# How Users React to Proactive Voice Assistant Behavior While Driving

**Maria Schmidt[1,2], Wolfgang Minker[2], Steffen Werner[3]**
[1]Mercedes-Benz AG, [2]Ulm University, [3]Daimler Trucks North America LLC
{maria.m.schmidt, steffen.s.werner}@daimler.com, wolfgang.minker@uni-ulm.de

## Abstract

Nowadays Personal Assistants (PAs) are available in multiple environments and become increasingly popular to use via voice. Therefore, we aim to provide proactive PA suggestions to car drivers via speech. These suggestions should be neither obtrusive nor increase the drivers' cognitive load, while enhancing user experience. To assess these factors, we conducted a usability study in which 42 participants perceive proactive voice output in a Wizard-of-Oz study in a driving simulator. Traffic density was varied during a highway drive and it included six in-car-specific use cases. The latter were presented by a proactive voice assistant and in a non-proactive control condition. We assessed the users' subjective cognitive load and their satisfaction in different questionnaires during the interaction with both PA variants. Furthermore, we analyze the user reactions: both regarding their content and the elapsed response times to PA actions. The results show that proactive assistant behavior is rated similarly positive as non-proactive behavior. Furthermore, the participants agreed to 73.8% of proactive suggestions. In line with previous research, driving-relevant use cases receive the best ratings, here we reach 82.5% acceptance. Finally, the users reacted significantly faster to proactive PA actions, which we interpret as less cognitive load compared to non-proactive behavior.

**Keywords:** voice assistant, proactivity, user reactions

## 1. Introduction

These days more and more people use Personal Assistants (PAs) via voice, such as Google Assistant and Amazon Alexa at home (Mittal et al., 2015), Apple Siri and Microsoft Cortana on the smartphone (Kepuska and Bohouta, 2018), or Mercedes-Benz MBUX Voice Assistant and BMW Intelligent Personal Assistant in the car (Braun et al., 2019). These are available on many different devices and offer convenient functionalities in different environments, such as setting reminders, navigating through traffic, or sending messages to friends and colleagues. While serving the users' needs, PAs constantly collect personal data in order to personalize their services and adapt their behavior. Adaptation needs not only to be performed in terms of adapting to the user, but also in terms of adapting to the situation, in which these assistants are used. Especially, when the user drives or is busy with another task at home (e.g., cooking), the interaction with a PA is only the secondary task. Thus, user experience designers need to focus on the user's cognitive load in such settings, too. (Gabaude et al., 2012; Villing, 2009b; Villing, 2009a) In order to investigate how users perceive proactive voice output while driving, we conducted a Wizard of Oz study in a driving simulator with 42 participants. We varied traffic density during a highway drive to induce different levels of cognitive load . Furthermore, we permuted six in-car specific use cases and added a non-proactive control condition with the same six use cases. By employing a subjective DALI questionnaire (Pauzié, 2008) we assessed the users' cognitive load during the interaction with the two PA variants. Additionally, we let the participants rate both PA variants through the SASSI questionnaire (Hone and Graham, 2000) both while driving and afterwards. The results show that the proactive assistant behavior has been rated similarly positively as for the non-proactive one, where users initiated the dialog. In line with previous research, the most driving-relevant use cases were rated the best.

## 2. Related Work

In this work, we take a look at proactivity in PAs, but relate it to the user's cognitive load during the interaction as well. Concerning proactivity, (Buss et al., 2011) and (Schrempf et al., 2005) focus on the proactive actions of robots, which are sometimes related to the user's recognized intention. But the proactivity lies in the proactive planning or execution of tasks and does not contain proactive dialog behavior. Regarding the latter, Nothdurft et al. (2015) declare appropriate interaction strategies for proactive dialogue systems as an open quest. L'Abbate (2007) suggested in his dissertation how to model proactive behavior of conversational interfaces: He defined that the assistant takes over the initiative in problematic and unclear situations in a virtual risk management advisor scenario. Concerning cognitive load, Lindström et al. (2008) have shown that there is an effect of cognitive load on disfluencies when the user speaks to in-vehicle spoken dialog systems. In (Angkititrakul et al., 2007) the topic is discussed in a broader manner, modeling driver-behavior and assessing distraction for these in-vehicle speech systems. Radlmayr et al. (2014) present how traffic situations and non-driving related tasks (such as talking to a PA) affect the take-over quality in highly automated driving, whereas the works by Villing (2009b; 2009a) as well as Fors and Villing (Fors and Villing, 2011) are exactly focusing on cognitive load while driving and talking to a dialog system or voice assistant. While Hamerich (2007) did not take cognitive load into account, but presented proactive dialogs relying on the context of real-time traffic situations (at that time transmitted via TMC). According to previous research by Schmidt et al. (2019b), proactivity and certain use cases that are closely related to tasks while driving are preferred by users during in-car HCI. Based on their findings and the prior work of Hamerich (2007), we designed the usability study presented in this work. To the best of our knowledge, we are the first ones to systematically combine all areas: proactive voice assistant behavior, cognitive load,

Figure 1: Driving simulator with operator desk (Schmidt et al., 2020)

and the subsequent user acceptance during the interaction (secondary task) while driving (primary task). In this work, we decided to assess our subjects' cognitive load while driving in a subjective manner. For this purpose, we rely on the DALI questionnaire as introduced in (Pauzié, 2008). Regarding evaluation, we are assessing both the proactive and the non-proactive assistant (control condition) ratings by means of the SASSI questionnaire (Hone and Graham, 2000).

## 3. Driving Simulator Study

In this study (cf. (Schmidt et al., 2019a; Schmidt et al., 2020)) 42 subjects completed the whole experiment in the driving simulator. The distribution of sexes was almost even with 22 male (52.4%) and 20 female (47.6%) subjects. Their age averaged out on 43.7 years, ranging from 22 to 65 years. We paid attention to a balanced distribution of yearly kilometrage among the participants. This is due to the fact that driving was the primary task during the experiment and driving habits could also influence the subjects' perceived cognitive load, though we did not ask them to perform challenging driving maneuvers. As shown in section 4., the yearly kilometrage does not have a significant effect on the subjects' response times.

Figure 1 depicts the setup of the driving experiment: it contained a fixed-base simulator with a $180°$ screen in a room with controlled light and temperature conditions. The operator desk was located in the same room, but could not be observed while the participants sat on the driver's seat.

Methodically, the study was designed as a two factor within-subject experiment. Figure 2 illustrates that each subject interacted with both a proactive (P) as well as a non-proactive (NP) voice assistant, separated by a short driving break in which the first assistant was rated. In between the interaction with each of the assistants, the traffic density was varied from low to high or vice versa. Consequently, every subject interacted with both assistants and experienced both traffic conditions during the respective interaction phases. The order in which the assistants and traffic conditions were presented was permuted, so that we created the following four different experiment procedure variants:

- Variant 1: starting with NP and low traffic, switching to high traffic; switching to P while remaining in high traffic, ending with P and low traffic

- Variant 2: starting with NP and high traffic, switching to low traffic; switching to P while remaining in low traffic, ending with P and high traffic

- Variant 3: starting with P and low traffic, switching to high traffic; switching to NP while remaining in high traffic, ending with NP and low traffic

- Variant 4: starting with P and high traffic, switching to low traffic; switching to NP while remaining in low traffic, ending with NP and high traffic

The subjects were only informed that they were interacting with assistant A or B, but they did neither know about the current interaction type (NP or P), nor about the traffic condition.

As mentioned beforehand, the tested voice assistant variants were operated in a Wizard of Oz (WOz) setup. All potential (permuted) dialog paths were modeled in a rule-based manner. We prepared up to four different possible responses, depending on the subject's input, and were also able to repeat selected phrases, if subjects requested for it. In addition to a synthesized female assistant voice (same as in current Mercedes-Benz models), we also integrated a synthesized male voice to pose questionnaire items about subjectively perceived cognitive load (DALI) and system ratings (SASSI). In this way, "he" acted as a standardized virtual co-examiner to further establish controlled conditions. Controlled conditions were also the main reason why we chose a WOz setup: when testing proactive dialog behavior, we did not want potentially varying speech recognition performance to influence our results. Furthermore, as we have to cope with subject's cognitive load, we did not want to try out proactive voice assistant behavior in the wild, because we did not know whether it would be too demanding for any subject.

In the following we describe an exemplary experiment procedure of our driving simulator study (cf. (Schmidt et al., 2019a)). It had three parts: briefing, main experiment (the drive), and debriefing. First, the examiner welcomed the subject and led them to the briefing room. Then the subject was informed about audio- and video-taping which was agreed by signing a respective form. Following this, the subjects should fill out a general questionnaire on experience with PAs, technical affinity, their own car etc. Afterwards, the subjects were led to the cabin and introduced to the car for the main experiment. The examiner informed about the video camera and the two-way intercommunication system inside the car. Furthermore, they offered assistance in case the subject needs help at any point during the study. The subjects were given driving instructions: stay on the right lane, drive around 110 km/h and follow the lead car, do not overtake. The examiner gave the Empatica E4 wristband to the subject and checked that it was worn correctly. After answering potential questions, the examiner took a seat at the operator's desk. They assured that the subject can hear them (and vice versa), that the Empatica E4 was recording properly. The simulated car was situated in
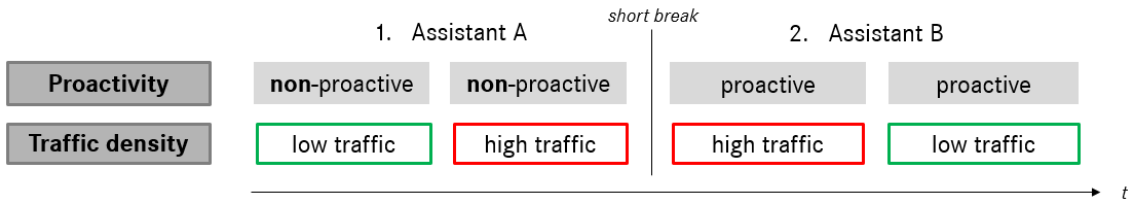
Figure 2: Exemplary experiment procedure (Schmidt et al., 2019a)

a service area next to a three lane highway. When the subjects were asked to start driving, they entered the highway with no other traffic (neither same nor opposite direction). After around one minute, the subject closed up to the lead vehicle, which they should follow at all times. It drove with a constant speed of 110 km/h. After around two minutes of the baseline drive the examiner reassured that the subject feels well (no motion sickness due to graphic projection). After this point the controlled experiment started and only the WOz assistant(s) talked to the subject for the remaining drive (exception: in the middle of the drive, when the subject stopped at a service area, the examiner checked again for the subject's well-being). Following the baseline drive (around five minutes), the traffic simulation started and cars in the same and opposite direction were shown. After the drive was finished, the examiner prepared the cabin and the simulation setup for the next subject. They took back the Empatica E4 wristband and led the current subject to the debriefing room. The examiner asked the subject to fill out a final short questionnaire about the usefulness of the presented use cases, and then saw them off.

To manipulate the subjects' cognitive load, we varied the traffic density during the experiment. After the baseline drive without any traffic, the neural network traffic simulation was being started. Depending on the variant, it started with a low or high traffic condition. In the low traffic condition, 10 cars were simulated per 1 km on the three-lane highway. In the high traffic condition, 40 cars were simulated per 1 km on the three-lane highway. We determined these numbers experimentally, taking the average speed and speed variations during these situations into account which influence the subjects' level of exposure and the total time spent driving. If we would have increased the number of cars from 10 to more than 50, there would have appeared highly demanding braking situations when traffic slows down, comparable to a real "stop and go" traffic. Because this might have caused many motion sick subjects, we limited the high traffic condition to 40 cars per 1 km. Additionally to the traffic density, the traffic simulation included different types of drivers (excluding very aggressive ones). As described above, our subjects interacted with two different assistants In order to be able to compare both interactions to each other, we controlled the experiment by applying the same six use cases to the P and NP assistant, respectively (see examples in Table 1). Overall we presented the subjects five driving-related and one not driving-related use cases. Most driving-related use cases were close to the navigation domain, such as refueling or rerouting. The order in which the use cases were presented was permuted across subjects and variants. After three use
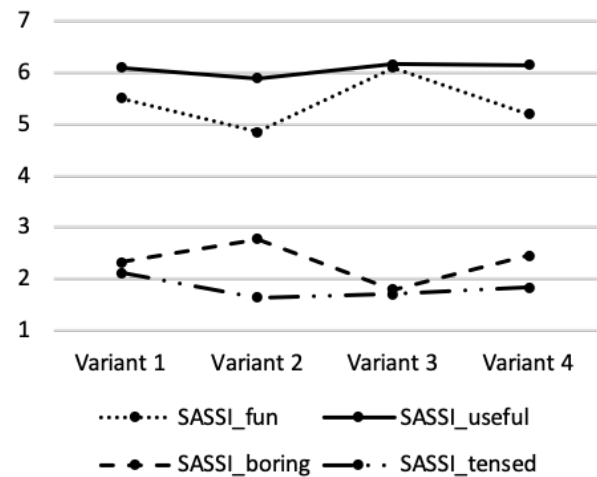


Figure 3: SASSI mean ratings by variant (Schmidt et al., 2020)

cases, i.e. when either the assistant or the traffic condition was changed, the virtual co-examiner posed the same five SASSI and six DALI questions.

## 4. Results & Evaluation

### 4.1. User Satisfaction & Cognitive Load

In this section we shortly summarize the results of the user satisfaction as well as subjective cognitive load ratings, elicited by means of SASSI and DALI questionnaire items. (Schmidt et al., 2020) Figure 3 illustrates the mean SASSI ratings across the 4 auditory items *having fun* using the system, finding the system *useful*, finding it *boring*, or *feeling tensed* while using the system, per variant. It shows that the negatively connoted items *boring* and *tensed* got relatively low ratings on the 7-point Likert scale from *I do not agree at all* to *I totally agree*. Coherently, the positive items *fun* and *useful* were rated relatively high. Generally speaking, there are no noteworthy effects, but there is a significant rise of *fun* between variant 2 and 3 ($p < 0.05$), which is reversely reflected in the negative item *boring*.

Similarly, the results from the DALI questionnaire (Schmidt et al., 2020) show only minor variations between the non-proactive and proactive assistants. There are also no big differences in the ratings relating to high and low traffic densities, respectively. For the latter, we would have expected more distinct results.

Table 1: Sample Dialogs (Schmidt et al., 2019a)

| Assistant Type | Interlocutor | Sample Dialogs |
|---|---|---|
| Non-Proactive | examiner | *Please express your request to refuel.* |
| | customer | Hey Mercedes, I need to refuel. |
| | vehicle | `The next gas station is located at a highway service area in 10 kilometers. Should I navigate you there?` |
| | customer | Yes, please. |
| | vehicle | `Ok, I set the gas station as an intermediate stop.` |
| Proactive | vehicle | `Your remaining fuel range is 150 kilometers. Should I already search for a gas station for you?` |
| | customer | Yes, please. |
| | vehicle | `Ok, the next gas station is located at a highway service area in 30 kilometers. Should I navigate you there?` |
| | customer | Yes, please. |
| | vehicle | `Ok, I set the gas station as an intermediate stop.` |
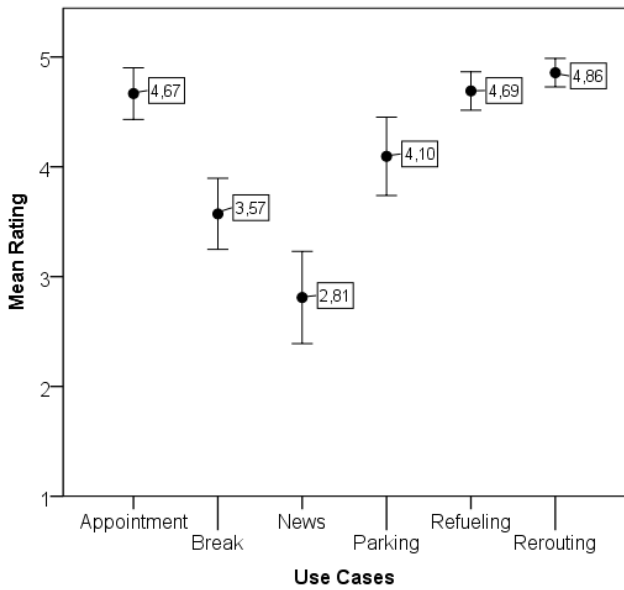


Figure 4: Average ratings of the six different use cases (Schmidt et al., 2020)



Figure 5: Users "Yes"/"No" responses to proactive suggestive questions

## 4.2. User Reactions to Use Cases

Collapsing the rating results from (Schmidt et al., 2020) regarding the use cases, the subjects had a clear preference towards driving-related use cases as shown in Figure 4.

The use cases *Rerouting (4.86)* and *Refueling (4.69)* were rated best and did not receive any low ratings at the same time, i.e. 1 or 2 on a 5-point Likert scale,. *Appointment (4.67)* was rated slightly lower and got ratings between 2 and 5. While all three remaining use cases *Parking (4.10)*, *Break (3.57)*, and *News (2.81)* got rated on the full scale from 1 to 5, *Parking* was clearly the preferred use cases among those three. While the suggestion to *take a break* because of car-detected tiredness of the driver was still perceived as a somewhat positive feature (probably because of safety reasons as shown in (Schmidt et al., 2019b)), informing about *news* was not rated as positively with an average below scale mean. In order to approve the subjects' preference of specific use cases, we performed the Wilcoxon Signed Ranks test for crossfold validation. The follow-
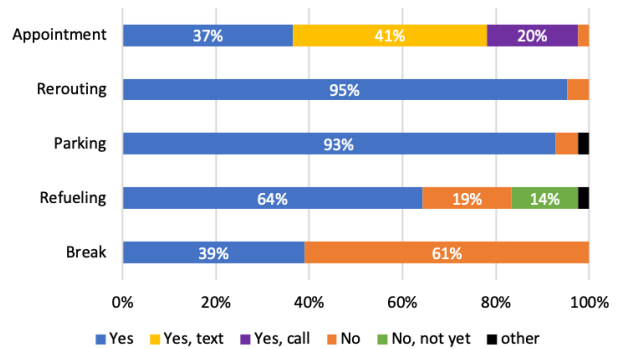
ing use case relations are rated significantly different with $p < 0.00\overline{3}$ (calculated Bonferroni adjustment for 95% confidence interval): Rerouting to News, Parking and Break. Refueling to Break, News, and Parking. Appointment to News and Break. Parking to News, and Break to News.

Moreover, we investigated how large the percentage was among the participants to accept a proactive use case by saying "Yes [...]" or "No [...]", respectively. Figure 5 depicts these percentages for the five most accepted use cases. As we can see compared to the previous numerical analysis, the overall preference of the use cases in their ranking is almost identical, only *Rerouting* and *Appointment* switched rank 1 and 2. Averaging out all five use cases, the participants agreed to 73.8% of proactive suggestions. In line with previous research, driving-relevant use cases receive the best ratings, here we even reach 82.5% acceptance for the top four use cases. Interestingly, in the case of *Appointment*, where the suggestion was to "call Anna or text her that you are running 15 minutes late", 37% just responded "yes" while 41% wanted the PA to send a text message, and 20% wanted the PA to call Anna. With regards to the dialog robustness of a PA to successfully process ambiguous user input, it is good to see that most participants fully specified their desired intent. For the well accepted use case *Refueling*, it is interesting to see that besides 64% acceptance and 19% denial, in 14% of the cases participants are just

not immediately convinced as they would personally like a refueling reminder that takes effect later, i.e., when the remaining range is lower than 150km. For the least accepted use case *Break*, it showed that 61% deny the PA suggestion to take a break, which is also in line with our previous findings.

### 4.3. User Response Times

In this subsection, we analyze the users' response times elapsed between a PA action and the users' respective response. Table 2 shows the results of a Friedman test, performing a pair-wise comparison between the respective proactive and non-proactive use case realizations. For the presented four use cases, the users reacted significantly faster to proactive PA actions, which we interpret as less cognitive load compared to non-proactive behavior. We performed the test on a Generalized Linear Mixed Model and give mean response times in seconds to the reader. Similar to the use case ratings presented in the previous section, this shows the actual impact on the driver's load during proactive voice interaction while driving.

Because other factors could have also been influencing our results regarding the participants' response time, we analyzed the latter as a dependent variable. None of the parameters sex, age and kilometrage have a significant effect on the response time as illustrated in Table 3.

### 4.4. Proactivity in General

In general, participants are satisfied with proactive suggestions by voice assistants. We can already derive this from the positive SASSI ratings, which in some cases even were more positive than the ratings for the NP variants (cf. (Schmidt et al., 2020)). To get a clear picture of the subjects' opinion on proactivity, we posed the following direct question in addition: *One of the two assistants you have experienced, has spoken to you unrequestedly (proactively). How important is it to you that a voice assistant makes suggestions by its own accord (proactively)?* The majority of subjects responded that proactive suggestions by a voice assistant are either rather or extremely important to them. In the free text areas we gave in the questionnaire, a few subjects wrote that proactive suggestions are the actual benefit for them and the assistant appears intelligent through these. As proactivity is a polarizing topic, we also asked the participants to rate whether they wish to be able to deactivate proactivity in a voice assistant. Table 4 shows that only 4 participants do not wish for an deactivating option. 5 participants wish to have a complete deactivation of proactive suggestions. The vast majority of 33 participants wishes to selectively switch the proactivity on or off depending on the respective content, such as appointments, navigation etc.

## 5. Conclusion

In this work, we presented how users perceive and react to proactive dialogs in a driving simulator WOz setting. As drivers are already cognitively occupied with the primary task of driving, proactively triggered interaction by the voice assistant has to remain unobtrusive to regard road safety. While the basic preconditions stayed the same among subjects, the order in which they were confronted

Table 2: Friedman Test: pair-wise comparison of P/NP use cases. Significance values adjusted through Bonferroni correction.

| Use Case | DOF | mean NP | mean P | $p$ |
|---|---|---|---|---|
| Appointment | 11 | 3.79 | 2.23 | 0.003 |
| Rerouting | 11 | 2.56 | 1.51 | 0.006 |
| Parking | 11 | 3.06 | 1.91 | 0.026 |
| Refueling | 11 | 2.58 | 1.60 | 0.001 |

Table 3: Proactivity Results from a Generalized Linear Mixed Model; dependent variable: Response Time

| Parameter | Std. Error | DOF | $p$ | CI 95% |
|---|---|---|---|---|
| sex | 0.193 | 36 | 0.872 | [-0.361,0.424] |
| age | 0.008 | 36 | 0.069 | [-0.031,0.001] |
| kilometrage | 0.000 | 36 | 0.951 | [-0.000,0.000] |

with high or low traffic density varied. We assessed the users cognitive load by means of subjective DALI ratings as well as their user satisfaction by means of SASSI questionnaire items. The results show that proactivity in this context is at least equally likable as non-proactive interaction behavior while driving. At the same time the study subjects significantly rate that they would like to be able to deactivate proactivity for specific functionality (e.g., appointments, navigation etc.). The cognitive load measured by means of DALI items was not diverging at all between variants or assistant/traffic conditions. Furthermore, we analyzed the user reactions: both regarding their content and the elapsed response times to PA actions. Regarding their response content, the participants agreed to 73.8% of proactive suggestions. In line with previous research, driving-relevant use cases receive the best ratings, here we reach 82.5% acceptance. Finally, the users reacted significantly faster to proactive PA actions, which we interpret as less cognitive load compared to non-proactive behavior. We conclude from these findings that though users want to deactivate proactivity, the majority sees it very positively while driving in a controlled condition with several different traffic densities. The vast majority even accepts the proactive suggestions by the respective PA. For future experiments we plan to implement a proactive assistant, while the driving task is taking place in the wild or at least is alternated in such a way that it contains more driving maneuvers

Table 4: "Do you want to be able to deactivate proactive suggestions?"

| | frequency | percent |
|---|---|---|
| no | 4 | 9.5 |
| yes, completely | 5 | 11.9 |
| yes, related to specific contents | 33 | 78.6 |

such as overtaking to be more realistic.

# 6. Bibliographical References

Angkititrakul, P., Kwak, D., Choi, S., Kim, J., PhucPhan, A., Sathyanarayana, A., and Hansen, J. H. (2007). Getting start with utdrive: Driver-behavior modeling and assessment of distraction for in-vehicle speech systems. In *Eighth Annual Conference of the International Speech Communication Association*.

Braun, M., Mainz, A., Chadowitz, R., Pfleging, B., and Alt, F. (2019). At your service: Designing voice assistant personalities to improve automotive user interfaces. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 40. ACM.

Buss, M., Carton, D., Gonsior, B., Kuehnlenz, K., Landsiedel, C., Mitsou, N., de Nijs, R., Zlotowski, J., Sosnowski, S., Strasser, E., et al. (2011). Towards proactive human-robot interaction in human environments. In *2011 2nd International Conference on Cognitive Infocommunications (CogInfoCom)*, pages 1–6. IEEE.

Fors, K. L. and Villing, J. (2011). Reducing cognitive load in in-vehicle dialogue system interaction. In *Proceedings of the 15th Workshop on the Semantics and Pragmatics of Dialogue, SemDial*, pages 55–62. Association for Computational Linguistics.

Gabaude, C., Baracat, B., Jallais, C., Bonniaud, M., and Fort, A. (2012). Cognitive load measurement while driving. in: Human factors: a view from an integrative perspective.

Hamerich, S. W. (2007). Towards advanced speech driven navigation systems for cars. In *2007 3rd IET International Conference on Intelligent Environments*, pages 247–250, Sep.

Hofmann, H., Hermanutz, M., Tobisch, V., Ehrlich, U., Berton, A., and Minker, W. (2016). Evaluation of in-car sds notification concepts for incoming proactive events. In *Situated Dialog in Speech-Based Human-Computer Interaction*, pages 111–124. Springer.

Hone, K. S. and Graham, R. (2000). Towards a tool for the subjective assessment of speech system interfaces (sassi). *Natural Language Engineering*, 6(3-4):287–303.

Kepuska, V. and Bohouta, G. (2018). Next-generation of virtual personal assistants (microsoft cortana, apple siri, amazon alexa and google home). In *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, pages 99–103. IEEE.

L'Abbate, M. (2007). *Modelling Proactive Behaviour of Conversational Interfaces*. Ph.D. thesis, Technische Universität, Darmstadt, November.

Lindström, A., Villing, J., Larsson, S., Seward, A., Åberg, N., and Holtelius, C. (2008). The effect of cognitive load on disfluencies during in-vehicle spoken dialogue. In *Ninth Annual Conference of the International Speech Communication Association*.

Mittal, Y., Toshniwal, P., Sharma, S., Singhal, D., Gupta, R., and Mittal, V. K. (2015). A voice-controlled multifunctional smart home automation system. In *2015 Annual IEEE India Conference (INDICON)*, pages 1–6. IEEE.

Nothdurft, F., Ultes, S., and Minker, W. (2015). Finding appropriate interaction strategies for proactive dialogue systems – an open quest. In *Proceedings of the 2nd European and the 5th Nordic Symposium on Multimodal Communication, August 6-8, 2014, Tartu, Estonia*, number 110, pages 73–80. Linköping University Electronic Press.

Pauzié, A. (2008). A method to assess the driver mental workload: The driving activity load index (dali). *IET Intelligent Transport Systems*, 2(4):315–322.

Radlmayr, J., Gold, C., Lorenz, L., Farid, M., and Bengler, K. (2014). How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 58, pages 2063–2067. Sage Publications Sage CA: Los Angeles, CA.

Schmidt, M., Helbig, D., Bhandare, O., Stier, D., Minker, W., and Werner, S. (2019a). Assessing objective indicators of users' cognitive load during proactive in-car dialogs. In *Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization*, UMAP'19 Adjunct, pages 87–91, New York, NY, USA. ACM.

Schmidt, M., Stier, D., Werner, S., and Minker, W. (2019b). Exploration and assessment of proactive use cases for an in-car voice assistant. *Elektronische Sprachsignalverarbeitung*.

Schmidt, M., Minker, W., and Werner, S. (2020). User acceptance of proactive voice assistant behavior. In *Studientexte zur Sprachkommunikation: Elektronische Sprachsignalverarbeitung 2020*. TUDpress, Dresden.

Schrempf, O. C., Hanebeck, U. D., Schmid, A. J., and Worn, H. (2005). A novel approach to proactive human-robot cooperation. In *ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005.*, pages 555–560, Aug.

Villing, J. (2009a). Dialogue behaviour under high cognitive load. In *Proceedings of the SIGDIAL 2009 Conference: The 10th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 322–325. Association for Computational Linguistics.

Villing, J. (2009b). In-vehicle dialogue management-towards distinguishing between different types of workload. In *Proceedings of SiMPE, Fourth Workshop on Speech in Mobile and Pervasive Environments*.