Turn-taking in Graphical Communication: an exploratory study

Atsue Takeoka[1,2], Atsushi Shimojima[1,2], Yasuhiro Katagiri[2]

[1] School of Knowledge Science
Japan Advanced Institute of Science and Technology
1-1, Asahidai, Tatsunokuchi-machi, Nomi-gun, Ishikawa, 923-1292, Japan
{takeoka, ashimoji}@jaist.ac.jp

[2] ATR Media Information Science Labs
2-2 Hikaridai Seika-cho, Soraku-gun, Kyoto, 619-0288 Japan
{katagiri}@atr.co.jp

Abstract

This paper reports an investigation of the turn-taking functions of drawings in graphical communication. Based on the examination of dialogue data collected that involve collaborative drawing interactions, as well as spoken dialogue interactions, in a joint problem solving task, we found that 'drawing turns,' in which the drawer presents a piece of information to her partner through drawing, had almost the same turn-keeping effects as 'speech turns.'

1 Introduction

Dialogue is characterized by taking turns, that is, who should talk next and when should they talk. Studies on turn-taking in dialogues have mostly focused on phenomena and processes of dialogue participants taking turns in spoken conversations. But speech is not the only means of interaction. We draw maps to give directions. We draw plans to discuss floor plans. We make diagrams to solve problems. It is crucial to take account of these non-speech modes to obtain a comprehensible model of taking turns as the basic mechanisms of human interactions. It is crucial to elucidate the nature of turn-taking in multi-modal setting, as communication tools become increasingly multi-modal to include such modes as free drawings, photographs, web page references and other visual information presentations.

Several researchers have looked at roles of non-verbal cues, e.g., eye gaze, gestures and other bodily movements, in coordinating turn-taking in spoken utterances. Non-verbal cues were regarded as ancillary to speech by providing additional signals to support orderly turn-takings in speech domain. In contrast, when we expand the scope of dialogues to include multi-modal interactions, we need to consider the possibilities of non-speech turns, such as drawing turns, and cross-modal turn changes, such as those between drawings and speech.

Based on graphical communication data we collected, we conducted a preliminary analysis on the existence and the functions of drawing turns. We found that drawing turns have almost the same turn-keeping power as speech turns.

2 Mechanism of dialogue interactions: turn-taking and grounding

Sacks, Schegloff and Jefferson (1974) noted surprisingly orderly nature of taking turns in our ordinary conversations, and proposed a set of rules to characterize the underlying mechanisms for the turn-taking structures in dialogues.

We conduct dialogues with turn-takings governed by the following rules.

At each transition-relevance place[1] of each turn:

- -If during this turn the current speaker has selected A as the next speaker, then A must speak next.
- -If the current speaker does not select the next speaker, any other speaker may take the next turn.
- -If no one else takes the next turn, the current speaker may take the next turn.

In this notion of taking turns, the current speaker presents a piece of information to the partner in a turn. When a turn consists of clauses, it presents a series of information. Even when a turn consists of a single word, it also presents a piece of information. To take a turn, the partner needs to

[1] Transition-relevance place is where the structure of the language allows speaker shifts to occur.

issue an utterance that is more than a particular signal of acceptance of the information presented by the current speaker.

From a different perspective, Clark (1996) developed the notion of *grounding* to capture processes with which conversants establish a set of information as a shared *common ground* for on-going (and possibly subsequent) dialogue. In a grounding process, a conversant presents a piece of information verbally or non-verbally, and the partner issues a particular public signal to show the receipt, understanding, or acceptance of that information.

On the basis of Clark's work, Traum (1994) substantiated the process of grounding in a finite-state transition model. The model specifies what contributions individual utterance in a dialogue make to a grounding process and how individual contributions combine themselves to make a sequence that completes the grounding. In Traum's terms, such contributions are *grounding acts*, the units of utterance that can perform grounding acts are *utterance units or UUs*—according to their contribution to a grounding process, the sequences of utterance units that complete the grounding are *discourse units or DUs*. Table 1 shows the seven categories of grounding acts and their definitions.

Table 1. Seven categories of grounding acts

Initiate (init)	An initial utterance component of a discourse unit.
Continue (cont)	A continuation of a previous act performed by the same speaker.
Acknowledgement (ack)	An acknowledgement claiming or demonstrating understanding of a previous utterance. It may be either a repetition or paraphrase of all or part of the utterance, an explicit signal of comprehension such as "ok" or "uh huh", or an implicit signaling of understanding.
Repair (repair)	Changes the content of the current DU.
ReqRepair (reqRepair)	A request for a repair by the other party.
ReqAck (reqAck)	Attempt to get the other agent to acknowledge the previous utterance.
Cancel (cancel)	Closes off the current DU as ungrounded.

Now, we can define a turn as a constituent of one or more DUs that initiate a piece of information. The following figure schematically shows speech turns consisting of DUs that are made of annotated UUs.

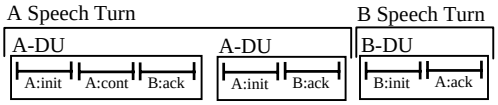

Figure 1. Speech turns defined by the notion of DUs and UUs

In graphical communication, especially in settings where one speaker gives information to a partner, communicator not only speak but also draw to present a piece of information. Thus, using the definition of speech turns mentioned above, natural questions arise as to 'drawing turns', namely, turns that have analogical characteristics attributed to speech turns.

This paper investigates this question in dialogues in which one is instructed to inform one's floor plans to a partner. Do drawing turns have similar turn-taking rules to speech turns? How do drawing turns keep their turns?

To address these questions, we collected a spoken dialogue data with drawings. However, since participants used only one pen at a time, no such phenomena as taking turns with a pen were observed[2]. Under this setting, our purpose in this study is to account for the effect of keeping turns when participants are taking drawing turns.

3 Method

3.1 Data

Our data consist of 2 dialogues. Both dyads were familiar with each other: in one case they were close friends and in the other they were mother and son. The objective of each dialogue is to inform each other of their floor plans. We prepared a pen and papers for the first dyad and two pens and papers for the second. However, the second dyad seldom drew at the same time. A video camera was used to videotape dyads and drawing surfaces from above. They participated on a voluntary basis.

3.2 Transcription

We transcribed 16 minutes and 30 seconds of

[2] The partner could take a turn by using their finger to draw transparently. However, this rarely occurred in our dialogues.

dialogues in total. According to the objective of each dialogue, a person in each dyad first represented their plans with a drawing, and then they communicated by pointing or adding drawings for further explanation. Then the other person in each dialogue did the same. Therefore, we had four drawers in two dialogues. Hereafter, we refer to a drawer as X and the other party as Y. We took one representative portion from each drawer. The length of the transcriptions varied from one minute and 35 seconds to 5 minutes and 42 seconds. All audible words and word fragments were transcribed, including overlapping speech, nonlexical fillers (such as "uh"), and other vocalizations (such as laughter and harrumph). Silences of more than 100 milliseconds were also written down and the time recorded.

After audible transcriptions, we added segmented drawings in transcriptions, at which point of the utterance the drawing segment starts and ends. We defined the starting point as the time the drawer begins to draw with a pen and the ending point as the time the drawer changes grip, does not move the pen on the shortest track, or takes their hand off of the drawing surface.

3.3 Coding

On the basis of Traum's finite-state transition model of grounding, we divided all speech in our dialogue data into *utterance units* (UUs), namely, "continuous speech by the same speaker, punctuated by prosodic boundaries (including pauses of significant length and boundary tones)" (Traum, 1994). One of the authors then classified each utterance unit into one of the seven categories of *grounding acts*.

4 Results

4.1 Summary of statistics

In this section, we show a summary of the statistics of speech and drawing data. First, we counted the number of drawing segments and divided them into two groups. One group of drawings were started during X's utterance, followed by X's utterances or X's utterances followed the drawings. Another group of drawings were started during Y's utterance or between Y's utterances. We call the former X-drawings and the latter Y-drawings. There were 149 X-drawings and 8 Y-drawings. We found four drawing segments by Y with Y's finger but we excluded these from this summary.

Second, we counted the number of utterances. Table 2 shows a summary of the statistics of the dialogue data.

Table 2. Number of Utterance Units and their breakdown of the seven categories of grounding acts and number of Discourse Units

	Total	X	Y
UU	1028	587	441
Grounding Acts			
init	290	234	56
cont	269	218	51
repair	8	6	2
reqRepair	6	0	6
reqAck	5	5	0
ack	344	70	274
ack init	108	54	52
DU	395	287	108

*We also excluded utterances relevant to Y's finger drawings.

In this setting of dialogues, only X informed Y by drawing. However, X initiated 2.7 times more DUs than Y did. The objective of the dialogues was for one party to inform their floor plan to the other party. These summary data showed that information flowed from X to Y not only by drawings but also by utterances.

4.2 Silences

We first looked at the occurrences of silence in speech. There were 436 periods of silence of more than 100 milliseconds in the portion analyzed. Silence was classified into 2 categories: during drawings and not during drawings. There were 164 of the former and 272 of the latter. Figure 2 showed the frequency distribution of silent periods for each group.

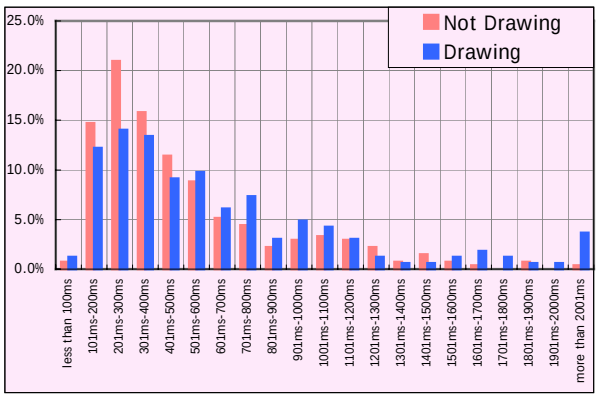

Figure 2. Frequency distribution of silent periods of more than 100ms

The average time of silences during drawing was 667ms and that of during not drawing was 519ms.

The data indicates that silences during drawing tend to be longer than those during not drawing. This suggests that drawings function as independent turns. An occurrence of a drawing turn can keep the next speech turns from starting, despite no speech is uttered at the time.

4.3 Overlaps

If drawing turns work similarly to speech turns in terms of their turn-keeping functions, interruptions by overlapping speech should occur about the same amount toward drawing turns and speech turns. Case 1 in Figure 3 schematically shows a regular speech turn change structure, whereas Case 2 and Case 3, respectively, show a speech turn and a drawing turn interrupted by following speech turns.

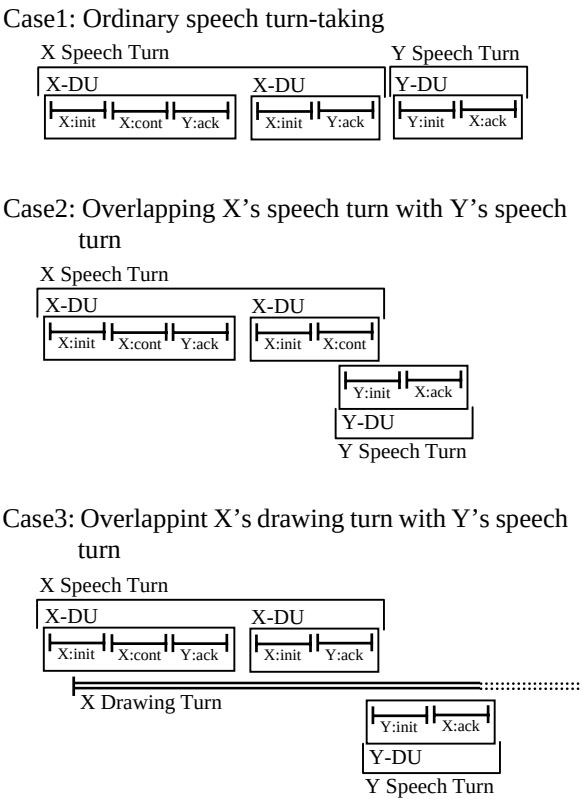

Figure 3. Speech turns and drawing turns interrupted by speech turns

Table 3 shows the relative amount of speech turns and drawing turns which were followed by overlapping next speech turns. The result appears to indicate that drawing turns are more amenable to interruption than speech turns, though the difference is not statistically significant (Z_0= 1.26, p=0.21 two-tailed).

Table 3. Number of speech turn during not drawing (see Case 2 in figure 3)

Total of X speech turns	153
Overlapping X speech turns	17
Ratio	11.1%

Table 4. Number of drawing turn during drawing (see Case 3 in figure 3)

Total of X drawing turns	149
Overlapping X drawing turns	25
Ratio	16.8%

Close examination of Case 3, in which drawing turns are interrupted by the following speech turns, revealed that some of the overlapping speech turns were actually responses, either acknowledgments or request for repairs, directed toward drawing contents. An example of an acknowledgment in the drawing context is found in the excerpt below.

act	UU		Utterance
$init_{26}$	41.1	X:	kou kocchino houni [mizumawariga ari]
			(like here there is a sink)
ack_{26} $init_{27}$	42.1	Y:	[uraomote na]
			(it is located back to back)

Here, X was drawing a line between two houses she already drew before, and directing Y's attention toward the location of kitchen sinks in them. Y could recognize the location of the kitchen sinks in the two houses from the drawing, and Y's subsequent short utterance "uraomote na" verbally confirmed the relative locations of them. Such utterances were directed toward drawings and worked to give acknowledgments toward their contents. These types of utterances might appear as instances of initiation acts if we only look at speech domain, but, actually, they do not constitute independent initiation acts if we admit grounding acts that work across different modalities.

Table 5 shows the result of recounting speech overlaps by excluding cross modal responses. The result shows that drawing turns and speech turns are exactly the same (Z_0=0.26, p=0.32 two-tailed)

in terms of their vulnerability toward interruptions.

Table 5. Number of drawing turn during drawing (see Case 3 in figure 3)

Total of X drawing turns	149
Overlapping X drawing turns	19
Ratio	12.8%

5 Discussions

Generally, when it is one's turn to speak, the other party is less likely to cut into the turn. They try to present their information clearly to avoid overlapping. In other words, the present speaker's speech turn is some form of resistance to the other party's interruptions. On the other hand, during one's drawing turn, even though part of the drawing period is filled with silence and during that silence the other party could easily utter 'init' utterances, the data on periods silence showed that longer silent periods were allowed during drawing. We suppose that drawing turns has some form of resistance to keep the other party from interrupting their to speak.

In this study, the percentage of overlapped X's speech turns was 11.1%. That is to say, X's speech turns had 89.9% of resistance to keep turns. In the same way, X's drawing turns was interrupted by Y's 'init' or 'ack init' utterances at 12.8%, which showed almost the same extent of resistance, 87.2%. Thus, our comparison suggests that the turn-keeping resistance of X's drawing turns is almost equivalent to that of X's speech turns.

The equivalence of the extent of resistance of speech and drawing turns also suggests that silence during drawing tended to be longer not because the other party could not speak while drawing but because he/she did not want to speak while drawing. In other words, drawing could make an excuse for not speaking while preserving the smoothness of turn-taking interactions.

Finally, within the data we analyzed, all the cases in which a speech turn overlaps a preceding drawing turns were accompanied by co-temporal speech by the drawer. This is schematically shown in Case 3 of Figure 3. In more than half of these cases, the following speech makes an overlaps to the preceding speech as well as to the drawing (Case 4 in Figure 4). In contrast, although theoretically possible, we did not find any cases in which bare drawing turns, e.g., without accompanying speech by the drawer, were interrupted by the following speech turns (Case 5 in Figure 4).

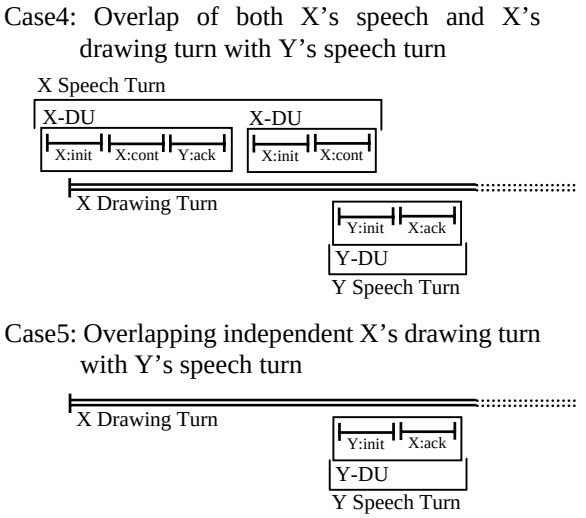

Figure 4. Drawing turns interrupted by a speech turn

6 Conclusion

We conducted a preliminary study on the drawing turns in graphical communication. Based on collected dialogue data that involve graphical as well as spoken interactions, we conducted an analysis of both speech and drawing turns. We found that drawing turns have almost the same turn-keeping effect as speech turns. Further investigations are necessary on both semantic and grounding contributions of drawings to develop fuller account of graphical interactions.

References

Clark, H. H. 1996. *Using language*. Cambridge: Cambridge University Press.

Sack, H., Schegloff, E. A. & Jefferson, G. 1974. A simplest systematics for the organization of turn-taking in conversation. Language, 50, 696-735.

Traum, D. R. 1994. A Computational Theory of Grounding in Natural Language Conversation. Technical Report 545. University of Rochester.