4-2 Cognitive Mechanisms of Preverbal Communication

KOZIMA Hideki and NAKAGAWA Cocoro

The goals of "understanding humanity" and "humanizing robots" tightly relate to each other. Infanoid Project has related robotics to human sciences in order to understand the underlying mechanism of social communication specific to humans and some species of primates. Early communication between a child and caregiver is mainly embodied through touch and eye contact. By investigating the developmental mechanism of the preverbal interaction, especially through our longitudinal observation of children with communication disorders, we investigated the core human communication capabilities and design principles for future info-communication systems with which we can make symbiotic relationships.

Keywords
Non-verbal communication, Eye contact, Joint attention, Interaction, Developmental disorders

1 Introduction

The activity of communication is supported by mutual access of minds between self and other Various social activities, cooperative or competitive alike, are maintained by reading or manipulating intentions (what the other person is trying to do) and/ or emotions (how the person evaluates a situation), thereby predicting or inducing actions in the other party[1]. Can such communication, which links the minds of humans, also be established between robots and humans? Under what conditions will humans feel that a robot has its own mind? What functions will robots need to be equipped with to be able to access the human mind? Answers to these and related questions will be essential if we are to create robots in true symbiotic relationships with humans.

The acquisition of the ability to communicate among humans first begins through eye contact and the exchange of touch, voice, and expressions between a child and a caregiver, typically the mother. These actions eventually progress to the stage of exchanged attention, including pointing and joint attention (in which the two are looking at the same object)[2][3]. Through such actions, the child and the caregiver establish a common perception of the surrounding world, and by exchanging information on each other's mental state — how the other perceives the world — using expression or voice, the existence of the other person and his or her relationship to the surrounding world are shared. In this way, a child learns to derive meaning from and ascribe value to his or her surroundings, eventually developing into an autonomous being with the acquisition of language skills and culture.

In the present paper, we will report on the Infanoid Project (http://www.infanoid.com), in which the above mechanism of communication development was used as a model for the design of an info-communication terminal that will instill in the user the feel of another's "mind", and that will enable the user to establish communications with this other mind.

First, we will focus on communication development in children to establish the theory that attention and the exchange of emotion is essential to "mental attribution"[4][5]. Next, we will introduce the "Infanoid" and "Keepon" robots developed by the authors, as an example of the elemental technology used to link to the user's attention and emotions. Discussions follow on the process of establishing mental attribution with respect to the robot based on the observation of interactions between children and robots in actual application in autism therapy[6][7]. Finally, we will propose a direction for future info-communication terminals based on the results of these discussions.

2 Findings from communication development in children

The foundation for empathetic understanding of another person's intentions or emotions — in other words, the other's "mental state" — is established during the first year after birth[2][3]. From there, various skills in social communication (such as language, use of tools, and culture) all stem from this basic ability to empathize with and understand others. Here we will examine the interactions that take place between the mother (the main caregiver) and her child (an infant) during this one-year period[1].

- From birth to 3 months of age: Eye contact, accompanied by exchange of voice and expression, is established. By looking at and away from the infant's eyes or by changes in the voice or expression, the rhythmic turn-taking or alternation of the sender/ receiver role (similar to the alternation of communication) is established between the mother and child. However, at this stage, the temporal pattern is created through the caregiver's reading of the infant's response pattern (Fig. 1a).
- From 3 months to 9 months of age: The caregiver begins actively to read and respond to the desires or negative emotions of the infant. Although asymmetrical, the interaction created resembles to a degree the joining of minds. Through such interactions, the infant gradually learns to predict the actions of the caregiver, and eventually actions that reflect such predictions begin to appear in the infant (Fig. 1b).
- From 9 months of age: "Joint attention" will begin to function, in which the infant will turn to view the same object as another person when eyes or fingers are used to indicate or point to the object. By giving joint attention to an object, the voice, expression, and body motions targeted at the object may also be referenced, and the meaning, value, and the method of handling the target may be shared between the mother and child. Through such target-centered interaction, the infant acquires experience in superimposing or comparing the self and the non-self (Fig. 1c).

In this manner, the infant and caregiver link attention and emotions through eye contact and joint attention, referencing and sharing relationships to various targets of interest. The object of attention may be a physical one such as a toy, or an event such as the presence

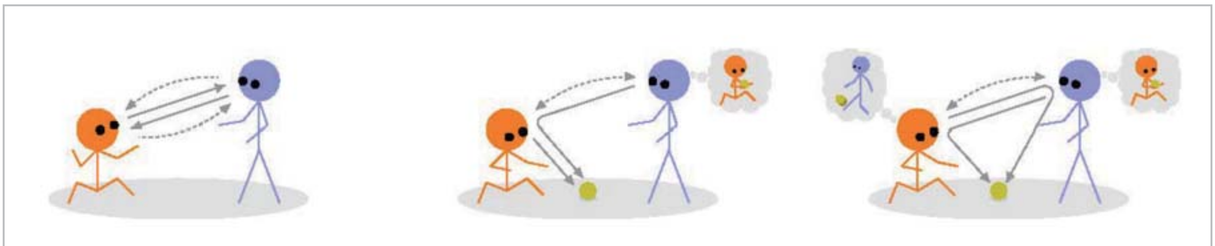

Fig. 1 *Development of infant-mother interaction*

(a) left: establishment of eye contact, (b) center: reading of desires or negative emotions of the infant by caregiver, (c) right: sharing of perceived target by joint attention or motion of, or relationship to, an object. By referring to the emotions (joy, surprise, discontent) roused by such targets, the infant is able to understand in an empathetic manner how the other person perceives the object or event. Through the establishment of such empathy, the infant learns from others — particularly through his or her attachment to the caregiver — how to respond to or handle various targets.

3 Sharing attention and emotion with robots

The section above on the development of mother-infant interactions elucidates the role of the interactions of attention and emotion through eye contact and joint attention in empathetic communication. Is it not reasonable to conclude that by giving robots the functions of eye contact and joint attention, empathetic communication between such robots and humans can be established? We developed two robots based on this concept: the "Infanoid" child-like robot[8][9] and "Keepon", a creature-like robot[9].

3.1 The "Infanoid" child-like robot

Infanoid (Fig. 2a)[8][9] is a robot with an upper half of a body, approximately the size of a 4-year old (at a seated height of 480 mm), and is equipped with 29 monitors and numerous sensors. There are five fingers on each hand, which is capable of pointing at something and grasping toys. There are two (left and right) eyes on its head, and it can perform quick saccade in the left-right and up-down directions. It is also capable of forming various "expressions" by moving its eyebrows and the upper and lower lips.

The eyeball is equipped with a wide-angle video camera (horizontal angle of view 120˚) for peripheral view and a telephotographic video camera (horizontal angle of view 25˚) for foveal view. The images acquired by the cameras are processed by computer to perform detection and tracking of human faces and toys. The two microphones set on the left and right sides of the head at the positions of the ears capture human voice, and the robot is capable of parroting what it hears by the extraction of prosodic (intonation) or phonemic (word fragment) information and input of this information into a voice synthesizer unit.

3.2 The "Keepon" creature-like robot

Keepon (Fig. 3a)[9] has a yellow snowman-shaped body (120 mm in height and 80 mm in diameter). The top ball corresponds to the head, and is equipped with two eyes (wide-angle video cameras) and a nose (in

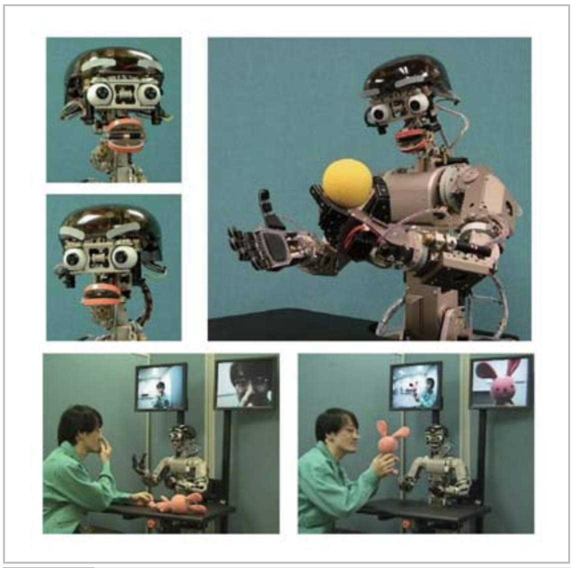

Fig.2 *Child-like "Infanoid" robot*

(a) top: expression and body motion, (b) bottom: eye contact and joint attention

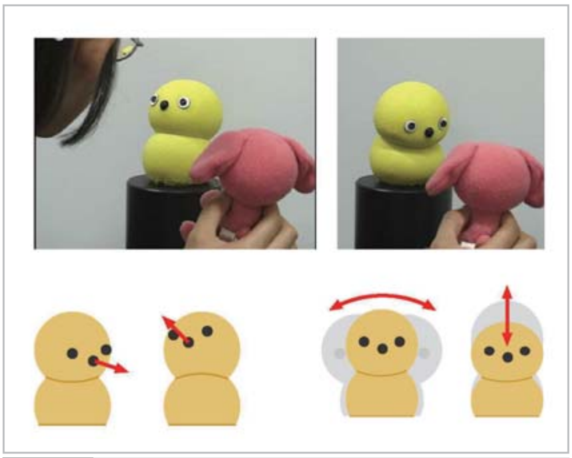

Fig.3 *Creature-like "Keepon" robot*

(a) top: eye contact and joint attention, (b) bottom: expressions of attention and emotion fact, a microphone). The robot's audio-visual perception is nearly at the same level as that of the Infanoid. The bottom ball serves as the body, and the electronic circuits and motors are all stored in the black columnar container at the base. Wires stretching from the base to the body control Keepon's movement. Only the framework, video cameras, and microphone are contained in the body; the head and body are both soft and deform to the touch, so the device is safe for children.

The simplicity of the body design comes from Keepon's limitation to two types of movements (Fig. 3b): (1) expression of attention — the head is moved up and down or to the sides to direct the face (in other words, the eyes) to the target in the surroundings, to relay information on the perceived target to the user; and (2) expression of emotion — the line of sight is maintained while the body rocks from left to right or bobs up and down to relay information on the machine's "mental state" (such as joy or excitement) to the user.

3.3 Mechanism for linking attention to emotion

Both robots are capable of exchanging information on attention and/or emotion with the user (especially infants) by establishing eye contact or joint attention (Figs. 2b & 3a). "Eye contact" is realized by detecting the frontal view of the human face from the video camera image and by turning the eyes (or both the head and hands) to keep track of the face.

"Joint attention" is established by identifying the direction of the human face based on the video image, searching for the target (such as a toy) in the detected direction, and then turning the eyes (or both the head and hands) to the identified target. The eye contact and joint attention functions allow users to identify on an intuitive level what the robot perceives.

4 Interactions with children

Will users feel the presence of a "mind" in robots having the functions of eye contact and joint attention? We chose children as subjects in this experiment, based on the assumption that children have little cultural bias with respect to robots, observing the voluntary actions the children displayed toward these robots.

4.1 Interactions with Infanoid

Children primarily in early childhood (average age of 5 years) were individually presented face-to-face with Infanoid without any advance knowledge or instructions. Infanoid's movements were set to automatic mode, in which the robot alternates between eye contact and joint attention to a toy. When required, the direction of Infanoid's attention (via line-of-sight or pointing) was remotely controlled by the experimenter. At first, the child was placed alone with Infanoid, and then joined by his or her caregiver 3 to 4 minutes later. The children enjoyed the interactions with the robot for approximately 30 minutes on average, until they displayed signs of boredom or weariness.

Based on these observations, the following changes were seen in the voluntary interactions between the child and the Infanoid with time (t), as shown in Fig. 4a.

- Phase Ⅰ (t = 0 - 3 min.): In the first 3 minutes, when the children are alone with Infanoid, they remain tense and keep their gazes fixed on the robot. The children appear puzzled as to how to act toward Infanoid; their gazes remain fixed and they seem to appear confused when the Infanoid shifts its line of sight. The children regard the robot as a moving object.
- Phase Ⅱ (t = 3 - 10 min.): The children use their caregiver as a secure base from which they can explore the Infanoid's response pattern. The children approach the robot with various actions, such as moving a toy in front of Infanoid's eyes or touching its hands. Whenever they discover a new response, they call out to their caregiver in a

Fig.4 *Interaction between children and robot*

(a) top: with Infanoid, (b) bottom: with Keepon (In both figures, indicated from left to right are the relationships established as a “moving object”, a “system with perception”, and an “agent having a mind”, respectively.)

social referencing action. At this stage the children regard the Infanoid as a system equipped with perception.

- Phase Ⅲ (t = 10 min. and later): Children attempt to read Infanoid’s emotions as they show or hand over a toy to the robot. They question Infanoid (“Which one do you like?”) or issue commands (“Hold it tight!”), deepening their social relationship with Infanoid under the supposition that the robot has intentions and desires. The children see the Infanoid as an agent, with a mind.

In this way, the manner in which the relationship formed between the children and Infanoid changed dynamically with time, as the robot changed for the children from an object to a system and finally to an agent[4][5].

4.2 Interactions with Keepon

Next, children in infancy (Under 1 year, 1- and 2-year-olds) were selected for a face-to-face interaction with Keepon accompanied by their caregiver, without any advance knowledge or instruction. Keepon was set to manual operation mode, in which the direction of the head and emotional expressions are remotely controlled by the experimenter. Keepon was controlled so that its gaze remained fixed either on the face of the infant or on a toy, and whenever there were any actions from the infant (such as eye contact or touch), the robot was made to display a positive emotional expression, such as bobbing the body up and down several times. Observation of these interactions continued until the infant displayed signs of boredom or weariness, and lasted for 10 minutes on average.

Based on these observations, the following trend was observed in the development of interactions between the infant and Keepon, as shown in Fig. 4b.

- Stage Ⅰ (0 to 2 year old): For 1 to 2 years-olds, infants treated Keepon as a moving object and kept their distance as they watched it, and, for 0 to 1 year-olds, infants used their hands and/ or mouths to feel it. They displayed almost no signs of confusion or tension, but remained indifferent to the direction of Keepon’s eyes. They did, however, react positively to emotional expression (particularly with the bobbing of the body). No pre 1-year-olds passed beyond this stage.
- Stage Ⅱ (1 to 2 years old): The infants regard Keepon as a system having perception. They explore Keepon’s response pattern by approaching it (touching or moving toys in

front of Keepon's eyes) while moving back and forth to and from the safe base of their caregiver. Compared to Stage I, the infants appear to keep a little more distance. No.1-year-olds passed beyond this stage.

• Stage Ⅲ (2 years old and above): The infants treat Keepon as an agent with a mind. They not only show it toys and greet it ("Hello!") with both words and body language, but display social and prosocial behavior, such as patting the Keepon's head when the robot displays an appropriate response.

In this way, a significant difference was observed in the manner in which infants of different ages related to Keepon — as an object among the youngest, then as a system and finally as an agent among the older groups[4][5].

5 Application to autism therapy

The interaction experiment introduced in the previous section was based only on ad-hoc observations and the results are not in themselves sufficient in understanding the establishment of communication skills in children. Accordingly, a long-term visit was made to a therapeutic daycare center for children with developmental disorders such as autism, with the aim of conducting a more practical experiment on communication development in the field[5]. At the daycare center, we longitudinally observed the interactions between children and a robot within the context of regular therapeutic activities[7].

At this daycare center, the children (mainly 2-4 years old) and their mothers and the therapists participate in various free play and group play. The behavior of the children is slowly given social meaning through these varied and dynamic, yet extremely routine, exercises.

5.1 Keepon at the therapeutic facility

We were given permission to place the creature-like robot Keepon in the playroom of this daycare center (Fig. 5a). During the therapy session, which lasts for approximately three hours, the children are allowed to play with Keepon whenever they like. In the free-play time, Keepon is one of the various toys available to the children. In the group-play time, Keepon is moved out of the way (to a corner of the playroom), but children who have become bored or weary of the group activity may return to play with Keepon.

The Keepon robot in the playroom is installed inside a plastic cover approximately 25-cm high (Fig. 5b). Batteries and wireless units are stored within the device to allow an operator in another room to control Keepon remotely in manual operation mode. Keepon was controlled so that its gaze remained fixed either on the face of the infant or on a toy, and in response to any actions on the part of the children (such as eye contact or touch), the robot was made to display positive emotional expressions, such as bobbing the body up and down several times while emitting a popping sound.

Fig.5 *Keepon in a therapeutic facility*

(a) left: scene from inside the playroom, (b) right: wireless version of Keepon

5.2 The children as seen from keepon

Observation of the interactions between Keepon and the children inside the playroom began in October 2003. Until the completion of the project in March 2006, approximately 100 sessions (with a total of over 700 person-session) were observed. To our knowledge there are no previous reports of such a longitudinal study of interactions between children and robots.

These observations allowed us to confirm first-hand the interaction between children and Keepon through the eyes of the latter, and we were able to record and analyze the expressions, body motions, voice, and words of the children in the various interactions from the first-person perspective of Keepon[7]. This first-person perspective is actually the subjective perspective of Keepon's operator, but nevertheless the interactions with the children through the simple Keepon structure (based on robotic motions) have been completely recorded and can be played back. In other words, Keepon may be regarded as a tool featuring both the phenomenological subjectivity of a self that interacts with the children, as well as the objectivity inherent in the ability of others to experience these interactions vicariously[4].

Various forms of interactions with Keepon were seen in the children through Keepon's eyes. Sometimes the children even displayed expressions that they normally would not show to others (even to their mothers). They also performed actions to assist Keepon, such as placing a hat on its head or pretending to feed it. Overall, the following observations were made[7].

- The fact that Keepon was neither complex human nor a simple toy enabled children lacking in interpersonal communication skills to approach Keepon with a sense of security and curiosity.
- The children established not only direct interaction with Keepon, but they also appeared to develop some extent of interpersonal relationship driven by a desire to share the joy and surprise experienced through this interaction with others (mothers, therapists, and other children).
- The form of interaction with Keepon and its development differed qualitatively between individuals and indicates that unique pathways exist in the development of each individual, which may override general diagnostic conditions such as pervasive developmental disorders (PDD), autism, or Down's syndrome.

We are now providing the "stories" of each child as seen from the eyes of Keepon to the parents and therapists in hopes that these scenes may contribute to improving the services at the therapeutic facilities or assisting the parents in raising each child.

6 Conclusions

Children first perceive robots as moving objects. However, when these robots are given the functions of eye movement and expressions, children begin to see an autonomous entity — life-likeness, in other words — in the robot. The child sees that robots can perceive humans and toys — "just like I do" — and that they can respond emotionally. Such changes in perception transform the robot from a formerly simple object into a system that can perceive and respond to its surroundings. Next, the children see that the attention and emotion of robots are associated (related temporally and spatially) with their own actions. This drives them to approach the robot with various actions, such as handing toys, patting the head, and talking to the robot. They discover and give social meaning to the responses from the robot. In this way, a joining of attention and emotion is felt between the robot and the child's sense of self, and the robot is transformed into an agent having a mind in the eyes of the child[4].

Why are there differences in the reactions observed in children towards Infanoid and Keepon? In contrast to the strong tension and confusion displayed initially by the children upon meeting Infanoid, children seemed to be voluntarily motivated to interact with Keepon. Initially, children seemed to perceive the eyes, hands, and mouth of Infanoid separately. While the information received from each organ of the body is abundant, the signs of autonomy and life-likeness that should be evident as a whole (gestalt) — the recognition that the robot perceives and interacts with the world in the same manner as "I" do — is lost among the flood of information, and cannot be instantly perceived. Only through careful analysis and integration of the motions of each of the body parts do the children realize coherence of the whole as a "living organism". On the other hand, Keepon's form is entirely different from that of humans, and due to the simple expressions of attention toward a target and emotion and also due to its soft texture, the children are able to recognize intuitively autonomy and life-likeness in Keepon, as a gestalt perception[5] (Fig. 6).

Based on the above, we can summarize that an info-communication terminal that allows users to perceive its "mind" does not necessarily have to take a human-like form. Rather, the robot should have a form that will render the expression of attention and emotion most obvious to the user. Further, it should be possible to realize empathetic communications[10] on the assumption that minds are present at both the terminal and in the user, if the sharing of attention and emotion can in fact be established between them through eye contact and joint attention. In the future, we hope to conduct a more practical system implementation and validation experiment based on these findings.

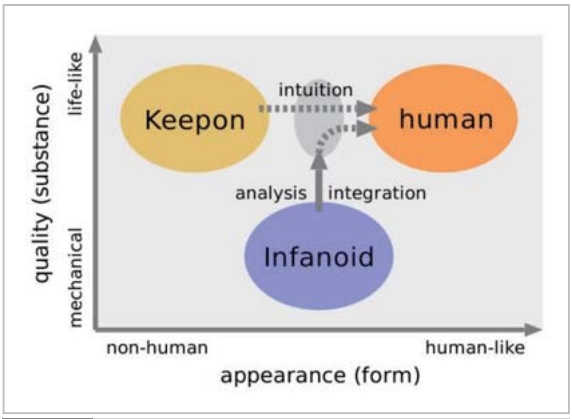

Fig.6 *Robot as seen by children — comparison of Infanoid, which requires conscious analysis and integration for recognition, and Keepon, which requires only intuitive interpretation*

References

1 Hideki Kozima and Akira Takada, "Developmental approach to social interaction", Journal of the Japanese Society of Artificial Intelligence, Vol.16, pp.812-818, 2001. (in Japanese)

2 Michael Tomasello, "The Cultural Origins of Human Cognition", Harvard University Press, Cambridge, MA, USA., 1999.

3 Colwyn Trevarthen, "Intrinsic motives for companionship in understanding: Their origin, development, and significance for infant mental health", Infant Mental Health Journal, Vol.22, pp.95-131, 2001.

4 Hideki Kozima, "Emergence of communication seen from an infant robot", Hattatsu (Human Development), Vol.24, pp.52-60, 2003. (in Japanese)

5 Hideki Kozima and Shoji Nagataki, "When we find "mind" in robots, when robots find mind in us", in Shoji Nagataki (ed.), Phenomenology and the Knowledge of 21st Century, Nakanishiya Publisher, pp.108-136 (and color leaf), 2004. (in Japanese)

6 Hideki Kozima, "What robots can do for special education", in Shin-ichi Watabe (ed.), Technology of 21st Century and Special Education, Gakuensha Publisher, pp.105-113, 2004. (in Japanese)

7 Hideki Kozima, Cocoro Nakagawa, and Yuriko Yasuda, "Children-robot interaction: a pilot study in autism therapy", in von Hofsten & Rosander (Eds.) Progress in Brain Research, Vol.164, pp.393-409, 2007.

8 Hideki Kozima, “Infanoid: A babybot that explores the social environment”, In: Dautenhahn K, et al. (eds), Socially Intelligent Agent, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp.157-164, 2002.

9 Hideki Kozima, Cocoro Nakagawa, and Hiroyuki Yano, “Can a robot empathize with people?”, International Journal of Artificial Life and Robotics, Vol.8, pp.83-88, 2004.

10 Dautenhahn, K., “I could be you: the phenomenological dimension of social understanding”, Cybernetic Systems Journal, Vol.28, pp.417-453 1997.

***KOZIMA Hideki**, Ph.D.*

Senior Researcher, Universal City Group, Knowledge Creating Communication Research Center (Former: Senior Researcher, Social Interaction Group, Keihanna Human Info-Communication Research Center, Information and Network Systems Department)

Cognitive Science, Developmental Psychology, Robotics, Human-Interface

***NAKAGAWA Cocoro**, Ph.D.*

Former: Researcher, Social Interaction Group, Keihanna Human Info-Communication Research Center, Information and Network Systems Department

Human Interface, Developmental Psychology, Therapy for Children with Developmental Disorders