Effects of Polite Behaviors Expressed by Robots: A Case Study in Japan

Tatsuya Nomura[1,2]
[1]Department of Media Informatics
Ryukoku University
Otsu, Shiga 520-2194, Japan
nomura@rins.ryukoku.ac.jp

Kazuma Saeki[1]
[2]ATR Intelligent Robotics and
Communication Laboratories
Kyoto, 619-0288, Japan

Abstract— **To explore the design of politeness in embodied conversational agents and social robots from the perspective of user studies, a case study based on a psychological experiment was conducted to straightforwardly investigate effects of polite behaviors by robots in Japan, using a small-sized humanoid robot. Results of the experiment suggested that differences of polite behaviors expressed by the robot affected participants' impression of and behaviors toward the robot.**

Keywords- Robots; Politenes; Psychological Experiement

I. INTRODUCTION

Polite behaviors are an important factor in human-human communication. It leads researchers on embodied conversational agents and social robots to exploration of polite behavior design for the agents and robots to engage with humans [1][2][3]. In particular, it is important when robots build initiation of engagement with humans [4]. Moreover, non-verbal polite behaviors should be focused on from multi-modality in human-agent and human-robot interaction, as well as verbal expression of politeness.

However, there are few studies that straightforwardly validated effects of polite behaviors by agents and robots from the perspective of user studies. Although some existing works included politeness as one of the user evaluation criteria for specific behavioral factors of agents and robots (e.g., [5]), these works did not focus on polite behaviors of the agents and robots. Although Rehm and André [3] implemented and validated an conversational agent's polite behaviors in experiments, their study did not have sufficient number of samples. Moreover, politeness depends on cultures [6]. In order to consider design of agents' and robots' polite behaviors suitable for individuals, it needs to explore its influential factors based on user evaluation in several countries.

As an attempt for the above aim, a case study based on a psychological experiment was conducted in Japan, by using a small-sized humanoid robot. This experiment had a simple design based on the research question: whether differences on polite behaviors in the robot affect human impression of and behaviors toward the robot. The paper reports results of the experiment to provide with basic information for design of polite behaviors in agents and robots.

II. METHOD

A. Subjects

The experiment was conducted from October to December, 2008. A total of forty two persons participated to the experiment. They were university students in the western area of Japan, and recruited with one thousand yen. Table 1 shows the demographic characteristics of the participants.

TABLE I. DEMOGRAPHIC CHARACTERISTICS OF SUBJECTS IN THE EXPERIMENT

		Male	Female
	Education		
	Social sciences	13	15
	Natural sciences or engineering	7	5
	Unknown	2	0
	Total	22	20
Age	**Mean**	19.7	19.9
	SD	1.2	1.3

B. Robots Used in the Experiment

The small-sized humanoid robot used in the experiment was "Robovie-X" shown in the left figure of Fig. 1, which has been developed by Vstone Corporation. This robot stands 34.3 cm tall and weighs about 1.3 kg. The robot has a total of 17 Degrees of Freedom (DOFs) at its feet, arms, and head. This large number of DOF allows it to execute various gestures such as walking, bowing, and a handstand. Moreover, this robot has a function of utterance based on audio data recorded in advance such as Windows WAV files, which is limited to 300 KB.

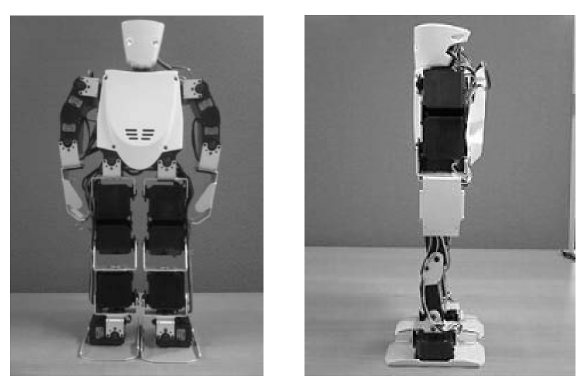
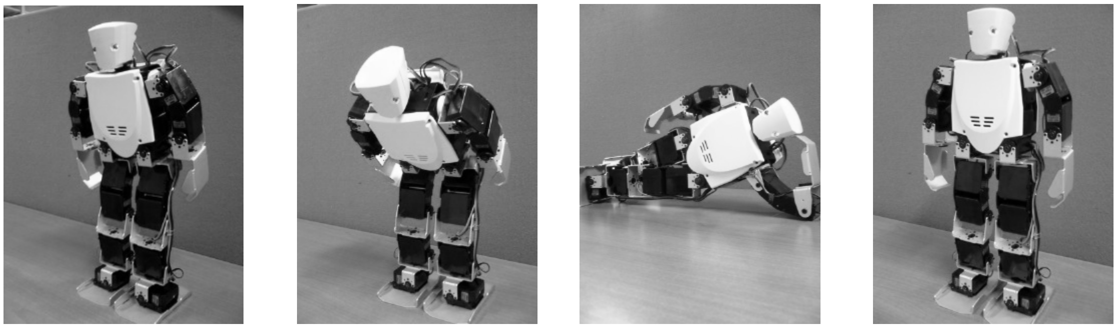


Figure 1. Robovie-X (left figure: front view, right figure: side view)



Type 1: Bowing Type 2: Deep Bowing Type 3: Lying Type 4: Just standing

Figure 2. Motion Types of the Robot Different on the Level of Politeness

C. *Task for Subjects and Voice Instructions from the Robots*

The task to be requested for the subjects in the experiment was manipulation of objects on a desk. This task is similar with the one conducted in the experiment on influences of robot physical appearances into human perception [7]. In the experiment, it was instructed by the robots with voice.

Voice data consisting of Japanese sentences was synthesized from text data by using “Easy Speech,” “Text-to-Speech Engine Japanese version,” “Sound Engine Free” (free software), Microsoft SPAI 4.0, and L & H TTS 3000. The quality of the voice was artificial and neutral independent on gender. Then, it was played by the robot as instructions from the robot to the subjects. The instructions were common in all the experiment conditions. They were presented with polite expression specific in the Japanese as follows:

“Hello, I am Robovie-X. Please fill the cup with tea in the plastic bottle in front of you, and then drink it. Before drinking, please separate the garbage produced from the empty plastic bottle. Thank you for your cooperation.”

D. *Polite Motions of the Robot*

The experiment focused only on politeness in motions of the robot to control other factors. As mentioned in the previous section, the speech data from the robot was common in all the motion conditions to control the level of politeness in the linguistic contents.

Barraquand and Crowley [1] described as follows; Social common sense refers to the shared rules for polite, social interaction that implicitly rule behavior within a social group. Based on this description, the experimenters prepared four types of motion different on the level of politeness, referring to the common sense assumed to be shared within the Japanese community.

In the first type of motion, the robot inclined its upper body forward at fifteen degree angle just after the utterances of “Hello” and “Thank you” in the instruction to subjects. In the second type of motion, the robot performed the same motion at forty five degree angle. These types of motion correspond to blowing motions different on the level of politeness in Japan. In the third type of motion, the robot stayed lying during utterance of the instructions. In the fourth type of motion, the robot kept standing without other motions during utterance of the instructions. These types of motion correspond to impolite attitudes, and were prepared to compare with the first and second types of motion on the level of politeness.

These motions were produced by the accessory software “RobovieMaker2” and installed in advance. Figure 2 shows samples of these motions.

E. *Procedures*

The experiment adopted a 2 x 4 between-subjects design of gender and types of the robot motion. Each session was conducted based on the following procedures:

1. Each subject was explained about the experiment and signed the consent form about dealing with data including video-recording. In this stage, the experimenters only indicated that the task in the experiment was interaction with a robot and they planed to video-record the scene in the experiment.

2. The subject was led to an experiment room, in which the robot, an empty cup, and a plastic bottle filled with tea were put on a desk, as shown in Figure 3. The experimenters instructed him/her to sit on the chair in front of the desk and wait in the room for a while, and left the room.

3. Just after the subject was left alone in the room, the robot started the instruction for him/her to drink tea and separate the garbage into three plastic bags assigned to the front of another desk, as mentioned in the previous section. It was remotely controlled by the experimenters out of the room.

4. When the subject finished the instructed task or two minutes passed without performing the task, the experimenters entered the room again, and indicated that the session finished. Then, the experimenters conducted debriefing about the actual aim of the experiment and the fact that the session was video-recorded by a camera concealed from the subject.

5. Then, the subject responded a questionnaire for measuring his/her impression of the robot. Finally, the experimenters interviewed with the subject about the robot and the experiment.

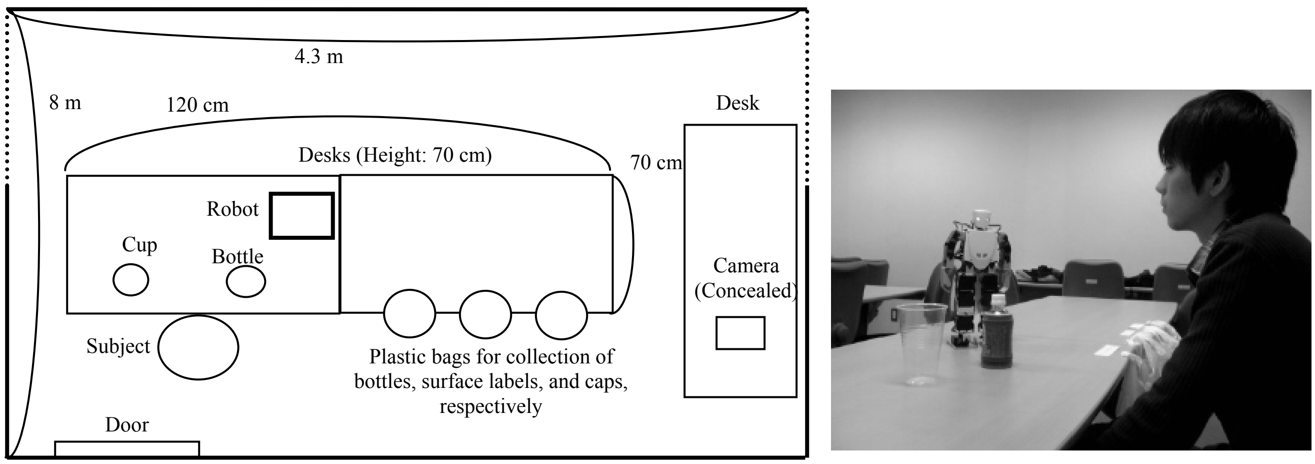


Figure 3. Overview of the room where the experiment was executed (a view from above) and a Scene of the Session

F. *Measures*

The scenes of the experiment were recorded with a digital video camera to extract the subjects' behaviors toward the robots.

The questionnaire for measuring the subjects' impressions of the robots consists of twenty pairs of adjectives shown in Table 2. The subjects were asked to respond to each pair of adjectives to present degrees to which they felt the impression represented by the pair of adjectives for the robots they experienced. These adjectives were selected from the ones used for measurement of subjects' impression in an experiment of interaction with a humanoid robot [8]. Moreover, the pair of "polite"--"impolite" was added. Each questionnaire item had a score for rating with seven intervals (1-7). On the questionnaire, it was randomized at each item which side the positive or negative adjective appeared at.

TABLE II. PAIRS OF ADJECTIVES FOR MEASURING SUBJECTS' IMPRESSIONS OF THE ROBOT

Positive		Negative
Polite	---	Impolite
Mild	---	Terrible
Fine	---	Ill
Familiar	---	Unfamiliar
Safe	---	Dangerous
Warm	---	Cold
Pretty	---	Hateful
Chatty	---	Formal
Comprehensible	---	Not comprehensible
Approachable	---	Unapproachable
Light	---	Dark
Funny	---	Boring
Pleasant	---	Unpleasant
Favorite	---	Disfavorite
Interesting	---	Tedious
Fast	---	Slow
Aggressive	---	Negative
Showy	---	Plain
Cheerful	---	Gloomy
Clever	---	Foolish

III. RESULTS

To investigate the effects of the robot's motion type into the subjects' impressions of and behaviors toward the robot, the following analyses were performed based on the video data and results of the questionnaire.

A. *Numbers of Subjects Who Performed the Instructed Task*

Table 3 shows the numbers of subjects assigned to the conditions in the experiment and results on whether they performed the task instructed by the robot. About 79% of the subjects performed the task of separating the garbage as requested by the robot. χ^2-tests found no relationships between the performance of the task and motion type, or gender (motion type: $\chi^2(3) = .586$, n.s., gender: $\chi^2(1) = .046$, n.s.)

B. *Impression of the Robot*

For each item of adjectives pair, the score of the seven-graded answer was coded from 1 to 7 so that higher score corresponded to the positive adjective of the pair. Then, exploratory factor analysis with Maximum-likelihood method and Promax rotation was performed to classify these items and extract subscales for measuring the subjects' impressions of the robots. As a result, five factors having eigen values more than 1 were extracted. Then, item analysis

TABLE III. NUMBERS OF SUBJECTS ASSIGNED TO THE CONDITIONS IN THE EXPERIMENT AND RESULTS ON WHETHER THEY PERFORMED THE TASK INSTRUCTED BY THE ROBOT

	Male		**Female**		**Total**
	P	**N**	**P**	**N**	
Bowing	3	2	4	1	10
Deep Bowing	6	0	3	2	11
Lying	4	2	5	0	11
Just standing	4	1	4	1	10

(P: Subjects who performed the task, N: Subjects who did not performed the task)

TABLE IV. RESULTS OF FACTOR ANALYSIS AND ITEM ANALYSIS FOR IMPRESSION ITEMS

Adjective (positive)	Factors I	II	III	IV	V	h^2	Note
Favorite	**1.137**	-.073	-.002	.031	-.206	.999	
Familiar	**.634**	-.120	.011	.113	.031	.387	
Pretty	**.575**	-.250	.359	.104	.129	.563	
Funny	**.519**	.254	.007	-.147	.180	.633	
Approachable	**.456**	.111	.074	-.165	.316	.591	
Clever	-.136	**.757**	.083	-.194	.037	.491	
Showy	-.156	**.641**	.164	.176	-.291	.379	
Pleasant	.480	**.607**	-.085	-.068	-.042	.830	
Light	.069	**.562**	-.074	.223	-.067	.437	
Comprehensible	-.169	**.475**	.292	.012	-.026	.279	Removed by item analysis
Warm	.246	**.437**	-.026	.194	.149	.599	
Polite	.083	.150	**.800**	-.203	-.118	.700	
Fine	.124	.252	**.546**	-.129	.113	.616	
Safe	-.108	-.003	**.546**	.313	.307	.534	Removed by item analysis
Aggressive	-.098	-.015	-.077	**.634**	-.082	.395	Removed by item analysis
Fast	.068	-.040	.374	**.566**	-.091	.435	
Cheerful	.123	.346	-.072	**.529**	.279	.813	
Chatty	.167	.060	-.120	**.451**	-.163	.266	
Mild	.122	-.066	.063	-.258	**.937**	.999	Removed by item analysis
Interesting	.388	-.293	-.049	.094	**.417**	.358	Removed by item analysis

Subscale	#. Item	Chronbach's α
I: Familiarity	5	.841
II: Extroversion	5	.790
III: Politeness	2	.763
IV: Activeness	3	.628

Correlations	II: Extroversion	III: Politeness	IV: Activeness
I: Familiarity	.596**	.462**	.381*
II: Extroversion		.376*	.503**
III: Politeness			.116

(*$p < .05$, **$p < .01$)

using Chronbach's α-coefficients and I-T correlations was performed for each factor to select items in the corresponding subscale. Table 4 shows the results of these analyses.

The first factor consisted of five items and the item analysis found no item to be removed. Based on the contents of these five items, the corresponding subscale was interpreted as "familiarity". The second factor consisted of six items and item analysis found one item to be removed. Based on the contents of these five items, the corresponding subscale was interpreted as "extroversion". The third factor consisted of three items and item analysis found one item to be removed. Based on the contents of the two items, the corresponding subscale was interpreted as "politeness". The fourth factor consisted of four items and the item analysis found one item to be removed. Based on the contents of these three items, the corresponding subscale was interpreted as "activeness". The fifth factor was removed from the analysis since it originally consisted of only two items and had low internal consistency (Chronbach's α = .593).

The score of each impression subscale was calculated as the sum of the scores of the corresponding items. Thus, the maximum and minimum scores are 35 and 5 for "familiarity" subscale, 35 and 5 for "extroversion" subscale, 14 and 2 for "politeness" subscale, and 21 and 3 for "activeness" subscale, respectively. Then, to compare the subjects' impressions of the robots between the conditions, two-way ANOVAs with robot motion x gender were performed for the scores of the four subscales. Figure 4 shows the means and standard deviations of these subscale scores and results of the ANOVAs.

As a result, the main effects of motion type were at statistically significant trends for the scores of "extroversion" and "politeness". Post-hoc analyses with Bonferroni's method revealed that the "extroversion" subscale scores of the subjects who experienced "deep bowing" robot motion were higher than the scores of those who experienced "just standing" robot motion at a statistically significant level. Moreover, it was found that the "politeness" subscale scores of the subjects who experienced "lying" robot motion were lower than the scores of those who experienced the other robot motions at statistically significant levels. For the scores of "activeness", the interaction effect was at a statistically significant trend. However, post-hoc analysis found no difference between the subject groups, except for that between the male and female subjects who experienced "just standing" robot motion, at a statistically significant trend.

Moreover, two-way ANOVAs with task performance/no-performance and gender were performed for these subscale scores, in order to investigate relationships between the subjects' impressions of the robot and their performance behaviors of the instructed task. Although these analyses should originally be included in three ways with motion type x gender x task performance/no-performance, some cells in

	Motion type	Gender	Interaction	Post-hoc
Familiarity	1.657	1.165	2.164	
Extroversion	4.031*	1.109	1.412	Deep Bowing > Just Standing*
Politeness	9.696***	.127	.339	Lying < Bowing**, Lying < Deep Bowing***, Lying < Just Standing*
Activeness	.746	.416	$2.364^{\dagger}$	Male > Female in "Just Standing" †

($^{\dagger}p < .1$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$)

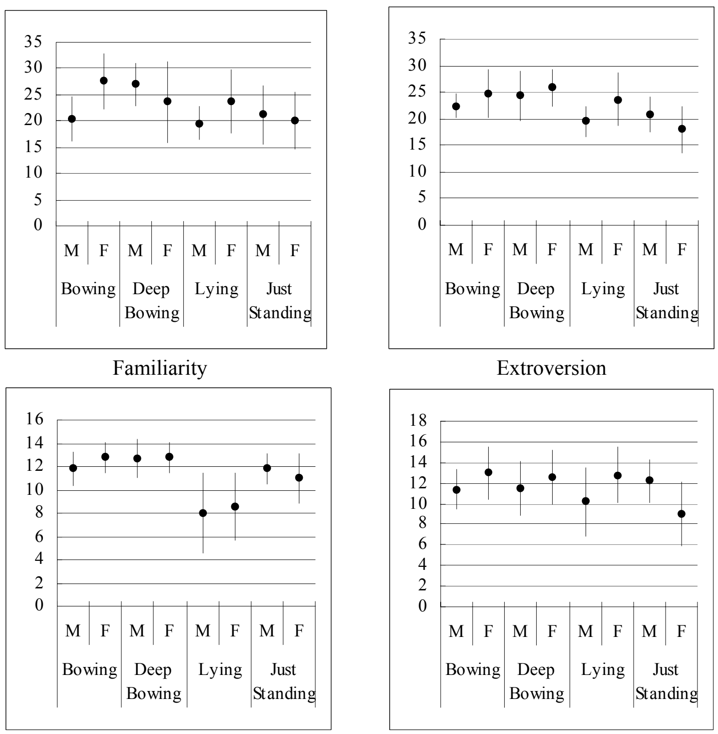


Figure 4. Results of ANOVAs with Motion Type and Gender and Means and Standard Deviations of Impression Subscale Scores

the three ways were empty as shown Table. 1, and thus they were performed independently.

The results revealed that the "familiarity" subscale scores of the subjects who performed the task were higher than those who did not perform at a statistically significant level. Moreover, an interaction effect at a statistically significant trend level was found for the "politeness" subscale scores. A post-hoc analysis found the scores of the male who performed the task were higher than those of the male subjects who did not perform at a statistically significant level, and the scores of the female subjects who did not perform the task were higher than those of the male subjects who did not perform at a statistically significant trend level. There were no statistically significant differences on the other scores. Figure 6 shows the results.

C. *Time for Reply and Correlations with Impression*

As an index of the subjects' behaviors toward the robot, we measured time spent until they replied to the instruction from the robot. Some subjects started the task of separating the garbage before the robot finished the utterance of the instruction (about twenty seconds). Thus, the time was defined as time that the subjects spent until they started the task since the robot had started the instruction utterance. The subjects who did not perform the task were removed from the analysis.

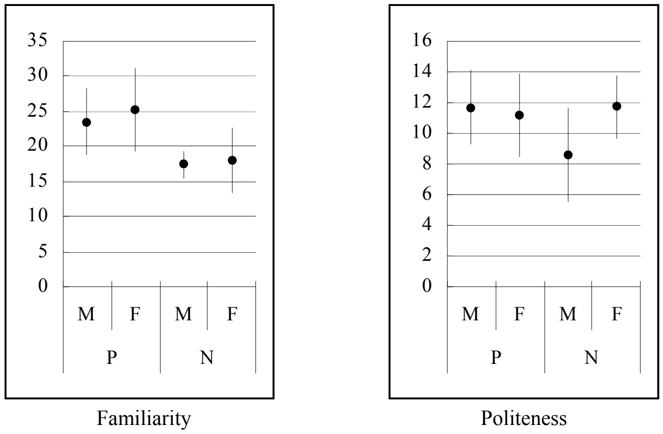
Figure 6 shows the means and standard deviations of time for reply to the robot. ANOVA with motion type x gender found no main or interaction effect (motion type: $F(3,25) = 1.180$, gender: $F(1,25) = .544$, interaction: $F(3,25) = .709$).

To investigate relationships with impression of the robot, correlation coefficients between the time spent for reply to the robot and impression subscale scores were calculated. Some existing studies suggested gender differences on relationships between psychological states, contexts, and behaviors about robots (e.g., [9][10]). Thus, these coefficients were calculated in each gender. Table 5 shows these coefficients and results of tests on equality of the correlation coefficients between the male and female samples.

The results revealed a moderate level of negative correlation between the time for reply to and "familiarity" impression of the robot regardless of gender. Moreover, it was found that there was a strong level of negative correlation between the time for reply to and "extroversion"

	Task Performance	Gender	Interaction	Post-hoc
Familiarity	12.070**	.363	.773	
Extroversion	1.961	.942	.186	
Politeness	1.722	1.843	3.581†	P > N in Male*, Male < Female in N †
Activeness	2.173	.073	1.912	

(†p < .1, *p < .05, **p < .01)



(P: Subjects who performed the task, N: Subjects who did not performed the task)

Figure 5. Results of ANOVAs with Task Performance/No-Performance and Gender and Means and Standard Deviations of "Familiarity" and "Politeness" Impression Subscale Scores

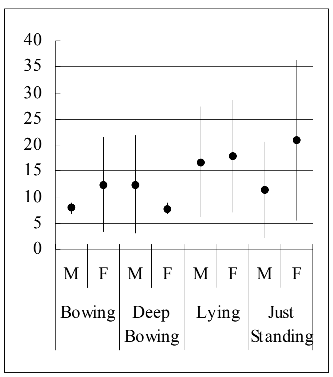


Figure 6. Means and Standard Deviations of Time for Reply to the Robot

TABLE V. CORRELATION COEFFICIENTS BETWEEN TIME FOR REPLY TO THE ROBOT AND IMPRESSION SUBSCALE SCORES, AND RESULTS OF TESTS ON EQUALITY BETWEEN MALE AND FEMALE SAMPLES

	Familiarity	Extroversion	Politeness	Activeness
Complete (N = 33)	-.319†	-.358*	-.266	-.069
Male (N = 17)	-.372	.088	-.377	.201
Female (N = 16)	-.340	-.649**	-.166	-.379
Z_0	.095	2.237*	.592	1.565†

(†p < .1, *p < .05, **p < .01)

impression of the robot in the female samples, although no correlation between them in the male samples. Furthermore, there was a difference at a statistically significant trend level between the male and female samples on the correlation between the time for reply to and "activeness" impression. Although these correlation coefficients were not at statistically significant levels, there was a moderate level of negative correlation in the female samples.

IV. DISCUSSION

The research is a case study and the results are limited to the Japanese samples experiencing a small-sized humanoid robot and specific polite behaviors. Thus, the generality of the findings is limited. However, the results have several implications on effects of polite behaviors by robots.

The results of the experiment did not show the direct effect of the robot's polite motions on human behaviors of task performance. However, they suggested their effects on human impressions of the robot, and some relationships between the impressions and behaviors toward the robots. It was found that the subjects who experienced "deep bowing" motion of the robot felt it more extrovert than those who experienced "just standing" motion, and the female subjects more strongly feeling the robot extrovert replied for the task instruction from the robot faster. The subjects who performed the task instructed by the robot felt the robot more familiar than those who did not perform, and those more strongly feeling the robot familiar replied for the task instruction faster. Moreover, the subjects who experienced "lying" motion of the robot felt the robot less polite than those who experienced the other motions, and the male subjects who did not performed the task felt the robot less polite than the male subjects who performed the task and the female subjects who did not perform the task.

The results of the experiment also suggested gender differences on relationships between polite behaviors of robots and human impressions of and behaviors toward robots. The extrovert impression of the robot correlated with the time for reply to the robot only in the female samples. Moreover, the difference on polite impression of the robot between task performance/no-performance appeared only in the male samples.

The results of the case study have some important implications. Even small differences on polite behaviors by a specific type of robot have a possibility of influencing human impressions of and behaviors toward robots. Moreover, even in a specific population such as Japanese university students, the effects may be dependent on personal traits such as gender. In other words, robotics designers should sufficiently take into account relationships between polite behaviors and user personal traits including age, gender, and cultures, or consider autonomous adaptation mechanisms of robot polite behaviors such as proposed by Barraquand and Crowley [1], including models of situations and contexts.

As future problems, the case study did not consider contexts and interaction styles of tasks that the robot and human co-worked, other subject demographics such as education and culture, or other characteristics of robots such as size and appearances. We are going to extend the experimental design to investigate relationships between robot polite behaviors and user characteristics, including age, gender, and cultural differences on politeness.

ACKNOWLEDGMENT

The research was supported in part by "High-Tech Research Center" project for private universities: matching fund subsidy from MEXT (Ministry of Education, Culture, Sports, Science and Technology), 2002–2006, and Ministry of Internal Affairs and Communications.

REFERENCES

[1] R. Barraquand and J. L. Crowley, "Learning Polite Behavior with Situation Models," in Proc. 3th ACM/IEEE International Conference on Human-Robot Interaction, pp.209-216, 2008.

[2] B. De Carolis and G. Cozzolongo, "Planning the Behavior of Social Robot Acting as a Majordomo in Public Environments," in AI*IA 2007, R. Basili and M. T. Pazienza, eds., Spinger, pp.805-812, 2007.

[3] M. Rehm and E. André, "More Than Just a Friendly Phrase: Multimodal Aspects of Polite Behavior in Agents," in Conversational Informatics: An Engineering Approach, T. Nishida, ed., Wiley, pp.69-84, 2007

[4] C. Sidner and C. Lee, "The Initiation of Engagement by a Humanoid Robot," in AAAI Spring Symposium on Multidisciplinary Collaboration for Socially Assistive Robotics, 2007.

[5] M. Lohse, M. Hanheide, B. Wrede, M. L. Walters, K. L. Koay, D. S. Syrdal, A. Green, H. Hüttenrauch, K. Dautenhahn, G. Sagerer, and K. Severinson-Eklundh, "Evaluating Extrovert and Introvert Behaviour of a Domestic robot: A Video Study," in Proc. 17th IEEE International Symposium on Robot and Human Interactive Communication, pp.488-493, 2008.

[6] Y. Taki, "Culture and Politeness: Differences of Apology Strategies of the British and Japanese People: Comparison of Young and Older Subjects," Studies in Language and Literature, vol.22, no.2, pp.27-63 2003.

[7] C. Kidd and C. Breazeal, "Effect of a robot on user perceptions," in Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 3559–3564, 2004.

[8] T. Kanda, H. Ishiguro, and T. Ishida, "Psychological Evaluation on Interactions between People and Robot," Journal of the Robotics Society of Japan, vol.19, no.3, pp.362-371, 2001. (in Japanese).

[9] B. Mutlu, S. Osman, J. Forlizzi, J. Hodgins, and S. Kiesler, "Task structure and user attributes as elements of human–robot interaction design," in Proc. 15th IEEE International Symposium on Robot and Human Interactive Communication., pp. 74-79, 2006.

[10] T. Nomura, T. Kanda, T. Suzuki, and K. Kato, "Prediction of Human Behavior in Human–Robot Interaction Using Psychological Scales for Anxiety and Negative Attitudes Toward Robots," IEEE Transactions on Robotics, vol.24, no.2, pp.442-451, 2008.