

■ Children Teach a Care-Receiving Robot to Promote Their Learning: Field Experiments in a Classroom for Vocabulary Learning

Fumihide Tanaka
University of Tsukuba / JST PRESTO

and

Shizuko Matsuzoe
University of Tsukuba

In contrast to conventional teaching agents (including robots) that were designed to play the role of human teachers or caregivers, we propose the opposite scenario in which robots receive instruction or care from children. We hypothesize that by using this care-receiving robot, we may construct a new educational framework whose goal is to promote children's spontaneous *learning by teaching* through their teaching the robot. In this paper, we describe the introduction of a care-receiving robot into a classroom at an English language school for Japanese children (3–6 years of age) and then conduct an experiment to evaluate if the care-receiving robot can promote their learning using English verbs. The results suggest that the idea of a care-receiving robot is feasible and that the robot can help children learn new English verbs efficiently. In addition, we report on investigations into several forms of teaching performed by children, which were revealed through observations of the children, parent interviews, and other useful knowledge. These can be used to improve the design of care-receiving robots for educational purposes.

Keywords: care-receiving robot, teachable robot, learning by teaching, child-robot interaction, early childhood education, learning support, direct teaching, robot ethics, NAO

1. Introduction

Childhood education is a globally important issue and many societies are seeking ways to improve their current method. Robotics researchers have been attempting to support and enrich childhood education by introducing robots into educational environments such as elementary schools (Kanda, Hirano, Eaton, & Ishiguro, 2004; Han, Jo, Park, & Kim, 2005; You, Shen, Chang, Liu, & Chen, 2006; Kanda, Sato, Saiwaki, & Ishiguro, 2007; Han, Jo, Jones, & Jo, 2008; S. Lee et al., 2011), nursery schools (Movellan, Tanaka, Fortenberry, & Aisaka, 2005; Tanaka, Fortenberry, Aisaka, & Movellan, 2005; Tanaka, Movellan, Fortenberry, & Aisaka, 2006; Tanaka, Cicourel, & Movellan, 2007; Ruvolo, Whitehill, Virnes, & Movellan, 2008; Movellan, Eckhardt, Virnes, & Rodriguez, 2009), and home (NEC, 2005; Yujin Robot, 2007) or other (Robins, Dautenhahn, Boekhorst, & Billard, 2005) environments. Although there still remain many practical issues such as safety con-

Authors retain copyright and grant the Journal of Human-Robot Interaction right of first publication with the work simultaneously licensed under a Creative Commons Attribution License that allows others to share the work with an acknowledgement of the work's authorship and initial publication in this journal.

cerns, children have generally been showing to enjoy interacting with robots and the interaction is considered to be effectively used for educational purposes.

Most educational robots for children so far have been designed and developed to play the role of human teachers or caregivers. In other words, they were caregiving robots developed to teach or care for children. In fact, some robots have already been named explicitly as “childcare robots” (NEC, 2005) or marketed as “the service of the teacher’s role” (Yujin Robot, 2007). In contrast, we consider the opposite scenario wherein robots receive instruction or care from children. By introducing a robot that is less capable of doing something than the children and allowing them to teach the robot, we aim to promote children’s spontaneous learning by teaching as a result.

The concept of learning by teaching has been studied and practiced widely in the field of education (Roscoe & Chi, 2007; Aronson, 2000). Moreover, it has been applied with virtual agents in computers (Brophy, Biswas, Kitzlberger, Bransford, & Schwartz, 1999; Biswas, Leelawong, Schwartz, Vye, & Vanderbilt, 2005). Lee classified computer-mediated learning applications into three types: teaching agents, teachable agents, and colearner agents (J.-E. R. Lee et al., 2007). As stated above, in case of educational robot applications, there have been attempts to introduce teacher-type robots (a few of them can also be considered as the colearner type), but there has been no attempt to introduce the teachable type, which receives instruction or care from humans, as a way of promoting the learning activities of the humans who interact with the robot. In addition, the concept of intentionally designing a weaker robotic entity has been employed with Sony’s entertainment robot AIBO (Sony, 1999) and Okada’s ecological creature Muu (Okada, Sakamoto, & Suzuki, 2000). Here we will consider applying the concept for educational purposes.

The paper describes the care-receiving robot project that started in 2009 (Tanaka & Kimura, 2009, 2010) and reports knowledge obtained from field trials that have been conducted since its inception. After spending about one year to prepare an experimental site, we began initial pilot trials in 2010 by introducing a care-receiving robot into a classroom of children (3–6 years of age) to investigate if it would induce children’s caretaking behaviors, as had been assumed. Then, following another set of pilot trials to identify and design an experimental protocol, we started conducting the main experiment in 2011. The purpose was to test whether it could contribute to the learning of English verbs by children by allowing them to teach the care-receiving robot.

The structure of the paper is as follows. First, we explain the basic idea of a care-receiving robot (Section 2), and then describe the overall research plan (Section 3) and the experimental setting (Section 4). Following a summary of the initial pilot trials (Section 5), the main experiment will be described in Section 6.

2. Care-Receiving Robot

A care-receiving robot is a robot that receives care from the people around it. The original concept of the care-receiving robot was first proposed in 2009 (Tanaka & Kimura, 2009, 2010). Conceptually, the meaning of care is broad¹; thus, we can imagine many types of care-receiving robots. We assume that the most typical scenario using a care-receiving robot is for learning support or learning reinforcement for children (Figure 1). In this scenario, teachers or parents decide on an educational topic and ask the children to teach the topic to a robot. Then, children teach the robot, which initially does not appear to be good at solving the topic, but shows an improvement based on the children’s instruction. As will be explained below, children tend to be strongly motivated to take care of robots. Therefore, it is expected that by introducing a care-receiving robot, children would,

¹In this study, we presuppose that teaching behaviors are included in caretaking behaviors. Sometimes we come across behavioral instances of children in which they were not just teaching a robot but caring for it (e.g., giving a word of encouragement). We assume that these behaviors also have some effect on promoting children’s learning, and thus we named the robot here as care-receiving robot without restricting it to teachable robot.

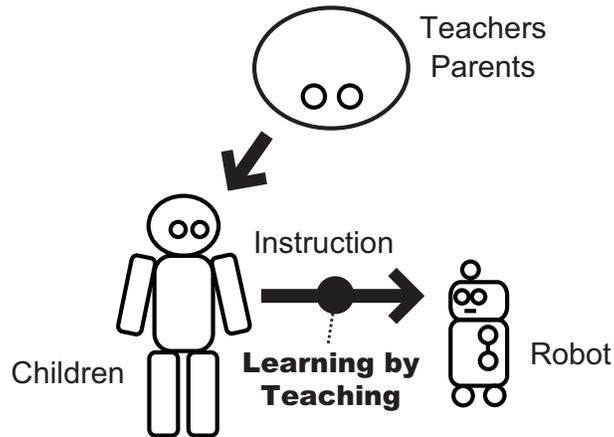


Figure 1. Conceptual diagram involving use of a care-receiving robot to promote children's learning by teaching

as a result, become highly motivated to complete the topic for learning, providing indirect practice (learning by teaching) for the children.

The idea of a care-receiving robot was derived from our past study of children - robot interaction at a nursery school in California. Between 2004 and 2007, the first author of this paper and his colleagues at the University of California, San Diego conducted a long-term field trial, which consisted of immersing a small humanoid robot into a classroom of children younger than 2 years of age (Tanaka et al., 2007; Movellan et al., 2005; Tanaka et al., 2006, 2005). From the observational study, we discovered significant evidence on the socialization process between the children and the robot (Tanaka et al., 2007). The behavioral data obtained from these children provided important insights for the care-receiving robot. It was found that the robot that we introduced into the classroom had encouraged the children's caretaking behaviors more than other toys did in the classroom. Furthermore, it was confirmed that the children's interactions involving caretaking for the robot lasted for longer periods of time. From those observations, we hypothesized that if children are motivated to take care of a robot, we could possibly exploit this characteristic to promote children's learning by teaching.

Another background regarding the care-receiving robot concerns a Robot Ethics debate. In a 2008 issue of *Science*, Sharkey warned that a series of childcare robots, some of which had already been on the market, could produce undesirable side effects such as children's unanticipated emotional development due to the lack of crucial attachment to human caregivers (Sharkey, 2008). As mentioned in the previous section, most educational robots developed so far have played the roles of human teachers or caregivers; therefore, they were considered to have the same issue as childcare robots. On the other hand, the idea behind the care-receiving robot is based on children's caretaking towards a robot. This is similar to their regular activities such as caretaking towards dolls or pets, and the idea of a care-receiving robot is therefore considered to be more acceptable to a wide range of societies (Tanaka & Kimura, 2009, 2010).

3. Overall Research Plan

The idea of a care-receiving robot that was explained in the previous section is based on several assumptions. Although we observed that children younger than 2 years of age frequently cared for the

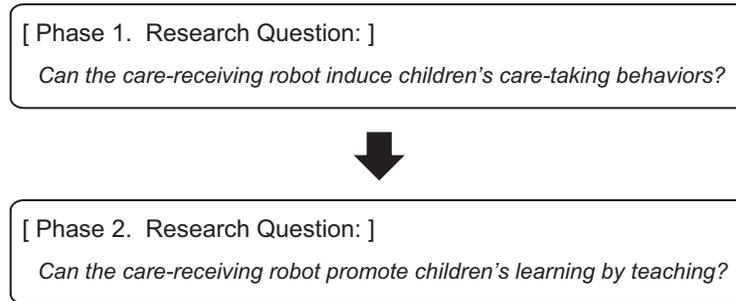


Figure 2. Two phases in testing the idea of a care-receiving robot

robot introduced into a classroom in California (Tanaka et al., 2007), we had not yet tested whether the same behavior would be replicated with different age groups. Since the learning activities of children older than 3 years are very different from those in younger age groups, we had no idea whether the framework of using a care-receiving robot to promote children's learning by teaching would be feasible and useful.

Thus, to test the idea of a care-receiving robot for children's learning support, we developed a research plan that comprised two phases (Figure 2): The goal of the first phase was to confirm the minimum feasibility of the care-receiving robot in a classroom of children older than 3 years of age. We wanted to investigate whether the robot introduced into the classroom was able to induce children's spontaneous caretaking behaviors. The first phase also assumes the role of a pilot study for the second phase. The second phase is the main part of the research with the goals of testing that the educational purpose of the care-receiving robot can be implemented in children's real learning activities, and of investigating if it can, in fact, contribute to their learning reinforcement.

4. Experimental Setting

The best experimental setting for implementing and testing the idea of a care-receiving robot is a classroom within the context of learning activities. Based on this consideration, we chose the venue of an English language school for Japanese children. With the kind cooperation of the Minerva Language Institute Co., Ltd., which manages 600 classrooms in Japan, we were fortunate to be able to conduct experiments in a classroom in Tsukuba. Our target participants were children from 3 to 6 years of age who attended classes in this classroom. This school is a private school specializing in English language instruction for children. Most children come to the school once per week for their lesson and this classroom hosts 5 to 10 lessons per week. After we received approval for this experiment from the Ethical Committee of the University of Tsukuba, we started recruiting participants by explaining and advertising our research to the parents of the children. After obtaining written consent from each parent, we started the initial pilot trials (Section 5) in 2010.

Figure 3 shows the classroom in which the experiments were conducted. The classroom was approximately 25 m^2 in size, and all children were very familiar with the room. Since children are very sensitive to changes in the classroom atmosphere, we needed to keep as natural as possible, even during our experiments. Sometimes that effect can be much more dominant than an experimental control; therefore, we always attempted to maintain a relaxed and natural atmosphere for the children. We taped the classroom activities during all trials with the care-receiving robot using two camcorders (one mounted in the ceiling and the other in a corner at ground level).



Figure 3. Classroom where care-receiving robot trials were conducted

5. Initial Pilot Trials

5.1 Goal of the Pilot Trials

Between September 2010 and February 2011, we conducted initial pilot trials that consisted of introducing a care-receiving robot into the classroom for a total of 6 days using 18 participants. Our goal was to implement the care-receiving robot and observe whether it could induce children's spontaneous caretaking behaviors during the context of English instruction in the classroom (the first phase in our research plan described in Section 3).

5.2 Method

After studying a typical class flow for the target children (ages 3–6 years), we prepared Aldebaran Robotics' Nao to behave as a care-receiving robot by programming the necessary movements (locomotion and gestures) and dialogs (a few common phrases used in the classroom), along with its teleoperation (wizard-of-Oz) interface. Then, we introduced the robot into several game activities in the classroom, such as the "wall touch game," where the children had to touch the appropriate flashcard on a wall, as instructed by the teacher, and the "favorite card game," where the children had to select one flashcard and announce the word that the flashcard represented.

During the 6 days of trials, we tested two types of Nao. One Nao was remote-controlled so that it always gave incorrect answers to all questions, thereby presenting many opportunities for children to teach it (error rate 100% condition). In contrast, the other Nao did not make any mistakes at all, answering all the questions correctly (error rate 0% condition). For the 100% condition, 9 participants interacted with Nao, while 9 different participants independently interacted with Nao for the 0% condition.

Using videos taken during the trials, we performed a behavioral analysis of the participants. First, we coded as many interaction behaviors of the participants with the robot as possible. After carefully reviewing the videos, we decided to adopt 6 behavioral categories: "Touch," "Speech," "Gesture," "Gaze," "Hand-delivering," and "Others." Then, each coded behavior was classified on the basis of whether it was a caretaking behavior or not. That way, we observed the number of caretaking behaviors that were induced at both conditions. The video coding was conducted by the second author of this paper whose inter-observer reliability, which was calculated with three external

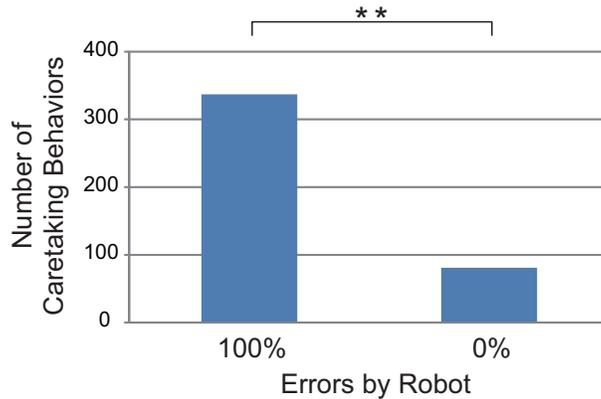


Figure 4. Effect of robot errors on the total of caretaking behaviors observed for each condition. The Chi-square test was conducted ($\chi^2 = 156.8, p < 0.001$).

coders, averaged 0.70.

5.3 Results and Discussion

Figure 4 shows the results. Significantly more caretaking behaviors were observed with the error rate 100% condition, which implies that the basic idea of a care-receiving robot would be feasible in the classroom. Having confirmed that the robot can induce children’s caregiving behaviors in a classroom of children that are older than 3 years of age, we were ready to proceed to the second phase to investigate its effects on children’s learning support. The initial pilot trials also provided us with some important insights regarding the design of experiments for the second phase. These are mentioned below.

First, the game activities used to introduce the care-receiving robot in the initial pilot trials were designed to be easy for children in this age group and were also easy for them to teach the care-receiving robot. However, those activities were not suitable for determining their learning process. Furthermore, because the activities were so easy, some children appeared to quickly get bored with them. When designing an experiment for the second phase, we determined that the experimental scenario needed to include more difficult learning tasks for the children.

Second, as shown in Figure 5, the most frequently observed behaviors of children toward the care-receiving robot were concerned with their “Touching” the robot. Thus, one promising scenario that could be utilized for teaching a care-receiving robot would be a case where children are required to take the robot by the hand and teach it step by step. As will be explained in Section 6, we adopted this *direct teaching* task for the second phase.

Third, it is very important that our experiments are not interpreted as being supportive of robots replacing human teachers or parents. The care-receiving robot proposed here is supposed to be a tool for teachers or parents in educational contexts. Thus, we consider it is appropriate to create an experimental scenario in which a care-receiving robot keeps attending a lesson given by a human teacher and its effect is assessed by comparing it with other educational tools within the lesson.

Finally, due to the classroom capacity and the enrollment of children, it would be difficult to recruit a sufficient number of unbiased participants to conduct experiments with multiple independent conditions. The maximum population size of participants expected for the next experiment was 15–20, which also includes children who participated in the initial pilot trials. It would therefore

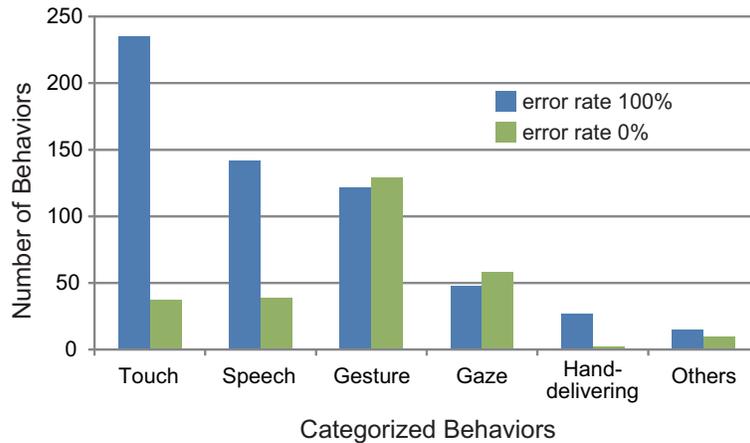


Figure 5. Frequency of six categorized behaviors in the error rate 100% and 0% conditions.

be difficult to control the exposure of each participant to the robot uniformly under all conditions. Therefore, for the actual experiment, it was necessary to give up the idea of dividing participants into multiple conditions, and instead, consider using a single experimental protocol that is common to all the participants from which we attempted to extract useful results.

6. Experiment Using the Care-Receiving Robot to Promote Learning by Teaching

6.1 Goal of the Experiment

Following the initial pilot trials described in Section 5, we proceeded to the second phase, which was the main phase of the overall research plan, as explained in Section 3.

Here, the goal was (1) to investigate whether the care-receiving robot that we introduced into the classroom could promote children’s learning by teaching and (2) to categorize the forms or types of their teaching. It was expected that children would teach the care-receiving robot in many ways. Some children might just give verbal instruction, while other children might take its hand and taught it step by step. We analyzed the behavior of the children, with a particular focus on their teaching behaviors. The analysis is expected to provide useful knowledge for designing a more teachable care-receiving robot in the future.

6.2 Method

6.2.1 Participants and apparatus After performing another set of pilot trials for 3 days to finalize the experimental protocol, we started the experiment in July 2011 in the same classroom described in Section 4. We went to the classroom almost every day for one month to set up, advertise, and conduct the experiment. We obtained permission to include 18 children in the experiment. However, the data for one participant was excluded because the participant could not complete the experimental session correctly. As a result, a total of 17 participants (3–6 years of age, 12 females and 5 males) provided the data for our study. Similar to the case with pilot trials, the experiment approved by both the Minerva Language Institute Co., Ltd. and the Ethical Committee of the University of Tsukuba was conducted after obtaining written consent from all parents of the participants.

Two experimenters were involved, both of whom were authors of this paper. Experimenter #1

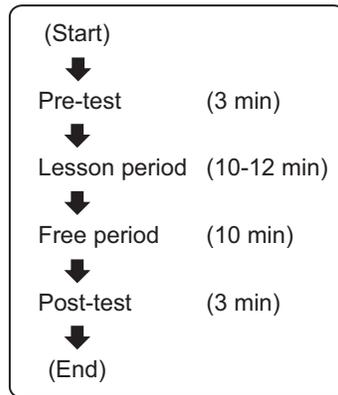


Figure 6. Experimental flow chart

had been studying and visiting the classroom for two years and therefore was very familiar with the regular class flow. He played the role of a human teacher. This individual was also in charge of maintaining safety during the experiment. Experimenter #2 remotely operated a care-receiving robot from outside the classroom. A monitoring camera and a microphone (LifeSize Passport) along with a window provided sufficient sensory information for the operator to execute the experiment. Similar to the case with pilot trials, we used Aldebaran Robotics' Nao to implement the care-receiving robot. During the experiment, it was fully teleoperated by Experimenter #2 using a teleoperation interface that had been developed.

6.2.2 Design We decided to adopt an experimental scenario in which a care-receiving robot attended a lesson given by a human teacher, and its effect was assessed by comparing it with other educational tools (such as vocabulary cards) used within the lesson. All participant children were exposed to a common protocol. In particular, each participant was asked to attend a lesson together with a care-receiving robot. The lesson was guided by a human teacher and it comprised four rounds of verb-learning games. For two randomly chosen rounds, the human teacher used the care-receiving robot during the verb-learning game. For the other two rounds, the human teacher conducted a regular verb-learning game. The number of teaching actions given from the human teacher to the participant during the experiment was controlled (equally) for both the cases with and without the care-receiving robot. The hypothesis tested was that the participants would learn more English verbs over exposure to the robot.

6.2.3 Procedure In this section, we will explain the experimental procedure in detail. Each session (one participant trial) lasted approximately 30 min. Figure 6 illustrates the flow chart of the experiment. Next, we describe each step sequentially.

Pre-test: Based on repeated consultation with professional teachers at the Minerva Language Institute Co., Ltd., we decided to adopt a verb-learning game. This was considered to be most appropriate for the age group of 3–6-year olds used in our study. Each graphic card (on the left in Figure 7) represented a verb whose spelling was printed on its corresponding word card (right in Figure 7). Most children in this age group could not read the verb printed on the word card when viewing it, but they were asked to pick up a corresponding graphic card on the floor when a teacher (Experimenter #1) pronounced a specific verb while showing them the word card.

At the beginning of each session, one participant entered the classroom with Experimenter #1.



Figure 7. Graphic cards (left) that represented verbs whose spelling was printed on a corresponding word card (right). By conducting a pre-test, four word cards were chosen for the main experiment.

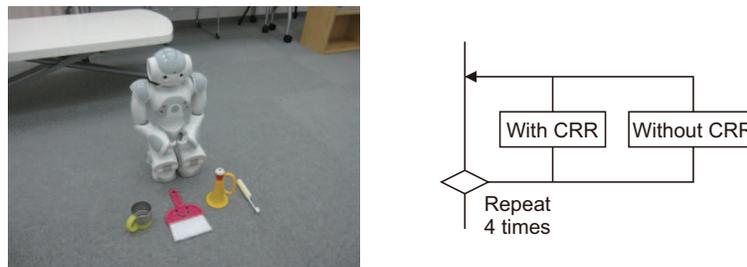


Figure 8. (Left) Care-receiving robot with four objects on the floor that represented the verbs “drink,” “sweep,” “play,” and “brush.” (Right) Lesson period flow chart

Then, the pre-test was conducted to identify and select four verbs that were unknown to the participant. The pre-test procedure was as follows. The participant and Experimenter #1 sit in front of six graphic cards placed on the floor (left in Figure 7). Those cards were randomly chosen from a set of 13 cards before each session starts. Experimenter #1, who had the word cards in his hand, chose one word card from them and then asked the participant to pick up the graphic card from the floor that matched the verb spoken by Experimenter #1 and was shown on the word card. If the participant picked up the correct card, Experimenter #1 said “That is correct” and gave the card pair to the participant. Then, Experimenter #1 added one new graphic card to the card circle on the floor, maintaining six graphic cards in the group at all times. If the participant picked up an incorrect card, Experimenter #1 just said “That is incorrect,” and moved on to the next question. This procedure was repeated until the participant made a mistake four times, indicating that four unknown verbs had been identified.

Lesson period: Then, Experimenter #1 went to a back room and brought the care-receiving robot into the classroom along with four objects that matched the four verbs identified during the pre-test. After the care-receiving robot was placed at the center of the classroom facing the four objects on the floor (left in Figure 8) and the participant, the lesson began.

The lesson period was initiated by Experimenter #1 who played the role of a classroom teacher. The flow chart of the lesson period was shown on the right in Figure 8. The scenario followed the pattern of a regular lesson in the classroom. First, Experimenter #1 would say the name of the participant and the care-receiving robot, “Nao,” which was teleoperated by Experimenter #2. The care-receiving robot greeted the participant cheerfully and introduced itself. The robot also attempted to bond with the participant by calling his or her name and shaking hands with the child.

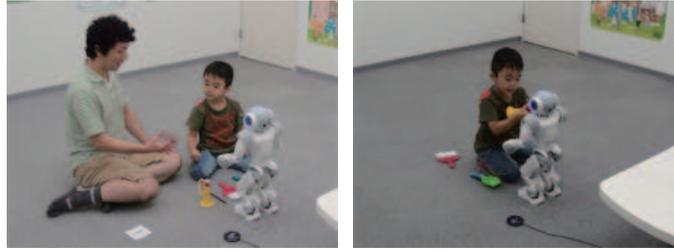


Figure 9. (Left) Lesson period, (Right) Free period

Then, Experimenter #1 started and guided a verb-learning game with the following procedure: (1) From the four word cards obtained during the pre-test, two cards were randomly chosen in advance and marked as “With CRR (care-receiving robot).” (2) From the four word cards, Experimenter #1 picked up one card randomly and asked the child its meaning saying, “Show us how to do <the verb>.” Since the verb was still unknown to the participant at this time, the participant could not answer the question correctly. (3) Then, Experimenter #1 showed the correct answer by picking up the corresponding object and pronouncing the verb. If the card was marked “With CRR,” the game continued to step (4), otherwise, they returned to step (2), repeating the steps with another card. (4) Experimenter #1 turned to the care-receiving robot and asked it the same question. The care-receiving robot responded “Yes” and then said “Please pick up” while opening its right hand. Experimenter #1 handed the corresponding object to the robot. The care-receiving robot then held the object and made a wrong movement. For the wrong movement, we designed four patterns in advance that appeared to be completely meaningless and were different from all the verbs used in the experiment. The care-receiving robot outputted one of the four patterns each time it made an incorrect movement. (5) Experimenter #1 said “No, that is incorrect.” The care-receiving robot responded with “Please teach me.” Then, Experimenter #1 took the robot by the hand and taught it step by step (direct teaching). As soon as Experimenter #1 started direct teaching, Experimenter #2 switched the care-receiving robot to a recording mode, wherein every servo movement was recorded in real time. After direct teaching was finished, the care-receiving robot played back the movement correctly while pronouncing the verb. Then, Experimenter #1 said “Yes, that is correct.”

Experimenter #1 repeated the procedure for all four word cards to finish the first part of the lesson. The second part of the lesson was performed in a similar manner using the same four word cards. However, in this case, in step (5), Experimenter #1 asked the child to teach the care-receiving robot instead of Experimenter #1. Since not all participants could immediately perform direct teaching, Experimenter #1 might be required to repeat the teaching as in the first part of the lesson. A snapshot picture taken during a lesson can be seen in Figure 9 (left).

Free period: After the lesson period was completed, the participant was told to play freely with the care-receiving robot by themselves (Figure 9, right). During that time, Experimenter #1 went to a bookshelf in the corner and pretended to read books (appearing to ignore the care-receiving robot and the participant). However, he was paying enough attention to maintain the safety of the participant. Experimenter #1 used a stopwatch to ensure that the free period was exactly 10 min long. Meanwhile, the care-receiving robot was teleoperated to follow steps (4) and (5) described above. When the participant came in front of the robot, it said “Please pick up” while opening its right hand. If the participant handed it an object, the robot made a wrong movement at first. Then, if the participant said “No” or exhibited any kind of behavior intending to express a negative response

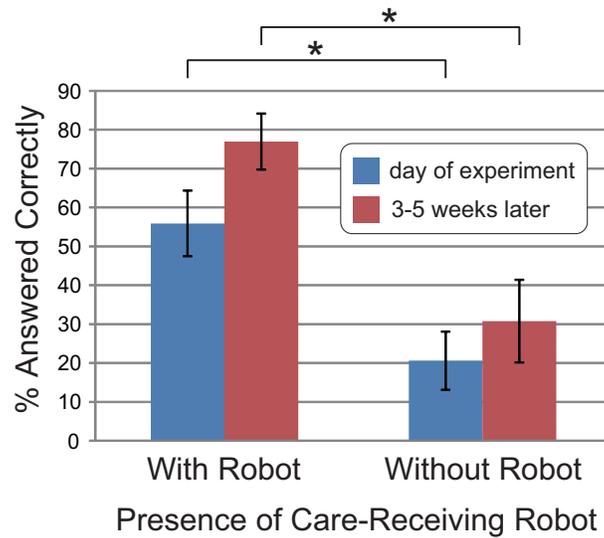


Figure 10. Post-test results: Each bar represents the average percentage of post-test questions answered correctly by 17 (13 at the second post-test) participants. The post-test was conducted two times: the initial post-test was conducted on the day of the experiment, and the second post-test was conducted 3–5 weeks later. Four participants joined another session with the care-receiving robot during the 3–5 week period; therefore, we excluded their data from the results of the second post-test. Wilcoxon signed-rank test and rank-sum test (Mann-Whitney U test) were conducted. There were significant differences ($p < 0.05$) between the average percentage answered correctly for verbs learned with and without the care-receiving robot. The differences between the results obtained on the day of the experiment and 3–5 weeks later were not statistically significant.

(e.g., shaking his or her head horizontally), the robot said, “Please teach me” and waited for the participant to teach it (as in step (5) above).

Post-test: At the end of each experimental session, a post-test was conducted to evaluate if the participant could answer the four questions that he or she could not answer during the pre-test. The basic procedure for the post-test was identical to that of the pre-test. Six graphic cards (four of which represented the verbs used during the lesson period) were placed on the floor, and the participant was asked to pick up the correct answer from them. Note that since no graphic card was used during the lesson period or free period, the participant had to be able to sufficiently generalize the knowledge of each verb such that he or she could identify the corresponding graphic card by only seeing the pictures and hearing the pronunciation.

6.3 Results

6.3.1 Post-test summary Figure 10 summarizes the results of the post-test. Each bar represents the average percentage of post-test questions answered correctly by 17 participants. We observe that the participants picked up more correct graphic cards for verbs learned with the care-receiving robot than without. As discussed in Section 6.4, this experiment was not designed to compare human teachers with the care-receiving robot, and therefore, the results should not be interpreted as showing the superiority of the care-receiving robot to human teaching.

One might feel that this result is natural because children probably spent more time interacting with the care-receiving robot than the time spent without the care-receiving robot (also the differ-

ence in interaction between the two cases could have influenced the result). However, from our viewpoint, it was difficult to know whether interaction with the care-receiving robot would result in enhanced learning, since this was the first experiment investigating this hypothesis. We consider it to be valid to interpret the results as showing that the idea of a care-receiving robot can be implemented to supplement the learning of English verbs in a classroom. It can be used to promote and enhance children's learning by teaching, even verbs that are difficult to learn with a minimum style of teaching (only about 20% were answered correctly without the care-receiving robot).

We also conducted another set of post-tests 3–5 weeks after the day of each experiment. The goal was to investigate the extent to which participants remembered the verbs learned during the experiment after a period of absence. We used exactly the same post-test procedure as that described in Section 6.2.3. Contrary to our initial assumption, the average percentage of questions answered correctly was in fact higher than the results obtained on the day of the experiment (we predicted that it would have been lower because the children might have forgotten the learned verbs). Subsequent interviews with parents (details in Section 6.3.3) provided us with some clues for this result. Most parents told us that their children appeared to enjoy the experience of teaching the care-receiving robot so much that he or she continued to play the game at home, even after several days, weeks, and wherever similar objects were found. This fact was very encouraging to us because this was exactly what we wanted to achieve — promoting children's spontaneous learning.

6.3.2 Forms of teaching One aspect that we did not foresee before conducting the experiment was that despite the form of direct teaching featured during the lesson period (Section 6.2.3), we observed the children using other forms of teaching the care-receiving robot during the free period. Investigating these forms was useful for designing a more effective care-receiving robot in the future; thus, we conducted a behavioral analysis of the participants during the free period. Video coding was conducted by the second author of this paper using ELAN software. The inter-observer reliability measured with three independent external coders was 0.70. From observation of the videos, we found that the general forms of teaching used by the participants with the care-receiving robot could be classified into the following three categories.

Direct teaching: The participant took the robot by the hand and taught it step by step (e.g., let the robot hold a toothbrush and moved its hand in front of the face). This involved more direct behavior by the participant where he or she acted directly on the robot (e.g., brushed the robot's mouth while holding a toothbrush in his or her hand).

Gesturing: The participant demonstrated a procedure to the robot by moving his or her body and showing it to the robot (e.g., brushed his or her mouth with or without holding a toothbrush in his or her hand).

Verbal teaching: The participant gave vocal instructions to the robot (e.g., just saying "Brush" to the robot).

Figure 11 shows histograms of the total durations (accumulated time) and total frequencies of participants' behaviors for the three categories. It was observed that gesturing and verbal teaching behaviors were performed frequently, although Experimenter #1 did not lead the participants to do so. The results may imply that gesturing and verbal teaching are also easy methods for the participants and are consistent with the observations made (Figure 5) in initial pilot trials.

Next, we investigated the extent to which each form of teaching resulted in the improved learning performance of the participants. To this end, we conducted ANOVA. The results showed that in cases of both the accumulated time and frequency, direct teaching had a significant effect ($p < 0.001$) on the learning performance (% answered correctly) of the participants in this experiment. Table

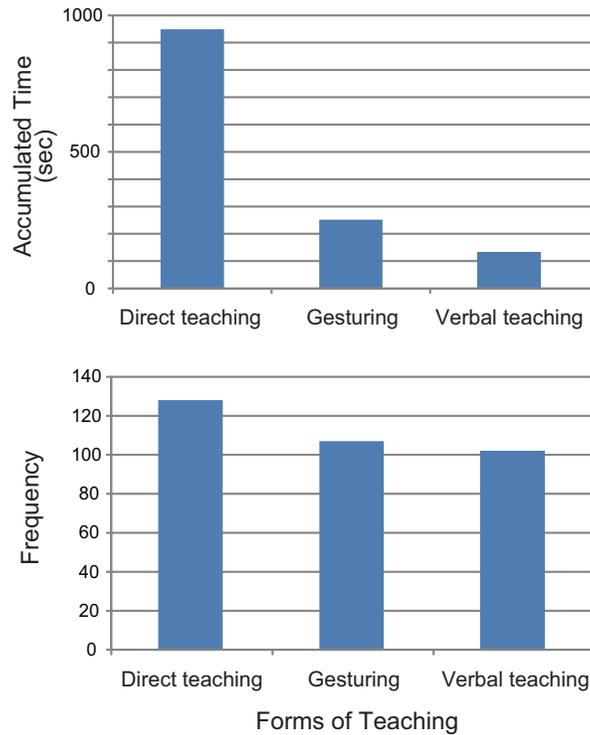


Figure 11. Histograms of the three forms of teaching observed by the 17 participants with the care-receiving robot

1 summarizes one of the results. Gesturing and verbal teaching may be easily performed by the participants, but the results indicate that they did not result in the reinforcement of the participants' learning as was the case with direct teaching. However, owing to experimental limitations, we cannot conclude that gesturing and verbal teaching do not have any effect on the learning performance of the participants in general. We will discuss this issue in Section 6.4.

6.3.3 *Parent interviews* At the time of the second post-test (3–5 weeks after the experiment), we interviewed each parent. The interview consisted of three yes/no questions followed by some casual conversations on two main topics:

- Q1: Does the participant have a brother or sister?
- Q2: Does the participant live with a pet?
- Q3: Did the participant have any prior interaction with robots?
- Q4: Was there any preference on teaching or caretaking?
- Q5: Were there any behavioral changes after the experiment?

The answers reveal that 10 of the 17 participants have brothers or sisters, 4 participants live with a dog, 7 participants had prior experiences interacting with robots at amusement parks, and 14 participants were said to enjoy teaching or taking care of dolls or humans. ANOVA showed a significant effect of Q4 on the learning performance of the participants in this experiment. We also felt that the preference for teaching or caretaking seemed to be a general preference for participants in this age group, which is consistent with the knowledge obtained from our past study (Tanaka et

Table 1: ANOVA results for the effect of each source factor on the learning performance of the participants (Generalized Linear Models, L.R.: Likelihood Ratio)

Source	df	L.R. χ^2	<i>p</i>
Frequency of Direct teaching	1	32.3	< 0.001
Frequency of Gesturing	1	1.4	<i>ns</i>
Frequency of Verbal teaching	1	0.3	<i>ns</i>

al., 2007).

With regard to the last question on behavioral changes (or any comment) of the participant after the experiment, many parents mentioned that their children had repeated the teaching game often even though there was no robot at home, particularly when he or she found objects that were similar to those used in the experiment. Since it was difficult to objectively measure this factor, we could not interpret based on statistical results. However, as we discussed in Section 6.3.1, we believe that children’s spontaneous activity during the 3–5 week interval could have likely contributed to the increase in the second post-test results. In fact, during the second post-test, 8 participants could correctly answer a question that he or she could not answer at the time of the initial post-test.

6.4 General Discussion and Limitations

First, it is important to recognize that the results do not suggest the superiority of the care-receiving robot to human teaching. Professional teachers can use various teaching methods flexibly based on the personality and capability of each student. In contrast, the role of the human teacher was significantly controlled in our experiment such that Experimenter #1 always performed the same minimum teaching scenario with all participants. The results appear to show that the idea of a care-receiving robot is feasible and has the effect of promoting and enhancing children’s learning by teaching if it is used within classroom activities guided by human teachers. We may also deduce that a care-receiving robot can help with their teaching activities. Usually, children in this age group are easily distracted; thus, it sometimes requires significant effort from teachers to keep them focused on a given learning activity. Since the care-receiving robot has the effect of inducing children’s spontaneous caretaking behaviors, teachers could use it when they want children to return to and focus on the learning activity.

This study has several limitations. First, with regard to the sample, it was moderately biased toward females (12 females and 5 males). Although we have not yet been able to statistically argue the gender difference, there may be a difference between female and male participants in interactions with the care-receiving robot. We (including professional teachers in the classroom) notice a general difference between the behaviors during the lessons of female and male children in this age group, but it has not been clarified.

Second, in this experiment, we did not control the extent of each participant’s exposure to the robot Nao before the experiment started. As was described in Section 5, we conducted several pilot trials before the main experiment. Of the 17 participants who participated in the main experiment, 10 participants had at least one previous experience of interacting with Nao. This was one of the reasons for adopting a single common protocol for all participants when designing the experiment (see Sections 6.2.2 and 5). Because the condition was common for all participants, it becomes valid to argue a difference within the condition (e.g., Figure 10). Nevertheless, we recognize that those results were obtained from participants whose exposure to the robot Nao was uncontrolled.

Third, there was a limitation with regard to the robot’s response to the participants. As we

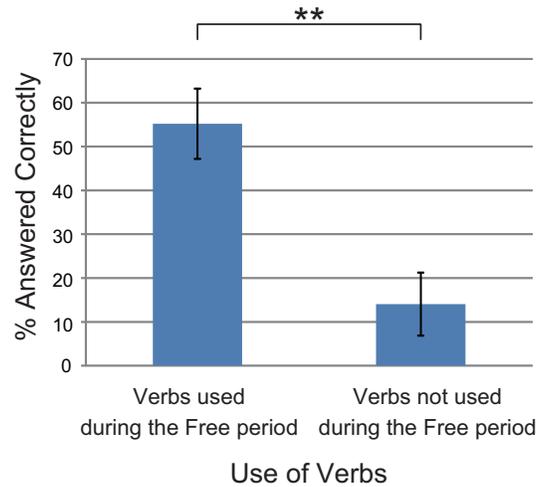


Figure 12. Post-test results (follow-up): Each bar represents the average percentage of post-test questions (that were conducted on the day of the experiment) answered correctly about the verbs used or not used during the free period. Wilcoxon rank-sum test was conducted. There was a significant difference ($p < 0.01$) between the two groups of the results.

described in Section 6.2.3, the care-receiving robot responded to the participants’ direct teaching. However, it did not respond to their other forms of teaching such as gesturing or verbal teaching. This was because for this experiment we originally decided to focus on children’s behaviors concerned with their “Touching” the robot, and then decided to adopt a direct teaching task (Sections 5 and 6). Because the robot did not respond to the participants’ gesturing or verbal teaching, it may have been unfair to compare the effects of the three forms of teaching on the learning performance of the participants in this experiment. Therefore, the ANOVA results in Section 6.3.2 have to be evaluated while considering this limitation. We are currently planning for subsequent experiments that focus mainly on the different approaches of teaching on care-receiving robots.

Fourth, the design of the experiment could not precisely isolate whether the participants learned by observing the experimenter teaching or demonstrating to the care-receiving robot, or by modeling the experimenter’s behaviors using the robot, or by teaching the robot. Therefore, the results in Figure 10 have to be evaluated considering this limitation. To gain an added understanding on the effect of the participants’ learning by teaching, we analyzed the free period once again. We first extracted the verbs used by each of the participants during the free period (in which the participant was teaching the verbs to the care-receiving robot), and then observed whether the participant could answer post-test questions on the verbs correctly. From the analysis, it was found that the average percentage of post-test questions answered correctly about the verbs used during the free period was significantly higher than that with the other verbs, which were used during the lesson period but not used during the free period (Figure 12). Considering that all the verbs were used an equal number of times during the lesson period, this result indicates that the participants learned further by teaching the care-receiving robot (or modeling the experimenter) than by just observing the experimenter during the lesson period.

The most significant disadvantage of the care-receiving robot used in the experiment is considered to be the limitation of *its* learning capability. This was the first formal experiment used to assess the care-receiving robot for the purpose of children’s learning by teaching; thus, we did

not incorporate a learning factor into the robot, i.e., the robot continued to make mistakes (it was only able to repeat a correct movement once just after the participant taught it). This likely caused some participants to become frustrated or irritated. After the experiment was over, we asked four participants to attend another session with the care-receiving robot. There, we followed the same experimental protocol as in the original experiment. Although the participants continued to learn new verbs with the care-receiving robot, we also noticed signs of habituation from two of the participants. If the care-receiving robot can continue to learn as humans, it may increase the motivation of the participants. The design of robot learners (Thomaz & Cakmak, 2009; Suay & Chernova, 2011) and the study of humans who interact with a teachable robot (Kim, Leyzberg, Tsui, & Scassellati, 2009) have been becoming major topics in HRI (Human-Robot Interaction) research field. These literatures are valuable in designing better and autonomous care-receiving robots in near future. Particularly, in our case, the most important issue would be the learning speed or dynamics of the robot. If it learns at a speed that is faster than the participants, then this may also cause a problem. Therefore, the appropriate learning speed or the appropriate learning dynamics of the care-receiving robot will become our next important research target.

7. Conclusions

The paper reported an experiment that introduced a care-receiving robot into a classroom of English instruction for children and observed its effect on children's learning by teaching. The results provided us with more knowledge than expected in our initial hypothesis. In particular, the increased performance of the second post-test seems to be one indicator that the care-receiving robot contributed to the enhancement of the children's spontaneous learning, which is a key concept of this educational framework, accelerating children's natural motivation to provide caretaking by introducing a robot.

We believe that this idea will provide a useful aspect for both education and robotics research fields. In terms of educational support, we believe that the basic idea of a care-receiving robot is very general. In the future, we will therefore explore cases that involve learning other than vocabulary learning, along with improving the requirements for designing better care-receiving robots. In terms of robotics, we consider that the philosophy of developing a weakly designed agent and a better design policy to support humans who interact with it will provide a new perspective to the robotics field in general.

Acknowledgments

We acknowledge the support provided by the Minerva Language Institute Co., Ltd., the JST PRESTO program, and KAKENHI (23680020). We thank the parents and children in the classroom, and the students of the University of Tsukuba for their cooperation. We also thank the editors and reviewers of this paper for their efforts and valuable comments.

References

- Aronson, E. (2000). *The Jigsaw Classroom*. Retrieved March 21, 2012, from <http://www.jigsaw.org>.
- Biswas, G., Leelawong, K., Schwartz, D., Vye, N., & Vanderbilt, T. T. A. G. at. (2005). Learning by teaching: A new agent paradigm for educational software. *Applied Artificial Intelligence: An International Journal*, 19(3-4), 363-392, <http://dx.doi.org/10.1080/08839510590910200>.
- Brophy, S., Biswas, G., Katzlberger, T., Bransford, J., & Schwartz, D. (1999). Teachable agents: Combining insights from learning theory and computer science. In *Artificial Intelligence in Education* (p. 21-28). Amsterdam, The Netherlands: IOS Press.

- Han, J., Jo, M., Jones, V., & Jo, J. H. (2008). Comparative study on the educational use of home robots for children. *Journal of Information Processing Systems*, 4(4), 159-168, <http://dx.doi.org/10.3745/JIPS.2008.4.4.159>.
- Han, J., Jo, M., Park, S., & Kim, S. (2005). The educational use of home robots for children. In *Proceedings of the 2005 IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2005)* (pp. 378-383, <http://dx.doi.org/10.1109/ROMAN.2005.1513808>). Nashville, TN, USA: IEEE.
- Kanda, T., Hirano, T., Eaton, D., & Ishiguro, H. (2004). Interactive robots as social partners and peer tutors for children: a field trial. *Human-Computer Interaction*, 19(1-2), 61-84.
- Kanda, T., Sato, R., Saiwaki, N., & Ishiguro, H. (2007). A two-month field trial in an elementary school for long-term human-robot interaction. *IEEE Transactions on Robotics*, 23(5), 962-971, <http://dx.doi.org/10.1109/TRO.2007.904904>.
- Kim, E. S., Leyzberg, D., Tsui, K. M., & Scassellati, B. (2009). How people talk when teaching a robot. In *Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2009)* (pp. 23-30, <http://dx.doi.org/10.1145/1514095.1514102>). La Jolla, CA, USA: ACM/IEEE.
- Lee, J.-E. R., Nass, C., Brave, S. B., Morishima, Y., Nakajima, H., & Yamada, R. (2007). The case for caring colearners: The effects of a computer-mediated colearner agent on trust and learning. *Journal of Communication*, 57(2), 183-204, <http://dx.doi.org/10.1111/j.1460-2466.2007.00339.x>.
- Lee, S., Noh, H., Lee, J., Lee, K., Lee, G. G., Sagong, S., et al. (2011). On the effectiveness of robot-assisted language learning. *ReCALL*, 23(1), 25-58, <http://dx.doi.org/10.1017/S0958344010000273>.
- Movellan, J. R., Eckhardt, M., Virnes, M., & Rodriguez, A. (2009). Sociable robot improves toddler vocabulary skills. In *Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2009)* (pp. 307-308, <http://dx.doi.org/10.1145/1514095.1514189>). La Jolla, CA, USA: ACM/IEEE.
- Movellan, J. R., Tanaka, F., Fortenberry, B., & Aisaka, K. (2005). The RUBI/QRIO project: origins, principles, and first steps. In *Proceedings of 4th IEEE International Conference on Development and Learning (ICDL 2005)* (pp. 80-86, <http://dx.doi.org/10.1109/DEVLRN.2005.1490948>). Osaka, Japan: IEEE.
- NEC. (2005). *Childcare Robot PaPeRo*. Retrieved March 21, 2012, from <http://en.wikipedia.org/wiki/PaPeRo>.
- Okada, M., Sakamoto, S., & Suzuki, N. (2000). Muu: artificial creatures as an embodied interface. In *SIGGRAPH 2000, the Emerging Technologies: Point of Departure* (p. 91). New Orleans, LA, USA: ACM.
- Robins, B., Dautenhahn, K., Boekhorst, R. T., & Billard, A. (2005). Robotic assistants in therapy and education of children with autism: can a small humanoid robot help encourage social interaction skills? *Universal Access in the Information Society*, 4(2), 105-120, <http://dx.doi.org/10.1007/s10209-005-0116-3>.
- Roscoe, R., & Chi, M. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. *Review of Educational Research*, 77(4), 534-574, <http://dx.doi.org/10.3102/0034654307309920>.
- Ruvolo, P., Whitehill, J., Virnes, M., & Movellan, J. R. (2008). Building a more effective teaching robot using apprenticeship learning. In *Proceedings of the 7th IEEE International Conference on Development and Learning (ICDL 2008)* (pp. 209-214, <http://dx.doi.org/10.1109/DEVLRN.2008.4640831>). Monterey, CA, USA: IEEE.
- Sharkey, N. E. (2008). The ethical frontiers of robotics. *Science*, 322, 1800-1801, <http://dx.doi.org/10.1126/science.1164582>.
- Sony. (1999). *AIBO*. Retrieved March 21, 2012, from <http://en.wikipedia.org/wiki/AIBO>.
- Suay, H. B., & Chernova, S. (2011). A comparison of two algorithms for robot learning from demonstration. In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC 2011)* (pp. 2495-2500, <http://dx.doi.org/10.1109/ICSMC.2011.6084052>). Anchorage, AK, USA: IEEE.
- Tanaka, F., Cicourel, A., & Movellan, J. R. (2007). Socialization between toddlers and robots at an early childhood education center. *Proceedings of the National Academy of Sciences of the U.S.A. (PNAS)*, 104(46), 17954-17958, <http://dx.doi.org/10.1073/pnas.0707769104>.
- Tanaka, F., Fortenberry, B., Aisaka, K., & Movellan, J. R. (2005). Developing dance interaction between QRIO and toddlers in a classroom environment: plans for the first steps. In *Proceedings of the 2005 IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2005)* (pp. 223-228, <http://dx.doi.org/10.1109/ROMAN.2005.1513783>). Nashville, TN, USA: IEEE.

- Tanaka, F., & Kimura, T. (2009). The use of robots in early education: a scenario based on ethical consideration. In *Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2009)* (pp. 558-560, <http://dx.doi.org/10.1109/ROMAN.2009.5326227>). Toyama, Japan: IEEE.
- Tanaka, F., & Kimura, T. (2010). Care-receiving robot as a tool of teachers in child education. *Interaction Studies*, 11(2), 263-268, <http://dx.doi.org/10.1075/is.11.2.14tan>.
- Tanaka, F., Movellan, J. R., Fortenberry, B., & Aisaka, K. (2006). Daily HRI evaluation at a classroom environment: reports from dance interaction experiments. In *Proceedings of the 1st Annual Conference on Human-Robot Interaction (HRI 2006)* (pp. 3-9, <http://dx.doi.org/10.1145/1121241.1121245>). Salt Lake City, UT, USA: ACM.
- Thomaz, A. L., & Cakmak, M. (2009). Learning about objects with human teachers. In *Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2009)* (pp. 15-22, <http://dx.doi.org/10.1145/1514095.1514101>). La Jolla, CA, USA: ACM/IEEE.
- You, Z.-J., Shen, C.-Y., Chang, C.-W., Liu, B.-J., & Chen, G.-D. (2006). A robot as a teaching assistant in an English class. In *Proceedings of the 6th IEEE International Conference on Advanced Learning Technologies (ICALT 2006)* (pp. 87-91, <http://doi.ieeecomputersociety.org/10.1109/ICALT.2006.31>). Kerkrade, The Netherlands: IEEE.
- Yujin Robot. (2007). *iRobiQ*. Retrieved March 21, 2012, from <http://www.irobibiz.com/english/>.

Fumihide Tanaka, Faculty of Engineering, Information and Systems, University of Tsukuba, Tsukuba, Japan. E-mail: fumihide.tanaka@gmail.com

Shizuko Matsuzoe, Graduate School of Systems and Information Engineering, University of Tsukuba, Tsukuba, Japan. E-mail: shi25tty@gmail.com