Embodied Synchrony in the Process of Consensus Building

○Robin Miao Sin Yap (Tokyo Institute of Technology),Yuki Inoue (Tokyo Institute of Technology), Shunsuke Miura (Tokyo Institute of Technology), Jinhwan Kwon (Tokyo Institute of Technology), Ken-ichiro Ogawa (Tokyo Institute of Technology), Yoshihiro Miyake (Tokyo Institute of Technology)

Embodied synchrony has been studied in dyadic interaction between adult-neonate, client-counseller and physician-patient in many past researches. However, the mechanism in which embodied synchrony occurs in human communication is still not well understood due to complex human dynamics. In this study, we investigated the relationship between the degree of head motion synchrony and empathy in dyadic interaction using consensus-building task to reveal the degree of dependence on differences between high and low context cultures. A total of 29 pairs of subjects participated in this research. The subjects can be broadly classified into two main groups, namely, native English speakers (Male 4 pairs, Female 4 pairs) and native Japanese speakers (Male 11 pairs, Female 10 pairs). The result showed that a positive correlation exists between the degree of head motion synchrony and empathy for both native Japanese and English speakers. In addition, head motion synchrony at low, medium and high frequency of 1.0 Hz, 2.5Hz and 5.0 Hz respectively are significant for both native Japanese and English speakers. However, at low head motion frequency of 1.0 Hz, the time lag for native Japanese speakers (0 sec to 0.3 sec) is shorter as compared with native English speakers (0.4 sec to 0.5 sec). At high head motion frequency of 5.0Hz, the time lag for native Japanese speakers is longer (0.4 sec to 0.5 sec) but the time lag for native English speakers varies from 0.1 sec to 0.5 sec. These differences may be attributed to different context cultures of head motion between native English and Japanese speakers.

Keyword: Head motion synchrony, Consensus building, Degree of empathy, High and low context cultures

1. Introduction

Consensus building is a decision-making process that seeks the consent of all participants. It is a typical process of dyadic interaction between people. There are many past researches, which investigated the effects of consensus building from a micro perspective (i.e. organizational level, Hyun 2009) to a macro perspective (i.e. regional level, Matthias and Marcel, 2006). Hyun (2009) studied consensus building through role-playing from the perspective of self-persuasion on simulated civic forum. Matthias and Marcel (2006) studied the effects of consensus building processes on regional collaboration in environmental psychology.

In addition, many past researches reported that body motion and utterance were synchronized during dyadic interaction between people. Condon and Sandar (1974) found that neonate`s motor behaviour is entrained and synchronized with the utterance of adults in the environment. Koss and Rosenthals' (1997) study supported the theoretical proposed link between positivity and interactional synchrony in physician-patient relationship.

However, there are very few research studies on embodied synchrony during consensus building. In this present study, we investigated the dyadic interaction in adult-adult interaction using consensus-building task based on the earlier work by Yoshida et al. (2008). Specifically, we focused on empathy as an indicator for the degree of consensus building, and researched the relationship between the degree of empathy and head motion synchrony in the process of consensus building with native Japanese and English speakers. We compared the results between native Japanese and English speakers to reveal the degree of dependence between high and low context cultures, which may have significant importance in the process of intercultural consensus decision building.

2. Materials and Methods
2.1 Experiment Task

The present study involves dialogue between two participants on a consensus building task using Japanese apartment information taken from the Japanese housing property website"SUUMO". In this task, each participant was given two different apartment information which are closely related to each other, namely Material A and Material B. Materials A and B are two different apartments which contains the following information such as subway line, bus station, monthly rental and room type etc. Material A contains all the essential details of a Japanese apartment except the monthly rental price in Japanese yen. Material B contains another apartment details, which are incomplete. Both participants need to exchange information with each other to complete the missing information on their copy of material B and use it

as a reference to make a guess on the unknown monthly rental price on material A. When the price has been agreed upon, a consensus was reached between both participants.

2.2 Participants

A total of 29 pairs of subjects participated in this present study. The subjects are either native English or Japanese speakers. For native English speakers, there are 4 pairs of male (age between 23 to 42) and 4 pairs of female (age 19 to 39). The native English-speaking participants come from the following countries: 6 from US, 4 from UK, 3 from Canada, 2 from New Zealand and 1 from Australia. For native Japanese subjects, there are 11 pairs of male (age between 17 to 26) and 10 pairs of female (age between 18 to 26). All the subjects of each pair belong to the same gender. All the Japanese subjects knew each other and were able to talk to each other naturally. However, 5 out of 8 pairs of English subjects do not know each other beforehand. In this case, a 45 minutes icebreaker activity was conducted beforehand for these subjects to get to know each other better before the start of the experiment.

2.3 Experiment Environment

Figs. 1 and 2 show the schematic diagram and snapshot of the experiment environment. The experiment was conducted in a conference room. The humidity, temperature and brightness level of the room were adjusted to a comfortable level for the participants. During the experiment, there were only two participants in the room. The two participants were seated at opposite end of the table (1.75 m by 0.90 m). Two video cameras (CMOS Xacti Full HD, SANYO, Japan) were located vertically at a distance of 2.5 m away from the left and right hand side of both participants. Two book stands used to support the dialogue materials were positioned on the table facing each participant. Two laptops with Sennheiser headsets attached were positioned on the table to record the real-time vocal data of each participant. In addition, a voice recorder (ICD-UX533F, SONY, Japan) was also positioned on the table to record the real-time dialogue of the participants for playback. A bell was also positioned on the table for the participant to ring and alert the experimenter at the end of the dialogue. An accelerometer (WAA-006, OMRON, Japan) was attached to the forehead and body of each participant to record the head and body motion data in real-time.
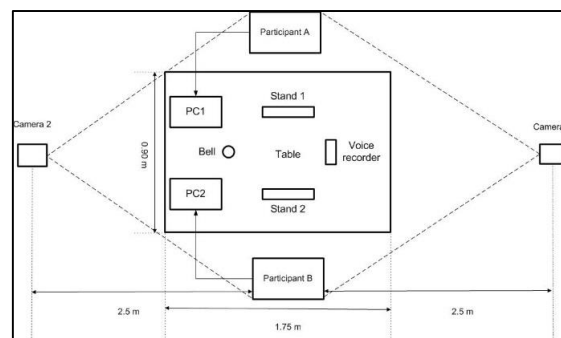

Fig. 1 Schematic diagram of experiment environment.

Fig. 2 Snapshot of experiment environment. Two digital video cameras capture the participants, two laptop with headsets record real-time vocal data, and accelerometers attached to forehead and body of participant record real-time head and body motion data.

2.4 Experiment Procedures

Before the experiment, the experiment procedures and rules were explained to the participants. The participants were informed that there was no time limit for the dialogue. When the participants had confirmed their understanding of the task, the participants were given a practice session before the actual session. During the practice session, the participants were given a set of practice material. Each participants exchange information with each other based on the apartment information given for the practice session and rang the bell when a consensus was reached. No questionnaire was administered to rate the degree of empathy between each participants during this session. After the practice session, the participants were given a short break of about 5 minutes, which was followed by the actual session. During the actual session, the participants were given a different set of materials.

Accelerometers were secured to the forehead and body of each participant using elastic bands and each participant put on their respective headsets. The experimenter will leave the room during the experiment. When a consensus had been reached, one participant rang the bell on the table to alert the experimenter. After the dialogue, the target monthly rental price was revealed to both participants. Finally, the participants were given a questionnaire to rate their degree of empathy with each other by listening to the playback from the voice recorder. As a form of incentive and motivation to the participants, the participants were informed that their monetary allowance will increase if there was close agreement (± 5000 yen) between the agreed price and target price just before the practice session of the experiment. This information was disseminated to all native Japanese speakers. However, 5 out of 8 native English speakers were not informed of the incentives before the experiment as the experiment were conducted on two different occasions.

2.5. Empathy Evaluation

Evaluation of the degree of empathy was done using questionnaire. A sample of the questionnaire was shown in Figure 3. The questionnaire uses a 5-level rating system, starting from 0 (very low empathy) to 4 (very high empathy). The participants were informed that if the content of the dialogue was based on facts, knowledge and logical thinking, the empathy level was low. However, if the content of the dialogue was based on mutual understanding and emotional feelings, the empathy level was high. Each participant listened to the playback from the voice recorder once and put a rating on the questionnaire every 30 seconds. However, the participants were allowed to listen to the playback twice at their own discretion. For confidential reasons, a partition was placed in between each participant, so that each participant was not able to see each other's questionnaire during this process.

Fig. 3 Questionnaire to measure degree of empathy: Evaluation was done every 30 sec from 0 (very low empathy) to 4 (very high empathy).

2.6. Head Motion Synchrony Analysis

The statistical analysis used to determine the synchrony of head motion between both participants is a four-step process: namely (1) calculation of head acceleration norm; (2) time-frequency analysis of head acceleration norm; (3) calculation of head motion indicator; and (4) calculation of time lag between head motion indicator of both participants using correlation analysis. The detailed mathematical modellings were explained in the following sections.

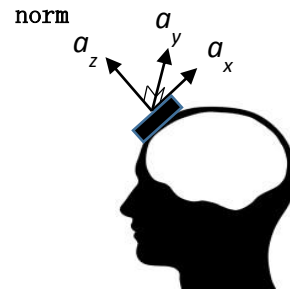### 2.6.1 Calculation of head acceleration norm

Fig. 4 Accelerometer attached to forehead of participant. a_x represents acceleration in vertical direction, a_y in lateral direction, a_z in forward-backward direction.

Fig. 4 shows the schematic diagram of the accelerometer attached to the forehead of the participant. Fig. 5 shows a typical head acceleration norm of one participant. The sampling period is 10 ms and the sampling frequency is 100 Hz. The head acceleration norm $a(t)$ of a participant in the vertical direction, $a_x(t)$, sideway direction, $a_y(t)$ and forward-backward direction, $a_z(t)$ was calculated as follows:

$$a(t) = \sqrt{a_x^2(t) + a_y^2(t) + a_z^2(t)} \qquad (1)$$

where $a_x(t)$ = Acceleration in vertical direction, $a_y(t)$ = Acceleration in lateral direction, $a_z(t)$ = Acceleration in forward-backward direction.

2.6.2 Time-frequency analysis of head acceleration norm

Time-frequency analysis of head acceleration norm was done using short-time Fourier transform (STFT). STFT of the head acceleration norm was calculated as follows:

$$F(\xi, t) = \int_{-\infty}^{\infty} f(t)\omega(t - \tau)\exp(-2\pi j\xi\tau)\, d\tau \quad (2)$$

where ξ = Frequency in Hz, $\omega(t)$ = Hamming window function, t = Central time of window function. The window length is 1.28 sec and frame shift is 0.1 sec. Fig. 6 shows a typical STFT plot of the head acceleration norm.
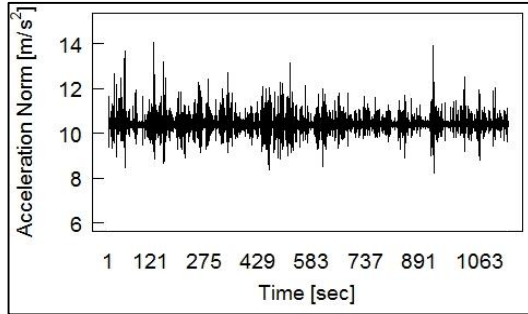

Fig. 5 Head acceleration norm. x-axis represents time in sec and y-axis represents acceleration norm in m/s^2.
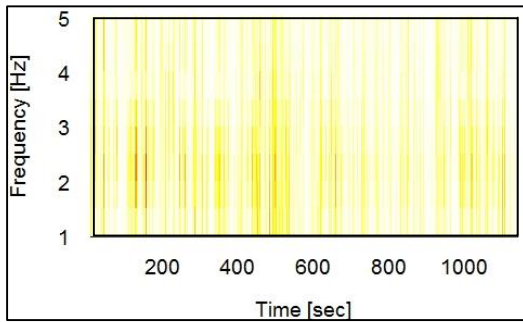

Fig. 6 Typical STFT plot. x-axis represents time in sec and y-axis represents frequency in Hz. Dark and light region represents high and low intensity regions respectively.

2.6.3 Calculation of head motion indicator

From the STFT results of acceleration norm of participant A and B, we extracted the data at 0.5 Hz interval between the frequency bandwidth of 1.0 Hz to 5.0 Hz and calculated the amplitude spectrum at each frequency band every 0.1 sec. Figure 7 shows a typical plot of the head motion indicator between both participants.
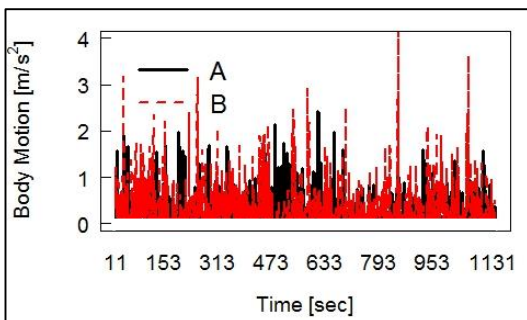

Fig. 7 Head motion indicator of both participants. x-axis represents time in sec and y-axis represents head motion indicator in m/s^2.

2.6.4 Calculation of time lag between head motion indicator of both participants.

Finally, head motion synchrony between both participants was analyzed using Spearman rank-order correlation coefficient. We conducted a correlation analysis of the head motion indicator at 0.5 Hz intervals between the bandwidth of 1.0 Hz to 5.0 Hz, using a window length of 1.8 sec and a frame shift of 0.1 sec to determine the head motion synchrony between both participants in the time lag interval between −0.5 sec and 0.5 sec. We used a threshold of 90 percentile of the population data as the criterion for head motion synchrony. When the peak of two consecutive head motion indicators was above the 90 percentile line with a time lag between −0.5 sec to 0.5 sec, this was counted as head motion synchrony. A positive time lag value indicates that participant A has a faster head acceleration motion as compared with participant B and a negative time lag value indicates otherwise. Figure 8 shows a typical plot of the head motion synchrony between both participants for native English speakers.
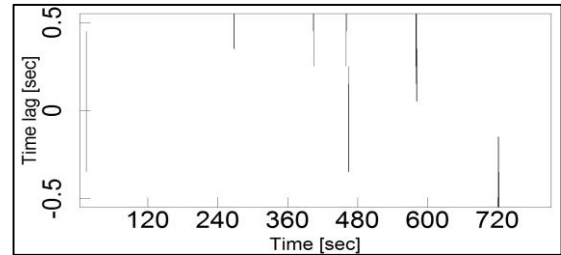

Fig. 8. Typical head motion synchrony plot between both participants. x-axis represents time in sec and y-axis represents time lag in sec.

3. Results
3.1 Empathy Analysis

Figs. 9 and 10 show the comparison of the degree of empathy for first half and second half of the dialogue for native Japanese and English speakers respectively. The data for the first 60 seconds of dialogue were removed as the criteria for consensus building task has not yet been defined. For both native Japanese and English speakers, the mean degree of empathy for the second half of dialogue was higher as compared with the first half. The mean degree of empathy for first and second half of dialogue for native Japanese and English speakers were (1.8, 2.4) and (2.4, 2.7) respectively. There was a significant difference between the first and second half of dialogue for

native Japanese speakers (P < 0.01) and for native English speakers (P < 0.1) using Wilcoxon signed rank test.
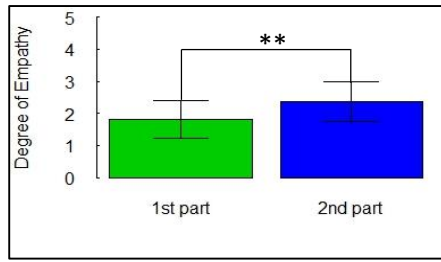

Fig. 9 Significant difference test for native Japanese speakers. Degree of empathy for second half of dialogue is higher as compared with first half (**: P < 0.01).
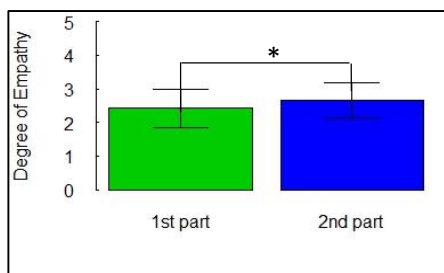

Fig. 10 Significant difference test for native English speakers. Degree of empathy for second half of dialogue is higher as compared with first half (*: P < 0.1).

3.2 Head Motion Synchrony Analysis

We calculated the mean value of head motion synchrony for the first and second half of dialogue when the frequency of head motion synchrony is significant for both native Japanese and English speakers. The mean value of head motion synchrony for the second half of dialogue is higher as compared with the first half for both native English and Japanese speakers when the frequency of head motion synchrony is significant. These results showed that head motion synchrony increases with time in the process of consensus building. In comparison with the results for degree of empathy for native Japanese and English speakers from Figs 9 and 10, the results showed that there is a positive correlation between head motion synchrony and degree of empathy in the process of consensus building for both native English and Japanese speakers. Tables 1 and 2 show the mean value of head motion synchrony between the first and second half of dialogue when the frequency of head motion synchrony is significant for native English and Japanese speakers respectively.

Table 1 Mean value of head motion synchrony for native English speakers.
(**: P < 0.05, *: P < 0.10)

Frequency (Hz)	Time Lag (sec)	Mean Head Motion Synchrony		P
		First	Second	
1.0	0.3	0.05	0.17	**
	0.4	0.10	0.28	**
2.0	0.3	0.24	0.52	*
2.5	0.3	0.16	0.51	*
	0.4	0.16	0.48	*
	0.5	0.17	0.49	*
4.0	0.3	0.20	0.37	**
4.5	0.3	0.18	0.43	*
	0.4	0.23	0.52	*
	0.5	0.28	0.50	*
5.0	0.1	0.13	0.35	*
	0.2	0.13	0.40	**
	0.3	0.15	0.43	**
	0.4	0.22	0.41	*
	0.5	0.25	0.45	*

Table 2 Mean value of head motion synchrony for native Japanese speakers.
(**: P < 0.05, *: P < 0.10)

Frequency (Hz)	Time Lag (sec)	Mean Head Motion Synchrony		P
		First	Second	
1.0	0	0.12	0.22	*
	0.1	0.13	0.24	*
	0.2	0.13	0.22	**
	0.3	0.11	0.19	**
2.5	0.5	0.34	0.50	*
4.5	0.3	0.33	0.51	*
	0.4	0.33	0.51	*
	0.5	0.30	0.46	*
5.0	0.4	0.32	0.50	*
	0.5	0.30	0.47	*

3.3 Comparison of Results

Tables 3 and 4 show the head motion synchrony between 1.0 Hz to 5.0 Hz at 0.5 Hz interval, with a time lag between 0 sec to 0.5 sec at 0.1 sec interval for native English and Japanese speakers respectively. The results showed that head motion synchrony at low, medium and high frequency of 1.0 Hz, 2.5 Hz and 5.0 Hz respectively were significant for both native English and Japanese speakers (Wilcoxon signed rank test, P < 0.05, P < 0.1). However, at low head motion frequency of 1.0 Hz, the time lag for native Japanese speakers are shorter (0 sec to 0.3 sec) (Wilcoxon signed rank test, P < 0.05, P < 0.1) as compared with native English speakers (0.4 sec to 0.5 sec) (Wilcoxon signed rank test, P < 0.05). Furthermore, at high head motion frequency of 5.0 Hz, the time lag for native Japanese speakers is longer (0.4 sec to 0.5 sec) (Wilcoxon signed rank

test, P < 0.1), whereas the time lag for native English speakers varies from 0.1 sec to 0.5 sec (Wilcoxon signed rank test, $P < 0.05$, $P < 0.1$). These indicates that these results are attributed to different context cultures of head motion between native English and Japanese speakers.

Table 3 Significance test of head motion synchrony for native English speakers.
(**: $P < 0.05$, *: $P < 0.10$, n.s.: non-significant)

Time Lag (sec)	Native English Head Motion Frequency (Hz)								
	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
0.0	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
0.1	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	*
0.2	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	**
0.3	n.s.	n.s.	*	*	n.s.	n.s.	**	*	**
0.4	**	n.s.	n.s.	*	n.s.	n.s.	n.s.	*	*
0.5	**	n.s.	n.s.	*	n.s.	n.s.	n.s.	*	*

Table 4 Significance test of head motion synchrony for native Japanese speakers.
(**: $P < 0.05$, *: $P < 0.10$, n.s.: non-significant)

Time Lag (sec)	Native Japanese Head Motion Frequency (Hz)								
	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
0.0	*	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
0.1	*	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
0.2	**	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	*	n.s.
0.3	**	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	*	n.s.
0.4	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	*
0.5	n.s.	n.s.	n.s.	*	n.s.	n.s.	n.s.	*	*

4. Conclusion

The present study investigated the correlation between degree of empathy and head motion in the process of consensus building with native English and Japanese speakers. For both native Japanese and English speakers, we found that the mean head motion synchrony for the second half of dialogue was higher as compared with the first half when the frequency of head motion synchrony is significant. In addition, the mean degree of empathy for the second half of dialogue was higher as compared with the first half with a significant difference for both native Japanese and English speakers. These results showed that a positive correlation exists between degree of empathy and head motion synchrony for both native English and Japanese speakers. These results were further supported by the research by Koss et al. (1997) and Yoshida et al. (2008). However, it should be noted that the significance test results for head motion synchrony of native English speakers was based on the combined data from the incentive group (3 pairs) and non-incentive group (5 pairs). As such, we were unable to determine the effects of incentive on head motion synchrony and degree of empathy as it is beyond the scope of this study. Also, the head motion synchrony results for native English speakers were based on the combined data of participants from different native English-speaking countries. It should be noted that even between native English-speakers, there exist different context culture of head motion. For future work, we will continue to collect more data from native English speakers under the same condition (i.e. with incentives) to reveal the degree of dependence of head motion synchrony and empathy on high and low context cultures between Japanese and English speakers. As an extension of this present framework, we wish to apply the timing results in engineering applications to realize a fluid human-robot communication in the near future.

5. References

[1] Buchecker, M., Hunziker, M, The Effect of Consensus Building Processes on Regional Collaboration, Agricultural Economic Review, Vol. 7 (1), pp. 72-83, 2006.

[2] Condon, W.S., Sander, L.W., Neonate Movement Is Synchronized with Adult Speech: Interactional Participation and Language Acquisition, Science, Vol. 183 (120), pp. 99-101, 1974.

[3] Hall, J.A., Harrigan, J.A., Rosenthal, R, Nonverbal Behaviour in Clinician-Patient Interaction, Applied and Preventive Psychology, Vol. 4, pp. 21-37, 1995.

[4] Hyun, S.P., Consensus Building through Role-Playing from Perspective of Self-persuasion, Journal of Media and Communication Studies, Vol. 1(2), pp. 23-32, 2009.

[5] Koss, T., Rosenthal, R., Interactional Synchrony, Positivity and Patient Satisfaction in the Physician-Patient Relationship, Medical Care, Vol. 35(11), pp. 1158-1163, 1997.

[6] Inoue, Y., Body Motion Synchronization in the Process of Consensus Building (In Japanese), Master Thesis, Tokyo Institute of Technology, 2014.

[7] Yoshida, M., Miyake, Y., Furuyama, N., Temporal Development of Pragmatics and Dynamics in Conversation for Building Consensus, IEEE International Conference on Systems, Man and Cybernetics, pp. 2419-2425, 2008.