Evaluation of Communication with Robot and Agent: Are robots better social actors than agents?

Junji Yamato, Kazuhiko Shinozawa, Futoshi Naya, & Kiyoshi Kogure
NTT Communication Science Laboratories,
2-4 Hikari-dai, Seika-cho, Kyoto, 619-0237 JAPAN
junji@cslab.kecl.ntt.co.jp

Abstract: We experimentally investigated the effects of an embodied conversational agent and a robot on user decision making by using a simple color-name selection problem. The experiments were conducted using 5 groups of 30 people each; each group was tested under a different condition. The agent or robot recommended which name the subject should select for each presented color. The agent had more effect on the user's choice even though the subjects felt a higher degree of familiarity with the robot. We also investigated the effects of two strategies used by agent to gain the user's confidence, and found one was more effective than another. These results indicate that the design rules for interactive robots are different from those for agents and that the effect of an agent or robot may not be evident to the users.

Keywords: Conversational agent, Robot personality, Decision making

1 Introduction

The recent rapid growth in computer power has enabled the development of embodied conversational agents that interact with their users in a natural and friendly manner by speech recognition, synthesized voice, and action display. Also being developed are personal robots that can serve as communication partners or ‘‘pets”. Because robots are corporal and exist in the real world, not on a screen, they can be effectively used for communication if they are correctly designed. A computer can be given a personality by using minimal superficial cues(Reeves & Nass, 1996). How can a robot be given a personality? What are the differences and similarities between agents and robots? To determine the key points for designing the interaction of communication robots, we experimentally evaluated the interactions between agents and users, and between robots and users. We focused on the effect on user decision making of the recommendations made by the agent or robot.

2 Method

The task we designed to quantitatively measure the effect of agent/robot recommendations was as follows. The subjects were shown color squares one at a time on a computer display and given two candidate names for each color. Most of the colors and candidate names are unfamiliar to ordinary people. The subjects were asked to name each color as it was displayed. The answer was not obvious, and most subjects had no prior reference. The agent or robot suggested which color to choose, and the subject could accept or reject the recommendation. The subjects were told that the experiment was a color-name recognition test, not an agent/robot effect test. After the subject named the color, the agent/robot showed pleasure if the choice was the same as the recommendation and disappointment if it was not. We expect the subject to more readily accept the recommendation if he or she felt a higher degree of familiarity with the agent/robot.

The agent system was developed at the NTT East and is based on the Microsoft Agent. It uses the ‘‘Fluet” Japanese speech synthesizer developed by the NTT Cyber Space Laboratories (Mizuno & Nakajima, 1998). The robot was developed by the NTT Cyber Solution Laboratories, and we developed the experiment system.

We conducted three experiments under different conditions using the same color-name selection problem. The subjects were 150 adults, divided into five equal-size groups A to E. They ranged in age from 18 to 55; there were 70 men and 80 women.

Experiment 1: No Agent vs. Agent

Group A received no assistance in naming the displayed color. Group B was aided by a conversational agent displayed on the screen that recommended one of the two candidate names.

Experiment 2: Different agent strategies

We tested two different agent strategies to determine which type of strategy is more effective for affecting users' decisions. In both strategies, a sequence of 30 color selections were grouped into a biased section and a common evaluation section. In one strategy, when the selections in the biased section were presented, the agent gradually decreased the similarity of its recommendations to the average user preferences, which had been measured by a pretest. We call this the ''down-type" strategy. In the other strategy, the agent increased the similarity. We call this the ''up-type" strategy(Shinozawa et al, 2001). We used the down-type strategy for group B and up-type strategy for group C.

Experiment 3: Comparing Robot and Agent

We used a small personal robot with seven degrees of freedom: two head motions and one motion for the mouth, each arm, the waist, and the eyelids. The actions were designed to be similar to those of the agent. Group D subjects were assisted by the robot dialog, and Group E subjects were assisted by the agent. Because a robot can not be equipped with a dialog balloon, we removed the balloon from the agent. The down-type strategy was used for both groups. The robot and the agent used exactly the same speech.

Figure 1 : Experimental setup for robot-user interaction

3 Results

In experiment 1, the number of answers that were the same as the recommendation was on average 16.2 for Group B and 14.3 for Group A, and the difference was significant ($p^*<0.02$). In experiment 2, the down-type strategy resulted in 18.6 matches on average and the up-type strategy resulted in 14.7 (Figure 2). The difference was significant ($p^*<0.05$). The answers to a post-test questionnaire showed that the ''familiarity" factor was the same for both groups, indicating that users were not aware of the agent's effect. The results for experiment 3 were not what we expected. With the robot, the number of matches was 14.8 on average, while with the agent it was 17.6. However, the familiarity factor was much stronger for the robot (4.0 and 3.4, $p^*<0.05$), as shown in Figure 3.

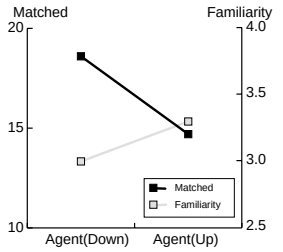

Figure 2: "Up" vs. "Down" strategies

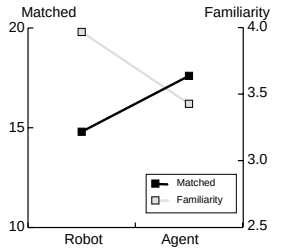

Figure 3: Robot vs. Agent

4. Discussion

These results show that an agent can affect a subject's decision and that a down-type recommendation affects decisions more than an up-type recommendation. Also, a robot does not necessarily have a larger effect than an agent, in spite of its embodiment and stronger familiarity factor Based on the answers of the post-test questionnaire, the subjects recognized the speech of the robot and that of the agents equally well, and the familiarity factor was stronger for the robot than for the agent. This suggests that robots can potentially communicate better with users, although the design rules for robot-user interaction may differ from those for agent-user interaction. To make the robot more effective, we should take advantages of its real embodiment . We plan to control the direction of its eyes to achieve better eye contact and to use color squares printed on paper for the robot experiment, instead of displaying on CRT. These modifications will enable shared attention to the real object that is most powerful advantage of the robot in the real world.

Acknowledgements

We thank Syun-ichi Yonemura and Mio Hosoya of the Cyber System Project Group of the NTT-East R&D Center for providing the agent system, the Business Innovation Project of NTT Cyber Solution Laboratories for developing ''CyberPerformer", the personal robot, and the Speech Synthesis Group of the NTT Cyber Space Laboratories for providing and fine tuning the ''Fluet''speech synthesis system.

References

Mizuno, O. & Nakajima, S., Synthetic Speech/Sound Control Language: MSCL, 3rd ESCA/ COCOSDA *Proceedings of International Workshop on Speech Synthesis*, 1998, pp. 21-26.

Reeves, B. & Nass, C. *The Media Equation.*, Cambridge University Press. 1996.

Shinozawa, K. et al., Quantitative evaluation of effect of embodied conversational agents on user decision, *Proceedings of HCI International 2001(to appear)*