Accident Analysis and Prevention 109 (2017) 18–28

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Full length article

Driver compliance to take-over requests with different auditory outputs in conditional automation

Yannick Forster[a,*], Frederik Naujoks[b], Alexandra Neukum[c], Lynn Huestegge[d]

[a] Wuerzburg Institute for Traffic Sciences (WIVW GmbH), Robert-Bosch-Str. 4, Veitshoechheim, 97209, Germany
[b] Interdisciplinary Center for Traffic Sciences (IZVW), University of Wuerzburg, Wuerzburg, 97070, Germany
[c] Interdisciplinary Center for Traffic Sciences (IZVW), University of Wuerzburg, Wuerzburg, 97070, Germany
[d] Institute of Psychology, University of Wuerzburg, Wuerzburg, 97070, Germany

ARTICLE INFO

Keywords:
Highly automated driving
Conditional automation
Human machine interface
Human-machine interaction

ABSTRACT

Conditionally automated driving (CAD) systems are expected to improve traffic safety. Whenever the CAD system exceeds its limit of operation, designers of the system need to ensure a safe and timely enough transition from automated to manual mode. An existing visual Human-Machine Interface (HMI) was supplemented by different auditory outputs. The present work compares the effects of different auditory outputs in form of (1) a generic warning tone and (2) additional semantic speech output on driver behavior for the announcement of an upcoming take-over request (TOR). We expect the information carried by means of speech output to lead to faster reactions and better subjective evaluations by the drivers compared to generic auditory output. To test this assumption, N = 17 drivers completed two simulator drives, once with a generic warning tone (‘Generic’) and once with additional speech output (‘Speech + generic’), while they were working on a non-driving related task (NDRT; i.e., reading a magazine). Each drive incorporated one transition from automated to manual mode when yellow secondary lanes emerged. Different reaction time measures, relevant for the take-over process, were assessed. Furthermore, drivers evaluated the complete HMI regarding usefulness, ease of use and perceived visual workload just after experiencing the take-over. They gave comparative ratings on usability and acceptance at the end of the experiment. Results revealed that reaction times, reflecting information processing time (i.e., hands on the steering wheel, termination of NDRT), were shorter for ‘Speech + generic’ compared to ‘Generic’ while reaction time, reflecting allocation of attention (i.e., first glance ahead), did not show this difference. Subjective ratings were in favor of the system with additional speech output.

1. Introduction

Automated driving systems are on the doorstep of the consumer market (Neville A Stanton et al., 2015). Conditionally Automated Driving (CAD) will soon follow already commercially available Partially Automated Driving systems. CAD characterizes systems are designed to assume vehicle control without the need for the human driver to continuously monitor the system. The driver is thus free to engage in non-driving related tasks (NDRT) such as writing emails or reading a newspaper. However, the driver is still required to be available in case the system exceeds its operational limits. According to SAE (2016), such a CAD system can be classified as a level 3 system.

Taking over vehicle control in such situations can be a demanding task for the human driver as automation removes drivers from both the physical and cognitive control loops, and he/she has to switch from executing an NDRT to manual driving within a relatively short timeframe. These so-called control transitions have thus attracted considerable research interest recently, focusing mostly on imminent situations in which manual vehicle control has to be regained within a few seconds (Flemisch et al., 2012; Gasser and Westhoff, 2012; Gold et al., 2013; Merat et al., 2014; Naujoks et al., 2014) but also on non-critical transitions of control with a large time budget up to 30 s for the drivers to take over vehicle control (Eriksson and Stanton, 2016; Payre et al., 2016). Examples for system limits are for example a broken vehicle on the lane ahead (Gold et al., 2013; Radlmayr et al., 2014), missing lane markings, emerging secondary lanes or a construction site with offset of lane markings (Forster et al., 2016). This paper seeks to extend existing findings by investigating how the safety of control transitions from automated to manual driving can be enhanced by the implementation of a visual auditory HMI that integrates data from car-

---

[*] corresponding author.

E-mail addresses: forster@wivw.de (Y. Forster), naujoks@psychologie.uni-wuerzburg.de (F. Naujoks), neukum@psychologie.uni-wuerzburg.de (A. Neukum), lynn.huestegge@uni-wuerzburg.de (L. Huestegge).

http://dx.doi.org/10.1016/j.aap.2017.09.019
Received 18 February 2017; Received in revised form 26 August 2017; Accepted 22 September 2017

to-car or car-to-X communication with vehicle localized environmental perception (Naujoks et al., 2015a; Rauch et al., 2012).

With such enhanced environmental perception, information about possible system limits (e.g., work zones, missing line markings, etc.) is – in principle – available early and information about upcoming mandatory transitions can be presented to the driver well in advance. However, there is a dearth research about how a suitable human-machine interface (HMI) for this purpose should be designed. We propose an HMI that integrates relevant information for a successful transition of control (Naujoks et al., 2016a) provided by cooperative perception technology into a visual interface and investigate how additional speech output can enhance the effectiveness of the proposed HMI using a motion-based simulator.

1.1. Background: supporting drivers through effective HMI communication

The transition from automated mode where the driver may be engaged in an NDRT to manual mode where she/he has to engage in safety relevant driving behavior can be characterized as a switch from a task A (i.e., NDRT) to another task B (i.e., full manual vehicle control). Research in the field of cognitive psychology has shown that any task switch (when compared with task repetitions) is accompanied by performance costs such as increased response times and error rates. These performance costs are either due to a time-consuming reconfiguration of mental task representations (e.g., assembling new stimulus-response rules) or due to conflict based on persisting activation of a previous task set after switching to a new task (Allport et al., 1994; Rogers and Monsell, 1995). Applying these cognitive psychological research findings to the area of CAD yields the prediction that by switching into a manual driving mode, drivers may also be prone to performance decrements. Any opportunity to prepare for this switch (e.g., by providing timely and maximally specific information regarding an upcoming switch) should be expected to counteract such performance costs (e.g., see Kiesel et al. (2010), for beneficial effects of advance preparation on task switching performance). The role of task switching in the context of driving automation has been discussed by Lorenz et al. (2015). The authors propose a three-staged process for the driver to get back in the loop. At first the driver has to allocate his/her attention away from the NDRT towards the relevant stimulus (e.g., head-up display or vehicle surroundings). Acquiring situation awareness and decision making, form the second step. Eventually the driver has to execute the maneuver that has the highest probability of success in the particular situation.

Thus, to ensure safety of control transitions during automated driving, there is a pressing need of investigation of HMI solutions for automated driving, that prepare the human driver as good as possible for regaining manual vehicle control. Consequently, besides information about the current status of the CAD system, drivers should be provided with sufficient information about upcoming events and actions by the system. Early information about upcoming conflict situations, so called *advisory warnings* (Lenné and Triggs, 2009) can be presented well in advance (Seeliger et al., 2014) without a need for the human driver to immediately react to the warning but rather to be ready to respond (Naujoks et al., 2015a). Wiedemann et al. (2015) have found evidence, that early announcements of the outline of interaction scenarios are beneficial for driver performance.

Relevant information can be carried not only through the visual channel alone but can further be supported through the auditory channel. Auditory interfaces provide advantages such as (1) omnidirectionality, (2) the possibility to be perceived at almost all times, (3) transient sound and (4) the possibility for humans to selectively focus on the content (Bazilinskyy and de Winter, 2015). In manual driving, speech based systems in particular can be beneficial for driver performance resulting in lower lane variation and steadier speed (Barón and Green, 2006; Neville Anthony Stanton and Edworthy, 1999). To date there has been little research on the implementation and design of semantic auditory interfaces for the transition of control in automated driving. Additional speech output could be very beneficial when a larger window of time is left to react to the particular system limit – which is precisely the benefit of cooperative perception technology – for several reasons.

- First, the CAD system could communicate its intention and status more explicitly and clearly so that drivers attain higher mode awareness and thus react appropriately (Naujoks et al., 2016b).
- Second, speech output could reduce visual workload during driver-automation interaction (Bazilinskyy and de Winter, 2015). For example, speech based auditory interfaces reduce the necessity to look away from the road to find out about information content of a system communicated through its HMI (Alvarez et al., 2011). A lot of information that needs to be decoded from the visual part of the automated driving HMI could be delivered quickly through a speech based system alone and therefore reduce visual workload and support driver comfort (Bazilinskyy and de Winter, 2015). Furthermore, Naujoks et al. (2017) have found first evidence that drivers tend to look less towards a visual HMI and the vehicle surroundings during the independent execution of a maneuver by the CAD function when additional speech output is presented.

Third, additional speech output could possibly speed up responses to TORs and enhance the safety of control transitions. Results by Politis et al., 2015 support the assumption that semantic speech output could be important for take over quality (i.e., lane keeping behavior). However, the authors could provide no definite guideline. They conclude that an advantage for superiority of language-based cues compared to abstract ones seemed to be present.

Naujoks et al. (2016b) have examined the effects of a visual HMI supplemented by speech output for the independent execution of a maneuver by the automated system. The study's results support the assumed superiority of semantic auditory output over generic information/warning tones to communicate upcoming automated maneuvers. The results, however, only apply to scenarios where the CAD system can handle the upcoming event by itself. Thus, effects on drivers' compliance to early system indications and their subjective evaluations during a transition from automated to manual are investigated in the present study.

Besides the benefits of auditory and especially speech based interfaces, it is also important to mention the potential downsides of these. There is the nuisance factor that could arise under frequent presentation of speech output or the erroneous perception of an indication as a warning which in turn results in mode confusion (Bliss and Acton, 2003; Cotté et al., 2001). For example, it has been shown that false alarms reduce compliance with urgent visual-auditory car-to-X-warnings, but that this compliance decrement can be prevented by using less urgent visual car-to-X-warnings (Naujoks et al., 2016c).

1.2. Research question

There are many challenges, which remain to be overcome before CAD can be commercially accessible without doubt about driver and passenger safety. All the factors mentioned above explicitly point towards the importance of take-over scenarios. The study was designed to advance knowledge about TORs in two aspects that have not been studied extensively yet.

First, it was of interest to develop an HMI containing useful information about upcoming transitions to manual driving that can be provided by cooperative perception technology. We thus followed the concept of so-called *situation announcements* that was put forward by Wiedemann et al. (2015). To date, research has widely ignored long take-over times of up to 20 s, which might be possible through cooperative perception. There is first evidence, that drivers take longer to resume control in non-critical scenarios when there is no time pressure at the onset of the TOR (Eriksson and Stanton, 2016). An exploratory approach shall bring forth further empirical evidence about driver reaction times under a window of time as long as 20 s.

Second, a previously developed visual HMI (Naujoks et al., 2016a) was supplemented with different auditory components. The research question here is whether there is a superiority of additional semantic speech output over generic auditory warnings for visual-auditory TORs. Building on the work of Naujoks et al. (2016b), we applied speech output to TOR scenarios. The applied speech output in this study did not have any information content going beyond that of the visual HMI. However, the semantic content of a message is suspected to be processed faster and understood more easily when presented additionally by means of speech output via the auditory channel compared to mere generic information on the auditory channel. To assess drivers' compliance with the CAD function during a TOR scenario, we applied the common definition from Meyer (2004) who stated that compliance is considered *"[…] the response when an operator acts according to a warning signal and takes some evasive action […]"* (p. 199). From here, we defined compliance as the reaction time until (1) first gaze towards the road ahead, (2) the drivers putting away a manual NDRT and have their hands free for the take-over action, (3) putting their hands on the steering wheel and (4) deactivation of the CAD system. To evaluate effectiveness of the two different HMIs we additionally compared *usefulness, visual workload, usability* and *acceptance* between the two experimental conditions.

2. Materials and methods

2.1. Driving simulation

The study was conducted in the motion-based driving simulator at the Wuerzburg Institute for Traffic Sciences (WIVW GmbH, see Fig. 1) using the institute's simulation software SILAB. The integrated vehicle's console contains all the necessary instrumentation and is identical to a production type BMW 520i with automatic transmission. To simulate a realistic steering torque, a servo motor based on a steering model was used. The motion system uses six degrees of freedom and can briefly display a linear acceleration up to $5\ \mathrm{m/s^2}$ or $100^\circ/\mathrm{s}^2$ on a rotary scale. It consists of six electro-pneumatic actuators (stroke ± 60 cm; inclination ± 10°). Three LCD projectors are installed in the dome of the simulator and provide the projection. Three channels provide a 180° screen image. LCD displays serve as exterior and interior mirrors.

2.2. Conditional automation specifics

The conditional automation could be activated by simultaneously pushing two buttons on the steering wheel. The two buttons could easily be reached with the thumbs when holding the steering wheel at "ten and two" position. Deactivation was possible by simultaneously pushing the same buttons again, by braking (i.e., 10% of the braking pedal) or by applying a torque on the steering wheel (i.e., overcoming 7 Nm with 45° change in the current steering wheel angle) so that the automated steering function was overridden. Once deactivated, the vehicle would not decelerate but drift slowly in the direction of the current steering wheel direction. When active, the CAD system executed both longitudinal and lateral vehicle guidance. A target speed was pre-set to 130 km/h. If possible, the CAD system accelerated until target speed was reached. In case there were slower vehicles ahead, it would stay within an acceptable distance by regulating speed down (pre-set time-headway: two seconds). In ambiguous situations (secondary lane markings) it could not guarantee safe vehicle guidance and indicated this by an announcement and a subsequent TOR (see Section 2.4 scenario layout and 2.5.1 *visual interface*).

2.3. Design and procedure

The study employed a within-subject design with two levels of auditory HMI output. Participants completed one drive with a system that provided unspecific warning tones (condition: 'Generic') while the other drive was completed with a system that additionally provided semantic speech output (condition: 'Speech + generic'). The generic warning tone was presented in both experimental conditions to make sure that possible differences in reaction times and subjective evaluations can be directly attributed to the semantic content of the additional speech output. Participants were randomly assigned to either the 'Speech + generic' or the 'Generic' condition in the first drive(Table 1). Each drive contained three automated system maneuvers and one take-over scenario (see Fig. 2) while driving was conditionally automated. Results of the system maneuver scenarios are reported in Naujoks et al. (2017). One drive lasted about 10 min and covered four scenarios where participants had to interact with the automated system. There were about two and a half minutes between each of the interaction scenarios. The order of interaction scenarios was balanced between participants. In each drive, they were asked to work on an NDRT in the form of reading pre-selected articles in a weekly German print magazine as instructed by the experimenter. To make sure drivers would actively engage in the NDRT, participants were told that they would be examined about the articles' content at the end of the experiment. The experimenter explicitly pointed out that the CAD system would take care of lateral and longitudinal vehicle guidance and that it would inform them with sufficient notice should manual control be needed.

Upon arrival, participants were welcomed and gave informed consent. The experimenter explained that the goal of the study was to test a new developed HMI for CAD in two drives. To get unbiased evaluations, drivers did not receive any preliminary information about the visual nor the auditory HMI. At the beginning of each drive participants had about 2 min where they could familiarize themselves with the simulator and the system. At the beginning of each drive all drivers were told to activate and deactivate the CAD system once. There is accumulating evidence that prior familiarization with system capabilities and limitations has a strong influence on driver's take-over behavior and system evaluations of automated driving systems (Hergeth et al., 2016). During the familiarization, drivers did not experience a TOR and evaluations of the TOR as well as behavioral measures should not be affected. Practicing the deactivation by pressing both buttons

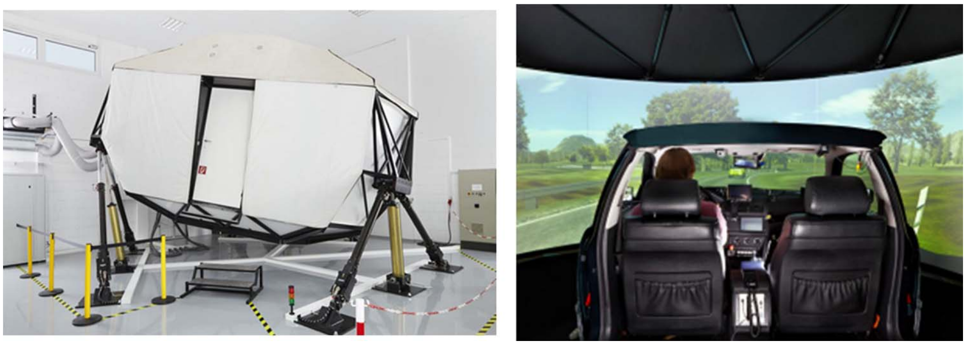


Fig. 1. The WIVW moving based driving simulator. Hexapod movement system (left) and Simulator interior with vehicle mock-up and video projection (right).

Table 1
Schematic outline of experimental procedure in the two drives for a participant with 'Generic' in the first drive and 'Speech + generic' in the second drive.

Drive 1: (e.g. 'Generic')					break	Drive 2: (e.g. 'Speech + Generic')				
Familia-rization	Syst. Man.1	TOR	Syst. Man.2	Syst. Man.3		Familia-rization	Syst. Man.1	TOR	Syst. Man.2	Syst. Man.3

simultaneously was necessary to avoid responses where drivers would try to only press one button and the CAD system is not deactivated.

2.4. Scenario layout

One experimental drive consisted of four scenarios in which the driver had to interact with the system. Three of these scenarios were system maneuver scenarios (i.e., independent execution of a maneuver by the CAD system) and one scenario was a take-over scenario (i.e., emerging secondary lanes). Since the present paper focuses on the investigation of visual-auditory HMIs for TORs, results for the system maneuver scenarios are reported in Naujoks et al. (2017). The *Secondary Lane* scenario differed from the system maneuver scenarios because the CAD system was not capable of safely guiding the vehicle and drivers were informed through the HMI as described in Section 2.5. Here, yellow lane markings besides the white lane markings emerged (see Fig. 2). The white markings turned slightly to the right and ended with a bumper after about 200 m. The yellow lane markings turned slightly to the left and color was switched back to white 200 m away. Participants were not given information about the system's behavior in case of non-responding to the TOR. The TOR was triggered at a constant velocity of 130 km/h with a time budget of 20 s left to the emerging secondary lanes. In case the driver did not take over vehicle control before reaching the yellow lane markings, the system would drive through the scenario with constant velocity and adapt position to the yellow lane markings.

2.5. Human-Machine interface

2.5.1. Visual interface

The visual interface was depicted in the head-up display. Fig. 3 shows the visual HMI in automated mode. The numbers in round grey shapes are explained below. The basic blue color layout indicates that the automation is active and working reliably. Blue lane symbols and the blue rectangle in front of the host vehicle show that both lateral and longitudinal vehicle guidance are executed by the automated system. Set speed (1a) and current speed (2) are displayed below. This part of the HMI resembles that of existing HMI solutions for ACC with additional steering assistance (Naujoks et al., 2015b). Information about upcoming events is depicted by a symbol on the right (4) and additionally through a textbox above (3). Distance to the traffic event is shown by a decreasing horizontal bar (5). Automated speed adaptation is depicted by marking a line through the set speed (1b) until the speed limitation event is over.

20 s before arriving at the system limit (i.e., emerging secondary lanes), the visual HMI displayed an announcement of the upcoming scenario and the color of the visual HMI switched from blue to orange to indicate that the CAD system would not be capable of managing the scenario (Fig. 4 left). Following this announcement, a non-imminent indication to take over manual control surrounded by a circle indicating a countdown was displayed (so-called Soft TOR; Fig. 4, middle) (Naujoks et al., 2015b) when there was still a comfortable time budget to take over manual driving. Eventually an imminent TOR (so-called Hard TOR, Fig. 4, right) (Naujoks et al., 2015b) followed the soft TOR in case the driver does not react in time to the first TOR. The aim of this graded take-over procedure is to provide enough time and sufficient information about the upcoming system limit to the driver, so that she/he can redirect her/his attention to the driving task and get ready to drive manually (Wiedemann et al., 2015). Durations for each stage were 7 s (announcement and soft TOR) and 6 s (hard TOR) adding up to a total time budget of 20 s. For a substantive description of the visual HMI see Naujoks et al. (2016a) and for an expert evaluation of the visual HMI preceding the present study see Forster et al. (2016).

2.5.2. Auditory interface

There were two different auditory HMI outputs. Generic output was presented as two high frequency warning tones before the announcement stage (duration: 150 ms; frequency: 1000 Hz sinus; interval: 150 ms). To create a more naturalistic speech output, a female voice (Bazilinskyy and de Winter, 2015) was recorded using a Dictaphone to get semantic speech output instead of generating machine-based speech output. The additional speech output verbalized the information provided by the visual HMI and was presented once at the onset of the announcement stage. Speech output during the drive was presented in the German language. The presented wording translated from German into English was: "Unclear lane ahead, please take over soon." The duration of this output presented in German language was 3.5 s.

Additionally, there were three high frequent warning tones (duration: 100 ms; frequency: 750 Hz sinus; interval: 100 ms) in both auditory HMI condition preceding the hard TOR to encourage immediate action. There was no auditory output at the onset of the soft TOR.

2.6. Dependent variables

2.6.1. Quantitative take-over behavior

For the present HMI it is of interest, in which HMI stage drivers would begin to actively get back in the loop (Lorenz et al., 2014). Therefore, we analysed the HMI stage in which participants would put their hands back on the steering wheel. Performance in TOR scenarios is most often assessed by different reaction times measures (Gold et al., 2013; Naujoks et al., 2012, 2015a). Since the present scenario was non-critical and drivers rather had to stabilize lateral and longitudinal vehicle guidance, common intervention metrics such as steering input larger than 2 ° or braking pedal position of more than 10% (Gold et al., 2013) cannot be considered here. When comparing two different auditory HMIs while both visual HMI and vehicle automation (i.e., the non-critical TOR) itself is kept constant, shorter reaction times should indicate more safety and thus a more appropriate HMI design for this

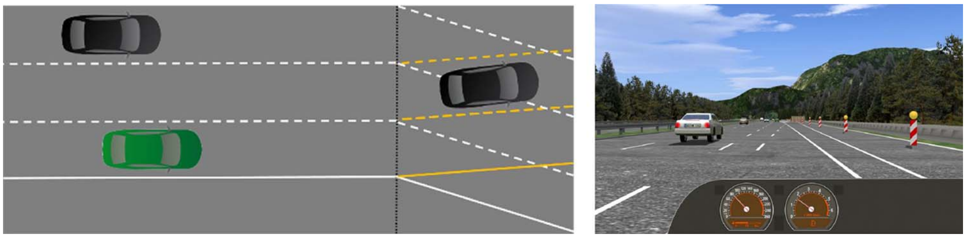


Fig. 2. Schematic outline of secondary lane scenario (left) with host vehicle (green) and surrounding traffic (black) and SILAB view (right). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

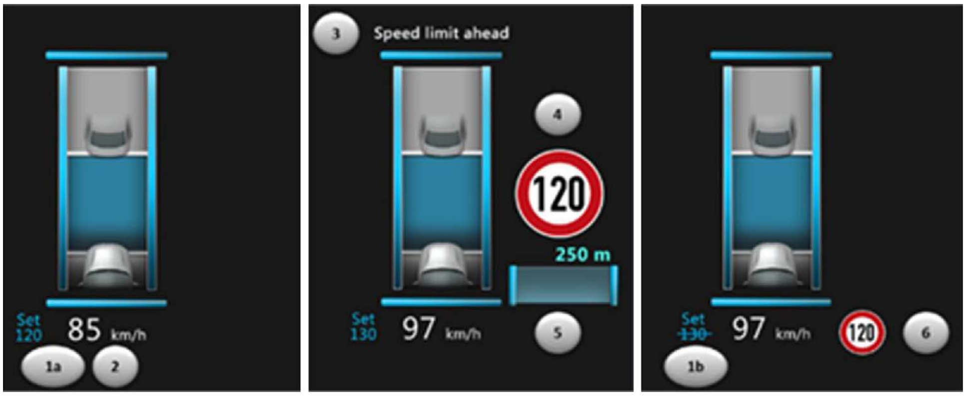


Fig. 3. Human-Machine Interface for automated mode and independent system maneuver execution.

particular interaction scenario. Reaction times can be calculated for different steps of the take-over process like gaze reaction (i.e., first fixation on the road after the onset of the TOR; Gold et al., 2013), hands-free (i.e., the NDRT has been terminated and both hands are available for the driving task), hands-on (i.e., at least one hand is touching/grabbing the steering wheel; Gold et al., 2013; Naujoks et al., 2014) and intervention or deactivation of the system (e.g., participant pushing the buttons; Gold et al., 2013). The first gaze on the road was recorded through video coding by a data reduction specialist. The video recordings had a resolution of 1280 × 720 Pixels. The video coding tool provided the possibility of slowing down playback speed to a minimum rate of 10 Hz. Accurate temporal resolution of the coded data was therefore supported. The reference point of the reaction times was the onset of the announcement stage. Pressure sensors that were installed underneath the surface of the steering wheel measured hands-on signals. The reaction time parameters, a brief description and the source of measure are shown in Table 2.

2.6.2. *Subjective evaluation*

To get an overall impression of the usability of the system in the two experimental conditions, participants completed a comparative German version of the System Usability Scale (SUS) (Brooke, 1996) after the second drive. A forced choice item operationalized acceptance of each of the system. Participants had to answer whether they preferred the system with additional speech output or the system with an unspecific warning tone during a follow-up interview. Furthermore, drivers were asked to evaluate usefulness of both visual and auditory system indications and visual workload from retrieving information from the HMI. These evaluations were recorded just after completing the TOR scenario during the drive. A 15 point Likert scale ranging from 1 (“*very little*”) to 15 (“*very much*”) with an additional category of 0 (“*not at all*”) was used. The 15 point scale includes five verbal categories (“*very little*”, “*little*”, “*medium*”, “much”, “*much*”) which are divided up into three sub-categories. The respective items were as shown in Table 3.

Table 2
Take-Over parameters, description and source of measure.

Parameter	Description	Source of measure
Button	Participant pushing the buttons to deactivate the CAD system	Signal from buttons in simulation
Hands-on	At least one hand touching/ grabbing the steering wheel	Signal from pressure sensors in simulation
Hands-free	NDRT interrupted, both hands available for the driving task	Video coding
Gaze-reaction	first fixation on the road after TOR onset	Video coding

Table 3
Measure and corresponding item wording of subjective evaluation.

Measure	Item	Range
Usefulness (visual)	How helpful was the display?	Likert [0–15]
Usefulness (auditory)	How helpful was the auditory output?	Likert [0–15]
Visual workload	How much attention did you pay to the display?	Likert [0–15]

2.6.3. *Statistical procedure*

Data was pre-processed using Mathworks Matlab and statistical tests were calculated using IBM SPSS Statistics Version 23. Before applying methods of inferential statistics, the obtained data was analysed descriptively. Depending on the respective inferential test, data was screened using a Kolmogorov-Smirnov test for normal distribution of data. Given the experimental setup (within-subject design) and descriptive data, inferential methods were applied as shown in Table 4.

2.7. *Participants*

A total of $N = 17$ participants completed the two drives. 10 participants were male and 7 female with a mean age of 29.0 years ($SD = 8.12$ years). The oldest driver was 56 years old and the youngest

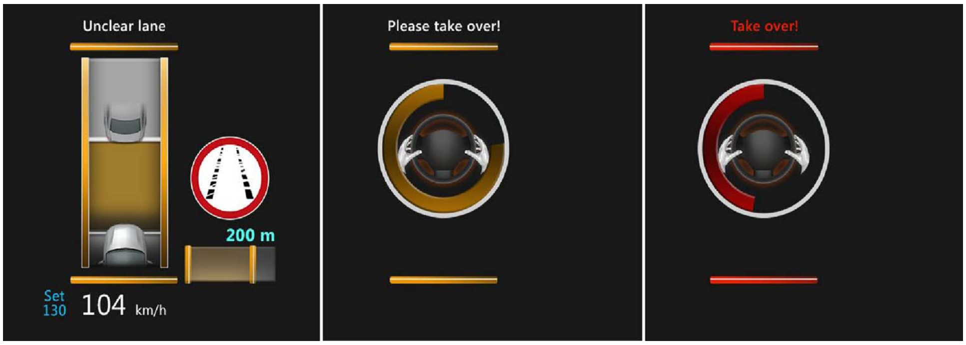


Fig. 4. HMI for Take-Over scenario: Announcement, soft TOR, hard TOR (from left to right).

Table 4
Measures and according statistical tests.

Measure	Statistical test
- Reaction Times	- Kolmogorov-Smirnoff (KS) test - 2 × 2 repeated measures ANOVA
- hands-on signals	- descriptive comparison
- Acceptance	- descriptive comparison
- Usability (SUS)	- *t*-test for paired samples
- Usefulness (visual & auditory HMI), visual workload	- *t*-test for paired samples

driver 22 years old. All participants were recruited from the test driver panel of the WIVW and had taken part in an extensive driving simulator training (Hoffmann et al., 2003) prior to the study.

3. Results

3.1. Reaction times

The number of hands-on signals was counted for the three stages of the HMI display. There were only a few drivers in both conditions who took over vehicle control when the hard TOR was displayed (‘Speech + generic’: $n = 3$; ‘Generic’: $n = 3$). While most drivers took over control in ‘Speech + generic’ ($n = 8$) in the announcement stage, the majority of drivers waited until the soft TOR was presented in the ‘Generic’ condition ($n = 8$).

There was $n = 1$ driver in the ‘Speech + generic’ condition that did not deactivate the system but drove through the TOR scenario with hands on the steering wheel. Reaction time data of this particular participant was not used for inferential data analysis due to the experimental within design plan. All other drivers deactivated the CAD system by simultaneously pressing the buttons on the steering wheel.

Prior to inferential analysis, the distribution characteristics of the reaction time data was examined. Fig. 5 shows the frequency distributions of the different reaction time parameters. It is important to note that there was one driver in the ‘Generic’ condition, who deactivated the system still holding the magazine in hands. This particular high reaction time value (25.3 s) is not displayed in the frequency plot below to keep the x-axis as long as the time budget (i.e., 20 s) was.

While the distribution for ‘Speech + generic’ seemed to be positively skewed and unimodal, the distribution for ‘Generic’ reaction times was bimodal with one peak at the lower end of the distribution and one peak close to the mean reaction time. Applying the Kolmogorov-Smirnoff test for violation for normality assumption, reaction times of button (KS d = 0.22), hands-on (KS d = 0.28) and hands-free (KS d = 0.22) in the ‘Speech + Generic’ condition revealed a significant difference to a normal distribution, while the normality assumption for the particular reaction times in the ‘Generic’ condition could not be rejected. Reaction times for gaze reaction followed a normal distribution in both experimental conditions.

Inferential (i.e., df_1, df_2, F-, p-, η^2-value) statistics for main effects of both order and HMI conditions and their interaction for the four reaction time measures can be seen in Table 5. Fig. 6 shows the main effect for auditory HMI condition for each reaction time. Besides the mentioned dropout all drivers managed to safely take over vehicle control within the time budget of 20 s. We applied 2 × 2 repeated measures ANOVA with order of experimental conditions as between factor and auditory HMI output as within factor. Reaction time for the first gaze towards the road did not show an advantage of ‘Speech + generic’ ($M = 1.42$, $SD = 0.42$) over ‘Generic’ ($M = 1.29$, $SD = 0.28$) condition. Hands-on reaction times in the ‘Speech + generic’ condition ($M = 5.66$, $SD = 1.96$) were significantly shorter than in the ‘Generic’ condition ($M = 7.84$, $SD = 3.16$). Hands-free reaction times in the ‘Speech + generic’ condition ($M = 6.13$, $SD = 2.79$) were also significantly shorter than in the ‘Generic’ condition ($M = 9.12$, $SD = 5.50$). Results showed that reaction times for deactivation of the system by pressing the buttons did not show a statistically significant difference (‘Speech + generic: $M = 8.26$, $SD = 3.01$; ‘Generic’: $M = 9.86$, $SD = 3.84$).

As participants were randomly assigned to one of the two experimental conditions in the first drive (and experienced the respective other condition in the second drive) and there were only two takeovers for each participant, we tested for order effects. Results showed that there are interaction effects for button press, hands-on and hands-free, but not for gaze reaction. Consequently, we applied paired samples *t*-tests for each reaction time parameter for each of the experimental order groups (i.e., ‘Speech + generic’ first vs. ‘Generic’ first). Results show that the significant interaction effects are due to the group, which

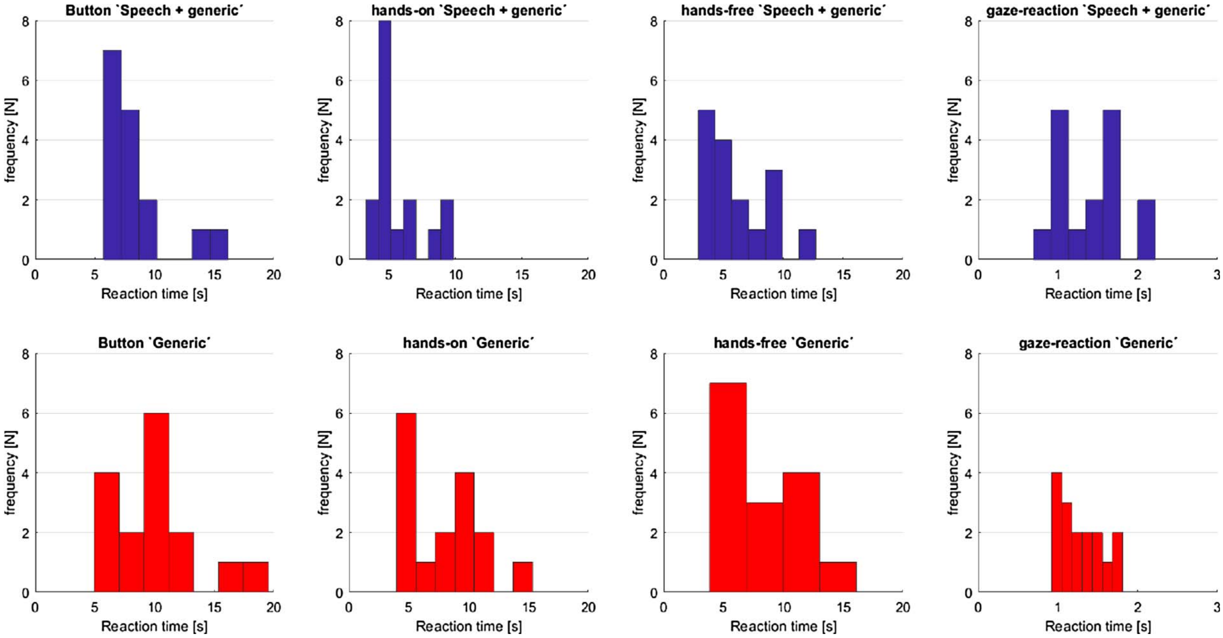


Fig. 5. Frequency distributions for reaction time parameters (blue: ‘Speech + generic’, red: ‘Generic’).

Table 5
Inferential results of repeated measures ANOVA for reaction time parameters. Statistically significant parameters are colored in grey.

Measure	Effect	df1	df2	F	p	η^2
Button	Order	1	14	0.542	.474	.037
	HMI condition	1	14	2.55	.133	.154
	Order * HMI condition	1	14	16.00	.001	.533
Hands-on	Order	1	14	1.50	.241	.097
	HMI condition	1	14	20.14	.001	.590
	Order * HMI condition	1	14	10.62	.006	.431
Hands-free	Order	1	14	2.91	.110	.172
	HMI condition	1	14	6.84	.020	.328
	Order * HMI condition	1	14	7.00	.019	0.333
Gaze-reaction	Order	1	14	1.21	.289	.080
	HMI condition	1	14	1.11	.311	.073
	Order * HMI condition	1	14	.70	.418	.047

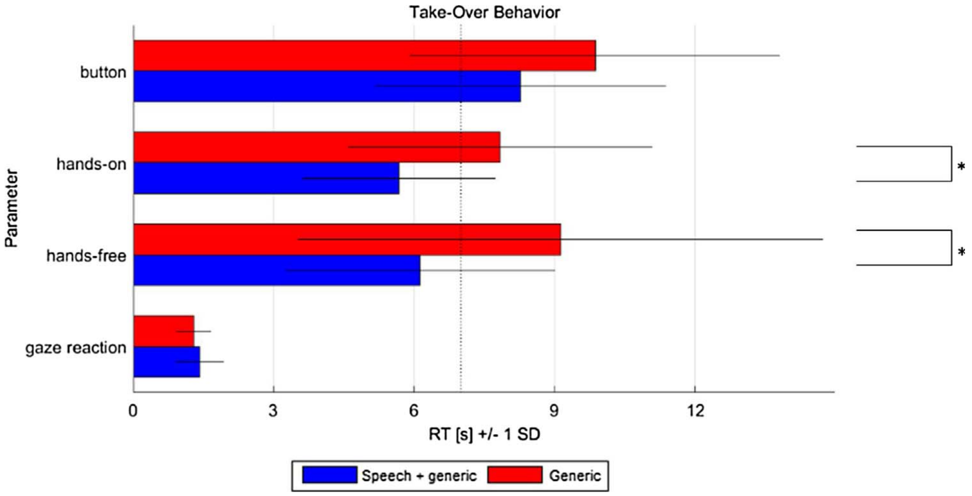


Fig. 6. Take over cascade for mean reaction times and SD of button, hands-on, hands-free and gaze reaction (from top to bottom) by auditory HMI output. Dashed vertical line indicates onset of soft TOR. * indicates a statistically significant difference.

experienced the generic warning tone in the first drive (Table 6). Drivers who experienced the generic tone in the first drive thus reacted significantly faster at terminating the NDRT, putting their hands at the steering wheel and deactivating the CAD system when experiencing speech output in the second drive. Consequently, drivers experiencing speech output in the first drive did not show a large difference between their first reaction time and the particular second reaction time with generic output.

3.1.1. Subjective evaluation

In terms of overall acceptance, participants clearly favored 'Speech + generic' over 'Generic' auditory HMI output (16 out of 17 participants). The only participant in favor of the 'Generic' auditory HMI indicated that additional speech output could be annoying over time when occurring too frequently.

Primarily, SUS scores were calculated as described in (Brooke, 1996). A *t*-test for paired samples showed, that drivers clearly favored the 'Speech + generic' system (M = 91.62, SD = 8.79) over the 'Generic' system (M = 74.41, SD = 17.15; t(16) = 4.988, $p < 0.001$ d = 1.26, two-tailed) in terms of its overall usability.

Subjective ratings of the usefulness of the visual display (Fig. 7, left) were on a medium to high level for both 'Speech + generic' and 'Generic', according to the verbal anchors of the scale. On a descriptive level, there were higher scores for the 'Generic' condition compared to 'Speech + generic' which did not reach statistical significance. Usefulness of the auditory output (Fig. 7, middle) was rated significantly higher for the 'Speech + generic' condition compared to the 'Generic' condition. The semantic information provided by additional speech output seemed to be important for drivers to initiate, prepare and execute the take-over process.

Drivers rated visual workload on a high level for both conditions without reaching a significant difference. To derive information from the HMI, drivers in the 'Speech + generic' condition could use the auditory channel, whereas in the 'Generic' conditions they had to use the visual channel. Therefore, mean ratings of attention towards the visual display are slightly higher in the 'Generic' condition (see Fig. 7, right). Table 7 shows descriptive and inferential data for situation specific evaluation of the TOR scenario.

4. Discussion

The current study investigated whether additional speech output could improve driver system interaction in Take-Over scenarios of CAD systems. Reading articles in a print magazine, *N* = 17 participants completed two 10 min drives that included one take-over scenario each. One condition was presented with a generic warning tone (condition 'Generic') and one condition was presented with additional speech output in form of a female human voice ('Speech + generic'). We

Table 6
Inferential *t*-test results of reaction time parameters for 'Speech + generic' and 'Generic' by order. Statistically significant parameters are colored in grey.

Order	Measure	Condition	n	M	SD	t	df	p	d
'Speech + generic' first	Button	'Speech + generic'	7	9.41	4.17	1.644	6	.15	.62
		'Generic'	7	7.57	2.21				
	Hands-on	'Speech + generic'	7	5.67	2.10	1.07	6	.33	.40
		'Generic'	7	6.21	2.57				
	Hands-free	'Speech + generic'	7	5.98	2.31	.11	6	.92	.04
		'Generic'	7	5.95	2.70				
	Gaze reaction	'Speech + generic'	7	1.28	0.38	.18	6	.86	.07
		'Generic'	7	1.26	0.23				
'Generic' first	Button	'Speech + generic'	9	7.36	1.39	4.16	8	<.01	1.39
		'Generic'	9	11.65	3.98				
	Hands-on	'Speech + generic'	9	5.66	1.89	5.12	8	<.01	1.71
		'Generic'	9	9.10	3.10				
	Hands-free	'Speech + generic'	9	6.24	3.25	3.03	8	.02	1.01
		'Generic'	9	11.59	5.97				
	Gaze reaction	'Speech + generic'	9	1.53	0.43	1.26	8	.24	.42
		'Generic'	9	1.32	0.32				

assessed different reaction time measures for the take-over process as well as subjective evaluations of acceptance, usability and usefulness of both visual and auditory HMI as well as perceived visual workload. Since semantic speech output has been shown to be beneficial for the effectiveness of indications by an automated system (Politis et al., 2015), we expected drivers to take over faster and evaluate the system as superior when additional speech output is presented compared to the presentation of merely generic auditory output.

4.1. Interaction behavior

We formulated two research questions regarding driver-system interaction behavior. First, we wanted to find out about potential benefits and describe interaction behavior with cooperative perception and a long window of time up to 20 s. Take-Over performance was assessed for four different measures of gaze reaction, hands-on, hands-free and system deactivation within the take-over process (Gold et al., 2013). Independent from the respective experimental condition, drivers showed fast reactions to the TOR by looking at the road scenery ahead. Drivers took most actions at the announcement stage and a few at the early soft TOR stage. As in many cases, they did not even wait until the hard TOR came up. For both experimental conditions, this suggests that the presentation of information with this particular HMI worked considerably well. Initial generic auditory output and the yellow coloring of the visual HMI followed by a soft TOR (if even necessary) seems to be an effective strategy to warn drivers that are currently engaged in a NDRT about upcoming control transitions.

Furthermore, frequency of hands-on signals for the three HMI stages revealed that independent from auditory HMI output condition, the majority of drivers took their hands back on the steering wheel during the announcement or soft TOR stage. This result is surprising since drivers knew through the countdown in the visual HMI (see Section 2.5.1 *visual interface*) that there was still time until the system limit would be reached. In contrast to studies where not all drivers successfully managed to regain manual control but collided with an object that they should have evaded (Gold et al., 2013; Naujoks et al., 2014), all

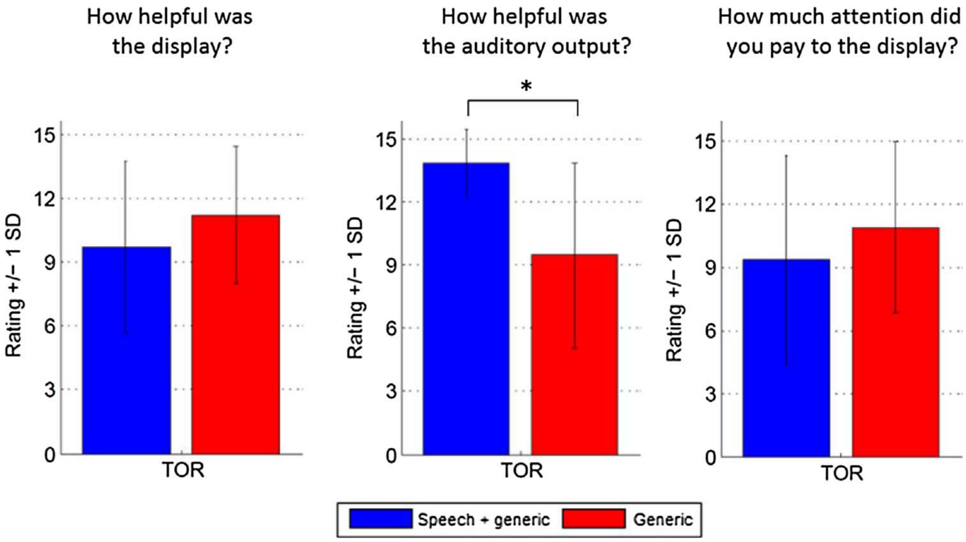
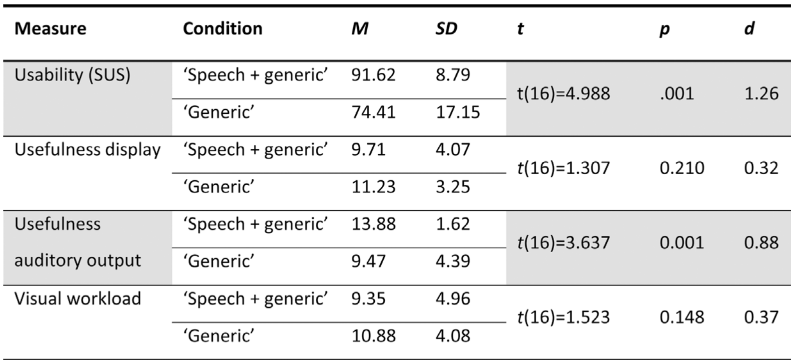


Fig. 7. Subjective evaluation for the TOR scenario: Mean usefulness ratings (± 1 SD) of the visual and auditory HMI components and visual workload (left to right). * indicates a statistically significant difference.

Table 7
Summary of descriptive and inferential data for subjective evaluations of the TOR scenario. Statistically significant variables are colored in grey.



Measure	Condition	*M*	*SD*	*t*	*p*	*d*
Usability (SUS)	'Speech + generic'	91.62	8.79	t(16)=4.988	.001	1.26
	'Generic'	74.41	17.15			
Usefulness display	'Speech + generic'	9.71	4.07	*t*(16)=1.307	0.210	0.32
	'Generic'	11.23	3.25			
Usefulness auditory output	'Speech + generic'	13.88	1.62	*t*(16)=3.637	0.001	0.88
	'Generic'	9.47	4.39			
Visual workload	'Speech + generic'	9.35	4.96	*t*(16)=1.523	0.148	0.37
	'Generic'	10.88	4.08			

drivers reacted in the desired manner to the system indication and took over manual control in time before the system's limit was reached. Still, the entire process lasted considerably long with a period of more than 4 s between the gaze reaction and the first actual intervention to the TOR (i.e., hands-on) and even longer until the final deactivation. This result suggests two important points:

- First, drivers were given enough time to put the magazine away in a timely comfortable manner, which in turn lowers the probability of potential risky steering or braking maneuvers that are necessary for time-critical TOR scenarios (Gold et al., 2013; Merat et al., 2014; Naujoks et al., 2014).
- Second, reaction time results showed that the transition from automated to manual driving occurred mostly in the first half of the available time budget of 20 s. With a lot of time available, all drivers took over safely within the time budget. Apparently, through a straightforward communication of the transition of driving task responsibility, *mode awareness* (Hakuli et al., 2012) was high for the present scenario.

The second research question was to investigate the benefits of additional speech output compared to the presentation of merely generic auditory output. The comparison of the two experimental conditions revealed significantly shorter hands-on and hands-free reaction times in the 'Speech + generic' condition compared to the 'Generic' condition. However, this overall difference was not observed for gaze reaction and deactivation of the CAD system. For the interpretation of the findings of shorter reaction times with additional speech output, one needs to consider the interaction of order of presentation and auditory HMI condition. It became apparent that drivers reacted faster in the 'Speech + generic' condition than in the 'Generic' condition when the HMI without additional speech output was presented in the first drive. This difference was not observed when speech output was presented in the first drive. Thus, the overall advantage of additional speech output over mere generic auditory output is due to the group with generic auditory output in the first drive and additional speech output in the second drive. The findings are discussed in detail below.

Take-Over behavior did not differ for the first reaction to the TOR (i.e., first glance to the road ahead). The process of allocation of visual attention in a TOR scenario (Lorenz et al., 2014) is guided by the initial system indication. From here, the result can be explained by the fact that both experimental conditions contained the generic warning tone which initiated an interaction between the driver and the system. Attention allocation was guided by the unspecific warning tone in both conditions in the same manner. The reaction time of the first gaze to the road ahead is therefore not discriminating between the two conditions and this research cannot recommend using this parameter for research on the compliance to visual-auditory TORs when a large time budget is available.

For subsequent reaction time measures of hands-free and hands-on, the 'Speech + generic' HMI was superior to the 'Generic' HMI with significantly shorter reaction times. This finding is in line with research from cognitive psychology indicating that specific preparation is better suited than generic preparation for counteracting adverse effects of task switches on performance (Kiesel et al., 2010). After attention had been allocated to the road ahead or the HMI, respectively, drivers had to process the information conveyed by the visual HMI (both experimental conditions) and additionally the auditory HMI ('Speech + generic'). This process corresponds to the second step (i.e., *building situation awareness*) in the transition from automated to manual mode (Lorenz et al., 2014) since attention had been allocated to the vehicle surroundings, but no evasive maneuver (e.g., steering, system deactivation) has been executed yet. Hands-free and hands-on reaction times reflect drivers' conclusion that they had to re-engage in the driving task and need to terminate the NDRT. Drivers were aware of the upcoming system limitation and initiated the take-over process by interrupting their NDRT engagement and putting their hands at the steering wheel earlier when additional speech output was presented. Speech output supported drivers in building situation awareness and get ready to intervene if necessary. Furthermore, with mean reaction times looming larger into the window of time than the presentation of speech output (i.e., 3.5 s), we can be sure that drivers' reaction towards the TOR are facilitated by the additional speech output. Had reaction times been shorter than the duration of the speech output, the effect could not have been attributed to the additional presentation of the semantic speech output.

The final step of getting back in the loop (Lorenz et al., 2014) is to deactivate the CAD system and stabilize longitudinal and lateral vehicle guidance. Here, drivers need to understand the situation and find the according pattern on the road ahead that was conveyed by the HMI. The actual intervention in this scenario (i.e., deactivation) was only facilitated in the group that experienced the generic auditory HMI output in the first drive. The significant interactions between auditory HMI output and order for the hands-free, hands-on and button press reaction time measures outline an important point. When drivers experienced additional speech output in the first take over scenario, they connected information from the auditory and visual HMI elements to take over control. When they experienced the TOR in the second drive with generic auditory information, they had already learned the visual HMI's content and showed similar reaction times compared to their first drive. Therefore, we suppose that the additional speech output presented in the first drive also facilitated reaction times in the second drive (so-called *carry-over effect*). Drivers exhibited the longest reaction times for generic output when presented in the first drive. The comparison with reaction times in the following drive with additional output revealed significantly shorter reaction times for hands-free, hands-on and deactivation. While in the first drive, drivers took long to extract information from the visual HMI and taking over vehicle control, the additional semantic speech output supported their take-over behavior by explicitly clarifying the content of the visual HMI.

From the present reaction time findings, we suppose that the process of building situation awareness (Endsley and Garland, 2000), additional speech output actively supports drivers in the formation of an appropriate mental model of the control transition scenario. Semantic information that is provided by means of additional speech output thus actively supported drivers in the initiation and execution process of the transition from automated to manual control.

4.2. *Subjective evaluation*

The investigation of acceptance of the two systems with and without additional speech output showed a clear picture. Results suggest that in general drivers prefer the 'Speech + generic' system. The only participant who preferred the 'Generic' system indicated that too frequent presentation of speech output could be annoying.

Usability of both systems assessed through SUS scores was rated as relatively high with drivers clearly favoring the HMI with additional speech output. Participants ascribed extremely high usability to the 'Speech + generic' HMI version with only small potential for improvement. Additional speech output seemed to be an important feature for enhancing usability of the CAD function. When presented with speech output, drivers do not have to monitor the visual parts of an HMI for a long time to extract relevant information (Naujoks et al., 2016b) and consequently such a system is easier and faster to understand, use and learn for the human driver. Inferential statistics for visual workload do not support this assumption, however, on a descriptive level drivers tended to look less at the visual HMI in the 'Speech + generic' condition than in the 'Generic' condition. Taking a closer look at glance behavior during system maneuver scenarios in CAD, Naujoks et al. (2017) found that additional speech output led to a decrease in observed monitoring behavior.

When presented with additional speech output the visual HMI component tended to be rated as less helpful in the 'Speech + generic' condition than in the 'Generic' condition. At the same time, the auditory HMI component was rated as more important in the 'Speech + generic' condition than in the 'Generic' condition. This suggests that when presented with semantic auditory information, the visual parts of an HMI become less important while the auditory message is of high importance. Designers of visual-auditory TORs need to consider this finding when conceptualizing an HMI.

The findings on high ratings for acceptance, usability and usefulness for the speech based system contradict the assumption of lower acceptance due to high nuisance for speech based systems (Bliss and Acton, 2003; Cotté et al., 2001). A nuisance could have been expected in the present experiment, since during the 10-min drive there were four interaction scenarios (see 2.4 *scenario layout*). On a descriptive level, results showed a slightly elevated level of visual workload when only generic auditory information was presented. It might be, drivers had the impression that they needed to attribute more visual resources towards the HUD in order to derive information about the upcoming transition. Along with the results on monitoring behavior from Naujoks et al. (2017), semantic speech output can support drivers by at least slightly reducing visual workload as proposed by Alvarez et al. (2011) or Vilimek and Hempel (2005).

4.3. *Practical implications*

The results of the study are directly relevant for the design of non-imminent take-over requests for automated vehicles. Adding speech output sped up processes involved in the transition to manual driving. This impact of speech-based TORs leads to an increase in safety and comfort of transitions to manual driving. Specifically, drivers were faster in disengaging from the non-driving related task and putting their hands back on the steering wheel. Consequently, they had a greater time budget to perceive and react to the driving situation. Taken together with the higher acceptance of the speech-based take-over concept, it can be recommended to use additional semantic speech output to prepare drivers to take over manual driving in non-imminent situations.

4.4. *Limitations*

The relatively small sample size can be considered as a constraint of this study concerning generalizability. The present results point towards a superiority of semantic versus generic auditory output concerning the assessed reaction times and subjective evaluations. However, a replication with a larger sample is necessary in order to validate these findings. Future studies on the examination of additional speech output should incorporate a larger sample in order to prevent potential type two errors.

The mapping of cognitive information processing to the reaction time measures is a limitation of this study. The generic warning tone presented in both experimental conditions seemed to guide visual attention towards the road in a "bottom-up" manner. Therefore, we suppose that such a gaze measure is not very suitable to find out about differences between the two HMIs. If this bottom-up process is considered a part of information processing, it is not possible to selectively differentiate between attention allocation and information processing anymore, because attention allocation is rather a top-down process (i.e., looking towards an area where one assumes important information). Future studies need to investigate the processes underlying subsequent measures from termination of the NDRT to full recovery of manual vehicle control.

It is possible to argue that the present evidence of the superiority of additional speech output over mere generic auditory output is attributable to an information redundancy effect rather than speech per se. As outlined above, the additional speech output completed before actions of NDRT interruption had begun. This makes it highly likely, that the semantic content of the information facilitated drivers' reactions towards the TOR. Future studies need to clarify which particular information content about upcoming system limitations (e.g., limitation characteristics, necessary driver actions) facilitates driver responses to TORs.

This study did not examine take-over quality as described in Gold et al. (2013). This is partly due to the fact that the emerging secondary lanes scenario and the large time budget did not require active input such as steering or braking by the driver. For a safe transition to manual it was sufficient to deactivate the system and stabilize lateral vehicle guidance which is primarily reflected by hands-on reaction times. In order to find out about take-over quality, scenarios, which require a more difficult input from the driver, need to be investigated.

5. Conclusion

The present study brought forth further evidence of driver's behavior under conditions of early announcements in take-over scenarios. Reaction time findings suggest that drivers react quickly to the primary system indication and deliberately came to the conclusion to terminate the NDRT and take over manual vehicle control. In the present study, the process of getting back in the loop (Lorenz et al., 2014) from being engaged in an NDRT was accomplished by the drivers within the first half of the available time budget.

Furthermore, these results could support existing findings for the superiority of semantic speech output over generic auditory output for visual-auditory HMIs from system maneuver scenarios (see Naujoks et al., 2016a,b,c) to TOR scenarios in the area of CAD. This superiority does not only apply to subjective evaluations by the participants but also to measures of take-over quantity which include the initiation of the take-over process (i.e., hands-on and hands-free), but not the first reaction to a TOR or the final deactivation of the CAD system.

Acknowledgment

This work results from the joint project Ko-HAF – Cooperative Highly Automated Driving and has been funded by the Federal Ministry for Economic Affairs and Energy based on a resolution of the German Bundestag. We would like to thank Samantha Fritzsch for her support with the graphical design and Martin Grein for the implementation of the CAD function in the driving simulation. Both Samantha Fritzsch and Martin Grein are with WIVW. Additionally, we would like to thank three anonymous reviewers for their valuable suggestions and remarks.

References

Allport, D.A., Styles, E.A., Hsieh, S., 1994. Exploring the Dynamic Control of Tasks Shifting Intentional Set.

Alvarez, I., Martin, A., Dunbar, J., Taiber, J., Wilson, D.-M., Gilbert, J.E., 2011. Designing driver-centric natural voice user interfaces. Adj. Proc. AutomotiveUI 11, 42–49.

Barón, A., 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.

Bazilinskyy, P., de Winter, J., 2015. Auditory interfaces in automated driving: an international survey. PeerJ Comput. Sci. 1.

Bliss, J.P., Acton, S.A., 2003. Alarm mistrust in automobiles: how collision alarm reliability affects driving. Appl. Ergon. 34 (6), 499–509.

Brooke, J., 1996. SUS-A quick and dirty usability scale. Usability Eval. Ind. 189 (194), 4–7.

Cotté, N., Meyer, J., Coughlin, J.F., 2001. Older and younger drivers' reliance on collision warning systems. Paper Presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Endsley, M., Garland, D.J., 2000. Theoretical underpinnings of situation awareness: a critical review. Situation Awareness Analysis and Measurement. pp. 3–32.

Eriksson, A., Stanton, N.A., 2016. Takeover time in highly automated vehicles: non-critical transitions to and from manual control. Hum. Factors 59 (4), 689–705.

Flemisch, F., Heesen, M., Hesse, T., Kelsch, J., Schieben, A., Beller, J., 2012. Towards a dynamic balance between humans and automation: authority, ability, responsibility and control in shared and cooperative control situations. Cognit. Technol. Work 14 (1), 3–18.

Forster, Y., Naujoks, F., Neukum, A., 2016. Your turn or my turn? Design of a Human-Machine Interface for Conditional Automation. In: Paper Presented at the 8th Conference of User Interfaces and Vehicular Applications. Ann Arbor, MI.

Gasser, T.M., Westhoff, D., 2012. BASt-study: definitions of automation and legal issues in Germany. Paper Presented at the Proceedings of the 2012 Road Vehicle Automation Workshop.

Gold, C., Damböck, D., Lorenz, L., Bengler, K., 2013. Take over! How long does it take to get the driver back into the loop? Paper Presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Hakuli, S., Bruder, R., Flemisch, F., Löper, C., Rausch, H., Schreiber, M., Winner, H., 2012. Kooperative automation Handbuch Fahrerassistenzsysteme. Springer, pp. 641–650.

Hergeth, S., Lorenz, L., Krems, J.F., 2016. Prior familiarization with takeover requests affects drivers' takeover performance and automation trust. Hum. Factors 0018720816678714.

Hoffmann, S., Kruger, H.P., Buld, S., 2003. Avoidance of simulator sickness by training the adaptation to the driving simulation. VDI BERICHTE 1745, 385–406.

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A.M., Koch, I., 2010. Control and interference in task switching—a review. Psychol. Bull. 136 (5), 849.

Lenné, M.G., Triggs, T.J., 2009. Warning drivers of approaching hazards: the importance of location cues and multi-sensory cues. Hum. Factors Secur. Safe. 203–211.

Lorenz, L., Kerschbaum, P., Hergeth, S., Gold, C., Radlmayr, J., 2014. Der fahrer im hochautomatisierten fahrzeug. vom dual task zum sequential task paradigma. Paper Presented at the 7. Tagung Fahrerassistenz.

Lorenz, L., Kerschbaum, P., Hergeth, S., Gold, C., Radlmayr, J., 2015. Vom Dual-Task zum Sequential-Task Paradigma: Ein R{ü}ckblick {ü}ber Fahrsimulatorstudien. Paper presented at the 7. Tagung Fahrerassistenz.

Merat, N., Jamson, H., Lai, F., Daly, M., Carsten, O., 2014. Transition to manual: driver behaviour when resuming control from a highly automated vehicle. Transp. Res. Part F: Traffic Psych. Behav. 27, 274–282 citeulike-article-id:14113872.

Meyer, J., 2004. Conceptual issues in the study of dynamic hazard warnings. Hum. Factors 46 (2), 196–204.

Naujoks, F., Grattenthaler, H., Neukum, A., 2012. Zeitliche Gestaltung Effektiver Fahrerinformationen Zur Kollisionsvermeidung Auf Der Basis Kooperativer Perzeption. Paper presented at the 8. Workshop FAS.

Naujoks, F., Mai, C., Neukum, A., 2014. The effect of urgency of take-over requests during highly automated driving under distraction conditions. Adv. Hum. Aspects Transp. (Part I) 431.

Naujoks, F., Grattenthaler, H., Neukum, A., Weidl, G., Petrich, D., 2015a. Effectiveness of advisory warnings based on cooperative perception. IET Intell. Transp. Syst. 9 (6), 606–617.

Naujoks, F., Purucker, C., Neukum, A., Wolter, S., Steiger, R., 2015b. Controllability of partially automated driving functions–does it matter whether drivers are allowed to take their hands off the steering wheel? Transp. Res. Part F: Traffic Psychol. Behav. 35, 185–198.

Naujoks, F., Forster, Y., Wiedemann, K., Neukum, A., 2016a. A human-Machine interface for cooperative highly automated driving. In: Paper Presented at the 7th International Conference on Applied Human Factors and Ergonomics. Orlando, FL.

Naujoks, F., Forster, Y., Wiedemann, K., Neukum, A., 2016b. Speech improves human-automation cooperation in automated driving. Mensch Und Computer 2016–Workshopband.

Naujoks, F., Kiesel, A., Neukum, A., 2016c. Cooperative warning systems: the impact of false and unnecessary alarms on drivers' compliance. Accident Anal. Prev. 97, 162–175.

Naujoks, F., Forster, Y., Wiedemann, K., Neukum, A., 2017. Improving usefulness of automated driving by lowering primary task interference through HMI design. J. Adv. Transp.

Payre, W., Cestac, J., Delhomme, P., 2016. Fully automated driving impact of trust and practice on manual control recovery human factors. J. Hum. Factors Ergon. Soc. 58 (2), 229–241.

Politis, I., Brewster, S., Pollick, F., 2015. Language-based multimodal displays for the handover of control in autonomous cars. Paper Presented at the Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications.

Radlmayr, J., Gold, C., Lorenz, L., Farid, M., Bengler, K., 2014. How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving. Paper Presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Rauch, A., Klanner, F., Rasshofer, R., Dietmayer, K., 2012. Car2x-based perception in a high-level fusion architecture for cooperative perception systems. Paper Presented at the Intelligent Vehicles Symposium (IV), 2012 IEEE.

Rogers, R.D., Monsell, S., 1995. Costs of a predictible switch between simple cognitive tasks. J. Exp. Psychol. Gen. 124 (2), 207.

SAE, 2016. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles J3016R. SAE International.

Seeliger, F., Weidl, G., Petrich, D., Naujoks, F., Breuel, G., Neukum, A., Dietmayer, K., 2014. Advisory warnings based on cooperative perception. Paper Presented at the Intelligent Vehicles Symposium Proceedings, 2014 IEEE.

Stanton, N.A., Edworthy, J., 1999. Human Factors in Auditory Warnings. Ashgate.

Stanton, N.A., Walker, G.H., Salmon, P.M., 2015. Human Factors in Automotive Engineering and Technology. Ashgate Publishing Ltd.

Vilimek, R., Hempel, T., 2005. Effects of Speech and Non-Speech Sounds on Short-Term Memory and Possible Implications for In-Vehicle Use.

Wiedemann, K., Schömig, N., Mai, C., Naujoks, F., Neukum, A., 2015. Drivers' monitoring behaviour and interaction with non-driving related tasks during driving with different automation levels. In: Paper Presented at the 6th International Conference on Applied Human Factors and Ergonomics (AHFE). Las Vegas, USA.