# Using Learned Predictions of User Utterances to Decrease Distraction

**Fredrik Kronlid**
Talkamatic AB
Första Långgatan 18
413 28 Göteborg
Sweden
fredrik@talkamatic.se

**Staffan Larsson**
Talkamatic AB
Första Långgatan 18
413 28 Göteborg
Sweden
sl@talkamatic.se

**Pontus Wärnestål**
Halmstad University
Box 823
301 18 Halmstad
Sweden
pontus.warnestal@hh.se

## Abstract

Driver distraction is one of the most common causes of accidents. By having a dialogue manager request predicted user answers from a user model instead of asking the user, we can reduce the number of utterances in the dialogue and thereby reduce the time that the user is distracted.

## 1 Background

### 1.1 Driver Distraction

As interaction complexity in the car increases due to more advanced infotainment systems, and peripheral technologies in the form of smartphones and tablets, drivers are often executing several tasks in parallel to the primary task of driving. The increased functionality of car information and entertainment systems has resulted in large hierarchical information architectures that prolong interaction time and that may thereby negatively affect safety as well as user experience (Dagmar & Albrecht, 2009). According to the 100-Car Study (Neale et al., 2002), non-primary task distraction is the largest cause of driver inattention.

The goal of the work described in this paper is to design an in-vehicle information system that enables shorter and more efficient interaction in the form of natural language dialogues. The basic assumption is that using apps and services in an in-vehicle context inherently leads to distraction, and that reducing interaction time will reduce driver distraction.

### 1.2 TDM

Based on Larsson (2002) and later work, Talkamatic AB has developed the Talkamatic Dialogue Manager (TDM). TDM provides a general interaction model based on patterns found in human-human dialogue, resulting in a high degree of naturalness and flexibility which increases usability. TDM offers integrated multi-modality which allows user to freely switch between modalities. The model is domain-independent which means that dialogue behaviour can be altered without touching application properties and vice versa.

### 1.3 Grounding in TDM

Grounding is, roughly, the process of making sure that dialogue participants agree on what has been said so far and what it meant. TDM has an extensive model of grounding (Larsson, 2002). It operates on different levels: *Perception*, *Semantic Understanding*, *Pragmatic Understanding* and *Acceptance*. System feedback (positive, negative and in some cases interrogative) can be generated on each level. Examples: “I didn’t hear” – negative perception; “Madonna, is that right?” – interrogative semantic understanding; “OK” – positive acceptance.

## 2 Learning and Classification

Many dialogue applications require the user to answer a number of questions. To make dialogue shorter, we have extended TDM so that it tries to predict user answers on the basis of a user model learned from observations of user behaviour. As an illustration, we use a road information application which tries to predict the user’s destination and thereby eliminate the need to ask the user about this.

### 2.1 Selection of learning Method

Initially, more complex learning methods (MDP, POMDP) were explored, but the KNN (K-Nearest Neighbours) were considered the best method. An important advantage is that KNN can learn from a relatively small set of observations. This is in contrast to the MDP and POMDP methods, which require large amounts of data to generate useful behaviour. A potential drawback of KNN is that this model cannot model sequences of user behaviours.

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

On the basis of user studies, it was decided that the most important user model parameters was position, day of the week and hour of the day. The training data were simulated and correspond to the behaviour of an archetypical persona provided by the user partner in the project.

The learning part of the system listens for a number of events, such as “start-car”, “stop-car” etc.. From these events and information about current position, the time of the day and the day of the week, the system creates new data instances. The system thus learns how the user’s destination varies depending on these parameters. When the dialogue manager requests a prediction of the destination, the KNN algorithm tries to find the K data points closest to the present data point, and the top alternatives are returned to the dialogue manager together with confidence scores indicating the reliability of the predictions.

## 3 Integration of Classifications into TDM

### 3.1 Grounding uncertain information

We treat the information emanating from the user model as uncertain information about a (predicted) user utterance. Hence, the same mechanisms used for grounding utterances have been adapted for integrating user model data.

### 3.2 Integrating Classifier Output

TDM is based on the Information State Update (ISU) approach to dialogue management. The rule for integrating the user model data is a standard ISU rule, consisting of preconditions and effects on the information state. The information state in TDM is based on that of the system described in Larsson (2002).

If the user model data is sufficiently reliable to be trusted, the ISU rule described informally below triggers:

**Preconditions** If the user is the latest speaker and if there is a propositional answer from the user model resolving a question in the current plan, and if the confidence score reported from the user model is above a certain level, the rule should be applied.

**Adaptation Effects** Applying the rule means that we should accept the propositional answer (include it into the shared commitments), and – depending on the confidence score – give feedback to the user by enqueuing an appropriate feedback move on the agenda. We isolate three different cases when it comes to the feedback:

- For highly probable answers, we embed the feedback move into the next system utterance, e.g. “Which route do you want to take to work?”. The user can always reject the predictionby requesting another destination.
- For relatively certain answers, the feedback move (positive understanding) can be realised as “I assume you’re going to work”. If the user says “no”, the answer is rejected, but silence is interpreted as acceptance.
- For uncertain answers the feedback would be “To work, is that correct?” (interrogative understanding). In this case, the user needs to explicitly accept the proposed answer. Otherwise, the user is prompted for an answer.

### 3.3 GUI Behaviour

If the ISU rule above does not apply because of too low confidence scores, user model information is still used in the GUI. When a Wh-question is raised by the system, the GUI always presents a list of possible alternatives. High-confidence alternatives are highlighted and sorted before the other alternatives in the list.

## 4 Conclusions and further work

We have designed and implemented a mechanism which, when exposed to repeated patterns of use, simplifies and shortens the dialogue. It remains for future work to establish that this actually reduces the distraction rate of drivers. We also want to test the performance of the learning mechanism by training it on real observations of user behaviour (as opposed to simulated data).

The current mechanism only predicts answers to individual system questions, which may result in suboptimal behaviour in cases where there are dependencies between the questions pertaining to some task. An interesting area for future work is to instead predict *sequences* of answers; however, this would require a more powerful learning and classification mechanisms.

## Acknowledgements

This work was carried out within the FFI project “Safe Speech by Knowledge” (2012-00941), funded by VINNOVA, Volvo Car Corporation and Talkamatic.

## References


Dagmar, K. and S. Albrecht *Design space for driver-based automotive user interfaces*. In Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications. 2009, ACM: Essen, Germany. 2009.

Staffan Larsson. *Issue-Based Dialogue Management*. PhD thesis, University of Gothenburg, 2002.

Staffan Larsson, Alexander Berman, and Jessica Villing. *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.

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

Sharon L. Oviatt, Rachel Coulston, and Rebecca Lunsford. *When do we interact multi- modally?: cognitive load and multimodal commu- nication patterns*. In ICMI , pages 129-136. 2004.