SPOKEN DIALOGUE INTERFACE IN A DUAL TASK SITUATION

Shuichi Tanaka, Shu Nakazato, Keiichiro Hoashi, and Katsuhiko Shirai
School of Science and Engineering, Waseda University

ABSTRACT

In this paper, we examined the effects of the spoken dialogue interface in a dual task situation. One of the well suited situations for the use of the spoken dialogue interface is that including dual tasks, e.g. when several tasks are being carried on at the same time. In order to develop a spoken dialogue system, it is important to consider such a situation and the effects of the spoken dialogue interface in it. It has been said that the merit of spoken dialogue interface is that it doesn't make the user's manual and visual modalities busy. Therefore, it is expected that this interface is effective in a dual task situation. We made experiments to prove this hypothesis. As the result of these experiments, it became clear that spoken dialogue interface was effective in a dual task situation and that considering the situation is important for the design and evaluation of the interface.

1. INTRODUCTION

The recent progress of speech recognition technology is remarkable, and the application of the technology is expected to improve man-machine interface. However, speech recognition is not a perfected technology yet; there is no speech recognizer that does not make recognition errors. Improvement of the system interface is necessary to compensate such errors. The multimodal interface is one of such improvements. Another idea of improvement is to find the application to a situation in which existing speech recognition technology is still effective. A simple example of this is a situation which speech is the only usable input method. It is important to consider such situations which existing input methods such as the keyboard or the mouse are impossible to use.

The situation in which a spoken dialogue system is used is important for the system to be used effectively. It is quite obvious that the demand of the user will differ whether if he/she can or cannot use input devices other than speech. For example, demands for a car navigation system interface should be different whether the user is driving or not.

Considering previously proposed merits of the spoken dialogue interface, input by speech can be effective when the user's hands are busy, and output by speech can be effective when the user's eyes are busy. This interface is expected to be especially effective in a situation when the user is working on several tasks concurrently. In this research, we considered the effectiveness of the spoken dialogue interface in this "dual task situation".

2. PROBLEMS OF SPOKEN DIALOGUE SYSTEMS

Numbers of research on spoken dialogue systems have been made in recent years. In many of such previous research, the main idea was: "What system or what user is the spoken dialogue interface effective for?" However, few researches have been made focusing on the situation which spoken dialogue interface is effective for. This is an important problem, since the user's satisfaction towards a system will differ depending on the situation he/she is using it. A realistic example of the system and the situation surrounding it is shown in Figure 1, where the user is using the system (a car navigation system) while driving a car.

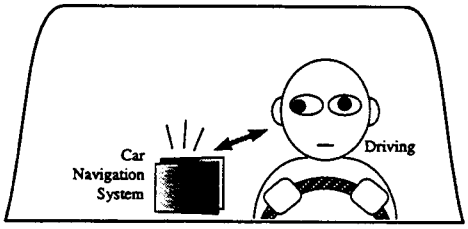

Figure 1 An example of the system and the surrounding situation

In this paper, we considered the situation surrounding the spoken dialogue system as an essential element of the system itself.

3. DUAL TASK SITUATION

One of the merits of the spoken dialogue interface is that it does not make the user's manual and visual modalities busy. Therefore, the spoken dialogue interface is expected to be effective in a dual task situation which manual and visual modalities are busy in one task. However, few researches have been made on the effectiveness of the spoken dialogue interface in such situations.

A "task" is to work towards a goal. A dual task situation

is a situation in which there are several goals, and one is proceeding on tasks towards those goals. Contrarily, the situation which one is proceeding on a single task towards a single goal is called a single task situation.

In this paper, we took up the dual task situation as a situation in which the spoken dialogue interface is effective, and examined its effectiveness in this situation. From the results, we considered the importance of the situation in which the spoken dialogue interface is used.

4. EXPERIMENT IN THE DUAL TASK SITUATION

In this paper, we will limit the number of tasks and goals in the dual task situation to two, and we will refer to the two tasks in this situation as the "Main Task", and the "Sub Task."

4.1. Method

Objects The object of this experiment is to prove the effectiveness of the spoken dialogue interface in a dual task situation. In order to do this, we changed the combination of modalities for the Main Task and the Dual Task.

Task As the task for this experiment, the subject inputs his/her response to the system's output, which is either a number or a color, by mouse or speech. The dual task situation in this experiment is made by making the subject work on two tasks as written concurrently.

We prepared two conditions for the pace of the system output, "force-paced", and "self-paced". Force-paced means the output is presented to the subject regardless of the pace of his/her response. Self-paced means the output is presented to the subject after his/her response.

Experiment

1. Single task (Main Task)
   The output is shown to the subject 10 times at the system's pace (force-paced). The subject inputs the output back to the system. This task was done 4 times, for all the combinations of input/output methods. (See Table 1)
2. Single Task (Sub Task)
   The output is shown to the subject 10 times at the subject's pace (self-paced). The subject inputs the output back to the system. This task was done 4 times, for all the combinations of input/output methods. (See Table 1)
3. Dual Task
   - Main Task
     The output is shown to the subject at the system's pace (force-paced). The subject inputs the output back to the system.
   - Sub Task
     The output is shown to the subject at the subject's pace (self-paced). The subject inputs the output back to the system.

     The task ends when the Sub Task is accomplished 10 times.

     This task is done 16 times, for all the combinations of input/output methods for the Main Task and the Sub Task.

Table 1 Combinations of input/output modalities for single task

Combination	Input	Output
MV	manual	visual
SV	speech	visual
MA	manual	auditory
SA	speech	auditory

4.2. Results

We evaluated the task efficiency of single tasks by the response time of the subject. There was no significant difference between the color output and the number output in force-paced conditions, but in self-paced conditions, the subject's response time was significantly shorter for the number output (p <0.01).

We evaluated the efficiency of Dual Tasks for each combination of modalities by comparing the response time of each task done under the Single Task or Dual Task conditions. There was no significant difference between the response time of Group A and B, since the time was compared to that of the Single Task conditions.

The result is represented on a single point in a performance operating characteristic (POC) space (Figure 2)[1].

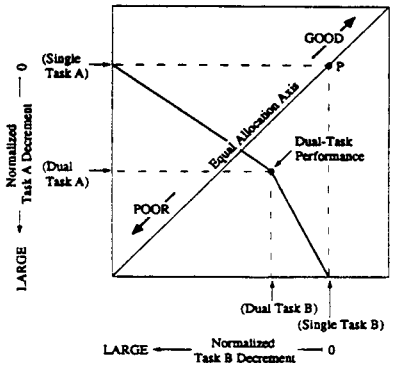

Figure 2 Hypothetical representation of dual task performance within the POC space

That point represents the performance in both tasks relative to their respective single-task performance levels. Shifts along the positive diagonal towards the upper right represent improvements in time-sharing efficiency.

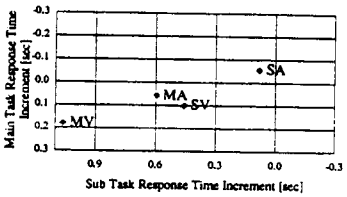

Figure 3 Change of response time(Main Task MV)

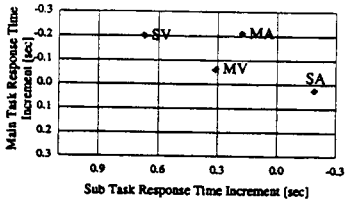

Figure 4 Change of response time(Main Task SV)

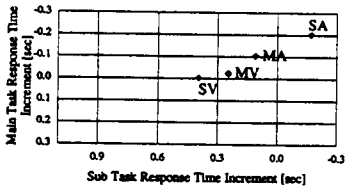

Figure 5 Change of response time(Main Task MA)

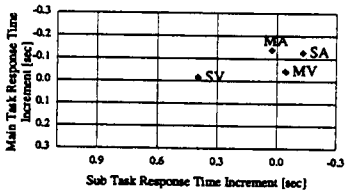

Figure 6 Change of response time(Main Task SA)

Of the 4 combinations of input/output modalities, SA is at the rightmost position, significantly different from MV and SV. From this result, it can be presumed that the spoken dialogue interface is the most efficient interface in a dual task situation.

Moreover, the Sub Task SA is at the rightmost position of all of the POC spaces. From this result, it can be presumed that the performance in the dual task situation is the best when using the spoken dialogue interface, regardless of the modalities in the Main Task.

4.3. Discussion

From the results of this experiment, it is clear that when visual and manual modalities are used in the Main Task, the use of the spoken dialogue interface in the Sub Task improves task performance. Therefore, in a dual task situation where the visual and manual modalities are busy in the Main Task, the spoken dialogue interface is the most effective interface to apply.

Furthermore, the spoken dialogue interface was found to be the most effective in all dual task situations, regardless of the modalities used in the Main Task. This means that the use of modalities which are not used in the other task of the dual task situation does not always improve task performance. The spoken dialogue interface proved to be an effective interface in the dual task situation, considering the fact that it did not decrease the performance much in the experiment.

5. APPLICATION TO CAR-NAVIGATION SYSTEM

In this section, we will explain the experiment made to examine the effectiveness of a multi-modal car navigation system[2] in the dual task situation.

5.1. Evaluation experiment

Object The object of this experiment is to evaluate the effectiveness of the multi-modal car-navigation system in a dual task situation.

Task The Main Task is driving while using a car-navigation system as the Sub Task. Instead of actual driving, the Main Task for this experiment is the same task as in the previous experiment. A color output is shown to the subject at the system's pace (force-paced), and the subject inputs it back to the system with the mouse (MV Task). The Sub Task is the use of the car-navigation system.

The task using the car-navigation system was to find a given destination and display it on the map screen. The subject first searches a point near the destination with the map search function. Then the subject searches the given destination with the map control function.

Subjects The subjects for this experiment were 12 students. They were divided into 3 groups, depending on the input method of the car-navigation system for each experiment. As seen in Table 2, the subjects did two Single Task experiments, and two Dual Task experiments.

Group A consists of 6 subjects. This group can use either manual or vocal input methods for all 4 experiments. Group B and C consists of 3 subjects each. Groups B and C are capable of using both manual and vocal input methods for the Single Task experiments, but can only use one of these methods for the Dual Task experiments. These two groups are divided by the order of the usable modality for the Dual Task experiments.

Experiment

1. Single Task
   The subject operates only the car-navigation system under the conditions in Table 2, to achieve the task previ-

Table 2 Conditions of Experiment

Group	Single Task(Exp1)		Dual Task(Exp2)	
	1st	2nd	1st	2nd
A	m-modal	m-modal	m-modal	m-modal
B	m-modal	m-modal	vocal	manual
C	m-modal	m-modal	manual	vocal

(*m-modal = multimodal*)

ously written.

2. Dual Task
   The subject does the same task as in the Single Task experiment, under the conditions in Table 2, while doing the Main Task MV.

5.2. Results

The ratios of manual(Ma) and vocal(Vo) input of the car-navigation system under the Single Task situation and the Dual Task situation when the subject is able to use both input methods are shown in Fig 7. It is clear from this Figure that the manual input is more frequently used in the Single Task situation(Vo: 11.5%, Ma:88.5%), and the vocal input is more frequently used in the Dual Task situation(Vo: 60.7%, Ma:39.3%). The difference of the modality used by the subject in these situations is undoubtful.

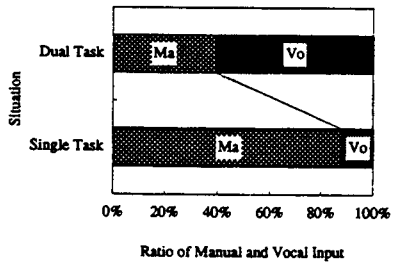

Figure 7 Ratio of manual and vocal input

The task achievement time for the manual or vocal input when the subject was restricted to use only one of the input methods is shown in Fig 8. The number of the change of glances from the Main Task to the Sub Task is shown in Fig 9. The error rate of the Main Task is shown in Fig 10.

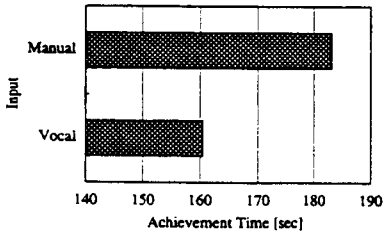

Figure 8 Task achievement time

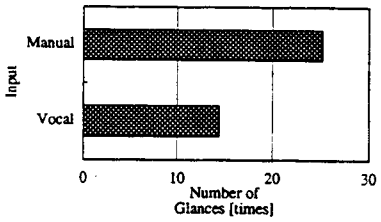

Figure 9 Number of change of glances

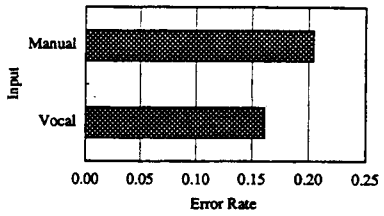

Figure 10 Error rate

The voice input showed better performance in the error rate(Vo:0.16%, Ma:0.20%) and the task achievement time(Vo:160.4*sec*, Ma:183.2*sec*). The number of glance changes was higher when using the manual input(Vo: 14.4, Ma:25.2). A significant difference was recognized for this result ($p < 0.05$).

6. CONCLUSION

We considered the dual task situation as a situation in which the spoken dialogue interface would be effectively used. We examined the effectiveness of the interface in this situation.

Based on this conclusion, we examined the operation of a car-navigation system while driving as a realistic dual task situation, and evaluated this system. The vocal input was used more frequently in the dual task situation than the single task situation in this experiment, meaning that the spoken dialogue interface is effective in a realistic dual task situation.

The results also show that the user's way of using a system changes by the surrounding situation. Therefore, when designing or evaluating a spoken dialogue system, it is necessary to consider not only the functions of the system, but also the situation in which the user will use the system.

7. REFERENCES

[1] Wickens,C.D., Mountford,S.J., Schreiner,W.: "Multiple Resources,Task-Hemispheric Integrity,and Individual Differences in Time-Sharing", Human Factors, 23(2), pp.211-229(1981).

[2] Kameyama,S., Nakazato,S., Shirai,K.: "Spoken Dialogue on Car Navigation System", Proc. Spring Meeting ASJ, pp 21-22(1994.3).(In Japanese)