International Journal of Humanoid Robotics

A HUMANOID IN COMPANY WITH CHILDREN

HIDEKI KOZIMA and COCORO NAKAGAWA

National Institute of Information and Communications Technology
Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0289, Japan
{*xkozima, cocoron*}*@nict.go.jp*

NOBUYUKI KAWAI

Graduate School of Information Science, Nagoya University
Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan
kawai@cog.human.nagoya-u.ac.jp

DAISUKE KOSUGI*

Graduate School of Letters, Kyoto University & JSPS Fellow
Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan
dkosugi@bun.kyoto-u.ac.jp

YOSHIO YANO

Faculty of Education, Kyoto University of Education
Fukakusa-fujimori-cho, Fushimi-ku, Kyoto 612-8522, Japan
yano@kyokyo-u.ac.jp

We describe how normal and autistic children interacted with a humanoid, *Infanoid*, focusing especially on the effect of the robot's attentiveness and emotiveness on the children's understanding of what the robot can do and how to interact with it. *Infanoid* is an upper-torso humanoid robot, which is capable of attentional and emotional interaction with humans through gaze, voice, facial expressions, and bodily/manual gestures. We observed 14 normal children (ranging from 6 months old to 9 years old) and one autistic child (6 years old) interact with *Infanoid*. Each of the children sat in front of *Infanoid* with his/her caregiver and gradually got into the loop of interaction, where he/she spontaneously played with the robot, gradually deepening the level of interaction, in a playful mood. We assume the relatively high predictability of *Infanoid*'s behavior enabled the autistic child, as well as the normal children, to get into the relaxed interaction with the robot, from which the children explore higher communicability of the robot. This study suggests potential applicability of humanoids to remedial services as well as psychological investigation of social interaction.

Keywords: Social interaction; developmental psychology; attribution of mind; autism and communication disorders.

*Current affiliation: Shizuoka Institute of Science and Technology (Shizuoka, Japan)

1. Introduction

We are building and evaluating a humanoid robot that is intended to help normal and handicapped children learn to socially communicate with other people. Embodied interaction is the key activity in the social communication, where one sees invisible mental states (e.g. intention and belief) in the visible posture and movement of others. This paper describes what we observed 14 normal children and one autistic child interact with an upper-torso humanoid robot, *Infanoid*. The children's rather relaxed manner of interaction suggests the advantage of using humanoids (neither complex humans nor simple toys) in pedagogical and therapeutic applications as well as cognitive studies on social interaction.

Although there has been some pioneering studies on child-robot interaction, not much has been done on that for the remedy for autism. Most of the early studies used small mobile/immobile robots,[1,2] giving minimum mental and physical impact to children. Probably for this reason, humanoids have barely been used in the research on child-robot interaction, especially that for therapeutic purposes, though a few successful works[3,4] used a simple humanoid doll (with 5 DOFs).

This paper describes our preliminary study on child-robot interaction using relatively sophisticated humanoid, *Infanoid* (with 29 DOFs). First, we introduce the structure and function of our robotic research platform *Infanoid*, in Section 2 and 3, respectively. Then, Section 4 describes what we observed in the interaction of the normal children and the autistic child with *Infanoid*. Finally Section 5 gives discussion and conclusion of this explorative study.

2. Infanoid, the Attentive and Expressive Humanoid

Infanoid, shown in Fig. 1, is an upper-torso humanoid robot, which is as big as a 3- to 4-year-old human child.[5] *Infanoid* is not only a research platform on which we model and implement human communicative development, but also a tool for psychological experiments to investigate how humans, especially children, interact with it. We are observing how children respond to the robot's social actions such as gazing and pointing, and also how they perform spontaneous actions such as showing and giving an object to the robot. It is worth noting that we can control the complexity of *Infanoid*'s behavior in order to meet our research objectives and the children's developmental stages.

Infanoid has 29 actuators (mostly DC motors with encoders and torque/current sensing devices) and a number of sensors arranged in this relatively small body . It has two hands, each of which has four fingers and a thumb that are capable of pointing, grasping, and a variety of hand gestures.

The head of *Infanoid* has two eyes, each of which contains two different color CCD cameras for peripheral and foveal view; the eyes can perform saccadic eye movements and smooth pursuit of a visual target. The video images taken by the cameras are fed into a cluster of PCs for real-time detection of human faces (by a skin-color filter and template matching) and of physical objects such as toys

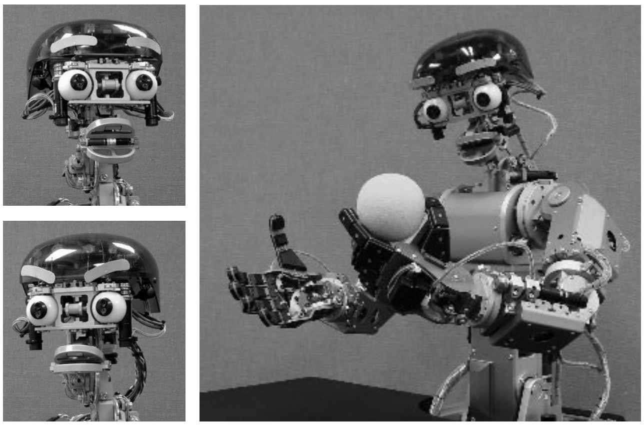

Fig. 1. *Infanoid*, the attentive/expressive humanoid.

(by color/motion segmentation). The distance to the faces and objects can also be computed from the disparity between the left and right images.

Infanoid has lips and eyebrows; by changing the shape of the lips and eyebrows, it expresses a variety of emotional expressions, like surprise and anger, as shown in Fig. 1 (left). The lips also move in accordance with the sound produced by a speech synthesizer, giving ventriloquism effects to the interactants.

From the microphones at the positions of the ears, *Infanoid* hears human voices and analyzes the sound into a sequence of phonemes — the robot does not have any prior knowledge about language (like lexicon or grammar). It also recognizes any change in the fundamental frequency to extract an emotional contour from human speech. By feeding the output of the speech analyzer into a speech synthesizer, *Infanoid* performs vocal imitations while sharing attention with the interactant, which we consider to be a precursor to the primordial phase of language acquisition.

3. Eye-Contact and Joint Attention

Psychological studies of communicative development[6,7,8] as well as some robotics studies on social interaction[9,10] suggest the importance of eye-contact and joint attention, by which a child's attention and action are coupled with those of the caregiver, forming a dynamic system of interaction:

- Eye-contact gives temporal synchronization to the interaction, making the child and the caregiver aware of each other's attentiveness. The interactants often exchange their facial expression and vocalization, through which they can monitor and share each other's emotional states. It also gives mutual acknowledgment that both are aware of having eye-contact.
- Joint attention gives spatial orientation (i.e. focusing) to the interaction, making the child and the caregiver share the perception of the target (some-

Fig. 2. Eye-contact and joint attention with *Infanoid.*

times from different view points). It also gives mutual acknowledgment that both are attending to the same target and sharing the almost same perception.

One's attentiveness suggests the existence of subjectivity, which selectively picks up a certain portion of the environment and makes the one ready to interact with that portion; synchronization and co-orientation in the interaction suggest the existence of a relationship or companionship between two subjective beings.

The discussion above implies that a robot capable of eye-contact and joint attention deserves to be a socially interactive partner, which leads us to believe the existence of a mind. In other words, an agent we relate to in terms of attention and emotion can be considered as a social interactant to which we attribute a mind. This is our basic motive for building a series of interactive robots, *Infanoid*s, capable of eye-contact and joint attention.

3.1. *Eye-contact in human and robot*

In normal infants, the onset of eye-contact appears right after birth[11], often with so-called neonatal imitation of facial gestures[12]. These presumably innate competences are the driving force to respond to the environment, through which the infant explore the physical and interpersonal world[8].

In our robot, *Infanoid*, the eye-contact capability is implemented as follows. First, from the real-time image streams (30 frames/sec) taken by the cameras, the robot searches for a human face looking straight using a skin-color filter and average-face templates. If a face is detected, the robot drives the motors to direct the gaze and face (sometimes together with the arms and upper-body) toward the detected face; also the eyes give an appropriate amount of convergence according to the distance to the face computed from the image disparity. Then the gaze of the robot and that of the human interactant are facing straight at each other, establishing eye-contact between the human and the robot. (See Fig. 2, left.)

3.2. *Joint attention in human and robot*

The joint attention capability in normal infants starts working before 6 months old and gradually becomes more sophisticated up to 18 months old[13]. At the first stage, infants can identify the attentional target in the rough direction (e.g. right or left side) of the interactant's head only when the infants could see both the agent and the target within their visual field; at the later stage, infants become able to utilize the head direction of the interactant to identify the right target of the attention; finally, they become able to identify targets behind them. Joint attention is also observed in some species of non-human primate,[14] although most of the evidences were obtained from captive subjects.

In our robot, *Infanoid*, the joint attention capability is implemented as follows. The robot first generates several hypotheses about the direction of the face being tracked. From the images taken by the cameras, the robot computes the likelihood for each of the hypotheses and selects the most likely direction of the face. Then, the robot starts searching in that direction and identifies the target of the human attention. Currently, the target object is segmented out from the background by a predetermined color (e.g. saturated colors of baby's toys) and motion. Finally, the robot drives the motors to direct the gaze and face (sometimes together with arms and upper-body) toward the target object, thus establishing joint attention with the human interactant. (See Fig. 2, right.)

4. Child-Robot Interactions

We first carried out a series of observations, where normal children interacted with *Infanoid* without any prior knowledge about the robot. From this observation of the children, we learned the developmental trend in the manner of their spontaneous interaction, and also we were convinced of the *Infanoid*'s safety in terms of mental and physical impact to the children.

4.1. *Infanoid and normal children*

We have observed 14 normal children (from 6 months to 9 years of age) interact with *Infanoid*. In these observations, the robot ran in Automatic Mode, in which it alternates between eye-contact and joint attention with pointing. If necessary, a remote operator made adjustments to the robot's attention (e.g. direction of the gaze/face/arms/body). First, each child was seated alone in front of the robot. About 3 to 4 minutes later, the child's mother came in and sat next to the child. Interaction continued until the child get tired or bored; on average, each child had an interaction of about 30 minutes.

From this observation, most of the children (especially those that were 3 to 6 years old) showed the following changes in the interaction. (See Fig. 3.)

- **Neophobia phase:** When the child interacted with the robot alone, he or she looked seriously into the robot's eyes. Even though the robot produced

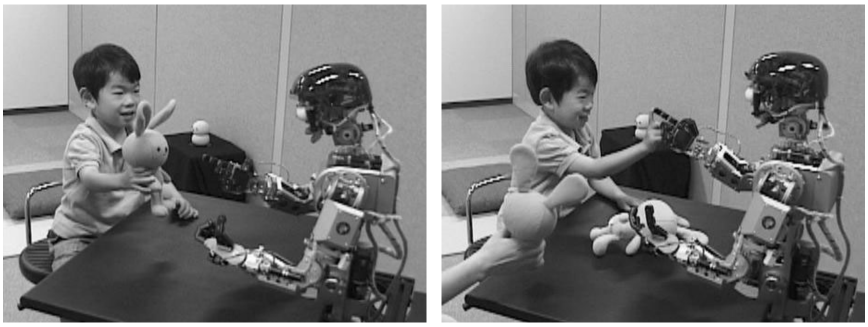

Fig. 3. Interaction with a normal child.

a mutual gaze or an aversive gaze, the child's eyes were locked onto the robot's eyes. The child then showed embarrassment, not knowing how to deal with this weird, moving thing.

- **Exploration phase:** Next, using his or her mother as a secure base, the child started exploring how the robot changed its attention and posture in response to various interference, such as showing toys and poking the robot. When the child elicited an interesting response from the robot, he or she often made referential-looking and comments to their mother. Through this exploration, the child would find that the robot was an autonomous agent that has attention and emotion.
- **Interaction phase:** The child then gradually got into a social interaction, where he or she pointed at the toys to share attention and gave the toys to the robot by putting them in the robot's hands. Verbal interaction also started by asking questions (e.g. *Which one do you want?*, showing two toys) or asking the robot to do something (e.g. *Grasp it like this!*, showing how to handle a toy). The child seemed to attribute mental states, such as desire and satisfaction, to the robot.

The children changed their recognition or expectation of *Infanoid* dynamically: first as an unknown, ambiguous "moving thing", then as an "autonomous agent" that has attentiveness and responsiveness, and finally as an "social being" that deserves to be involved in a social interaction including verbal one.

4.2. *Autism as communication disorder*

Autism (or autistic spectrum disorders), which used to be considered as an enigmatic mental disease of infants and children, is actually a neurophysiological disorder caused by specific and mainly hereditary brain dysfunction.[15] High co-occurrence is found in twins, which implies its genetic background. People with autism often have abnormal brain waves (sometimes with epilepsy), which implies its neurophysiological background. Most autistic people have mental retardation in various degrees. A recent statistical study reported that about 0.1 to 0.2% of the population are said to be autistic in any country.[16] The ratio of males to females is about 4 : 1.

People with autism have difficulties in social interaction, verbal communication, and maintaining diversity of interest and behavior,[17] which make autistic people difficult to establish and maintain social relationship with others. Major symptoms of autism can be described as follows:

- **Social interaction:** Difficulty in the use of gaze, pointing, and facial expressions in socially meaningful ways; inability to share interests and activities with others
- **Communication:** Delay or lack of language development; use of stereotyped and repetitive speech; impairment in pragmatic and conversational use of language and gesture.
- **Imagination:** Stereotyped and restricted pattern of interest and behavior; adherence to specific things and aimless routines; difficulty in coping with novel situations.

In spite of these symptoms, people with autism often retain other cognitive skills like spatial recognition and rote memory.

Autistic infants and children, in general, are less likely to engage in eye-contact or joint attention with others (even with their caregivers),[15,18] which is actually one of the significant clues for the diagnosis of autism. However, being instructed by an experimenter, they look into the eyes and often identify the attentional target;[19] this implies that their perception is intact, but they seem rather lacking in motivation to read a certain information from other people's gaze and face.

4.3. *Infanoid and an autistic child*

We carried out two observations of a 6-year-old high-functioning autistic boy (hereafter *S*) interacting with *Infanoid*. *S*'s verbal and non-verbal intelligence was in the normal range; however, he had difficulty in interpersonal communication and in adaptation to unfamiliar situations. In order to minimize his possible robotophobia (analogous to homophobia to strangers), we had sent his family beforehand a photo of *Infanoid*, asking his parents to tell him that he was going to play with the robot. Other conditions and experimental procedure were equivalent to those with the normal children described above. The first session with *S* was taken in March 2003, and second in December 2003 at his own request.

In the first session, *S* had an interaction with *Infanoid* for about 45 minutes, which was far beyond our expectation, so we had to stop the observation before *S* got bored or tired. As with the normal children, *S* first showed a neophobia phase, where he seemed to feel embarrassed at the first encounter with the robot, not understanding how to do with it (Fig. 4, left). Several minutes later, however, *S* started exploring gradually how *Infanoid* responded to various stimuli, such as

Fig. 4. An autistic child S interacting with *Infanoid*.

Fig. 5. S drawing the gaze-line of *Infanoid*.

showing stuffed animals, touching the robot's hands, and looking at various parts of the body (Fig. 4, middle). Sometimes S took a very close look at *Infanoid*'s skull, which is made of semi-transparent acrylic, so that he could see the mechatronics stuff inside.

After the 10- to 15-minute exploration phrase, S started interacting with *Infanoid* in a social way, where S attributed desire and likes/dislikes to the robot. For example, S verbally asked *Infanoid* "*Which one do you want?*", showing a number of stuffed animals to the robot (Fig. 4, right). When he tried to find the answer, he often referred to *Infanoid*'s reaching and pointing direction, not the face/gaze direction. Another interesting anecdote is his drawing a gaze-line of *Infanoid* (Fig. 5). When his father came into his right side (not shown in Fig. 5), his mother asked S, "*Where is he looking at?*". S did not reply to the question quickly, but he drew the gaze-line of *Infanoid* and finally he moved to his father's position, confirmed the straight gaze of *Infanoid*, and finally said "*He is looking at my dad!*".

In the second session (about 10 months after the first one), S started socially interacting with *Infanoid* from the very beginning. This session lasted for about 40 minutes until S showed a slight sign of boredom. Throughout the session, S initiated a pretense game with *Infanoid*, telling the robot a rule something like "*You must put it in this box, otherwise it would explode!*", putting a stuffed animal in *Infanoid*'s hand and then showing a plastic box to the robot. S also played hide-and-seek with *Infanoid*, ducking his head under the table, so as not to be looked at by the robot. His back was, however, partially seen by *Infanoid*; he talked with (not whispered to) his mother while he was under the table, although he knew that *Infanoid* could hear the voice.

5. Discussions and Conclusion

From the interactions with the normal children, we found the following stages unfolding toward the emergence of social interaction. First, children recognize the robot as a moving thing; then, observing the robot's motion in response to various environmental disturbances, they recognize that the robot is an autonomous, subjective system that possesses attention and emotion as an initiator of the motion. Next, they find that the robot's response (in terms of attention and emotion) has a spatio-temporal relation with what they have done to the robot; then, they recognize the robots as a companion with which they can exchange or coordinate their attention, emotion, and action.

From the interaction with the autistic child, we found that the child naturally got into a social interaction with the robot keeping himself relaxed and comfortable. This would be mainly because of the relatively high predictability of *Infanoid*'s behavior — e.g. it never stands up or walks over to the child. From the child's point of view, the gamut of the possible actions of *Infanoid* was much smaller and easier to estimate than that of humans, probably because of the robot's anthropomorphic structure (e.g. maximum distal reach and range of rotational motion) and semi-repetitive motion (e.g. alternating between eye-contact and joint attention). This enabled the child to get into the relaxed interaction, ranging from social actions (such as giving, showing, and asking) to an *ad lib* game and hide-and-seek. Although his way of playing with *Infanoid* is still in his own style, he attributed a mind (e.g. desire and likes/dislikes) to the robot and also shared pleasure and excitement with the robot.

To summarize, attentive and expressive humanoids have high potential for helping normal and handicapped children learn (or get used to) social communication. Especially when interacting with an autistic child, the medium level of complexity or predictability of the robot's behavior enables the child to be involved in relaxed and playful interaction with the robot. In addition, we can control the complexity so as to meet the child's developmental profile, giving the child an appropriate learning environment (zone of proximal development).

References

1. Michaud, F. & Théberge-Turmel, C. (2002) Mobile robotic toys and autism, in Dautenhahn, K. *et al.* (eds), *Socially Intelligent Agent*, Kluwer Academic Publishers, pp. 125–132.
2. Dautenhahn, K. (1999) Robots as social actors: Aurora and the case of autism, in *Proceedings of the Internatonal Cognitive Technology Conference*, pp. 359–374.
3. Billard, A. (2003) Robota: Clever toy and educational tool, *Robotics and Autonomous Systems*, Vol. 42, pp. 259–269.

4. B. Robins, K. Dautenhahn, R. te Boekhorst, & A. Billard (2004) Effects of repeated exposure to a humanoid robot on children with autism, in Keates, S. *et al.* (eds), *Designing a More Inclusive World*, Springer Verlag, pp. 225–236.
5. Kozima, H. (2002) Infanoid: A babybot that explores the social environment, in Dautenhahn, K. *et al.* (eds), *Socially Intelligent Agent*, Kluwer Academic Publishers, pp. 157–164.
6. Kaye, K. (1982) *The Mental and Social Life of Babies*, University of Chicago Press.
7. Tomasello, M. (1999) *The Cultural Origins of Human Cognition*, Harvard University Press.
8. Trevarthen, C. (2001) Intrinsic motives for companionship in understanding: their origin, development, and significance for infant mental health, *Infant Mental Health Journal*, Vol. 22, pp. 95–131.
9. Breazeal, C. & Scassellati, B. (2000) Infant-like social interactions between a robot and a human caretaker, *Adaptive Behavior*, Vol. 8, pp. 49–74.
10. Kozima, H., Nakagawa, C., and Yano, H. (2003) Attentional interaction as a prerequisite for social interaction, in *Proceedings of the IEEE International Workshop on Robot and Human Interactive Communication*.
11. Fantz, R. L. (1961) The origin of form perception, *Scientific American*, Vol. 204, pp. 66–72.
12. Meltzoff, A. & Moore, M. K. (1999) Persons and representation: why infant imitation is important for theories of human development, In J. Nadel and G. Butterworth (eds), *Imitation in Infancy*, pp. 9–35, Cambridge University Press.
13. Butterworth, G. & Jarrett, N. (1991) What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy, *British Journal of Developmental Psychology*, Vol. 9, pp.55–72.
14. Itakura, S. (1996) An exploratory study of gaze-monitoring in nonhuman primates, *Japanese Psychological Research*, Vol. 38, pp. 174–180.
15. Frith, U. (1989) *Autism: Explaining the Enigma*, Blackwell.
16. Rapin, I. & Katzman, R. (1998) Neurobiology of autism, *Annals of neurology*, Vol. 43, pp. 7–14.
17. *The ICD-10 Classification of Mental and Behavioral Disorders* (1993), World Health Organization.
18. Hobson, R. P. (1993) *Autism and the Development Mind*, Psychology Press.
19. Baron-Cohen, S. (1995) *Mindblindness: An Essay on Autism and Theory of Mind*, MIT Press.