Similarities and Differences in Users' Interaction with a Humanoid and a Pet Robot

Anja Austermann *1, Seiji Yamada *1,2
*1 The Graduate University for Advanced Studies (SOKENDAI)
*2National Institute of Informatics
Tokyo, Japan
{anja, seiji}@nii.ac.jp

Kotaro Funakoshi, Mikio Nakano
Honda Research Institute Japan
Wako, Japan
{funakoshi, nakano}@jp.honda-ri.com

***Abstract*—In this paper, we compare user behavior towards the humanoid robot ASIMO and the dog-shaped robot AIBO in a simple task, in which the users has to teach commands and feedback to the robot.** *(Abstract)*

Keywords: Human-Robot Interaction; Humanoid; User Study

I. INTRODUCTION

Different factors influence how people interact with robots. Studies such as [2][3] showed that the appearance of the robot plays an important role for the behavior of users. As a part of our work on learning commands and feedback for human-robot interaction, we conducted a user study on how participants give commands and feedback to the robots AIBO and ASIMO. AIBO is a dog-shaped robot, made by Sony, which has roughly the size of a cat or small dog. ASIMO is a 1,30m tall humanoid robot created by Honda.

II. OUTLINE OF THE STUDY

The goal of our study is to find differences and similarities in user behavior when the participants give commands and feedback to ASIMO or AIBO. The users interacted with either ASIMO or AIBO and instructed the robot perform different household tasks like bringing a coffee, switching on the light or the TV, tidying up objects etc. and gave feedback to the robot for correct or incorrect performance. In order to avoid time-consuming and error-prone task execution and because of the different physical capabilities of the two different robots, we decided to use "virtual training tasks". The tasks are shown on a screen, and the robot acts in front of the screen using gestures and speech. On the screen, the robot's actions are visualized with a hand or paw icon, so that the user can easily understand the connection between the robot's motions and the changes, happening in the scene. Details on the training tasks can be found in [1]. While the robots differed in shape and size, we kept all other parameters as similar as possible, using the same synthesized speech utterances, similar gestures, same simulated learning rate etc.

Figure 1. Experimental Setting

We have conducted experiments with 16 participants aged from 22 to 52. Ten participants interacted with ASIMO (3 males, 7 females) six with AIBO (2 males, 3 females) for roughly 45 minutes. The language used in the experiments was Japanese. All participants were employees of the Honda Research Institute Japan. Fig. 1 shows the experimental setting.

III. RESULTS

In our user study, we obtained two different kinds of results: We asked the participants to answer a questionnaire about their subjective impression of the interaction and we annotated the data, which was recorded during the interaction, to find objective similarities and differences in the participants' behavior. We used the T-Test to determine the statistical significance of the observed differences.

A. Questionnaire results

From the results of the questionnaire, we can see a slight tendency towards better ratings for the interaction with AIBO. However, none of the differences is statistically significant.

B. User behavior

We analyzed different aspects of the users' commands and feedback that we assumed to be related to the perceived intelligence and human-likeness of the robot. We compared the speaking speed (in seconds per word) and the number of words per command/feedback, as we estimated that people talk slower

TABLE I. USERS' RATINGS OF THE INTERACTION

Question *(5: fully agree – 1: do not agree)*	**ASIMO** *mean (stdev)*	**AIBO** *mean (stdev)*
I enjoyed teaching the robot through the given task	3.5 (0.8)	4 (0.8)
The robot understood my feedback	3.6 (0.9)	4.3 (1.1)
The robot learned through my feedback	3.2 (1.3)	4.3 (0.5)
The robot adapted to my way of teaching	3.2 (1.1)	3.8 (1.3)
I was able to instruct the robot in a natural way	3.6 (1.1)	3.5 (1.5)
The robot took too much time to learn	3.6 (1.4)	2.7 (0.9)
The robot is intelligent	2.7 (1.3)	2.8 (1.5)
The robot behaves autonomously	2.7 (1.4)	2.8 (0.9)
The robot behaves cooperatively	3.7 (0.8)	3.3 (0.7)

978-1-4244-4893-7/10/$25.00 © 2010 IEEE

and in simpler sentences, when they consider the robot less intelligent. However, we found, that the length of commands was almost the same for both robots. An average command for ASIMO was 3.75 (sd=0.42) words long, while an average command for AIBO was 3.72 (sd=0.71) words long. The speaking speed was also similar for AIBO with 0.45 (sd=0.09) seconds per word, and ASIMO with 0.42 (sd=0.07) seconds per word. This is in line with the participants' subjective evaluation of the robots' intelligence as shown in Table 1.

1) Multimodality

During the interaction with both robots, we did not observe pointing gestures from any of the users. A possible explanation is that all objects were very easy to distinguish verbally, so that pointing gestures would have been redundant. We observed touch-based rewards for only one out of ten participants for ASIMO but for five out of six participants, who interacted with AIBO. As touch is frequently used with real dogs, we assume that users considered touch to be appropriate for giving feedback to AIBO because of its dog-like appearance.

2) Verbal commands

We analyzed, how many commands had explanations or polite expressions and how many commands were put as a question. We assumed that users might be more polite, explain more and use more questions when talking to a humanoid robot, while they rather give plain commands to a dog-like robot. We considered commands that contain words like "…kudasai", "…kureru?", "…moraeru?" etc., which are similar to the English word "please" as polite commands. We also analyzed, how many commands were given implicitly, e.g. saying "it is too dark here" when the user wants the robot to switch the light on, and in how many commands some expected parameters were left out like in "put away the toy car" instead of "put *the toy car* into *the box*", because we assumed, that this kind of verbal behavior might be related to the perceived intelligence of the robot. The results can be found in Table 2.

TABLE II. COMMAND TYPES

Type	ASIMO	AIBO
Plain commands	75.01 (14.00)	60.83 (41.04)
Polite commands	9.86 (10.88)	26.23 (41.99)
Questions in commands	10.23 (3.51)	8.34 (6.73)
Implicit commands	3.40 (4.82)	4.10 (7.23)
Parameters left out	6.78 (2.25)	4.13 (4.77)
Explanations in commands	1.81 (3.90)	0.95 (2.32)

All values in percent, values in brackets show the standard deviations

The values do not add up to 100% because not all types of commands are mutually exclusive. While we observed quite different utterances for different users, the differences seemed to be rather caused by personal preferences, than by the appearance of the robots. This assumption is supported by the high standard deviations between users. None of the observed differences was statistically significant.

3) Verbal positive and negative feedback

We distinguished three different types of feedback: Personal rewards like "Thank you", feedback which directly comments on the performance of the robot, like "Well done." or "That was wrong." and explanations used as rewards like "That is not a toy car, it is a ball." or "That is a toy car".

TABLE III. REWARD TYPES

Type	ASIMO	AIBO
Personal	52.78 (17.99)	24.83 (27.41)
Performance evaluation	38.39 (18.28)	70.02 (28.16)
Explanations	11.10 (14.29)	3.56 (3.90)

All values in percent, values in brackets show the standard deviations

We found statistically significant differences (5%-level) for the usage of personal rewards and rewards, which comment on the robots performance. While the users usually gave feedback like "well done (yoku dekimashita)" or "good (ii yo)" to AIBO, they used personal rewards like "Thank you (arigatou)" for ASIMO, especially for positive reward. While the participants gave more explanations when talking to ASIMO, especially for negative rewards, the difference between both robots was not statistically significant

IV. DISCUSSION

While especially the way of uttering commands seems to depend rather on personal preferences of the user, than on the appearance of the robot, we found robot-dependent differences in the feedback, given by the participants. The most obvious one was the frequent use of touch for giving feedback to AIBO, while touch was almost not used for ASIMO. Moreover, we found, that users tended to give personal feedback like "Thank you" to ASIMO, while they rather commented on the performance when giving feedback AIBO. These findings suggest that people actually use their experience with real dogs as a guideline when giving feedback to AIBO.

The users' subjective evaluation did not reveal significant differences between ASIMO and AIBO. As both robots were programmed to behave in the same way on the same task, we assume, that the users' impression of the robot's behavior depends more on its actual performance than on its appearance.

There are different possible explanations, why no significant differences were observed for giving commands. One of them is that both robots used speech to communicate with the user. As speech is a typical human modality of interacting, differences might have been stronger, if AIBO had communicated with the user in a more dog-like non-verbal way.

V. REFERENCES

[1] A. Austermann, S. Yamada, Learning to Understand Parameterized Commands through a Human-Robot Training Task, IEEE International Symposium on Robot and Human Interactive Communication (ROMAN'09), 2009

[2] Takayuki Kanda, Takahiro Miyashita, Taku Osada, Yuji Haikawa, and Hiroshi Ishiguro, Analysis of Humanoid Appearances in Human–Robot Interaction, IEEE TRANSACTIONS ON ROBOTICS, VOL. 24, NO. 3, JUNE 2008

[3] J. Goetz, S. Kiesler, and A. Powers, Matching robot appearance and behavior to tasks to improve human-robot cooperation, IEEE Workshop on Robot and Human Interactive Communication (ROMAN'03), 2003.