

Head motion synchronization in the process of consensus building

Yuki Inoue, Eisuke Ono, Jinhwan Kwon, Masanari Motohashi,
Daisuke Ikari, Ken-ichiro Ogawa and Yoshihiro Miyake

Abstract— Human communication contains not only explicit factors like the meaning of utterance but also implicit factors like body motion. We presume that the synchronization of body motions has a long time with an increased consensus degree during consensus building, and conducted a conversation task to verify this presumption. The present study focuses on the positive correlation of head motions as an indicator of body motion synchronization and questionnaire evaluation as an indicator of consensus building to analyze the relationship between head motion synchronization and consensus degree. Our experimental results showed that two participants' head motions synchronized with each other and the synchronization time of the head motions has a long time during the period of high consensus evaluation. The results suggest that the synchronization period has a long time as consensus degree increases in the process of consensus building.

I. INTRODUCTION

HUMAN communication contains not only explicit factors like the meaning of utterance but also implicit factors like body motion. We have thus investigated the dual relationship between the explicit factors and the implicit factors in human communication [1]. In a previous study, we took up conversation during consensus building as a typical example of co-emergence communication. We analyzed the relationship between response time to partner's utterance as an implicit factor and consensus degree as an explicit factor [2]. As a result, we observed that synchronization of response time became high along high consensus evaluation. However, previous studies had not still investigated synchronization of body motions in consensus building. Some previous researches suggested that synchronization of body motions had positive effects in face-to-face communication [3] [4]. These studies suggest that body motion synchronization has an important role in conversation.

Accordingly, we presumed that the synchronization time of body motions has a long time with an increased consensus degree in the process of consensus building. The present study aims to verify this assumption. We set a conversation task in

consensus building to analyze consensus degree. In particular, the present study focused on head motion including nodding as an indicator of analysis of body motion. The reason is that some previous studies suggest that head nodding had various functions: head nodding manages interactional processes in face-to-face contact events, and simultaneous nod sequences seem to give persons positive affects [5] [6].

II. METHODS

A. Task of Consensus Building

The present study did a conversation task for analysis of the process of consensus building. We referenced a previous study for the design of the task [2]. The content of the task was "guess of house rent of a rental apartment". We provided two participants with common two materials (Material A and B). Material A was a main document for guess of house-rent. Material B was a support document for guess of house-rent. Material A described information about a target apartment, that is, the number of the floor, the area, the arrangement of the room, etc. However, the house-rent wasn't included in it. Material B described another apartment having the conditions near to the apartment of material A. Material B included the house-rent. Accordingly, the participants could guess house-rent of the target apartment. The participants discussed and decided one price together.

B. Participants

Total 10 students participated in our experiment (4 males (age: 22-24) and 6 females (age: 21-22)). With reference to the previous study, we made pairs of participants as follows for promotion of active conversation [2]. All the participants were native Japanese. The participants of each pair were the same sex, and knew well each other and could talk to each other naturally.

C. Experimental Environment

The experimental environment was as follows. (1) The conversation was conducted in a conference room. During the conversation task, there were only two participants in the room. (2) The conditions of brightness, noise, temperature and humidity were adequate for the participants. (3) Each participant took a seat at a distance of 1.5 [m] against a partner. (4) One video camera (Xacti, SANYO Corp.) was located at a distance of 3 m vertically from participants. (5) The materials were fixed on a bookstand on the table. (6) A voice recorder

Manuscript received August 30, 2013.

Y. Inoue, E. Ono, J. Kwon, M. Motohashi, D. Ikari, K. Ogawa, Y. Miyake are with the Department of Computational Intelligence and Systems Science, Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology, 4259, Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa, Japan. (e-mail: inoue@myk.dis.titech.ac.jp, ono@myk.dis.titech.ac.jp, motohashi@myk.dis.titech.ac.jp, ogawa@dis.titech.ac.jp and miyake@dis.titech.ac.jp) .

was put on the table for evaluation of consensus degree. (7) An accelerometer (Wireless Technologies Inc.) was equipped with participant's forehead with a rubber band to measure the head motion. Fig. 1 shows an actual experimental environment.

D. Process of Experiment

Before the conversation task, an experimenter explained the detail of the task. The experimenter confirmed their consents for the experiment, and the experimenter confirmed their understanding on the task. Just after the participants started the conversation, their voices are recorded. After the participants decided one price and ended their conversation task, they called the experimenter.

E. Evaluation of Consensus Degree

After the conversation task, some evaluators evaluated the consensus degree in conversation. They filled in questionnaires by listening the recorded voice in the midst of their conversation. Evaluations were done every 1 [min]. The experimenter stopped the voice recorder and instructed evaluators to fill in the questionnaires. Recorded voice was played one time. If an evaluator would like to hear it again, the voice recorder was replayed. Note here that we did not use the image recorded by the video camera for the evaluation. Consensus degree was evaluated from 1 (very low) to 5 (very high). Fig. 2 illustrates a questionnaire sheet. The evaluators were instructed to mark up intuitively on the sheet. In the present study, three evaluators evaluated consensus degree. We calculated the average of three evaluation results and used this as the consensus degree of the conversation.



Fig. 1 Experimental Environment

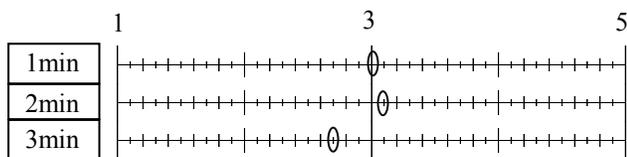


Fig.2 Questionnaire of Consensus Degree: Evaluators evaluated consensus degree every 1 [min] from 1 (consensus degree was very low) to 5 (consensus degree was very high) and marked appropriate points.

F. Calculation of Head Motion Indicator

As shown in Fig. 3, each participant was equipped with the accelerometer. The sampling frequency was 100 [Hz]. The present study conducted a frequency analysis to characterize the head nodding. The procedure is as follows.

1) *Norm of head acceleration*: At first, we calculated the synthesis norm of head acceleration in the vertical direction (x) and front-back direction (z) of the accelerometer [7] as

$$f(t) = \sqrt{x^2(t) + z^2(t)}. \quad (1)$$

Here, the time resolution of $f(t)$ was 0.01 [sec].

2) *Short-term Fourier transform*: We used the short-term Fourier transform (STFT) to analyze the time series data of the acceleration norm. The Hamming window function ($\omega(t)$) was used, of which the range was set to 128 points (1.28 [sec]) as

$$F(\xi, t) = \int_{t-0.64}^{t+0.64} f(t')\omega(t') \exp(-2\pi i t' \xi) dt' \quad (2)$$

where ξ represents frequency and t represents a central time of window function. We shifted the window function and computed STFT every 0.1 [sec] to create time series data.

3) *Definition of natural frequency band of head nodding*: The present study calculated the natural frequency band of head nodding to characterize nodding before we analyze entire data of conversation. We collected 10 actual waveforms of head nodding and computed STFT to obtain the natural frequency of head nodding. Additionally, we calculated the average of the amplitude spectrum in each frequency to investigate the high intensity area. As a result, the maximum point of the amplitude spectrum was around 3 [Hz], and the amplitude spectrums within the range of about 1.5 [Hz] to 4.0 [Hz] were over the half value of the maximum amplitude spectrum. We removed the frequency band below 2 [Hz] because the amplitude spectrum of neighboring 1.5 [Hz] might include much information of other body motion (ex. posture change) in conversation. Accordingly, the present study defined the natural frequency band of head nodding as one within a range of 2 [Hz] to 4 [Hz].

4) *Sum of amplitude spectrum*: Based on the natural frequency calculated in section II.F.3), we extracted the data which was calculated by STFT in section II.F.2) within the range of 2 [Hz] to 4 [Hz] and calculated the sum of the amplitude spectrum every 0.1 [sec] as

$$Sum(t) = \int_2^4 |F(\xi, t)| d\xi. \quad (3)$$

According to the Parseval theorem, the accumulation value of the amplitude spectrum in this frequency band means to the sum of the amplitude of head nodding. Therefore, the head nodding can be characterized by the accumulation value of the amplitude spectrum. Accordingly, we defined the accumulation value as head motion indicator.

Fig. 4 is a typical example of this analysis. Fig. 4 (a) represents the norm of head acceleration in the vertical and front-back direction of the accelerometer. Here X-axis represents time and Y-axis represents calculated norm of acceleration of vertical and front-back direction of

accelerometer. Fig. 4 (b) represents the data calculated by STFT. Here X-axis represents time and Y-axis represents frequency. Light and shade represent the intensity of the amplitude spectrum. The color becomes white as the intensity becomes higher. Fig. 4 (c) represents the head motion indicator calculated. Here X-axis represents time and Y-axis represents head motion indicator. These data are parallel in the time-series. In Fig. 4 (a), the waveform around 749 [sec] and 753 [sec] denotes head nodding, that is, the values of the head motion indicator are high around these points.

G. Statistical Analysis of Head Motion Synchronization

We analyze the synchronization between two participants' head motions using correlation analysis [2]. Since the head motion indicator was not normal distributions, we used the Spearman rank-order correlation coefficient.

At first, we calculated the average duration time of head nodding to utilize as the range of window in correlation analysis. Head nodding is relatively strong motion compared with other head motions. Accordingly, when the value of the head motion indicator was over 90 percentile of the population of participant's head motion indicator, we set the time as head nodding happened time. After computed the duration time of head nodding in all participants, we calculated the average of the duration time. We set this average duration time as the range of correlation analysis window.

Based on the average duration time calculated, we conducted a correlation analysis every 0.1 [sec]. When the correlation was significantly positive and the medians of each participant's data were over 90 percentile of each population, we regarded this state is synchronization of head motions.

Fig. 5 is a typical example of this analysis. Fig. 5 (a) represents the norm of head acceleration of two participants. Here X-axis represents time and Y-axis represents the acceleration norm calculated in the vertical and front-back direction of the accelerometer. Fig. 5 (b) represents the two participants' head motion indicators calculated. Here X-axis represents time and Y-axis represents head motion indicator. Fig. 5 (c) represents the detection result of the synchronization between two participants' head motions. Here X-axis represents times and Y-axis represents synchronization. The value "1" of Y-axis means that head motion synchronization was detected. In Fig. 5 (a), the waveform around 749 [sec] represents the simultaneous occurrence of head nodding.

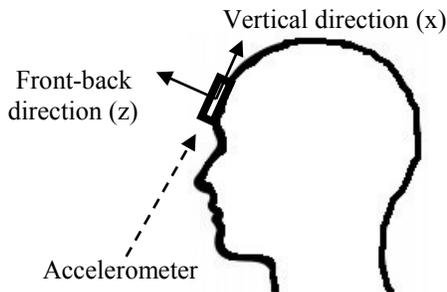
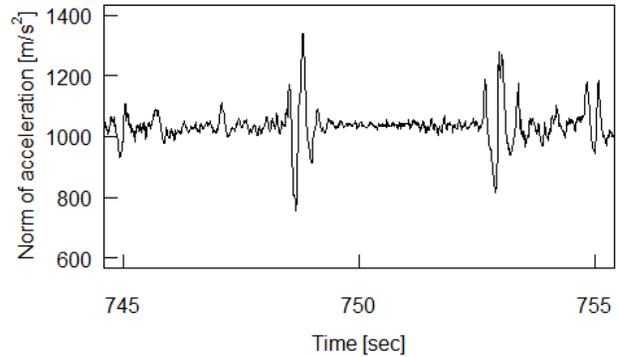
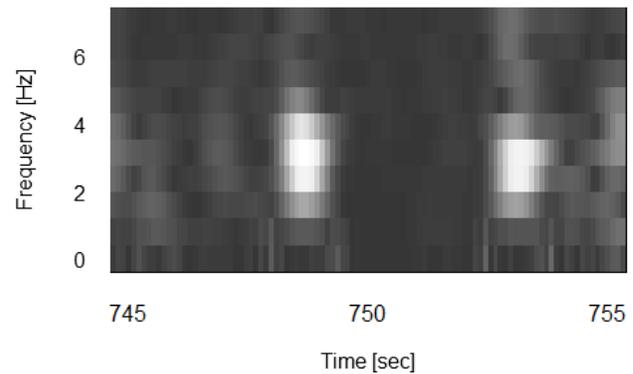


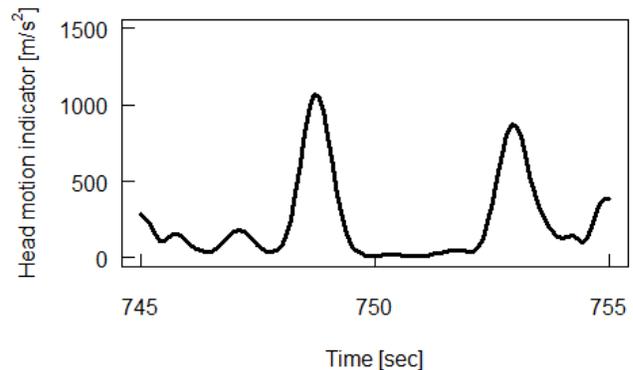
Fig.3 Equipment of an Accelerometer: The accelerometer was equipped on the participant's forehead. Sampling time of the accelerometer was 100 [Hz].



(a) Norm of Head Acceleration: X-axis represents time. Y-axis represents an acceleration norm in the vertical and front-back direction of the accelerometer.

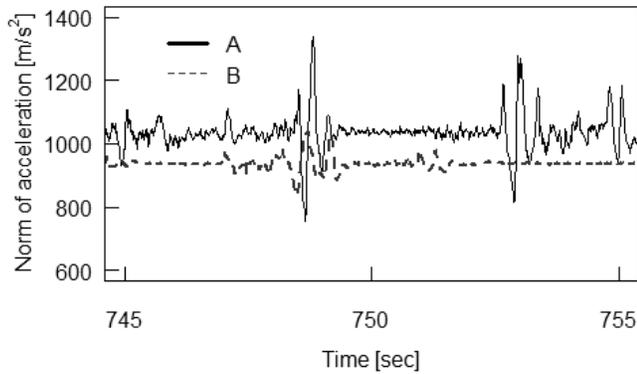


(b) Result of the Short-time Fourier Transform: X-axis represents time. Y-axis represents frequency. Light and shade represent the intensity of the amplitude spectrum.

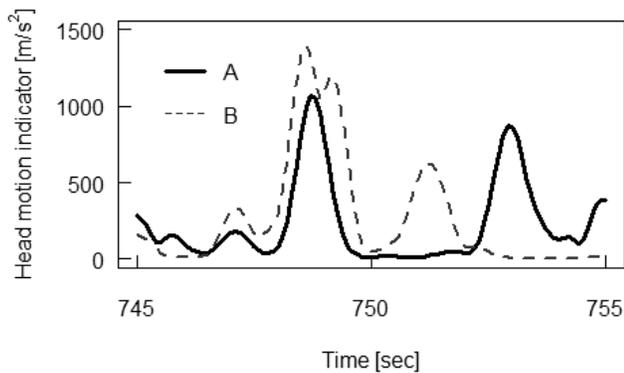


(c) Calculated Head Motion Indicator: X-axis represents time. Y-axis represents the head motion indicator calculated.

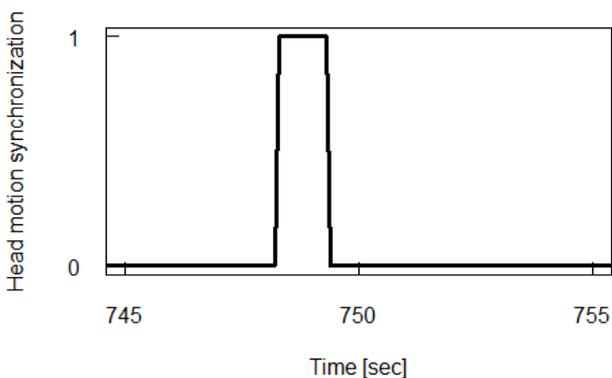
Fig. 4 Typical Example of Head Motion Calculation: (a), (b) and (c) are parallel in the time-series.



(a) Norm of Head Acceleration of Two Participants: X-axis represents time. Y-axis represents calculated acceleration norm in the vertical and front-back direction of the accelerometer.



(b) Calculated Head Motion Indicator of Two Participants: X-axis represents time. Y-axis represents the head motion indicator calculated.



(c) Judgment Result of Synchronization: X-axis represents times. Y-axis represents the judgment results of synchronization. The value “1” of Y-axis means that head motion synchronization was detected.

Fig. 5 Analysis of Synchronization of Two Participants' Head Motion: (a), (b) and (c) are parallel in the time-series.

A. Analysis of Consensus Degree

Fig. 6 is a typical example of the consensus degree evaluated. Here X-axis represents time [sec] and Y-axis represents the average of evaluation values. In this example, consensus degree started from around 3 and gradually increased as the advance of the conversation. We extracted opening 5 minutes and last 5 minutes from the conversation time and compared them. Then we removed first 1 minute because the participants might not define the criteria of evaluation yet. Table 1 shows this result. The average consensus degree in the last phase was larger than average consensus degree in the opening phase in all conversations. There was significant difference by the Wilcoxon signed-rank test ($p < .01$).

B. Analysis of Head Motion Synchrony

Fig. 7 is a typical example of the head motion indicator calculated. Here X-axis represents time [sec] and Y-axis represents head motion indicator. We find some coincidence areas in the waveform of the head motion indicator of the participants A and B. Fig.8 is a typical example of the head motion synchronization computed. Here X-axis represents times and Y-axis represents the judgment results of synchronization. The value “1” of Y-axis means that head motion synchronization was detected. In this figure, some synchronization periods of head motion are detected. In addition, we find that the synchronization period of head motion increases. These examples are data of run number 5 and parallel in the time-series.

As the result of section III.A, the present study suggested that consensus degree in the last phase was larger than consensus degree in the opening phase in conversation. Accordingly, we extracted opening 5 minutes and last 5 minutes from the conversation time and compared them similarly to section III.A.

The average duration time of head nodding was approximately 0.8 [sec] in cases of 10 participants. Accordingly, we set that the range of window in correlation analysis was 0.8 [sec]. After we analyzed the synchronization of head motion indicator along method of II.G, we calculated the synchronization time of head motion in the opening phase and last phase. Table 2 shows this result. The synchronization time in the last phase was longer than the synchronization time in the opening phase in cases of all conversations. Additionally, we analyzed the difference of the synchrony time between the opening phase and the last phase by the Wilcoxon signed-rank test. The result has a significant difference ($p < .01$). From these results, the present study suggested that the synchronization of the head motion indicator was strong when the evaluated consensus degree was high.

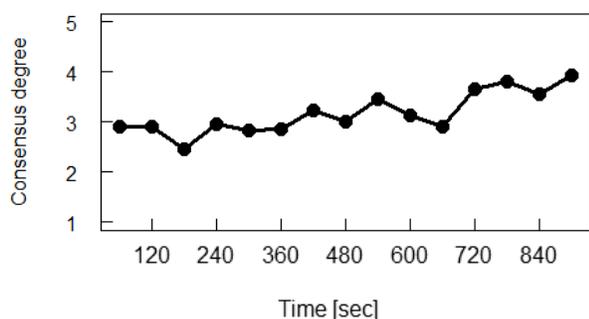


Fig.6 Time Series Data of Consensus Degree: X-axis represents times. Y-axis represents the evaluated value of consensus degree. This example is data of run number 5

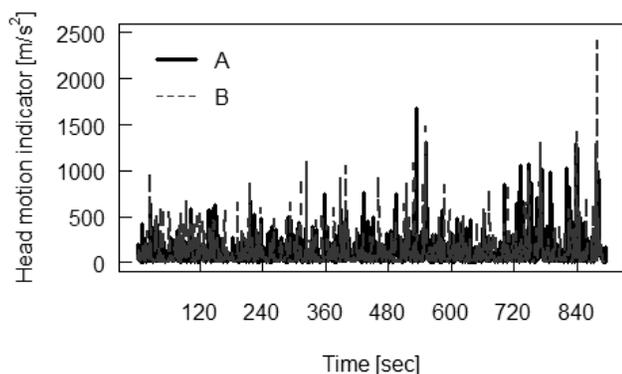


Fig.7 Time Series Data of Head Motion Indicator: X-axis represents times. Y-axis represents head motion indicator. This example is the data of run number 5

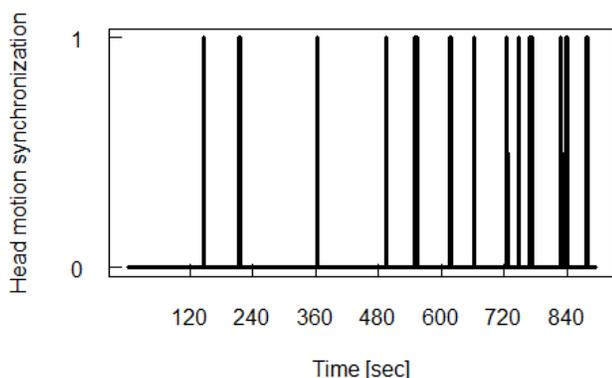


Fig.8 Time Series Data of Head Motion Synchronization: X-axis represents times. Y-axis represents the judgment results of synchronization. The value “1” of Y-axis means that head motion synchronization was detected. This example is the data of run number 5.

IV. DISCUSSION

The present study analyzed the head motion synchronization in the process of consensus building. As a result, the last 5 minutes got higher consensus degree than the first 5 minutes. Based on this result, we analyzed the head motion synchronization of the participants. As the result, the

Table 1 Significant Difference Test Results of Consensus Degree (**: $p < .01$)

Run number	Opening Phase	Last Phase	p value
1	2.6	3.7	
2	3.0	3.9	
3	2.7	3.4	
4	3.0	3.4	
5	2.8	3.6	
Mean	2.8	3.6	**

Table 2 Synchronization Time of Head Motions: The time unit is second. (**: $p < .01$)

Run number	Opening Phase	Last Phase	p value
1	2.7	5.6	
2	5.0	6.3	
3	4.3	9.8	
4	4.6	7.6	
5	2.5	13.1	
Mean	3.8	8.5	**

synchronization time of head motion in the last phase was longer than that in the opening phase on the all experiments. Accordingly, we suggested that the trend of head motion synchronization became strong as consensus degree increased during the conversation task.

A previous study reported that the synchronization of response time became strong as increased consensus degree [2]. Another previous study, which focused on body motion, reported that a counselor obtained high evaluation when strong body motion synchronization between the counselor and a client was observed [3]. In terms of conversation analysis, a previous study suggested that simultaneous nod sequences seem to be associated with positive affects [6]. Accordingly, the results of the present study are supported by these previous studies.

We assumed that the trend of body motion synchronization got strong as consensus degree increased in the process of consensus building. We argue that this assumption is supported by our experimental results and statistical analysis. Accordingly, the present study suggested the presence of the dual relationship between an implicit factor (consensus degree) and an explicit factor (head motion synchrony) in face-to-face conversation.

However, the present study did not clarify whether head motion synchronization was emerged as an embodiment of high consensus degree or head motion synchronization promoted consensus building. As one way to address this problem, we consider the implementation of an interventional experiment. We did not intervene on conversation in the present study. Hence, we have an experimental plan that one

participant never nod or never gain consensus with a partner in conversation. We consider that these experiments are effective for investigation of the dual relationship between head motion synchronization and consensus degree. Also, we need to perform more detailed time-series. The states of human conversation are changed shorter than 5 minutes. Accordingly, analysis in shorter time interval is needed in the future.

REFERENCES

- [1] Y. Miyake et al., "Man-Machine Interaction as Co-emergence Communication," *Trans. Of the Society of Instrument and Control Engineers*, vol. E-2, no.1, pp. 195-206 (2002).
- [2] M. Yoshida, Y. Miyake, N. Furuyama, "Temporal Development of Pragmatics and Dynamics for Consensus Building," *SICE journal (in Japanese)*, vol.47, no.8, pp.337-345, 2011.
- [3] C. Nagaoka et al., "Body Movement Synchrony in Psychotherapeutic Counseling: A Study Using the Video-Based Quantification Method," *IEICE TRANS. & SYST.*, vol. E91-D, no. 6 (2008).
- [4] M. Komori et al., "The Relationship between Classroom Seating Locations and Instructor-Student Entrainment: A Video Analysis Study," *International Conference on Biometric and Kansei Engineering* (2011).
- [5] S. K. Maynard, "Interactional Functions of a Nonverbal Sign: Head Movement in Japanese Dyadic Casual Conversation," *Journal of Pragmatics*, vol. 11, pp 586-606 (1987).
- [6] Kita, S., et al., "Nodding, aizuchi, and final particles in Japanese conversation: How conversation reflects the ideology of communication and social relationships," *Journal of Pragmatics*, 39, pp 1242-1254 (2007).
- [7] H. Saiga et al., "Function analysis of nodding for conversation adjustment in multi-party conversation," *IPSJ SIG Technical Report (in Japanese)*, vol.2010-UBI-26 no.1 (2010).