Better Driving and Recall When In-car Information Presentation Uses Situationally-Aware Incremental Speech Output Generation

Casey Kennington
CITEC, Bielefeld University
Universitaetstrasse 25
Bielefeld, Germany
ckennington@cit-ec.uni-bielefeld.de

Spyros Kousidis
Bielefeld University
Universitaetstrasse 25
Bielefeld, Germany
spyros.kousidis@uni-bielefeld.de

Timo Baumann
Hamburg University
Vogt-Koelln-Strasse 30
Hamburg, Germany
baumann@informatik.uni-hamburg.de

Hendrik Buschmeier
CITEC, Bielefeld University
Universitaetstrasse 25
Bielefeld, Germany
hbuschme@uni-bielefeld.de

Stefan Kopp
CITEC, Bielefeld University
Universitaetstrasse 25
Bielefeld, Germany
skopp@uni-bielefeld.de

David Schlangen
Bielefeld University
Universitaetstrasse 25
Bielefeld, Germany
david.schlangen@uni-bielefeld.de

ABSTRACT

It is established that driver distraction is the result of sharing cognitive resources between the primary task (driving) and any other secondary task. In the case of holding conversations, a human passenger who is aware of the driving conditions can choose to interrupt his speech in situations potentially requiring more attention from the driver, but in-car information systems typically do not exhibit such sensitivity. We have designed and tested such a system in a driving simulation environment. Unlike other systems, our system delivers information via speech (calendar entries with scheduled meetings) but is able to react to signals from the environment to interrupt when the driver needs to be fully attentive to the driving task and subsequently resume its delivery. Distraction is measured by a secondary short-term memory task. In both tasks, drivers perform significantly worse when the system does not adapt its speech, while they perform equally well to control conditions (no concurrent task) when the system intelligently interrupts and resumes.

Author Keywords

Spoken Dialogue Systems, Incremental Dialogue, In-car Dialogue, Speech Output Generation

ACM Classification Keywords

H.5.2 User Interfaces: Information Systems

AutomotiveUI '14, September 17 - 19 2014, Seattle, WA, USA

ACM 978-1-4503-3212-5/14/09 $15.00
http://dx.doi.org/10.1145/2667317.2667332

1. INTRODUCTION

The risks of holding conversations on a mobile phone while driving are by now well established [17, 21]. This is commonly attributed erroneously to the handling of the actual device [8], although it has been shown that hands-free devices rarely improve driver performance while talking or texting [17, 9, 10, 8]. Holding conversations with a passenger, in contrast, is found to be much safer. Drews et al. [6], found a much smaller impact on driving performance, perhaps due to the fact that "surrounding traffic becomes a topic of the conversation, helping driver and passenger to share situation awareness, and mitigating the potential effects of conversation on driving" [6, p. 2210]. There is evidence that the discriminating factor is *awareness* of the driving situation, which allows the passengers to adopt strategies that relieve the driver from attending to the conversation in difficult driving situations [7]. In other words, *co-location* is a requirement for risk-free in-car interaction, regardless of the use of manual or speech-based interface. Interestingly, co-location can be achieved via *telepresence*, e.g. in [16], cell-phone conversations were safer when the partners had real-time visual information and could thus assess the driving conditions.

These findings on in-car conversations can be carried over to in-car information systems research, especially in the case of speech-based interfaces and Spoken Dialogue Systems (SDS), as paying attention to speech induces additional cognitive load to the driver [5]. Currently, such systems do not exhibit situational awareness. When they expect voice input, they expect it to come within a certain time window, regardless of whether or not the driver should have focus elsewhere, and when they produce voice output, the implicit assumption is that it will be equally well understood at all times, regardless of driving situation. A recent study, using a simulated interactive voice system, consequently found that their system put an even

higher cognitive load [20] on drivers than conversing on a cell phone.

We hypothesised that incremental output generation (which for us, following [4], covers both the incremental generation of language as well as of speech) can adapt the speech presentation such that a spoken dialogue system has some awareness of the surroundings and can interrupt its own speech, thus reducing cognitive burden on the user. Using a driving simulation setup, we implemented a dialogue system that realises this strategy. By employing incremental output generation, the system can interrupt and flexibly resume its output. We tested the system using a variation of a standard driving task, and found that it improved both driving performance and recall, as compared to a non-adaptive baseline system.

In this paper, we first explain incremental dialogue, argue why it is the right approach for in-car SDS, and detail our component for incremental language and speech generation. We then describe our system setup, experiment design and tasks, the conditions and the variables. Following this, we give results of our experiments, with discussion, and conclude.[1]

2. INCREMENTAL LANGUAGE GENERATION

Incremental SDS process input and produce output as much as possible; it does not wait until the end of an utterance to begin processing. In this section we explain a component of SDS that is the focus of this paper: speech output generation.

Making the output of an in-car SDS situationally aware requires its output generation modules – speech synthesis and natural language generation – to be able to (1) timely and plausibly interrupt and resume speech output, and (2) to flexibly adapt or even reformulate the content of its utterances, taking into account a preceding delivery interruption. Both requirements call for incremental processing in these modules.

For speech synthesis, incrementality allows for shorter response times (i.e., it can resume faster) as the system can start speech output while still synthesising the rest of an utterance [4]. It also enables changes to the prosody of an ongoing utterance [2], allowing the system to add a prosodic marker to signal the system's awareness to the word preceding the interruption. For natural language generation, incrementality makes it possible to change those parts of an utterance that have not been delivered yet. The continuation of an interrupted utterance can thus differ from planned but yet undelivered parts by choosing a continuation point that, e.g., re-states some of the content but does not repeat more than is needed.

Our work builds on the existing incremental output generation system of [4] that fulfills the requirements specified above and is partially available in the open source incremental dialogue processing toolkit INPROTK [3], explained below.[2] It consists of incremental components for natural language generation and speech synthesis that are integrated in such a way that timely interruptions and adaptive continuations are possible.

[1]This paper is a more in-depth report on the system design and provides an extended analysis of the results compared to the preliminary report in [13].

[2]`http://inprotk.sourceforge.net/`

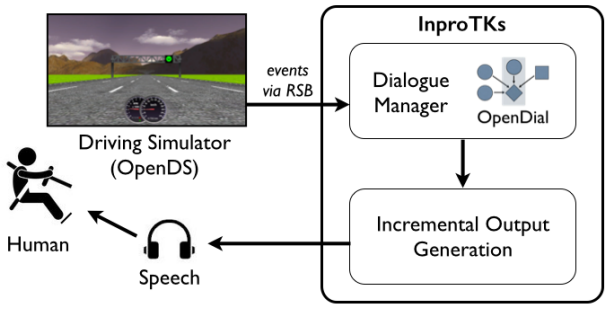

Figure 1. System overview: the human participant controls a steering wheel and a pedal in front of a large screen that shows the OpenDS simulator. Events are sent via RSB to INPROTK$_S$ where the DM, Opendial, controls the incremental speech output.

The system's language generation component creates utterances in two processes [4]. The first of these plans the overall utterance by laying out a sequence of *chunks* which determine what will be said when; the second, based on the SPUD microplanning framework [19], computes how each of these chunks is realised linguistically. Utterances are incrementally generated chunk by chunk. Adaptations to an ongoing utterance are also constrained on the chunk level. The chunk-planning process can change the sequence of chunks, repeat one or several chunks, or leave some out. The microplanning process can change how a chunk is realised, e.g., by inserting or leaving out cue words, by providing information that has been mentioned before, or by making information conveyed implicitly explicit – or vice versa. Our system made use of adaptations resulting from both processes.

Incremental speech synthesis [2] performs all computationally expensive processing steps such as waveform synthesis as late as possible while performing prosodic processing (which has non-local effects) as early as necessary [1], resulting in fast response times without sacrificing quality. Ongoing synthesis can be changed, and adapted prosodically with minimal latency, and provides detailed progress information on various linguistic levels. Our system uses the incremental capabilities to stop synthesis at word boundaries when interrupted, to generate new sentence onset intonations for continuations, and to drive the generation processes just-in-time.

3. SYSTEM SETUP

The overall layout of our system is depicted in Figure 1. Our driving simulation scenario consists of a 40-inch 16:9 screen with a Thrustmaster PC Racing Wheels Ferrari GT Experience steering wheel and pedal. Audio is passed to the participant via headphones (see also Figure 4).

For the driving simulator, we used the OpenDS Toolkit.[3] We developed our own simple driving scenarios (derived from the "ReactionTest" task, which is distributed together with OpenDS) that specified the driving task and timing of the concurrent speech, as described below. We modified OpenDS to pass real-time data (e.g. car position/velocity/events in the

[3]`http://www.opends.eu/`

simulation, such as a gate becoming visible or a lane change) using the *mint*.tools architecture [14].

For the SDS, we use INPROTK [3] which realises the *IU*-model of incremental processing [18]. To extend INPROTK to handle situated, multimodal input, we used a recently extended version, INPROTK$_S$ [12]. The system we implemented does not represent what is generally termed as a dialogue system, however, we used the same modularisation as in more typical dialogue systems by using a dialogue management (DM) component that controls the system actions based on user actions. We integrated OpenDial [15] as the DM into INPROTK$_S$,[4] though we only used it to make simple, deterministic decisions. We used the incremental output generation capabilities as described previously, as a module in INPROTK$_S$.

Combining the tools described above, we are able to pass messages from the driving simulation to the dialogue system using real-time interprocess communication protocols over Gigabit LAN. The messages passed are *event triggers* that simulate the capability of an intelligent system to be aware of the driving conditions: when an event occurs in the driving simulation (e.g. a signal becomes visible on the road, the car changes lane) this event triggers a message to the dialogue manager. Since such events in the driving simulation can have unique identifiers, we use the latter to *script* dialogue manager behaviour. Several of these signals are invisible in the simulation: the driver cannot see them, but passing through them initiates an event (e.g. the system starts speaking).

4. EXPERIMENT DESIGN

The goal of the experiment is two-fold: first, we want participants to be able to perform a **driving task** as a responsible driver would; second, we want to explore how well they pay attention to and **recall speech** during driving, under two possible presentations of speech. One presentation is that the speech is *adaptive*, in that when a "dangerous" situation is detected in the scene, the incremental speech output is interrupted and later resumed after the dangerous situation is no longer present. This mimicks a situated dialogue participant which is aware of the physical surroundings and driving conditions. The second presentation of speech is a *non-adaptive*, non-incremental system that does not stop speaking when a dangerous driving condition is detected. Both tasks (driving and memory) are explained below.

The Driving Task

For the driving task we used a variant of the well-known lane-change task (LCT), which is standardised in [11]. The task requires the driver to react to a green light positioned on a signal gate above the road (see Figure 3). The driver, otherwise instructed to remain in the middle lane of a straight, 5-lane road, must move to the lane indicated by the green light, remain there until a tone is sounded, and then return again to the middle lane. OpenDS gives a *success* or *fail* result to this task depending on whether the target lane was reached within 10 seconds (if at all) and the car was in the middle lane when the signal became visible. In addition, OpenDS reports a *reaction time*, which is the time between the moment the signal to change lane becomes visible and the moment the lane has been reached.

In pre-experiments it was determined that the task was too easy, so we added an additional constraint to slightly increase the cognitive load: during a lane-change, the driver was to maintain a speed of 60 km/h, where the car maintained 40 km/h when the pedal was not pressed, with a top speed of 70 km/h when fully pressed. We calculate a further response variable to measure performance in this last task, namely the root mean square error (RMSE) of the car velocity *difference* from 60 km/h during the lane-change manoeuvre.

The Memory task

We tested the attention of the drivers to the generated speech using a simple true/false memory task. The dialogue system generated utterances such as *"Am Samstag den siebzehnten Mai 12 uhr 15 bis 14 uhr 15 hast du gemeinsam Essen im Westend mit Martin"* (On Saturday the 17th of May from 12:15 to 14:15 you are meeting Martin for lunch). These utterances always had 5 information tokens in a specified order: day, time, activity, location, and partner (the date was excluded) and were spoken by a female voice. Soon after the utterance was complete, and while no driving distraction occurred, a true/false confirmation question about one of the uttered tokens was asked by a male voice, e.g. *"Richtig oder Falsch?–Freitag"* (Right or wrong?–Friday). The subject was then required to answer true or false by pressing one of two respective buttons on the steering wheel.

The token of the confirmation question was chosen randomly, although tokens near the beginning of the utterance (day and time) were given a higher probability of occurrence, as we observed in pilot experiments it is generally easier to remember the latter tokens of the utterance in comparison to early tokens. Especially in the case of an interruption/resumption, tokens spoken after the resumption can be more easily remembered than those given before the interruption. Giving the early tokens higher probability of occurrence biases the design *against* the adaptive system since the question tends to refer to tokens spoken *before* the interruption more often than not.

Interaction between tasks

A lane-change is defined as a "dangerous" situation; driving in the middle lane is a "normal" situation. Under adaptive speech

Figure 3. Lane signal as presented on screen in our experiments.

[4]OpenDial is available at `http://opendial.googlecode.com/`

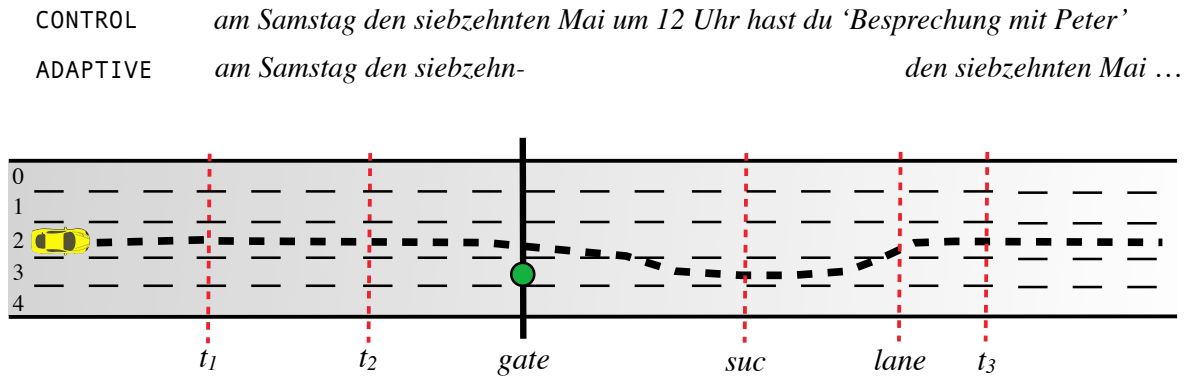

Figure 2. Top view of driving task: as the car moves to the right over time, speech begins at t_1, the gate with the lane-change indicator becomes visible at t_2, where in the adaptive version speech pauses. Successful lane change is detected at `suc`; successful change back to the middle lane is detected at `lane`, and resumes. (If no change back is detected, the interruption times out at t_3). All red-dashed lines denote events sent from OpenDS to the Dialogue Manager.

presentation, the speech was interrupted during a lane-change. Under the non-adaptive speech presentation, no interruption occurred and speech continued during the lane-change. Figure 2 shows how the task unfolds over time when changing a lane: all red-dashed lines represent invisible event triggers or simply events of the simulation that trigger unique messages to be sent to the dialogue system. At the t_1 marker, A trigger is sent to the DM to start the speech. In order to vary which tokens are uttered during a lane change, a random delay (0–4 sec) is inserted before the speech begins (non-adaptive setting). In the case of an adaptive setting, a random delay (4–7 sec) is inserted to vary how many tokens are presented before the interruption is triggered by the lane change. At t_2, the gate is in view (as seen from Figure 3) and a gate light is visible. In the adaptive setting, at this point the speech would be interrupted; in the non-adaptive setting the speech would continue until complete. At *suc*, the target lane has been reached (the tone signal is sounded), but the speech still does not resume. At *lane*, the car has returned to the middle lane, at which point the adaptive speech would resume the interrupted utterance. In case the task was not completed correctly (i.e., the target lane change did not happen), a timeout at t_3 would trigger the adaptive speech to continue. Three seconds after the calendar event speech is completed, the true/false question is asked. There is ample time for the participant to respond before any other distraction (on the road or from speech) occurs.

5. CONDITIONS

We consider all permutations between the two tasks: speech delivery can be adaptive, non-adaptive or absent (driving test by itself). Non-adaptive speech can also occur outside dangerous situations (memory test by itself). This yields four possible conditions, shown in Table 1.

Presentation order

In order to balance the presentation order we design the scenario as follows: In the driving simulator scenario we place 44 gates with a condition attached to them (11 gates/condition) along a straight stretch of road. We shuffle the order of these gates randomly, with no constraints. We also place two empty gates in between each, which have no condition attached to them (and thus no lane-change signal). For each participant, a new scenario is generated, thus each participant faces a unique presentation order.

The above arrangement serves a second purpose, namely that it is not possible for participants to predict what the condition of the next gate will be. Even the number of empty in-between gates varies, as one of the conditions (*ControlSpeech*) is a gate with no lane-change signal, and if several of them happen to be in a row, then the number of consecutive gates with no signal could be high. This prohibits participant strategies based on local repetition of conditions (if the conditions had instead been presented in contiguous blocks), which would allow them to expect when/if the speech delivery is going to be started. This, coupled with the random time delay before speech initiates, creates an impression of complete randomness and independence of the two tasks.

6. EXPERIMENTAL PROCEDURE

Figure 4 shows a participant during the driving simulation experiment. At first, the participant signed a consent form and was then seated in the chair in front of the steering wheel and screen (seat adjustments made, if necessary). The participant was then given headphones to put on, after which the audio levels were tested and the task was explained. Then, the OpenDS scene was started, showing the driving simulation on the large screen, at which point the participant was instructed to control the steering wheel and pedal. In the beginning of the simulation, 10 signal gates were presented for practice in using the controls and performing the task. During this practice stretch

Speech Condition	Driving condition	
	Lane change	No lane change
No speech	CONTROL_DRIVE	—
Non-adaptive	NO_ADAPT_DRIVE	CONTROL_SPEECH
Adaptive	ADAPT_DRIVE	—

Table 1. Condition labels of combined conditions from driving and memory task.

Figure 4. Example of task: user is seated in front of steering wheel and a large screen, the speech is presented via the headphones.

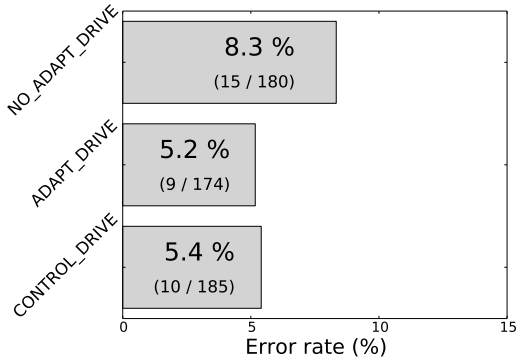

Figure 5. Error rate in three conditions for driving task.

of road an experimenter was sitting next to the participant in order to clarify any questions that could be asked during this phase (the simulation could be paused, if necessary, to answer difficult questions or make adjustments). When the participants confirmed that they had understood the task, the experimenter left the scene.

Immediately after the practice gates, without any interruption, a clearly marked **START** gate signaled the beginning of the experiment. The participant was presented the four conditions, as explained above, over the course of 44 gates. The end of the experiment was signaled with a clearly marked **FINISH** gate, at which point the simulation stopped. In total, the driving simulation took around 30 minutes, including practice time. The participant was then given a post-task questionnaire.

Difficulty	Freq.
4 (easy)	8
3	7
2	1
1 (hard)	1

Table 2. Subjects' judgment of task difficulty.

Preference	Freq.
ADAPTIVE	3
CONTROL	9
Neither	5

Table 3. Subjects' system preference.

In total, 17 participants (8 male, 9 female, aged 19–36) participated in the study. All of the participants were native German speakers affiliated with Bielefeld University and holders of a (at least EU class B, which is standard) driving license. Two participants had previous experience with driving simulators and only one had previous experience with spoken dialogue systems.

7. RESULTS AND DISCUSSION

We first present results from the post-experiment survey. As seen in Table 2 the majority of participants found the task relatively easy. The one who found it extremely difficult did not perform worse or better than average. Table 3 shows the *preference* of participants between the different speech delivery strategies on the system.[5] We observe that the *non-adaptive* strategy is preferred by the majority, followed by the neutral response (no preference).

[5]All participants noticed after completion that the system had two presentation methods; it was not explained before the experiment.

Driving task

In terms of successful trials – successful lane change – in the driving task we compare three of the four conditions (the condition in which no lane change occurs is of course omitted here). Figure 5 shows the percentage error rate per condition, across all participants.

We find that the error rate is higher (a greater percentage of failed trials) in the condition of the non-adaptive system. The performance in the adaptive system condition is identical to that of the control condition, in which no concurrent speech occurs during the driving task.

We have tested the significance of the results using a generalized linear mixed model (GLMM) with CONDITION and SUBJECT as factors, which yields a p-value of 0.01231 when compared against a null model in which only SUBJECT is a factor (condition is the within-subject factor). No significant effects of between-subject factors GENDER, DIFFICULTY or PREFERENCE were found. In addition, the within-subject variable *time* did not have any significant effect (i.e., subjects do not improve in the driving task with time). This finding meets our expectation that an adaptive speech delivery strategy, aware of the driving conditions, does not noticeably distract the driver while the non-adaptive strategy clearly does.

The ability of subjects to keep a constant velocity of 60 km/h while overtaking was not affected by CONDITION. However, participants got better at this task over time (see Figure 6). This learning effect was found to be significant (repeated measures ANOVA, 2x2 factorial design, $F_{verr} = 20.464, p < 0.001$). None of the between-subject variables GENDER, DIFFICULTY or PREFERENCE showed significant effects. Finally, neither CONDITION, TIME, nor any between-subject factors showed any effect on the reaction time to the Lane change task signal.

Memory task

The percentage of wrong answers to the system's recall questions (across all participants) are shown in Figure 7. Here we compare across the three conditions in which speech is present (changing lane without concurrent speech is of course not considered). As in the case of the driving task, we observe that the adaptive system outperforms the non-adaptive version significantly (same GLMM approach as above yields a p-value of 0.027 when compared against the simpler model with only

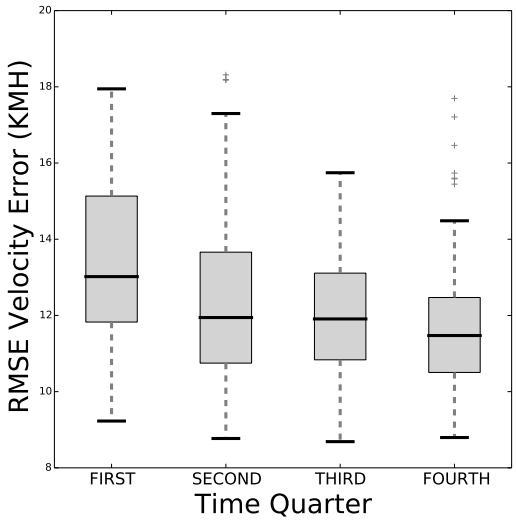

Figure 6. Root-mean-square error from a reference velocity of 60 km/h during lane change.

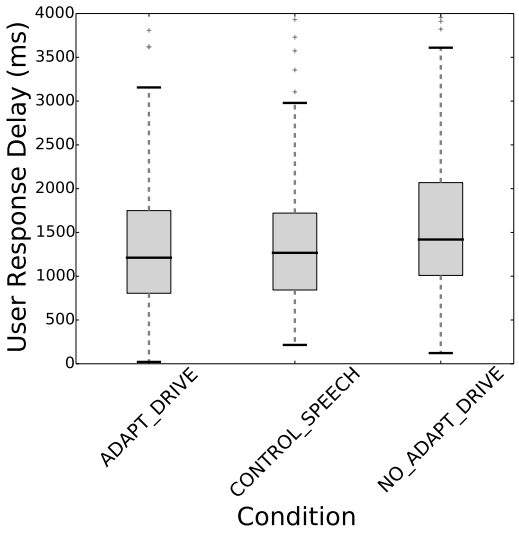

Figure 8. User answer response delay under three conditions.

SUBJECT as factor). Our main hypothesis is again evaluated, namely that the adaptive nature of the speech delivery naturally allows the driver to focus better on the information encoded in the speech while there are no concurrent distractions on the road. The ability of the incremental language generation to resume by appropriately rephrasing the remaining tokens adds to the quality of the experience and presumably to the performance, as opposed to pausing/resuming the raw audio, which could result in undesirable clipping and half-word tokens that could hinder language perception and thus degrade performance.

The within-subject variable TIME was not found to be significant; participants did not improve in the memory task over time. It may be that although participants can get used to the task and the un-changing syntactic ordering of the sentences, *fatigue* could become a factor over time, canceling out the learning effect. Also, none of the between-subject factors GENDER, DIFFICULTY or PREFERENCE were found to have any significant effect.

In the case of the average response delay (from the end of the recall question to the button press), we observe that both

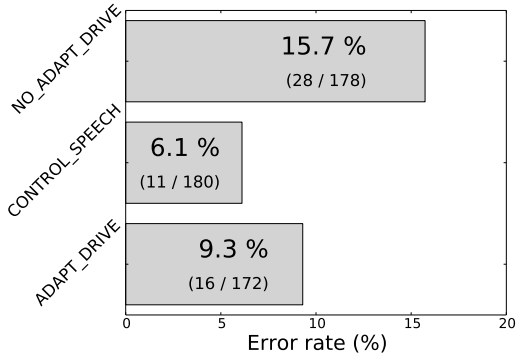

Figure 7. Answer error rate in three conditions for memory task.

CONDITION (Figure 8), and TIME are important factors. The response delay to recall questions is significantly higher in the non-adaptive system condition, while no variation is observed between the adaptive condition and the control condition (recall task without concurrent lane change). In addition, response delay decreases with time, possibly showing a learning effect with respect to the hands finding the button on the steering wheel more automatically, but also to the structure of the prompt sentence and the type of tokens it contains. Both factors (CONDITION and TIME) are significant (repeated measures ANOVA, 2x2 factorial design, $F_{condition} = 3.858, p = 0.0359, F_{time} = 4.672, p = 0.00662$). No significant effects were found for any of the between-subject factors (GENDER, DIFFICULTY, PREFERENCE).

It is interesting that the gains in performance and safety are lost in user preference, as the non-adaptive preference was overwhelmingly rated more favourably, while the adaptation strategy of the system was sometimes understood as a malfunction. Mostly, however, the participants stated that they would like more control over the adaptation strategy and, indeed, it would be better if the interruption/resumption signals could be more customised, allowing some kind of user input to override the default behaviour.

8. CONCLUSION AND FUTURE WORK

We have presented a situationally-aware in-car SDS. It was shown that adapting speech delivery to the road conditions, made possible by incremental SDS technology, improves performance in both the primary driving task and the secondary short-term memory recall task. This is in agreement with relevant evidence from the literature that situational-aware, or co-located conversations do not contribute to driver distraction. This finding has important implications, as current industrial speech-based information systems (such as navigators) are not co-located. Our system would potentially benefit from added functionality of driver control, e.g., of when to resume interrupted speech. It would also benefit from some kind of verbal

cue, signaling to the driver that her attention is required (e.g., "um" preceding the beginning of an utterance or resumption).

For our next steps, we plan to incorporate functionality that will allow users to have some control over the interruption/resumption of speech delivery, using either speech (INPROTK$_S$ provides incremental speech recognition), head gestures, or manual control.

ACKNOWLEDGMENTS

This research was partly supported by the Deutsche Forschungsgemeinschaft (DFG) in the CRC 673 "Alignment in Communication" the Center of Excellence in "Cognitive Interaction Technology" (CITEC), and a PostDoc grant by Daimler and Benz Foundation to the 3rd author. The authors would like to thank Oliver Eckmeier and Michael Bartholdt for helping implement the system setup, as well as Gerdis Anderson and Fabian Wohlgemuth for assisting as experimenters. Thanks also to the anonymous reviewers.

REFERENCES

1. Baumann, T., and Schlangen, D. Evaluating prosodic processing for incremental speech synthesis. In *Proceedings of Interspeech* (Portland, USA, 2012), 438–441.
2. Baumann, T., and Schlangen, D. Inpro_iSS: A component for just-in-time incremental speech synthesis. In *Proc. ACL2012 System Demonstrations* (Jeju Island, Korea, 2012), 103–108.
3. Baumann, T., and Schlangen, D. The InproTK 2012 release. In *NAACL-HLT Workshop SDCTD* (Montréal, Canada, 2012), 29–32.
4. Buschmeier, H., Baumann, T., Dosch, B., Kopp, S., and Schlangen, D. Combining incremental language generation and incremental speech synthesis for adaptive information presentation. In *Proceedings of SigDial* (Seoul, Korea, 2012), 295–303.
5. Demberg, V., Sayeed, A., Mahr, A., and Müller, C. Measuring linguistically-induced cognitive load during driving using the ConTRe task. In *Proceedings of Automotive'UI* (Eindhoven, The Netherlands, 2013), 176–183.
6. Drews, F. A., Pasupathi, M., and Strayer, D. L. Passenger and cell-phone conversations in simulated driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 48* (2004), 2210–2212.
7. Drews, F. A., Pasupathi, M., and Strayer, D. L. Passenger and cell phone conversations in simulated driving. *Journal of Experimental Psychology: Applied 14* (2008), 392.
8. He, J., Chaparro, A., Nguyen, B., Burge, R., Crandall, J., Chaparro, B., Ni, R., and Cao, S. Texting while driving: Is speech-based texting less risky than handheld texting? In *Proceedings of Automotive'UI* (2013), 124–130.
9. Horrey, W. J., and Wickens, C. D. Examining the impact of cell phone conversations on driving using meta-analytic techniques. *Human Factors 48* (2006), 196–205.
10. Ishigami, Y., and Klein, R. M. Is a hands-free phone safer than a handheld phone? *Journal of Safety Research 40* (2009), 157 – 164.
11. ISO. Road vehicles – Ergonomic aspects of transport information and control systems – Simulated lane change test to assess in-vehicle secondary task demand. ISO 26022:2010, 2010.
12. Kennington, C., Kousidis, S., and Schlangen, D. InproTKs: A toolkit for incremental situated processing. In *Proceedings of SigDial*, ACL (Philadelphia, USA, 2014), 84–88.
13. Kousidis, S., Kennington, C., Baumann, T., Buschmeier, H., Kopp, S., and Schlangen, D. Situationally aware in-car information presentation using incremental speech generation: Safer, and more effective. In *Proceedings of the EACL 2014 Workshop on Dialog in Motion* (Gothenburg, Sweden, 2014), 68–72.
14. Kousidis, S., Pfeiffer, T., and Schlangen, D. MINT.tools: Tools and adaptors supporting acquisition, annotation and analysis of multimodal corpora. In *Proceedings of Interspeech* (Lyon, France, 2013), 2649–2653.
15. Lison, P. Probabilistic dialogue models with prior domain knowledge. In *Proceedings of SigDial* (Seoul, Korea, 2012), 179–188.
16. Maciej, J., Nitsch, M., and Vollrath, M. Conversing while driving: The importance of visual information for conversation modulation. *Transportation Research Part F: Traffic Psychology and Behaviour 14* (2011), 512–524.
17. McEvoy, S. P., Stevenson, M. R., McCartt, A. T., Woodward, M., Haworth, C., Palamara, P., and Cercarelli, R. Role of mobile phones in motor vehicle crashes resulting in hospital attendance: A case-crossover study. *BMJ 331* (2005), 428.
18. Schlangen, D., and Skantze, G. A general, abstract model of incremental dialogue processing. *Dialogue & Discourse 2* (2011), 83–111.
19. Stone, M., Doran, C., Webber, B., Bleam, T., and Palmer, M. Microplanning with communicative intentions: The SPUD system. *Computational Intelligence 19* (2003), 311–381.
20. Strayer, D. L., Cooper, J. M., Turrill, J., Coleman, J., and Medeiros, N. Measuring cognitive distraction in the automobile. Tech. rep., AAA Foundation for Traffice Safety, 2013.
21. Strayer, D. L., Drews, F. A., and Crouch, D. J. A comparison of the cell phone driver and the drunk driver. *Human Factors 48* (2006), 381–91.