ROBOMO: Towards an Accompanying Mobile Robot

Khaoula Youssef, Kouki Yamagiwa, Ravindra Silva, and Michio Okada

Interaction and Communication Design Lab, Toyohashi University of Technology
1-1 Hibarigaoka, Tempaku, Toyohashi, Aichi
{youssef-yamagiwa}@icd.cs.tut.ac.jp, {ravi,okada}@tut.jp
www.icd.cs.tut.ac.jp

Abstract. Many of the most daily uses of robots require them to work alongside users as cooperative and socially adaptive partners. To provide the human with the better suited assistance at a convenient time, a robot must assimilate the user's behaviors and afford an adaptive response within the context of the interaction they share. We try to understand how a robot communication that is based on inarticulate sounds and iconic gestures is capable to help on the establishment of the attachment process and can enhance the social bonding between the human and our accompanying mobile robot (ROBOMO). In this paper, we draw on inarticulate sound and iconic gestures in order to design our robot and ground the attachment process. We showed that using simple inarticulate sounds and iconic gestures, the attachment process can evolve incrementally which significantly helped to acquire the meaning of the robot's behaviors.

Keywords: Social Bonding, Inarticulate Sound, Minimal Design, Iconic Gestures.

1 Introduction

Social bonding suggests that taking part in a communication increases the attachment and consequently the adaptation capability which may enhance the meaning acquisition process [1]. As an example, infants who form a social bond with their caregivers establish a better sense of their surroundings. In fact, slowed voice tones and physical contact, help the child to establish a preference for the caregiver and a mutual interest in communication evolves [2]. In such scenarios, children distinguish the different voices, and turn their heads to pick up the tones. They can intentionally generate imitations of hand gestures and voice sounds, with different expressions transferring a knowledge, an interest, an excitement, etc. [3]. Meanwhile, caregivers, excited by the infant's expressions, respond with affectionate behaviors by using rhythms of speech and slowed gesture with a soft voice and a moderate modulation of pitch [4]. Incrementally, the attachment evolves and the mutual understanding occurred by mirroring the patterns of each others' expressions [5]. Another similar example that involves the attachment process is the human-pet relationship. Many studies [6][7][8] investigated

M. Beetz et al. (Eds.): ICSR 2014, LNAI 8755, pp. 196–205, 2014.

the beneficial effects of pet ownership on human's interpersonal relationships and explored the importance of the human-animal interaction for the human's relational development [7][8][9]. Sparks et al [9] defines the behavioral attachment during the human pet interaction [1] as a prominent factor that helps the human to understand the pet's signals. It is then reasonable to presume that attachment between the human and others plays a unique role that helps on understanding others and the environment.

In this vein, we are interested in understanding whether inarticulate sounds and simple gestures help to establish the attachment process between the human and our mobile accompanying robot. We believe that we can use them to create a social bond just like in the caregiver-child or the human-pet scenarios and then enhance the adaptation within the human-robot interaction. Designing a robot that is not related to any language or any special cultural behaviors, will afford the chance to create a universal form of communication for the human-robot interaction just as in the child-caregiver scenario that is based on the attachment between both parties and the use of simple cues to establish online the customized social rules. To measure the social bonding, we intend to assess the values of five factors : the degree of adaptation to the social creature, the stress felt by the subject, the friendliness of the robot, the cooperation and the achievement degrees. In our paper work, we will afford a brief explanation about ROBOMO's design and architecture, explain the experimental setup, expose the results and finally we will give a brief discussion.

2 Background

Many studies investigated the attachment of humans to social robots [10][11]. Sung et al [11] indicated that people had a tendency to name their robots. Findings such as this suggested that people may treat robots like they treat a child or a pet [12]. In fact, if the robot exhibits a social behavior, a social bond will be formed and then people feel more comfortable with robots [13]. As an example, Samani et al [13] proposed Lovotics, a robot that uses audio and touch channels along with internal state parameters in order to establish long standing bonds with individuals. Lovotics afforded for the users an intimate relationship and people felt so comfortable that they even hugged the robot. Hiolle et al [14] used the Sony AIBO robot during their experiment where they showed that people tend to form a social bonding with needy robots that demanded assistance from users. The latter study suggests that robots do not need multimodal communication to develop the attachment process and that exhibiting a simple behavior can be attractive enough for the human to feel attached to the robot and to embark on a positive constructive relationship with its. In our study, we will use similar simple behaviors that can be assembled under the immediacy

[1] Behavioral attachment: It consists on the human's involvement in different tasks with their pets such as play or teaching them new instructions where the pets are using their inarticulate sounds and their bodies movements to transfer the meaning to the owner.

cues category: the gestures and inarticulate sounds. We want to explore whether these two social cues can help to ground the attachment process and explore the social bonding's effect on the interaction's meaning acquisition. Inarticulate sounds were used to establish playground language with autistic children [15] and were studied in the context of the human-computer interaction [16] where it was proved that it can lead to a compassionate effect. Iconic gestures [17] [2] facilitates the human-robot interaction [18] and were used in different contexts such as hosting activity [19], showing hesitation [20], etc... In our current work, we intend to ground the attachment process that may evolve between ROBOMO and the participants. We want to verify whether a social bonding can emerge in the context of the human-ROBOMO interaction and whether it can guarantee to transfer the meaning once meshed with the iconic gestures and the inarticulate sounds.

3 ROBOMO Design

We respected the minimal design paradigm which consists on reducing the robot's design and preserving only the most elementary components [16]. ROBOMO has a long shaped body with an attractive container (made of plush) and has no arms. We had intentionally given ROBOMO a pitcher plant (Nepenthe) appearance to encourage people to interact with it, much as one might with a young child or a pet. We believe that exposing a half hairy head (Fig. 1), makes the robot looks cute and affords a starting point for the social bonding process formation. Although used for personal navigation, our accompanying mobile robot is not designed to walk which may create a sort of an empathetic feeling towards ROBOMO. Inarticulate sounds were produced according to Okada et al's [21] generation method of inarticulate sounds. Three types of behaviors were exhibited (i) the inarticulate sounds with meaning (ii) the nodding (iii) gestures (table 1).

Fig. 1. ROBOMO's design

[2] They are speech-related gestures that mention concrete objects for example showing the direction for the human.

Table 1. The different behaviors that ROBOMO can exhibit

Code of the Behavior	Behavior	Description of the Behavior
IS	inarticulate sounds	yes, no, right, left, forward
ND	nodding	en..well, thank you, I'm not sure
GS	gestures	turning left, turning right

4 ROBOMO Architecture

ROBOMO consisted of a micro PC, five servo-motors (AX-12+) for the body movement and a speaker as an output for the robot's inarticulated sounds. A web camera helped to recognize the person's face and a microphone detected the user's requests that was recognized by Julius (a software for Japanese word recognition) (Fig.2).

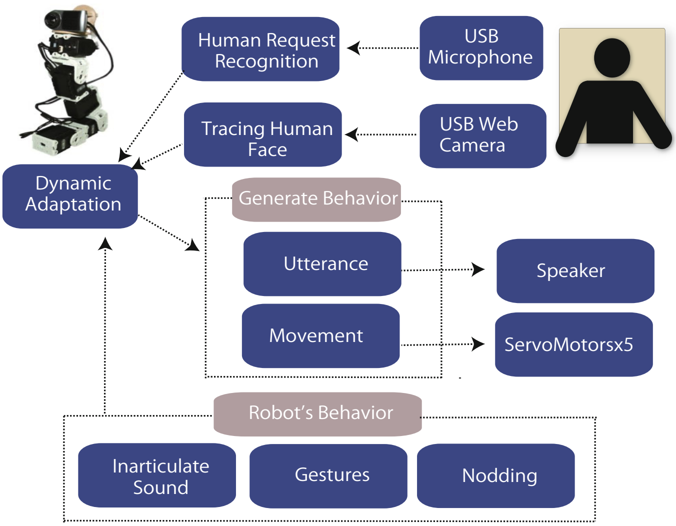

Fig. 2. The system architecture of ROBOMO

5 Experimental Protocol

The main objective is to explore the effectiveness of the attachment process and its impact on the meaning acquisition within a human-robot interaction. We expect that gradually, the communication will be clearer. We setup an indoor ground for navigation task that contains cross points (Fig.3). To pick the right behavior, the participant is instructed by the robot. We asked the participant to talk to ROBOMO with simple words and slowly. 12 participants with age varying in $[22-30]$, take part in 3 sessions. We have chosen several configurations during the 3 sessions to guarantee the diversity of the participant's responses. It helps also to ensure that any successful meaning guessing of ROBOMO's behaviors is not related to the fact that we are using the same configuration but it is related to the social bonding which enhances the participants' adaptation. In our scenario, if the human does not perceive the robot's response, he will repeat his question within a short period for direction's confirmation. In such case, the robot exhibits a body behavior such as pointing to the left or right direction using its upper body part combined with the right inarticulate sound as a response. On the other hand, in the short periods of silence (when the user is not addressing any request), a nodding behavior is displayed. Each student interacts with ROBOMO for 2 minutes and then answers the same 5-Likert Scale questionnaire (13 questions). The table 2 contains the different questions.

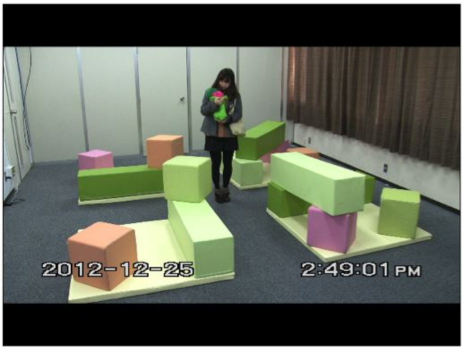

Fig. 3. A snapshot of our mobile accompanying robot interacting with a participant during the experiment

Table 2. The questionnaire evaluating the social bonding's five factors

Factors	Code	Questions
Cooperation	Q1	Has ROBOMO tried the best it can to help you?
	Q2	Do you feel that ROBOMO needed your help?
	Q3	Have you wanted to help ROBOMO?
Achievement	Q4	Had you recognized the direction indicated?
	Q5	Can you distinguish ROBOMO behaviors' different meanings?
	Q6	Do you think that you established a good relational contact?
Friendliness	Q7	Can you consider ROBOMO as a friend?
	Q8	Have you felt that ROBOMO was familiar for you?
Stress-Free	Q9	Was it hard for you to understand ROBOMO?
	Q10	Can you get the feeling of ROBOMO?
Adaptability	Q11	Do you think that ROBOMO is a smart robot?
	Q12	Can you feel that ROBOMO showed some animacy?
	Q13	Do you think that ROBOMO behaved like a baby?

Our evaluation of the social bonding process is articulated around five factors: the adaptation, the stress, the friendliness, the cooperation and the achievement. We tried to record on log files the participants' requests and the robot's instructions. We recorded also the interaction videos that helped us to detect the spatial points when the gestures were used.

6 Results

6.1 Questionnaire Based Results

To statistically identify the most ameliorated social bonding factors, we applied ANOVA based on the users' answers. Table 3 exhibits the different p-values and the Fig.4 displays the average mean opinion score (MOS) values of the different subjects per session where the horizontal axis shows the social bonding five factors combined with their related questions during the three sessions and the vertical axis shows the MOS values for 12 subjects. The MOS is the arithmetic mean of all the individual scores, that ranges from 0 (worst) to 5 (best) where a value that is equal to 3 is acceptable. Based on the Figure4, we can see that cooperation, achievement and stress-free factors slightly went up by means of sessions. Table 3 showed that, the questions Q1, Q2 and Q3 which evaluate the cooperation factor were statistically significant with p-values respectively equal to ***$p=0.0024<0.005$; *$p=0.0927<0.1$ and *$p=0.0993<0.1$. The questions evaluating the achievement (Q4, Q5 and Q6) showed also significant results with p-values respectively equal to ***$p=0.001<0.005$, *$p=0.0615<0.1$ and **$p=0.0137<0.05$. Finally, the questions that concern the stress-free (i) Q9: **$p=0.0391<0.05$ (ii) Q10: **$p=0.0185<0.05$ showed also that there were statistically significant results. These results suggest that the robot's cooperation capability using the inarticulate sounds and the gestures helped on achieving the task and leaded to stress reduction while interacting with ROBOMO.

Based on the Figure.4, we can see that friendliness and adaptability increase slightly while statistically there was no significant differences between the different sessions with respectively (i) Q7: $p=0.2439$ (ii) Q8: $p=0.1573$ for friendliness and (i) Q11: $p=0.2038$ (ii) Q12: $p=0.2875$ (iii) Q13: $p=0.4785$ for adaptability.

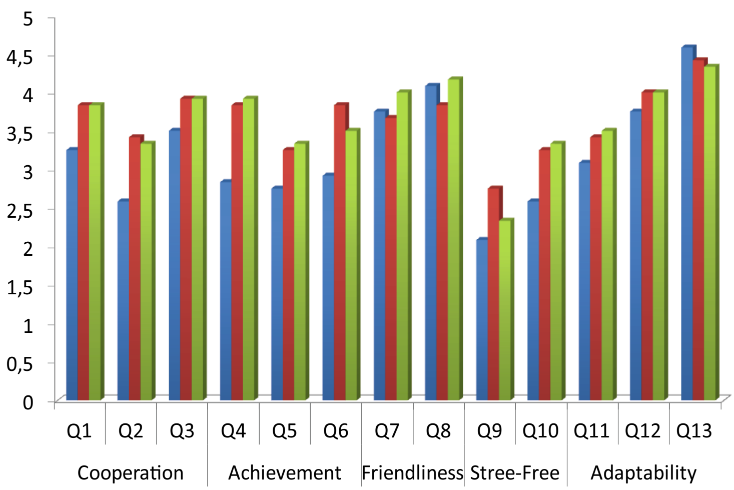

Fig. 4. Results of the average mean opinion score (MOS) based on the 13 questions' answers and for the 3 sessions of the experiment

Table 3. ANOVA evaluation of the questionnaire results

Factors	Code	P-value	Results
Cooperation	Q1	$*p = 0.0927 < 0.1$, d.f=11	significant
	Q2	$* * *p = 0.0024 < 0.005$, d.f=11	significant
	Q3	$*p = 0.0993 < 0.1$, d.f=11	significant
Achievement	Q4	$* * *p = 0.001 < 0.005$, d.f=11	significant
	Q5	$*p = 0.0615 < 0.1$, d.f=11	significant
	Q6	$* * p = 0.0137 < 0.05$, d.f=11	significant
Friendliness	Q7	$p = 0.2439$, d.f=11	not significant
	Q8	$p = 0.1573$, d.f=11	not significant
Stress-Free	Q9	$* * p = 0.0391 < 0.05$, d.f=11	significant
	Q10	$* * p = 0.0185 < 0.05$, d.f=11	significant
Adaptability	Q11	$p = 0.2038$, d.f=11	not significant
	Q12	$p = 0.2875$, d.f=11	not significant
	Q13	$p = 0.4785$, d.f=11	not significant

We asked from people to write down their opinions before and after experiment. We analyzed the participants' different subjective answers and we found out that users confirm that it is easy to adapt with ROBOMO. They found its friendly and cute before even starting the experiment. Thus, the robot's appearance played a key role to reduce the adaptation gap and to give a good first impression.

6.2 Real Time Interaction Results

Based on the stored log files of the speech recognition system and the recorded videos, we counted the user's picked directions based on the robot's indications and the related robot's behaviors (getures, nodding, inarticulate sounds) (table 4) We used the data of the table 4 to evaluate the relationship between participants' behaviors and robot's behaviors. Table 5 shows the different Chi-square test's results where we can see that gradually the p-value increases by means of sessions: $p1 < p2 < p3$ with a statistical significance during the third session. We noticed also that there was no significant results during the two initial sessions. This incremental p-value increase suggests that gradually a strong relationship evolves between the human and the robot's behaviors.

Table 4. The contingency table integrating the human behavior and the related robot's behavior during the 1st, 2nd and 3rd sessions

	Session 1			Session 2			Session 3		
	Human Behaviors			Human Behaviors			Human Behaviors		
Robot's Behaviors	Forward	Left	Right	Forward	Left	Right	Forward	Left	Right
Inarticulate Sounds	9	13	12	13	20	32	9	12	27
Nodding	13	12	18	14	7	11	16	12	13
Gestures	12	6	21	11	11	10	7	17	11

Table 5. Chi-Square test of independency and the corresponding P-values evaluating the relationship between the human behaviors and the robot's behaviors during the different sessions of the experiment

Sessions	Chi-Square Values	P-Values	Results
Session 1	χ^2=5.21, dof=4	$p = 0.266$	not significant
Session 2	χ^2=7.53, dof=4	$p = 0.110$	not significant
Session 3	χ^2=12.2, dof=4	$p = 0.016 < 0.05$	significant

6.3 Correspondence Analysis Results

In order to visualize the relationship between the robot and the users' behaviors, we used a visual approach which is the correspondence analysis. The bi-dimensional map exposed the relationship among categories spatially on empirically derived dimensions. The frequency for each category (*forward*, *right*, *left*) and for each variable (nodding, inarticulate sounds (IS) or gestures) is considered in order to expose the Euclidean distance in two dimensions. Figure 5 depicts the associations between categories of robot's behaviors and participants' picked directions during the three trials. The red triangles represent the participants' chosen directions and the blue dots represent the robot's behaviors. Considering the first trial's correspondence analysis Fig.5 (left), we can see that there was no clear relationship between the robot's behaviors and the human's chosen directions. By analyzing the second session results Fig.5 (center), we can see that the robot's behaviors starts to be mapped with the human chosen directions. In fact, there is a tendency to attribute the nodding behavior with the *left* direction, the inarticulate sounds with the *right* direction while the gestures were associated with the *forward* direction. During the final session Fig.5 (right), the Euclidean distance between the robot's behaviors and the human chosen directions becomes shorter and the tendency to associate for each direction a specific robot's behavior becomes clearer. In fact, human turning *right* behavior was related to inarticulate sounds, turning *left* was associated with the nodding, while going *forward* occurred when the robot exposes gestures.

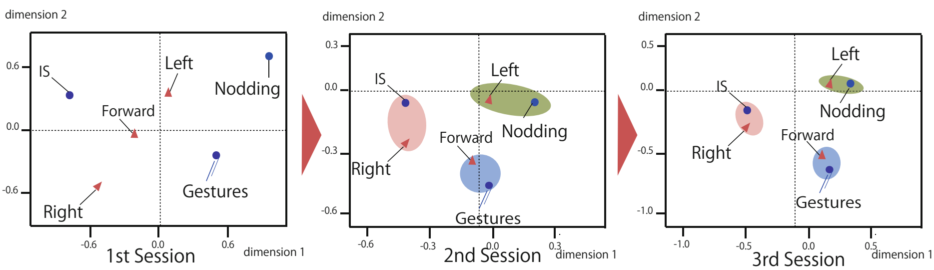

Fig. 5. The correspondence analysis of the trial 1 (left), trial 2 (center) and trial 3 depicting the association between the robot's behaviors (inarticulate sound and gestures) and the directions (forward, right, left)

7 Discussion

Based on the questionnaire results (Fig.4 and table3), we noticed a gradual amelioration on the human's attachment process. The stress was decreasing during the interaction (Fig.4) which explains the different significant p-values (p=0.0391, p=0.0185). Cooperation had also significant values with p=0.0024, p=0.0927 while achievement p=0.001, p=0.0615. This highlights the effectiveness of using inarticulate sounds and iconic gestures to decrease the stress, encourage the human to cooperate with the robot in order to achieve the task and thus helps on creating a social bonding during the human-robot interaction which may facilitate the meaning's acquisition. In fact, we remarked a common interest on finding the frequent successful patterns combining for each robot's behavior a particular direction. Based on the table 5, we remarked that there was an increasing tendency to associate the robot's behaviors with the available directions during the navigation task ($p1 < p2 < p3$). The incremental formation of attuned patterns which maps the robot's behaviors with the human's chosen direction was clearer during the sessions 2 and 3 as the Fig.5 shows. Our experiment leads us to the conclusion that our accompanying mobile robot succeeded in eliciting positive and affectionate behavior from participants. We conclude then that the inarticulate sounds and gestures that were used by ROBOMO during this dyadic interaction appeared sufficient for the attachment evolvement and helped on acquiring the meaning of the robot's behaviors.

8 Conclusion

Our study explored the human's attachment toward our accompanying robot. ROBOMO used inarticulate sound and iconic gestures in order to help people navigating in a block-based environment. It was surprising to see no anxiety-avoidance type of attachment existing in the participants towards ROBOMO which helped to decrease the stress and strengthens the human-robot cooperation in order to achieve the task. The results showed that inarticulate sounds and iconic gestures helped on grounding the attachment process during the experiment and that the participants gradually acquire the meaning of the robot's behaviors. In our future work, we intend to integrate in ROBOMO a self-learning mechanism to improve its adaptation capability and measure the attachment process during the human-robot interaction.

Acknowledgments. Authors would like to express their gratitude to robotics designing group at Kyoto University of Art and Design. This research has been supported by Grant-in-Aid for scientific research of KIBAN-B (26280102) from the Japan Society for the Promotion of science (JSPS).

References

1. Hirschi, T.: Causes of Delinquency. University of California Press, Berkeley (1969)
2. Aitken, K., Trevarthen, C.: Self/other organization in human psychological development. In: Developmental Psychopathology, pp. 653–677 (1997)
3. Custodero, L., Fenichel, E.: The Musical Lives of Babies and Families. Zero to three (2002)
4. Papousek, M.: Intuitive parenting: a hidden source of musical stimulation in infancy. In: Musical Beginnings: Origins and Development of Musical Competence (1996)
5. Beebe, B., Lachmann, F.M.: Infant Research and Adult Treatment: Co-Constructing Interactions. Analytic Press (2005)
6. June, M., Gilbey, A., Rennie, A., Ahmedzai, S., Dono, J., Ormerod, E.: Pet ownership and human health: a brief review of evidence and issues In: BMJ (2005)
7. Kidd, A.H., Kidd, R.M.: Seeking a theory of the human/companion animal bond. In: Anthrozoos, pp. 140–157 (1987)
8. Melson, G.F.: A multidisciplinary journal of the interactions of people and animals. Berg Journals, 45–52 (1987)
9. Melson, G.F.: Children's attachment to their pets: Links to socio-emotional development. Children's Environments Quarterly, 55–65 (1991)
10. Forlizzi, J., DiSalvo, C.: Service robots in the domestic environment: A study of the roomba vacuum in the home. In: ACM SIGCHI/SIGART Conference on Human-robot Interaction, pp. 258–265 (2006)
11. Sung, J., Guo, L., Grinter, R.E., Christensen, H.I.: my roomba is rambo": Intimate home appliances. In: Sung, J., Guo, L., Grinter, R.E., Christensen, H.I. (eds.) Proceedings of the 9th International Conference on Ubiquitous Computing, pp. 145–162 (2007)
12. Jones, K.S., Schmidlin, E.A.: Human-robot interaction toward usable personal service robots. Reviews of Human Factors and Ergonomics, 100–148 (2011)
13. Samani, H.A., Cheok, A.D., Foo, W.N., Nagpal, A., Mingde, Q.: Towards a formulation of love in human - robot interaction. In: RO-MAN, pp. 94–99 (2010)
14. Hiolle, A., Canamero, L., Davila-Ross, M., Bard, K.A.: Eliciting caregiving behavior in dyadic human-robot attachment-like interactions. Interactive Intelligent Systems (2012)
15. Kina, N., Tanaka, D., Ohshima, N., De Silva, R., Okada, M.: Culot: Sociable creature for child's playground. In: Human-robot Interaction, pp. 407–408 (2013)
16. Matsumoto, N., Fujii, H., Okada, M.: Minimal design for human agent communication. In: Artificial Life and Robotics, pp. 49–54 (2006)
17. Cassell, J., Steedman, M., Badler, N., Pelachaud, C., Stone, M., Douville, B., Prevost, S.: Modeling the interaction between speech and gesture (1994)
18. Kanda, T., Ishiguro, H., Ono, T., Imai, M., Nakatsu, R.: Development and evaluation of an interactive humanoid robot"robovie". In: ICRA, pp. 1848–1855 (2002)
19. Sidner, C.L., Lee, C., Kidd, C.D., Lesh, N., Rich, C.: Explorations in engagement for humans and robots. In: Artificial Intelligence, pp. 140–164 (2005)
20. Moon, A.J., Parker, C., Croft, E., der Loos, H.F.: Did you see it hesitate? - empirically grounded design of hesitation trajectories for collaborative robots. In: IROS, pp. 1994–1999 (2011)
21. Noriko, S., Yugo, T., Kazuo, I., Okada, M.: Effects of echoic mimicry using hummed sounds on human computer interaction. Speech Communication, 559–573 (2003)