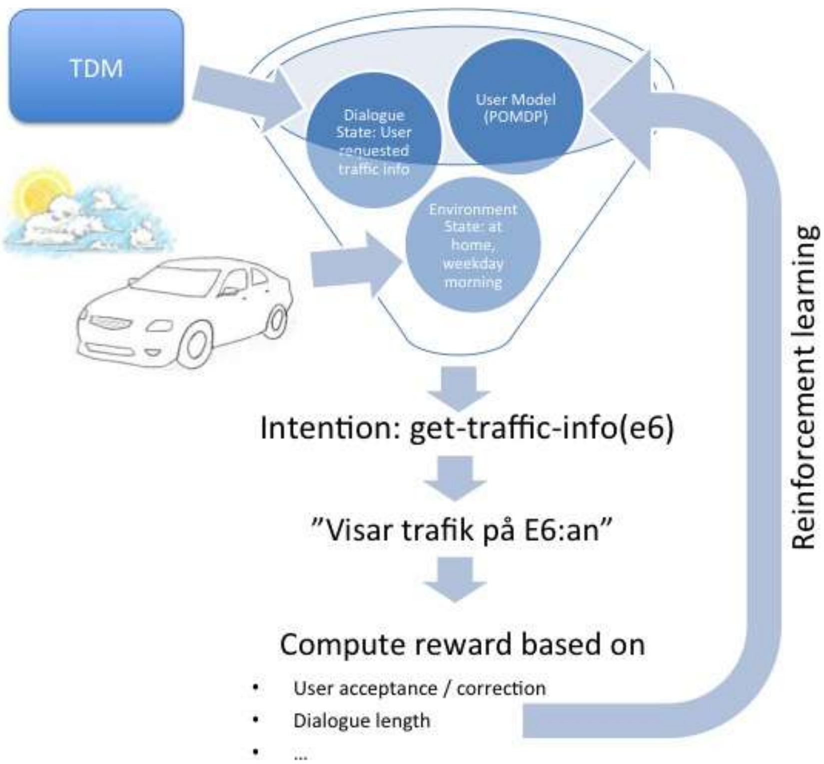
FORDONSSTRATEGISK
FORSKNING OCH INNOVATION

Safe Speech by Knowledge



Fredrik Kronlid
2013-09-24
Vehicle safety

FFI

Fordonsstrategisk Forskning och Innovation

Table of Contents

1 Executive Summary

The driving environment of today’s cars has become increasingly complex and raises the needs on the interaction design. Also, secondary tasks are more frequently carried out by the driver. The secondary tasks consist of smartphone use and use of other technologies, which adds to distraction, which causes an increased risk for accidents. To design interactions which helps the user to become less distracted and to create a driving environment where the driver can keep her hands on the wheel and her eyes on the road becomes an important task.

The objective of this project is to design, implement and test a multimodal, adaptive dialogue system in order to reduce the time spent on the interaction, which in turn reduces the distraction and increase the safety in the driving environment. The project also aims at building knowledge about how predictive models can be used in order to design spoken interaction with a high user experience ca be reached.

The project has been carried out in three steps: analysis, user study and implementation and has resulted in a working impleentation which reduces the dialogue based on the user’s driving behaviour. The system also reduces the number of repetitive questions from the system, which according to the user study of the project not only reduces the time of interaction, but also contributes to a positive user experience for the driver.

Also, the user study has resulted in a set of user-centered design principles for spiken adaptive interaction, which can be used for future development projects. Finally, a method of user testing optimised for rapid prototyping of in-vehicle spoken interaction has been created during the project.

2 Background

Driver distraction is a common cause of accidents, and is often caused by the driver interacting with technologies such as mobile phones, media players or navigation systems. A study, commonly referred to as the”100 car study” (Neale et al., 2005) revealed that secondary task distraction is the largest cause of driver inattention, and that the handling of wireless devices is the most common secondary task. The U.S. National Highway Traffic Safety Administration (NHTSA) expect that 16% of deadly accidents and 20% of accidents leading to serious injury in 2009 were connected to distraction.

In the vehicle industry, we talk about Human Machine Interfaces (HMI's) or Driver-Vehicle interfaces, some of which include speech. In academia, many researchers instead talk about *Multimodal Dialogue Systems* (or MDS's for short). A multimodal dialogue system enable spoken communication between humans and machines, but complements the spoken modality with traditional human-machine interaction modalities such as visual output (screen, head-up display) and haptic input (scroll wheels, buttons, etc.).

TDM is Talkamatic's product for dialogue management and is the central component in a spoken dialogue system. It features Free Dialogue, Multi-Modality and Rapid Development. TDM is to our knowledge the most capable dialogue manager available on the market. The current focus of development is low distraction and faster development.

State-of-the-art infotainment systems typically do not include user models at all, which means that they treat all users alike. Apple's Siri, which is not an infotainment system per se but which can be expected to be used by drivers for playing music etc., has some primitive user modeling capabilities in that it can make use of personal information explicitly provided by the user (home address, etc.). In the research field of used modeling, this is sometimes referred to as a "static user model". For example, if the user says "I want to take a taxi", Siri will assume that the intended destination is the user's home address. In contrast, this project will explore the use of state-of-the-art technologies for *adaptive* user models, which learn users preferences and behaviour patterns from observing their interactions with the infotainment system and the context in which these interactions take place.
.

3 Objective

The primary purpose of tthis project has been to build a multimodal dialogue system with a dynamic user model, where the user model affects the dialogue in such a way that the length of the interaction needed to perform a certain task is minimised: the user does not need to answer repetitive questions where the system itself can find out the answer from the user model. By minimising the length of the nteraction also the time during which the user is distracted is minimised.

A secondary purpose has been to win knowledge about predictive systems and spoken interaction with such systems. How should they be designed? What does the user expect from them? Do they create problems: does the user have a feeling of being under surveillance when the system knows of details in her or his life that has not been explicitly shared with the car?

4 Project Realisation

The project realisation had three components: Analysis, User Study and Implementation. The goal of the analysis component was to create an understanding of how and in what contexts the technology gives the best results: Which types of patterns are the easiest ones to learn, and what types of applications are linked to these kinds of patterns? The purpose of the user study was to answer questions in the UX area: What does the user expect from predictive systems, how should the systems be designed, what usability problems and usability gains can be expected etc. The results of these two components were fed into the implementation component, which had the purpose of building the actual system with user model, ability to learn, support for reasoning and dialogue about uncertain knowledge, and finally also applications.

4.1 Analysis

The analysis work consisted of a literature study of adaptive systems, natural-language interaction in the in-vehicle setting as well as driver/user experience. This work directly supported the design choices, and was also a foundation for how the collection of empirical data was to be carried out (see section **Error! Reference source not found.**~~4.2 below~~).

The analysis work also included analysis of driving patterns as well as consumer behaviour with regard to what technological platforms and services are used in different segments – for instance smartphone use in different age groups. This kind of data influences the design of the ecosystem that the final service will be a part of.

Two series of workshops were completed on location at VMCC. The first workshop saw participants from Talkamatic and VMCC and focused on how to implement the technology for reducing the users' "visual struggle" with handling the technology. The workshop resulted in a number of suggestions of use cases in line with this focus. It was also stressed that the user should not feel as being under surveillance. The week after, there was a second workshop with participants from Högskolan i Halmstad and VMCC, focusing on user experience and creation of scenarios. The main scenario contained two primary personas, representing archetypical users of the service.

During these workshops the goal of simply reducing the dialogue in number of turns was elaborated. It was questioned whether just reducing the number of turns is always an improvement. The result of the discussion was that dialogue on the right level is better than a short dialogue – The conclusion was that the challenge of interacting with the system should be somewhere between (not including) boredom and stress. This means that repeated entry of the same information should be avoided, in preference of the user simply accepting system suggestions. It also means that the system should have really good reasons to actually remove utterances – because removing utterances in an inexplicable way could cause stress.

4.2 User Study

Högskolan i Halmstad performed a qualitative user study, where a prototype of the services that had been sketched up in the scenarios of the second workshop were tested. Seven respondents participated in the study where they drove a car prepared with a prototype of an adaptive navigation service. The scenario consisted of a progression of three parts, where the systems learned from the user driving patterns and adapted the dialogue to these patterns.

The study was conducted as a Wizard of Oz study, where the user was unaware that a human (the "wizard") was conducting the dialogue. The "wizard" who perform the dialogue turns of the system follows a predefined interaction protocol based on a theory of adaptive interaction, best practice in spoken interaction and also the scenarios sketched in the 2nd workshop. Based on the interaction protocol and the user input, the wizard sends tailor-made text strings to the car, which are read out to the user via a state-of-the-art text-to-speech module. Between the scenarios, the interaction protocol were adapted in order to simulate an adaption to the users' driving behaviour and preferences.

A semi-structured interview was performed with each user after each scenario part about her experience. The purpose was to make the user discuss different aspects of usability, user experience and perceived stress, safety, cognitive load and emotional response from the interaction with the system.

4.3 Implementation

4.3.1 Dialogue functionality

The implementation of the dialogue functionality was conducted as an agile software development project, and was completely run by Talkamatic. Högskolan i Halmstad and VCC regularly contributed with feedback on the system, and shared insights and experiences primarily collected during the user studies.

As every agile project, the implementation was characterised by fast releases, short iterations, focus on user needs (via Högskolan i Halmstad) and test-driven development.

During the implementation project, five releases were made:

- Release 1 – Basic functionality. User model learned from interaction. Mocked traffic information application.

- Release 2 – Elaborated application with search and navigation. Better quality of data. Feedback from the system when integrating uncertain data from the user model.
- Release 3 – Tools for simulating user models, including graphical user interface. Better quality of feedback on uncertain data from the user model.
- Release 4 – Higher quality data which supports use cases from personas. Possibility to visualise user models. Final model for system feedback on user model data.
- Release 5 – System running on a *BD-SL-i.MX6* card.

4.3.2 Rapid Prototyping platform

Parallel to the development of the dialogue functionality, VCC conducted development activities on a number of issues: A general solution architecture for the type of voice interaction that is the subject of this project was designed, the rapid prototyping platform of VCC was adapted for use with the Talkamatic platform on the IMX-6 board, and a prototype application for use with the talkamatic system was conceived.

5 Results

5.1 Demonstrator/Implementation

There is a demonstrator showing interaction with a learning dialogue system, which handles (for instance) the example dialogue from the application:

```
User: Traffic information
Car: Ok. What road?
User: E6.
Car: Showing traffic on the E6
```

This is repeated on a number of occasions, and one morning...

```
User: Traffic information
Car: Showing traffic on the E6
```

The dialogue system learns from the user interaction with the system and from the user's actual driving behaviour. Events are classified after time, weekday and position.

The system reduces the need for repetitive and information-scarce utterances from the user. Even before the system has learned as much as needed for not asking the user for information, the dialogue becomes better. When the system has started identifying a pattern, it will start to suggest the most probable alternatives. To start, the most probable answers are presented to the user as the top items in a list. The alternatives are also marked in a different color to make them more visible to the user:

```
User: Traffic information
Car: Ok. What road?
Displaying:     [E6]
                [E45]
                [E20]
                [155]
                [40]
User: E6.
Car: Showing traffic on the E6
```

After some further use, the system has identified a pattern which is prominent anough for the system to make a suggestion:

```
User: Traffic information
Car: E6, is that right?
User: Yes.
Car: Showing traffic on the E6
```

After further support for the hypothesis of the system, it will start to inform the user that an assumption has been made. If the user is satisfied with the assumption, the user does not need to do anything, but can confirm it.

```
User: Traffic information
Car: I assume E6.
User: [silence]
Car: Showing traffic on the E6
```

The user always has the possibility to reject the system suggestion.

```
User: Traffic information
Car: I assume E6.
User: No, E45.
Car: Showing traffic on the E45
```

If the user rejects the system suggestion without giving another answer, the system will show a menu where the most probable choices are the topmost ones, and marked in a distinct colour.

```
User: Traffic information
Car: I assume E6.
```

```
User: No.
Car: What road?
```

```
Displaying:

    [E6]
    [E45]
    [E20]
    [155]
    [40]
```

When the system is entirely certain about its hypothesis, the system will provide the user with the desired information without asking the user for parameters:

```
User: Traffic information
Car: Showing traffic on the E6
```

These behaviours altogether should reduce the cognitive load and the distraction for the user in most cases.

5.2 Building UX knowledge about predictive dialogue

The primary task of a driver is to conduct their vehicle in a safe way in traffic (primary interaction). In today's driver environment there is also secondary interaction with for instance infotainment systems, climate control and navigation. In Human Machine Interaction for driver environment, a great effort is made to ensure that the secondary interaction does not affect the traffic safety in a negative way. By the study of use and user situation, it is easy to understand that there is also tertiary interaction in the driver environment, for instance in the form of smartphone use for navigation or communication (text messages or voice calls). These three levels of interactions are constantly increasing the demands on the user's cognitive resources, which can affect the driver's ability to perform the primary task in a negative way. Aside from the safety effects, this also affects the user experience negatively. In this project we have studied aspects of user experience (UX) in the use of predictive dialogues in order to a) not to charge the visual-manual interaction modalities in order to free eyes and hands for the primary interaction and b) to shorten the secondary and tertiary interaction in order for the driver to free cognitive resources for the primary task.

By prototyping and testing dialogue systems solutions (see section **Error! Reference source not found.**~~4.2~~) with adaptive and predictive qualities on a representative set of end-users in an authentic driving environment, we have built knowledge on experience-based qualities regarding spoken interaction in a driver environment. The analysis of the study verified positive features and indicated problems mainly regarding six aspects of interaction: Antropomorfism, trust, predictability, understanding, integrity and control. In the debriefing interviews of the study, several solutions to the shortcomings of the system were discussed. These discussions resulted in a model for UX-oriented design of spoken adaptive interaction in an in-vehicle setting. (A manuscript reporting this is currently in preparation).

Further results under this heading is the development of the method which addresses the problem of quickly and for a low cost test functionality and user experience in an authentical user situation. The method is based on the Wizard-of-Oz methodology and reinforced with a set of environments and tools specific for the task of prototyping and testing adaptive in-vehicle dialogue interaction.

5.3 TDM on embedded platform

As a part of the project, TDM has been ported to an embedded platform, a card from Boundary Devices (*BD-SL-i.MX6)* which complies to Freescale's Light Sabre specification. The card is, in terms of processing power, memory, connectivity etc. representative for the plattforms used in today's infotainment systems. This means that we have now verified that TDM can run on current state-of-the-art infotainment platforms without problems.

5.4 Create practical knowledge on introduction of new phone based applications in a way that does not compromise safety..

Within vision 2020 we know that more than 90 % of the accidents are caused by driver mistakes. Driver Distraction is a contributing cause of these accidents, so it is of utter importance to continuously strive to improve the interaction with new functionalities in the car. We need to understand how we create customer value with the aid of new technologies that comes with connected functions and at the same time tailor the experience in order to reduce the distraction from using these new systems.

Within the project VCC has, step by step, built knowledge about how customer value can be created by the use of predictive functionality in speech systems, and how distraction can be reduced for recurring use-cases and about how functionality can be realised in an in-vehicle setting.

The work has consisted of three distinct phases:

1. analysis of customer behaviour
    a. Data-mining of customer behaviour from NHTS 2009 for quantitative analysis and clustring of driving patterns and defining relevant input data for event-logging.
2. Innovation of use cases and functionality testing
    a. Workshops with focus of defining functionalities to implement.

   b. Evaluation of performance of implemented functionality in the Talkamatic platform.
3. Development of prototype platform and adaption to in-vehicle implementation
   a. Definition of solution architecture

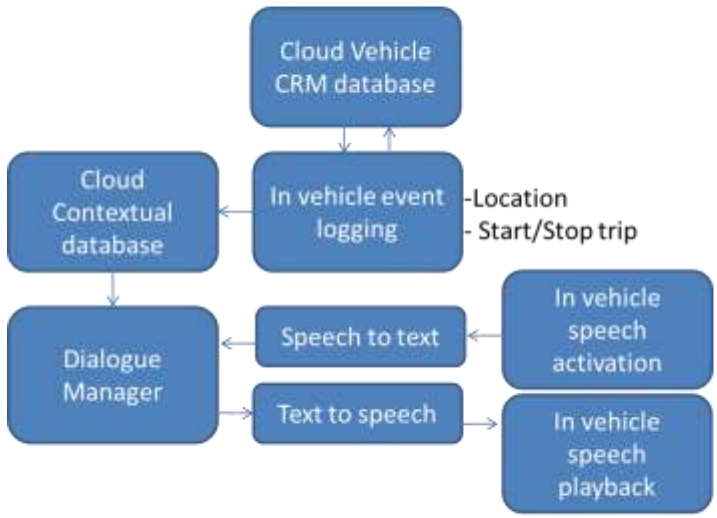


Figure 1 Lösningsarkitektur

   b. Development of IMX-6 prototype platform for integration with Talkamatic solution
   c. Definition of application of commonly implemented prototype in other development projects at VCC.

5.5 Delivery to the FFI goals

5.5.1 Implement a learning spoken interaction system

The main goal of this project is to design and implement support for learning in TDM. This goal delivers both to the target of having a competitive Swedish vehicle industry and the target of reducing the number of persons killed or injured in traffic.

5.5.2 Build knowledge about UX in the area of predictive dialogue.

This goal delivers both to the target of having a competitive Swedish vehicle industry and the target of reducing the number of persons killed or injured in traffic. The project has contributed to two knowledge-development blocks within the frame of interaction design and user experience (UX). The work has resulted in theory development regarding principles for design of adaptive in-vehicle dialogue systems (a manuscript reporting this is currently in preparation). The project has also contributed to increased knowledge about rapid prototyping in order to efficiently study UX qualities for end users by a specific Wizard of Oz method and environment (manuscript in preparation). This knowledge contributes to concrete improvement of production in the participating companies. The project has also contributed to the goal of strengthened co-operation between the vehicle industry and the academia through the work on UX qualities.

5.5.3 TDM on embedded platform

Delivers to the target of having a competitive Swedish vehicle industry.

5.6 Create practical knowledge on introduction of new phone based applications in a way that does not compromise safety..

This goal delivers both to the target of having a competitive Swedish vehicle industry and the target of reducing the number of persons killed or injured in traffic.

5.6.1 Patent application

A Swedish provisional patent application has been filed by Talkamatic. The application is requesting a patent for the technical solution for integrating into a dialoge knowledge of different strengths from a user model.

6 Dissemination and publications

6.1 Knowledge and results dissemination

Talkamatic has established contact with a group at Volvo Cars who is working on route prediction. It is a fair assumption that the work that has been conducted in that group and SSK can mutually enrich one another. Talkamatic is also working with Volvo Cars to find a continuation for the work on predictive dialogue. It is also possible that the results from the projects are integrated into the prototype which is built in the SIMSI project (FFI).

6.2 Publikationer

– Holdaj Pettersson, K., Thunberg, M. 2013. Adaptiva talbaserade system i fordon: designförslag för att främja user experience. Kandidatuppsats, Högskolan i Halmstad.

– Wärneståål, P., Kronlid, F. 2013. Towards a User Experience Design Framework for Adaptive Spoken Dialogue in Automotive Contexts. *Manuscript in preparation*.
– Kronlid, F. Wärneståål, P. 2013 Integrating Uncertain Information in an Information State Update Dialogue Manager. *In preparation.*

7 Conclusions and future research

We have built a technical solution, which when exposed to repeated patterns of use, simplifies and shortens the dialogue, so that the time that the user is occupied with dialogue is shortened. It is however not established that this actually reduces the distraction rate of the user. This would be interesting to research in future projects. Since a simulator environment, with advanced equipment for measuring cognitive load and distraction, has been built at the Centre for Language Technology at the University of Gothenburg in connection to the SIMSI project, there are good possibilities to perform such studies.

The technical solution which has been built in the project models a user's answers to the questions of the system, and not the actual preferences of the user expressed by these answers. This may be considered a superficial difference, but can be of significance in some cases – where it may lead to the system asking questions which it could have answered by itself, had it modelled the user in a more elaborate way. It is an interesting research track to further develop the user model so that it encodes the user preferences in order to research if the user distraction can be further reduced.

From a design perspective the lack of design principles and standards for adaptive in-vehicle dialogue interaction is obvious, and this work is important both from a safety perspective and from a user experience perspective. This work has, based on findings from an authentic user situation, started to scratch this surface. Future research is needed in order to fill in the details in order to create a more complete design framework.

8 Participating parties and contact person

Talkamatic
FREE DIALOGUE

Talkamatic AB, Fredrik Kronlid, fredrik@talkamatic.se, +46703602190

Volvo Cars, Jonas Söderqvist, jsoderqv@volvocars.com, Mobile: 1 805 603 0997

Högskolan i Halmstad, Pontus Wärnestål, pontus.warnestal@hh.se, +4635167696