

Effects of Speech-based vs Handheld E-mailing and Texting on Driving Performance and Experience

Jacques Terken
Technische Universiteit Eindhoven
P.O. Box 513
5600 MB Eindhoven NL
+31 40 247 5254
j.m.b.terken@tue.nl

Henk-Jan Visser
Ikspreek.nl
P.O. Box 6165
5600 HD Eindhoven
+31 40 402 1125
henkjan@ikspreek.nl

Andrew Tokmakoff
Ikspreek.nl
P.O. Box 6165
5600 HD Eindhoven
+31 40 402 1125
tokmakoff@ikspreek.nl

ABSTRACT

In this paper we present a voice-enabled service for handling e-mail and SMS messages while driving, and an evaluation of the service. In the evaluation, driving performance was compared in three conditions with a highway driving scenario in a driving simulator: driving only, driving in combination with voice-enabled handling of e-mail/SMS messages, and driving in combination with handheld handling of e-mail/SMS messages. Both objective measurements and subjective judgments about driving performance were collected. The results showed that drivers increased the headway when performing an additional task. With respect to the subjective measures, drivers felt that driving only was safest, and that voice-enabled interaction was safer, enabled better concentration and a better driving performance than handheld interaction. We conclude that handheld interaction is felt to have a stronger impact on driver performance and workload than voice-enabled interaction.

Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation]: User Interfaces

General Terms

Human Factors

Keywords

Voice interfaces; Driving performance; Multitasking

1. INTRODUCTION

With the advent of smart phones, drivers have increasingly started to engage in e-mailing and texting while driving. A recent investigation indicates that as many as 40% of drivers owning smart phones engage in e-mailing and texting while driving [10]. From an interaction point of view, this is a setback compared to handsfree calling, since it forces the driver again to interact manually with the phone and to direct his/her visual attention away from the traffic. One may hypothesize that speech-enabled e-mailing and texting would reduce driver distraction in comparison to manual e-mailing and texting. In this paper, we compare the effects of speech-based and handheld emailing and texting on driving performance. In addition, we compare how the different ways of emailing and texting

compare to driving only, in order to assess potential decrements in driving performance due to multitasking.

Numerous studies have investigated the effects of speech-based interaction whilst driving relative to driving only, mostly focusing on handsfree calling (among others [11][4][5]). The general conclusion from these studies is that “a speech-based interface is not a panacea that eliminates the potential distraction of in-vehicle computers” [6], in other words, that multi-tasking leads to poorer driver performance. Relatively few studies have compared handheld and handsfree calling directly. The conclusion of a meta-analysis of over thirty studies reported in [3], four of which compared handheld and handsfree calling directly, is that both handsfree and handheld calling have considerable impact on driving performance and driver attention, while the difference in performance between handsfree calling and handheld calling is relatively small. On the other hand, a somewhat smaller review led to the conclusion that speech interfaces result in less workload and reduced eyes-off-the-road times compared to manual interfaces [1].

The reason for conducting the current study is two-fold. In the first place, the present study was conducted to evaluate one particular speech interface for handling e-mail and SMS messages. Also, the current research focused on a highway driving context, which seems a more suitable and obvious context for e-mailing and texting while driving than driving in the city. Since the studies reported in the literature impose different contexts of use (different tasks, different interfaces and different testing conditions), direct comparison and prediction of effects for a specific interface and a specific driving context is usually difficult. In the second place, as noted above, there are relatively few studies directly comparing handheld and speech-based interaction, or both conditions and the baseline condition of driving only; [3] mentions four studies and concludes that multi-tasking leads to a performance decrement, while there is little difference between handheld and handsfree interaction. Additional studies; showing a somewhat different picture, are reported in [7] and [8]. [7] investigated manual vs speech-based interaction with a PDA while driving, and found that speech-based interaction was in between manual interaction and driving only in terms of subjectively experienced workload and braking response times. [8] compared handheld and speech-based control of phoning and portable music players. It was found that handheld interaction resulted in poorer performance and more subjective mental demand than speech-based interaction for music players but not for phone conversations.

Given the relative scarcity of studies comparing speech-enabled and manual time-sharing conditions and a driving only condition, the differences between tasks and the sometimes contradictory outcomes, it appeared relevant to add further results and extend the body of knowledge. In the remainder of the paper we first characterize an application that was developed for speech-enabled handling of e-mail and SMS messages, and then present a study in which the effects of speech-enabled handling of e-mail and SMS messages while driving were compared with handheld handling and with driving only.

2. THE SYSTEM

An interface was developed giving access to e-mail and SMS messages on a smart phone through speech. Design principles were chosen to create a simple and easy-to-learn interface. The interface gives access to e-mail and test messages through a limited set of simple commands such as “e-mail”, “next”, “previous”, “repeat” and “done”. The header and content of the message is then converted to speech by text-to-speech (TTS) software. Furthermore, the User Interface also presents a simple menu for responding to received messages. By saying a digit name between 1 and 5, the driver can reply to the message. Option 1 activates a function allowing the driver to send a voice response. Options 2 to 5 result in replying through a predefined test message such as “I’m driving now” or “I will get back to you later”. In all cases the available commands are shown on the display of the smartphone to improve the ease of use and prevent the driver from having to memorize the commands in advance. Clearly, this forces the driver to look at the screen in the initial stages of use, but the assumption is that the glances at the screen will be shorter for a speech interface than when the whole interaction is hand-based. In addition, displaying the available commands on the screen will enable the driver to learn the commands while using the system, so we may assume that the need to consult the screen will decrease with prolonged use.

3. EVALUATION

3.1 Tasks

Tests were run with a fixed-base driving simulator with standard hand-gearing (common in European cars) (see Figure 1). The primary target users for the system are people who drive longer stretches between cities on the highway, such as business people and commuters. Typically, in such contexts the traffic condition is rather busy. Therefore, a highway driving scenario was chosen with busy traffic conditions, requiring drivers to stay attentive. The highway driving scenario limited the required operational tasks to steering and speed control through pushing the accelerator and braking, while excluding hand gearing operations. Participants drove the highway scenario once without additional tasks, once while handling incoming e-mail and SMS messages manually, and once while handling incoming e-mail and SMS messages through a speech interface. E-mail and SMS messages were simple messages of approximately 25 words such as proposing the rescheduling of an appointment. All participants received the same messages.

3.2 Participants

Forty-three participants were recruited from the student population of the department of Industrial Design of Eindhoven University of Technology, both male and female, with ages ranging from 20 to 29 years old. Conditions for participation

were being in possession of a driver’s license and a smartphone with access to e-mail and SMS. Participants were awarded a small fee for their participation.

3.3 Procedure

Participants took the test individually. Upon arrival, each participant was provided with a brief explanation of the test and instructions, read and signed a consent form and filled in a form to provide personal information such as driving history and experience with e-mailing and texting. Next, the participant drove the highway scenario to familiarize with driving in a driving simulator until s/he felt sufficiently familiar with driving in a simulator. Next, the participant received a smart phone on which the speech interface for handling e-mail and SMS messages was installed and a quick reference card, and performed a few tasks involving handling and replying to e-mail and SMS messages. When the participant felt sufficiently familiar controlling the application by voice, the actual test began. The participant performed three runs with the highway scenario: once without additional tasks, once while handling incoming e-mail and SMS messages manually, and once while handling incoming e-mail and SMS messages through a speech interface. The order of runs was varied between subjects, so that different orders were balanced across subjects. Figure 1 shows how the mobile phone providing the speech-based service was mounted in the driving simulator. In each multitasking run, the participant received two e-mail messages and two SMS messages. S/he was instructed to send a simple reply such as “I’m underway” (implying that a more extensive reply is not possible). In the case of the speech interface this act was executed by saying a digit. The participant was instructed to comply with the traffic rules. For the multitasking runs, the participant was told to decide him/herself how to divide attention between the driving task and the additional task, but to act as s/he would do normally (implying that safety would be the highest priority). After each run, the participant filled a questionnaire, providing his/her opinion about the difficulty of the task. Each run lasted about seven minutes. The total duration of the test was approximately 45 minutes

3.4 Data collection

Both objective measurements and subjective judgements were collected. Objective measures were number of collisions; number of near collisions (characterized by a short headway to the preceding car followed by abrupt braking), average speed and speed variability, and average headway and headway variability; average speed and headway, and speed and headway variability are commonly used metrics in research on driver performance and distraction [11][4][5]. For the calculation of speed and speed variability and average headway and headway variability onsets and offsets were removed from the logfiles. Onsets involved the acceleration from 0 km/h to cruise speed, offsets concerned the slowing down phase at the end of run. The cutoff criterion for onsets and offsets was 30m/s (108 km/h). The durations of the onset and offset were on average approximately 50 s and less than 10 s, respectively. The durations of the e-mail and SMS-tasks were measured by hand, from the moment of arrival to the moment of sending the answer. Subjective judgements were obtained through a questionnaire asking how safe the participants felt s/he could drive, how well s/he could concentrate on the driving task and how the driving performance compared to his/her usual driving

performance on 7-point scales, using Likert type questions (statements to which the participants have to express their agreement). Also, they filled the Rating Scale for Mental Effort [12], expressing mental effort on a 150 point labelled scale, ranging from absolutely no effort to extreme effort.



Fig. 1. Left: The driving simulator used in the evaluation. Right: Detail, showing the placement of the mobile phone running the speech interface.

4. RESULTS

Data from two participants were discarded because of technical problems, so that the analyses are based on data from 41 participants. The average total duration of the e-mail and SMS tasks was 3.50 min in the handheld condition and 4.48 min in the speech-based condition.

4.1 Objective measurements

Table 1: Means for objective measures. Speed in metre/s, speed variability: SD of speed. Headway in metres and seconds, headway variability: SD of headway (m)

	Driving	Driving + Handheld	Driving + Speech-based
Nr of collisions	0.12	0.24	0.12
Nr of near collisions	0.37	0.41	0.61
Speed (m/s)	30.3	30.1	30.1
Speed variability	3.3	3.3	3.3
Headway (m)	56.1	62.8	61.6
Headway (s)	1.85	2.08	2.05
Headway variability	40.0	44.8	43.3

Table 1 shows the means for the objective measurements. The data were subjected to statistical tests for significance, using Analyses of Variance with Repeated Measurements, with the following outcomes. Number of collisions: $F_{2,80}=0,486$, n.s. ($p=.63$); Number of near collisions: $F_{2,80}=2,37$, n.s. ($p=.10$); Speed: $F_{2,80}=1,43$, n.s. ($p=.246$); Speed variability: $F_{2,80}=0,35$, $p=.966$, Headway: $F_{2,80}=3,12$, $p=.05$, Headway variability: $F_{2,80}=1,62$, n.s. ($p=.30$) (although the differences for Headway variability were in the expected direction). Posthoc tests applying Bonferroni correction on the means for Headway showed that the Driving only and Driving+Handheld conditions were significantly different ($p=.046$). The difference between driving only and Driving+Speech-based handling of e-mail and SMS messages was not statistically significant.

4.2 Subjective measurements

Table 2 shows the mean ratings for the three questions: “I felt I could drive safely”, “I felt I could concentrate well on the driving task”, “My driving performance compared well to my normal driving performance” and for the subjective estimate of mental effort, using the 150 point Rating Scale for Mental Effort [12]. The data were subjected to statistical tests for significance, using Analyses of Variance with Repeated Measurements. For all aspects the effect was statistically significant. Safety: $F_{2,80}=68,69$, $p<005$; Concentration: $F_{2,80}=62,48$, $p<.005$; Driving Performance: $F_{2,80}=45,52$, $p<.005$; Mental workload: $F_{2,80}=71,48$, $p<.005$. Posthoc tests applying Bonferroni correction showed that for all aspects all pairwise differences between individual conditions were statistically significant.

Table 2: Means for subjective judgements. Judgements for Safety, Concentration and Performance are expressed on a 7-pt scale, with 1 bad, 7 good. For mental effort judgements are expressed on a 150 pt scale (RSME) with 1 low effort, 150 high effort.

	Driving	Driving + Handheld	Driving + Speech-based
Safety	5.8	2.7	4.3
Concentration	6.1	2.9	3.9
Performance	4.9	2.2	3.3
Mental Effort	35	77	64

5. CONCLUSION AND DISCUSSION

The results for the objective measures show that the task significantly affects the headway: the mean headway for driving without additional task is smaller than for driving in combination with handheld handling of e-mail and SMS text messages (in the case of speech-based handling of e-mail and SMS messages this difference was almost as large but not statistically significant). That is, when handling e-mail and SMS messages manually, drivers on average maintain a larger distance to a preceding car than when only driving. In general, this suggests that drivers increase the distance to a preceding car when engaging in a distracting activity in order to reduce the risk of a collision. This tendency to compensate has been noted before ([1][3]), although the present study indicates that this compensation may not always be sufficient, as the number of collisions was slightly higher (but not significantly so) in the handheld condition. None of the other objective measures shows a significant effect of driving condition. Although the results for headway variability are in the expected direction, the difference is not significant.

From the subjective judgements we see that participants felt that driving only was safest and allowed them to concentrate and perform better than when driving while performing an additional task. Also, the mental effort for driving only was felt to be lower than for driving with an additional task. If we compare the two conditions where driving is combined with an additional task, participants felt that driving combined with speech-based e-mail/SMS handling was safer, enabled them to concentrate better and to perform better than driving combined with handheld e-mail/SMS handling. Also, the mental effort was felt to be lower for driving combined with speech-based e-mail/SMS

handling than for driving combined with handheld handling. It needs to be pointed out that people had very limited experience in using the speech-based interface. Also, the performance of the automatic speech recognition was not always at a sufficient level (which may also account for the task completion times). It may be assumed that more experience and a better ASR performance would have further reduced the subjectively experienced workload for the speech interface condition.

There are several possible explanations for this pattern of result. In the first place, the significant difference in headway and the absence of significant effects for the other objective measures may stem from a driver strategy. We instructed participants to decide for themselves how to divide their attention between the driving task and the e-mail/SMS handling task. The consideration here was ecological validity, because in reality drivers will also prioritize tasks themselves. The consequence of this may have been that drivers gave priority to the driving task in order to drive safely (as most drivers would do in regular traffic too), and that therefore we didn't observe the performance decrements observed in other studies. In fact, the headways (whether measured in meters or seconds) are relatively large compared to what may be observed on the highway during rush hours. In the second place, the present study employed a driving simulator, and it has been suggested that driving on real roads is affected more by performing additional tasks than driving in a simulator [9]. Finally, the subjective judgements indicate that the Driving Only condition induced only moderate workload. We cannot exclude the possibility, therefore, that people just gave more effort in the multitasking conditions. While such a strategy would account for the increase in subjective effort in the multitasking conditions, at the same time it would reduce differences in objective performance between the Driving only and the multitasking conditions, because drivers would just work harder to maintain an adequate level of performance.

In terms of objective measurements, our results do not confirm the well-attested finding in many other studies that driving performance is adversely affected by multi-tasking, whether it concerns speech-enabled or manual interaction [3]. Instead, we observed a tendency for drivers to increase the headway, so that they were able to maintain their driving performance as measured by number of collisions and near collisions, therewith compensating for safety hazards which may result from multi-tasking. In many other studies reaction time is taken as a measure for driving performance. In our case, a probe reaction time task might still have shown an effect of multi-tasking, but the effect of increasing the headway is precisely to compensate for slower response times in case of multi-tasking. The current study partially agrees with the conclusion in [3] that "drivers adjust their headway when using a handheld phone, but not [or less so – *our insertion*] with a hands-free phone" (p. 1283). This observation is in agreement with the fact that the subjectively experienced workload for speech-enabled interaction is less for with manual interaction, leaving more mental resources for adequate reaction to the traffic situation.

6. ACKNOWLEDGMENTS

The research was made possible through a grant of the Dutch Ministry of Economic Affairs through grant G0102347. We

thank Eduard Helmann, Bart Naaijken and Josha Woudstra for assistance in the data collection.

7. REFERENCES

- [1] Alm, H. and Nilsson, L. (1995). The effects of a mobile telephone task on driver behaviour in a car following situation, *Accident Analysis & Prevention* 27, 707-715
- [2] Baron, A. and Green, P. (2006) Safety and usability of speech interfaces for in-vehicle tasks while driving: A brief literature review. University of Michigan, Transportation Research Institute: Technical Report UMTRI-2006-5.
- [3] Caird, J.K., Willness, C.R., Steel P. and Scialfa, C. (2008) A meta-analysis of the effects of cell phones on driver performance. *Accident Analysis & Prevention*, 40(4): 1282-1293
- [4] Hamish-Jamson, A.H., Westerman, S.J., Hockey, G.R.J. and Carsten, O.M.J (2004). Speech-Based Email and Driver Behavior: Effects of In-vehicle Message System Interface. *Human Factors*, 46(4), pp. 625-639
- [5] Kubose, T.T., Bock, K., Dell, G.S., Garnsey, S.M., Kramer, A.F. and Mayhugh, J. (2006), The effects of speech production and speech comprehension on simulated driving performance. *Applied Cognitive Psychology*, 20, 43-63.
- [6] Lee, J.D., Caven, B., Haake, S., Brown, T.L. (2001). Speech-Based Interaction with In-Vehicle Computers: The Effect of Speech-Based E-Mail on Drivers' Attention to the Roadway. *Human Factors*, 43, 631-640
- [7] McCallum, M.C., Campbell, J.L., Richman, J.B. and Brown, J.L. (2004) Speech Recognition and In-Vehicle Telematics Devices: Potential Reductions in Driver Distraction. *International Journal of Speech Technology* 7, 25-33
- [8] Owens, J., McLaughlin, S., and Sudweeks, J., (2010) "On-Road Comparison of Driving Performance Measures When Using Handheld and Voice-Control Interfaces for Mobile Phones and Portable Music Players," *SAE Int. J. Passeng. Cars - Mech. Syst.* 3(1):734-743, doi:10.4271/2010-01-1036.
- [9] Ranney, T.A. (2008) Driver distraction: a review of the current state-of-knowledge. Washington, D.C.: U.S. Dept. of Transportation, National Highway Traffic Safety Administration. <http://permanent.access.gpo.gov/lps97805/810787.pdf>
- [10] Schreiner, C. (2011) Widespread use of SMS and Email while driving. Strategy Analytics. Report. <http://www.strategyanalytics.com/default.aspx?mod=report&abstractviewer&a0=6149>
- [11] Waard, D. de, Hernandez-Gress, N., Brookhuis, K.A. (2001). The feasibility of detecting phone-use related driver distraction. *International Journal of Vehicle Design*, 26, 85-95.
- [12] Zijlstra, F. R. H. (1993). Efficiency in Work Behaviour: A Design Approach for Modern Tools. Ph. D. Thesis, Delft University of Technology.