(Manuscript in review for JHRI, 2012)

Care-Receiving Robot to Promote Children's Learning by Teaching: Field Experiments at a Classroom for Vocabulary Learning

Fumihide Tanaka
University of Tsukuba / JST PRESTO
and
Shizuko Matsuzoe
University of Tsukuba

In contrast to the conventional teaching agents (including robots) that are designed to play the role of human teachers or caregivers, we propose the opposite scenario in which robots are *carereceivers* from children. We hypothesize that by using this care-receiving robot (CRR), we may construct a new educational framework whose goal is to promote children's spontaneous *learning by teaching* through teaching the CRR. In this paper, we describe the introduction of a CRR into a real classroom at an English language school for Japanese children (3–6 years of age) and then conduct an experiment to evaluate if the CRR can promote their learning using English verbs. The results suggest that the idea of the CRR is feasible and that the CRR 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 a CRR for educational purposes.

Keywords: Care-receiving robot (CRR), learning by teaching, child-robot interaction, early childhood education, vocabulary 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; 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 environments (NEC, 2005; Yujin Robot, 2007). Although there still remain many practical issues such as safety concerns, children have generally been shown 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

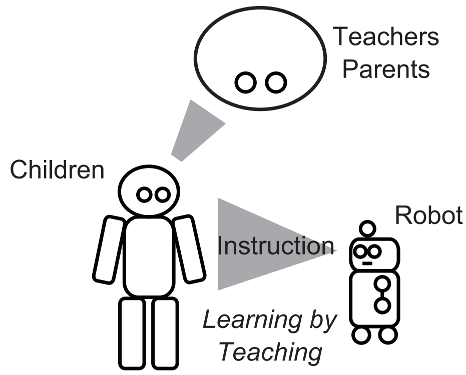

Figure 1. Conceptual diagram involving use of a CRR to promote children's learning by teaching

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 are *carereceiver* from children. By introducing a robot that is somewhat *weaker* 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 intentionally designing a weaker robotic entity has also 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 this concept for educational purposes.

The paper describes the care-receiving robot (CRR) 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 CRR 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 with the CRR in 2011. The purpose was to verify whether it could contribute to the learning of English verbs by children by allowing them to teach the CRR.

The structure of the paper is as follows. First, we explain the basic idea of the CRR (Section 2), and then describe the overall research plan (Section 3) and the experimental field (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 (CRR)

A CRR is a robot that receives care from the people around it. The original concept of the CRR was first proposed in 2009 by the first author of the paper (Tanaka & Kimura, 2009, 2010). Conceptually, the meaning of *care* is broad; thus, we can imagine many types of CRRs. We assume that the most typical scenario using a CRR is for learning support or learning reinforcement for children (Figure 1). In this scenario, teachers/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 CRR, children would, as a result, become highly motivated to complete the topic for learning, providing indirect practice (learning by teaching) for the children.

The idea of the CRR 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 CRR. 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 care of a robot, we could possibly exploit this characteristic to promote children's learning by teaching.

A second background regarding the CRR 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 CRR 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 the CRR is therefore considered to be ethically safer and more acceptable to a wide range of societies (Tanaka & Kimura, 2009, 2010).

Since childhood education is a highly interesting topic to a wide range of people including non-researchers, it is necessary that we carefully report our activities and experimental results to avoid misinterpretation. As will be described in Section 6.2.2, this is the reason why we avoided conducting a typical control experiment, in which the resulting performance after introducing a CRR was assessed on the basis of a comparison between the CRR and human teaching. These types of experiments can be interpreted easily as being supportive of robots replacing human teachers, which contradicts the original philosophy of the CRR. We consider that it is important to recognize that these educational robots are designed to support human teachers and not replace them.

3. Overall Research Plan

The idea of a CRR 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 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 the CRR to promote children's learning by teaching would be feasible and useful.

Thus, to verify the idea of the CRR 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 CRR 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

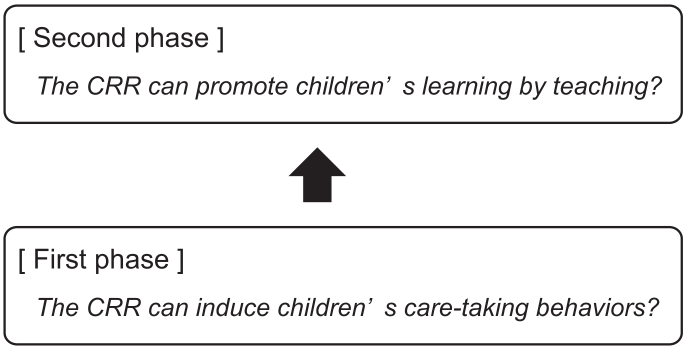

Figure 2. Two phases in verifying the idea of CRR

role of a pilot study for the second phase. The second phase is the main part of the research with the goals of verifying that the educational purpose of the CRR can be implemented in children's real learning activities, and of investigating if it can, in fact, contribute to their learning reinforcement.

4. Experimental Field

The best experimental field for implementing and testing the idea of the CRR is a real classroom within the context of learning activities. Therefore, 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 subjects were children from 3 to 6 years of age who attend 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 is approximately 25 m^2 in size, and all children are 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 CRR using two camcorders (one mounted in the ceiling and the other in a corner at ground level).

5. Initial Pilot Trials

Between September 2010 and February 2011, we conducted initial pilot trials that consisted of introducing a CRR into the classroom for a total of 6 days using 18 subjects. Our goal was to implement the CRR and observe whether it could induce children's spontaneous caretaking behaviors during the context of English instruction in the classroom (the first

Figure 3. Classroom where CRR trials were conducted

phase in our research plan described in Section 3). In this section, we will summarize the findings from the trials.

After studying a typical class flow for the target children (ages 3–6 years), we prepared Aldebaran Robotics' Nao to behave as a CRR by programming the necessary movements and dialogs, 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 subjects interacted with Nao, while 9 different subjects independently interacted with Nao for the 0% condition.

Using videos taken during the trials, we performed a behavioral analysis of the subjects. First, we tried to code as many interaction behaviors of the subjects 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 behavioral video coding was conducted by the second author of this paper whose inter-observer reliability, which was calculated with three external coders, averaged 0.70.

Figure 4 shows the results. Significantly more caretaking behaviors were observed with the error rate 100% condition, which implies that the basic idea of the CRR 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.

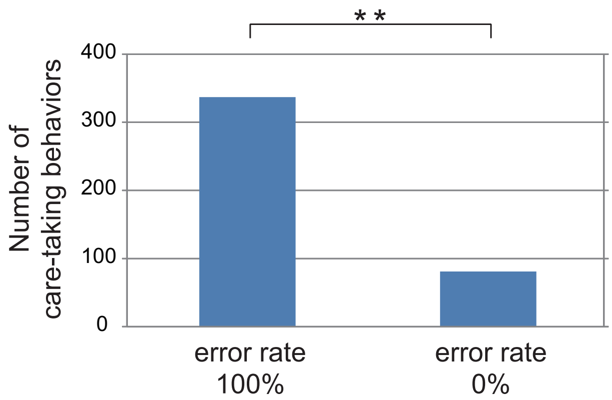

Figure 4. Cumulative total of caretaking behaviors observed for each condition. The Chi-square test was conducted ($\chi^2 = 156.8$, $P < 0.001$).

- The game activities used to introduce the CRR 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 CRR. 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.
- As shown in Figure 5, the most frequently observed behaviors of children toward the CRR were concerned with their "Touching" the CRR. Thus, one promising scenario that could be utilized for teaching the CRR would be a case where children are required to take the CRR by the hand and teach it step by step. As will be explained in Section 6, we decided to adopt this *direct teaching* task for the second phase.
- As described in Section 2, it is very important that our experiments are not interpreted as being supportive of robots replacing human teachers/parents. The CRR proposed here is supposed to be a tool for teachers/parents in educational contexts. Thus, we consider it is appropriate to create an experimental scenario in which the CRR keeps attending a lesson given by a human teacher and its effect is assessed by comparing it with other educational tools within the lesson.
- Due to the classroom capacity and the enrollment of children, it would be difficult to recruit a sufficient number of unbiased subjects to conduct experiments with multiple independent conditions. The maximum population size of subjects expected for the next experiment was 15–20, which also includes children who participated in the initial pilot trials. It would therefore be difficult to control the exposure of each subject to the robot uniformly under all conditions. Therefore, for the actual experiment, it was necessary to give up the idea of dividing subjects into multiple conditions, and instead, consider using a single experimental protocol that is common to all the subjects from which we attempted to extract useful results.

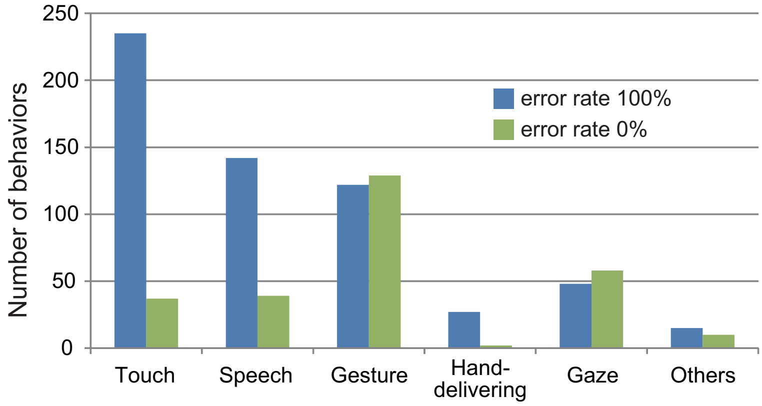

Figure 5. Histograms of 6-categorized behaviors for both conditions (error rates 100% and 0%).

6. Effect on Promoting Learning by Teaching

6.1 Goal of the Experiment

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

Here, the goal is (1) to investigate whether the CRR that we introduce into the classroom can promote children's learning by teaching and (2) to categorize the forms or types of their teaching. It is expected that children will teach the CRR in many ways. Some children may just give verbal instruction, while other children may take its hand and teach 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 CRR in the future.

6.2 Method

6.2.1 Subjects 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 subject was excluded because the subject was too shy to complete the experimental session without the attendance of her parents. As a result, a total of 17 subjects (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 subjects.

Two experimenters were involved, both of whom are authors of this paper. Experimenter #1 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 the CRR 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 CRR. During the experiment, it was fully tele-

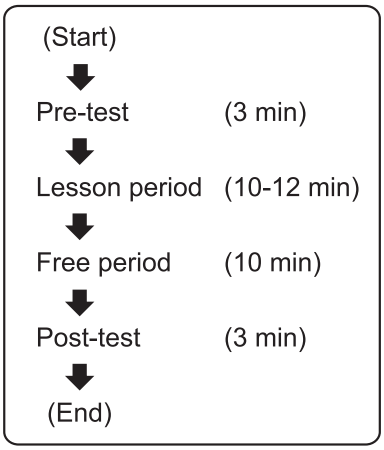

Figure 6. Experimental flow chart

operated (wizard-of-Oz) by Experimenter #2 using a teleoperation interface that had been developed.

6.2.2 Design As emphasized in both Sections 2 and 5, our experimental design had to be such that it is not misinterpreted as supporting the use of robots to replace human teaching. Therefore, we decided to avoid designing a straightforward control experiment in which introducing the CRR was independently compared with regular human teaching. Instead, we decided to adopt an experimental scenario in which the CRR kept attending 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.

Also, because of the difficulty with controlling the amount of exposure received by each subject to the robot in this experiment (as was explained in Section 5), we decided to adopt a single common protocol to all subjects (single condition design), from which we attempted to extract possible results.

In particular, each subject was asked to attend a lesson together with a CRR. 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 CRR 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 subject during the experiment was controlled (equally) for both the cases with and without the CRR.

Then, the experimental hypothesis being tested was that the subjects would learn more verbs with the CRR, which was assessed by a pre/post-test framework.

6.2.3 Procedure In this section, we will explain the experimental procedure in detail. Each session (one subject 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 using cards. 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) represents a verb whose

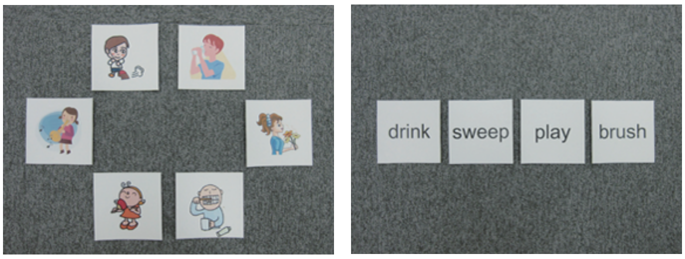

Figure 7. Graphic cards (left) that represent verbs whose spelling is printed on a corresponding word card (right). By conducting a pre-test, four word cards are chosen for the main experiment.

spelling is printed on its corresponding word card (right in Figure 7). Most children in this age group cannot read the verb printed on the word card when viewing it, but they are asked to pick up a corresponding graphic card on the floor when a teacher (Experimenter #1) pronounces a specific verb while showing them the word card.

At the beginning of each session, one subject enters the classroom with Experimenter #1. Then, the pre-test is conducted to identify and select four verbs that are unknown to the subject. The pre-test procedure is as follows. The subject and Experimenter #1 sit in front of six graphic cards placed on the floor (left in Figure 7). Those cards are randomly chosen from a set of 13 cards before each session starts. Experimenter #1, who has the word cards in his hand, chooses one word card from them and then asks the subject to pick up the graphic card from the floor that matches the verb spoken by Experimenter #1 and is shown on the word card. If the subject picks up the correct card, Experimenter #1 says "That is correct" and gives the card pair to the subject. Then, Experimenter #1 adds one new graphic card to the card circle on the floor, maintaining six graphic cards in the group at all times. If the subject picks up an incorrect card, Experimenter #1 just says "That is incorrect," and moves on to the next question. This procedure is repeated until the subject makes a mistake four times, indicating that four unknown verbs have been identified.

Lesson period Then, Experimenter #1 goes to a back room and brings the CRR into the classroom along with four objects that match the four verbs identified during the pre-test. After the CRR is placed at the center of the classroom facing the four objects on the floor (left in Figure 8) and the subject, the lesson begins.

The lesson period is initiated by Experimenter #1 who plays the role of a classroom teacher. The flow chart of the lesson period is shown on the right in Figure 8. The scenario follows the pattern of a regular lesson in the classroom. First, Experimenter #1 would say the name of the subject and the CRR, "Nao," which is teleoperated by Experimenter #2. The CRR would greet the subject cheerfully and introduce itself. The CRR would also attempt to bond with the subject by calling his or her name and shaking hands with the child. Then, Experimenter #1 starts and guides a verb-learning game with the following procedure: (1) From the four word cards obtained during the pre-test, two cards are randomly chosen in advance and marked as "With CRR." (2) From the four word cards, Experimenter #1 picks up one card randomly and asks the child its meaning saying, "Show us how to do <the verb>." Since the verb is still unknown to the subject at this time, the subject cannot answer the

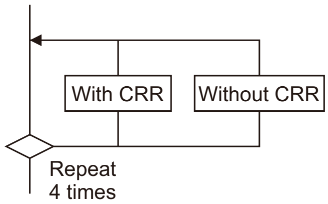

Figure 8. (Left) CRR with four objects on the floor that represent the verbs "drink," "sweep," "play," and "brush." (Right) Lesson period flow chart

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

question correctly. (3) Then, Experimenter #1 shows the correct answer by picking up the corresponding object and pronouncing the verb. If the card is marked "With CRR," the game continues to step (4), otherwise, they return to step (2), repeating the steps with another card. (4) Experimenter #1 turns to the CRR and asks it the same question. The CRR responds "Yes" and then says "Please pick up" while opening its right hand. Experimenter #1 hands the corresponding object to the CRR. The CRR then holds the object and makes a wrong movement. For the wrong movement, we designed four patterns in advance that appear to be completely meaningless and are different from all the verbs used in the experiment. The CRR outputs one of the four patterns each time it makes an incorrect movement. (5) Experimenter #1 says "No, that is incorrect." The CRR responds with "Please teach me." Then, Experimenter #1 takes the CRR by the hand and teaches it step by step (direct teaching). As soon as Experimenter #1 starts direct teaching, Experimenter #2 switches the CRR to a recording mode, wherein every servo movement is recorded in real time. After direct teaching is finished, the CRR plays back the movement correctly while pronouncing the verb. Then, Experimenter #1 says "Yes, that is correct."

Experimenter #1 repeats the procedure for all four word cards to finish the first part of the lesson. The second part of the lesson is performed in a similar manner using the same four word cards. However, in this case, in step (5), Experimenter #1 asks the child to teach the CRR instead of Experimenter #1. Since not all subjects can immediately perform direct teaching, Experimenter #1 may 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 is completed, the subject is told to play freely with the CRR by themselves (Figure 9, right). During that time, Experimenter #1 goes to a bookshelf in the corner and pretends to read books (appearing to ignore the CRR and the subject). However, he is paying enough attention to maintain the safety of the subject. Experimenter #1 uses a stopwatch to ensure that the free period is exactly 10 min long. Meanwhile, the CRR is teleoperated to follow steps (4) and (5) described above. When the subject comes in front of the CRR, it says "Please pick up" while opening its right hand. If the subject hands it an object, the CRR makes a wrong movement at first. Then, if the subject says "No" or exhibits any kind of behavior intending to express a negative response (e.g., shaking his or her head horizontally), the CRR says, "Please teach me" and waits for the subject to teach it (as in step (5) above).

Post-test At the end of each experimental session, a post-test is conducted to evaluate if the subject can answer the questions that he or she could not answer during the pre-test. The procedure for the post-test is identical to that of the pre-test. Note that since no graphic card was used during the lesson period or free period, the subject has to be able to sufficiently generalize the knowledge of each verb such that he or she can 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 subjects. We observe that the subjects picked up more correct graphic cards for verbs With CRR than without. As discussed in Section 6.4, this experiment was not designed to compare human teachers with the CRR, and therefore, the results should not be interpreted as showing the superiority of the CRR to human teaching.

One might feel that this result is natural because children probably spent more time interacting with the CRR than the time spent without the CRR. However, from our viewpoint, it was difficult to know whether interaction with the CRR 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 the CRR can be implemented to supplement the learning of English verbs in a real 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 CRR).

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 subjects 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 CRR 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 is very encouraging to us because this is exactly what we wanted to achieve — promoting children's spontaneous learning.

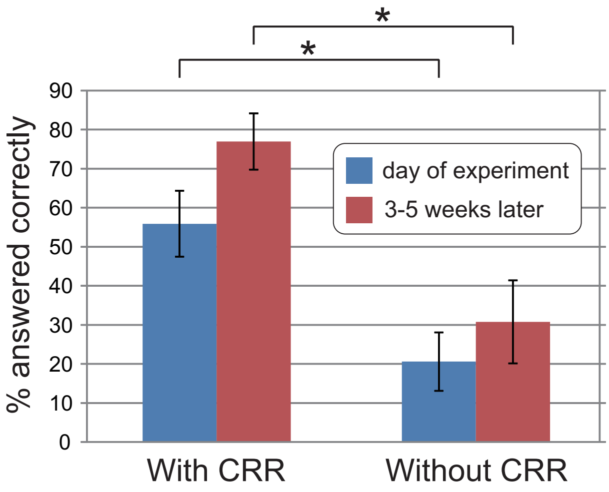

Figure 10. Post-test results: Each bar represents the average percentage of post-test questions answered correctly by 17 (13 at the second post-test) subjects. The post-test was conducted two times: the initial post-test was conducted on the day of the experiment, and the second post-test was conducted 3–5 weeks later. Four subjects joined another session with the CRR 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 With CRR and Without CRR. The differences between the results obtained on the day of the experiment and 3–5 weeks later were not statistically significant.

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 CRR during the free period. Investigating these forms was useful for designing a more effective CRR in the future; thus, we conducted a behavioral analysis of the subjects during the free period. Video coding was conducted by the second author of the 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 subjects with the CRR could be classified into the following three categories.

Direct teaching: The subject takes the CRR by the hand and teaches it step by step (e.g., let the CRR hold a toothbrush and move its hand in front of the face). This also involves more direct behavior by the subjects where he or she acts directly on the CRR (e.g., brush the CRR's mouth while holding a toothbrush in his/her hand).

Gesturing: The subject demonstrates a procedure to the CRR by moving his or her body and showing it to the CRR (e.g., brush his or her mouth with/without holding a toothbrush in his or her hand).

Verbal teaching: The subject gives vocal instructions to the CRR (e.g., just saying "Brush" to the CRR).

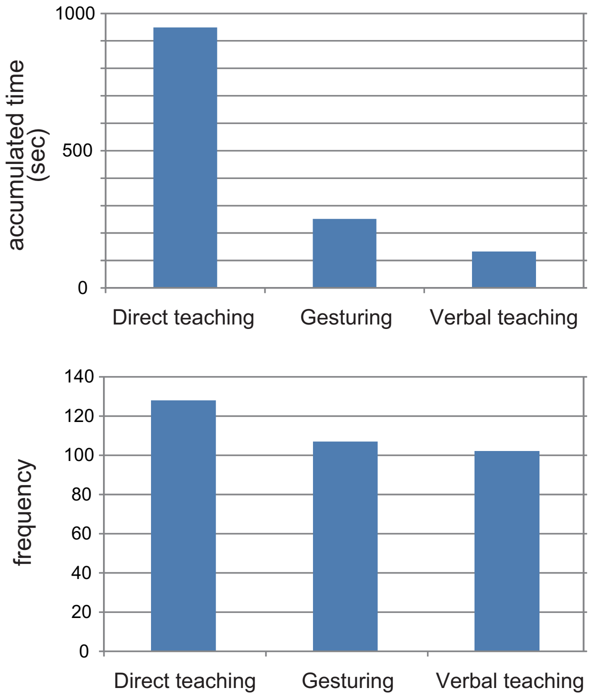

Figure 11. Histograms of the three forms of teaching observed by the 17 subjects with the CRR

Figure 11 shows histograms of the total durations (accumulated time) and total frequencies of subjects' behaviors for the three categories. It was observed that gesturing and verbal teaching behaviors were performed frequently, although Experimenter #1 did not lead the subjects to do so. The results may imply that gesturing and verbal teaching are also easy methods for the subjects 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 subjects. 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 subjects in this experiment. Table 1 summarizes one of the results. Gesturing and verbal teaching may be easily performed by the subjects, but the results indicate that they did not result in the reinforcement of the subjects' 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 subjects 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 subject have a brother or sister?
- Q2: Does the subject live with a pet?

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

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

- Q3: Did the subject 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 subjects have brothers or sisters, 4 subjects live with a dog, 7 subjects had prior experiences interacting with robots at amusement parks, and 14 subjects were said to enjoy teaching or taking care of dolls/humans. ANOVA showed a significant effect of Q4 on the learning performance of the subjects in this experiment. We also felt that the preference for teaching or caretaking seemed to be a general preference for subjects in this age group, which is consistent with the knowledge obtained from our past study (Tanaka et al., 2007).

With regard to the last question on behavioral changes (or any comment) of the subject after the CRR experiment, many parents mentioned that their children had repeated the teaching game often even though there was no CRR 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, eight subjects could correctly answer a question that he or she could not answer at the time of the initial post-test.

6.4 Discussions and Limitations

First, it is important to recognize that the results do not suggest the superiority of the CRR 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 subjects. The results appear to show that the idea of the CRR 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 CRR 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 CRR 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 is 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 subjects in interactions with the CRR. 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 subject'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 subjects who participated in the main experiment, 10 subjects had at least one previous experience of interacting with Nao. This was one of the reasons for adopting a single common protocol for all subjects when designing the experiment (see Sections 6.2.2 and 5). Because the condition was common for all subjects, it becomes valid to argue a difference within the condition (e.g., Figure 10). Nevertheless, we recognize that those results were obtained from subjects whose exposure to the robot Nao was uncontrolled. Third, there was a limitation with regard to the CRR's response to the subjects. As we described in Section 6.2.3, the CRR responded to the subjects' 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 CRR, and then decided to adopt a direct teaching task (Sections 5 and 6). Because the CRR did not respond to the subjects' 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 subjects 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 the CRR.

The most significant disadvantage of the CRR used in the experiment is considered to be the limitation of *its* learning capability. This was the first formal experiment used to assess the CRR for the purpose of children's learning by teaching; thus, we did not incorporate a learning factor into the robot, i.e., the CRR continued to make mistakes (it was only able to repeat a correct movement once just after the subject taught it). This likely caused some subjects to become frustrated or irritated. After the experiment was over, we asked four subjects to attend another session with the CRR. There, we followed the same experimental protocol as in the original experiment. Although the subjects continued to learn new verbs with the CRR, we also noticed signs of habituation from two of the subjects. If the CRR can continue to learn as humans, it may increase the motivation of the subjects. In this case, the most important issue would be the learning speed/dynamics of the CRR. If it learns at a speed that is faster than the subjects, then this may also cause a problem. Therefore, the appropriate learning speed (or the appropriate learning dynamics) of the CRR will become our next important research target.

7. Conclusions

The paper reported an experiment that introduced a CRR into a real 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 CRR 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 the CRR.

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 the CRR 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 CRRs. 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.

8. 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 for their support. We also thank the students of the University of Tsukuba for their help.

References

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.

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). *Papero.* 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.

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.* http://en.wikipedia.org/wiki/AIBO.

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.

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.* http://www.irobibiz.com/english/index.php.

Fumihide Tanaka, Faculty of Engineering, Information and Systems, University of Tsukuba, Japan. E-mail: fumihide@iit.tsukuba.ac.jp

Shizuko Matsuzoe, Graduate School of Systems and Information Engineering, University of Tsukuba, Japan. E-mail: s1130184@u.tsukuba.ac.jp