Human-Computer Interaction -- INTERACT'03
M. Rauterberg et al. (Eds.)
Published by IOS Press, (c) IFIP, 2003, pp. 713-716

An analysis of participation structure in conversation based on Interaction Corpus of ubiquitous sensor data

Mayumi Bono† ‡, Noriko Suzuki† & Yasuhiro Katagiri† ‡

† ATR Media Information Science Laboratories,

2-2-2 Hikaridai, Keihanna Science City, Kyoto, JAPAN 619-0288

‡ Graduate School of Cultural Studies and Human Science, Kobe University

{bono, noriko, katagiri}@atr.co.jp

Abstract: We propose the application of ubiquitous sensor technology to analyze dynamic structure transitions and the exchange of non-verbal information in multi-party human-to-human conversational interactions in open situations. We present an analysis of conversational interactions that took place in an open interaction space of poster presentation sessions and were collected as an Interaction Corpus. Based on a model of conversational participation structure, we hypothesized that the audience diversity in conversations can be understood by observing the role of each person standing in front of a poster and an exhibitor. We argue that the conversational interactions observed in the Interaction Corpus can effectively be captured by the differences in the patterns of transition of the roles each participant plays in the conversational participation structure.

Keywords: participation structure, ubiquitous sensor, interaction corpus, non-verbal information

1 Introduction

A natural human conversation requires more than a mere exchange of utterances between conversational participants. Conversational participants first need to be established and mutually admitted into the conversation before they can engage in it. Participants play certain roles in the conversation, and these roles change during the course of the conversation. New participants may join and old participants may leave. These dynamic changes are signaled and managed through the use of various verbal and non-verbal cues. Many studies have been conducted in the fields of sociology and anthropology on the structure of conversations and interactive communities. However, very little work has been done in the CHI community, despite the increasing recognition of the importance of non-verbal information and its functions in human-to-human interaction. For example, elucidating the dynamic transitions of conversation participation structures is essential to developing robotic/electronic agents that can interact with and help humans in daily life situations. Dynamic information on the structure of conversation participation can also be used to organize and summarize one's personal history of social interaction (Hagita et al., 2003).

We propose in this paper that ubiquitous sensor technology can be used to collect and analyze non-verbal, as well as verbal, information exchanges in human conversations, thereby providing a source for constructing and testing models of human dynamic conversation participation structures.

2 Background

2.1 Participation structure

Some sociologists and anthropologists (Goffman, 1981) who have analyzed the phases of interaction have defined the internal structure of conversational space as 'participation structure' or 'participation framework.' In conversation, the participants exchange their roles, such as 'Speaker' and

'Addressee,' by exchanging the right of utterance for a moment. The structure of conversation (participation structure) consists of components (participants) and is not static but dynamic. Clark (1996) proposed a model of participant relationships as follows.

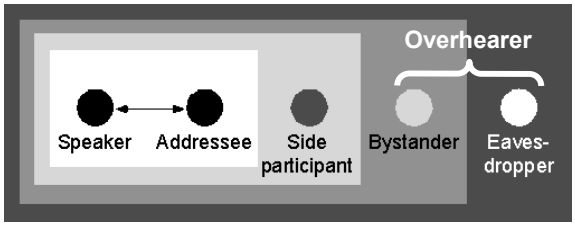

Figure 1. Participation structure (Clark, 1996)

Clark defined speaker as the agent of the illocutionary act and the speaker designs addressee as the participant who is the partner of the joint action the speaker is projecting for them to perform. Side participants take part in the conversation but are not currently being addressed. All other listeners are overhearers who have no rights or responsibilities in the conversation, that is, they don't take part in it. There are two main types of overhearers: bystanders are openly present but not part of the conversation and eavesdroppers listen in without the speaker's awareness.

2.2 Audience design

The term 'design' that we have referred to in the last section involves 'Audience design' (Clark & Carlson, 1982). In conversation with three or more parties, one person can speak to the others, that is, those who are the audience. Though a simple gathering of listeners is called an 'audience,' they don't have equal rights and responsibilities in listening to the speaker's utterance. Clark & Carlson (1982) suggested that a speaker designs an utterance for a specified listener who is assigned the role of addressee. The speaker designs his/her utterance by including topics that only the listener knows, namely mutual knowledge, through a common background with him or her. Although studies have been made on these issues, such as knowledge and information in the participant's mind, by observing the contents and forms of utterances to derive a pragmatic interpretation, there has been little work on non-verbal behaviors.

In this study, we examine the relevance of non-verbal information, specifically human posture and view. We explore the possibilities of analyzing these types of non-verbal information in directing utterances to someone; in concrete terms, we investigate whether the speaker's body posture is directed to the addressee and whether the speaker tends to put the addressee into his or her view.

3 Materials and Corpus Design

Elucidating the above issues requires the collection of data for encounter and leaving events in conversation. We used the Interaction Corpus (Hagita et al., 2003) constructed during the ATR Exhibition 2002, when a wide variety of people from outside ATR visited poster and oral presentations and joined demonstrations. This corpus consists of the data that were gathered in the Ubiquitous Sensor Room, which had numerous sensors and cameras for recording the behaviors of exhibition participants. This room had five booths where the exhibitors held poster presentations as shown in Figure 2.

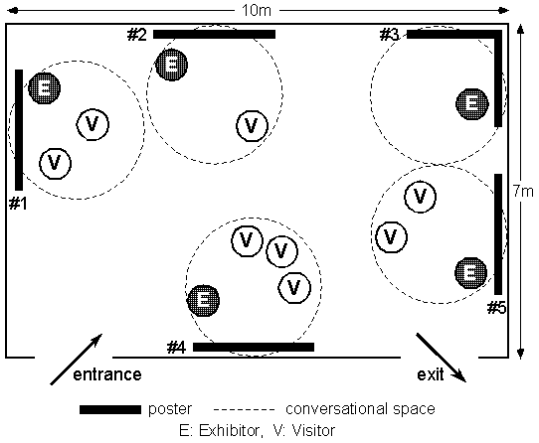

Figure 2. Ubiquitous sensor room.

Two cameras were fixed to the ceiling of each booth to capture human behavior. We assume that there are five 'conversational spaces' where people conduct interaction in this room, corresponding to each poster presentation booth.

Furthermore, all of the participants wore headset microphones with sensors and a camera that detected their postures and views. We observed the relationships of their postures and views in the conversation by using the recorded data, which indicate patterns of interaction in the participants' body behaviors.

4 Hypothesis

In general, it has been proposed that the duration of staying in front of an exhibit is related to the participant's taking an interest in it. However, we assume that it is not possible to fully understand a participant's interests simply by the duration of

staying. Beyond this measure, we should also observe the activities of each participant.

4.1 Two phases of participation

We assume that Clark's model of participation structure indicates a natural organization of two phases of participation. One is participation in the conversational space and the other is participation in the conversation itself. For example, in a poster presentation conversation, the visitor reduces physical distance to the poster to hear the exhibitor's speech and to look at the poster contents in detail, which amounts to participation in the conversational space (i.e., being promoted from non-participant to bystander).

4.2 Conversational roles

Beyond participation in the conversational space, the participant needs to be further promoted to enter the conversation itself. We assume that the participant either takes the floor of the conversation himself/herself (i.e., being promoted to speaker), is assigned the role of addressee by the current speaker (i.e., being promoted to addressee), or is admitted to join by receiving the speaker's gaze, namely, the recognition of his/her existence by the existing participants (i.e., being promoted to side participant). These are some of the possibilities explaining how participation progresses in conversation.

4.3 Definitions

Participation in conversational space

Non-participant to bystander:

Promotion depending on audience's active behavior

Intentionally coming close and staying for a while.

Participation in conversation

Bystander to side participant:

Promotion depending on Audience's passive behavior

Receiving the gaze of existing participants

Side participant to addressee:

Promotion depending on Audience's passive behavior

Receiving the utterance that the current speaker designs and produces for him/her

Side participant to speaker:

Promotion depending on audience's active behavior

Intentionally speaking to existing participants

The process of participation in conversation is closely tied to audience design by the speaker, in the sense that some of the transitions (e.g., bystander to side participant and side participant to addressee) need to be sanctioned by the speaker.

5 Case studies

Here, we present our initial results of analyzing the Interaction Corpus (Hagita et al., 2003) by using concepts of participation structure. We focus on situations in which a third participant (second visitor) joins the already established conversation between two participants (poster exhibitor and first visitor).

5.1 Case 1

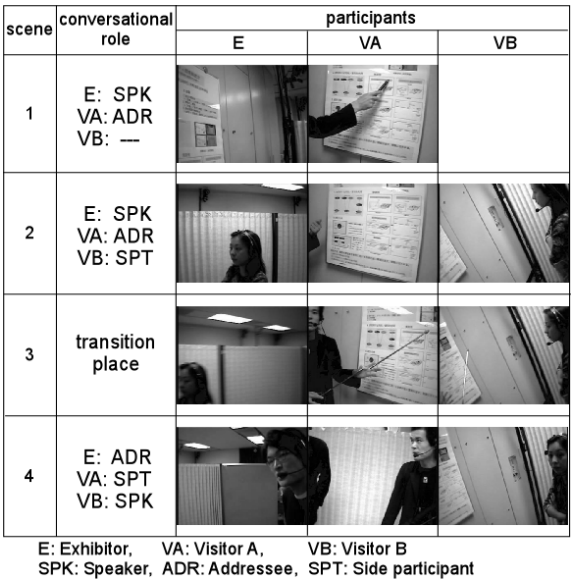

Figure 3. Case 1

Figure 3 shows the sequence of events that took place in Case 1. First, the person standing in front of the poster is the Exhibitor (E), who has an absolute right to be a speaker owing to his knowing the contents of the poster in this situation. The others are visitors who came to listen to the presentation. One of them (Visitor A: VA) came to this booth before the other visitor (Visitor B: VB) (in scene 1). After E and VA exchanged utterances and shared the floor of the conversation for a while, VB arrived (in scene 2). The second column from the left in Figure 3 indicates the role of each participant in this conversation. The next three columns shows the subjects' view data taken by the wearable cameras in this order of participants: E, VA and VB. Whenever they exchange turns, we add a new row. In scene 2, E is speaker (SPK), and we assume he directs his body posture to VA by observing E's view image data at scene 2, that is, VA is addressee (ADR) and VB is side participant (SPT) in Clark's categorization.

In this sequence of events, audience behaviors, that is, behaviors of VA and VB, are exactly opposite: VA is passive and VB is active. The role of VA was demoted from ADR to SPT, and she did not have the right of turn. On the other hand, VB

was promoted from SPT to SPK, so he produced speech and directed it to E. Even though VA stays at this booth longer than VB, VB is more active than VA. This case study indicates that duration of staying time alone is not sufficient information for understanding participant activities and interests.

5.2 Case 2

P: poster, E: Exhibitor, VC: Visitor C, VD: Visitor D

Figure 4. Case 2

Figure 4 shows a top view taken by a fixed camera during a poster presentation, where the participants are different from those in Case 1. This case is similar to Case 1, since a third participant (VD) joins the already established conversation between two participants (E and VC). Although the video of Case 2 shows that VD came close to E and VC, VD didn't produce any utterance. After about 27 seconds passed, VD left the area. In this scene, VC is more active than VD, and their activities seemed quite different. We could interpret VC as ADR and VD as SPT by observing this three-party conversation, especially with E's body direction oriented toward VC and VC's body being directed to E. They appear to hold their postures to as a way of preventing the newcomer from cutting off their previous talk. The roles of VC and VD are quite different for E.

We suggest that it is necessary to focus on audience diversity for the speaker by labelling each participant with a role of participation.

5.3 Discussions and future works

Up to this point, we have presented an overview of how the dynamic transitions in human-to-human conversational interaction can be described as instantaneous changes. The results of our case studies clearly show that participation structure has a dynamic movement. This movement was shown by the transition we observed among the roles of the audience; addressee, side participant and bystander. The audience roles and their transitions might be used as a measure for capturing the interest of the audience, which should have a wide range of applications to interface technologies.

We are now working on developing a method of identifying participation structure automatically by using ubiquitous sensor technology that recognizes human non-verbal information data like the body posture and view of each participant.

6 Conclusions

We proposed an application of ubiquitous sensor technology to analyzing the dynamics of multi-party human-to-human conversational interactions in open situations. We presented our analysis results of conversational interactions carried out in an open interaction space of poster presentation sessions. We argued for this method's ability to explain the variety of possible transitions of participation structure by utilizing not only duration of staying at an exhibit but also by dynamic shifts in the roles of participation. The implications of our study, although still preliminary and restricted to a small range of interaction types, are promising for future expansion. We believe the method itself has huge potential for elucidating the use of non-verbal cues in human-to-human interactions, particularly when combined with various automatic signal processing technologies.

Acknowledgement

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

References

Clark, H. H. and Carlson, T. B. 1982 Hearers and speech acts. *Language,* 58: 332-373.

Clark, H. H. 1996 *Using language.* Cambridge University Press.

Goffman, E. 1981 *Forms of talks.* University of Pennsylvania Press.

Goodwin, C. 1981 *Conversational organization: Interaction between speakers and hearers.* New York: Academic Press.

Goodwin, C. 1986 Audience diversity, participation and interaction. *Text,* 6(3): 283-316.

Hagita, N., Kogure, K., Mase, K. and Sumi, Y. 2003 Collaborative Capturing of Experiences with Ubiquitous Sensors and Communication Robots. *2003 IEEE International Conference on Robotics and Automation (IEEE ICRA 2003).*

Kendon, A. 1990 *Conducting interaction.* Cambridge University Press.