambrosio & Coughlin, 2010)

Also Engstroem, Johansson and Oetslund (2005) report generally consistent results in field and simulator studies. However, an apparent difference was a higher result in the physiological workload and steering activity during the field study. The authors interpret this difference, as an indication of increased effort, which seems reasonable given the higher actual risk in real traffic.

4 Voice Interaction

On the basis of the above reported studies, voice interaction can be considered as a safer means to perform secondary tasks while driving. The presented publications on distraction (section 2), driver performance and accident hazard (section 3) clearly indicate that voice interaction provides many advantages for in-car interaction when compared to other currently available control methods. This section summarizes the most important findings on voice interaction in the car from the reviewed literature. It focuses on the following two main questions:

– How does voice interaction compare to other interaction modes in terms of driver distraction, driver performance and accident hazard influence driver attention and performance?

– Can voice interaction help to reduce the level of distraction and prevent accidents?

4.1 Voice interaction in comparison with other interaction modes

As pointed out in section 1.4, attention theory suggests that voice interaction is less distracting than interactions with a visual display because the driving task is primarily visual (cf. Wickens, 1984). This means that drivers can better divide attention across the visual and acoustic modalities than intra-modally. Numerous empirical studies confirm this assumption:

Jensen, Skov, and Thiruravichandran's study (2010) provides strong evidence for the advantage of speech output compared with visual output (see also section 3.1). In a real traffic driving study, they report significantly better driving performance when the participants receive audio instructions from a navigation system than with visual instructions. The substantial increase in longitudinal (i.e. speed violations) and lateral control errors (i.e. lane excursions) in the visual display condition can be explained by an increased frequency and duration of eyes-off-the-road glances for the visual and the audio-visual experimental conditions.

Barón and Green (2006) provide a review of fifteen experiments to investigate the use of speech interfaces for typical in-vehicle tasks like music selection, email processing, dialling and destination entry. They summarize that speech interaction leads to better driving performance, decreased workload and less time with eyes off the road as opposed to manual interfaces. Most tasks can usually be performed better with speech. An exception is dialling which is usually better in manual control mode. A crucial factor is the design and implementation of the user interface - especially the speech recognition accuracy. Finally, the current driving situation and also driver characteristics (especially driver's age) play an important role for the suitability of speech vs. manual interaction modes.

As described above, Maciej and Vollrath (2009) found substantial safety improvements for speech interaction in their comparison of manual and speech-based in-vehicle interfaces. Speech interaction led to significant improvements of the driving performance (measured via lane-keeping and reaction times) and in the visual attention (measured via gaze time towards the driving task) as well as to a significant reduction of the subjective distraction (rating scales) in their 30 participants lane change experiment. Given the great potentials for increased driver safety, the authors conclude that "speech control is a must in the car of the future".

McCallum et al. (2004) found significantly decreased cognitive workload levels when their participants interacted with a speech-based PDA compared to a manual PDA (cf. section 2.2.3 Cognitive distraction).

Gärtner et al. (2002) report significantly less driving errors for the speech input condition than for manual interaction with a driver information system in realistic driving situations. Similar results from realistic driving situations were found by Jensen et al (2010) who compared audio vs. visual vs. audio-visual output modality conditions for a navigation system (cf. section 3.1 Longitudinal and lateral control).

Castronovo, Mahr, Pentcheva and Müller (2010) examined the primary driving task performance with three different systems including manual, speech-only and multimodal interaction. 24 participants carried out the lane change task. Figure 11 shows that the distraction in the manual condition is significantly higher than in speech-only and multimodal conditions. Although the drivers were able to perform more of the experimental tasks in the manual condition, their driving was significantly safer with using the speech-only or multimodal dialogue.

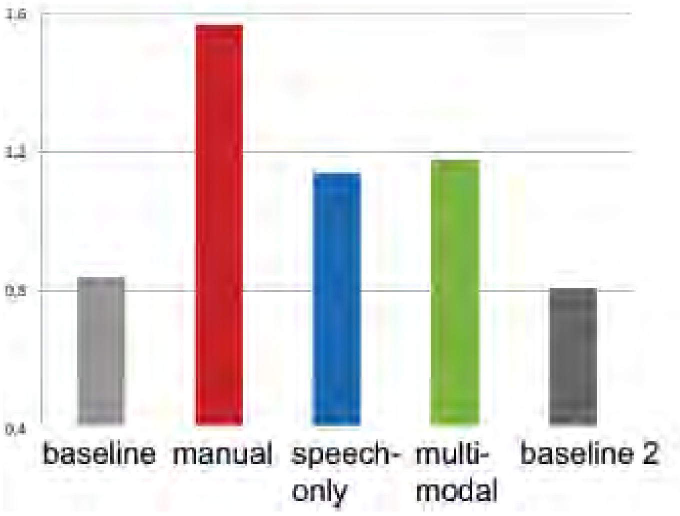

Figure 11 Driving performance measured by mean deviation in meters between a normative model and the actual driving: Distraction in the manual condition was significantly higher than both in speech-only and multimodal (Source: Castronovo, Mahr, Pentcheva & Mueller, 2010).

The Virginia Tech Transportation Institute (Ford Motor Company, 2009) studied the use of a speech based system "SYNC" developed by the Ford Motor Company. This study shows that speech-based interaction enables a quicker task completion and less eyes-off-road time as compared to using a hand-held device. VTTI's real-world study tracked 109 drivers collecting over 2 million miles driven within one year. The "100-Car-Study" concluded that manually dialling, a task that requires a person to have the eyes off the road, was almost 2.8 times riskier than normal driving. VTTI also explains that "'Headset' cell phone use is not substantially safer than 'hand-held' use because the primary risk is associated with both tasks is answering, dialling, and other tasks that require your eyes to be off the road. In contrast, 'true hands-free' phone use, such as voice activated systems, are less risky if they are designed well enough so the driver does not have to take their eyes off the road often or for long periods." (Box, 2009).

4.2 Success criteria for effective voice interaction solutions

Most above cited authors point out that the quality of the design and implementation of the voice user interface have a substantial influence on the potentials for increasing the driver safety by using voice interaction instead of visual-manual interaction methods.

A first crucial quality factor of speech-based in-vehicle systems is the performance of the automatic speech recognition (ASR). Kun, Peak and Medenica (2007) compared the effects of high and low ASR accuracy levels on driving performance in a high-fidelity driving simulator. They measured the steering wheel angle variance of their 20 participants as a main indicator for lateral control performance. They found highly significant effects of the recognition accuracy (see Figure 12).

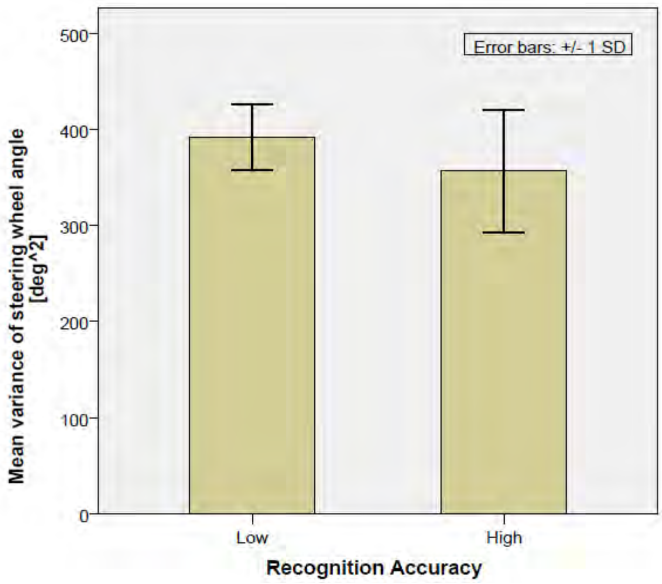

Figure 12 Steering wheel angle variance is affected by different levels of speech recognition accuracy (Source: Kun, Peak & Medenica, 2007)

Zhang and Wei (2010) emphasize the importance of the voice user interface design for keeping the complexity of today's in-vehicle applications at a manageable level in order to reduce the cognitive workload and contribute to safer driving. On the basis of a literature review and a case study with an existing speech recognition application they propose guidelines for the design of ergonomic in-vehicle speech recognition interfaces.

Finally, the type and content of an application heavily influences the suitability of using voice interaction. The nature of voice interaction is different from graphical user interfaces. For example, voice user interfaces are restricted to sequential input and output whereas graphical displays can convey a lot of information in parallel on one screen. On the other hand, voice input can avoid tedious menu trees or interaction processes which require a number of subsequent steps if the speech-based system can process more complex user utterances or if the user can just say what he wants. Although voice interaction will therefore not be the best modality for all tasks (cf. also Castronovo et al., 2010), it provides rich opportunities for innovative services in the car – also for the sake of driving safety. Alvarez, Martin, Dunbar, Taiber, Wilson and Gilbert (2010) present a voice interfaced driver manual which allows users to pose naturally spoken questions on the usage and maintenance of the vehicle and returns the most relevant answer from the manual database. On the basis of previous studies, the authors anticipate high recognition accuracy and expect the system to help reducing driver distraction and increasing driver satisfaction and manual usability. In the future, Alvarez and his colleagues want „to make the user more aware of the vehicle and the vehicle more aware of the surroundings". This might lead to providing the driver with information about the route conditions (traffic and weather), warnings about his physical state (fatigue) or similar, as shown in Figure 13.

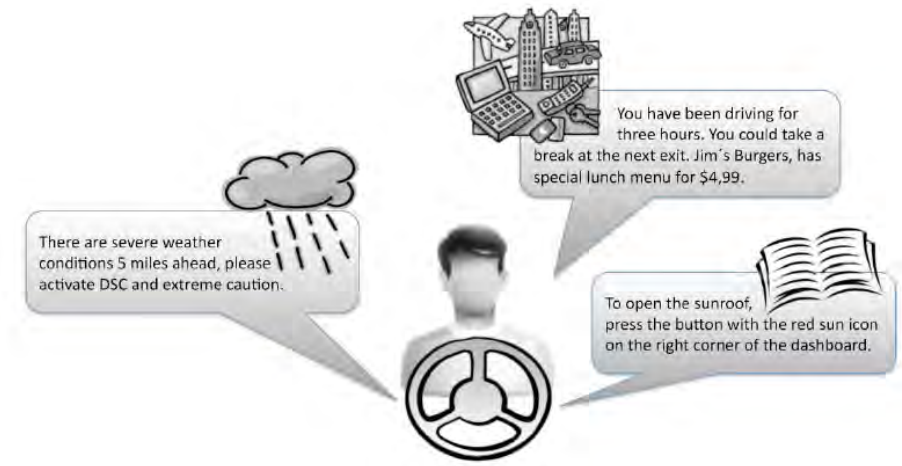

Figure 13 Vision of context-sensitive interactive manual user interactions
(Source: Alvarez, Martin, Dunbar, Taiber, Wilson & Gilbert, 2010)

4.3 Future applications for in-vehicle voice interaction

Visions like the one presented by Alvarez et al. (2010) build on a permanent internet access in cars. In fall 2008, Mercedes-Benz presented a fully internet-based infotainment system, which is called »myCOMAND«. It runs on a 4G wireless high-speed Long Term Evolution (LET) network. In Germany, a LTE network test-ground has been built up including an eight times greater bandwidth than today's telecommunication networks to test the system functionalities. The functions included map load, audio and video stream, street views satellite images, world radio, internet telephony via VoIP, up-to-the-minute information to find the cheapest gas station also as make a hotel or restaurant reservation. (Fischer & Nürnberger, 2010 and http://www.daimler.com).

The Continental Cooperation is predicting to release a network-based vehicle in 2012. The communication system called »Auto-LinQ« will combine navigation functions, internet and online-applications with safe driving. The system will allow drivers to access to vehicle centric information from almost everywhere and at any time. Also receiving and processing emails without distracting the driver will be possible due to speech interaction (cf. http://www.conti-online.com).

Permanent in-vehicle internet access will bring also social networks and other extremely fast communication services like twitter into the car. Even without having any well-founded studies on safety issues of using these services in the car, one can expect that they pose big challenges on the interaction design and driver attention. Voice interaction might play a major role in integrating new functionalities into a safe and comfortable driving environment of the future.

5 Conclusion

This paper reviews much of the existing knowledge on driver distraction, its causes and correlates, its effects on driving performance and accident hazard, and the role of voice interaction for a better driving safety. On the basis of the reviewed material, several conclusions can be drawn.

Based on the in-depth literature review at hand, voice interaction can be regarded as a promising approach to increase driving safety by avoiding the main factors of driver distraction. In comparison to visual-manual control modes, voice interaction generally improves the driving performance, reduces the subjective workload and allows drivers to keep their eyes on the road and their hands on the wheel. Voice interaction will become more and more common for in-vehicle systems.

For further improvements of voice interaction systems it will be important to carefully consider the relevant safety aspects. Safe use of in-vehicle systems depends on whether the interactions interfere with driving, whether drivers recognize the interference, and whether drivers are able to modulate their attention to minimize the consequence of this interference. When designing an in-vehicle information system it is important to ensure that the information and the method to convey this information do not negatively affect cognitive processing and driving performance. Reducing the interaction complexity will be a main issue for an effective voice user interface design. And finally, high speech recognition accuracy in all driving conditions (i.e. robustness) has to be a priority of speech-based interfaces in the future.

6 References

Alvarez, I., Martin, A., Dunbar, J., Taiber, J., Wilson, D.-M., & Gilbert, J. E. (2010). Voice Interfaced Vehicle User Help. In: Proceedings of the Second International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2010), November 11-12, 2010, Pittsburgh, Pennsylvania, USA

Barón, A., & Green, P. (2006). Safety and Usability of Speech Interfaces for In-Vehicle Tasks while Driving: A Brief Literature Review. Technical Report UMTRI.

Box, S. (2009). New Data from VTTI Provides Insight into Cell Phone Use and Driving Distraction. Virginia Tech Transportation Institute. Downloaded from: www.vtti.vt.edu/.../7-22-09-VTTI-Press_Release_Cell_phones_and_Driver_Distraction.pdf

Castro, C. (2009). Human factors of visual and cognitive performance in driving. Boca Raton, Fla.: CRC Press.

Castronovo, S., Mahr, A., Pentcheva, M., & Mueller, C. (2010). Multimodal Dialog in the Car: Combining Speech and Turn-and-Push Dial to Control Comfort Functions. In: Proceedings of Interspeech (2010). Conference in the Annual Series of Interspeech Events (INTERSPEECH-2010), September 26-30, Makuhari, Japan, ISCA, Makuhari, Japan, 9/2010.

Chang, J. C., Lien, A., Lathrop, B., & Hess, H. (2009). Usability Evaluation of a Volkswagen Group In-Vehicle Speech System. In: Proceedings of the First International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2009), Sep 21-22 2009, Essen, Germany.

Chisholm, S., Caird, J.K., & Lockhart, J. (2008). The effects of practice with MP3 players on driving performance. Accident Analysis and Prevention 40, 704–713.

Engstroem, J., Johansson, E., & Oetslund, J. (2005). Effects of visual and cognitive load in real and simulated motorway driving. Transportation Research Part F 8 (2005). 97–120.

Eyben, F., Wöllmer, M., Poitschke, T., Schuller, B., Blaschke, C., Färber, B., & Nguyen-Thien, N. (2010). Emotion on the Road—Necessity, Acceptance, and Feasibility of Affective Computing in the Car. In: Advances in Human-Computer Interaction - Special issue on emotion-aware natural interaction. Volume 2010. Hindawi Publishing Corp. New York, NY, United States.

Fischer, P., & Nürnberger, A. (2010). myCOMAND automotive user interface: Personalized interaction with multimedia content based on fuzzy preference modeling. In: User Modeling, Adaptation, and Personalization. Lecture Notes in Computer Science, 2010, Volume 6075/2010, 315-326.

Ford Motor Company (2009). Virginia Tech Transportation Institute Study Shows Ford SYNC's Voice Technology Helps Drivers Keep Eyes on Road. http://media.ford.com/article_display.cfm?article_id=32381.

Gärtner U., König, W. & Wittig, T. (2002). Evaluation of manual vs. speech input when using a driver information system in real traffic. Proceedings of the First International Driving Symposium on Human Factors in Driver Assessment (Ed.).

Gopher, D. & Donchin, E. (1985). Workload – An examination of the concept. In: Boff, K.R., Kaufmann, L., & Thomas, J.R. (Eds.): Handbook of Perception and Human Performance. Vol. 2, pp. 411-418. New York: Wiley

Green P. (2000). Crashes Induced by Driver Information Systems and What Can Be Done to Reduce Them. Society of Automotive Engineers, Inc.

Harbluk, J.L., Noy, Y.I., & Eizenman, M. (2002). The impact of cognitive distraction on driver visual behaviour and vehicle control. Report no. TP13889E. Transport Canada, Ontario.

Hosking, S. G., Young, K. L., & Regan, M. A. (2009). The Effects of Text Messaging on Young Drivers. SAGE Psychological Science.

Insurance Institute for Highway Safety (2010). IIHS Status Report newsletter: Phoning while driving. Insurance Institute for Highway Safety.

Jameson, A.H., Westerman, S.J., Hockey, G.R., & Carsten, M.J. (2004). Speech based e-mail and driver behaviour: effects of an in-vehicle message system interface. Human Factors 46 (4), 625–639.

Jensen, B. S., Skov, M. B., & Thiruravichandran, N. (2010). Studying driver attention and behaviour for three configurations of GPS navigation in real traffic driving. In: CHI '10 Proceedings of the 28th international conference on Human factors in computing systems. ACM, New York, NY, USA.

Kern, D., & Schmidt, A. (2009). Design Space for Driver-based Automotive User Interfaces. In: Proceedings of the First International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2009), Sep 21-22 2009, Essen, Germany.

Klauer, S. G., Dingus T. A., Neale V. L., Sudweeks J. D., & Ramsey D. J. (2006). The Impact of Driver Inattention On Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data. Virginia Tech Transportation Institute, National Highway Traffic Safety Administration.

Kun, A., Peak, T., & Medenica Z. (2007). The Effect of Speech Interface Accuracy on Driving Performance. Interspeech 2007.

Lee, J.D., Caven, B., Haake, S., & Brown, T.L. (2001). Speech-based interactions with invehicle computers. The effect of speech-based e-mail on drivers' attention to the roadway. Human Factors 43, 631–640.

Liang, Y., & Lee, J. D. (2010). Combining cognitive and visual distraction: Less than the sum of its parts. Accident Analysis and Prevention 42 (2010) 881–890

Maciej, J., & Vollrath M. (2009). Comparison of manual vs. speech-based interaction with in-vehicle information system. Accident Analysis and Prevention 41 (2009) 924–930

Mattes, S., 2003. The lane-change-task as a tool for driver distraction evaluation. In: Strasser, H., Kluth, K., Rausch, H., Bubb, H. (Eds.), Quality of Work and Products in Enterprises of the Future. Ergonomia Verlag, Stuttgart, pp. 57–60.

Mc Callum, M. C., Campbell, J. B., Richman, J.B., Brown, J.L. & Wiese, E. (2004). Speech Recognition and In-Vehicle Telematics Devices: Potential Reductions in Driver Distraction. International Journal of Speech Technology, 7(1), 25-33.

National Safety Council (2010). Understanding the distracted brain: Why driving while using hands-free cell phones is risky behavior. National Safety Council.

NHTSA National Highway Traffic Safety Administration (2010). Driving Distracted | Stats and Facts | Driving and Texting. U.S. Department of Transportation.

Olson, R. L., Hanowski, R. J., Hickman, J. S., & Bocanegra, J. (2009). Driver Distraction in Commercial Vehicle Operations. Virginia Tech Transportation Institute. U.S. Department of Transportation.

Rakauskas, M. E., Gugerty, L. J., & Ward, N. J. (2004). Effects of naturalistic cell phone conversations on driving performance. Journal of Safety Research 35 (2004), 453– 464.

Regan, M. A., Lee, J.D. & Young, K. (2009). Driver distraction: Theory, effects, and mitigation. Boca Raton, Fla.: CRC Press.

Stutts, J. C., Reinfurt, D. W., Staplin, L., & Rodgman, E. A. (2001). The Role of Driver Distraction in Traffic Crashes. AAA Foundation for Traffic Safety.

Wang, Y., Mehler, B., Reimer, B., Lammers, V., D'Ambrosio, L. A., & Coughlin, J. F. (2010). The validity of driving simulation for assessing differences between invehicle informational interfaces: A comparison with field testing. Ergonomics, 53: 3, 404 — 420.

Wickens, C. (1984). Processing Resources and Attention, Varieties of Attention. Academic Press: New York.

Wickens, C. D., & Hollands, J. G. (2000). Engineering psychology and human performance (3. ed.). Upper Saddle River, NJ: Prentice-Hall.

Young K., Regan, M., & Hammer, M. (2003). Driver Distraction: A Review of the Literature. Tech. Rep., Monash University Accident Research Center.

Zhang, H., Smith, M. R. H., & Witt, G. J. (2006). Identification of Real-Time Diagnostic Measures of Visual Distraction With an Automatic Eye-Tracking System. Human Factors, Vol. 48, No. 4, Winter 2006, pp. 805–821.

Zhang, H. & Wei, L. N. (2010). Speech Recognition Interface Design for In-Vehicle System. Proceedings of the Second International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2010), November 11-12, 2010, Pittsburgh, Pennsylvania, USA