Effects of Conversational Agent and Robot on User Decision

Junji Yamato, Kazuhiko Shinozawa, Futoshi Naya, and Kiyoshi Kogure

NTT Communication Science Laboratories
2-4 Hikari-dai, Seika-cho
Kyoto, JAPAN 619-0237
{junji, shino, naya, kogure}@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 recommendation is not based on logical persuasion nor objective evidence, but only on wish of agent or robot. 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. The one in which the agent gradually reduced the similarity of its recommendations to the average user preferences resulted in 18.6 matches out of 30. The opposite strategy resulted in 14.7. 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.

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 [1]. Reeves and Nass gave many examples that people tend to take a humanlike attitude toward the media. People find a personality in the computer according to the characteristic of voices and displayed images and feel the different veracity for information from computers in spite of same information as text.

Comparing the robot with these media or agents, we expect that robot can be more easily given a personality, because robots have body and it makes uses to feel that robots exist in the same world as them. However, there has been no concrete comparison study. To determine if robots are actually better social actors, and to find what are the differences and similarities between agents and 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.

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, and were not given any accounts of agents or robots. 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 or robot.

The agent system was developed at the NTT East R&D Center and is based on the Microsoft Agent. It uses the ``Fluet" Japanese speech synthesizer developed by the NTT Cyber Space Laboratories [2]. 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: Agent and strategies

Group A received no assistance in naming the displayed color. Group B and C were aided by a conversational agent displayed on the screen that recommended one of the two candidate names. 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[3]. We used the down-type strategy for group B and up-type strategy for group C.

Figure 1: Experiments with Agent

Experiment 2: 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 robot system, "CyberPerformer" was developped at NTT Cyber Solution Laboratories. 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. These groups were compared with group A, no-agent condition.

Figure 2: Experiments with Robot

Results

In experiment 1, the number of answers that were the same as the recommendation in the last ten questions were compared (Fig.2). The sections were same for all three conditions and we can measure the effect of the agents.

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 ($F=5.98$, $p^*<0.05$) by the analysis of variance. Group C, the agent with the down-type strategy, was more effective than both Group A and Group C from Scheffe's post-hoc test ($p^*<0.05$).

The results for experiment 2 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 (Figure.3). The difference was significant ($F=7.59$, $p^{**}<0.01$), and the group E was more effective than both Group A and Group D from Scheffe's post-hoc test ($p^*<0.05$).

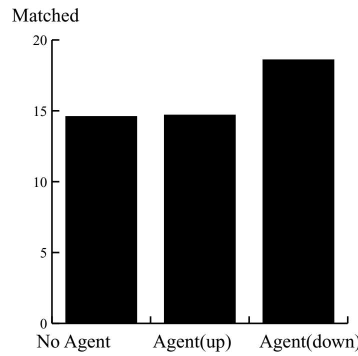

Figure 3: "Up" and "Down" strategies Agents

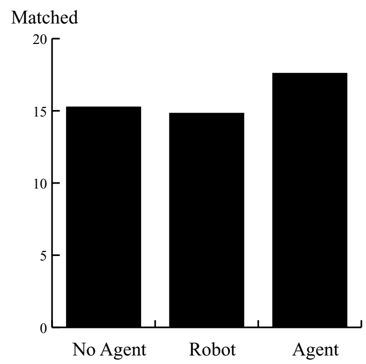

Figure 4: Robot vs. Agent

The answers to a post-test questionnaire showed that the subjective ``familiarity'' factor were different from the behavior of subjects. There was no significant difference between group B and C in familiarity, while the behaviors of those two groups were significantly different (Figure 4). This may imply that users were not aware of the agent's effect to their behavior. Also the familiarity factor for group E, the robot group, was much stronger than for group D (4.0 and 3.4, $p^*<0.05$), while subjects accepted agent's recommendation more (Figure 5).

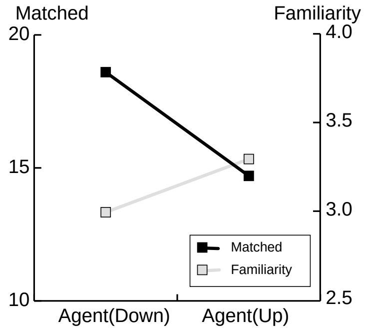

Figure 5: "Up" vs. "Down" strategies

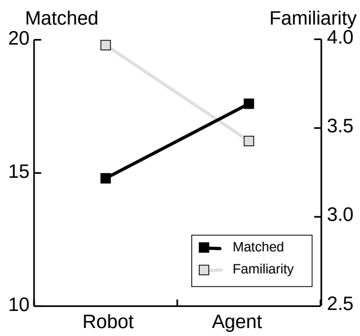

Figure 6: Robot vs. Agent

Discussion

In summary, these results of experiment 1 showed that an agent can affect a subject's decision and a down-type recommendation affects more than an up-type recommendation. These two strategies seem to correspond to two types of compliance-gaining-strategy in social psychology; door-in-the-face technique and foot-in-the-door technique [cite: social psychology]. In the foot-in-the-door technique, small request that is easy to accept is made first, then requests change to big one gradually. On the contrary, in the door-in-the-face technique, big request that is difficult to accept is made at first, then requests gradually change to small one. “Up-type” recommendation strategy corresponds to foot-in-the door technique. Each recommendation in “up-type” changes from color names that many people don’t select to one that many people tend to select. It is similar to the foot-in-the-door technique. “Down-type” recommendation strategy corresponds to door-in-the-face technique. When request is made by those who is not familiar to the people, foot-in-the-door technique works better than door-in-the-face technique, because people refuse first big request made by the stranger and the attitude tends to continue. In the experiments, the agent and the robot were stranger for the subjects at first. This is possibly the reason that “down-type” strategy worked better than “up-type” strategy in the experiments.

In the experiment 2, we found that 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. But the design rules for robot-user interaction may differ from those for agent-user interaction. This means that the experimental conditions were not good enough for robot to bring its capability into full play. There are two issues. One is position of the robot and another is eye contact.

The robot was sitting on the desk, while the agent was acting on the screen. The subjects were concentrating on the color square on the screen so the agent was in a better position to catch their attention. The agent seems to be in the same “world” as the color-square, while the robot is out side of the screen, same as the subjects. The biggest advantage of the robot is its embodiment, so it can refer to a real object and share the attention with a person. However, this advantage was not well utilized in the experimental set-up. To make the robot more effective, actual printed color paper should have been used to display color for the experiment 2 in this point of view. This makes consistent condition for the robot and agent. Both refer the color in the same world. Then we can compare two conditions: the agent and color square are in the screen and subject is in the real world, and the robot and color square are in the real world and subject is also in the real world. This enables us to measure the effect of embodiment of the robot.

Another issue is eye contact. It is important aspect of the attitude when social actor communicates with people. The agent in the two-dimensional screen can easily achieve the eye contact with the subject by illusion that subject feels like the agent is looking at him/her even if he/she is in the oblique position to the screen. But the robot is in the three-dimensional space, and subjects feel the robot is not making eye contact when its gaze is not to him/her. The gaze of the robot has to be controlled precisely to look at the subjects in order to feel them being looked at.

Based on these discussions, we have designed and developed new experimental set-up to solve these two issues. We employed desert survival problem as a task for subjects and developed gaze control system for the robot using image processing and created item card set for robot to refer. These improvements expected to 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

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

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

Shinozawa, K. Yamato, J., Naya, F., Kogure, K. 2001. Quantitative evaluation of effect of embodied conversational agents on user decision. *Proceedings of HCI International 2001*