How Appearance of Robotic Agents Affects How People Interpret the Agents' Attitudes

Takanori Komatsu
Future University-Hakodate
116-2 Kamedanakano
Hakodate 041-8655, Japan
+81-138-34-6220
komatsu@fun.ac.jp

Seiji Yamada
National Institute of Informatics
2-1-2 Hitotsubashi
Tokyo 101-8430, Japan
+81-3-4212-2562
seiji@nii.ac.jp

ABSTRACT

An experimental investigation of how the appearance of robotic agents affects interpretations people make of the agents' attitudes is described. We conducted a psychological experiment where participants were presented artificial sounds that can make people estimate specific agents' primitive attitudes from three kinds of agents, e.g., a Mindstorms robot, AIBO robot, and normal laptop PC. They were also asked to select the correct attitudes based on the sounds expressed by these three agents. The results showed that the participants had higher interpretation rates when a PC presented the sounds, while they had lower rates when Mindstorms and AIBO robots presented the sounds, even though the artificial sounds expressed by these agents were completely the same.

Categories and Subject Descriptors

J.4 [Computer Applications]: Social and Behavioral Sciences – *Psychology.*

General Terms

Experimentation and Human Factors

Keywords

Appearance of agents, agents' attitudes, subtle expressions, human-agent interaction

1. INTRODUCTION

Recently, one of the hottest topics in human-computer interaction or human-agent interaction studies is "what kind of appearance should robotic agents have in order to interact with people effectively." People are said to determine agents' behavior model based on the appearance of agents. Therefore, the agents' appearance significantly affects the interaction with these people [1]. For example, when people encounter a dog-like robot, they expect dog-like behaviors from this robot, and they would naturally speak to it using commands and other utterances intended for dogs, such as "sit", "lie down", and "fetch". However, they do not act this way toward a cat-like robot.

Several studies have focused on the effects of appearance of agents on interactions with people [2,3,5]. Keisler [2] conducted a psychological experiment where participants were asked to play a prisoner's dilemma game with virtual characters (human and dog) that appeared on a computer display. The results showed that participants who had some experience with owning dogs interacted significantly more effectively with the dog-like virtual agent, e.g., cooperating with this agent significantly. Goetz et al. [5] investigated the effects on people's impressions of different designs of the head for humanoid robots. The results showed that the participants answered that the robots with human-like heads are good at social tasks, while robots with machine-like heads are good at industrial tasks.

On the basis of the findings of these studies, we are currently focusing on the issue of "What kinds of appearance should robotic agents have in order to interact with people effectively? And which kind of information should these agents express to people?" As part of the first step to finding a comprehensive solution to this issue, we investigated how people interpret presented sounds to see if they could determine specific attitudes from agents that have different appearance.

2. AGENTS' ATTITUDES, APPEARANCE, AND EXPRESSED INFORMATION

Several approaches can be used to tackle with this issue. In this study, we experimentally investigated the effects of the basic psychological relationship between the appearance of agents and the information expressed on how people interpret the agents' attitudes.

We selected positive and negative attitudes corresponding to valence values as the primitive attitudes that robotic agents should express. Informing people of these two values is quite important if the agents are to interact effectively with people. We selected three artifacts for agents having different appearance: a Mindstorms robot [7], AIBO robot [8], and a normal laptop PC (Let's note W2, product of Panasonic Inc.). These artifacts correspond to making mechanical impressions of people, familial impressions, and non-agent-like impressions, respectively (Figure 2).

Concerning the effects of the relationship between the appearance of the agents and their expressed information and how they affect users' interpretations of the agents' attitudes, we hypothesized that *agents with a lifelike appearance (quite similar to people or pet animals) expressing true-to-life information (verbal information or animal-like behaviors) are actually more confusing to users and are not really effective for interacting with people: Instead, agents without a lifelike appearance expressing simple but intuitive information (e.g., subtle expressions* [6]*) are readily understood and are much more effective for interaction* (see Figure 1) [9]. If our hypothesis holds true, we would be able to better facilitate people's comprehension of the agents' primitive attitudes without having to develop dexterous and complex robotic or computer graphic systems, which would be quite expensive. Our hypothesis may eventually contribute to creating robotic agents more easily and to establishing deeper and more natural interactions with people.

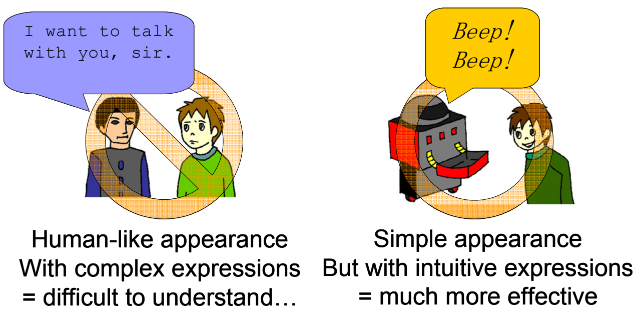

Figure 1. Comparison of human-like and simple robots and their impact on intuitive interactions.

3. PRELIMINARY EXPERIMENT

Before the actual experiment, we conducted a preliminary experiment to determine what kinds of subtle expressions are effective to evoke certain attitudes, positive or negative, from people. In this study, we focused on artificial sounds that acted as subtle expressions in previous studies [4]. We then investigated what kinds of sounds were interpreted as being positive or negative attitudes.

We prepared 44 different types of triangle wave sounds with four different durations and 11 different F0 (fundamental frequency) contours. Specifically, the four durations were 189, 418, 639, and 819 ms. The 11 F0 contours were set so that the transition range of F0 values between the onset and endpoint in the sound stimuli were 0, ±25, ±50, ±75, ±100, or ±125, and these were linearly downward or upward (Figure 2). All these 44 stimuli have the same F0 average of 131 Hz. And these sounds have a tone that sounds like a computer's beeping.

3.1 Participants

Ten Japanese university students (6 men and 4 women; 19 – 23 years old) participated.

3.2 Procedure

First, an experimenter gave the instruction "please determine the attitude of this laptop PC based on the sounds it makes." They were then asked to select one of the three attitudes "positive," "negative," or "undistinguishable" after presenting one of 44 prepared sounds. These three attitudes were described to participants as follows:

- **Positive**: the PC's internal state appears to be good.
- **Negative**: the PC's internal state appears to be bad.
- **Undistinguishable**: it is unclear whether the PC's internal state is positive or negative.

As part of the procedure, one randomly selected sound among 44 prepared sounds was presented to the participants. Afterwards, the participants were asked to select one of the three aforementioned attitudes. Each participant heard all 44 prepared sounds. The order of the sounds was counterbalanced for all 10 participants. Actually, these sounds were presented by a normal laptop PC (Let's note W2, product of Panasonic Inc.).

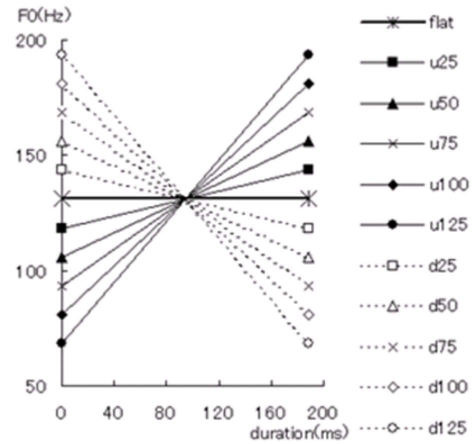

Figure 2. 11 different F0 contours (duration: 189 ms). For example, "189u25" indicates that duration was 189 ms, and F0 transition range was 25 Hz with upward slope (increasing intonation).

3.3 Results

The results of this preliminary experiment show that all 10 participants believed the PC had positive attitudes for four sounds, that is, 189 ms with an upward slope range of 125 Hz (189u125), 418u125, 639u125, and 868u125. Also, all 10 participants believed the PC had negative attitudes for four sounds, that is, 418 ms with a downward slope range of 25 Hz (418d25), 418d50, 418d75, and 639d50. Thus, sounds with increasing intonation regardless of the duration were interpreted as being positive attitudes, while sounds that have around 500 ms duration with slightly decreasing intonation were interpreted as being negative attitudes. These eight sounds were then selected for agents with different appearance in the next experiment.

4. EXPERIMENT

The purpose of this experiment was to investigate the effects of the agents' appearance on how participants interpreted the agents' attitudes. Specifically, the participants were presented the selected sounds used in the preliminary experiment by agents that have different appearances, the Mindstorms robot, AIBO robot, and the normal laptop PC (see Figure 3). They were then asked to select the correct attitudes (positive, negative, or undistinguishable) based on the expressed sounds made by these three agents.

4.1 Participants

Nine Japanese university students (8 men and 1 woman; 21 – 24 years old) participated. These participants were not familiar with robots or these toys and had not participated in the preliminary experiment.

4.2 Procedure

First, the participants were told that the concrete task of this experiment was to select one of the three attitudes (positive, negative, or undistinguishable) based on the sounds these agents made.

Figure 3. AIBO robot, Mindstorms robot, and the laptop PC (from left to right).

All participants experienced the following four conditions.

1. Eight sounds expressed by Mindstorms (**MS-sound condition**): the eight sounds came from an FM radio tuner placed on the Mindstorms. This radio tuner received the transmitted sounds from a sound expressing PC (Figure 4).
2. Eight sounds expressed by AIBO (**AIBO-sound condition**): the sounds were presented using AIBO's operating software "AIBO entertainment player" that was installed in an AIBO operating PC.
3. Eight of AIBO's prepared behaviors (**AIBO-motion condition**): The AIBO entertainment player has about 80 prepared behaviors, such as "good morning" or "delightful." We selected the following eight behaviors for expressing primitive attitudes to the participants. The positive behaviors were (four behaviors) "cheer1," "cheer3," "cheer4," and "cheer5." The negative behaviors were (four behaviors) "angry1," "angry2," "sad1," and "sad2." These behaviors were selected based on the verbal labels of these behaviors.
4. Eight sounds expressed by the laptop PC (**PC-sound condition**): The same conditions as those in the preliminary experiment were used. This laptop PC was remotely operated by a sound expressing PC.

First, participants experienced MS-sounds, AIBO–sounds, and AIBO-motion conditions in a random order. Then, they heard the PC-sound condition. In all four conditions, the eight sounds or behaviors were randomly presented to the participants. The participants were asked to select the correct attitudes from the three attitudes (positive, negative, or undistinguishable) after the agent expressed certain information. Therefore, this procedure was nearly the same as the one in the preliminary experiment. The order of the experimental conditions and the order of eight sounds or behaviors were counterbalanced for the participants.

4.3 Results

We calculated the interpretation rates, which indicated how many times the participants succeeded in correctly determining the agents' attitudes in all four experimental conditions. The results were that the participants has interpretation rates of 3.33 for eight experimental stimuli in the MS-sound condition, 2.89 in the AIBO-sound condition, 3.33 in the AIBO-motion condition and 6.44 in the PC-sound condition (Figure 5). The results of an ANOVA showed significant differences in these four experimental conditions ($F(3,24)=8.26$, $p<.01$(**)), and a multiple comparison using an LSD test showed significant differences between the PC-sound condition and the other three conditions (Mse=2.9421, 5% level).

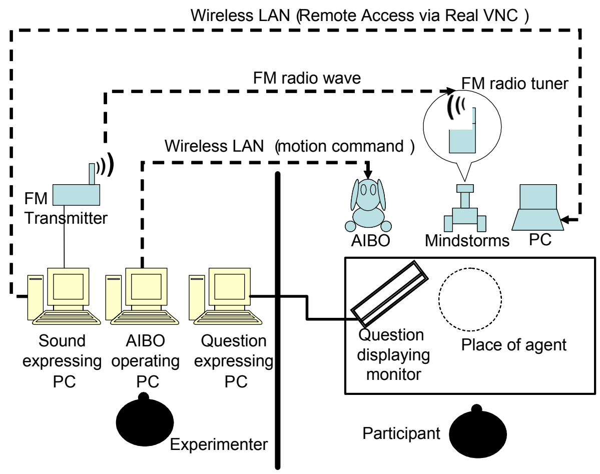

Figure 4. Experimental Setting.

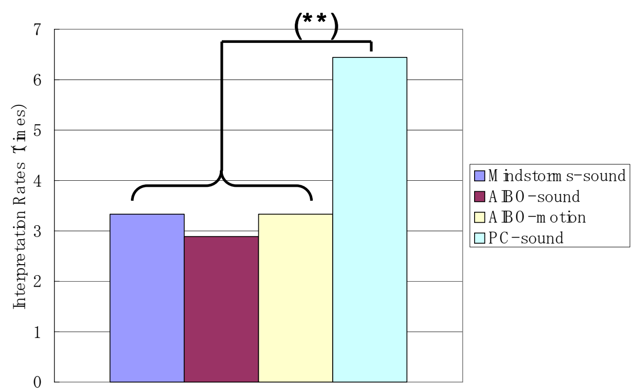

Figure 5. Participants' interpretation rates in four experimental conditions.

These results showed that the participants' interpretation rates for the same sound stimuli differed based on each agent's appearance. Simply stated, the participants showed higher interpretation rates in the PC-sound condition, while they showed significantly lower rates in the MS-sounds and AIBO-sounds conditions, even though the same sounds were presented to participants in these three conditions. Moreover, the interpretation rates observed in the AIBO-motion condition were also significantly lower than the PC-sound condition, and these interpretation rates were nearly the same as those of the MS-sound and AIBO-sound conditions, even though these motions were prepared for a commercial product.

5. DISCUSSION AND CONCLUSIONS

Eight artificial sounds selected in a preliminary experiment were presented to participants by a Mindstorms, AIBO, and PC. The participants' interpretation rates, indicating how many times the participants selected the correct attitudes, were then investigated. The results were that the interpretation rates when a PC expressed these sounds were significantly higher than the rates when a Mindstorms and AIBO expressed them. This difference revealed that the agents' different appearances affected people's interpretations of the agents' attitudes, even though these agents expressed information that was completely the same.

Let us consider why the PC-sound condition showed higher rates compared with the other conditions. One reason is that these eight sounds were selected in the preliminary experiment when these sounds were presented by the laptop PC, which was also used in the PC-sound condition. Thus, these sounds may have been effective only for informing people of primitive attitudes when they were presented by the laptop PC. This phenomenon may be rooted in the fact that the PC expressing beep-like sounds was very familiar for all participants.

In any case, these results indicate that robotic agents' different appearances may have affected the users' interpretation of the agents' attitudes. Was this actually why the interpretation rates in the PC-sound condition were higher than those in the other conditions? To answer this question, we need to conduct other experiments to consider the relationship between appearance and expressed information.

The results of this experiment revealed an interesting phenomenon where the interpretation rates in the AIBO-motion condition were lower than those in the PC-sound condition, just as they were for the MS-sound and AIBO-sound conditions. This indicates that the AIBO's prepared behaviors were not really efficient in informing the participants of its primitive attitudes, positive or negative. This result would support our hypothesis described in the Introduction: "agents with a lifelike appearance (quite similar to people or pet animals) expressing true-to-life information (verbal information or animal-like behaviors) are actually more confusing to users and are not really effective for interacting with people: Instead, agents without such a lifelike appearance expressing simple but intuitive information (e.g., subtle expressions [6]) are readily understood and are much more effective for interaction (see Figure 1) [9]." Of course, the behaviors of AIBO were not designed for informing people of the primitive attitudes we estimated. However, our results suggest a design policy is needed to inform people of certain attitudes effectively.

We intend to pursue a series of follow-up studies based on our results in this study and subsequently establish a design policy for the most appropriate information based on the agents' appearance. For example, what information is appropriate for the Mindstorms robot to inform people of its primitive attitudes? Are Starwars' R2D2 like behaviors appropriate? We expect that these follow-up studies will contribute to establishing a design policy that can clarify an effective coupling between the appearance of a robotic agent and the information it expresses so that interactive agents can be created readily and easily.

6. REFERENCES

[1] S. Yamada and T. Yamaguchi: Training AIBO like a Dog, the 13th International Workshop on Robot and Human Interactive Communication, pp. 431-436 (2004).

[2] S. Kiesler, L. Sproull, & K. Waters. A Prisoner's Dilemma Experiment on Cooperation with People and Human-like Computers. Journal of Personality and Social Psychology, 79(1), 47-65 (1995).

[3] S. Kiesler & L. Sproull. Social Responses to "Social" Computers. In B. Friedman, Human values and the design of technology, CLSI Publications (1997).

[4] T. Komatsu. Audio Subtle Expressions Affecting User's Perceptions, the 9th International Conference on Intelligent User Interfaces, p. 306-308 (2006).

[5] J. Goetz, S. Kiesler, & A. Powers: Matching Robot Appearance and Behavior to Tasks to Improve Human-Robot Cooperation, 12th IEEE Workshop on Robot and Human Interactive Communication (2003).

[6] K. Liu & W. R. Picard. Subtle expressivity in a Robotic Computer, CHI2003 Workshop on Subtle Expressivity for Characters and Robots (2003).

[7] MindStorms: http://mindstorms.lego.com/

[8] AIBO: http://www.jp.aibo.com

[9] S. Yamada & T. Komatsu. Designing Simple and Effective Expressions of Robot's Primitive Minds to a Human, 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 2614-2619 (2006).