Can User-paced, Menu-free Spoken Language Interfaces Improve Dual Task Handling While Driving?

Alexander Eriksson[1, 3], Anders Lindström[2], Albert Seward[2]
Alexander Seward[2], and Katja Kircher[3]

[1]Linköping University, Department of Computer and Information Science, Linköping, Sweden
aleer821@me.com
[2]Veridict AB, Stockholm, Sweden
{anders.lindstrom, albert.seward, alexander.seward}@veridict.com
[3]Swedish National Road and Transport Research Institute, Linköping, Sweden
katja.kircher@vti.se

Abstract. The use of speech-based interaction over traditional means of interaction in secondary tasks may increase safety in demanding environments with high requirements on operator attention. Speech interfaces have suffered from issues similar to those of visual displays, as they often rely on a complex menu structure that corresponds to that of visual systems. Recent advances in speech technology allow the use of natural language, eliminating the need for menu structures and offering a tighter coupling between the intention to act and the completion of the action. Modern speech technology may not only make already existing types of interaction safer, but also opens up for new applications, which may enhance safety. One such application is a speech-based hazard reporting system. A small fixed-base simulator study showed that drivers adapt the timing of the hazard reports to the situation at hand, such that an increase in reported workload was avoided.

Keywords: speech-based interface, natural language, compensatory behaviour, hazard reporting, human factors, VUI, strategic driving behaviour, simulated driving, IVIS

1 Introduction

The use of speech-based interaction over traditional means of interaction in secondary tasks may increase safety in demanding situations, like when driving an automobile or flying an aircraft, where the requirements on operator attention and vigilance are high. Traditional means of interaction such as screens, buttons and touchscreens usually involve rather complex menu structures, input fields and controls. Using those means while performing a high workload spatial task such as driving decreases overall performance significantly as this type of interaction competes for our limited resources [1-3].

Speech interfaces have until recently suffered from issues similar to those of visual displays, as they often rely on a complex menu structure corresponding to that of

visual systems, with the added drawback that it is normally inappropriate or even impossible to make said structure immediately visible on a screen. Instead, it has to be envisioned and/or remembered by the user. Typical examples are modern in-car navigation systems and speech recognition-based telephone services. These systems typically employ a rigid menu structure with clear expectations on the next input from the user, such that the input method is very similar to a traditional visual/manual input, except that verbal commands replace the hand movements.

Recent advances in speech technology allow the use of natural language and thereby eliminate the need for menu structures. The use of a natural language based speech interface offers a tighter coupling between the intention to act and the completion of the action as the user can jump between topics and deal with several conversation threads in parallel, as well as spontaneously introduce novel topics, without having to allocate any resources to accommodate the system structure. The conversation is essentially user-paced, which means that it can be interrupted, suspended and resumed just as the user sees fit. Modern speech technology has many potential application areas in driving, both within the domain of comfort systems and of safety systems. It may not only make already existing types of interaction safer, but also opens up for new applications, which may enhance safety.

In this article, although mainly theoretical, we will therefore also test a concept for a traffic hazard reporting system that is based on novel types of spoken language interfaces, which are characterised by being user-paced and menu-free, as opposed to more traditional types of speech interfaces. The main point of this paper is the assumption is that this new type of interface can be handled by drivers with only negligible interference with the driving task. Thus, its net effect is expected to be an increase in traffic safety, as drivers will be warned of hazards, while the action of reporting the hazard does not have any measurable negative impact.

1.1 Spoken Language Interfaces in Driving

There is ample evidence from different fields of research that adding a secondary task to a primary task usually deteriorates performance in the primary task [4-7]. Controlled studies of driving behaviour have shown time and again that adding a task like using a mobile phone or navigation system to the driving task often leads to increased reaction times [e.g. 8, 9]. Given that driving is generally agreed to have a large visual component [10], this concern is especially pronounced for tasks classified as having heavy visual/manual components, that is, where the driver needs to look away from the traffic and manipulate dials or screens. At the same time, contemporary naturalistic studies show that the act of talking on the phone was not associated with an increased crash risk when separated from dialling and reaching for the phone [11, 12].

These data therefore indicate that behavioural effects observed in simulator trials, in controlled test track tests or even in on-road studies cannot be translated one to one to behaviour in real traffic. Drivers in controlled studies are typically assigned pre-specified tasks to be carried out under certain conditions and are required to execute the secondary task in a given situation. In real traffic, however, the same drivers would have much greater room to compensate for their expected temporary attention-

al deficits. It is probably partly therefore that crash numbers have not increased substantially as the entertainment technology increasingly has found its way into our vehicles. Drivers do not use the technology completely uncritically, but they adapt their behaviour on the operational, the tactical and even on the strategic level, for example by slowing down, by choosing not to overtake, and by selecting situations of low complexity for secondary task execution [11, 13]. They also adapt how they execute the secondary task, for example by keeping telephone calls brief [14]. Some studies exist, although admittedly impressionistic in nature, indicating that drivers employ a whole range of linguistic devices and other communicative strategies in their spoken interaction within the vehicle or over a mobile phone line, in order to accommodate the cognitive demands of simultaneous talking and driving [15]. In a more recent study it was also investigated how cognitive load affects the degree of disfluencies during in-vehicle spoken dialogue between drivers and passengers performing a consciously demanding interview task while navigating in real traffic. The passengers acted as interviewers while also giving navigation instructions. On a side note, the authors made the collateral finding that all passengers actually spoke *less* disfluently when their drivers experienced high workload, which could also be seen as a tell-tale sign of co-operative adaptation from the side of the passengers [16].

While it is good news that drivers adopt strategies to improve their safety, it is of course advisable to offer methods of interaction that interfere as little as possible with the driving task. It may very well be that a certain type of secondary task can be impossible to perform with one type of interface, but very easily with another.

In order to make predictions about how different interface solutions affect the driving task, we lean on the concepts provided by the theory of threaded cognition [3]. In this theory it is assumed that different tasks within a multitasking environment are made up of different *threads*. Threads compete for the same cognitive, perceptual and motor resources. Different resources can operate in parallel, but each resource can only handle one request at a time. This explains why certain multitasking processes lead to degraded performance in either one or several tasks, how this can change over time with additional practice, and, furthermore, why multitasking in some cases does not have to lead to degraded performance in any of the subtasks.

According to the theory of threaded cognition, as well as other multiple resource theories [e.g. 2, 17, 18, 19], task interference resulting in degraded performance is more likely to occur and persist when different task threads contend for the same resources. Procedural resources are central and used frequently by any thread, but by avoiding the use of the same peripheral resources, parallel processing of different threads can be achieved. A further improvement is achieved by a reduction of the use of declarative resources, which occurs for example when instructions for actions have to be retrieved from memory. As shown by Salvucci and Taatgen [3], a visual dialling task produced a higher level of interference with driving than did a comparable voice-dialling task. Similarly, a study by Levy et al. [20] showed that resource interference degrades performance. Based on this notion the NHTSA published guidelines on how to test built-in visual/manual interfaces for their suitability while driving [21].

This would lead to the seemingly easy conclusion that the vocal/auditory channel should be the interface of choice for tasks that are to be processed in parallel with a

visual task like driving. However, empirical data show that this is not necessarily the case. Yager [22] had her participants type and send text messages manually, and verbally with two commercial speech-based mobile services (Siri and Vlingo) while driving. She did not find any differences between the input methods with respect to a number of performance indicators, like eye gaze to the forward roadway, standard deviation of lateral position, mean speed and speed variation. A closer examination shows, however, that the allegedly verbal input methods still included a rather intensive usage of visual resources, and as a result, a comparison with manual text input is not as clear-cut as it would first seem to be. Furthermore, the two speech-based input systems are also likely to have incurred an intensive use of declarative resources, since the participants were not necessarily familiar with the voice protocol that has to be used for the two services. Thus, they often had to retrieve information memorised during instructions and initial practicing.

This points to one of the major problems which Yager's study shares with many similar reports, namely that the attempted over-all comparison between two types of interfaces (here: "direct manipulation vs. speech-based") based on the direct comparison of specific systems fails to generalise from the specific case to the canonical. Furthermore, as rightly pointed out by Green [23], "the demand characteristics of in-vehicle tasks in question are not well quantified", and even if they are split up (as they often are, following an idea popularised by McCracken and Aldrich [24]) into visual, auditory, cognitive and psychomotor demands (VACP), the question of what exact levels of task demands should be considered to be excessive still remains largely unanswered. Also according to Green [23], this is of course further aggravated by the fact that the workload of the main task of driving is not well quantified, which, in turn, is why most investigations resort to some sort of indirect comparison between the influence of different secondary tasks, as exemplified by Yager's study.

It is worth pointing out that many pre-existing commercial and research systems involving spoken input for mobile use, both in and outside of vehicles, in fact do not profit from the main advantages of spoken language interaction, but do little more than "push buttons using voice". For reasons unknown, the hierarchical menu design brought about by WIMP[1]-style computer GUIs already more than three decades ago, has had a remarkable but undesirable tendency to carry over to Voice User Interface (VUI) design, with many awkward and unintuitive system designs as a result. We would like to argue that making decisions and developing safety guidelines based on studies of such systems is misleading. Instead, we suggest that studies be made involving (real or simulated) spoken language systems where:

1. The (secondary) task lends itself to verbal interaction
2. The system design takes advantage of the intrinsic benefits of using human language and properly exploits the verbal and auditory channels
3. Users are allowed the benefit of training
4. Users are allowed the benefit of (tactical) planning

---

[1] windows, icons, menus, pointer

The first issue might seem trivial, but is in fact often overlooked. There are many cases where operating a button via direct manipulation and with instant tactile, visual or auditory feedback is optimal, such as when turning on the headlights or honking the horn. On the other hand, dashboards would soon be completely cluttered if each and every function stemming from the infotainment escalation in recent years had got a button of its own. The second item is perhaps the most generally ignored, and consequently holds the most development potential. We will come back to that later in the choice of experimental task. The third point reminds us that it is otherwise considered acceptable to allow considerable amounts of training, which is normally required for example when learning how to operate a stick-shift transmission, or when adapting to the levers and buttons of a new car. The fourth bullet highlights that the leeway introduced by having a user-paced scenario may be enough to accommodate a range of secondary tasks without negative interference with the primary task of driving.

1.2   A Tentative Service Scenario

Today, many radio channels and TV stations provide live (non-critical) traffic congestion, hazard, and obstacle reporting. Reporting is typically done by telephone. Eventually the information will be relayed to a large number of road users, for example by public broadcasting on the radio, possibly with prioritised reception by virtue of the Traffic Announcement bulletin handling present in RDS-enabled radio receivers since the late 1980s.

As both driving and making the phone call depend on procedural and visual resources, dual task performance is likely to be degraded. There are further technical disadvantages associated with making hazard reports by phone. The driver needs to give the precise location of the hazard, which puts demands on declarative memory. The hazard report has to be processed by a human, causing an unavoidable delay from reporting to broadcasting. Finally, distribution via broadcast radio inevitably precludes the possibility to individualize the message and to convey it only to those affected by the hazard.

Given all these issues it is worth investigating whether a traffic hazard reporting system could be automated. The idea is to use verbal reports from drivers, connect them to the position at which the report was made, have a backend that evaluates all reports, and then distribute appropriate information, warnings and alerts selectively to drivers in the affected area. As a first step, in this study we investigate how a simulated voice controlled reporting system would be operated by a driver in traffic. For such a system to be safe for use in traffic it should neither have a substantial impact on reported workload nor on driving performance. Given the low fidelity of the simulator that was available for the study, we did not consider it meaningful to assess driving performance directly. Instead, we decided to investigate how drivers would time their reports with respect to the hazards and obstacles on the road, in keeping with the notion of tactical self-regulation. Based on this it may be possible to draw conclusions about the likelihood that voice based hazard reporting will affect driving performance. We therefore specifically included a comparison between the drivers making reports at their own discretion and a forced "report-as-soon-as-possible" condition.

1.3 Requirements on a Speech-Based Hazard Reporting System

Based on the arguments given in the introduction, we would like to suggest that the following requirements on system design and system performance need to be fulfilled to avoid an excessive increase in workload and to gain user acceptance:

Satisfactory Speech Recognition. In a study by Kun, et al. [25] drivers were given a speech interaction task with a system capable of a simulated high recognition rate (89% of system dialogue turns) and with a system with a simulated poor recognition rate (44%). Their results show that the system with the low recognition rate caused a significant increase in lane position variance when using the push-to-talk (PTT) button, but this effect did not transfer to the better-performing system. These results may be caused by increased workload in the auditory system imposed by the system's poor recognition in combination with the visuospatial task of reaching for the PTT button in the centre console, resulting in an overall increase in workload.

Responsive, Intuitive and Effortless. Auditive and verbal feedback needs to be immediate, just as in human-human conversation. Reporting should be as easy as putting words on thoughts. Verbal reporting is already the primary means today, although it is done by talking over the phone with a human operator, so the task is obviously feasible. Furthermore, as Green [26] has pointed out, issuing a brief verbal command (like reporting a traffic congestion – in his example a voice-controlled radio is operated) is likely to require minimal thought.

Training. The envisioned mobile speech-based system covers many functions with a similar and consistent VUI, and users of the system are typically guided through interactive tutorials for each function specifically for training purposes. Users can also re-visit these tutorials and should have ample time to practice reporting prior to use.

Tactical Choice of Reporting Time. The system should possess functionality for the interruption of on-going dialogues as workload increases and should be able to resume the dialogue where the driver left off as the workload level decreases [27, 28]. We therefore suggest a speech-based system that is completely user-paced, such that users are free to choose the time of reporting to suit their driving pattern.

2 Method

2.1 Participants and Equipment

A convenience sample of 17 participants between 20 and 27 years of age (Mean=24, SD=2.3) took part in the study. All participants were required to have a valid driving license and a minimum of 2 years of driving experience on Swedish roads. The participants were students at Linköping University. Each participant was given two cinema tickets as compensation for participating in the study.

The VTI fixed-base simulator was used for the purposes of this study. The driver environment is constructed from parts of a Ford Focus and has all the essential controls such as the transmission stick, clutch, wheel and brakes. The simulation software was executed on a distributed computer system and the graphics were rendered with a resolution of 720x1280 pixels at 60 Hz and displayed on three 40" flat-screen TVs

with a 1080x1920 resolution providing an approximated 120° field of view. Sound effects were provided through a 5.1 Logitech® surround sound system. A 7" resistive touchscreen was placed on the centre console representing a full screen button activating the speech interface.

2.2　Primary and Secondary Task

The primary task was to drive approximately 20 km on a rural road with a village in the middle. The following nine hazards and obstacles were placed along the road:

- a truck trailer parked on the shoulder of the road
- a moose moving towards the road and then stopping at the road side
- a car parked at a bus stop
- a broken down car in an intersection
- a road construction site on an urban road
- a road construction site in an intersection
- a broken down car in an intersection with a cyclist crossing the road with oncoming traffic
- a truck partially parked on the shoulder of the road and a connecting road
- oncoming traffic and a cyclist in the participant's lane, cycling in the opposite direction

The secondary task used in this experiment was a speech interface for reporting roadside hazards and hindrances. The system was simulated using the Wizard of Oz method [29] where a hidden experimenter provided feedback using a synthesised voice. The reason for this was to ensure that VUI design issues would not adversely affect the drivers' performance. The system was activated by pressing the touchscreen, which resulted in auditory and visual feedback. The subject then provided a verbal hazard report that the system associated with the vehicle's physical position along the simulated road. Upon completing a report the subject received auditory feedback by a synthesised voice triggered by the experimenter.

2.3　Procedure and Design

The participants were asked to drive as they would normally do while paying attention to the traffic regulations. They had an initial training phase in the simulator for about five minutes before proceeding to the experimental conditions. There were four different experimental conditions:

- a *baseline* in which participants drove along the route without any additional tasks
- a *self-paced* experimental run in which the participants were asked to use the speech-based hazard reporting system as they saw fit
- a *video* run in which the drivers watched a film of a run in the simulator and used the speech-based reporting system as instructed in the self-paced run
- a *forced* experimental run in which the drivers were instructed to use the system as soon as they detected any traffic hazard or obstacle (externally paced)

It was decided that the order of the reporting conditions should go from least to most specific and enforcing, such that reporting behaviour in the self-paced conditions would not be influenced by the more externally paced condition. This meant that baseline driving always came first, followed by the self-paced and the video condition, the two latter of which were counterbalanced. The forced condition always came last. The ensuing risk for learning effects was considered, but viewed as less problematic than a carry-over effect of the forced reporting behaviour to the self-paced behaviour.

2.4   Data Collection and Analysis

A log was kept of when the participants started reporting each hazard/obstacle. Audio was recorded during the experimental runs using a microphone mounted on top of the dashboard. All participants were asked to fill out the NASA-RTLX form [30] after each run to obtain self-reported workload measurements.

The road around each obstacle was divided into *Area 1* (before the hazard/obstacle), *Area 2* (next to the hazard/obstacle) and *Area 3* (behind the hazard/obstacle). Area 1 started as soon as the obstacle became visible and ended 20 m in front of the obstacle. Area 2 lasted from 20 m in front of the obstacle to 20 m behind the obstacle. Area 3 started 20 m behind the obstacle and ended 150 m behind it. The number of hazard reports per obstacle, area and condition were counted based on where the report was initiated by pressing the touch screen.

3   Results

3.1   Reporting Strategy

For the self-paced condition a total of 115 obstacles were reported, in the video condition 110 obstacles were reported, and for the forced condition 127 obstacles were reported. The obstacle that was least likely to be reported across conditions was a road construction in town, followed by a road construction in a crossing and a bicyclist cycling on the wrong side of the road.

The number of reports per area and condition is displayed in

. In the self-paced condition 57% of the reports were initiated in Area 1, that is, before the area immediately surrounding the obstacle, was reached. More than half of the 22% of reports initiated in Area 2 in the self-paced condition were associated with the moose standing next to the road and with the bicyclist crossing the intersection. The share of reports (21%) that were initiated in Area 3, after the obstacle, were most frequently connected to the truck trailer and the truck parked on the roadside.

In the video condition 63% of the reports were initiated in Area 2 and 31% were initiated in Area 1. Only 6% were initiated in Area 3.

In the forced condition the vast majority of reports (89%) was initiated in Area 1, 7% of the reports were initiated in Area 2, and 4% were initiated in Area 3.
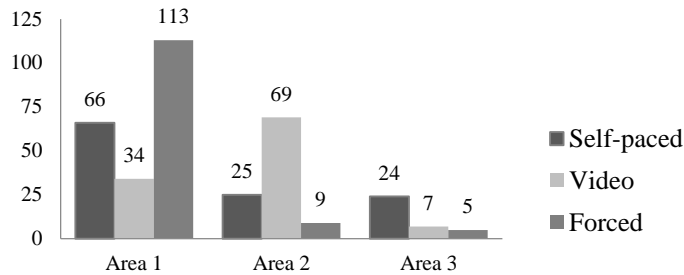
Fig. 1. Report count per road area and reporting condition for 17 drivers performing a simulated speech-based hazard-reporting task

Workload Measurements

Differences in workload ratings between conditions were analysed using an ANOVA, the results of which are displayed in Table 1. There was a significant difference between conditions ($F_{(3, 42)}$=5.565, p=.003, ω^2=.284, power=.985). Post-hoc tests at the uncorrected α=.05 level showed that the reported workload for baseline was significantly higher than for the self-paced and the forced condition, workload was also higher for the video condition than for the forced condition.

Table 1. Mean workload ratings on the NASA-RTLX and p-values for the post-hoc comparisons in the ANOVA with the factor condition

		baseline	video	self-paced	forced
mean ± sd		164.0±65.0	146.3±64.3	117.7±52.3	105.0±53.4
post-hoc test results	video	p = .091			
	self-paced	p = .021	p = .092		
	forced	p = .011	p = .025	p = .401	

4 Discussion

There is a large contrast between the reporting strategies employed in the different driving conditions, indicating that drivers adapt flexibly to the situation at hand. In the self-paced and in the video condition the drivers were instructed to use the system in the same way, except that in the video condition the driving task was excluded. In the self-paced condition drivers tended to make their reports in Area 1, which could be a strategy to avoid an increase in workload when passing the obstacles. Notable exceptions are the moose on the roadside as well as the bicyclist in the crossing. Here, the workload may be higher in Area 1, as drivers need to brake and assess the somewhat unpredictable behaviour of the moose or bicyclist. In those two cases reporting is often delayed to Area 2, where there is no longer any immediate threat. This claim of

the drivers' self-regulation is further supported by the fact that the general strategy changes in the video condition when no workload from driving is contributing to the overall workload. As the drivers do not have to focus on driving they tend to use the reporting system in close physical proximity to the obstacle.

In the forced condition practically all reports are initiated in Area 1, which shows that the participants followed the instructions. As this reporting behaviour does not correspond completely to the drivers' natural strategy, it is important to consider how instructions are phrased in further studies. It is recommended to give participants more leeway in executing secondary tasks, as this enables the employment of compensatory strategies and has more ecological validity.

The ANOVA on workload unexpectedly showed significant differences between the different driving conditions. It is worth noting, however, that there is a decrease in workload as the drivers work their way through the different conditions. This most likely reflects a learning effect with the drivers familiarizing themselves with the road conditions, the obstacles, the simulator and the reporting task over time. Still, the addition of the reporting task does not lead to reported workload levels above baseline, which is promising. This could possibly be explained in light of Salvucci's theory, which suggests that verbal tasks should lend themselves to being integrated with driving. In future studies learning effects have to be addressed more carefully, for example by working with highly trained participants, or by employing a between-group design. Also, performing workload assessment more frequently could provide further insight into how drivers are affected by driving and/or the reporting task.

5   Conclusions

The results show that drivers employ compensatory strategies when making speech-based hazard reports, and that the strategy seems to be dependent on the complexity of the traffic situation. The results also indicate that overall reported workload does not increase when using a speech-based hazard reporting system, providing initial support for the presented theory. However, the scope of the present study was limited, therefore it is recommended to follow up on the results and expand them, taking the recommendations given here into account.

References

[1]   Maciej, J., Vollrath, M.: Comparison of manual vs. speech-based interaction with in-vehicle information systems. Accident Analysis and Prevention, 41, 924-930 (2009).

[2]   Wickens, C.D.: Situation awareness and workload in aviation. Current Directions in Psychological Science, 11, 128-133 (2002)

[3]   Salvucci, D.D., Taatgen, N.A.: Threaded cognition: an integrated theory of concurrent multitasking. Psychological Review, 115, 101-130 (2008)

[4]   Wu, X., Li, Z.: Secondary task method for workload measurement in alarm monitoring and identification tasks. In: Rau, P.L.P. (ed.) Cross-Cultural Design. Methods, Practice, and Case Studies. vol. 8023, pp. 346-354. Springer Berlin Heidelberg (2013)

[5]   Grant, R.C., Carswell, C.M., Lio, C.H., Seales, W.B.: Measuring surgeons' mental workload with a time-based secondary task. Ergonomics in Design: The Quarterly of Human Factors Applications, 21, 7-11 (2013)

[6]   Alm, H., Nilsson, L.: The effects of a mobile telephone task on driver behavior in a car following situation. Accident Analysis and Prevention, 27, 707-715 (1995)

[7]   Beede, K.E., Kass, S.J.: Engrossed in conversation: The impact of cell phones on simulated driving performance. Accident Analysis and Prevention, 38, 415-421 (2006)

[8]   Horrey, W.J., Wickens, C.D.: Examining the impact of cell phone conversations on driving using meta-analytic techniques. Human Factors, 48, 196-205 (2006)

[9]   Caird, J.K., Willness, C.R., Steel, P., Scialfa, C.: A meta-analysis of the effects of cell phones on driver performance. Accident Analysis and Prevention, 40, 1282-1293 (2008)

[10]   Sivak, M.: The information that drivers use: Is it indeed 90 % visual? Perception, 25, 1081-1089 (1996)

[11]   Fitch, G.M., Hanowski, R.J.: The risk of a safety-critical event associated with mobile device use as a function of driving task demands. Second Conference on Driver Distraction and Inattention, Gothenburg, Sweden (2011)

[12]   Klauer, S.G, Guo, F., Simons-Morton, B.G., Ouimet, M.C., Lee, S.E., Dingus, T.A.: Distracted driving and risk of road crashes among novice and experienced Drivers. New England Journal of Medicine, 370, 54-59 (2014)

[13]   Cooper, J.M., Vladisavljevic, I., Medeiros-Ward, N., Martin, P.T., Strayer, D.L.: An investigation of driver distraction near the tipping point of traffic flow stability. Human Factors, 51, 261-268 (2009)

[14]   O'Brien, N.P., Goodwin, A.H., Foss, R.D.: Talking and texting among teenage drivers: A glass half empty or half full? Traffic Injury Prevention, 11, 549-554, (2010)

[15]   Esbjörnsson, M., Juhlin, O., Weilenmann, A.: Drivers using mobile phones in traffic: An ethnographic study of interactional adaptation. International Journal of Human-Computer Interaction, 22, 37-58, (2007)

[16]   Lindström, A., Villing, J., Larsson, S., Seward, A., Åberg, N., Holtelius, C.: The effect of cognitive load on disfluencies during in-vehicle spoken dialogue. Interspeech, Brisbane, Australia (2008)

[17]   Wickens, C.D.: Processing resources in attention. In: Parasuraman, R., Davies, D.R. (eds.) Varieties of attention, pp.63-102. New York: Academic Press (1984)

[18]   Wickens, C.D.: Multiple resources and mental workload. Human Factors, 50, 449-55 (2008)

[19]   Derrick, W.L.: Dimensions of operator workload. Human Factors: The Journal of the Human Factors and Ergonomics Society, 30, 95-110 (1988)

[20]   Levy, J., Pashler, H., Boer, E.: Central interference in driving: is there any stopping the psychological refractory period? Psychological Science, 17, 228-35, (2006)

[21]   NHTSA: Visual-manual NHTSA driver distraction guidelines for in-vehicle electronic devices. Docket NHTSA-2010-0053 (2012)

[22]   Yager, C.: An evaluation of the effectiveness of voice-to-text programs at reducing incidences of distracted driving. Texas A&M Transportation Institute, The Texas A&M University System, College Station, Texas 77843-3135, Technical Report SWUTC/13/600451-00011-1 (2013)

[23]   Green, P.A.: Driver interface/HMI standards to minimize driver distraction/ overload. UMTRI, SAE Paper 2008-21-2002 (2008)

[24]   McCracken, J.H., Aldrich, T.B.: Analyses of selected LHX mission functions: Implications for operator workload and system automation goals. Anacapa Sciences Inc., Research note ASI-479-024-84B, (1984)

[25]   Kun, A.L., Paek, T., Zeljko, M.: The effect of speech interface accuracy on driving performance. Interspeech, Antwerp, Belgium (2007)

[26]   Green, P.A.: Crashes induced by driver information systems and what can be done to reduce them. Conference of the Society of Automotive Engineers (SAE), Warrendale, PA, USA (1999)

[27]   Shioya, M., Nishimoto, T., Takahashi, J., Daigo, H.: A study of dialogue management principles corresponding to the driver's workload. In: Abut, H., Hansen, J.L., Takeda, K. (eds) Advances for In-Vehicle and Mobile Systems, pp. 251-265, Springer US (2007)

[28]   Villing, J., Larsson, S.: Speech, buttons or both? A comparative study of an in-car dialogue system. Third International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Salzburg, Austria (2011)

[29]   Dahlbäck, N., Jönsson, A., Ahrenberg, L.: Wizard of Oz studies — why and how. Knowledge-Based Systems, 6, 258-266 (1993)

[30]   Byers, J.C., Bittner, A.C., Hill, S.G.: Traditional and raw task load index (TLX) correlations: are paired comparisons necessary? International Industrial Ergonomics and Safety Conference, Cincinnati, Ohio (1989)