# EVALUATING THE PROJECT54 SPEECH USER INTERFACE

Laslo Turner [1]) Andrew L. Kun [2])


### *Abstract*

*The Project54 system integrates multiple electronic devices in police cruisers into a single system with a speech user interface (SUI). The system has been deployed in over 240 cruisers in the state of New Hampshire. We collected over 49,000 samples of speech commands, along with the corresponding SUI responses, from 27 officers using the Project54 system during their everyday work. We found that the system recognized officer utterances around 85% of the time. About one third of recognition errors were due to speech recognizer errors and about two thirds were due to user errors. Three types of user errors were identified: issuing a command that is not valid in any context of the SUI (54% of user errors), issuing a command that is valid in some context but not in the current one (34%) and operating the push-to-talk button incorrectly (12%).*


## 1. Introduction

In his article "The computer for the 21$^{st}$ century" Mark Weiser describes a world where ubiquitous computers blend into the background [5]. The Project54 system, designed at the University of New Hampshire makes a small part of our world look more like the world in Weiser's article: the inside of a police cruiser. The Project54 system [4] allows police officers to control electronic devices in the cruiser, as well as access records databases, through one standard user interface. As of December, 2004, over 240 cruisers in the state of New Hampshire are equipped with the Project54 system. In these cruisers the Project54 SUI is in everyday use. The user interface allows the officer to have control over all the electronic devices either through a touch screen, a keyboard or by voice commands. Controlling the system by voice allows the officer to keep his/her eyes on the road and hands on the wheel while driving. The goal of the research presented here was to quantitatively evaluate the performance of the Project54 system's speech user interface (SUI).

## 2. Background

In-car speech user interfaces have been developed by a number of groups. Hunt [3] describes work on in-car speech recognition at Dragon Systems UK R&D. Kun et al [4] describe a system that integrates multiple in-car de-vices and lets the user interact with them using a SUI. Using a SUI to interact with in-car devices promises to make such interactions safe, however poor SUI design can result in unsafe interactions which in turn can lead to accidents [1]. Developing SUIs for in-car use


[1] Empirix, 20 Crosby Drive, Bedford, MA 01730; email: lturner@engineer.com

[2] Electrical and Computer Engineering Department, University of New Hampshire, Kingsbury Hall, Durham, NH 03824; email: andrew.kun@unh.edu

requires in-car speech corpora. While multiple in-car speech corpora exist [2] we are not aware of any corpora recorded in police cruisers during everyday operation.

## 3. The Project54 Speech User Interface

The Project54 system user interface is shown in Fig. 1. The touch screen can be used as a second input and output modality in case the speech interface is not available or when speech interaction is not practical or desirable. The SUI uses a directional microphone on the visor, as well as a push-to-talk button on the steering wheel.

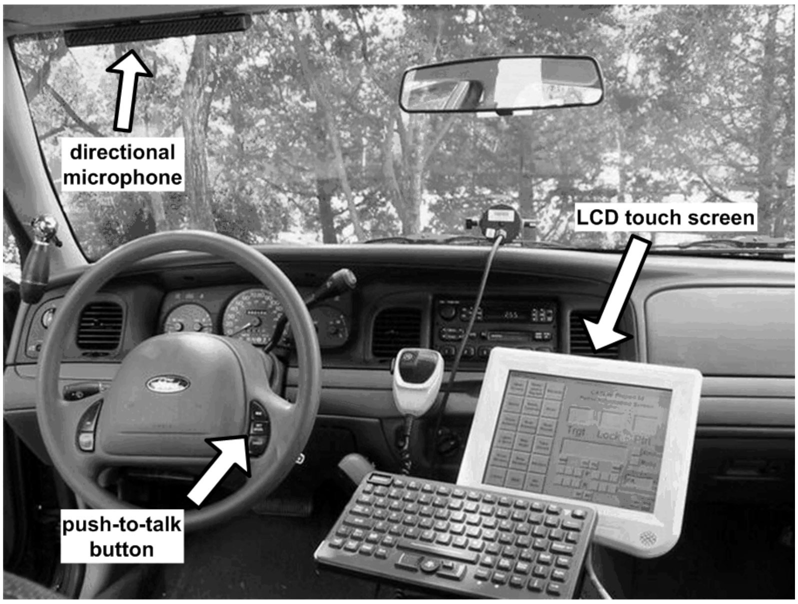


**Fig. 1. Project54 system user interface in a cruiser**

In our system the officer utters a phrase and the SUI reacts to this utterance. The SUI may execute a command, fill in a data field or initiate data retrieval. The SUI does not initiate interaction. The SUI uses a commercial SR engine (currently the Microsoft SAPI 5 recognizer) and a commercial text-to-speech (TTS) engine. Utterances that the SUI does not recognize are either misrecognized or unrecognized. Misrecognized commands lead to the execution of the wrong command which usually has to be undone. Unrecognized commands are rejected by the SUI.

SR performance is improved by using a push-to-talk (PTT) button. Recognition is started when the PTT button is pressed and it is stopped when the button is released. SR performance is also improved through the introduction of grammars. Grammars describe rules that the SR engine "believes" govern utterance generation. Grammars are switched by the SUI in response to officer utterances. For example, the officer may be filling in a form in which one of the fields describes the gender of a person. The officer may say 'gender' to signal that he/she will tell the SUI the gender of the person next. At this point the SUI loads a simple grammar which only lists the two genders and one or two other utterances (such as 'cancel').

When the user utters a command the SUI responds with the recognized utterance. This allows the officer to check if the recognition was correct. While this is a very useful feature it also slows down the interaction. In many cases the officers do not need, or have the patience, to listen to the SUI repeat the entire command they just issued. The officers can cut off any SUI utterance by pressing the PTT button. This makes sense when even a fragment of the SUI's utterance makes it clear that the officer's utterance was recognized correctly. However, cutting off the SUI is risky when the utterances are novel, e.g. when the officer is entering a plate number. In such cases there is no way to know if the utterance was recognized correctly without listening to the entire SUI utterance (or looking at the GUI).

## 4. Data Collection

We collected 49,177 utterances from twenty-seven police officers, between June 2003 and August 2004. All officers were male. Officers were informed that their utterances will only be recorded when the push-to-talk button is pressed. Nothing was recorded when the PTT button was not pressed. The utterances were saved as individual sound files. The recognition results corresponding to the utterances were also saved.

## 5. Results

In this paper we discuss results that addressed two of the questions that we asked in our study.

### 5.1 What is the Recognition Rate of the SUI?

The average recognition rate of the SUI, over the entire corpus, is 85.34%. For one of the officers it dips to 73.97% and for three officers it is over 90%. Another statistic that we wanted to see was how the SUI performance changes over time. We expected that with the passage of time officers will be more familiar with the SUI and consequently will make fewer errors. We expected this to result in better SUI recognition rate. We evaluated SUI recognition rate over several months for several of the officers and we did not find an improvement in SUI recognition rate. This result raised a new question to be answered in the future: why has the users' performance not improved with the usage of the SUI? It is hard to believe that users were not aware of their mistakes and that they repeated the same errors over time.

### 5.2 What are the Reasons for Imperfect Recognition?

Imperfect SUI recognition can be due to the SR engine's imperfect performance and user error. In our corpus, 37.02% of the utterances that were not recognized by the SUI were not recognized due to SR engine error and 62.98% were not recognized due to human error. We encountered the following three groups of user errors:

- Uttered command was not in any of the SUI's grammars (54.19% of user errors);
- The uttered command was not in the grammar that was loaded at the time of the utterance (33.58%);
- Utterance was cut off (at the beginning, at the end or both) because the user pushed the PTT button too late and/or released it too early (12.23%).

When the officers uttered a command without an error the SR engine performed with 94.02% accuracy. This is particularly good since it reflects SR engine errors due to noises such as voice output from the police radio.

## 6. Conclusion

Considering the harsh conditions where the speech recognizer was used and that most police officers who participated in our study used this type of technology for the first time, it is not surprising that the SUI recognition rate is below 90%. It is also important to note that most of the officers we have interacted with were pleased with the performance of the SUI. While we currently do not have quantitative data (we are designing a study to determine the level of user satisfaction with the SUI), our sense is that, in general, officers were not frustrated by the problems that they encountered using the SUI. They were able to apply the time-tested approach of interacting with electronic devices: if something failed once (e.g. an utterance was not recognized), try the same thing again (e.g. repeat the utterance). This often eliminated problems due to noise or the incorrect operation of the PTT button.

## 7. Acknowledgements

This work was supported by the U.S. Department of Justice under grant 2003CKWX0151.

## 8. References

[1] GREEN, P.A., Crashes Induced by Driver Inforation Systems and What Can Be Done to Reduce Them. Convergence (2000).

[2] HEEMAN, P., The US SpeechDat-Car Data Collection. Eurospeech (2001).

[3] HUNT, M.J., Some Exprerience in In-Car Speech Recognition. IEE Colloquium on Interactive Spoken Dialogue Systems for Telephony Applications (1999).

[4] KUN, A.L., et al. Project54: Computers in Police Cruisers. IEEE Pervasive Computing, 3, 4 (2002), 23-41.

[5] WEISER, M., The Computer for the 21st Century. IEEE Pervasive Computing, 1, 1 (2002), 19-25.