Evaluation of Speech Dialog Strategies for Internet Applications in the Car

Hansjörg Hofmann
Ute Ehrlich
André Berton
Daimler AG / Ulm, Germany,
hansjoerg.hofmann@daimler.com

Angela Mahr
Rafael Math
Christian Müller
DFKI / Saarbrücken, Germany
angela.mahr@dfki.de

Abstract

Due to the mobile Internet revolution, people tend to browse the Web while driving their car which puts the driver's safety at risk. Therefore, an intuitive and non-distractive in-car speech interface to the Web needs to be developed. Before developing a new speech dialog system in a new domain developers have to examine what the user's preferred interaction style is in order to use such a system. This paper reports from a very recent driving simulation study and its preliminary results which are conducted in order to compare different speech dialog strategies. The use of command-based and conversational SDS prototypes while driving is evaluated on usability and driving performance. Different GUIs are designed in order to support the respective dialog strategy the most and to evaluate the effect of the GUI on usability and driver distraction. The preliminary results show that the conversational speech dialog performs more efficient than the command-based dialog. However, the conversational dialog distracts more from driving than the command-based. Furthermore, the results indicate that an SDS supported by a GUI is more efficient and better accepted by the user than without GUI.

1 Introduction

The pervasive use of smartphones in daily situations impacts the automotive environment. In order to stay "always connected" people tend to use their smartphone's Internet functions manually while driving. However, using a smartphone manually while driving, distracts the driver and endangers the driver's safety. According to Governors Highway Safety Association (2011) 25% of U.S. car crashes are related to drivers using their cellphones while driving. Therefore, the development of an intuitive and non-distractive in-car speech interface to the Web is essential in order to increase driver safety (Peissner et al., 2011).

Before developing a new speech dialog system (SDS) in a new domain developers have to examine how users would interact with such a system. An Internet user study by Hofmann et al. (2012a) in which the subjects had to solve Internet tasks orally, revealed that concerning communicational (e.g. sending an Email) and transactional tasks (e.g. booking a hotel) conversational and command-based speaking styles were used with equal frequency. Because of the equal frequency of occurrence you have to examine which speech dialog strategy - the command-based or the conversational - is the most suitable for these tasks.

First studies on the evaluation of dialog strategies have been conducted by Devillers and Bonneau-Maynard (1998) who compare two SDS allowing the user to retrieve touristic information. One dialog strategy guides the user via system suggestions, the other does not. The evaluated dialog strategies comprise the fundamental ideas the command-based and conversational dialog strategy consist of. By applying qualitative and quantitative criteria they conclude that user guidance is suitable for novices and appreciated by all kinds of users. However, there was no GUI involved and the speech interaction was performed as primary task. Considering the driving use case other results may be achieved since the primary task is driving. Furthermore, the use of these SDS among advanced users needs to be investigated.

In the TALK project, Mutschler et al. (2007) compared a command-based speech dialog to a conversational dialog where the driver had to control the in-car mp3-player by speech while driving. The same graphical user interface (GUI) was used for both dialog strategies. Although the conver-

Proceedings of the SIGDIAL 2013 Conference, pages 233–241,
Metz, France, 22-24 August 2013. ©2013 Association for Computational Linguistics

sational dialog was more efficient the command-based dialog was more appreciated by the subjects. According to Mutschler et al. the high error rate of the conversational strategy was the reason for the higher acceptance of the command-based dialog. There were no significant differences in the driving performance revealed when using the different SDS.

The speech recognizer quality has improved enormously within the last five years. Therefore, the weak speech recognition performance of Mutschler et al.'s conversational dialog may be nowadays less significant. Furthermore, the use of the same GUI for different dialog strategies could have additionally influenced the result. The GUI should be adapted to the particular dialog strategy in order to benefit from the advantages of the respective strategy the most and to allow for a comparison of optimal systems.

This paper reports from a very recent driving simulation study and its preliminary results which are conducted in order to compare different speech dialog strategies. The use of command-based and conversational SDS prototypes while driving is evaluated on usability and driving performance. The systems have been developed for German and allows users to perform a hotel booking by speech. Different GUIs are designed in order to support the respective dialog strategy the most and to evaluate the effect of the GUI on usability and driver distraction. The experiments have been conducted at DFKI, Saarbrücken using the OpenDS[1] driving simulation. The research work is performed within the scope of the EU FP7 funding project GetHomeSafe[2].

The remainder of the paper is structured as follows: In Section 2, the developed SDS prototypes are briefly described. Section 3 presents the experimental setup and its results and finally, conclusions are drawn.

2 SDS Prototype Concepts

The chosen use case for the design of the SDS concepts is booking a hotel by speech while driving since it covers many different subdialog types (parameter input, list presentation and browsing, etc.). For this purpose, the online hotel booking service HRS[3] has been used as data provider for the SDS.

Each SDS prototype concept offers the same functionality: First, the user has to input his search parameter to retrieve a list of hotels. The user can browse the list and ask for detailed information about a certain hotel. If the hotel matches his needs he is able to book the hotel. In addition, the user can change the search parameters.

In the following, the different speech dialog strategies and the corresponding GUI designs are briefly decribed. A detailed description of the human-machine interface (HMI) concepts can be found in Hofmann et al. (2012b).

2.1 Speech Dialog Strategy Design

SDS Prototypes for German language have been developed including the following SDS features: In order to speak to the system the driver has to press a Push-To-Activate (PTA) button. Furthermore, the driver is able to interrupt the system while prompting the user ("barge-in"). When designing the different dialog strategies we particularly focused our attention on the dialog initiative, the possibility to enter multiple input parameters and the acoustic feedback.

2.1.1 Command-based Speech Dialog Strategy

The dialog behavior of the command-based dialog strategy corresponds to the voice-control which can be found in current state-of-the-art in-car SDS. By calling explicit speech commands the speech dialog is initiated and the requested information is delivered or the demanded task is executed. There are several synonyms available for each command. By using implicit feedback in the voice prompts the driver is informed about what the system has understood. After the first command the user is guided by the system and executes the steps which are suggested and displayed by the system. The GUI supports the speech dialog by showing the "speakable" commands as widgets on the screen (see Section 2.2). A sample dialog is illustrated in the following:

Driver:	*Book a hotel.*
System:	*Where would you like to book a hotel?*
Driver:	*In Berlin.*
System:	*When do you want to arrive in Berlin?*
Driver:	*Tomorrow.*
System:	*How long would you like to stay in Berlin?*
Driver:	*Until the day after tomorrow.*

[1]http://www.opends.eu/
[2]http://www.gethomesafe-fp7.eu
[3]http://www.hrs.com

2.1.2 Conversational Speech Dialog Strategy

In the conversational dialog strategy the dialog initiative switches during the speech interaction. The driver is able to speak whole sentences where multiple parameters can be set within one single utterance. Thereby, the dialog can be more natural, flexible and efficient. The driver is informed about what the system has understood by using implicit feedback. The GUI does not present the "speakable" commands on the screen. In order to indicate the possible functions icons are displayed (see Section 2.2). A sample dialog is presented in the the following:

Driver:	*I would like to book a hotel in Berlin.*
System:	*When do you arrive in Berlin?*
Driver:	*I'll arrive tomorrow and leave the day after tomorrow.*

As illustrated in the example the driver can already indicate some input parameters when addressing the system for the first time. The system verifies which input parameters are missing in order to send a hotel request. The system prompts the user and collects the missing information. Although the system asks for only one parameter, the user is able to give more or other information than requested.

2.2 GUI Design

The different GUIs have been designed in order to support the speech dialog strategies and to evaluate the effect of the GUI on usability and driving performance. The different GUIs have been customized corresponding to the dialog strategies only as much as necessary since an objective comparison is targeted. When designing the screens we followed the international standardized AAM-Guidelines (Driver Focus-Telematics Working Group, 2002).

2.2.1 Command-based GUI Design

In the command-based dialog strategy the driver uses commands to speak to the system. In order to give the driver an understanding of the "speakable" commands, the speech dialog is supported by the GUI. For that reason the currently possible speech commands are displayed on the screen at all times which may lead to a high visual distraction. Hence, in automotive terms the command-based speech dialog strategy is also called "speak-what-you-see" strategy.

Figure 1(a) illustrates the main screen of the hotel booking application at the beginning of the hotel booking dialog. Here, the first input parameter "destination" ("Ziel" in German) is requested by the system. Afterwards the user is guided step-by-step by the system. When the driver has given the requested information, a new widget appears on the screen and the system asks the driver for the corresponding input.

2.2.2 Conversational GUI Design

In the conversational dialog strategy the driver can speak freely and does not have to use certain commands. There is no need to give the driver a visual feedback of the currently "speakable" commands whereby the visual distraction may be lowered. For that reason, the content on the head unit screen does not have to indicate the possible options to proceed with the speech dialog. The sub-function line which was used to indicate the available commands is replaced by only few symbols which resemble the current GUI state. Figure 1(b) shows the form filling main screen at the beginning of the speech interaction where the user is already able to input several parameters at once.

2.2.3 Without GUI

We also investigated the need for a visual feedback, why the two speech dialog strategies are also evaluated "without GUI". In this case, without GUI means that no content information is displayed on the screen. However, a visual feedback which indicates if the user is allowed to talk is presented in the top bar of the screen (see Figure 1(c)).

3 Evaluation

3.1 Method

3.1.1 Participants

The experiment was conducted at DFKI, Saarbrücken. In total, 24 German participants (mainly students) participated in the experiment. All participants received a monetary expense allowance and possessed a valid driver's license. Due to missing data recordings during the experiment data of 1 participant had to be excluded from the analyses. The remaining participants comprised 9 male and 14 female subjects and the average age was 26 years (standard deviation (SD) = 4,1). 56,5% of the participants were driving their car at least once a day. 56,5% had little to no experience with speech-controlled devices.

(a) Command-based GUI (b) Conversational GUI (c) “without” GUI

Figure 1: Main Screens at the Beginning of the Interaction.

3.1.2 Experimental Design

Four different HMI concept variants were evaluated in a 2x2 (speech dialog strategy: command-based vs. conversational, GUI: with vs. without) design. The Command-based and Conversational GUI were only used with the corresponding dialog strategy. The 4 HMI concepts were the following:

- Command-based speech dialog (“Comm”)
  - with GUI (“CommGUI”) and
  - without GUI (“CommNoGUI”)
- Conversational speech dialog (“Conv”)
  - with GUI (“ConvGUI”) and
  - without GUI (“ConvNoGUI”)

Each participant encountered all four conditions (“within-design”). For each condition, two tasks had to be accomplished. We investigated the participants speech dialog performance and influences on driving performance while using the SDS.

3.1.3 Materials

Speech Dialog Prototypes: In the experiment, the speech dialog prototypes described in Section 2 have been used. In order to explain the functionality and the control of the SDS prototypes to the user, instruction videos for each speech dialog strategy were presented. By presenting tutorial videos, we ensured that each participant was given identical instructions.

During the experiment, participants had to solve several tasks: They had to book a certain hotel according to given search parameters. The tasks were verbalized as little stories which contained the necessary parameters in a memorable manner. A sample task in English is presented below:

> *Imagine, you and your colleague are on the way to Cologne for a two-day meeting right now. You need two single rooms for these two nights which you have not booked, yet. Your appointment takes place in the city center of Cologne, where you would like to spend your night. Please look for a matching hotel for those nights.*

In total, participants had to perform 16 tasks. Four tasks were used as sample tasks to familiarize participants with the respective speech dialog strategy after showing the instruction video. The remaining eight tasks were used for the data collection.

Questionnaires: During the experiment different questionnaires were used:

- Preliminary Interview: In a preliminary questionnaire we collected demographical information (age, gender, etc.) about the participants. Furthermore, we surveyed driving habits, experience with speech-controlled devices, and hotel booking habits.
- SASSI questionnaire (Hone and Graham, 2001): The SASSI questionnaire covering 6 dimensions consists of 34 questions and is widely used to measure subjective usability evaluation of SDS.
- DALI questionnaire (Pauzie, 2008): The DALI questionnaire covers 6 dimensions in order to evaluate the user’s cognitive load. The applied questionnaire consisted of 7 questions covering each dimension and an additional question addressing the manual demand.
- Final Interview: This questionnaire was designed to allow for a direct comparison of the respective SDS prototypes at the end of the experiment. Each participant had to rate the different SDS on a scale from 1 - 10 regarding several subjective measures. For each of the six SASSI dimensions, one question was asked. Additionally, we asked questions to directly compare cognitive load and to get information about the participants’ personal preference of interaction style with the system at different sub dialogs.

Driving Simulation Setup: The experiment was conducted in the driving simulator at DFKI’s “future lab” (see Figure 2). The participants were sitting on the driver's seat in a car which was placed in front of a canvas onto which the driving simulation was projected. The participants controlled the driving simulation by the car steering wheel and pedals. During the experiment the examiner was sitting on the passenger seat.

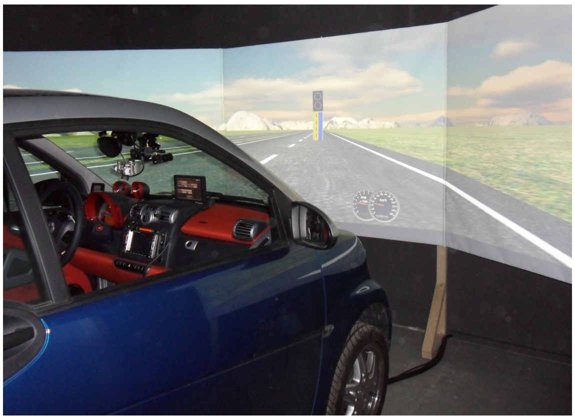

Figure 2: DFKI Driving Simulator Setup.

Previous driving simulation studies employ the standard Lane Change Test (LCT) by Mattes (2003). However, this driving task does not continuously mentally demand the user and thus, does not reflect the real cognitive load while driving. Furthermore, LCT is based on single tracks which limits the recordings to a certain time. We employed the ConTRe (Continuous Tracking and Reaction) task as part of the OpenDS[1] driving simulation software which complements the de-facto standard LCT including higher sensitivity and a more flexible driving task without restart interruptions. The steering task for lateral control resembles a continuous follow drive which will help to receive more detailed results about the two diverse dialog strategies. Furthermore, mental demand is addressed explicitly by employing an additional reaction task implemented as longitudinal control. A detailed description of the ConTRe task can be found in Mahr et al. (2012).

In the experiment, after giving the participant the hotel booking task instructions, the experimenter started the driving simulation. When the participant has crossed the start sign in the simulation he had to begin the speech dialog. When the hotel booking was completed, the experimenter stopped the driving simulation. Thereby, driving performance was only recorded during the speech dialog.

3.1.4 Procedure

In the experiment, 4 conditions were evaluated: The conversational speech dialog (with and without GUI) and the command-based speech dialog (with and without GUI). We did not randomize all four conditions, because the participants might have been confused if the speech dialog styles vary too often. Therefore, we decided to employ dialog styles blockwise (see Figure 3). In one block, only one speech dialog variant with the two GUI conditions was tested. The order of the two blocks was counterbalanced between participants to control for learning and order effects. Thereby, half of the participants were first introduced to the command-based dialog, whereas the other half of the participants started with the conversational dialog. Furthermore, the order of GUI conditions within one block was balanced between participants. In each of the four conditions, the participants had to perform two tasks. The order of the tasks was the same for all participants regardless of the system condition. Hence, all tasks were encountered in all dialog and GUI combinations. When the second task was finished, participants had to fill out the SASSI and the DALI questionnaire for each condition.

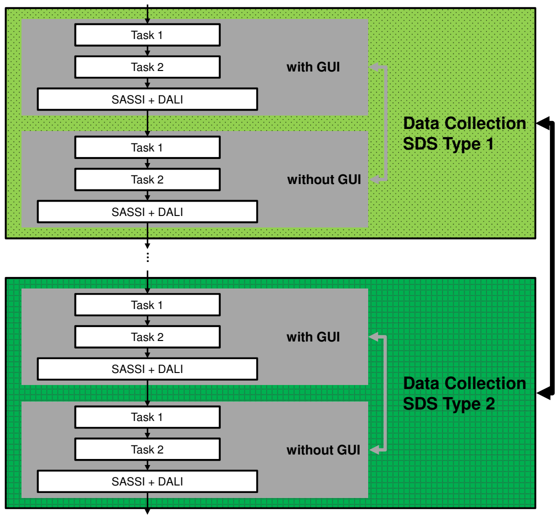

Figure 3: Experiment Structure.

The overall procedure of the experiment is illustrated in Figure 4. At the beginning of the experiment, participants had to fill out the preliminary questionnaire. Afterwards they had the possibility to get to know the driving simulation in a test drive lasting at least 4 minutes. After the test drive, the participants completed a 4 minutes baseline drive and had to fill out the DALI questionnaire afterwards to assess driving performance without secondary task. Next, the participants were shown the video of their first speech dialog variant and became familiar with the SDS by performing the 4 explorative tasks. Subsequently, participants performed the first SDS condition (SDS Type 1) both with and without GUI. After testing SDS Type 1, SDS Type 2 was introduced by presenting its instruction video and again the explorative tasks were performed. Participants performed the second SDS condition (SDS Type 2) also with and without GUI. Finally, participants completed a second baseline drive and filled out the final questionnaire.

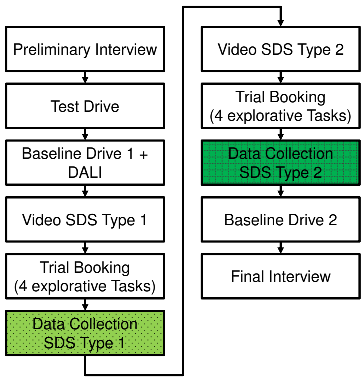

Figure 4: Overall Procedure of the Experiment.

3.1.5 Dependent Variables

In the experiment, we collected several types of data to evaluate the speech dialog and the driving performance data. During speech interaction the SDS produces log files, which contain the link to the recorded audio file of the spoken user utterance, the speech recognizer result, the interpretation of the natural language understanding (NLU) module, and the text-to-speech (TTS) output. Based on the log file, the whole speech dialog can be reconstructed. The driving simulation OpenDS also produces log files at runtime, which contain the steering wheel deviation for lateral control and the reaction times for longitudinal control for each recorded time frame. During the experiment, the examiner was observing the test procedure in order to take notes on task success. Based on the collected data, the measures illustrated in Table 1 were computed in order to evaluate the speech dialog and the driving performance. A detailed description and definition of the measures can be found in (Möller, 2005).

In this preliminary analysis, due to time constraints, only the first block of each participant could be transcribed and analyzed. In this report,

	Measure	**Data Source**
Speech Dialog Performance Measures	TS	Observations
	NoT	SDS logs
	DD	SDS logs
	CER	SDS logs
	Subjective Usability Assessment	SASSI, Final Interview
Driving Performance Measures	MDev	OpenDS logs
	Subjective Assessment of Cognitive Load	DALI, Final Interview

Table 1: Evaluation Measures of the Experiment.

we focus on the SDS performance. Based on the observations the task success (TS) of each speech dialog is assessed. The speech dialog logs are used to compute the Number of Turns (NoT) and the dialog duration (DD) of each dialog. We assess the concept error rate (CER) of each user utterance within a dialog instead of the word error rate (WER) since this value is crucial to a successful speech dialog. A subjective usability assessment is achieved by employing the SASSI questionnaire. Based on the OpenDS logs we compute the mean deviation (MDev) of the steering wheel. In the next step, the reaction time, the DALI questionnaire and the final interview are analyzed.

Overall, we expect better usability evaluation for the conversational dialog conditions compared with the command-based condition. The participants will accept the conversational dialog better than the command-based dialog because if reflects the human-human communication. Furthermore, we expect the conversational dialog to distract less than the command-based dialog because it is easier to control. Generally, a visual feedback makes it more comfortable to interact with an SDS. Therefore, we expect the participants to accept the SDS with GUI better than without GUI. However, concerning the influence of the GUI on the driving performance, we expect the GUI to cause more driver distraction due to the glances onto the GUI screen.

3.2 Results

In the following, the preliminary results concerning SDS quality and driving performance are presented. In total, 48 command-based dialogs and 44 conversational dialogs were transcribed and analyzed. First, the results of the speech dialog evaluation are described, followed by the results of the driving performance evaluation. When comparing the two speech dialog strategies (“Comm” vs. “Conv”) dependent t-tests for paired examples have been applied. Concerning the comparison of the 4 GUI conditions (“CommGUI” vs. “CommNoGUI”, “ConvGUI” vs. “ConvNoGUI”) the repeated measures anova test was applied. For each comparison, a significance level α =0,05 was assumed.

3.2.1 Speech Dialog

In this Section, first, the results of the speech dialog performance measures are presented, followed by the results of the questionnaires.

Task Success: In the first block of each experiment, each participant had to solve 4 tasks while data was recorded. Each of the 92 dialogs were finished with a hotel booking. If the participant booked a hotel, which did not match the task requirements the task was annotated as failed. Figure 5 shows the percentage of solved tasks for both speech dialog strategies (left) and additionally split according to the two GUI conditions (right). Using the command-based SDS prototype, participants were able to solve 95,8% of the tasks. 93,8% of the tasks could be solved when using the conversational prototype. Participants solved tasks more effective when using the command-based prototype with GUI than without GUI. In contrast, the participants solved more tasks successfully when using the conversational prototype without GUI than with GUI. However, none of the differences was found to be significant.

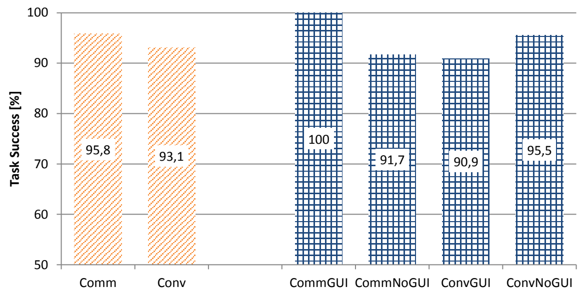

Figure 5: Overall TS rates.

Number of Turns: Figure 6 presents the average NoT. The high number of turns is due to the list browsing the user has to perform in order to find the matching hotel. Using the conversational SDS prototype, significantly fewer dialog turns were needed than using the command-based SDS prototype (p=0,047). The conditions without GUI needed less turns than the conditions with GUI. However, no significant differences were found when comparing the conditions with GUI with the conditions without GUI.

Dialog Duration: In Figure 7 the average DD is illustrated. The dialogs of the conversational

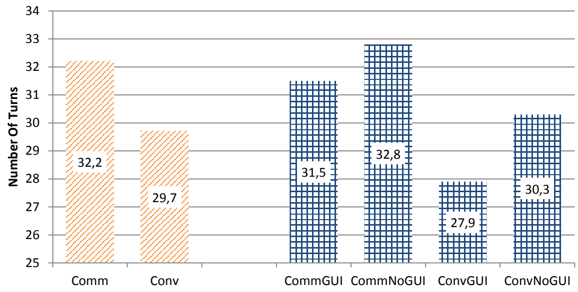

Figure 6: Average NoT per speech dialog.

speech dialogs were significantly shorter than the command-based speech dialogs (p=0,003). Comparing the GUI conditions within one speech dialog strategy, it seems that participants using the conversational speech dialog needed less time to accomplish a task when they could use the GUI. However, there was no significant difference revealed. Concerning the GUI conditions of the command-based dialog, no significant differences could be found, too.

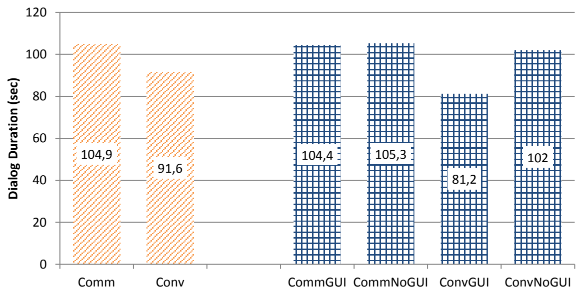

Figure 7: Average DD per speech dialog.

Concept Error Rate: The average CER per dialog is significantly smaller in the command-based speech dialog compared to the conversational speech dialog strategy (p=0,02) (see Figure 8). When comparing the GUI conditions within one speech dialog strategy, it seems that less concept errors occurred when the participants used the SDS prototypes supported by a GUI. However, no significant differences were found.

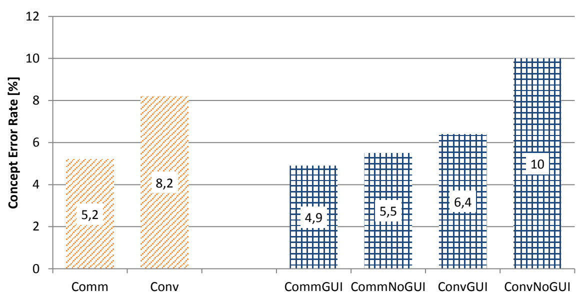

Figure 8: Average CER per speech dialog.

SASSI: The overall result of the SASSI questionnaire is illustrated in Figure 9. All SDS achieve a positive usability assessment. The conversational dialog is slightly better accepted by the user. It seems that the users accept the SDS supported by a GUI better than without a GUI. However, for none of the comparisons significant differences were found.

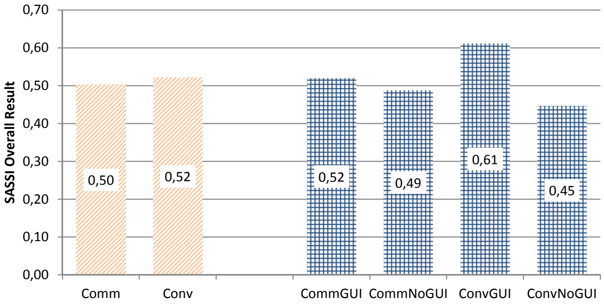

Figure 9: Overall SASSI result per speech dialog.

3.2.2 Driving Performance

In this Section a preliminary driving performance result is presented.

Mean Deviation: Figure 10 shows the MDev of the baseline drive (left), both speech dialog strategies (middle) and additionally split according to the two GUI conditions (right). The MDev of the baseline drive is 0,1. The MDev was significantly smaller when the participants used the command-based speech dialog (p=0,01) while driving compared to the conversational dialog. No significant differences were found when comparing the conditions with GUI with the conditions without GUI.

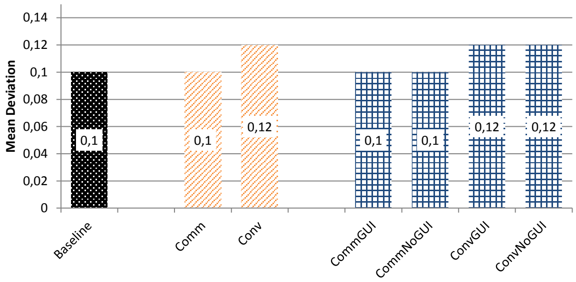

Figure 10: Average MDev per speech dialog.

3.3 Discussion

The preliminary results show that the participants were able to successfully finish the tasks with both SDS prototype variants. All SDS prototypes achieved a positive subjective usability assessment. Although the CER is higher when using the conversational dialog, it performs more efficient than the command-based dialog which is due to the possibility to input multiple parameters at once. The MDev of the baseline drive is as high as when using the command-based speech dialog while driving. Usually, one would expect a better driving performance when performing no secondary task. However, the ConTRe task is a quite difficult task since it continuously mentally demands the user. Therefore, the MDev is relatively high when only the driving task is performed. The conversational speech dialog distracts more from driving than the command-based dialog. Using the command-based dialog, the user is guided by the system step-by-step, which makes it easier to use. The mental demand when using the command-based SDS might be lower and therefore, this dialog strategy might be less distractive.

Concerning the comparison of the GUI conditions the results indicate that the conditions with GUI are more user-friendly than the conditions without GUI. However, we did not find any significant differences, yet, since the data set is too small when comparing the GUI conditions. When the whole data set of the experiment is analyzed further significances might be revealed.

4 Conclusions

This paper reports from a very recent driving simulation study and its preliminary results which are conducted in order to compare different speech dialog strategies. The use of command-based and conversational SDS prototypes while driving is evaluated on usability and driving performance. Different GUIs are designed in order to support the respective dialog strategy the most and to evaluate the effect of the GUI on usability and driver distraction. The preliminary results show that the conversational speech dialog performs more efficient than the command-based dialog. However, the conversational dialog distracts more from driving than the command-based. Furthermore, the results indicate that an SDS supported by a GUI is more efficient and better accepted by the user than without GUI.

In the next step, the data set will be analyzed on all mentioned usability and driving performance measures. The different subdialog types of each dialog will be investigated in detail on dialog performance and speaking styles. Furthermore, cross-links between subdialogs and the driving performance measures are analyzed.

References

L. Devillers and H. Bonneau-Maynard. 1998. Evaluation of dialog strategies for a tourist information retrieval system. In *Proc. ICSLP*, pages 1187–1190.

Driver Focus-Telematics Working Group. 2002. Statement of principles, criteria and verification procedures on driver interactions with advanced in-vehicle information and communication systems. *Alliance of Automotive Manufacturers*.

Governors Highway Safety Association. 2011. Distracted driving: What research shows and what states can do. Technical report, U.S. Department of Transportation.

H. Hofmann, U. Ehrlich, A. Berton, and W. Minker. 2012a. Speech interaction with the internet - a user study. In *Proceedings of Intelligent Environments*, Guanajuato, Mexico, June.

H. Hofmann, Anna Silberstein, U. Ehrlich, A. Berton, and A. Mahr. 2012b. Development of speech-based in-car hmi concepts for information exchange internet apps. In *Proceedings of International Workshop on Spoken Dialogue Systems*, Paris, France, December.

K. S. Hone and R. Graham. 2001. Subjective assessment of speech-system interface usability. In *Proceedings of Eurospeech*.

Angela Mahr, Michael Feld, Mohammad Mehdi Moniri, and Rafael Math. 2012. The ConTRe (continuous tracking and reaction) task: A flexible approach for assessing driver cognitive workload with high sensitivity. In *Adjunct Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pages 88–91, Portsmouth, United States.

Stefan Mattes. 2003. The lane-change-task as a tool for driver distraction evaluation. *Proceedings of IGfA*, pages 1–30.

Sebastian Möller. 2005. Parameters describing the interaction with spoken dialogue systems. ITU-T Recommendation Supplement 24 to P-Series, International Telecommunication Union, Geneva, Switzerland, October. Based on ITU-T Contr. COM 12-17 (2009).

Hartmut Mutschler, Frank Steffens, and Andreas Korthauer. 2007. Final report on multimodal experiments - part 1: Evaluation of the sammie system. d6.4. talk public deliverables. Technical report.

Annie Pauzie. 2008. Evaluating driver mental workload using the driving activity load index (DALI). In *Proceedings of European Conference on Human Interface Design for Intelligent Transport Systems*, pages 67–77.

Matthias Peissner, Vanessa Doebler, and Florian Metze. 2011. Can voice interaction help reducing the level of distraction and prevent accidents? meta-study on driver distraction and voice interaction. Technical report, Fraunhofer-Institute for Industrial Engineering (IAO) and Carnegie Mellon University.