UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Nonverbal Behaviors in Cooperative Work: A Case Study of Successful and Unsuccessful Team

Permalink
https://escholarship.org/uc/item/5296n150

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 29(29)

ISSN
1069-7977

Authors
Suzuki, Noriko
Umata, Ichiro
Kamiya, Toshiro
et al.

Publication Date
2007

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Nonverbal Behaviors in Cooperative Work: A Case Study of Successful and Unsuccessful Team

Noriko Suzuki (noriko@nict.go.jp)[(1)], **Ichiro Umata (umata@nict.go.jp)**[(1)],
Toshiro Kamiya (to_kamiya@hotmail.com)[(2)], **Sadanori Ito (sito@cc.tuat.ac.jp)**[(3,1,4)]
Shoichiro Iwasawa (shoichiro.iwasawa@nict.go.jp)[(1)], **Naomi Inoue (na-inoue@nict.go.jp)**[(1)]
Tomoji Toriyama (toriyama@atr.jp)[(4)], **Kiyoshi Kogure (kogure@atr.jp)**[(4)]
(1) National Institute of Informathon and Communication Technology (NICT)/
ATR Cognitive Information Science Laboratories (ATR-CIS)
2-2-2 Hikaridai, Keihanna Science City, Kyoto, 619-0288 JAPAN
(2) Tokyo University of Foreign Studies
3-11-1, Asahi-cho, Fuchu-shi, Tokyo, 183-8534, JAPAN
(3) Tokyo University of Agriculture and Technology/NICT/ATR-KSL
2-24-16 Naka-cho, Koganei-shi, Tokyo 184-8588, JAPAN
(4) ATR Knowledge Science Laboratories (ATR-KSL)
2-2-2 Hikaridai, Keihanna Science City, Kyoto, 619-0288 JAPAN

Abstract

This paper examines the differences in nonverbal behaviors between successful and unsuccessful cases of cooperative work in a team setting. A preliminary experiment was conducted on the task of assembling a large structure in a shared physical space. The speech, gaze and body motion of all participants were captured with ubiquitous sensors. Speech and gaze data in both successful and unsuccessful cases were analyzed. The results suggest that mutual gaze and shared gaze play an important role in facilitating interaction in triad cooperative work.

Keywords: Nonverbal behaviors; gaze; speech; teamwork; triad cooperative work; capturing with ubiquitous sensors.

Introduction

Cooperative activities are sometimes carried out correctly and effectively even when the participants are not familiar with each other. Their behaviors coordinate with each other during the work without the participants exchanging many words. What is the key for smooth and responsive interaction in such an unfamiliar environment? What is the difference in nonverbal behaviors between successful and unsuccessful cases of teamwork?

This paper focuses on how nonverbal behaviors, especially gaze, contribute to the development of teamwork through cooperative work in a team setting. In other words, we are interested in how nonverbal behaviors express partnership during interaction, especially in large spaces.

Some studies of human-human interaction have pointed out that nonverbal behaviors, including gaze, play an important role in interaction (Clark, 1996; Jarmon, 1996; Kendon, 1967; Otsuka, Yamato, Takemae, & Murase, 2006). In particular, mutual gaze has been regarded as a key to starting and maintaining interaction (Argyle & Cook, 1976; Vertegaal & Ding, 2002). On the other hand, the development of information commutation technology, including ubiquitous sensors, has made it possible to capture nonverbal behaviors objectively and automatically. For example, using eye-tracker data showed that eye movement expresses thought processes in human problem solving (T. Ito, Matsubara, & Grimbergen, 2005; S. Ito, Iwasawa, Umata, & Kogure, 2006). However, few research works have captured and analyzed the nonverbal behaviors of participants in teamwork while moving.

In the present task, people meeting each other for the first time were instructed to assemble a large structure consisting of pipe components. We chose a large structure similar to ”pipe furniture” by referring to the TV-cart assembly task (Lozano & Tversky, 2004). However, in this task, the pipe structure was preferred to the TV cart because it permits easy control of the level of configuration difficulty in assembly; furthermore, this structure facilitates the attachment and tracking of ubiquitous sensors. The participants in the task could move in a shared physical space to assemble the large structure, while their gaze, speech, and motion data were captured by sensors. This paper presents the differences in nonverbal behaviors between successful and unsuccessful cases of teamwork. To this end, the contributions of speech and gaze, including mutual and shared gaze, were analyzed in the process of assembling the target structure.

Method

Participants

Six male workers from 25 to 35 years old were recruited as participants. They had a semi-professional level of skill in assembling structures using pipes for setting up event sites. Two groups, each consisting of three participants, took part in this preliminary experiment.

Materials

We prepared a large structure consisting of pipes, similar to ”pipe furniture,” as the target object of this cooperative work task. Figure 1 (a) shows a picture of the completed large structure used in this experiment. Experimental materials (Figure 1 (b)) consisted of twelve types of components, including 30 pipes ranging from 35 to 150 cm and 26 joints (Figure 1 (d) and (f)). Each component was constructed from

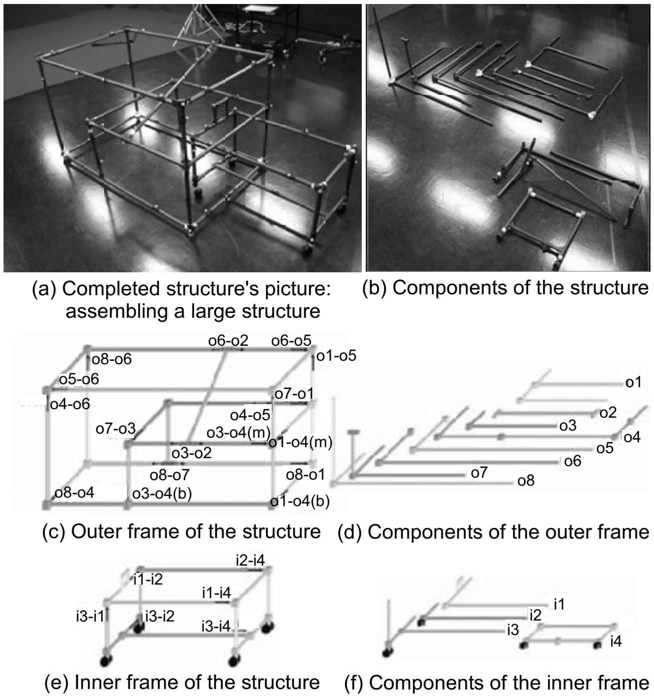

Figure 1: Materials used in the task of assembling a structure: completed picture and twelve kinds of components; Construction materials: outer frame, inner frame and pipes.

different kinds and numbers of pipes and joints. The structure was composed of two main frames, i.e. the inner and outer frames (Figure 1 (c) and (e)).

Design of the task

The focus of our observations was the effect of the participants' gaze and motion during cooperative work. In designing the experimental setting for cooperative work, we constructed a large structure made from assembled pipes while considering the following points:

1. **Size of structure:** We adopted structural components that were so large that it was difficult for a single participant to construct the structure on his own. When the participants joint or disjoint two components in assembling the structure, two or more of them should be engaged in this work at the same time. The size of the components as well as the structure make collaborative assembly work a practical necessity.

2. **Configuration of structure:** The large structure consists of two main frames: the outer and the inner frame. Three participants are assigned to construct two frames. Thus the correspondence between the number of participants and the number of frames is not a one-to-one relationship. This configuration of the structure is designed to let us observe the conditions of collaborative relationships among the participants.

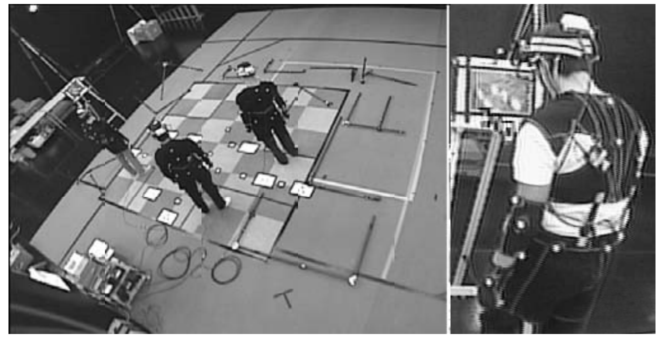

Figure 2: Large-structure assembly task in shared physical space and devices worn by participants.

3. **Complexity of assembly procedure:** We designed the assembly procedure of the large structure to be like solving a wire puzzle. The participants could not complete the large structure by assembling the two frames after constructing them independently. They needed to start construction of the inner frame during construction of the outer frame. The complexity of the assembly procedure was designed to require the participants to reassemble frames at least once.

4. **Type of components:** Pipes are selected as the components of the structure in this experiment because their shape makes it easy to observe the movement of the participants' gaze, since they do not include any planar surface. Moreover, attaching the sensors on the components makes it easy to identify the position of the participants' gaze on the part of a pipe or joint.

5. **Similarity among components:** the detailed shape of the twelve components are different except for pipe-o5 and pipe-o6. However, it is difficult to distinguish them at a glance because these components consist of similar pipes and joints with similar color and location of sensors. The similarity in shape among the components is expected to promote communication among participants in order to confirm their activities toward task completion.

Capturing environment

A Vicon Motion Capture System with 60-Hz time resolution and 1-mm space resolution captured body motions and locations using optical markers. In this preliminary experiment's setting these factors were captured within the following space: 750 x 650 x 250 (width/depth/height) (see Figure 2, left). The directions of the participants' gazes were measured by head-mounted eye trackers (EMR-8B, Nac Image Technology Inc.) with 30-Hz time resolution and 0.15-degree angle resolution. The speech of participants was recorded using close-proximity microphones.

Each participant wore a cap with an eye tracker, a close-proximity microphone, and body suits with optical markers for the motion capture device (Figure 2, right). The following

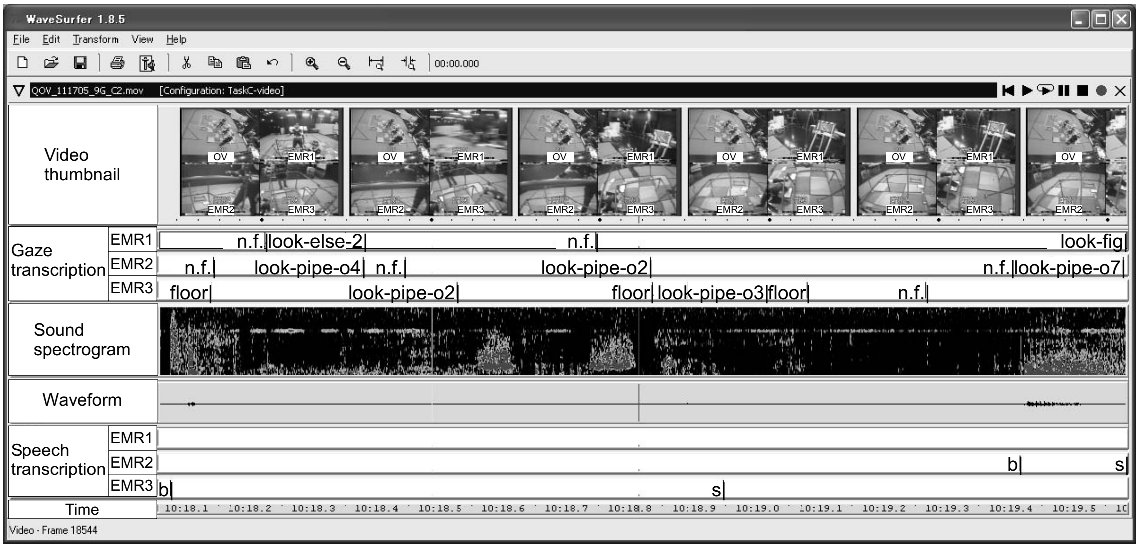

Figure 3: Sample image of data coding.

information was recorded synchronously: visual images, 2-D coordinates of gaze, 3-D coordinates of motion, and the voices of the participants.

Procedures

Capturing devices were attached to participants as described in the previous section. After calibration of the capturing devices, participants were instructed to assemble a large structure by using twelve kinds of components and by referring to the completed picture (Figure 1(a)). They were also informed that the task's time limit was twenty minutes. No participant was assigned a particular role, such as leader or director. At the start of this task, all participants stood on even ground.

Predictions

A successful team, as opposed to an unsuccessful team, is predicted to communicate with one another by using the following nonverbal behaviors in assembling the large structure smoothly:

(a) A successful team is expected to exchange more speech than an unsuccessful team.

(b) A successful team is expected to look mutually at task objects more frequently than an unsuccessful team.

(c) A successful team is expected to look at the same object, such as the picture of the completed structure or the components, more frequently than an unsuccessful team.

Coding

We analyzed the speech and gaze units of each participant in two groups: successful and unsuccessful cases of task achievement (e.g., Figure 3). Two labelers coded each unit by using the transcription labels in Table1 as follows:

Speech: A speech unit in cooperative work is the duration of a single participant's speech bounded by pauses longer than 200 ms. The speech unit includes back-channel responses, i.e., hai (yeah.), sou sou (that's right.). However, it does not include laughing, coughs and breathing sound.

Gaze: The directions of the participants' gazes are estimated from the positions of eye mark characters superimposed on a view image of the participants.

Table 1: Transcription label for coding.

Speech label	
Speech	s
No speech	b
Gaze label	
Participants	look-face[1-3] look-hand-[1-3] look-else-[1-3]
Parts of large structure	look-joint-i[1-4]-i[1-4] look-joint-o[1-8]-o[1-4] look-pipe-i[1-4] look-pipe-o[1-8]
Complete figure	look-fig
Others	cable floor n.f.

Results

Two group, C1 and C2, took part in the task of assembling the large structure. Group C2 succeeded in assembling the large structure within 787 sec, while group C1 failed to complete the task within 20 minutes.

Figure 4 shows the following coding results by the labelers; phase of the task, speech distribution, gaze distribution

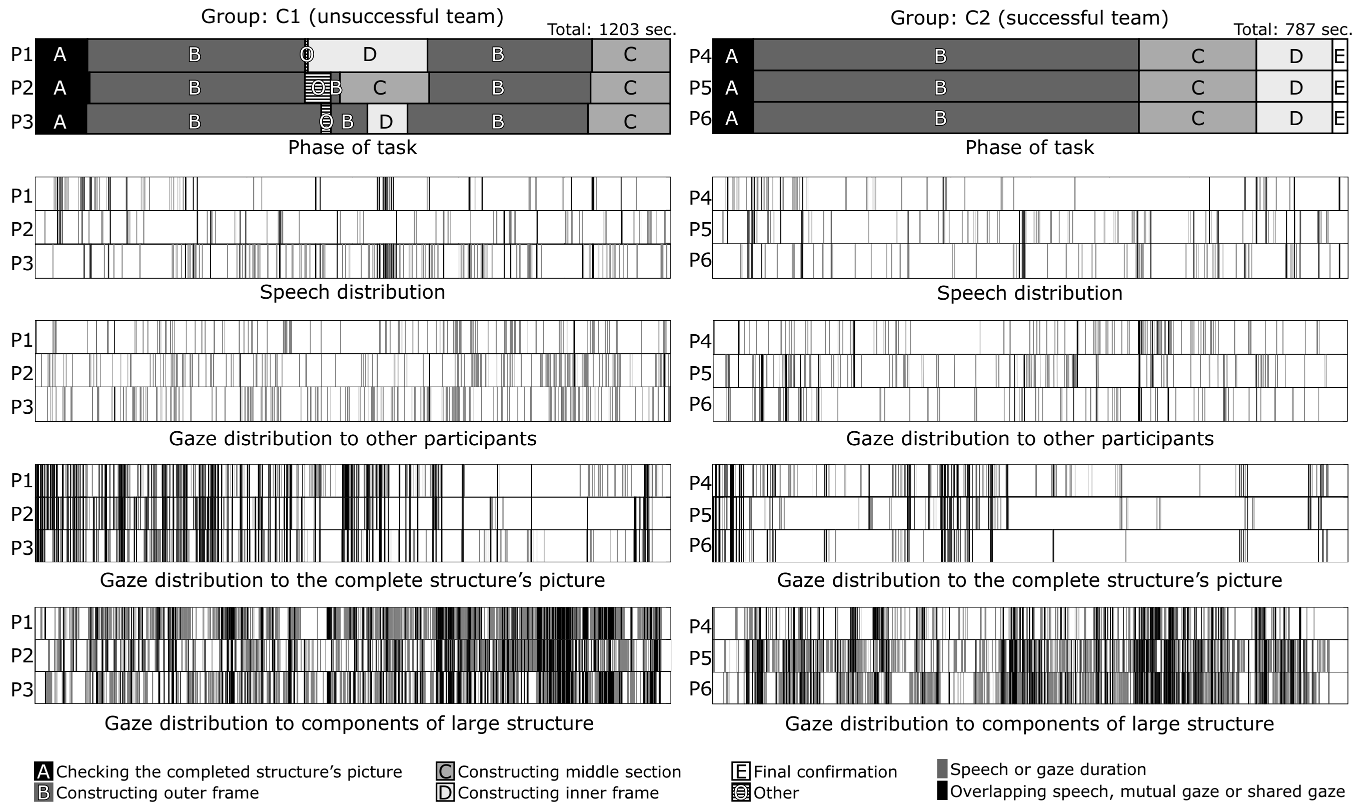

Figure 4: Process of assembling a large structure: utterance and gaze distributions

to other participants, gaze distribution to the completed structure's picture, and gaze distribution to components of the large structure.

Observation results: phase of cooperative work

Two labelers classified the task of the two groups into the following phases:
A. Checking the completed structure's picture.
B. Constructing the outer frame.
C. Constructing the middle section of the outer frame.
D. Constructing the inner frame.
E. Final confirmation.
O. Others.

The upper part of Figure 4 shows phases of the task for both groups, C1 and C2. From the data on the C1 group, every participant constructed a different frame of the structure, especially in the middle of the task. On the other hand, from the data on the C2 group, three participants constructed the same frame together.

Speech

The second figures from the top in Figure 4 show the speech distribution of the C1 and C2 groups. Table 2 shows the ratio (%) and frequency (times/minute) of speech of each group during the task. P1 to P3 indicate members of the C1 group, while P4 to P6 show the C2 group's members. No significant difference was found by using the Mann-Whitney U test for ratio and frequency of either single speech or overlapping speech. These results did not demonstrate that the members of the successful C2 group exchanged more speech than did those of the unsuccessful C1 group. Consequently, prediction (a) was not verified.

Table 2: Ratio and frequency of speech.

Speech					
	C1		C2		p-value
Ratio (%)	P1	7.02	P4	9.89	.8273
	P2	6.14	P5	8.85	
	P3	12.73	P6	6.92	
Frequency (times/min.)	P1	3.14	P4	5.71	.8273
	P2	3.64	P5	5.25	
	P3	6.48	P6	3.63	
Overlapping speech					
	C1		C2		p-value
Ratio (%)	P1	1.47	P4	1.71	.8273
	P2	.82	P5	1.38	
	P3	1.53	P6	1.08	
Frequency (times/min.)	P1	1.49	P4	2.01	.2752
	P2	1.05	P5	1.77	
	P3	1.64	P6	1.47	

Gaze to other participants

The middle figures in Figure 4 show the gaze distribution to other participants for both C1 and C2 groups. Table 3 shows the ratio (%) and frequency (times/minute) for each group during the task. No significant difference was found by using the Mann-Whitney U test for either ratio or frequency of gaze to other participants or ratio of shared gaze to the same participant. On the other hand, there were significant differences in both ratio and frequency of mutual gaze between two participants (Mann-Whitney U test: p = **.0495** and p = **.0463**) and frequency of shared gaze to the same participant (p = **.0495**). From these results, the members of the successful C2 team might look mutually at one another and also might look at the same person at the same time more frequently than do the members of the unsuccessful C1 team. Therefore, prediction (b) was partly verified.

Table 3: Ratio and frequency of gaze to other participants.

	C1		C2		p-value
Gaze to other participants					
Ratio (%)	P1	4.74	P4	8.05	.5127
	P2	6.30	P5	7.94	
	P3	6.16	P6	4.37	
Frequency (times/min.)	P1	6.98	P4	8.87	.8273
	P2	9.68	P5	11.57	
	P3	11.67	P6	8.18	
Mutual gaze between two participants					
Ratio (%)	P1	.02	P4	.13	**.0495**
	P2	.04	P5	.18	
	P3	.03	P6	.27	
Frequency (times/min.)	P1	.10	P4	.31	**.0463**
	P2	.20	P5	.46	
	P3	.20	P6	.62	
Shared gaze to the same participant					
Ratio (%)	P1	.25	P4	.33	.1266
	P2	.24	P5	.41	
	P3	.30	P6	.28	
Frequency (times/min.)	P1	.65	P4	.69	**.0495**
	P2	.60	P5	1.08	
	P3	.55	P6	.85	

Gaze to the completed structure's picture

The fourth figures from the top in Figure 4 show the gaze distribution to the completed structure's picture for both C1 and C2 groups. Table 4 shows the ratio (%) and frequency (times/minute) of each group during the task. No significant difference was found by using the Mann-Whitney U test in either ratio or frequency of gaze to the picture of the completed structure. On the other hand, there were significant differences in both ratio and frequency of mutual gaze to the completed structure's picture (Mann-Whitney U test: p = **.0495** and p = **.0495**). From these results, the members of the unsuccessful C1 team might have looked at the completed structure's picture at the same time rather more frequently than did those of the successful C2 team. Accordingly, prediction (c) was not verified.

Gaze to components of the large structure

The fifth figures from the top in Figure 4 show the gaze distribution to the components of the large structure for both C1 and C2 groups. Table 5 shows the ratio (%) and frequency

Table 4: Ratio and frequency of gaze to the completed structure's picture.

	C1		C2		p-value
Gaze to the completed structure's picture					
Ratio (%)	P1	37.31	P4	26.27	.2752
	P2	18.47	P5	15.27	
	P3	26.21	P6	16.12	
Frequency (times/min.)	P1	9.03	P4	8.02	.2752
	P2	4.89	P5	5.40	
	P3	13.22	P6	2.70	
Shared gaze to the completed structure's picture					
Ratio (%)	P1	22.57	P4	19.38	**.0495**
	P2	30.04	P5	15.79	
	P3	20.06	P6	17.13	
Frequency (times/min.)	P1	9.28	P4	6.33	**.0495**
	P2	13.22	P5	4.71	
	P3	9.83	P6	5.17	

(times/minute) for each group during the task. No significant difference was found by using the Mann-Whitney U test for ratio and frequency of either gaze to the components of the large structure or mutual gaze to the same components. These results did not indicate that the members of the successful C2 team looked at the same components' picture at the same time as did those of the unsuccessful C1 team. Again, prediction (c) was not verified.

Table 5: Ratio and frequency of gaze to components of the large structure.

	C1		C2		p-value
Gaze to components of the structure					
Ratio (%)	P1	42.76	P4	29.20	.8273
	P2	56.53	P5	47.50	
	P3	40.27	P6	64.37	
Frequency (times/min.)	P1	44.69	P4	25.99	.5127
	P2	65.04	P5	55.16	
	P3	57.96	P6	64.95	
Shared gaze to the same component of the structure					
Ratio (%)	P1	10.06	P4	9.97	.2752
	P2	10.36	P5	13.33	
	P3	9.22	P6	14.77	
Frequency (times/min.)	P1	12.27	P4	11.11	.5127
	P2	13.27	P5	16.82	
	P3	13.37	P6	17.13	

Discussion

Two groups assembled a large structure, similar to "pipe furniture," by using twelve kinds of pipe components. One group succeeded in building the structure within the time limit of twenty minutes. On the other hand, the other group failed to complete the task. This case study showed that the successful and unsuccessful teams used different nonverbal behaviors, especially gaze, as they assembled the structure. The members of the successful team looked mutually and looked at the same participant at the same time more frequently and longer than do the members of the unsuccessful team. The members of the unsuccessful team looked at the same picture of the completed structure at the same time more frequently than the successful team. From the observation results, all members of the successful team assembled as a group the same frame of the structure, while members of the unsuccessful team assembled different frames in the middle of the task. These results suggest that mutual and shared gazes to the other participants play an important role in the smooth progression of the task of assembling a large structure. However, due to the small scale of the preliminary experiment, the results cannot be generalized to other types of cooperative work. This is because the use of nonverbal behavior may vary according to the size of the task space, the materials used, and the number of participants.

Our findings have two main implications. First, mutual gaze and shared gaze to the same worker seemed to facilitate the process of the cooperative work. In this task, the participants of neither group seemed to make conversation with each other. The ratio and frequency of both speech and gaze to other participants were not much different between the groups (Table 2 and Table 3). It's possible that the participants pay attention to and monitor the behaviors of other participants by using mutual gaze and shared gaze to the same participant as well as using conversation. As a second implication, shared gaze to the same object seemed to confirm the process of cooperative work. In this task, the participants of both groups seemed to look at the objects, including the pictures of the completed structure and its components, rather than at the other participants. The ratio and frequency of gaze to the completed structure's picture and to the components of the architecture were much different (Table 3 from the results in Table 5). In particular, the unsuccessful team more often seemed to look at the same picture of the completed structure at the same time than did the successful team. They might use this shared gaze to correct their assembly errors and to check the status of the task.

Previous research has regarded gaze as playing an important supporting role to the main activity of conversation (Clark, 1996; Jarmon, 1996; Kendon, 1967; Otsuka et al., 2006). Mutual gaze is also considered to support turn-taking in conversation and to improve one's ability to give intimate impressions to the interaction partner (Argyle & Cook, 1976; Vertegaal & Ding, 2002). On the other hand, authors note another case study of visual measurement, i.e. virtual tracing of component size and shape, using the same task as used in this paper (Suzuki et al., 2006). Gaze is assumed to take precedence in interaction not only for deciding the next move in non-conversation games, such as Shogi (Japanese Chess), played on desktop space (T. Ito et al., 2005) but also for engaging others in cooperative work carried out in a large space.

Some of the authors have joined the Ultra-realistic Communication System Group in the National Institute of Information and Communication Technology of Japan (NICT). The purpose of the project is to construct a system for embodied interaction with remote partners by improving the reality of interaction, i.e., using 3-D video, 3-D audio, olfaction and tactile sense. A multi-point cooperative work system using a large space is one of the candidates for an ultra-realistic communication system. Implications from this face-to-face cooperative work task could be applied to the design and evaluation method of such a system with remote partners.

Acknowledgments

The research reported here was supported in part by a contract with the National Institute of Information and Communication Technology of Japan (NICT) entitled, "A study of innovational interaction media toward a coming high functioned network society." The authors thank Takugo Fukaya, Kazumi Asai, Kohei Suzuki and Dr. Kenji Susami for their technical support.

References

Argyle, M., & Cook, M. (1976). *Gaze and mutual gaze*. London: Cambridge University Press.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Ito, S., Iwasawa, S., Umata, I., & Kogure, K. (2006). A model for interest measurement by observable nonverbal behavior. In *Proceedings of cogsci2006* (pp. 396–374).

Ito, T., Matsubara, H., & Grimbergen, H. (2005). A cognitive science approach to shogi playing processes some results on next move test experiment. In *Proceedings of icga2005/acg11*.

Jarmon, L. H. (1996). *An ecology of embodied interaction: turn-taking and interactional syntax in face-to-face encounters*. Ph. D. Dissertation: University of Texas at Austin.

Kendon, A. (1967). Some functions of gaze direction in social interaction. *Acta Psychologica*, *32*, 1–25.

Lozano, S. C., & Tversky, B. (2004). Communicative gestures benefit communicators. In *Cogsci2004.*

McNeil, D. (2006). Gesture, gaze, and ground. In *Mlmi2005, lncs3869* (pp. 1–14). Springer-Verlag.

Otsuka, K., Yamato, J., Takemae, Y., & Murase, H. (2006). Quantifying interpersonal influence in face-to-face conversations based on visual attention patterns. In *Proceedings of chi2006* (pp. 1175–1179).

Suzuki, N., Ito, S., Umata, I., Iwasawa, S., Kamiya, T., Toriyama, T., et al. (2006). Nonverbal behaviors in cooperative work: interaction in assembling a large object. In *Proceedings of iccs2006* (pp. 195–196).

Vertegaal, R., & Ding, Y. (2002). Explaining effects of eye gaze on mediated group conversations: amount or synchronization? In *Proceedings of cscw02* (pp. 41–48).