Does Disturbance Discourage People from Communicating with a Robot?

Claudia Muhl and Yukie Nagai

Applied Computer Science, Faculty of Technology, Bielefeld University, 33594 Bielefeld, Germany
e-mail: {cmuhl, yukie}@techfak.uni-bielefeld.de

Abstract— **We suggest that people's responses to a robot of which attention starts to be distracted show whether they accept the robot as an intentional communication partner or not. Human-robot interaction (HRI) as well as human-human interaction (HHI) is sometimes interrupted by disturbing factors. However, in HHI people continue to communicate with a partner because they presuppose that the partner may shift his/her interactive orientation based on his/her internal state. We designed a communication robot equipped with a mechanism of saliency-based visual attention and evaluated it in an observational experiment of HRI. Our sociological analysis of people's responses to our robot showed that it was accepted as a proactive communication agent. When the robot shifted its attention to an irrelative target, the human partners, for example, followed the line of the robot's gaze and tried to regain its attention by exaggerating their actions and increasing their communication channels as they would do toward a human partner. Based on these results, we conclude that disturbance can be an encouraging factor for human activity in HRI. The results are discussed from both a sociological and an engineering point of view.**

I. Introduction

What makes human-robot interaction (HRI) more social? Although measuring the sociality of interaction seems to be difficult, many researchers have been attempting to find the factors influencing the human impression of robots [1]. Goetz et al. [2], for example, suggested that matching a task and the robot's behavior to it improves human-robot cooperation. They designed two types of robots, a playful and a serious one, and compared the people's acceptance of the robots working either on an entertaining or a serious task. Their results showed that people interacted longer with and accepted more the robot which acted in an appropriate manner to the task. Minato et al. [3] defined the familiarity of a robot with respect to its appearance and behavior. They extended the uncanny valley proposed by Mori [4] and described how the above two factors synergistically affect the people's impression of a robot. Their experiments using their android, in which the human response of breaking eye contact was measured, showed that people dealt with the android as a human-like agent. These studies had a major impact on the design of communication robots, however, they focused only on the factors directly relevant to HRI.

Our key interest is to find out additional factors, which are not the primal elements of HRI but could evaluate and encourage it. We propose that disturbance in interaction can be such a factor. Imagine human-human interaction (HHI): Our interaction is sometimes interrupted by an irrelative element. When two people are talking on the telephone and one of them starts to watch television, for example, the conversation could be interrupted because the attention of whom watching television might be engaged by it. Although HHI can be disturbed like this, people seldom give up communicating but try to repair it. One reason is that we assume that the partner's attention is driven by his/her internal states, e.g., belief, desire, or intention, and does not always respond to us. On the other hand, if we consider that a partner does not have its internal state like a machine and a computer, we would stop to communicate and not try to regain the partner's attention. We therefore suggest that investigating people's reactions to a robot of which attention starts to be distracted allows us to evaluate whether the robot is accepted as an intentional communication partner or not.

We furthermore point out that changes in people's responses caused by disturbances can aid a robot communicating with and learning from them. In the example mentioned above, if the partner's attention is engaged by the television, the other would probably start to talk louder and highlight the tone of his/her voice to regain the partner's attention. He/She may even move his/her hands and body although he/she knows this motion cannot be recognized by the partner. We consider these phenomena as positive effects caused by disturbance because the reinforced and additional actions may help the partner to presume the meaning and the intention of the actions.

We investigate from a sociological perspective how disturbance in HRI affects the behavior of human partners. As nowadays an interdisciplinary discourse is continuously emerging in the field of robotics (e.g., [5], [6]), we wish to add an analytical perspective from the field of sociology. We believe that a qualitative approach helps to reveal effects induced by disturbance, to discover how people act and react toward a robot, and thereby to contribute to the development of social robots.

In Section II, we describe sociological phenomena of HHI and disturbance in communication. Our robot simulation as a communication partner is explained in Section III. In Section IV, we show an experiment of HRI using the robot, and discuss the results from both a sociological and an engineering point of view in Section V. Finally, the conclusion is given in Section VI.

II. Sociological Discussion on HHI

A. Participation in Social Interaction

Every contact which takes place between humans, who are acting toward others, is social. Interaction means communicating with partners in attendance. One criteria for communication is the reciprocal percipience, whereas the presence or media mediated utterances of both partners is a prerequisite to be acquired [7]. The reciprocity of awareness of the coparticipants means both are sharing contextual perceptions which enable them to construct a common sense and to build a situated common ground. These operations open an intersubjective space of social actions and expressions. Participating in an interaction means to be engaged in the emerging course of action with interactive practices and to shape it [8]. Therefore an interaction system can be understood as a set of reciprocal addressed behavior [9].

Sequences of contributions of speech and actions like mimic, gesture, and body movements are aligned with the partner in each interaction [10], [11]. All elements of the dialog are organized as reciprocal turns which are successively arranged in a turn-taking set. Each participant of an interaction is oriented toward the partner by considering his/her individual situated involvement. This phenomenon includes a sensitivity to the coparticipant and his/her situatedness. Sacks et al. [12] named the context-aware possibility of referencing to partners' actions and utterances in HHI *recipient design*. Sociological reflections since Garfinkel [13] take into account that interactions are situated in a specific context constructed by each interaction partner employing his/her own category systems, commonsense knowledge, and practical reasoning to the actual experience. Though the interaction partners achieve mutual understanding. As a consequence of this individual construction of the specific social situation, actors are able to act within their circumstances and interprete others.

B. Dealing with Disturbances

Communications are fragile and their alternating follow-up is often disrupted by surrounding factors. What happens in case of addressing someone who has lost concentration and is occupied with processing information derived from a third person's perspective? This shift of the attention will be recognized by the partner and cause some reaction. Spontaneously appearing reasons might effectuate severe irritations that can lead to discontinuity in the dialog processing and the interaction might be terminated. By lifelong practice humans learn to deal with such disturbances. They can defy the problem, and thus an originally negative cause leads to positive effects.

In HHI the problem of focusing the attention in a communication has been investigated intensively. Social interactions can be studied in everyday life, and such analysis revealed that humans moreover often implement disturbances in communications themselves. Goodwin and colleagues [7], [10] analyzed multiple face-to-face interactions of humans in different contexts. They discovered that the ideal turn-taking in *talk-in-interaction*s is often disturbed by the coparticipants themselves [7]. Such strategic elements are used in order to evaluate whether the partner's co-orientation is still focused on the ongoing interaction [10].

Those techniques are applied to organize the exchange of speech and gaze, and, if inconvenience is discovered, *repairing mechanisms* are initiated [14]. For example, the speaker often stops him-/herself and restarts the sentence with identical words. Such explicitly evoked breaks in the verbal flow ensure about the interaction partner's concentration on the mutual topic. As a positive effect, those disturbances affirm the interaction and the dialog can be continued.

III. Design of a Communication Robot

To investigate how disturbance affects HRI and whether it causes positive effects as in HHI or not, we developed a communication robot of which attention can be naturally distracted by a visual disturbance.

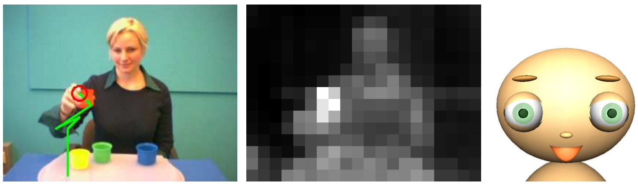
A. Robot's Face Simulation

Fig. 1 (c) shows a simulation of a robot's face used in our experiment, which was originally developed by Ogino et al. [15] for studying the emotional state of a robot. The robot can be controlled by using four parameters of the eyes, four of the eyebrows, eight of the upper and lower eyelids, and two of the mouth. Thus, it can move the eyes so that a human partner recognizes that it is gazing at a certain location in the environment. The eyelids, the eyebrows, and the mouth are used for blinking and showing emotional expressions of the robot, which makes the communication more natural. In the experiment described in Section IV, the robot is displayed on a computer screen with a camera for the robot's vision. No microphone or speaker is used, meaning the robot can only respond visually but not acoustically.

B. Mechanism of Visual Attention

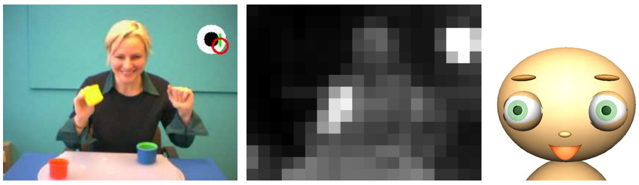
We adopt a model of saliency-based visual attention [16], [17] as the mechanism for the robot's vision. The model, inspired by the behavior and the neuronal mechanism of primates, imitates the primary attention of humans, who can rapidly detect and gaze at salient locations in their views. A salient location is here defined as a spot which locally stands out from the surroundings because of its color, intensity, orientation, flicker, and motion [17]. The model therefore works in a bottom-up manner without any knowledge about the environment or a task but allows the robot to detect likely important information in communication. The effectiveness of the model has been demonstrated in the studies of social robot learning and social robot interaction (e.g., [18], [19]).

Fig. 1 (a) shows a sample scene from the experiment, in which a human partner is picking up and showing a red cup to the robot, and (b) shows the corresponding saliency map, in which the degree of saliency is represented by the



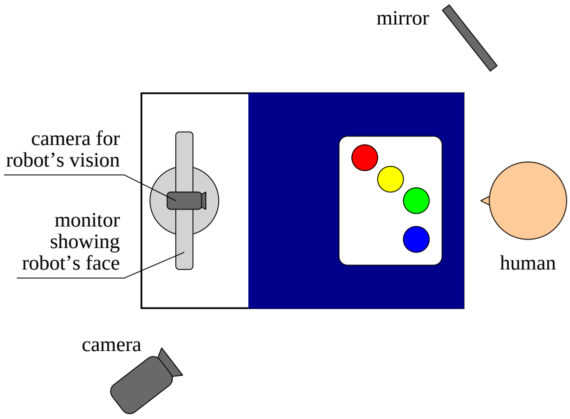
(a) input image (b) saliency map (c) robot's face

Fig. 1. A scene without disturbance, in which the robot is looking at the red cup held by the human partner

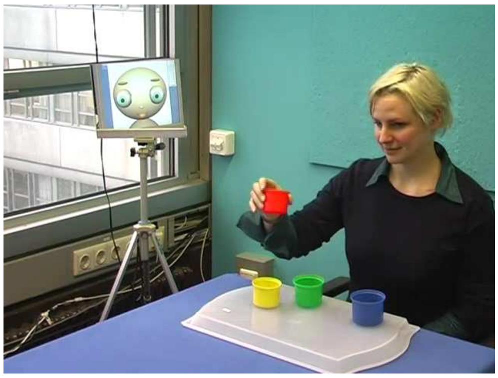


(a) input image (b) saliency map (c) robot's face

Fig. 2. A scene with disturbance (a black and white circle) superimposed in the image, which attracts the robot's attention more than the yellow cup because of the highly contrasted intensity and the motion



(a) top-view of experimental setup



(b) videotaped HRI, in which the robot is reflected on the left mirror

Fig. 3. Experimental setup

brightness of the pixels. Using the map, the robot selects the most salient location to attend to frame by frame. The attended location is denoted by a red circle in (a), and its trajectory over the last five image frames is denoted by green lines. In this scene, the red cup held in the right hand of the human partner is gazed at and has been tracked for a while because of its outstanding color and motion. Fig. 1 (c) shows the robot's face captured when it looks at the red cup. The robot's eyes are controlled so that human partners can recognize it is responding to their actions and is looking at an interesting location. Note that human partners can only see the simulation of the robot's face, but not the camera image or the saliency map.

C. Disturbance in Robot's Vision

To create a visual disturbance for the robot, we superimpose a salient feature in the robot's camera image. Fig. 2 shows a scene captured while a salient feature is put at the upper-right corner of the image. The disturbance is a white circle with a smaller black circle, which vibrates randomly. Because of its highly contrasted intensity and the motion, the disturbance is detected mostly, but not certainly, as the most salient location in a scene. In Fig. 2, the robot gazes at the disturbance although the human partner tries to attract its attention by showing the yellow cup and shaking her left hand. Note that human partners cannot see the disturbing feature in the real environment. The initial position of the disturbance is fixed at the upper-left or upper-right corner of the image.

IV. Experiment of HRI

A. Setting and Procedure

We conducted an observational experiment of HRI using the robot simulation. Fig. 3 (a) illustrates the experimental setup, and (b) shows a scene of videotaped interactions. A human partner was sitting in front of a monitor displaying the robot, and a camera for the robot's vision was placed on the monitor. Another camera beside the monitor recorded the interaction between a human partner and the robot reflected on a mirror.

Subjects were 22 university students, 16 of them study computer science and the others sociology or linguistics. They were instructed to teach some tasks, e.g., stacking cups and sweeping on the table, to the robot. Nothing about the usage of gesture or speech was mentioned although the robot's capability was explained, i.e., it could not respond acoustically but only visually. Note that the detailed mechanism of the robot's vision was not shown to the subjects. The interaction took five to thirty minutes depending on the subject. An experimenter controlled the timing to insert and to remove the disturbance responding to the subjects' actions toward the robot, which means that the subjects' efforts do not directly effect to regain the robot's attention. Another experimenter asked some questions about the impressions of the robot afterwards.

B. Methodology of Sociological Analysis

Qualitative sociological methodology helps to identify concrete human behavior and social interaction in a contextual setting. It seeks to describe the underlying social patterns which occur as concrete phenomena in the real world. In this experiment we make use of ethnomethodological conversation analysis for investigating the video data of the HRI.

Conversation analysis is a qualitative method to evaluate the speech and action processes of individuals in a continuous interaction situation [7]. This close grained analytical technique starts with describing prominent elements from the empirical data. With the categorization of action patterns, the interaction structure can be revealed.

The goal of the sociological reasoning in our HRI experiment is to evaluate the interactive potential of irritation. The disturbance of the robot becomes part of the interaction system, meaning it causes irritation in the human partners that leads to a change in their behavior. Therefore their most common reactions, while a disturbance occurred to the robot, have been collected and classified.

C. Categorization of People's Responses

In the beginning of the interaction, most of the subjects concentrated on their own actions. They were showing different strategies concerning eye-contact in order to check the robot's attention: some inspected the robot's gaze just after a fulfilled task, and others checked more often. When they recognized extraordinary changes in the robot's gaze behavior, they got irritated. Ascertaining a much more differentiated set of actions in case of disturbance than expected primarily, we searched for specific features in the human behavior. Here we focused on aspects that occurred during the interactions affected by the interruption.

Analyzing the individual performances, we found a set of main strategies. They were directed toward the robot and attended to evaluate the cause of its behavior. By bringing these observations to a more abstract level, we propose a two-dimensional map shown in Fig. 4, which illustrates the categories of the subjects' responses caused by the robot's disturbance.

We have five categories on two dimensions: the physical and psychological distance to the robot and the implied change in the subjects' activity, strong enough to recover the relationship in the ongoing HRI.

1) Building triadic interaction: While interacting with the robot, some subjects followed the line of the robot's gaze when it was disturbed and tried to achieve joint attention (see Fig. 5 (a)). At the same time, they often commented verbally on the expected direction of the robot's gaze, although there would not be anything to discover. This reaction shows situated involvement. A human is following the robot's action and attributing to it a participant's role in the interaction. This phenomenon marks the evolvement of a triadic interaction, which includes the surrounding context.

2) Attracting the robot's attention to oneself: The next category represents a huge variety in the reactive intensity. The subjects began exaggerating their already performed actions. Their gesture and movement became larger (see Fig. 5 (b)). Others called the robot, just started to talk to the

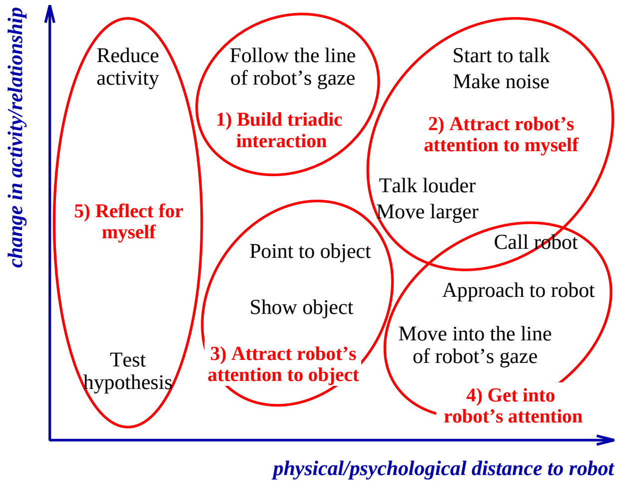


Fig. 4. Five categories of people's responses caused by disturbance in two-dimensional space: physical/psychological distance to robot and change in activity/relationship

robot or made noise, even though they already had tested the robot would not react to acoustic signals (see Fig. 5 (c)). They seemed to try to attract the robot's attention to themselves.

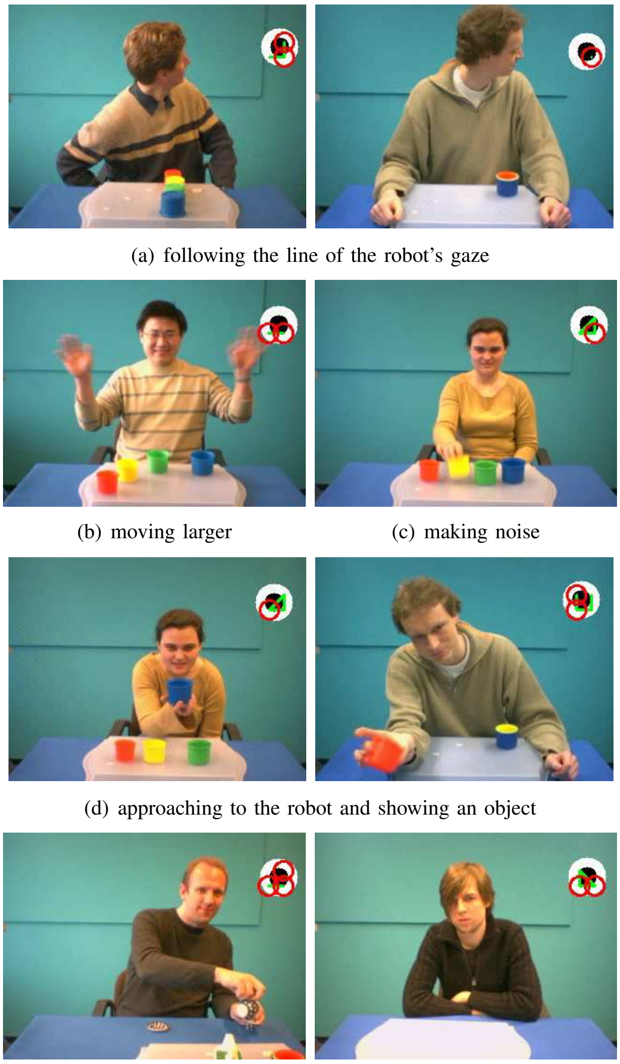
3) Attracting the robot's attention to an object: The third category assembles strategies that could possibly attire the robot's attention back to the object, i.e., getting closer to the robot while demonstrating the object (see Fig. 5 (d)). The object has been shaken or closely presented to the robot. Some subjects also pointed to the object to re-attract the robot's gaze.

4) Getting into the robot's attention: Reaching closer to the robot builds the fourth category of action. Here we sum movements like a physical approach of the subjects to the robot. Some of them even spatially moved into the line of the robot's gaze (see Fig. 5 (e)). As a consequence, they became more present to the robot and decreased the psychological distance to it.

5) Reflecting to oneself: The fifth category assembles the biggest and smallest change in the human activity compared to their former way of action toward the robot. Some of them tested their hypothesis on the robot's functionality by increasing and others by reducing the intensity of their activity (see Fig. 5 (f)). This included the sequential variation of the subject's former applied action patterns toward the robot.

In this experiment, humans performed social actions toward the robot. Ordering the categories of observed behavior depending on the complexity of the interaction and its direction, we consider three groups: a triadic relation (category 1), a dyadic one (categories 2 to 4), and other (category 5), in which people do not direct interactive utterances but rather go along with inner reflections.

As expected, people interacting with the robot, which was affected by an emerging disturbance, showed an immediate change in their behavior when they realized the interaction



(a) following the line of the robot's gaze

(b) moving larger

(c) making noise

(d) approaching to the robot and showing an object

(e) moving into the line of the robot's gaze

(f) reducing his activity and testing his hypothesis

Fig. 5. Samples of people's responses caused by disturbance

has been concerned. They intended to repair the situated disorder. Our findings prove that human action is likely to be varied in case they do not relieve the expected results. The reactions demonstrate a renewed conceptualization of the situation and the modification of the human hypothesis on the robot's functions, which reminds of recipient design in HHI. All of these strategies tend to refresh and repair the irritated flow of communication. After disturbance, the completion of the primordial task was often abandoned, but the communicative process has been reestablished and the interaction mostly even intensified.

V. Discussions

A. Sociological Perspectives on People's Interaction Strategies

Our qualitative analysis revealed a set of multiple human reactions in the HRI setting. The differences occurring among individuals might be taken into account carefully as they show the social process and its situatedness. We need to reflect diverse terms belonging the interaction. What do the participants take into account, and therefore what do their expectations toward the robot's actions consist of? Which knowledge do they bring in? The strategies varied for example based on the subject's familiarity with robot systems. Experts rather than laymen intensified an informed testing about what could have caused the fault. The practical strategy is therefore based on individual reasons as background knowledge and projections which derive from expectations by humans who apply recipient design. Our evaluation demands further investigations concerning the variety of the interactive behavior influenced by the background knowledge of humans.

In the focus of these interactions between a human and a robot, the problem in acting, speaking, hearing and interpreting became evident. The human partners used social repairing mechanisms as known in sociological conversation analysis. One effect of the disturbance is the encouragement of gestures and utterances. The irritation of the subject caused by the robot's miss-linked focus can motivate to reveal strategies to regain the robot's attention and to recover a turn-taking process. This phenomenon can also be detected in participants in HHI. They often establish turn-taking systems, which organize the exchange of information efficiently, and also apply repair mechanisms to vanquish distractions. Further sociological work should investigate the interactive relations in HRI and concentrate on the directed attention which can effect mutual talk-in-interaction between a human and a robot.

B. Engineering Findings on Robot's Attention Mechanism

Our analysis on HRI offers perspectives on designing social learning robots. Our robot equipped with the ability of saliency-based visual attention succeeded in motivating people to continue to communicate longer. In the experiment, most subjects held on to the communication for more than ten minutes, and some even continued for more than thirty minutes although the robot responded only visually. They seemed to try to uncover the attention mechanism of the robot through the long interaction. This result indicates that the primal but adaptable attention mechanism allowed people to partially understand the robot's behavior, and consequently motivated them to continue to communicate.

The attention mechanism also enabled the robot to take the initiative of communication. When the robot was distracted by a visual disturbance, some subjects tried to follow the line of its gaze in order to achieve joint attention. In contrast to the current studies on robotic joint attention (e.g., [20]–[22]), in which a robot could only follow the human gaze, our robot was able to lead the communication. Moreover, the people's response of following the robot's initiative indicates that they assumed some sort of intention of the robot. Thus, we suggest that the adaptable attention behavior of the robot enabled it to be accepted as an intentional and proactive communication agent.

Furthermore, the attention mechanism had the effect of encouraging people to modify their actions as in motionese [23]. Motionese, which is parental modifications in their infant-directed actions, is suggested to help robots as well as infants to detect the meaningful structure of the actions [19], [24], [25]. Exaggerating actions like moving larger and closely showing objects, which were observed in our experiment, is a typical phenomenon of motionese. Increasing communication channels is also suggested to support the understanding of actions [26], [27]. Nagai and Rohlfing [19] showed that the same attention mechanism as used in our experiment for the robot's vision can leverage the benefit of motionese. We will further investigate how the attention behavior of the robot influences people's demonstration of actions.

VI. Conclusion

Interaction with robotic partners will become more common in the near future. A promising approach to designing HRI is to learn from HHI. Sociological studies on HHI indicate that an interruption in the communication can cause positive effects. Thus we focused on disturbance, an occurrence that is usually not appreciated in communications, and examined how it affects HRI. Our experiment showed that human partners started to exaggerate their communicative activity and to increase their communication channels when realizing that the robot shifted its interactive orientation. Human partners moreover tried to achieve joint attention by following the line of the robot's gaze. We conclude with these results that disturbance does not always discourage people but can rather encourage them to make the communication with a robot more social.

Acknowledgment

This work has partly been funded by a fellowship of the Sozialwerk Bielefelder Freimaurer e.V.

We thank Masaki Ogino, JST ERATO Asada Synergistic Intelligence Project, Japan, for kindly providing the robot's face simulation.

References

[1] Terrence Fong, Illah Nourbakhsh, and Kerstin Dautenhahn. A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42:143–166, 2003.

[2] Jennifer Goetz, Sara Kiesler, and Aaron Powers. Matching robot appearance and behavior to tasks to improve human-robot cooperation. In *Proceedings of the 12th IEEE International Workshop on Robot and Human Interactive Communication*, pages 55–60, 2003.

[3] Takashi Minato, Michihiro Shimada, Shoji Itakura, Kang Lee, and Hiroshi Ishiguro. Evaluating the human likeness of an android by comparing gaze behaviors elicited by the android and a person. *Advanced Robotics*, 20(10):1147–1163, 2006.

[4] Masahiro Mori. Bukimi no tani (the uncanny valley). *Energy*, 7(4):33–35, 1970.

[5] Minoru Asada, Karl F. MacDorman, Hiroshi Ishiguro, and Yasuo Kuniyoshi. Cognitive developmental robotics as a new paradigm for the design of humanoid robots. *Robotics and Autonomous Systems*, 37:185–193, 2001.

[6] Max Lungarella, Giorgio Metta, Rolf Pfeifer, and Giulio Sandini. Developmental robotics: a survey. *Connection Science*, 15(4):151–190, 2003.

[7] Charles Goodwin and John Heritage. Conversation analysis. *Annual Reviw of Anthropology*, 19:283–307, 1990.

[8] Alfred Schütz. *Collected Papers*, volume 1: The problem of the social reality. The Hague: Nijhoff, 1962.

[9] Niklas Luhmann. *Social Systems*. Stanford University Press, Stanford, CA, 1995.

[10] Charles Goodwin and Majorie Harness Goodwin. Conurrent operations on talk: Notes on the interactive organization of assessments. In *IPRA Papers in Pragmatics*, volume 1, pages 1–54, 1987.

[11] Charles Goodwin. *Discourse of the Body*, pages 19–42. New York: Palgrave/Macmillan, 2003.

[12] Harvey Sacks, Emanuel A. Schegloff, and Gail Jefferson. A simplest systematics for the organization of turn-taking for conversation. *Language*, 50:696–735, 1974.

[13] Harold Garfinkel. *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice Hall, 1967.

[14] Charles Goodwin. *Notes on story structure and the organization of participation*. Cambridge: Cambridge University Press, 1984.

[15] Masaki Ogino, Ayako Watanabe, and Minoru Asada. Mapping from facial expression to internal state based on intuitive parenting. In *Proceedings of the Sixth International Workshop on Epigenetic Robotics*, pages 182–183, 2006.

[16] Laurent Itti, Christof Koch, and Ernst Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1254–1259, 1998.

[17] L. Itti, N. Dhavale, and F. Pighin. Realistic avatar eye and head animation using a neurobiological model of visual attention. In *Proceedings of the SPIE 48th Annual International Symposium on Optical Science and Technology*, pages 64–78, 2003.

[18] Charles C. Kemp and Aaron Edsinger. What can i control?: The development of visual categories for a robot's body and the world that it influences. In *Proceedings of the 5th International Conference on Development and Learning*, 2006.

[19] Yukie Nagai and Katharina J. Rohlfing. Can motionese tell infants and robots "what to imitate"? In *Proceedings of the 4th International Symposium on Imitation in Animals and Artifacts*, pages 299–306, 2007.

[20] Brian Scassellati. Theory of mind for a humanoid robot. *Autonomous Robots*, 12:13–24, 2002.

[21] Yukie Nagai, Koh Hosoda, Akio Morita, and Minoru Asada. A constructive model for the development of joint attention. *Connection Science*, 15(4):211–229, 2003.

[22] Yukie Nagai, Minoru Asada, and Koh Hosoda. Learning for joint attention helped by functional development. *Advanced Robotics*, 20(10):1165–1181, 2006.

[23] Rebecca J. Brand, Dare A. Baldwin, and Leslie A. Ashburn. Evidence for 'motionese': modifications in mothers' infant-directed action. *Developmental Science*, 5(1):72–83, 2002.

[24] Jannik Fritsch, Nils Hofemann, and Katharina Rohlfing. Detecting 'when to imitate' in a social context with a human caregiver. In *Proceedings of the ICRA Workshop on Social Mechanisms of Robot Programming by Demonstration*, 2005.

[25] Katharina J. Rohlfing, Jannik Fritsch, Britta Wrede, and Tanja Jungmann. How can multimodal cues from child-directed interaction reduce learning complexity in robot? *Advanced Robotics*, 20(10):1183–1199, 2006.

[26] Jana M. Iverson, Olga Capirci, Emiddia Longobardi, and M. Cristina Caselli. Gesturing in mother-child interactions. *Cognitive Development*, 14:57–75, 1999.

[27] Lakshmi J. Gogate, Lorraine E. Bahrick, and Jilayne D. Watson. A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child Development*, 71(4):878–894, 2000.