Is a Face-to-Face Conversation Model Applicable to Chat Conversations?

Kanayo Ogura †

† School of Knowledge Science,
Japan Advanced Institute of
Science and Technology

Kazushi Nishimoto ‡

‡ Center for Knowledge Science,
Japan Advanced Institute of
Science and Technology

1-1, Asahidai, Tatunokuchi, Nomi, Ishikawa, 923-1292 Japan
{k-ogura,knishi}@jaist.ac.jp

Abstract

In this paper, we discuss turn-taking in chat conversation by analyzing typing histories. From these typing histories, it emerges that users of chat systems begin making their next utterances before their partners finish making their comments, and that they tend to continue making their present utterance even if the other's utterance, which isn't semantically related to their utterance, is shown on their display. These facts suggest that chat system users tend to make utterances in their own way without being limited by others. Although other researchers have explained that the lack of timing information leads to complexity in the history of chat and difficulty in pursuing topics, chat users may create multiple threads on purpose, in order to develop many topics simultaneously.

1 Introduction

The Internet now makes it possible for us to communicate with many people using new communications mechanisms like Computer Mediated Communication (CMC). In particular, chat systems have gained popularity as tools for real-time conversations. Because chat conversations are synchronous communication, like face-to-face (FTF) communication, many people think that they can communicate in chat systems in the same way that they do in FTF communication. However, chat conversation differs from FTF conversations in the process of making utterances. In standard chat systems, chat users cannot share the situation of typing and reading messages because the companions are not in a shared space. Especially, a lack of timing information makes it difficult to decide the timing for sending messages. To solve this problem, many chat systems have a function for transmitting awareness during a conversation, for example information about whether chat users are typing or not and on what keys chat users are typing. However, these systems have not managed to reduce overlap-like phenomena and the construction complexity of conversations. Additionally, these systems have seemingly not been accepted by chat users. We attribute these problems to sufficiency of conversational awareness in chat because we believe that the model of FTF conversations can be applied to on-line chat conversations. However, the properties of chat conversations may not be the same as those of FTF conversations. In chat conversations the characteristic features of the process of making utterances, the timing of sending messages and the style of chat can be different than in FTF conversations.

In this paper, we discuss turn-taking in chat conversation by analyzing typing histories of typing, in order to explain how chat users develop their conversations. The rest of the paper is organized as follows. Section two explains the process of making utterances in chat conversations and some chat systems that have a function for transmitting conversational awareness. Section three explains the experiment, how we recorded chat conversations and processing of the typing history data. Section four shows the results of our turn-

taking analysis, and section five concludes with some final remarks.

2 Related Work

2.1 Differences in FTF and chat conversation processes

Though there are some similarities between FTF conversations and synchronous chat conversations, there is a difference in terms of whether the companions share space or not.

In FTF conversations, we can always monitor the communication process verbal and nonverbal information expressed by eyes, gestures, intervals between utterances, intonations, modulations, inflections, tones and subject matter. In addition, we can understand what our partner will talk about before they finish speaking because the speaker's utterance can be heard sound by sound.

In chat conversations, we cannot monitor the process of making utterances in the same way, because the other chat participants are not in front of us and we cannot see them. We can share only displayed messages that other users have finished typing and sent. We cannot see the message making process; therefore, we have difficulty in starting to make our next utterances, because we cannot obtain information fast enough for smooth turn-taking. This causes the phenomenon of complexity in developing topics, which seldom occurs in FTF conversations.

2.2 Chat systems transmit conversational awareness

To help users tackle the difficulty of deciding the timing in making and sending their messages, many chat systems have been developed which give information on 1) whether other users are typing a message or not and 2) what the other users are typing.

2.2.1. Chat systems with the function of giving whether the other users are typing a message or not

Two widely used chat systems, MSN Messenger[1] and Yahoo Messenger[2], have a function that indicates whether the other users are typing a message or not. These systems visualize information at the bottom of the window about whether the chat users are logged into the system and whether they are typing a message or not. However, these systems don't tell us what the others are typing. It is hard for us to praise these systems because we cannot clearly know when the other participants will actually send their messages, even though the systems display the message "your partner is typing a message now!!"

Another system, Tangible Chat [Yamada et.al. 03] uses another non-visual method to indicate whether or not the other users are typing a message. Tangible Chat communicates the state of the other users' typing using keystroke vibrations, including dynamics. This system was developed to communicate emotion using stress of the keystrokes. As a result, it is useful not only for communicating emotion, but also for determining the timing of composing and sending utterances. In evaluating the effectiveness of timing utterances by analyzing the semantic relations of adjacent turns in chat histories, we have difficulty in concluding that this solves the timing problem, because the number of semantically irrelevant turns does not decrease significantly.

2.2.2. Chat systems which show what the other users are typing

In the systems that indicate what the other users are typing, one which makes the typing state visible is UNIX's Talk[3] and a system for sharing timing information [Ogura et.al. 03]. The interface of this system has a main window and a sub window. The main window has a historical function in the chat conversation. In the sub window, each user's utterances are displayed on a keystroke-by-keystroke basis. In evaluating this system, as with Tangible Chat, we have difficulty in solving the problem by analyzing the semantic relations of adjacent turns. In addition to making the typing state visible, this system preserves the user's place at the start of typing by displaying his/her taking the floor. In evaluating the place-saving function, they reported that changing the sequence of turns though these systems has the function of commu-

[1] MSN Messenger: http://messenger.microsoft.com/

[2] Yahoo Messenger: http://messenger .yahoo.co.jp/

[3] This program communicates what the user's types in real time.

nicating what others are typing in real time. This result suggests communicating what others are typing in real time does not solve the timing problem for making utterances.

Fugue [Rosenberger et.al. 00] and Free Turn Chatting System [Yamada&Takeuchi 03] have make typing visible in a main window in which each character of the utterance is noted along the horizontal continuum, left to right. Though one of the advantages of chat conversations is that they enable us to easily see histories of the conversations, these particular systems make it difficult to see the history of a conversation because the range for displaying utterances is limited and they have an interface design problem. The Free Turn Chatting system accommodates only Japanese chat conversations. The standard process of making utterances in Japanese requires transforming hiragana into kanji, but in order to make the typing visible in real time this system uses only hiragana, thus increasing the burden of understanding utterances. In an analysis of conversations by Fugue, they reported that the conversation was chorus-like, with short turns that continuously overlapped one to the next. These systems seem to be effective in their approach to solving the problem of timing, because they allow us to make utterances simultaneously and to ignore turn-taking. But we should make a distinction between these systems and standard chat systems, because there are big differences in interface design and conversational behavior between them and the standard systems in common use.

As we know, many chat systems include a function which communicates awareness of the conversational situation, in order to tackle the problem of timing utterances in chat. But they don't address the basic causes of the timing problem and have not been popularized as part of a standard system. It seems that the correct process of making utterances and the actual problems involved in making them are not completely understood and that the conversational model of FTF communication might not be applicable to chat conversations.

3 Data collection of chat conversations

3.1 The experiment for collecting conversational data in chat

We obtained conversation data by engaging in free talks, involving two participants, for about twenty-five minutes on a standard chat system. We got two types of data for ten conversations. The first data was a conversational history including time of receiving message by host server, name of participants, and utterances. The second was a typing history from each computer including the keys that the participants typed, time of typing, and type of working program.

All subjects consist of fourteen persons (ten men and four women) and graduate students[4]. They are computer users at an inter-mediate and higher level, with experience in using chat systems (including Instant Messenger) and no problems with typing. The partners in these experimental conversations were acquainted with each other, so they could avoid concentrating on the timing of their utterances while trying to discover who their partner was.

3.2 Method of segmentation

For analysis, we segmented each utterance in the typing history in each computer. Figures 1 and 2 show how we segmented the utterances.

20:tes2 (PM 2:33:48): USJつれてくよん
Shall I guide you in USJ?
21:tes1 (PM 2:33:55): 案内してして！
Yes, of course, please guide me!
22:tes2 (PM 2:34:02): OK～
OK～
23:tes1 (PM 2:34:09): もちろん美味しいところもね。
Of course, lead me to where delicious foods are served.
24:tes2 (PM 2:34:15): たこ焼きか
It might be TAKOYAKI.

Figure1: Conversational history (English Translation)

[4] All of pairs in this experiment are different. (There are no same pairs.)

```
---------------------------------------------------------------------
14:33:48 N
14:33:48 [ENTER]  ←finish in 20
14:33:59 [SHIFT]  ←start in 22
14:33:59 O
14:33:59 K
14:33:59 [SHIFT]
14:34:00 [ENTER]
14:34:01 [BS]
14:34:01 [SHIFT]
14:34:02 [ENTER]
14:34:02 [ENTER]  ←finish in 22
14:34:07 N        ←start in 24
14:34:08 A
---------------------------------------------------------------------
```

Fugure2: Typing history of user 'tes2'

The completion of an utterance is usually carried out by pressing the ENTER key. We apply time information on [ENTER] in Figure2 to time information for receiving the message in Figure1. In these results, the completion of utterance numbers 20 and 22 is recorded. The next typing after an utterance has been completed is regarded as the start of the next utterance. Incidentally, the time of utterance 24 in Figure1 is different from typing information given for the start of the utterances Figure 2 because utterance 24 was modified half way through the typing. We will explain this phenomenon later.

After we segment each utterance by each participant, we created Figure3 and analyzed the timing of the utterances.

4 Analysis of the process of making utterances

4.1 The start of making utterances

At the start of typing an utterance we find one of the following two situations.

1) Participant A waits for participant B to finish making an utterance, and then starts typing after B's utterance has been displayed.
2) Participant A does not wait for participant B to finish, that is, he/she starts typing before B's utterance is displayed.

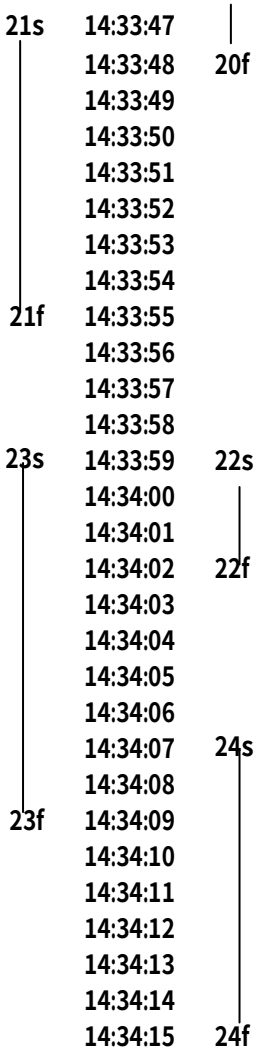


Figure3: The timing of making utterances (the left: user 'tes1', the right: user 'tes2')

The completion point in turn-taking during FTF conversations doesn't always come after speaking completeness [Kida et.al. 01] [Enomoto 03]. A speaker usually starts speaking after understanding the intention of the other's utterance. Situation 1) corresponds to the process in FTF conversations.

We classified all of the data shown in Figure3 into either situation 1) or 2). Table1 shows these results.

	Number of utterances (percentage)
1) after	476 (46.0%)
2) before	559 (54.0%)

Table1: The result of classifying the timing of utterances

The result shows that participants tend to begin making utterances before they understand the other's utterance. We suppose that chat participants might not intend to take turns.

4.2 Modifying utterances with incoherence for previous utterance

While one chat participant (A) is making an utterance before the other participant (B) finishes making one, as in situation 2) in section 4.1, B may continue typing, to finish making his/her utterance. In this situation, there can be the following two possibilities.

2-1) What participant A is typing and what participant B is finishing is coherent, and they can maintain a semantic sequence in adjacent turns.
2-2) What participant A is typing and what participant B is finishing is incoherent, and they cannot maintain a semantic sequence of adjacent turns.

	the number of utterances (percentage)
2-1) coherent	385(68.9%)
2-2) incoherent	174(31.1%)

Table2: The result of semantic relations between complete making utterances and half-making utterances

We classified situation 2) from Table1 into either situation 2-1) or 2-2); the result is shown in Table2.

In almost 70% of cases, what A starts typing is coherent to what B is already typing, even if he/she starts typing without waiting for B to complete the utterance. We cannot see whether half-made utterances are semantically related to the preceding utterance or to a previous utterance, because we don't make a strict distinction among these types of utterances.

When what participant A is typing is semantically incoherent with what participant B is completing, he/she has to modify the half-made utterances to avoid disturbing conversations. We classified situation 2-2) from Table2 into the following situations.

2-2-1) Participant A continues making utterances without modification, although he/she should modify them.
2-2-2) Participant A stops making utterances in order to modify them because he/she should modify what he/she is typing.

	The number of utterances (percentage)
2-2-1) no modification	96(55.2%)
2-2-2) modification	78(44.8%)

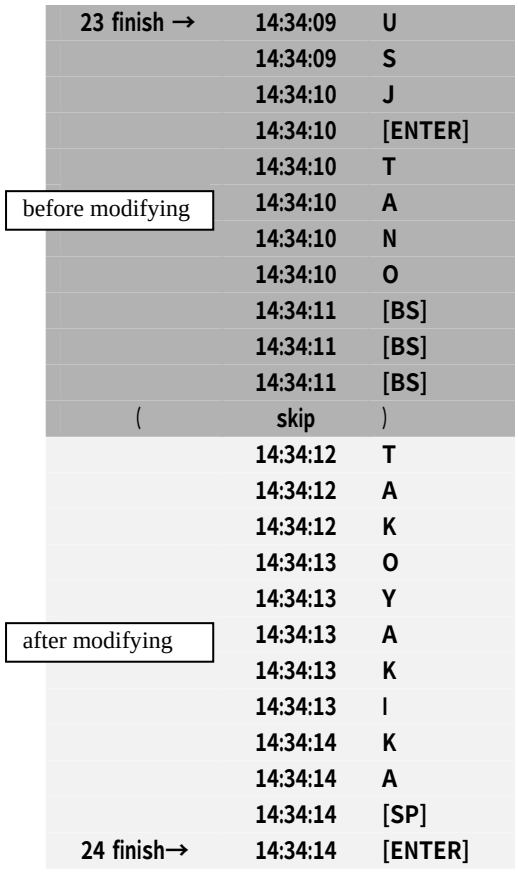
Table 3: The result of whether a participant modifies half-made utterances or not when the utterance is semantically incoherent.

Table 3 shows that the chat participants tended to continue making utterances without modification, even when the other participants finish making utterances and the utterances is semantically incoherent with the half-made utterance. We suppose that some participants might not intend to take turns; they develop conversations at their own pace because they can control conversations easily.

Whether chat participants modify half-made utterances or not is determined by comparing the typing history with the conversational history. We dealt with typing errors as utterances which should not be modified.

We checked the time of the preceding utterance in a conversational history with the typing history. If what A types before B finishes is different from after the time, we consider that the user modified his/her utterances. We can understand that participant A (tes1) completed making utterances at time of 14:34:09 when we compare Figure 4, the typing history of the participant 'tes2', with Figure1.

What 'tes2' types is different between colored and uncolored area in Figure 4. This participant modified the colored area to the uncolored area and finished making utterances by selecting that in the uncolored areas. In this case, user 'tes2' read the other user 'tes1's' completed utterance while typing his/her own utterance and understood 'tes1's utterance not about "USJ" but about "Food". The user 'tes2' noticed that what he/she was typing was not related to "Food", therefore deleted the utterance regarding "USJ" and made anew utterance about "Food" (especially "TAKOYAKI") and completed this utterance.



Fugure4: The process of modifying by user 'tes2'

5 Conclusion and future work

In this paper, we examined turn-taking in chat conversation by analyzing typing histories.

The result of our analysis shows that participants tend to start making utterances before they understand the other's utterance. We suppose that chat participants might not intend to continue in a turn-taking manner.

In addition we analyzed whether participants' utterances are semantically coherent or not when they begin making the utterance before the other participant has finished. In almost 70% of cases, what participant A starts typing is coherent to what participant B finishes typing, even if he/she start typing without waiting for B to finish. We could not tell whether half-made utterances were semantically related to the preceding utterance or to a previous utterance, because we don't make a strict distinction among these types of utterances.

Moreover, we analyzed whether A modified half-made utterances when what he/she was typing was semantically incoherent with what B completed. These results showed that chat participants tend to keep on making utterances without modification, even when the other participants had finished making utterances and it was clear that those utterances were semantically incoherent with half-made utterance. We therefore conclude that the participants do not intend to take turns, and they develop conversations at their own pace because they can control conversations easily.

References

[Enomoto 03] Enomoto, M.: When Does the Start His Turn? The Turn-Taking Rules in Japanese Conversation Apply Retrospectively After a Possible Completion Point Has Passed, Cognitive Studies, 10(2), pp.291-303. 2003.

[Kida et.al. 01] Kida, A., Inui, H., Kanzaki, K., Takanashi, K. and Isahara, H.: Grammar in Dialogue: Smooth Turn-Taking using Japanese Sentence Construction, SIG-SLUD-A102-6, pp.33-38. 2001.

[Ogura et.al 03] Ogura, K., Masuda, T. and Ishizaki, M.: Building a New Internet Chat System for Sharing Timing Information, Proc. of the 4th SIGdial Workshop on Discourse and Dialogue, pp.97-104. , 2003.

[Rosenberger et.al. 00] Rosenberger, T. M. and Smith, B. K.: Fugue: A Conversational Interface that Supports Turn-Taking Coordination, In Proc. of the 33rd Hawaii International Conference on Systems, 2000.

[Yamada&Takeuchi 03] Yamada, Yuji and Takeuchi, Y.: Development of Free Turn Chatting System and Analysis of the Dynamics of Social Dialog, Technical Report of IEICE, HCS2002-50, pp.19-24. 2003.

[Yamada et.al 03] Yamada, Yuko, Hirano, T. and Nishimoto, K.: Tangible Chat: A Chat-system that Conveys Conversation-context-awareness by Transmitting Vibration Produced by Key-stroke-act, IPSJ Journal, Vol.44, No.5, pp.1392-1403. 2003.