

Person Identification and Interaction of Social Robots by Using Wireless Tags

Takayuki Kanda¹, Takayuki Hirano¹, Daniel Eaton¹, and Hiroshi Ishiguro^{1&2}

¹ATR Intelligent Robotics and Communication Labs.
Seika-cho, Soraku-gun, Kyoto 619-0288
Japan

²Osaka University
Suita, Osaka 565-0871
Japan

E-mail: kanda@atr.co.jp

Abstract – This paper reports a trial of immersing interactive humanoid robots into a real human society, where they are charged with the communication task of foreign language education. For this purpose, we developed interactive humanoid robots equipped with wireless person identification technology for facilitating interaction with multiple persons. We believe that the ability to identify persons allows more meaningful social relationships between robots and humans to develop. This paper discusses the fundamental mechanisms of the multi-person identification ability, and how the robots established and sustained social relationships among children in an elementary school.

1. INTRODUCTION

The recent development of humanoid and interactive robots such as those described in the studies by Hirai et al. [1] and Fujita [2] has created a new research direction, based on the concept of *partner robots* (robots acting as human peers in everyday life). These robots are capable of effective multi-modal communication with humans, in order to perform a multi-faceted set of tasks together. Clearly, a robot that is skilled at a single or limited set of tasks cannot satisfy the designation of partner. For example, the museum tour guide robot [3] is equipped with robust navigational skills, which are crucial to its role; however, humans do not perceive such a robot as a partner, but merely as a museum orientation tool. While the ability to skillfully perform many types of tasks is a desirable attribute for a partner robot, this alone does not lead a human to consider the robot as their partner. Instead, we believe that to possess human-like body properties and have the capacity for interacting with multiple persons are fundamental requirements for partner robots.

Many humanoid robots have been developed in the past several years. To develop robots that work in our daily life, researchers believe that a humanoid robot body should be used for communication. A human can easily communicate with other humans by making various gestures, and likewise a human-like robot body allows these fundamental, non-verbal communicative modes to be used. Previous research on human-robot communication, which is often motivated by cognitive science and psychology, has found various interactive mechanisms that a robot's body should feature. For example, Scassellati developed a robot as a test-bed for verifying the effect of joint-attention [5]. Matsusaka and his colleagues developed a robot that can gaze at the person who is talking with it [6]. Nakadai and his colleagues developed a robot that tracks a speaking person [7]. Our robots also utilize such body properties for facilitating interaction with humans [8].

With multi-person interaction, however, it is difficult to develop robots that work in daily life, such as in the home and office, using only visual and auditory sensors. With respect to audition, there are many people talking at the same time. With respect to vision, lighting conditions are unpredictable, and the shapes and colors of objects in a real scene are not simple. For these reasons, existing computer vision techniques have difficulty in recognition. A useful human identification system needs to be robust. Mistakes in identification spoil communicative relationships between the robot and humans. For example, if a robot talks with a person and uses another person's name, it will negatively impact their relationship. To make matters worse, robots that work in public places may have to distinguish between hundreds of humans at once, and simultaneously identify the ones nearby. For example, thousands of people work together in office buildings, schools, and hospitals.

To solve the human recognition problem, we utilized wireless sensors. The people who should be identified in a specific environment wear wireless ID tags embedded in nameplates. Recent RFID (radio frequency identification) technology enabled us to use these contact-less ID tags in practical situations. Several companies have already adopted such tags to identify employees. As cellular phones have come into wide use, many people already carry wireless equipment, especially in areas such as Japan, Hong Kong, and Northern Europe. We predict that wireless identification will become the standard for person identification. Using wireless systems, robots have a robust means to identify many people simultaneously.

In this paper, we report on an interactive humanoid robot and its fundamental mechanisms for multi-person identification in the real human world, and an experiment employing this robot in an elementary school, to perform foreign language education. This is the first trial of such an experiment, which applies interactive humanoid robots in a real human society on a long-term. The robot's role is not that of a human language teacher, but instead, it behaves like a foreign child who speaks only the foreign language (in this experiment, English). Our expectation is that the robot's human-like form and behavior will evoke spontaneous communication from the children, which is more than what is possible with computer agent teaching tools or a self-teaching method. This task is motivated by the Japanese weakness in conversational English, which we believe stems from a lack of motivation and opportunity to speak the language.

II. HARDWARE MECHANISM

A. An Interactive Humanoid Robot “Robovie”

Figure 1 displays the humanoid robot named “Robovie.” The robot has human-like expressive abilities and various sensors. The humanoid body, consisting of a head, eyes, arms, etc., produces the body movements required for communicating with humans. The various sensors, such as auditory, tactile, ultrasonic, and vision sensors allow it to autonomously behave and to interact with humans. All the processing and control equipment is contained in the body. This includes a Pentium III PC, which processes sensory data (including image processing and speech recognition) and generates behaviors.

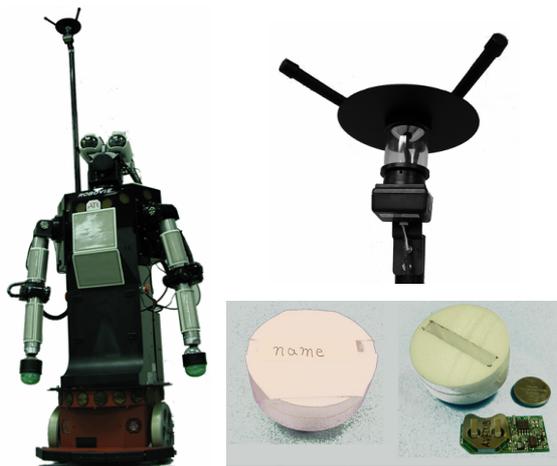


Figure 1: Robovie (left) and Wireless tags

B. Wireless Person Identification System

We adopted the Spider Tag system [11] for wireless person identification. In this system, a tag (shown in Figure 1) periodically broadcasts its ID (identification), which is received by the reader and, in turn, sent to the robot’s computer. The length of the tags is about 6 cm, so they are easily carried. They broadcast their ID over radio at 303 MHz. We attached the reader’s antenna on top of the omnidirectional camera, since the robot’s body radiates a large amount of radio noise. The reader’s attenuation can be adjusted to alter the reception range.

III. SOFTWARE ARCHITECTURE

Figure 2 outlines the software systems which enable the robot to simultaneously identify multiple persons and autonomously interact with them based on a memory for each person. The basic components of the system are “situated modules” and “episode rules.” The robot system sequentially executes situated modules according to execution orders defined by the episode rules. This is an extension of our previous architecture [9]. The basic idea of the system is that a large number of appropriately chosen interactive behaviors generate intelligent and dynamic interactions with humans [10]. We have verified the phenomena through various psychological experimentations, such as that conducted in our 2002 study [8].

With respect to person identification, the architecture utilizes four kinds of databases (DB): Person ID DB to remember internal IDs for each person, episode rules to control the execution orders of the situated modules, “public” and “private” episodes to maintain communication with each person, and “long-term individual memory” to memorize personal information. The “module control” controls the total execution of the situated modules by referring to the episode rules and episodes (history of communication).

Each situated module consists of “communicative units.” The communicative units are principal elements of interactive behaviors, such as eye contact and arm movements synchronized with the utterance. By combining communicative units, the developer can easily and quickly implement new situated modules. “Reactive modules” handle emergencies in both movement and communication. For

example, the robot stops when it collides with a wall, and then returns to the original episode. In the situated and reactive modules, inputs from sensors are pre-processed by sensor modules such as speech recognition. Actuator modules perform low-level control of actuators.

A. Communicative Units

Humans use eye contact and arm gestures for smooth interaction, as discussed in the psychology and cognitive science literature. The communicative unit is an elemental unit for body movement in human-robot communication. Each communicative unit is a sensor-action unit. Specifically, we have implemented ‘eye contact’, ‘nod’, ‘positional relationship’, ‘joint attention (gaze and point object)’, and so forth. Situated modules are implemented by connecting the communicative units with other sensor-action units needed for the behavior, such as a particular utterance, and positional movements.

B. Situated Modules

In linguistics, an “adjacency pair” is a well-known term for a unit of conversation where the first expression of the pair requires the second expression to be of a certain type. For example, ‘greeting and response’ and ‘question and answer’ are considered pairs. Similarly, human-robot interaction can be divided into action-reaction pairs. That is, when a human takes an action toward a robot, the robot reacts to the human’s action; and when the robot takes an action toward the human, the human reacts to the action. In other words, the continuation of the actions and reactions

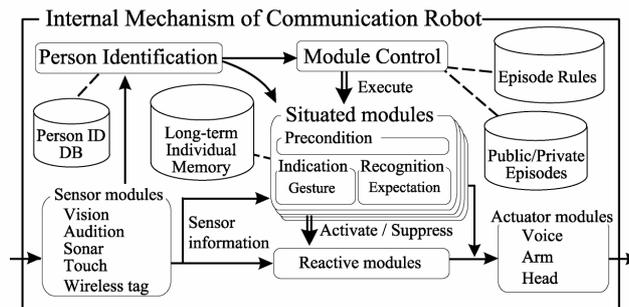


Figure 2: Software architecture

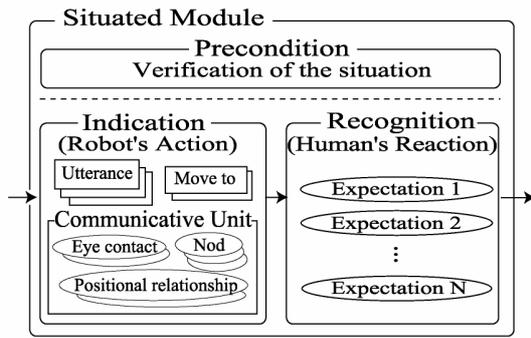


Figure 3: Sited module

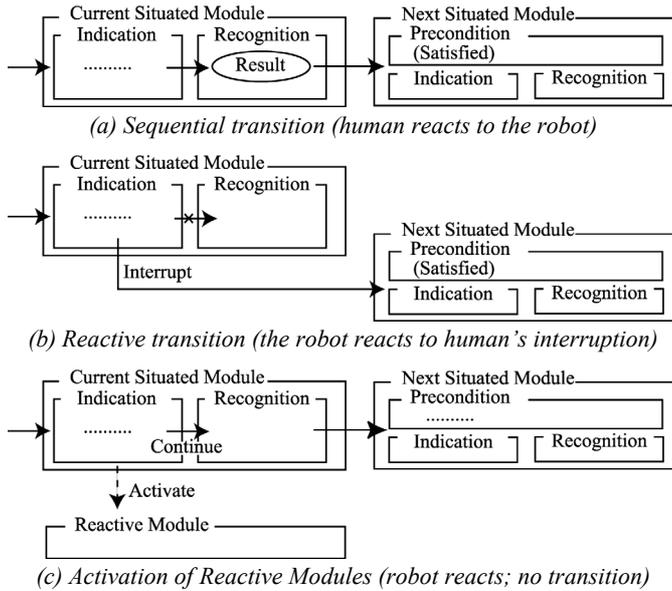


Figure 4: Transition of situated modules

forms the interaction.

Although the number of actions and reactions between humans and robots should be equal, at present, the recognition ability of the robot is not as powerful as the humans'. Therefore, the robot actively takes actions rather than making reactions in order to sustain communication with the human. Each situated module is designed to realize a certain action-reaction pair in a particular situation, where the robot mainly takes an action and recognizes the humans' reaction. Deviation from the basis is treated by reactive transition and reactive modules.

Precondition, Indication, and Recognition Parts

Each situated module consists of precondition, indication, and recognition parts, as shown in Figure 3. By checking its precondition, the robot knows whether the situated module is executable or not. For example, the situated module that talks about the weather by retrieving weather information from the Internet is not executable (precondition is not satisfied) when the robot cannot access the Internet. The situated module that asks to shake hands is executable when a human (a moving object located near the robot) is in front of the robot.

By executing the indication part, the robot takes an action to interact with humans. For example, the robot says

“Let’s shake hands” and offers its hand in the hand shake module. This behavior is achieved by combining communicative units for eye contact and for maintaining positional relationships (move its body towards the human), with speaking the sentence “Let’s shake hands” and making a particular body movement to offer its hand. The recognition part is designed to recognize humans’ reactions affected by the indication part. The situated module creates the particular situation between the robot and the human; therefore, the recognition part can predict certain human responses which are highly probable for the situation. By expecting a specific set of responses, the necessary sensory processing can be tuned to the situation. Thus, the robot can recognize complex human behaviors with simple sensory data processes. When the robot performs situated recognition by sight, we call it “Sited Vision”.

Sequential and Reactive Transition of Sited Modules, and Reactive Modules

After the robot executes the indication part of the current situated module, it recognizes the human’s reaction by the recognition part. It then records a result value corresponding to the recognition result, and moves to the next executable situated module (Figure 4 (a)). The next module is selected using result values and the execution history of the situated modules (episode). This sequential transition is defined by the episode rules.

Episode rules allow for consistent transitions between the situated modules. Sequential transition according to episode rules does not represent all transition patterns needed for human-robot communication. There are two other types of transitions: interruption and deviation. Let us consider the following situation. When two persons are talking, a telephone suddenly rings. They will stop talking and respond to the telephone call. On the robot, interruption and deviation such as this is dealt with as a reactive transition. Reactive transitions are also defined by some episode rules (Figure 4 (b)). If a reactive transition is assigned for the current situation and the precondition of the assigned succeeding situated module is satisfied, the robot stops executing the current situated module and immediately moves to the next situated module.

The reactive modules are also prepared for an interruption, but in this case, the robot does not quit the execution of the current situated module (Figure 4 (c)). Instead, the robot executes the reactive module in parallel with the current situated module. For example, we implemented a reactive module to gaze at body parts of the robot when they are touched. When a human touches the arm of the robot while it is speaking, the robot gazes at the arm while continuing to speak. This is a similar control to the Subsumption architecture [12]. Upper hierarchy modules (situated modules) suppress lower ones (reactive modules).

C. Distinction of Participant and Observers

In linguistics, Clark classifies talking people into two categories: participants and listeners [13]. Participants are mainly speakers and hearers, and listeners just listen to the conversation and take active role in it. Similarly, we classify humans located around the robot into two categories: participants and observers. Since we are concerned only with humans within the robot’s awareness, the categories are similar to Clark’s definitions, but our observer category does not include eavesdroppers (persons listening in without the speaker’s awareness).

The person identification software simultaneously identifies persons and separates them into participant and observer categories. The distance between the robot and the humans also enables the robot to categorize them. As Hall discussed, there are several distances between talking humans [14]. According to his theory, a distance of less than 1.2 m is “conversational,” and a distance from 1.2 m to 3.5 m is “social.” Persons who have met each other for the first time, often talk in the social distance. Our robot recognizes the nearest person within a distance of 1.2 m as a participant and others located within a readable distance of the wireless identification system as observers.

D. Episodes and Episode Rules

Public and Private Episodes

We define an episode as a sequence of interactive behaviors produced by the robot and humans. Internally, it is represented as a sequence of situated modules. There are two types of episodes as shown in Figure 5: “public” and “private.” The public episode is the sequence of all executed situated modules. That is, the robot exhibited those behaviors to the public. On the other hand, the private episode is a private history for each person. By memorizing each person’s history, the robot adaptively behaves to the person who is participating in or observing the communication.

Episode Rules for Public Episodes

The episode rules direct the robot into a new episode of interaction with humans by controlling transitions among situated modules. They also give consistency to the episode. When the robot switches the situated modules, all episode rules are checked with the current situated module and the episodes to determine the next one.

Table 1 indicates the basic grammar of the episode rule. Each situated module has a unique identifier called a Module ID. “<Module ID=result_value>” is the rule to refer to the execution history and the result value of the situated modules, then “<ModuleID1=result_value 1> <ModuleID2=result_value 2>•••” means the referring rule of the previously executed sequence of situated modules (Table 1-1). “<...>|<...>” means a selective-group (OR) of the executed situated modules, and “(...)” means the block that consists of a situated module, a sequence of situated modules, or a selective-group of situated modules (Table 1-2). Similar to regular expression, we can describe the repetition of the block as “(•••){n,m}”, where n gives the minimum number of times matching the block and m gives the maxi-

Table 1: Grammar of episode rules

1. <ModuleID=result_value>...<...>NextModule
 2. (<ModuleID1=result_value1>|<ModuleID2=result_value2>)...
 3. (...){n,m}...
 4. !<...>NextModule
 5. ^<ModuleID=result_value>NextModule
- (1:basic structure of describing executed sequence,
2: “OR”, 3: repetitions, 4: negation of *episode rule*,
5: negation of Module ID and result value)

mum (Table 1-3). We can specify the negation of the whole episode rule with an exclamation mark “!”. For example, “!<...>...<...>NextModuleID” (Table 1-4) means the module of NextModuleID will not be executed when the episode rule matches the current situation specified by “<...>...<...>”. The negation of a ModuleID or a result value is written with a caret character “^” (Table 1-5).

Episode Rules for Private Episodes

Here, we introduce two characters “P” and “O” to specify participation and observation, respectively. If there is a character of “P” or “O” at the beginning of the episode rule, the episode rule refers to the private episodes of the current participant or observers. Otherwise, the episode rules refer to public episodes. If the first character in the angle bracket is “P” or “O,” it indicates that the person experienced the module as a participant or an observer. Thus, “<P ModuleID=result_value>” is a rule to represent “if the person participated in the execution of ModuleID and it resulted in the result value.” Omission of the first character means, “The person participated or observed it.”

Examples

Figure 5 is an example of public and private episodes, episode rulesw and their relationships. The robot memorizes the public episode and the private episodes that correspond to each person. Episode rule 1 and 2 refer to the public episode, which realizes self-consistent behaviors of the robot. More concretely, episode rule 1 realizes sequential transition that “the robot will execute the situated module SING next, if it is executing GREET and it results in Greeted.” Similarly, episode rule 2 realizes the reactive transition “if persons touch the shoulder, the precondition of TURN is satisfied and then the robot stops execution of SING to start TURN.”

There are episode rules that refer to private episodes. Episode rule 3 means that if all modules in the participant’s individual episode are different with GREET, it will execute GREET next. Episode rule 4 represents “if once the person hears the robot’s song, it does not sing the same song for a while.” As in these examples, the episode rule lets the robot adaptively behave toward individuals by referring to the private episodes.

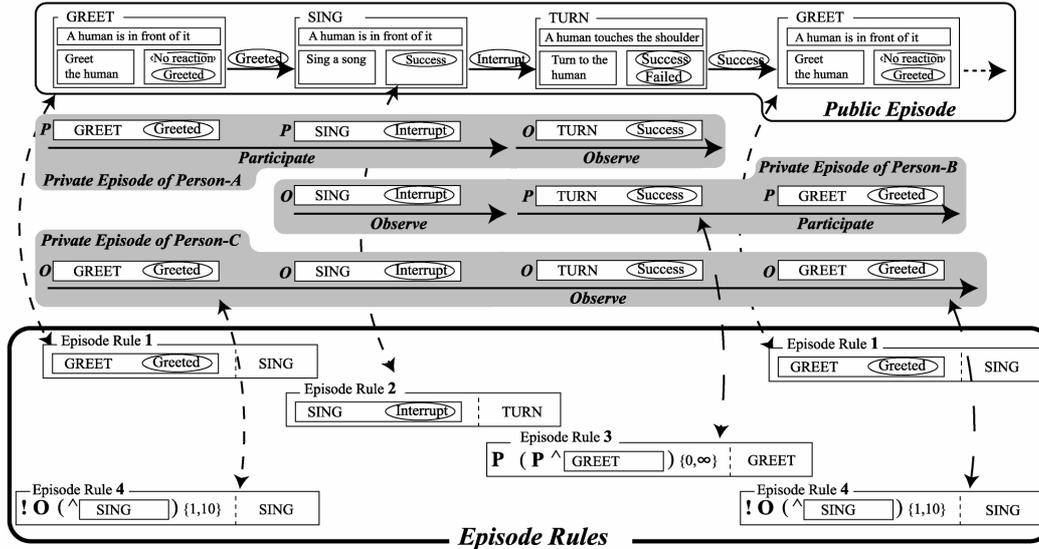


Figure 5: Illustrated example of episodes and episode rules

Episode rules refer to private episodes of participants and observers to adaptively interact with them as well as public episodes to realize consistent behavior.

E. Long-term Individual Memory

The long-term individual memory is a memory related to the situated modules. It is used to memorize local information given by executing particular situated modules as well as personal information such as a person's name. For example, the robot that teaches foreign language needs to manage the students' learning progress, such as the previous answer to game-like questions posed by a situated module. This long-term memory is not only associated with a particular situated module, but is also used for sharing data among several situated modules. For example, although the robot knows the person's name from the beginning, the situated module that calls the person's name will not be executable unless another situated module that asks the name is executed successfully.

IV. IMPLEMENTATIONS

A. Implementation of Interactive Behaviors

We installed this mechanism on "Robovie" for the experiment. The robot's task is to perform daily communication in the same manner as a child. One hundred situated modules have been developed: 70 of them are interactive behaviors such as handshake, hugging, playing paper-scissors-rock, exercising, greeting, kissing, singing a song, short conversation, and pointing to an object in the surroundings; 20 are idling behaviors such as scratching its head and folding its arms; and 10 are moving-around behaviors. For the English education task, every situated module only utters in and recognizes English. The robot speaks more than 300 sentences and recognizes about 50 words.

Several situated modules use person identification. For example, there is a situated module that calls a person's name at a certain distance, which is useful to encourage that person to interact with the robot. Another module plays a body-part game (the robot asks a person to touch its body

parts by saying the parts' names) and remembers the children's answers.

We prepared 800 episode rules to govern the transition among situated modules as follows: the robot sometimes asks humans for interaction by saying "Let's play, touch me," and exhibits idling or moving-around behaviors until a human responds; once a human reacts, it begins and continues the friendly behaviors while the human responds to them. When the human stops reacting, it stops the friendly behaviors, says "good bye" and re-starts its idling or moving-around behaviors.

B. Verification of Read Distance

To verify the performance of the person identification, we performed a preliminary experiment. In the experiment, a subject held a tag at various distances from the robot in an indoor environment. Then, we measured how often the system could detect the tag. As the results show in Figure 6, the system can stably detect subjects within 1.5 m. The reader has eight steps of attenuation that reduce the maximum gain of the receiver by 12.5 % with each step. As the attenuation parameter setting is increased, the readable area decreases. This is represented by the R in the graph, where the gain is R/8 of the maximum. This indicates that we can detect the nearest people. Since the readable area with the attenuation R=8 is smaller than R=5,6,7. We did not use the least attenuation level (R=8) because it seemed the tag system became oversensitive to the noise radiated by the robot itself.

C. Verification of Multiple Person Identification

We also verified the participant-observer distinction with three subjects. The distances between the subjects and the robot are measured by using a motion capture system. The subjects put on the tags and moved around the space, sometimes interacting with the robot.

Figure 7 is the result. The upper graph displays the distance between the three subjects and the robot. The lower graph is for the detected person. The bold line indicates when the subjects were detected as participants, and the fine line indicates when they were detected as observers. The subjects within 1.2 m (the conversational distance among adults) are always detected, and the nearest subject is considered as the participant. The robot also detected almost all subjects within 3.0 m. The time needed to detect a person was a little bit slow since the attenuation parameter was frequently changed to categorize the participant. The delay is about 10 seconds. However, this is sufficient for the robot, because it is on the same order as the execution time of each interactive behavior.

Through these two experiments, we verified the basic performance of the system for interacting with multiple people.

V. EXPERIMENT IN AN ELEMENTARY SCHOOL

A. Settings

We performed two sessions of the experiment in an elementary school in Japan, where each session lasted for two weeks. The subjects were the students of three first grade classes and three sixth grade classes. There were 119 first grade students (6 - 7 years old, 59 male and 60 female) for the first session and 109 sixth grade students (11-12 years old, 53 male and 56 female) for the second session. Each session encompassed nine school days. Two identical robots were put in a corridor connecting the three classrooms. Children could freely interact with both robots during recess. Each child had a nameplate with an embedded wireless tag so that each robot could identify the child during interaction.

B. Results for Long-Term Relationships

First, we analyzed the changes in relationships among the children and the robots during the two weeks for the first grade class. We divided the two weeks into the following three phases: (a) first day, (b) first week (except first day), and (c) second week.

(a) First Day: Big Excitement

On the first day, up to 37 children gathered around each robot (Figure 8-left). They pushed one another to gain position in front of the robot, tried to touch the robot, and spoke to it in loud voices. Since the corridor and classrooms were filled with their loud voices, it was not always possible to understand what the robots and children said. It seemed that almost all of the children wanted to interact with the robots. There were many children watching the excitement around the robots and they would often join the interaction by switching places with children near to the robot. In total, 116 students interacted with the robot out of the 119 students on the first day.

(b) First Week: Stable Interaction

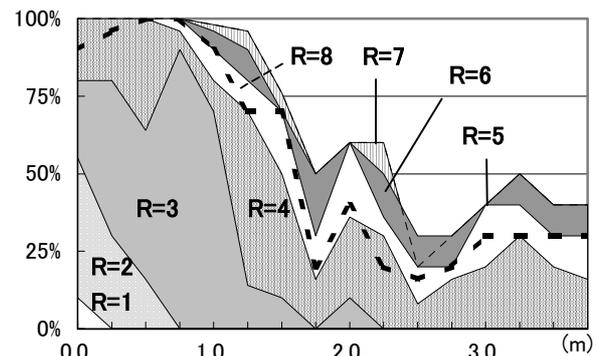


Figure 6: Read distance with different attenuation. In the graph, R indicates the attenuation parameter, the vertical axis corresponds with the rate that the robot found tags, and the horizontal is the distance from the robot.

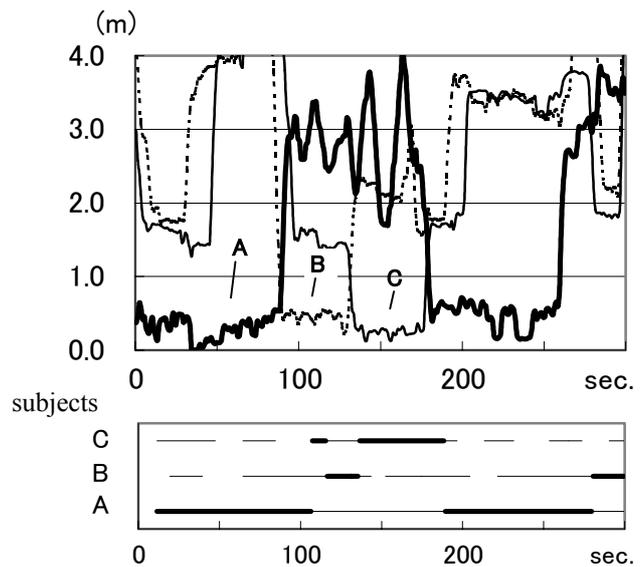


Figure 7: Performance of the detection and identification upper: transition of subjects' distance, lower: the detected participant (bold line) and observers (fine line) by the wireless tag system

The excitement on the first day soon quieted down. The average number of simultaneously interacting children gradually decreased (graph in Figure 10-upper). In the first week, someone was always interacting with the robots so the rate of vacant time was still quite low. The interaction between the children and the robots became more like inter-human conversation. Several children got in front of the robot, touched it, and watched its response.

(c) Second Week: Satiation

Figure 8-right shows a scene at the beginning of the second week. It seemed that satiation had occurred. At the beginning, the vacancy time around the robots suddenly increased, and the number of children who played with the robots decreased. Near the end, there were no children around the robot during half of the daily experiment time. On average there were 2.0 children simultaneously interacting with the robot during the second week. This seemed to be advantageous to the robot since it was easy for it to talk with a few children simultaneously. The way they played with the robots seemed similar to the play style in the first

week. Thus, only the frequency of children playing with the robot decreased.

Comparison with Sixth Grade

Regarding the sixth grade class, there were at most 17 children simultaneously around the robot on the first day, as shown in Figure 9-left. It seemed that the robots were less fascinating for sixth graders than for first graders. Then, similar to the first grade, the vacant time increased and the number of interacting children decreased at the beginning of the second week (Figure 10-bottom). Therefore, the three phases ‘first day, first week, and second week’ exist for the sixth grade students as well as the first grade students. In the second week (Figure 9-right), the average number of simultaneously interacting children was 4.4, which was larger than for the first grade. This is because many sixth grade students seemed to interact with the robot while accompanying their friends, which will be analyzed in a later section.

The results suggest that, in general, the communicative relationships between the children and the robots did not endure for more than one week. However, some children developed sympathetic emotions for the robot. Child A said, “I feel sorry for the robot because there are no other children playing with it,” and child B played with the robot for the same reason. We consider this to be an early form of a long-term relationship, which is similar to the sympathy extended to a new transfer student who has no friends.

Observation of Children’s Behavior

By observing their interaction with the robots, we found several interesting cases.

- Child C did not seem to understand English at all. However, once she heard her name uttered by the robot, she seemed very pleased and began to often interact with the robot.
- Children D and E counted how many times the robot called their respective names. D’s name was called more often, so D proudly told E that the robot preferred D.
- Child F passed by the robot. He did not intend to play with the robot, but since he saw another child G playing with the robot, he joined the interaction.

These behaviors suggest that the robot’s behavior of calling names significantly affected and attracted children. Additionally, observation of successful interaction is related to the desire to participate in interaction.

C. Results for Speaking Opportunity

During the experiment, many children spoke English sentences and listened to the robot’s English. We analyzed the spoken sentences. Mainly, it was simple daily conversation, and the robot used basic English phrases, such as “Hello,” “How are you,” “Bye-bye,” “I’m sorry,” “I love you,” and “See you again.” Since the duration of the experiment was different each day, we compared the average number of English utterances per minute. Figure 11 illustrates the transition of the children’s English utterances for



Figure 8: Scene of the experiment for first graders
(see video)



Figure 9: Scene of the experiment for sixth graders

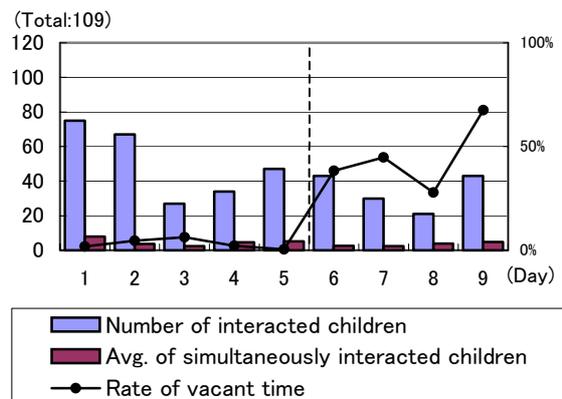
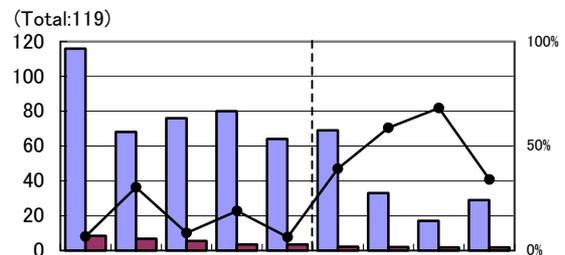


Figure 10: Transition of interaction with children

both the first grade and sixth grade students. In the first grade, there were 4.84 – 5.84 utterances per minute during the first three days. This rate gradually decreased as the vacant time increased. As a result, 59% of English utterances occurred during the first three days.

In the sixth grade, there were about 1.04 – 2.72 utterances per minute during the first week, and this decreased to 0.76 – 1.14 during the second week. This also seems to correspond to the vacancy time of the robot. That is, children talked to the robot when they wanted to interact with the robot. After they became used to the robot, they did not speak or even greet it very often.

VI. DISCUSSIONS AND CONCLUSION

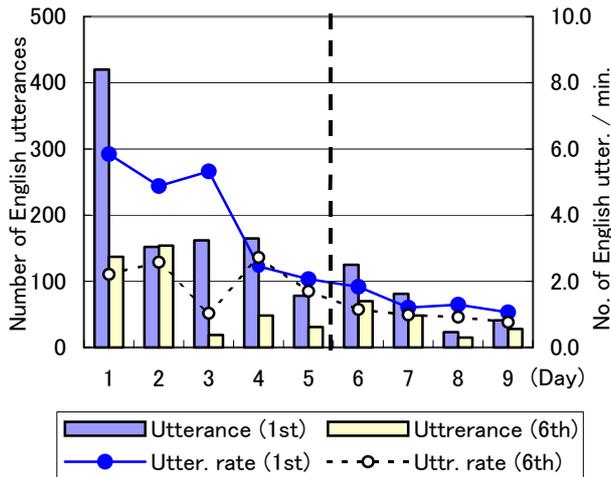


Figure 11: Transition of children's English utterance (Utterance means the total number of utterance every child among the 1st or 6th grade made, and Utterance rate means the average of the total number per minute)

With wireless tags embedded in nameplates, the developed robots socially interacted with multiple persons simultaneously. The preliminary experiments verified the basic ability of the person identification and distinction of participant. Then, these robots were used in an explorative experiment at an elementary school for foreign language education.

The experiment results indicate the robot's interactive ability in real human society as well as the effectiveness of the wireless tags distributed among humans. Regarding the interactive ability, the experiment results, such as the children's English utterances toward the robots, show the possibility of applying these interactive robots to communicative tasks in real human society. Meanwhile, we feel that the most difficult challenge in this experiment was coping with the loss of desire to interact with the robot on a long-term scale. It is necessary to create new mechanisms for long-term interaction.

The wireless tags proved excellent both at generating the robots' behaviors and for analyzing the humans' social behaviors. Processing complex sensory data from a real human environment was the other big challenge of the experiment. In the classroom, many children ran around and spoke in very loud voices; however, the wireless person identification worked well. For example, the name-calling behaviors impressively attracted children. Moreover, it was quite helpful for analysis after the experiments; the wireless system recorded the children's interaction log, which enabled us to evaluate the long-term aspects of their interaction.

VII. ACKNOWLEDGEMENT

This research was supported by the Telecommunications Advancement Organization of Japan.

VIII. REFERENCES

- [1] K. Hirai, M. Hirose, Y. Haikawa, and T. Takenaka, The Development of Honda Humanoid Robot, *Proc. IEEE Int. Conference on Robotics and Automation*, 1998.
- [2] M. Fujita, AIBO; Towards the Era of Digital Creatures, *Int. J. of Robotics Research*, 20(10):781-794, 2001.
- [3] W. Burgard, A. B. Cremers, D. Fox, D. Hähnel, G. Lakemeyer, D. Schulz, W. Steiner, S. Thrun, The Interactive Museum Tour-Guide Robot, *Proc. of National Conference on Artificial Intelligence*, 1998.
- [4] H. Asoh, S. Hayamizu, I. Hara, Y. Motomura, S. Akaho, and T. Matsui, Socially Embedded Learning of the Office-Conversant Mobile Robot Jijo-2, *Proc. of Int. Joint Conf. on Artificial Intelligence*, 1997.
- [5] B. Scassellati, Investigating Models of Social Development Using a Humanoid Robot, *Biorobotics*, MIT Press, 2000.
- [6] Y. Matsusaka, et al., Multi-person Conversation Robot using Multi-modal Interface, *Proc. World Multiconference on Systems, Cybernetics and Informatics*, Vol.7, pp. 450-455, 1999.
- [7] K. Nakadai, K. Hidai, H. Mizoguchi, H. G. Okuno, and H. Kitano, Real-Time Auditory and Visual Multiple-Object Tracking for Robots, *Proc. Int. Joint Conf. on Artificial Intelligence*, pp. 1425-1432, 2001.
- [8] T. Kanda, H. Ishiguro, T. Ono, M. Imai, and R. Nakatsu, Development and Evaluation of an Interactive Humanoid Robot Robovie, *Proc. IEEE Int. Conf. on Robotics and Automation*, 2002.
- [9] H. Ishiguro, T. Kanda, K. Kimoto, and T. Ishida, A Robot Architecture Based on Situated Modules, *Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp.1617-1623, 1999.
- [10] T. Kanda, H. Ishiguro, M. Imai, T. Ono, and K. Mase, A Constructive Approach for Developing Interactive Humanoid Robots, *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2002.
- [11] Spider Tag, <http://www.rfcode.com/includes/spiderTags.pdf>
- [12] R. A. Brooks, A Robust Layered Control System for a Mobile Robot, *IEEE J. of Robotics and Automation*, 1986.
- [13] H. H. Clark, *Using Language*, Cambridge University Press, 1996.
- [14] E. Hall, *The Hidden Dimension*, Anchor Books/Doubleday, 1990.