Entrainment in Human-Agent Text Communication

Ikuko Kanashiro[1], Kazuki Kobayashi[1], and Yasuhiko Kitamura[2]

[1] Graduate School of Science and Technology, Kwansei Gakuin University
2-1 Gakuen, Sanda, Hyogo, 669-1337 JAPAN
[2] School of Science and Technology, Kwansei Gakuin University,
2-1 Gakuen, Sanda, Hyogo, 669-1337 JAPAN
{i_kanashiro, kby, ykitamura}@ksc.kwansei.ac.jp

Abstract. Non-verbal information such as utterance speed and switching pause create an impression of the speaker. If intelligent agents could handle such non-verbal information properly, the quality of interactions between agents and human users would improve. Entrainment is a phenomenon in which brainwave synchronization is established by means of periodic stimulus. It is reported that non-verbal information expressed by an individual entrains that expressed by another in voice communication. We have interest in how an agent can affect people though entraining the non-verbal information in text communication. Text is much easier for agents to handle than voice. Through experiments, we show that the utterance speed of an agent can entrain the duration of switching pauses used by human subjects.

Keywords: human-agent text communication, switching pause, utterance speed, entrainment

1 Introduction

People communicate by not only verbal information but also the non-verbal information such as the volume, pitch, and speed of the utterances and switching pause which is the silence between turns [7,8]. The non-verbal information, as well as the verbal information, creates an impression of the speaker [13]. For example, many bank clerks speak slowly and clearly. This way of speaking creates an impression of credibility which is an important attribute for bank clerks. On the other hand, TV shopping casters speak rhythmically and rapidly to introduce commercial products. The viewers receive a lot of commercial information in a short time and are tempted to buy the products. This is an effective way for TV casters to sell products.

Nagaoka et al. [1] show that the duration of the switching pauses used by a speaker in voice communication can determine his/her impression. If the duration is short, he/she appears to be unkind, and if it is long, he/she appears to be unreliable. Thus, the non-verbal information is important in creating an impression of the speaker and to make the listener take the action desired by the speaker. An intelligent agent that could handle non-verbal information properly would improve the quality of interactions between the agent and human users [10,11].

This concept also permits non-verbal information to be used to entrain the partner. In voice communication, Watanabe et al. [4] report that breathing frequency

establishes entrainment between two persons when they talk to each other smoothly. Nagaoka[2, 3] showed that switching pauses are entrained in voice communication. Entrainments between an artifact and people have also been observed. Komatsu et al. [5] report that utterance speed is entrained in voice communication between a person and a computer given the situation that the person and the computer interchangeably read sentences written in a script. If we can utilize entrainment properly, we can develop an artificial agent that can create desirable impressions by its non-verbal information [9].

Conventional works deal with entrainment in voice communication, but this paper deals with text communication. Text is much easier to be handled by agents than voice. Conventional text communication systems such as chat systems or messengers fail to well represent non-verbal information. Our solution is a chat system that changes that rate at which it displays individual characters in a message to alter the utterance speed.

This paper shows how an agent can entrain non-verbal information such as utterance speed and switching pauses in text communication with humans. We believe that the result will lead to the development of agents that can create affective impressions by utilizing non-verbal information.

Section 2 introduces a chat system that can control utterance speed. Section 3 shows how an agent can entrain the utterance speed and the switching pause of human users and we discuss the result in Section 4. We conclude this paper in Section 5 with our future work.

2 Chat system that reflects utterance speed

Users receive no impression of utterance speed in conventional chat systems because the systems display the whole message at once when received. We developed a chat system that can control the utterance speed.

The utterance time per character (or just "utterance time") T [msec/char] is defined as

$$T = \frac{1}{V} = \frac{(T_2 - T_1)}{C} \qquad (1)$$

where T_1 is the time when the sender starts typing his/her message, T_2 is the time when he/she finishes typing it, and C is the number of characters in the message. The utterance speed V [char/msec] is defined as the inverse of T.

Fig. 1 shows the components of the chat system that reflects the utterance speed of the message creator. Users A and B send messages to each other through client programs A and B, respectively. When user A inputs a message using client program A, the client sends it with the user's name and the utterance time to client program B through the server program. When client program B receives the message, it types the characters in the message one by one according to utterance time T. User B sends back his/her messages to user A and the message is processed likewise. This system allows each user to feel the utterance speed of his/her counterpart. The server program which mediates the client programs stores the history of chat data between users in a log file.

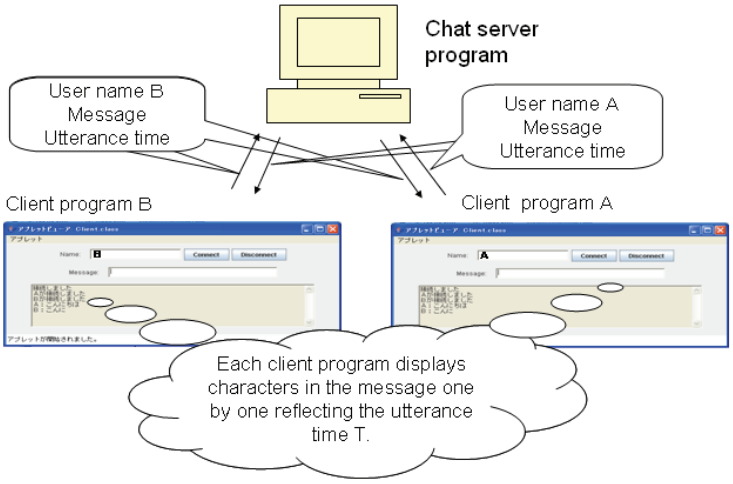


Fig. 1. Components of chat system that reflects utterance speed.

3 Experiment

We here examine how an agent can entrain the non-verbal information in text communication with subjects. The non-verbal information in voice communication has various modalities such as volume, pitch, and speed, but text communication is limited to just speed and switching pause. We examine how these two factors can establish entrainment.

3.1 Experiment system

We replaced a client program of the chat system shown in Fig. 1 by an agent as shown in Fig. 2. The agent automatically sends messages to a subject. It doesn't interpret the content of replies from the subject but merely uses them as cues for creating its next message. When the agent receives a cue, it reads the message (m_i), utterance time (ut_i), and switching pause (sp_i) stored in the setting file. It then applies the switching pause and sends the message with the utterance time specified to the subject.

The agent has no physical appearance like an avatar and all communication is through a chat interface as shown in Fig. 3.

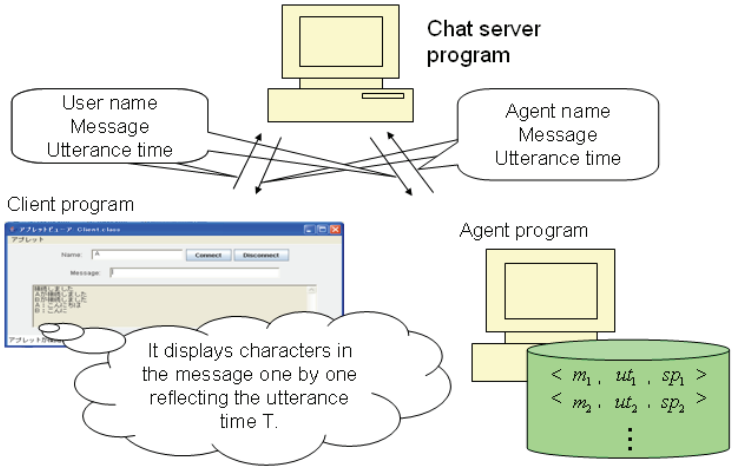


Fig. 2. Experimental system.

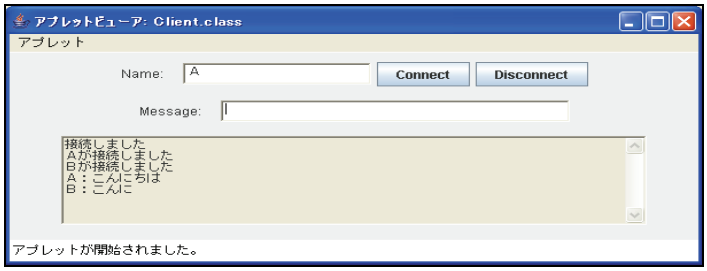


Fig. 3. User interface.

The subject receives a series of questions from the agent. We prepared only easy questions such as asking for the subject's name, age, preferences, and so on, in order to make him/her reply quickly. We didn't use tough questions that would need time to reply because they may hinder smooth communication between the agent and the subject. As shown in Table 1, we prepared 52 questions and stored them with corresponding utterance time and switching pause in a setting file.

Table 1. Example of questions from the agent.

Question	Utterance Speed	Switching Pause
(1) Hello. May I ask you some questions? What is your name?	500.00	0.00
(2) What is your nickname?	462.50	0.00
(3) How old are you?	425.00	0.00
(4) …	…	…

3.2 Procedure

In this experiment, we examined how an agent can entrain non-verbal information expressed by subjects. Non-verbal information in text communication is limited to utterance speed and switching pause, so we examined how these factors establish entrainment. We assessed entrainment by measuring the following correlations.

(1) Correlation between utterance speed of the agent and that of the subjects
(2) Correlation between utterance speed of the agent and switching pause of the subjects
(3) Correlation between switching pause of the agent and that of the subjects
(4) Correlation between switching pause of the agent and utterance speed of the subjects

40 university students participated in this experiment as subjects. We separated them into two groups; variable utterance speed group and variable switching pause group; each group had 20 subjects.

Each subject in the variable utterance speed group chatted with the agent whose utterance time varied from 50[msec/character] to 500[msec/character] in 13 steps of 37.5[msec/character] and with two peaks as shown in Fig. 4. The switching pause was fixed at 0[msec].

Each subject of the variable switching pause group chatted with the agent whose switching pause varied from 0[msec] to 9600[msec] with 13 levels separated by 800[msec] with two peaks as shown in Fig. 5. The utterance time was fixed at 0[msec/character] in this group so each message was displayed in its entirety.

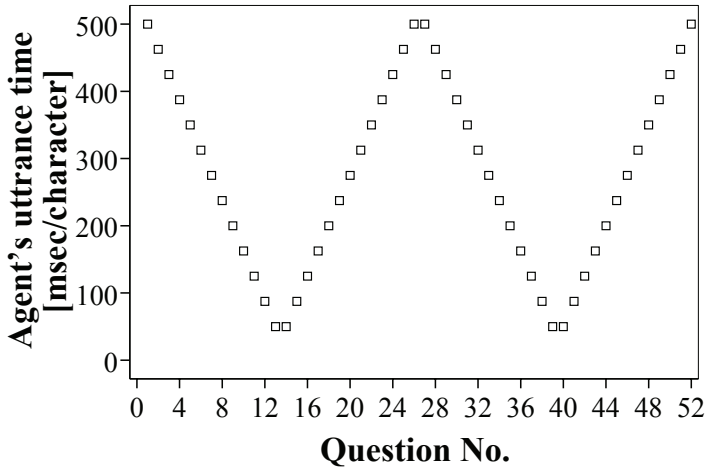


Fig. 4. Variable utterance time: agent.

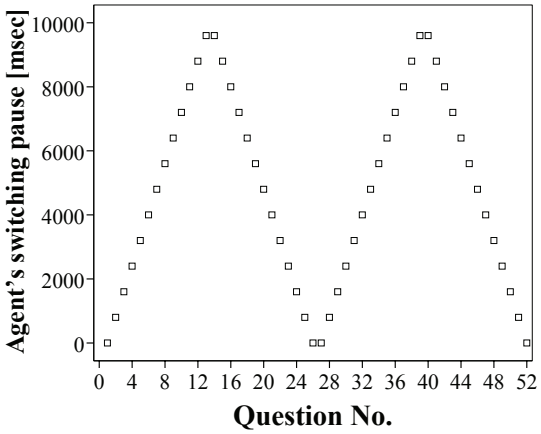


Fig. 5. Variable switching pause: agent.

3.2 Results

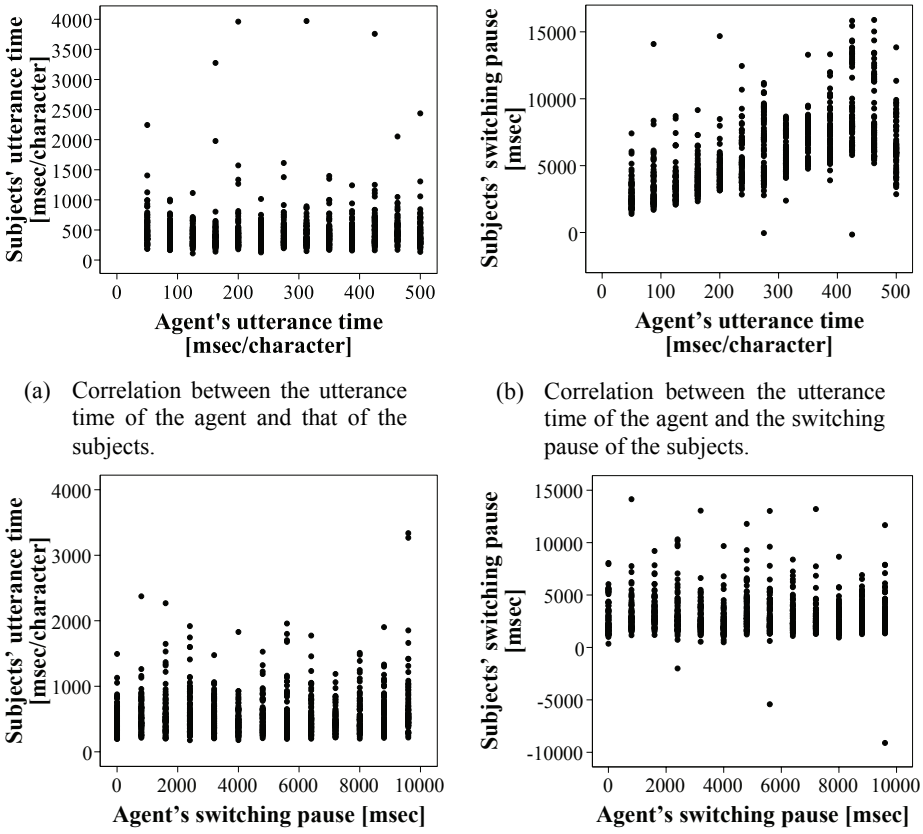
We plot variations between the utterance speed of the agent and that of subjects in Fig. 6(a). The vertical axis represents the utterance time of the agent and the horizontal axis represents that of the subjects.

Multiple-regression was used to determine the variations between the utterance speed of the agent and that of a subject [12]. We used the utterance speed of subjects as outcome variable, that of the agent and subjects as predictor variable. Subject was treated as categorical factor using dummy variable with 19 degrees of freedom. The *p* value from *t* test for the regression slope of the utterance speed of the agent was used to determine the probability of the analysis. The magnitude of correlation coefficient between the utterance speed of the agent and that of subjects was calculated as square root of (sum of squares for the utterance speed of the agent) / (sum of squares for the utterance speed of the agent + residual sum of squares). The sign of the correlation coefficient was given by that of the regression coefficient for the utterance speed of the agent. The correlation with correlation coefficient of 0.049 ($p=0.111$) was found, but this correlation is not significant.

Fig. 6(b) shows variations between the utterance time of the agent and the switching pause of the subjects. The correlation coefficient is 0.661 ($p<0.01$), so the correlation is significant and relatively strong.

Fig. 6(d) shows variations between the switching pause of the agent and that of the subjects. The correlation coefficient is -0.05 ($p=0.109$), so the correlation is not significant. Fig. 6(c) shows variations between the switching pause of the agent and the utterance time of the subjects. The correlation coefficient is 0.071 ($p<0.05$), so this correlation is significant but weak.

Table 2 summarizes the experiment's results. The utterance speed of the agent entrains the switching pause of the subjects but the other combinations do not yield any significant entrainment.



(a) Correlation between the utterance time of the agent and that of the subjects.

(b) Correlation between the utterance time of the agent and the switching pause of the subjects.

(c) Correlation between the switching pause of the agent and the utterance time of the subjects.

(d) Correlation between the switching pause of the agent and that of the subjects.

Fig. 6. Correlations of the utterance time and the switching pause between the agent and the subjects.

Table 2. Summary of experiment's results

		Subjects	
		utterance speed	switching pause
Agent	utterance speed	no significant correlation	relatively strong correlation
	switching pause	weak correlation	no significant correlation

4 Discussion

4.1 Typing ability of subjects

The above results show there is a relatively strong correlation between the utterance time of the agent and the switching pause of the subjects, but there is no significant correlation between the utterance time of the agent and that of the subjects. We discuss here the failure of the utterance time to establish entrainment.

Table 3. Correlation between the utterance time of the agent and that of each subject.

No.	Correlation coefficient	p
1	0.009	0.955
2	0.215	0.161
3	0.084	0.607
4	0.306	0.043*
5	0.481	0.002**
6	0.198	0.197
7	0.300	0.041*
8	0.009	0.951
9	-0.055	0.724
10	0.102	0.517
11	0.061	0.698
12	0.326	0.018*
13	-0.088	0.538
14	0.382	0.010*
15	0.206	0.150
16	0.203	0.209
17	0.062	0.684
18	-0.012	0.942
19	0.211	0.155
20	0.143	0.320

* $p < 0.05$ ** $p < 0.01$

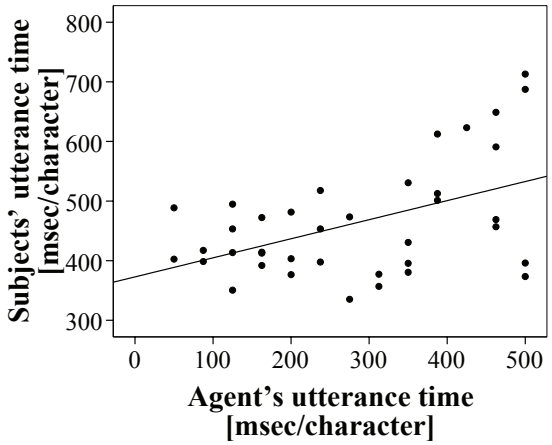


Fig. 7. Correlation between the utterance time of the agent and that of subject No.5 whose correlation is strongest in the group.

Table 3 shows the correlations between utterance time of the agent and that of each subject. It indicates that 5 of the 20 subjects showed significant correlations ($p<0.05$),

though the correlation was weak. Fig. 7 shows the correlation between the utterance time of the agent and that of the subject (No.5) who showed the strongest correlation.

The reason why most subjects were not entrained by the agent may be due to their limited typing ability. Although the subjects may have been entrained with regard to utterance speed, they may not have been able to express it because of their limited typing ability. In future work, we would like to perform this experiment again using some approach to offset the subjects' typing skills.

4.2 Switching pause of the agent

The experiment showed that the switching pause of the agent failed to establish entrainment. The reason may reflect how the subjects interpreted the switching pause of the agent. After this experiment, some subjects said they thought the chat system or the network was out of order. The subjects thought that the change in switching pause was merely due to a system problem but not due to the agent.

4.3 Affective influence on subjects

Changes in the agent's utterance speed influenced the subjects. In the experiment, the utterance speed of the agent entrained the switching pause of the subjects. The increased utterance speed of the agent goaded the subjects to replying more quickly. In future work, we need to examine the affective influence of the agent's utterance speed on the subjects in detail, for example by using a questionnaire survey or by measuring biometrics such as sweating or cardiac rate.

4 Summary

Communication between people involves not only verbal information but also non-verbal information, and the non-verbal information also creates an impression of the speaker. We find that entrainment can be established by non-verbal information.

Conventional studies on entrainment consider voice communication between two people or between a person and an agent, but we deal with entrainment in text communication. In this paper, we examined how the two factors of non-verbal information, utterance speed and switching pause, can entrain the responses of human subjects. For this experiment, we developed a chat system which uses typing speed to be used as an indicator of utterance speed.

We conducted an experiment to show that the utterance speed of an agent entrains the switching pause of subjects. We could not observe any other significant entrainment among the other combinations of utterance speed and switching pause. We can conclude that the utterance speed is effective in establishing entrainment, and that switching pause is entrained by the utterance speed.

In future work, we will examine the affective influence of utterance speed on people. We expect this work is lead to that development of interface agents that

affectively interact with human users through the use of non-verbal information in text communication.

References

1. Nagaoka, C., Draguna, M., Komori, M., Nakamura, T.: The Influence of Switching Pauses on Interpersonal Perception in Dialogues. In: Proc. of Human Interface 2002, (2002) 171-174 (in Japanese)
2. Nagaoka, C., Draguna, M., Komori, M., Kawase, S., Nakamura, T.: The Mutual Influence of the Speakers' Switching Pauses in Dialogues. In: Proc. of Human Interface 2001 (2001) 221-224 (in Japanese)
3. Nagaoka, C., Komori, M., Nakamura, T.: The interspeaker influence of the switching pauses in dialogue. Japanese Journal of Ergonomics **38**(6) (2002) 316-323
4. Watanabe, T., Okubo, M.: Physiological Analysis of Entrainment in Face-to-Face Communication. Proceedings of the Seventh International Conference on Human-Computer Interaction (1997) 411-414
5. Komatsu, T., Morikawa, K.: Entrainment in the Rate of Utterance in Speech Dialogs Between Users and an Auto Response System, Journal of Universal Computer Science **13**(2) (2007) 186-198
6. Fogg, B. J.: Persuasive Technology: Using Computers to Change What We Think and Do. Morgan Kaufmann (2003)
7. Vargas, M. F., Words, L. T.: An Introduction to Nonverbal Communication. Iowa State University Press (1987)
8. Patterson, M. L.: Nonverbal Behavior: A Functional Perspective. Springer-Verlag, Berlin Heidelberg New York (1983)
9. Picard, R.: Affective Computing. MIT Press (1997)
10. Cassell, J., Sullivan, J., Prevost, S., Churchill, E.: Embodied Conversational Agents. MIT Press (2000)
11. Prendinger, H., Ishizuka, M.(Eds.): Life-Like Characters. Springer (2004)
12. Bland J. M. and Altman D. G.: Calculating correlation coefficients with repeated observations: Part 1 – correlation within subjects. *Brit Med J*, 310 (1995) 446
13. Reeves, B. and Nass, C.: The Media Equation. CSLI Publications (1998)