Silence, Please! Interrupting In-Car Phone Conversations

Soledad López Gambino[1], Casey Kennington[2], and David Schlangen[1]

[1] CITEC, Bielefeld University, Universitätsstraße 25, Bielefeld, Germany
[2] Boise State University, 1910 University Dr., Boise, Idaho, USA

Abstract. Holding phone conversations while driving is dangerous not only because it occupies the hands, but also because it requires attention. Where driver and passenger can adapt their conversational behavior to the demands of the situation, and e.g. interrupt themselves when more attention is needed, an interlocutor on the phone cannot adjust as easily. We present a dialogue assistant which acts as 'bystander' in phone conversations between a driver and an interlocutor, interrupting them and temporarily cutting the line during potentially dangerous situations. The assistant also informs both conversation partners when the line has been cut, as well as when it has been reestablished. We show that this intervention improves drivers' performance in a standard driving task.

Keywords: in-car dialogue · driver distraction · cell phone · interruptions

1 Introduction

Talking on the phone while driving introduces risks which may result in accidents [11,14] and research has also shown that cognitive load is higher when talking on a phone than when talking to a passenger [2]. This difference seems to correlate with co-location: Whereas passengers are aware of the surroundings and can adapt their speech to the demands of the driving situation, a non-situated interlocutor in a telephone conversation does not have enough information to perform this type of adjustment. It has been suggested that this lack of situational awareness can be addressed, to some extent, by providing the interlocutor with real-time visual information of the driving situation [10]. This kind of telepresence results in speech which interferes less with the task of driving and it could be described as a way of "bringing the interlocutor into the scene".

Other efforts have focused on the potential usefulness of interrupting conversations when the circumstances on the road require it. In a Wizard-of-Oz experiment, [5] explored the effects of putting phone calls on hold while the driver needs to perform a more demanding maneuver, as well as of uttering spoken alerts about upcoming situations. The latter alerts proved effective in reducing errors when turning left/right. In the area of human-system dialogue, [8] implemented an information-providing system which interrupts its speech in the case of a demanding driving situation and resumes once the situation has passed. This system not only reduced impact on driving performance in comparison to a system which did not pause its speech, but it also enabled drivers to better remember the information presented by the system.

In the light cast by these studies, we explore the effects of employing an actual system to achieve this adaptive dynamics in human/human phone conversations: a "by-stander" agent that interrupts the conversation in potentially dangerous situations. Subsequently, we test the impact of this system on performance in a driving task.

2 Method

2.1 The System

We developed an *Interrupting Dialogue Assistant* (IDA) which mediates between two participants in a telephone conversation when required by the driving situation. Interruption is triggered by a signal from a component that evaluates the driving situation and judges that the undivided attention of the driver is needed. This immediately cuts the audio line between driver and caller (D and C, respectively). The system then informs D and C about this, as described below. Until it receives a signal indicating that the situation is clear again, the system periodically re-informs C about the line status, and also suppresses any attempts to speak. Finally, it notifies D and C when the line is open again.

The states of the system are shown in Figure 1 in a diagram, with events that trigger transitions, as well as the actions performed by the system at each one of these transitions.
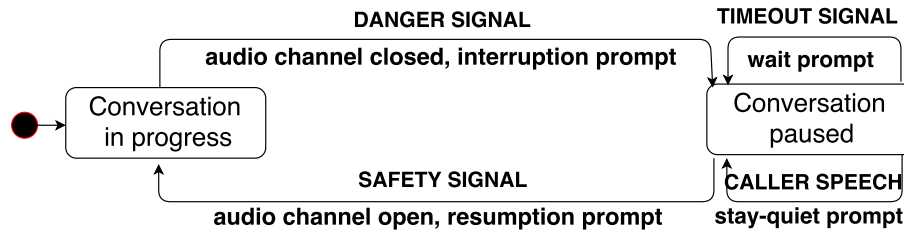

DANGER SIGNAL
audio channel closed, interruption prompt

TIMEOUT SIGNAL
wait prompt

Conversation
in progress

Conversation
paused

SAFETY SIGNAL
audio channel open, resumption prompt

CALLER SPEECH
stay-quiet prompt

Fig. 1. Interaction states

Informing the Caller The system informs the caller of the state of the interaction (line open / closed) through a set of utterances which are synthesized with MaryTTS[3] [13]. Corresponding to the states of the system, there are four types of system acts (see Table 2.1 for some example utterances):

– **Interrupting the conversation:** As soon as the line is cut, the IDA informs C by stating the need for a pause in the dialogue and/or the fact that the driver is busy.
– **Asking for more time:** While the audio line between both participants remains cut, the IDA regularly reminds C to continue waiting, in order to avoid long periods of silence and ensure clarity about the state of the line.
– **Preventing the caller from speaking:** If C speaks at any time while the line is cut, the assistant detects this through Voice Activity Detection (VAD) and informs C of the need to wait for a few more seconds.

---

[3] `http://mary.dfki.de/index.html`

– **Resuming the conversation:** On receiving the appropriate signal, the system announces that the line is open again.

Table 1. System utterances informing the caller about the state of the interaction (German original and English translation)

Interrupting the conversation (interruption prompt)	
Das Gespräch muss einige Sekunden unterbrochen werden.	The conversation has to be interrupted for a few seconds.
Ihr Gesprächspartner ist gerade wieder beschäftigt.	Your conversational partner is busy again.
Diese Unterhaltung muss nochmal kurz pausiert werden.	This conversation needs to be paused briefly again.
Asking for more time (wait prompt)	
Bitte eine Sekunde mehr Geduld.	Please be patient for one more second.
Ihr Gesprächspartner kann Sie noch nicht hören.	Your conversational partner can't hear you yet.
Die Leitung ist bald wieder offen.	The line will soon be reconnected.
Preventing the caller from speaking (stay-quiet prompt)	
Moment bitte.	One moment, please.
Noch nicht.	Not yet.
Bitte warten.	Please wait.
Resuming the conversation (resumption prompt)	
Sie können weiter sprechen.	You can go on speaking.
Jetzt kann der Fahrer wieder hören	The driver can now hear you again.
Die Unterhaltung geht jetzt weiter	The conversation now continues.

Informing the Driver The system also provides information to the driver, although it does so in a different way. Interruption of the audio line is communicated through a bell sound instead of verbally, as we considered that additional speech would be more distracting than a sound [3]. Once drivers have finished maneuvering and are able to resume the conversation, the system produces a short utterance such as *Los geht's* ("Off we go").

2.2 Tasks

Driving Task To test driving performance, we use a variant of the standard Lane Change Task (LCT) [4], implemented in a driving simulation environment (OpenDS[4]). This task consists in reacting to a signal positioned on a gate above the road. The driver, otherwise instructed to stay in the middle lane of a straight five-lane road, must move to the lane indicated by the light, remain there until a tone is sounded, and then return again to the middle lane. Following [6], we introduce an extra level of difficulty, by instructing drivers to perform lane changes at a speed of 60 km/h while the default

---

[4] http://www.opends.eu

driving speed was 40 km/h and the maximum possible speed was 70 km/h. The driving equipment consisted of a 40-inch 16:9 screen and a Thrustmaster PC Racing Wheels Ferrari GT Experience steering wheel and pedal.

Fig. 2. Lane change signal as presented on screen; the green light above the far right lane informs the driver to move into that lane.

Speaking Task To ensure that a lively, continuous conversation would take place between our experiment participants (driver and caller), we instructed them to engage in a role play activity. They were provided with discussion topics beforehand, and instructed to express opposing opinions about them, i.e. to contradict each other. Discussion topics were selected to be related to the experience and interests of the subject population (university students) and to be engaging but not extremely sensitive. The caller was given responsibility for the flow of the discussion and instructed to keep it as entertaining as possible and to switch between topics when necessary.

2.3   Experiment Structure and Conditions

To effectively evaluate our hypothesis that an assistant such as the one described above would result in better driving, we designed three experimental conditions:

NO-TALK Driving only (control condition; including lane changes), no conversation
UNINTERRUPTED Simultaneous driving (including lane changes) and conversation.
INTERRUPTED Simultaneous driving and conversation, but the latter gets interrupted as soon as a lane change is announced and resumed when this maneuver is completed.

The conditions were presented in blocks, as shown in Figure 3. The first block was always NO-TALK, whereas the order of the second and third conditions varied: For half of the participants, the second block corresponded to the UNINTERRUPTED condition and the third one, to the INTERRUPTED condition whereas, for the other half, these two blocks were inverted. Each of the three blocks lasted approximately 10 minutes and was made up of 11 trials: three practice trials and eight experiment trials. Only experiment trials are considered in the results.

The approximate duration of a whole experiment was 40 minutes. Before each phase, participants were given instructions. After completion of all three phases, they filled out a questionnaire. Finally, they swapped roles (the caller became the driver and viceversa) and the whole process was repeated.

Condition: NO-TALK		Condition: **INTERRUPTED**		Condition: **NON-INTERRUPTED**	
training	experiment trials	training	experiment trials	training	experiment trials

Condition: NO-TALK		Condition: **NON-INTERRUPTED**		Condition: **INTERRUPTED**	
training	experiment trials	training	experiment trials	training	experiment trials

Fig. 3. Experiment stages

2.4 Participants and Setup

Sixteen subjects participated in the study, which results in eight pairs participating twice each (due to role-swapping). Driver and caller were placed in two separate rooms, and audio was sent between them through networked computers via Robotic Service Bus [RSB] [15].[5] All participants were students between 20 and 29 years of age and native speakers of German. Ten were female and six male. All of them except for one had a driver's license.

The interrupter was developed using the control component of OpenDial [9] incorporated into InproTK[6,7] [1,7], implementing the state machine described above.

3 Results

The total number of trials recorded was 528: eleven in each of the three conditions, for each of the 16 participants. We excluded training trials from the analysis of driving performance, which left us with 384. In addition, it was necessary to exclude some episodes where the driver never reached the target lane, since this made it impossible to calculate lane-changing time. This resulted in 365 trials useable for analysis. Furthermore, given that two of the road lanes are adjacent to the middle whereas the other two (the external lanes) are not, some further episodes had to be excluded in order to ensure an equal number of changes to adjacent and non-adjacent lanes in all conditions. This left us with 342 trials: 114 for each condition, out of which 64 were changes to adjacent lanes and 50, to non-adjacent ones.

---

[5] https://code.cor-lab.org/projects/rsb

[6] http://opendial.googlecode.com

[7] https://bitbucket.org/inpro/inprotk

3.1 Interruptions

There were eleven interruptions for each driver-caller pair: Three correspond to the training phase and eight, to the experiment trials. Out of a total of 176 interruptions for the 16 participants, the driver was speaking at the moment of the interruption in 75 instances and the caller, in 88; both were speaking simultaneously in three cases, and both were silent in 10 cases. From the moment when callers started hearing the interruption prompt, it took them an average of 1.01 seconds to stop talking (*SD* 1.07). In addition, the caller spoke during the interrupted phase and was told by the system to wait in 26 instances. When callers were interrupted, they left the ongoing word incomplete in 22.7% of the cases, finished the word but left an incomplete syntactic clause in 34.1%, and produced a full clause in 43.2%. The mean duration of the interrupted periods (from the interruption prompt to the resumption prompt) was 11.321 seconds (*SD* 0.637); this was, of course, subject to how fast the driver was able to complete the maneuver.

3.2 Driving Performance

For every trial, we calculated *lane changing time*, which we defined as the time from the moment the lane changing signal appears until the driver reaches the target lane. Lane changing times were almost half a second shorter for the INTERRUPTED condition (4.059 s., *SD* 1.349) than for the NON-INTERRUPTED condition (4.552 s., *SD* 1.646), i.e. drivers were able to complete the change faster when the interrupter was employed. This difference is significant (t-test, $t(15) = 3.37$, $p < 0.01$). On the other hand, no statistically significant difference was found between lane changing times in the INTERRUPTED and in the NO-TALK (4.381 s., *SD* 1.423) conditions, which shows performance when the interrupter was employed was as fast as during the driving-only task, in which no speech was involved. These results can be interpreted as suggesting that our interrupting assistant enabled drivers to complete lane changes sooner by granting them the possibility to concentrate only on the driving, which would constitute an advantage in real life, since swiftness is normally associated with minimization of risks in overtaking maneuvers. On the other hand, it is also possible that the presence of an auditory stimulus (the bell) simultaneous with the visual lane changing signal (the arrow) might have contributed to a faster reaction in the interrupted condition than in the non-interrupted one, in which the moment to change lanes is only announced visually.

3.3 Subjective Evaluation

At the end of each phase of the experiment, participants filled out a questionnaire, rating how pleasant they had found the interruptions, on a scale with 1 meaning *extremely unpleasant* and 5 meaning *extremely pleasant*. They rated both the experience of being interrupted as caller and that of being interrupted as driver. The results are shown in Figure 4. Whereas ratings for drivers were varied (*M* = 2.94, *SD* = 0.97), the majority of subjects rated the interruptions for the caller as a 2 (*M* = 2.13, *SD* = 0.7), and hence were more unanimously displeased with them.

In an open post-experiment question, some participants suggested that other interruption modes (such as a sound signal only), interruption manners (with more forewarning), or interruption utterances might be more acceptable. This remains to be evaluated.
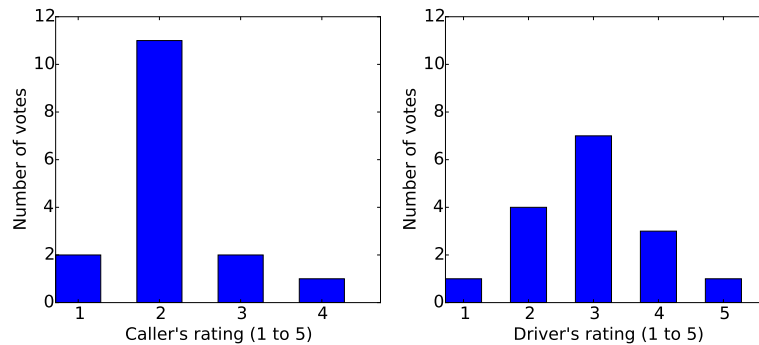

Fig. 4. Scores for pleasantness of interruptions, as rated by drivers and callers

It is also important to note that interruptions leave callers temporarily without any tasks to perform, whereas drivers still have their main task to concentrate on: This could also be a reason why callers get more frustrated by interruptions. Finally, pleasantness scores for drivers were correlated neither with number of successful trials nor with the frequency with which the driver had been interrupted (as opposed to the caller).

4   Discussion and Further Work

The results presented show that verbal interjections coming from a system do efficiently interrupt an ongoing conversation. They also show that, in a driving situation, doing so improves performance during difficult driving tasks. Further research needs to be done in order to better understand its influence on performance and to find ways in which it can be enhanced. It is essential to cast more light on users' emotional responses to these interruptions (as callers as well as drivers) and to find ways to minimize frustration and stress.

The feedback obtained through the questionnaires raises issues regarding both the content and the mode of interruptions. Among the suggestions, it is possible to identify two trends: Some participants recommended being more explicit as to the reasons behind the need for the interruption, whereas others suggested ideas which might appear (at least initially) precisely the opposite, such as shortening the utterances or using a sound signal instead of speech. Exploration of these different strategies and their effects on users will be a next step. Furthermore, it might be possible to combine these seemingly opposed suggestions, for example by producing shorter utterances which

still convey more precise information about the situation of the driver (*driver overtaking, please wait*) or sounds which stand for specific driving events.

Secondly, it is necessary to explore ways in which speakers can be helped, when resuming the conversation, to remember what was being said before they were interrupted. This kind of assistance might also be beneficial for driving performance, as some drivers in our experiment reported that they had to make a considerable effort during interrupted lane changes to keep the state of the conversation in their minds. It might here prove beneficial to monitor the speaker's production and, depending on the severity of the danger situation, decide whether to interrupt as soon as the alert signal becomes available or to wait for a specific moment in the dialogue, such as a transition-relevance place [12], in a similar way to [5].

Finally, some participants reported a desire for increased control over the system, for example by giving drivers the possibility to activate the interruptions themselves. This is related to [8], who found that granting users control over when a dialogue system resumes its speech after an interruption can improve user satisfaction without harming driving performance. It clearly remains a challenge to find ways in which this can be done without introducing too much additional cognitive effort for the driver.

5 Acknowledgments

References

1. Baumann, T., Schlangen, D.: The InproTK 2012 release. In: Proceedings of the NAACL-HLT Workshop on Future directions and needs in the Spoken Dialog Community: Tools and Data (SDCTD 2012). pp. 29–32. ACL (2012)
2. Drews, F., Pasupathi, M., Strayer, D.: Passenger and cell phone conversations in simulated driving. Journal of Experimental Psychology 14(4), 392–400 (2008)
3. Graham, R.: Use of auditory icons as emergency warnings: evaluation within a vehicle collision avoidance application. Ergonomics 42, 1233–1248 (1999)
4. International Organization for Standardization (ISO): Road vehicles – ergonomic aspects of transport information and control systems – simulated lane change test to assess in-vehicle secondary task demand. Standard (2010)
5. Iqbal, S.T., Horvitz, E., Ju, Y.C., Mathews, E.: Hang on a sec!: effects of proactive mediation of phone conversations while driving. In: CHI (2011)
6. Kennington, C., Kousidis, S., Baumann, T., Buschmeier, H., Kopp, S., Schlangen, D.: Better Driving and Recall When In-car Information Presentation Uses Situationally-Aware Incremental Speech Output Generation. In: AutomotiveUI 2014: Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. pp. 7:1–7:7 (2014)

7. Kennington, C., Kousidis, S., Schlangen, D.: Inprotks: A toolkit for incremental situated processing. In: Proceedings of the 15th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL). pp. 84–88. Association for Computational Linguistics, Philadelphia, PA, U.S.A. (June 2014), `http://www.aclweb.org/anthology/W14-4312`

8. Kousidis, S., Kennington, C., Baumann, T., Buschmeier, H., Kopp, S., Schlangen, D.: A Multimodal In-Car Dialogue System That Tracks The Driver's Attention. In: Proceedings of the 16th International Conference on Multimodal Interfaces. pp. 26–33 (2014)

9. Lison, P.: A hybrid approach to dialogue management based on probabilistic rules. Computer Speech & Language 34(1), 232 – 255 (2015)

10. Maciej, J., Nitsch, M., Vollrath, M.: Conversing while driving: The importance of visual information for conversation modulation. Traffic Psychology and Behaviour 14, 512–524 (2011)

11. McEvoy, S.P., Stevenson, M.R., McCartt, A.T., Woodward, M., Haworth, C., Palamara, P., Cercarelli, R.: Role of mobile phones in motor vehicle crashes resulting in hospital attendance: A case-crossover study. BMJ 331, 428 (2005)

12. Sacks, H., Schegloff, E.A., Jefferson, G.: A simplest systematics for the organization of turn-taking for conversation. Language 50(4), 696–735 (1974), `http://www.jstor.org/stable/412243`

13. Schröder, M., Trouvain, J.: The German text-to-speech synthesis system MARY: A tool for research, development and teaching. International Journal of Speech Technology 6, 365–377 (2003)

14. Strayer, D.L., Drews, F.A., Crouch, D.J.: A comparison of the cell phone driver and the drunk driver. Human Factors 48, 381–91 (2006)

15. Wienke, J., Wrede, S.: A Middleware for Collaborative Research in Experimental Robotics. In: IEEE/SICE International Symposium on System Integration (SII2011). pp. 1183–1190. IEEE (2011)