

Analyzing Structure of Multiparty Interaction: Group Size Effect in Story-Retelling Task

Noriko Suzuki¹, Mamiko Sakata², and Noriko Ito²

¹ Tezukayama University, Japan

² Doshisha University, Japan

7-1-1 Tezukayama, Nara-city, Nara, 631-8501 Japan

nsuzuki@tezukayama-u.ac.jp

Abstract. This paper examines the differences in verbal and nonverbal behaviors between different group sizes, specifically two-party and three-party interactions. An experiment was conducted using a story-retelling task. The interaction data were recorded with a video camera. Speech, gaze, and gesture data in both group sizes were analyzed. The results suggest that participants in three-party interaction change speakers more frequently by turning their gaze to each other than do those in two-party interaction.

Keywords: group size effect, multiparty interaction, verbal and non-verbal behavior, story-retelling task.

1 Introduction

In our daily life, we have many chances to collaborate face-to-face with each other in different group sizes, from two individuals to parties of ten or more individuals. The participants coordinate with each other through their behaviors during interaction. How does the difference in group size affect their verbal and nonverbal behaviors?

Conventional studies on conversational analysis have mostly focused on two-party face-to-face interaction for developing human-computer interaction system. Previous research on gesture has also analyzed conversations of a single speaker [1, 2] or two speakers [3]. However, our daily activities involve not only two participants but also larger groups. The latter type of activity is called multiparty interaction.

This paper investigates the nature of multiparty interaction in comparison with that of two-party interaction by using a quantitative approach. Few studies have compared the nature of interaction between two-party and multiparty interaction, although the work of Ishizaki [4] and Anderson [5] have indicated a difference in verbal behavior. On the other hand, there has been little focus on the effect of group size on nonverbal behaviors.

In the present task, speakers jointly retell to a listener the story from a short movie they just finished watching. We call this “the story-retelling task,” and it’s a kind of narrative interaction task. In this task, we prepared a short movie

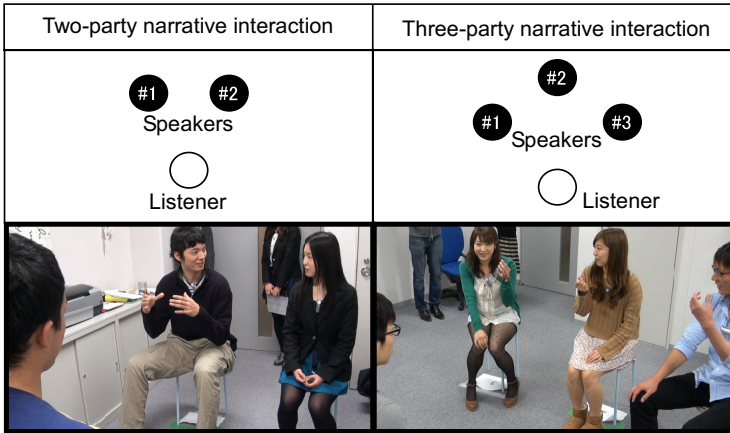


Fig. 1. Experimental settings in story-retelling task: two-party (left) and three-party interaction (right)

entitled “TSUMIKI NO IE (La maison en petits cubes)” [6] as the stimulus for speakers. This movie consists of dialog-free scenes. Neither characters nor objects in this short movie move very much.

In this experiment, we prepared two group sizes for speakers: two-party and three-party groups (Fig 1). Verbal and nonverbal behaviors of the speakers were video-recorded and analyzed. This paper presents the differences in verbal and nonverbal behaviors between the two group sizes. To this end, the contributions of speech and gaze were analyzed in the process of carrying out the story-retelling task.

2 Method

2.1 Participants

Twenty-three undergraduate students from 20 to 22 years old took part in the experiment as participants. They were divided into five groups consisting of three participants and four groups consisting of two participants. Table 1 shows the composition of each group in detail, where, in addition to a breakdown by gender, “H” means the participant felt familiar with the other or others when they talked to a teammate before the task. On the other hand, “L” means she did not feel such familiarity, since she did not previously talk to the other participant(s), or in other words, they were just on “nodding” terms with each other; “M” in this column means that the participant felt familiar with one member but not with the other.

2.2 Story-Retelling Task

The task used in this experiment was the story-retelling task. It requires just two steps: (i) two or three people together watch a short movie as stimulus for the

Table 1. Participants and groups

ID	Group size	Gender	Familiarity
G1	3	M	L
		F	L
		F	L
G2	3	M	L
		F	L
		F	L
G3	3	F	H
		F	H
		M	H
G4	3	F	L
		F	H
		F	M
G5	3	F	L
		F	L
		F	L
G6	2	F	H
		F	H
G7	2	M	L
		F	L
G8	2	F	H
		M	H
G9	2	F	L
		M	L

F: Female, M: Male; H: High, L: Low, M: H+L

story, and (ii) they jointly retell the detailed contents of the movie to a listener who has not watched it.

We chose a listener who was unfamiliar with the people retelling the story. He was asked to remain quiet while listening to the story and just nod without asking any question. Furthermore, he was called to the lab just after the other people watched the movie.

In this experiment, we used a DVD entitled “TSUMIKI NO IE (La maison en petits cubes)” [6]. Neither characters nor objects move very much in this 12-minute movie. Related studies of analyzing human behaviors in narrative interaction employed similar tasks and cartoons as stimulus [1–3].

In such studies, the storylines of cartoons are also explained through the cuts in action and the motions of the characters, without using any speech. From these stimuli, participants must express to each other spontaneous verbal and nonverbal behaviors in the course of their narration. In other words, they must narrate the storyline of the cartoon in their own words and gestures in collaboration with other participants. Accordingly, among the participants the story-retelling task has a convergent rather than a divergent tendency.

2.3 Procedure

First, two experimenters instructed the participants on the aim of the story-retelling task. This aim is to analyze the mechanism of human behavior as they narrate from memory the story of a short cartoon to a listener right after watching it. Second, participants were asked about their familiarity with the other members of their group. Then, the participants watched the movie on a 50-inch

plasma display in front of them. Just after the participants watched the movie, the experimenters called a listener to the lab. The experimenters asked the participants to immediately retell the story of the stimulus from memory to the listener, in as much detail as possible according to the storyline.

During the retelling of the story, we recorded the participants' performances with a video camera. The participants finished their narration with the expression "*Korede owaridesu (That's all)*" when they agreed that they had finished explaining the entire storyline.

After the story-retelling task, speakers answered the Affective Communication Test (ACT) in Japan, which is used for measurement of nonverbal emotional expressiveness [7], as a post-experiment questionnaire.

2.4 Coding

We analyzed the speech, gaze and gesture units of each participant. All participants coded each unit using the annotation software ELAN [8].

Speech: A speech unit in cooperative work is the duration of a single participant's speech bounded by pauses. The speech units were divided into three categories: turn; speech filler, including 'um' and 'ah'; and back-channel responses, including 'yeah' and 'that's right.' However, these units do not include laughing, coughs and breathing sounds.

Gaze: The directions of the participants' gazes are estimated from the direction of the eye's dark region of the pupil and iris. The gaze units were divided into three categories by gaze direction: other group members; the listener; and other directions.

Gesture: The gestures of the participants were labeled from their motion. The gesture units were also divided into three types: representational gestures; self-adaptor gestures; and nodding.

Total time: The overall duration of a narrative interaction, total time, was defined by its start and end times as follows. Start time is when one of the speakers starts to talk, i.e., the beginning of the first speech; end time is when one of the speakers finishes talking, i.e., the end of the last speech.

2.5 Predictions

The participants in the larger group size, the three-party interaction, are predicted to communicate more actively with one another, as a result of their increased behaviors of mutual concern, than those in a smaller group. Furthermore, they are expected to express the following verbal and nonverbal behaviors in the story-retelling task:

- (a) The participants in three-party interaction take turns more frequently than the participants in two-party interaction.
- (b) The participants in three-party interaction look at one another more often than the participants in two-party interaction.

3 Results

3.1 Verbal and Nonverbal Behaviors

Tables 2 and 3 show the results of annotation for verbal and nonverbal behaviors, including the score of the Affective Communication Test (ACT) in Japan [7].

In table 2, *SP ratio* means the ratio of speaking per minute. *NUM of turn* means the number of turns within the interaction duration. *Turn DUR* means speech duration per turn (sec). *BCR* means the duration of back channel response per minute (sec). *FIL* means the duration of speech filler per minute (sec). *SP-LTC* means the speech latency between two utterances per minute (sec). *SP-OVP* means the speech overlap duration between two utterances per minute (sec). *FRQ-SP-OVP* means the number of speech overlaps per minute.

In table 3, *GZ-OTR* means the occurrence ratio of gaze to other group members per interaction duration. *GZ-LSNR* means the occurrence ratio of gaze to the listener. *MG* means the duration of mutual gaze per minute (sec). *FRQ-MG* means the frequency of mutual gaze per minute. *R-GST* means the occurrence ratio of representational gestures per interaction duration. *S-ADP* means the occurrence ratio of self-adaptor gestures per interaction duration. *NOD* means the frequency of nodding per minute. *ACT* means the score of the affective communication test (ACT) in Japan of each participant.

3.2 Group-Size Effect

Figures 2 – 4 show the results for the group size effect. Regarding the difference in the number of turns between three-party and two-party interactions, a U test shows a significant difference ($p = 0.016$) (Fig. 2). For the difference in speech duration per turn, the U test also shows a significant difference ($p = 0.047$) (Fig. 3). Again, the U test found a significant difference ($p = 0.016$) in the occurrence ratio of gaze to other group members per interaction duration (Fig. 4). These results suggest that the participants in three-party interaction change more frequently with shorter turns while giving a look to other members than do

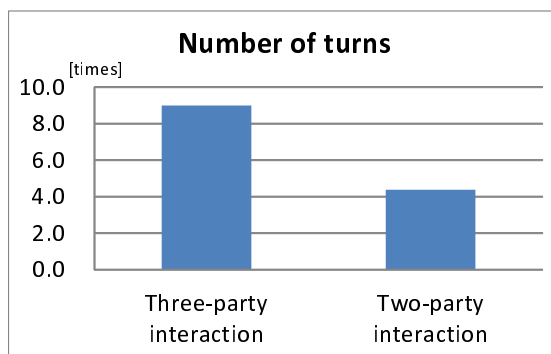


Fig. 2. Results for group size effect (1): Number of turns

Table 2. Results of verbal behaviors

ID	Gender	Seat	Total time	Familiarity	Verbal behaviors							
					SP ratio	Num of turn	Turn DUR	BCR	FIL	SPLTC	SP-OVP	FRQ-SP-OVP
G1	M	C1	185.4	L	0.322	4	14.937	0.000	3.660	4.164	5.115	4.207
G1	F	C2	185.4	L	0.473	11	7.979	0.971	1.294	4.164	9.449	7.443
G1	F	C3	185.4	L	0.277	6	8.558	0.324	0.324	4.164	6.328	6.149
G2	M	C1	181.0	L	0.175	15	2.113	4.310	0.663	0.163	8.275	12.267
G2	F	C2	181.0	L	0.672	11	11.064	2.321	3.647	0.163	8.574	11.604
G2	F	C3	181.0	L	0.289	12	4.358	2.652	0.332	0.163	9.311	11.272
G3	F	C1	187.0	H	0.531	9	11.036	0.642	7.059	11.309	2.695	4.813
G3	F	C2	187.0	H	0.010	5	0.380	1.604	0.000	11.309	0.887	2.246
G3	M	C3	187.0	H	0.431	16	5.035	2.888	4.171	11.309	3.088	4.813
G4	F	C1	265.0	L	0.367	4	24.310	2.717	0.679	0.634	2.780	2.491
G4	F	C2	265.0	H	0.583	5	30.894	0.906	2.038	0.634	4.181	4.755
G4	F	C3	265.0	M	0.034	2	4.509	2.038	0.000	0.634	1.761	3.396
G5	F	C1	230.0	L	0.272	10	6.260	2.089	0.552	3.022	0.357	1.304
G5	F	C2	230.0	L	0.253	19	3.063	0.048	0.004	3.022	0.321	1.565
G5	F	C3	230.0	L	0.223	6	8.558	1.828	1.567	3.022	0.162	0.783
G6	F	C1	234.0	H	0.159	4	9.325	2.500	1.500	2.436	3.088	4.615
G6	F	C2	234.0	H	0.829	4	48.500	1.250	2.500	2.436	3.088	4.615
G7	M	C1	298.0	L	0.571	5	34.029	0.201	2.215	1.100	6.220	4.631
G7	F	C2	298.0	L	0.464	4	34.533	4.228	3.423	1.100	6.220	4.631
G8	F	C1	238.7	H	0.910	2	108.646	1.006	5.530	5.110	3.610	1.257
G8	M	C2	238.7	H	0.075	7	2.558	8.296	0.000	5.110	3.610	1.257
G9	F	C1	129.9	L	0.224	4	7.285	0.000	1.385	6.963	1.649	2.771
G9	M	C2	129.9	L	0.519	5	13.488	0.000	0.924	6.963	1.649	2.771

Table 3. Results of nonverbal behaviors and ACT

ID	Gender	Seat	Total time	Familiarity	nonverbal behaviors							ACT
					GZ-OTR	GZ-LSNR	MG	FRQ-MG	R-GST	S-ADP	NOD	
G1	M	C1	185.4	L	0.036	0.777	0.267	0.647	0.131	0.000	0.000	61
G1	F	C2	185.4	L	0.085	0.169	0.945	1.942	0.152	0.175	6.796	88
G1	F	C3	185.4	L	0.155	0.214	0.678	1.294	0.184	0.310	0.647	54
G2	M	C1	181.0	L	0.228	0.099	7.192	5.305	0.149	0.000	3.315	75
G2	F	C2	181.0	L	0.154	0.092	4.214	3.978	0.358	0.034	4.973	64
G2	F	C3	181.0	L	0.314	0.121	6.636	5.968	0.114	0.044	5.968	56
G3	F	C1	187.0	H	0.299	0.206	7.697	10.909	0.302	0.044	1.925	68
G3	F	C2	187.0	H	0.085	0.146	4.588	6.417	0.014	0.066	4.813	60
G3	M	C3	187.0	H	0.205	0.077	8.576	10.267	0.321	0.127	2.888	79
G4	F	C1	265.0	L	0.093	0.167	2.059	3.170	0.367	0.032	7.925	72
G4	F	C2	265.0	H	0.095	0.498	2.138	3.396	0.418	0.012	6.566	81
G4	F	C3	265.0	M	0.084	0.373	0.805	1.585	0.000	0.016	4.075	66
G5	F	C1	230.0	L	0.526	0.278	0.000	0.000	0.144	0.113	1.826	81
G5	F	C2	230.0	L	0.409	0.387	0.000	0.000	0.050	0.015	4.696	37
G5	F	C3	230.0	L	0.348	0.439	0.000	0.000	0.328	0.012	1.043	57
G6	F	C1	234.0	H	0.203	0.485	3.462	3.846	0.039	0.056	8.205	69
G6	F	C2	234.0	H	0.107	0.154	3.462	3.846	0.415	0.115	2.308	63
G7	M	C1	298.0	L	0.019	0.080	0.394	1.007	0.239	0.245	3.624	75
G7	F	C2	298.0	L	0.125	0.187	0.394	1.007	0.239	0.076	6.644	72
G8	F	C1	238.7	H	0.074	0.251	0.595	1.006	0.449	0.110	2.011	83
G8	M	C2	238.7	H	0.013	0.946	0.595	1.006	0.005	0.009	7.541	55
G9	F	C1	129.9	L	0.048	0.350	1.397	1.385	0.238	0.000	2.771	52
G9	M	C2	129.9	L	0.051	0.408	1.397	1.385	0.178	0.041	1.385	51

the participants in two-party interaction. Therefore, our predictions were partly supported.

3.3 Familiarity vs. Mutual Gaze

Before the story-retelling task, we investigated the familiarity among group members through a pre-experiment questionnaire. We analyzed the relation between familiarity among the participants of a group and verbal/nonverbal behaviors of the participants, except for G_4 . From the results of the U test, we found significant tendencies between familiarity and the duration of mutual gaze ($p = 0.050$) as well as the frequency of mutual gaze ($p = 0.070$) among group members. In this experiment, mutual gaze means that the participants look at one another during narrative interaction, with no eye contact between any participant and the listener. Figure 5 shows the relation between familiarity and the duration

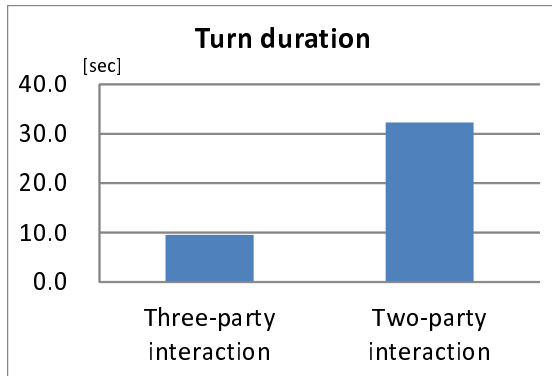


Fig. 3. Results for group size effect (2): Turn duration

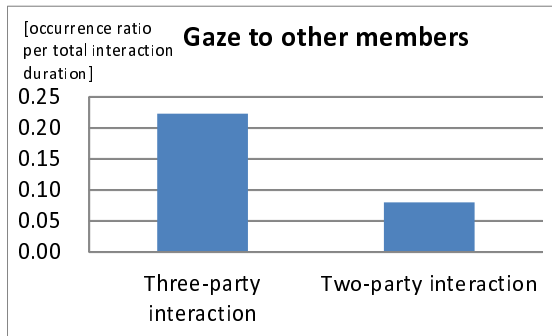


Fig. 4. Results of group size effect (3): Gaze to other group members

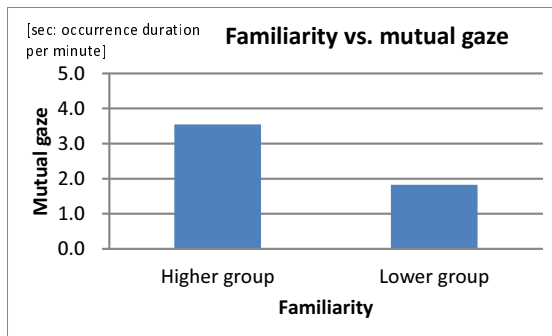


Fig. 5. Familiarity vs. mutual gaze

of mutual gaze per minute (sec), and Fig. 6 shows the relation between familiarity and the frequency of mutual gaze per minute. These results suggest that the participants with higher familiarity looked at each other much more than did the participants with lower familiarity. In other words, the participants with

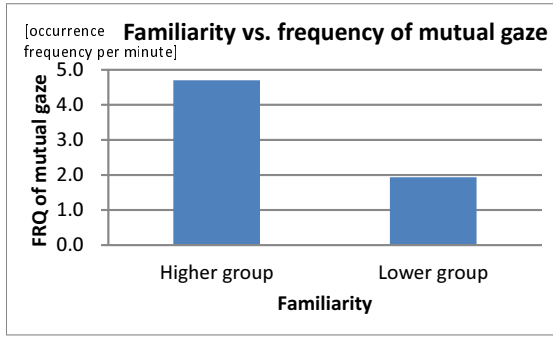


Fig. 6. Familiarity vs. frequency of mutual gaze

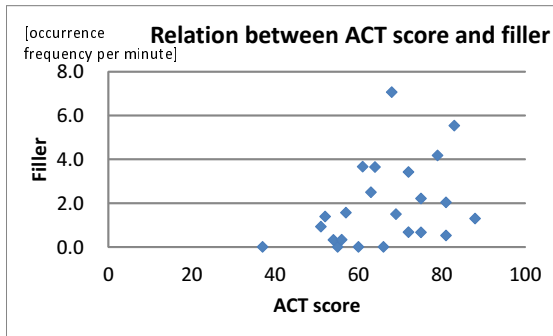


Fig. 7. Relation between ACT score and speech filler

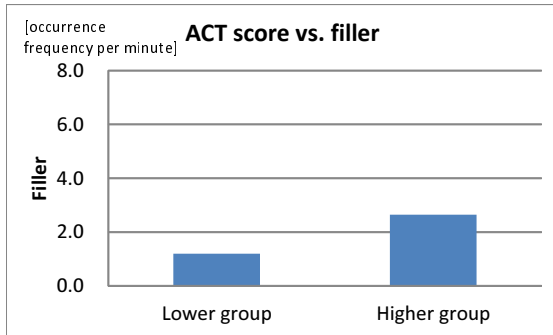


Fig. 8. ACT score vs. frequency of speech filler by score group

lower familiarity did not make nearly as much eye contact as the participants with higher familiarity.

3.4 ACT Score vs. Speech Filler

After the story-retelling task, we investigated the results of the Affective Communication Task (ACT) in Japan [7]. Originally, ACT was developed to explore

nonverbal emotional expressiveness in the U.S.A [9]. It is clearly positively related to acting ability and to receiving nonverbal messages. Simply put, an ACT score is suggestive of a person's nonverbal communication skills. We use the ACT in Japan, which is an adaptation of the U.S.A version for Japanese subjects. Here, the participants used a 9-point scale to respond to 13 statements, such as "when I hear good dance music, I can hardly keep still" and "my laugh is soft and subdued." The ACT score thus highlights the nonverbal communication skills of a person.

We analyzed the relation between the ACT scores of participants and their verbal/nonverbal behaviors. We found positive correlation between the ACT score and speech filler per minute (Fig. 7). We next compared the results of two groups of participants, one scoring higher and the other lower on ACT based on the median score, for the relation between ACT score and the frequency of using speech filler per minute (Fig. 8). These results suggest that the participants with higher skills of nonverbal, emotional expressiveness tend to produce speech filler during interaction because they try to continue talking to a listener.

Unfortunately, there is no significant difference either between the ACT score and group size or between the ACT score and other verbal and nonverbal behaviors, including gaze to other participants or the listener.

4 Conclusions

This paper investigates the nature of multiparty interaction in comparison with those of two-party interaction by using a quantitative approach. Nine groups participated in the story-retelling task. Five groups consisted of three persons and the other four groups consisted of two persons. We video-recorded and analyzed their verbal and nonverbal behaviors during narrative interaction. We also investigated their familiarity among group members and the affective communication score (ACT) [7] recorded before and after the task. The results suggest that there might be a group size effect. The participants in three-party narrative interaction might change more frequently with shorter turns while giving a look to other members than do the participants in two-party interaction. In other words, the difference in group size might affect the collaborative strategy in the story-retelling task. The results also suggest that the participants with higher familiarity look at each other much more than the participants with lower familiarity. The familiarity among group members might affect their gaze interaction in the task. From the results of the affective communication test (ACT) in Japan [7], the participants with higher ACT scores might produce speech filler much more than do the participants with lower ACT scores. The higher skills of nonverbal, emotional expressiveness seemed to affect the speech interaction in the task.

As our future work, we will carry out qualitative analysis as well as quantitative analysis. Furthermore, we will conduct the same story-retelling task while using both a greater number of participants and more groups.

Acknowledgments. The authors thank all of the students in the class assigned to analyze multimodal interaction: Hisashi Fujita, Miho Harada, Ami Hattori, Naoya Hirata, Naoko Ieshiro, Mari Igei, Reia Kadono, Takayuki Kadota, Chikara Kando, Yuya Kouzuki, Hirona Kuwata, Rina Matsuura, Rina Mitsukawa, Nozomi Miyazaki, Yoshitaka Miyazaki, Kana Ogawa, Naho Oguri, Misao Shibata, Eikei Shu, Miho Takesue, Ai Tatsutomi, Hinata Uehama, and Kaori Yamagami. This work was partly supported by a fund from the Grant-in-Aid for JSPS Fellows (Grant No. 24-40261).

References

1. McNeill, D.: Hand and mind. The University of Chicago Press, Chicago (1992)
2. Kita, S., Ozyurek, A.: What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language* 48, 16–32 (2003)
3. Okada, S., Bono, M., Takanashi, K., Sumi, Y., Nitta, K.: Context-based conversational hand gesture classification in narrative interaction. In: *Proceedings of ICMI 2013*, pp. 303–310 (2013)
4. Ishizaki, M., Kato, T.: Exploring the characteristics of multi-party dialogues. In: *Proceedings of the 17th International Conference on Computational Linguistics*, pp. 583–589 (1998)
5. Anderson, A., Mullin, J., Katsavras, E., McEwan, R., Grattan, E., Brundell, P., O’Malley, C.: Multi-mediating multi-party interactions. In: *Proceedings of INTER-ACT 1999*, pp. 313–330 (1999)
6. Kato, K.: *Tsumiki no ie (La maison en petits cubes)*. Robot Communications Inc. (2008)
7. Daibo, I.: Measurement of nonverbal expressiveness: construction of the affective communication test (act) in japan. *Bulletin of Faculty of Literature in Hokusei Gakuen University* 28, 1–12 (1990) (in Japanese)
8. ELAN4.5.2 (2013), <http://tla.mpi.nl/tools/tla-tools/elan/>
9. Friedman, H., Prince, L., Riggio, R., DiMatteo, M.: Understanding and assessing nonverbal expressiveness: the affective communication test. *Journal of Personality and Social Psychology* 39, 333–351 (1983)