Visual distraction test setup for an multimodal in-vehicle dialogue system

Staffan Larsson
Sebastian Berlin
University of Gothenburg
Box 200
SE-405 30 Gothenburg
Sweden
sl@ling.gu.se
sebastian.berlin@gu.se

Anders Eliasson
Mecel AB
Box 140 44
SE-400 20 Gothenburg
Sweden
anders.eliasson@mecel.se

Fredrik Kronlid
Talkamatic AB
Första långgatan 18
SE-413 28 Gothenburg
Sweden
fredrik@talkamatic.se

Abstract

The goal of the SIMSI (Safe In-vehicle Multimodal Speech Interaction) project is threefold. Firstly, to integrate a dialogue system for menu-based dialogue with a GUI-driven in-vehicle infotainment system. Secondly, to further improve the integrated system with respect to driver distraction, thus making the system safer to use while driving. Thirdly, to verify that the resulting system decreases visual distraction and cognitive load during interaction. This demo paper describes the test environment designed to enable evaluation of the system, and the planned visual distraction tests.

1 Background

1.1 Driver distraction and safety

Driver distraction is one common cause of accidents, and is often caused by the driver interacting with technologies such as mobile phones, media players or navigation systems. The so-called 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 goal of SIMSI is to design systems which enable safe interaction with technologies in vehicles, by reducing the cognitive load imposed by the interaction and minimizing head-down time.

1.2 The SIMSI Dialogue System

Based on Larsson (2002) and later work, Talkamatic AB has developed the Talkamatic Dialogue Manager (TDM) with the goal of being the most competent and usable dialogue manager on the market, both from the perspective of the user and from the perspective of the HMI developer. TDM provides a general interaction model founded in human interaction patterns, resulting in a high degree of naturalness and flexibility which increases usability. Also, TDM reduces complexity for developers and users, helping them to reach their goals faster and at a lower cost.

TDM supports multi-modal interaction where voice output and input (VUI) is combined with a traditional menu-based GUI with graphical output and haptic input. In cases where a GUI already exists, TDM can replace the GUI-internal interaction engine, thus adding speech while keeping the original GUI design. All system output is realized both verbally and graphically, and the user can switch freely between uni-modal (voice or screen/keys) and multi-modal interaction.

To facilitate the browsing of lists (a well known interaction problem for dialogue systems), Talkamatic has developed its Voice-Cursor technology[1] (Larsson et al., 2011). It allows a user to browse a list in a multi-modal dialogue system without looking at a screen and without being exposed to large chunks of readout information. A crucial property of TDM's integrated multimodality is the fact that it enables the driver of a vehicle to carry out all interactions without ever looking at the screen, either by speaking to the system, by providing haptic input, or by combining the two. We are not aware of any current multimodal in-vehicle dialogue system offering this capability.

While TDM offers full menu-based multimodal interaction, the GUI itself is fairly basic and does not match the state of the art when it comes to graphical design. By contrast, Mecel Populus is an commercial-grade HMI (Human Machine Interface) with professionally designed visual output. We have previously produced an integration of the TDM and Mecel Populus platforms(Larsson et al., 2013) to establish a commercial-grade HMI for experiments and demonstrations.

[1]Patent Pending

Proceedings of the 17th Workshop on the Semantics and Pragmatics of Dialogue, December 16-18, 2013, Amsterdam, The Netherlands.

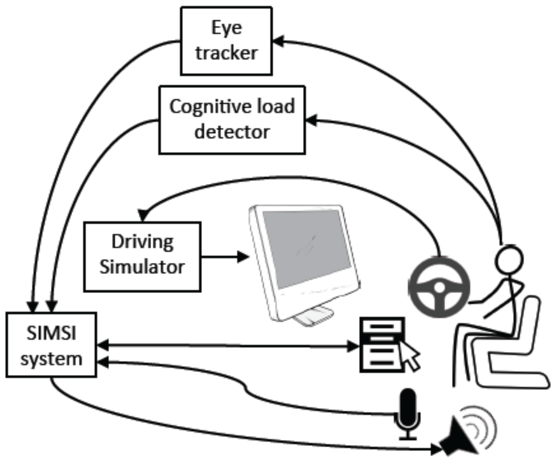

Figure 1: SIMSI test environment overview

2 Test environment

One goal of SIMSI is to conduct ecologically valid test of the applications, and to let the results of these tests feed back into the development of the system. Basically, we want to find the best interaction solutions and to verify these experimentally, especially in cases where it is not intuitively clear what is best. This involves implementing variants of a behaviour, testing them on naive users, collecting data from these interactions, and establishing statistically significant results based on the collected data.

The test environment consists of two parts, apart from the dialogue system: a driving simulator (SCANeR from Oktal) and an eye tracker (Smart Eye Pro from Smarteye). In later tests we will also include instruments for measuring cognitive load.

3 Visual distraction tests

The main point of the visual distraction tests is to investigate how the "eyes-on-road" time during interaction varies between different modality conditions. The eyetracker equipment will be used for capturing where the driver is looking. In addition, driving behaviour (including lane deviation) and dialogue state (including task success) is continously logged.

The following three conditions will be tested:

- GUI only (haptic only in, graphics only out)
- Multimodal with voice cursor (haptics and speech in, graphichs and speech out)
- GUI with voice cursor (haptics only in, graphichs and speech out)

Resources permitting, we may also test two additional conditions:

- multimodal without voice cursor (haptics and speech in, graphichs and speech out)
- speech-only with voice cursor (haptics and speech in, speech only out)

For each condition, we will be using two difficulty levels: easy and difficult. For both levels, the task is to drive along a softly curving road while keeping distance to one car in front of you and one car behind you. In the easy condition, the other cars have a constant speed. In the difficult condition, the other cars are speeding up and braking erratically, and the car behind you may indicate (by honking its horn) that you're going too slow.

This way of testing, which we informally refer to as the "annoying cars" setup, differs from existing experimental setups such as the ConTRe task (Engonopoulos et al., 2008). In the latter, the driver tries to match two vertical lines representing the vehicles position and the target (reference) position. Our setup has the advantage of being more realistic, although we acknowledge that it is still far from driving in real traffic. (On the negative side, our setup does require a full driving simulator environment, which the ConTRe task does not). Initial tests will be carried out to verify the adequacy if the "annoying cars" setup for our purpouses.

The application used in the tests has very basic phone functionality: browsing a list of contacts, and calling people up. At regular intervals, the driver receives a spoken instruction (with a voice different from the dialogue system), e.g. "You just remembered you need to call up Ashley on her mobile number.". The user should then carry out this instruction as efficiently and completely as possible.

We hypothesise that in the GUI only condition, there will be less eyes-on-road time than in the other two conditions, since the driver does not have to look at the screen in order to complete the task. Apart from testing this hypothesis, we are generally interested in which condition(s) gives the best results with respect to eyes-on-road time, task success, task completion time and usability (rated subjectively using a questionnaire).

We will demonstrate the SIMSI system, the three test conditions, and parts of the test environment.

References

Nikolaos Engonopoulos, Asad Sayeed, and V Demberg. 2008. Language and cognitive load in a dual task environment. In *CogSci 2013*, pages 2249–2254.

Staffan Larsson, Alexander Berman, and Jessica Villing. 2011. Adding a speech cursor to a multimodal dialogue system. In *INTERSPEECH 2011, 12th Annual Conference of the International Speech Communication Association, Florence, Italy, 2011*, pages 3319–3320.

Staffan Larsson, Sebastian Berlin, Anders Eliasson, and Fredrik Kronlid. 2013. Integration and test environment for an in-vehicle dialogue system in the SIMSI project. In *Proceedings of the 14th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 2002–2004.

Staffan Larsson. 2002. *Issue-based Dialogue Management*. Ph.D. thesis, Göteborg University.

Vicki L. Neale, Thomas A. Dingus, Sheila G. Klauer, Jeremy Sudweeks, and Michael Goodman. 2005. An overview of the 100-car naturalistic study and findings. Technical report.