New England University Transportation Center

NE University Transportation Center
77 Massachusetts Avenue, E40-279
Cambridge, MA 02139
Phone: 617-253-0753
Fax: 617-258-7570
web.mit.edu/utc

Principal Investigator:	Jim Glass	**Co-Principal Investigator:**	
Title:	Senior Research Scientist (CSAIL)	**Title:**	
University:	MIT	**University:**	
Email:	glass@mit.edu	**Email:**	
Phone:	(617) 253-1640	**Phone:**	

Final Report

Project Title:

A Study of Speech Interfaces for the Vehicle Environment

Project Number:	Project End Date:	Submission Date:
MITR21-5	3/31/2013	5/10/2013

Over the past few years, there has been a shift in automotive human machine interfaces from visual-manual interactions (pushing buttons and rotating knobs) to speech interaction. In terms of distraction, the industry views speech interaction as a "low" or "no" cost way of providing advanced information to the driver. However, limited data exist on how older and more technologically adverse operators currently experience difficulty learning command based syntaxes for speech interaction. These syntaxes are often unfamiliar, cumbersome and stepwise. The ultimate acceptability of speech based systems for in-vehicle interactions therefore lies in the development of more intuitive naturalistic modes of interaction. For example, when considering in vehicle-based speech interfaces, one would like to know how speech, language, and dialogue can be most effectively used by drivers of different ages and technological backgrounds to perform information retrieval tasks. To optimize this technology for groups with different technological backgrounds and demographics, this project investigated the extent to which a conversational spoken language speech-interface can enhance the driver and passenger experience while minimizing distraction and confusion.

This project aimed to study the effectiveness of a conversational interface in the vehicle environment. Compared to conventional command-based speech interfaces, conversational interfaces allow for more natural spoken input, and can understand an individual's queries in the context of the overall interaction. They have an ability to verbally summarize information that may be displayed on a screen, and engage in a dialogue to negotiate or clarify queries as needed during the course of an interaction. These systems are used by aging operators with differing technological skill and experience. Although conversational interfaces have the potential to increase the functionality of in-vehicle systems for operators across the lifespan, the design of these potentially disruptive technologies needs focus on more intuitive interactions to satisfy the needs of a diversifying user group. The demands for mobility from aging Baby Boomers will only seek to further enhance the importance of designing interfaces that are usable and understandable for the older driver.

By allowing drivers and passengers alike to more easily and flexibly access a wider variety of information, conversational technologies offer various benefits but also have the potential to be a disruptive technology. The objective of this project was to assess the potential applicably of conversational technology for in-vehicle telematics across two different demographic user groups. One of the important features of this study was to evaluate if an effective deployment of conversational technology can be created that is both safe and intuitive.

A BMW 5 series sedan was modified such that the vehicle display, iDrive controller, and recording buttons could communicate with a prototype conversational system (deployed on a laptop in the vehicle) Details on the implementation of the interface *City Browser* can be found in Gruenstein, Orszulak, Liu, Roberts, Zabel, Reimer, Mehler, Seneff, Glass, and Coughlin (2009). Overall, the system allows users to verbally enter a query to a database of local restaurants, hotels, museums and subway stations. In addition to the voice entry of information users could manipulate the graphical user interface through the vehicle's iDrive controller. The multi-modal interface, comprised of a graphical user interface with a conversational speech interface, provide a context-sensitive speech suggestions generator that aimed to reduce the difficulty users have with learning to use the system. This type of conversational interface was

expected to provide users with a more natural interaction commercially available vehicle interfaces at the time.

Data was collected from 94 participates across three age groups (25-34, 45-54, and 65-74 years). The final sample presented in Roberts, Mehler, Orszulak, Reimer, Coughlin, and Glass (2011) represents data from 72 cases. Results suggest that men were slightly more successful in completing tasks then women and gave the system, higher ratings on ease of getting information, feeling that the system understood what they were saying, and the extent to which they enjoyed using the system. Although the group means for the task completion rate decreased slightly with age and participants' rating of the difficulty of the tasks compared to other driving tasks increased with age, neither trend was statistically significant. However, across all participants ratings of most aspects of the system were positive. Overall results of this effort helped frame further efforts in the development of more usable voice based in-vehicle interfaces and have led to the development of a second generation system now under evaluation. Publications that directly resulted from this project include:

Roberts, S., Mehler, B., Orszulak, J., Reimer, B., Coughlin, J.F. & Glass, J. (2011). An Evaluation of Age, Gender, and Technology Experience in User Performance and Impressions of a Multimodal Human-Machine Interface. In T. Doolen E. Van Aken (Eds.), Proceedings of the 2011 Industrial Engineering Annual Research Conference (IERC), Reno, NV.

Abstract: The impact of age, gender, and technology experience on acceptance and quality of interaction was evaluated using an informational retrieval system combining manual control elements and a visual display with a naturalistic conversational speech based interface (City Browser). In addition to the technical challenges of developing useful human machine interfaces (HMIs), there is increasing recognition that individual characteristics can greatly influence potential users' interaction with new technologies and thus impact adoption. However, there is not a clear consensus as to what individual factors are most significant and under what conditions. Data was analyzed from 72 participants drawn from three age groups (25-34, 45-54, and 65-74 years) and closely balanced by gender. While there was a nominal decrease in task completion with age, the difference between age groups was not statistically significant. Gender significantly impacted performance and was also reflected in more positive ratings of various features by males. Overall, younger and older adults alike reported generally positive evaluations of the HMI with interactions between age and previous technology experience in various ratings. It is suggested that one reason for the apparent lack of a major age effect can be traced to the training provided to introduce users to the HMI.

Liu, S. (2010). Multimodal Speech Interfaces for Map-based Applications. M.Eng. Thesis, MIT Electrical Engineering and Computer Science Department, Cambridge, MA.

Gruenstein, A., Orszulak, J., Liu, S., Roberts, S., Zabel, J., Reimer, B., Mehler, B., Seneff, S., Glass, J. & Coughlin, J.F. (2009). City Browser: Developing a conversational automotive HMI. Proc. ACM Conference on Human Factors in Computing Systems (CHI 2009), Boston, MA.

Gruenstein, A. (2008). Toward Widely-Available and Usable Multimodal Conversational Interfaces. Ph.D. Thesis, MIT Electrical Engineering and Computer Science Department, Cambridge, MA.