A STUDY OF DIALOGUE MANAGEMENT PRINCIPLES CORRESPONDING TO THE DRIVER'S WORKLOAD

*Takuya Nishimoto*1, Makoto Shioya*2, Juhei Takahashi*3 and Hideharu Daigo*3*

*1 The University of Tokyo
*2 Hitachi, Ltd.
*3 Japan Automobile Research Institute (JARI)

ABSTRACT

We conducted a study in fiscal 2000-2002 concerning a network-distributed voice-activated telematics service system and another study in fiscal 2003-2004 concerning a voice-activated system and driver distraction. Based on those original studies, this paper presents dialogue management corresponding to the driver's workload and other factors, with the aim of help to develop consensus for voice-activated in-vehicle systems.

1. INTRODUCTION

Drivers are tending to operate their in-vehicle systems ever more frequently these days in order to access car navigation service and other telematics services that are being made possible by the ongoing progress of Intelligent Transport System (ITS) technologies. Accordingly, the operating ease of an in-vehicle system is one of the critical issues involved in efforts to improve driver comfort and convenience while assuring safe vehicle operation.

Voice-activated car navigation systems have been implemented in production vehicles by applying voice recognition and voice synthesis technologies. However, using voice-activated controls in situations involving a heavy driving workload, such as when turning at intersections, passing or merging with traffic, entails the possibility of causing driver's workload. One conceivable way of reducing that possibility would be to suspend voice interaction when there is a heavy driving workload. However, if the criteria for suspending voice interaction differed among the manufacturers of car navigation systems or among the systems themselves, it could affect safety and might detract from the convenience of telematics services.

Against that backdrop, the authors conducted a study in fiscal 2000-2002 concerning a network-distributed voice-activated telematics system and another study in fiscal 2003-2004 regarding a voice-activated system and driver distraction. Based on the results of those studies, this paper describes the aspects that should be shared or develop consensus among car navigation equipment manufacturers and among different systems with respect to dialogue control in voice-activated systems designed to reduce driver's workload. Furthermore, it presents the results of an examination of dialogue management principles and compliance procedures in dialogue control corresponding to the driver's workload, as considered from the standpoint of voice technology researchers.

2. NETWORK-DISTRIBUTED VOICE-ACTIVATED SYSTEM

This research envisioned a network-distributed voice-activated system [1,2] capable of providing over a network the various types of information that drivers need, in a timely manner and without interfering with driving operations. Studies have been conducted for the purpose of developing what might be called a "voice-activated driving partner" that supports conversational interaction while discerning the driver's circumstances.

Toward that end, one objective was to optimize the collaboration and functional distribution between an in-vehicle system that uses voice recognition/synthesis technologies and the network. Figure 1 shows a conceptual diagram of such a network-distributed voice-activated system. Furthermore, it was confirmed on the basis of experiments that a voiced-activated in-vehicle system needs to incorporate dialogue control capability that takes into account driver's workload and information presentation functionality that considers the priority level of information.

The following insights were gained through this research concerning a safe human-machine interface (HMI):

(1) Commonality of voice commands and sound effects as well as dialogue control corresponding to the driver's workload are effective in reducing impacts during driving.

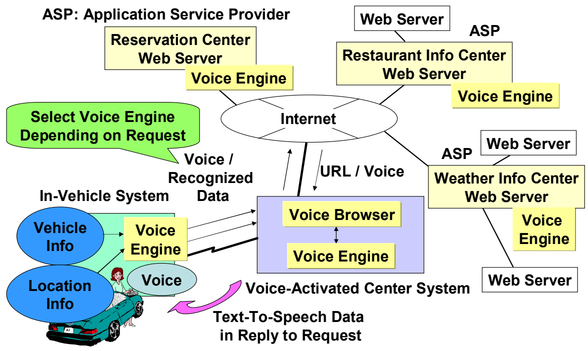

Fig. 1 Concept of Network-Distributed Voice-Activated Telematics Service System

(2) Information must be presented to drivers in a way that considers the priority level, and it is important to strike a balance between the timing for presenting information and its priority.
(3) Items (1) and (2) above differ depending on the driver's driving skill and degree of proficiency with a voice-activated system. It is necessary to have a function for adaptively varying the response of a voice-activated system to match such factors.

3. VOICE-ACTIVATED SYSTEM AND DRIVER DISTRACTION

The results of the above-mentioned studies confirmed the necessity of having dialogue control that takes into account driver distraction [3-5].

Accordingly, it was decided to continue with research that included the possible consensus of voice-activated technology, and for that purpose, a study was launched concerning voice-activated systems and driver distraction. This work was conducted by the Voice Telematics Working Group, consisting mainly of voice-related researchers who engaged in 3 limited Japanese electronics equipment manufacturers and the research of a network-distributed voice-activated system, which was newly formed under the Mobile Systems Committee that was established to support standardization activities at the Japan Automobile Research Institute (JARI).

The following objectives were determined for this research project, taking into account the possible consensus of voice-interaction control and other measures for reducing impacts during driving.

(1) To make clear the items that should be shared by voice-activated in-vehicle systems
(2) To prepare a draft proposal of the requirements for voice-activated systems

4. TENTATIVE PROPOSAL FOR DIALOGUE MANAGEMENT

As the first step, desirable dialogue management in the use of voice activation was provisionally examined, based on the insights gained in the research concerning a network-distributed voice-activated system.

4.1. Scope of dialogue management principles and compliance procedures

Dialogue management principles and compliance procedures were examined within the following scope.

(1) Dialogues between the driver and an in-vehicle telematics system while a vehicle is in motion.
(2) In such dialogues, voice is used as the main medium for inputs from the driver to the in-vehicle system. Voice inputs are primarily recognized by using voice recognition technology.
(3) A synthesized voice or sounds are used as the principal media for conveying the outputs of the system to the driver in such dialogues.
(4) The in-vehicle telematics system not only provides traffic information, points of interest (POI) information, news and weather reports, it also supports various popular Internet services such as e-mail, e-commerce, information searches and entertainment.
(5) The aim of the study concerning dialogue management principles and compliance procedures is to show examples of the basic principles, requirements and recommendations with respect to accessing telematics services within the scope of the driver's spare mental capacity after processing the driving workload (i.e., without the occurrence of any driver's workload).

4.2. Basic principles and requirements with regard to dialogue management proposed in this study

4.2.1 Proposing principles

A critical basic principle of dialogues for obtaining telematics services is that first priority is given to the processing of the driving workload as the primary task. This principle for putting priority on the processing of the driving workload can be explained as follows:

- Priority is given to the processing of the driving workload, so long as the driver always has sufficient mental capacity for handling the workload involved.
- The driver should keep his/her mind on driving until the processing of the driving workload is finished. In other words, so long as attention is not directed to things other than driving, priority is given to the processing of the driving workload.

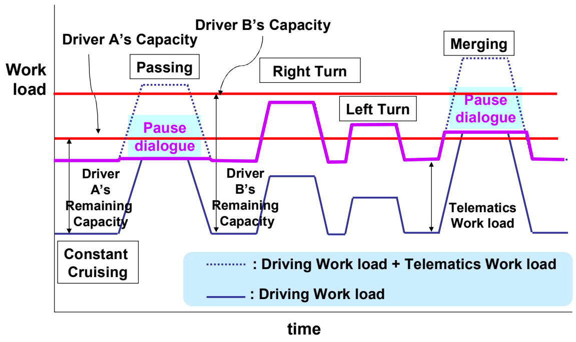

Fig. 2 An Example of Dialogue Control for Driver B

Or the principle can also be expressed as follows:

- If the driver engages in a dialogue for accessing a telematics service (secondary task) within the scope of the person's spare mental capacity after having processed the ever-present driving workload, then priority is being given to the primary task.

Figure 2 shows one example of the concept of dialogue control corresponding to the driver's spare capacity.

4.2.2 Proposing requirements

The following aspects were considered as fundamental requirements for complying with the basic principles.

- Dialogues for accessing telematics services must be controlled such that the driver's capacity for processing the workload involved in driving is secured first, and the workload for obtaining the telematics service is processed within the scope of the driver's remaining spare capacity.
- The workload involved in driving and the workload for accessing a telematics service should either be estimated in advance or measured in real time.
- The driver's capacity for processing these workloads should either be estimated in advance or measured in real time.
- Dialogue control should be executed in a way that takes into account the driver's state as well as various constantly changing driving conditions such as the road environment, traffic conditions, vehicle status, trip status and so on.
- Dialogue control should be executed in a way that gives consideration to the priority level of various types of services.
- The driver should be informed in a suitable manner of the status of the in-vehicle telematics system.
- The driver should have ultimate control over the dialogues.

4.3. Requirements of in-vehicle telematics system proposed in this study

4.3.1 Proposing system concept

The basic system concept is that the use of telematics services should be controlled in situations involving a heavy driving workload in order to ensure safe vehicle operation. In other words, in situations where the driving workload plus the telematics service workload is equal to or greater than the driver's capacity to handle these workloads, access to telematics services is controlled.

The control procedure takes into account adaptive functionality whereby control is performed in a manner that depends on each driving situation and the capacity of individual drivers.

4.3.2 Proposing system requirements

The essential elements making up an in-vehicle telematics system are listed below.

(1) A means of conveying indexes of the driving workload, telematics service workload and the driver's capacity for handling them to the system.
(2) A means of judging situations where the driving workload plus the telematics service workload is equal to or greater than the driver's capacity to handle these workloads.
(3) A means of conveying the situation to the driver.
(4) A means of controlling the telematics service workload depending on the situation and the driver's capacity.
(5) The elements above should allow for personal customizing through selection and adjustment according to the individual driver's driving skill and degree of proficiency with a voice-activated system.

In order to improve the transparency of the system to drivers, it will be important to:

(a) make the system easy to understand,
(b) make it easy for drivers to become accustomed to the system, and
(c) include a tutorial.

4.3.3 Recommended system functions

The in-vehicle system should incorporate the following functions in order to meet the basic requirements.

(1) A means of detecting various driving situations.
(2) A means of judging the possibility of driver's workload occurring.
(3) A means of conveying the situation to the driver.
(4) A means of controlling the telematics service workload in a manner that depends on the situation and the driver's capacity.

Specific examples of these functions and their applications are described below.

(1) The following are among the possible means of detecting various driving situations.

(a) Examples of methods of detecting the driving workload index

- Through the detection of driving operations by using information on the driver's operation of the brake pedal, accelerator pedal, steering wheel, turn signals and other devices.
- By using information output by in-vehicle equipment such as a speed sensor, accelerometer, angular velocity sensor, radar unit, camera or other devices.
- By using position information or map information such as information on the road geometry, road grade, road width, traffic restrictions (maximum speed, one-way traffic, no right turn, etc.) or road type (city street, school zone street, expressway, etc.).
- By using the time of day or weather conditions such as the outputs of a road surface sensor, sunlight sensor, raindrop sensor, windshield wiper sensor, headlight sensor, cross wind sensor or other devices.

(b) Examples of methods of detecting the telematics service workload index

- By different types of service such as according to the respective index for entry and other operating procedures and the difficulty of the dialogue involved in route guidance, parking searches, restaurant searches, payment, e-mail, online quizzes, karaoke, music title searches and so on.
- By using the telematics information tags such as the respective index for traffic restrictions, accident/disaster information, congestion information, route guidance and other items.

(c) Examples of methods of detecting the driver's capacity

- Based on age
- Based on the number of years since obtaining a driver's license or the number of years of driving experience
- Based on the number of years of accident-free driving
- Based on a self-assessment such as a declaration of one's level of driving ability, motor skill and information processing ability and a ranking of HMI ability
- Based on physiological abilities such as by having various physiological functions measured at the time of driver's license renewal and by using those self-declared values

(d) Examples of ways of obtaining temporary self-assessments from drivers

- By having a driver perform a barge-in task by voice input or touch screen input and using the result to temporarily revise reported data on the person's capacity.

(2) Examples of methods of judging the possibility of driver's workload occurring

(a) Based on the results of situation detection

Calculation and judgment of whether the driving workload plus the telematics service workload is equal to or exceeds the driver's capacity to handle the workloads

(3) Examples of methods of conveying the situation to the driver, assuming that voice and sounds are the main media used

(a) For conveying recommended actions to the driver from the system

- Use of voice announcements according to the situation.
- Use of sound effects or musical effects according to the situation.
- Basic separation into about three levels (utterance accepted, request for utterance, system malfunction) which are advised to the driver.
- Designing of easy-to-understand prompts in terms of the number of words (number of syllables), speed and voice range.
- Designing of easy-to-understand voice announcements in terms of the number of words (number of syllables), speed and voice range, and provision of a backup function for advising the driver of the system status.
- The timing and interval of a recommended action should match each type of situation in order to be effective.

(b) It is desirable to convey information to drivers via a multimodal means using sound effects, voice, images (videos, photos, graphics, colors, etc.), text or other media suitable to the situation.

If the media are also to be used while a vehicle is in motion, each one should be expressed in a way that is easy for drivers to understand. Visual, auditory and tactile means should be used suitably depending on the level of emergency or priority.

(c) Means of conveying information to drivers should take into account the ranking of HMI devices.

Because the ability to handle HMI devices differs among individuals, the devices should be ranked and information conveyed to drivers accordingly. HMI devices should be ranked according to the information processing levels of each user age group, including older drivers, and their proficiency with the processing procedures.

(4) Examples of methods of controlling telematics service workloads depending on the situation and the driver's capacity

(a) Provision of functions for controlling operations or dialogues according to the situation

Controlling and limiting dialogues depending on the level of risk, emergency or priority.

(b) Details of dialog control

- The system should possess a function for interrupting a dialogue when the driving workload becomes heavy and for resuming the dialogue again when the workload becomes lighter. The system should have functions for controlling operations and dialogues separately according the nature of the utterances.
- Several modes of dialogue interruption should be provided, including instantaneous interruption and interruption at a natural break in an utterance. Dialogue sentences and styles should be convertible to ones that are easily understood by drivers, including the use of key words, main points, summaries, menus, headings, full text, chapters, sections, pages and paragraphs.
- Visual, auditory and tactile means should be used suitably depending on the level of emergency or priority. Presentation timing should take into account ease of understanding by drivers. The level of consistency (inconsistency) with the content of other information should be indicated. The system should display the symbols and icons for indicating congestion, road construction, travel time and information.

(c) Dialogue control commands

Around twelve different types of voice commands, for which there is 99% recognition, should be prepared. For in-vehicle systems in Japan, voice commands should be in both Japanese and English. Voice commands should be provided that enable drivers to control the progress of a dialogue.

(d) Dialogue control taking into account the ranking of HMI devices

Drivers' capacity to handle HMI devices varies from individual to individual and also depending on personal preferences and the situation. Accordingly, it should be possible to control dialogues according to individual capacities to handle HMI devices, based on a ranking of the devices.

- It should be possible to select an HMI device level that matches user profiles, including personal preferences.
- It should be possible to select an HMI device level corresponding to the situation. It should be possible to make adjustments within HMI device levels according to user profiles and the situation.

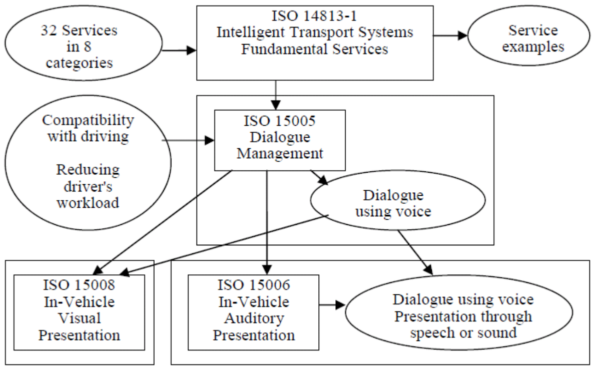

Fig. 3 Typical HMI-Related ISO Standards

(e) Consideration of task success rates and reduction of user dissatisfaction until task accomplishment in the case of HMI devices based primarily on voice interaction

- Total task time required to complete a task should be reduced as much as possible, as should the total number of steps required.
- The voice recognition rate should be improved. Synthesized voice quality should be enhanced. Response should be improved.
- The number of operations, number of voice recognition failures and number of recognized candidates should all be reduced as much as possible (see [3] for further details).

5. INTERNATIONAL STANDARDS POSSIBLY RELATING TO THIS STUDY

A survey was made of international standards related to dialogue management with the aim of developing a dialogue control standard for voice-activated systems designed to reduce driver's workload. The results made the following points clear.

(1) There are relatively few standards that directly govern voice-based dialogue control, but there are many standards related to visual aspects.

(2) The standard that is the closest to dialogue control is ISO 15005 [6], and the title and scope of application of this standard seem to be identical to what we are considering in the sense of dialogue management in the telematics field.

The correlation among the four international standards that were surveyed is outlined in Fig. 3. The standards are mentioned in the boxes, and the arrows indicate quotations or directions to take for references. The ISO 15005 standard is organized around the aim of reducing drivers' workloads due to the operation of an in-vehicle system so as to achieve compatibility between telematics service use and driving.

It is thought that a dialogue management standard, including a voice interaction standard as well, could be developed by simply complementing the dialogue principles of ISO 15005 as follows:

> *A citation from ISO 15005:*
> 5.2.4.3.4 Recommendation
> TICS should provide timely visual information to the driver.
>
> *An example of a supplemental recommendation:*
> 5.2.4.3.4-1 Recommendation
> Voice-activated telematics systems will provide the driver with timely audible information.
> Example 1: In situations where the driving workload increases such as when turning right or left at an intersection, the driver will be advised sufficiently in advance of the maneuver so that the voice-based telematics service dialogue can be performed safely.
> Figure 4 shows an example of dialogue control during the execution of a right turn, which is one situation involving a heavy driving workload. The dialogue is interrupted and resumed based on the On/Off status of the turn signal lever, respectively. "Interrupt" and "resume" commands are also provided so as to give precedence to the driver's judgment.

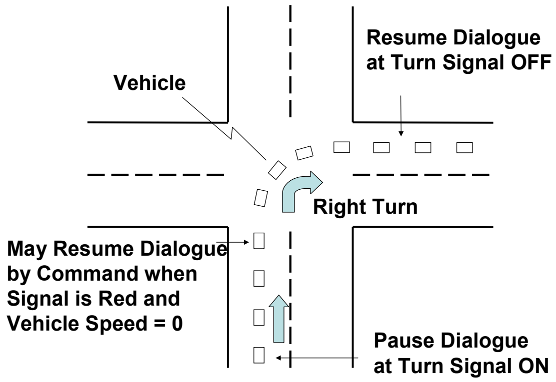

Fig. 4 Dialogue Control in a Right Turn

6. CONCLUSION

This study has made clear various aspects that should be shared or develop consensus among car navigation equipment manufacturers and among different systems with respect to dialogue control in voice-activated systems designed to reduce impacts during driving. Desirable dialogue management corresponding to the driver's workload were also examined and tentatively proposed from the viewpoint of voice technology researchers. This study has been done as part of general research toward the development of consensus for voice-activated telematics systems.

It is also important to choose suitable methods for measurement of workload for voice-user interface systems [7] and to assess the appropriateness of our proposal.

It is hoped that this research will serve as a first step toward collaborative activities between parties involved with automotive HMI technologies and those involved with electronics-related voice technologies, with the aim of developing voice activation consensus for original and after-market in-vehicle systems.

7. ACKNOWLEDGMENTS

This paper compiles the results of research conducted by the Voice Telematics Working Group under the Mobile Systems Committee at the Japan Automobile Research Institute. The authors would like to express their deep appreciation to all of the members involved in this project.

8. REFERENCES

[1] Association of Electronic Technology for Automobile Traffic and Safety, Feasibility Study Concerning the Infrastructure for Network-distributed Voice-activated In-vehicle Telematics Systems, a study commissioned by The Mechanical Social Systems Foundation, March 2003.

[2] Makoto Shioya, et al. Research on Network-Distributed Voice-Activated System Architecture for Telematics Services. 10th ITS World Congress, Madrid, 2003.

[3] Japan Automobile Research Institute, Mobile Systems Committee, Voice Telematics Working Group, A Study of Voice-activated Systems and Driver Distraction, March 2005.

[4] T. Nishimoto, Symposium on Ergonomics and the Usefulness of Mobile Phones and Car Navigation Systems, Japan Ergonomics Society, pp.125-128, Kyoto, March 2004 (in Japanese).

[5] Makoto Shioya, et al. Study on Reference Models for HMI in Voice Telematics to meet Driver's Mind Distraction. 11th ITS World Congress, Nagoya, 2004.

[6] ISO15005 Road vehicles - Ergonomic aspects of transport information and control systems - Dialogue management principles and compliance procedures.

[7] Takuya Nishimoto, Motoki Takayama, Haruaki Sakurai, Masahiro Araki: Measurement of Workload for Voice User Interface Systems, Systems and Computers in Japan, Volume 36, Issue 8, pp.81-89, May 2005.