



COMMUNICATIVE MODE DEPENDENT CONTRIBUTION FROM THE RECIPIENT IN INFORMATION PROVIDING DIALOGUE

Yukiko Ishikawa

NTT Network Information Systems Laboratories
1-2356, Take, Yokosuka, Kanagawa, 238-03 Japan
e-mail yukiko@nttnly.ntt.jp

Abstract

This paper claims that the contribution of the recipient in an information providing dialogue depends on the communicative mode. First, we analyze conversation data and show that the confirmations uttered by the information recipients and the response to them by the information giver differ in three different communicative modes: telephone conversation without map (tel-non-map), telephone conversation with maps (tel-map), and face-to-face conversation with map sharing (face-to-face). The results of an empirical study show that the recipient acknowledges the giver's utterance equally in all modes, but literal confirmation was most frequently used for tel-non-map while inferential confirmation was used for tel-map. The giver's response to the confirmation utterances also depends on the communicative mode. The giver responds more collaboratively in tel-non-map than the other two modes. Second, we discuss the empirical results by using two factors that characterize the three communicative modes; "the communication channel" and the "information available to the participants". Most of the results can be interpreted consistently by combining these two factors.

1 INTRODUCTION

Conversation is a collaborative joint action performed by its participants in order to accumulate common ground between them [2]. In "information providing dialogues" such as explaining how to use some equipment, how to cook something, or how to go somewhere, the conversation participants play one of two roles in the dialogue: either "the information giver (giver)" who explains how to do it or "the information recipient (recipient)" who is told how to do it by the giver.

In this kind of dialogue, the initiative mainly rests with the giver. It is, however, important to investigate the effect of the recipient's utterances because the recipient not only listens to the giver's explanation but also actively contributes to the conversation, for example, he appropriately displays his level of understanding to the giver according to the communication situation.

One of the most important contributions to the conversation performed by the recipient is issuing confirmation utterances that show that the recipient is paying attention to, recognizing to, and understanding the giver's utterances. In this paper, we analyze the recipient's confirmation utterances and the giver's response to them in order to investigate how the giver and the recipient accumulate common ground through recipient initiated events and how interac-

tions depends on the communicative mode.

Some examples of confirmation utterances are given below.

Dialogue 1

A: *and, there is a book store on the right hand side if you go straight down the street.*

To this piece of information there were three different utterances:

B1: *Yeah.*

B2: *Yeah, a book store on the right hand side.*

B3: *So, I go toward the City University?*

Each utterance **B1**, **B2** and **B3**, confirms A's utterance. **B1** returns only a back-channel response. **B2** repeats a part of A's explanation. **B3** refers to something inferred from A's explanation. As shown in the examples, the recipient can generate a variety of confirmation utterances.

One goal of this paper is to confirm that the type of confirmation used by the recipient as well as the giver's response depends on the conversation mode. For the route explanation task, the conversation mode can be characterized according to the following two factors: (a) bandwidth of communication channel (face-to-face vs. not-face-to-face) and (b) information available (with-map vs. without-map). A second goal is to show that the empirical results can be accounted for by these two factors.

Previous studies on communicative modes have compared the characteristics of discourse in face-to-face, telephone, and keyboard communication. They have focused mainly on the giver's explanation utterances [4, 9], or taken no account of the participants' roles [1, 5, 8]. On the other hand, the present research focuses on the recipient's contribution. Therefore, it can reveal other aspects of discourse characteristics that depend on the communicative mode.

2 EXPERIMENT

Task: The task for the giver was to explain to the recipient the way to his home from the nearest station. The task for the recipient was to listen to the giver's explanation and understand the route completely.

Method: Sixteen subjects (eight male and eight female subjects) explained the route using three different modes: telephone conversation without map (tel-non-map) where the participants can not see each other and do not have a map, telephone conversation with maps (tel-map) where the participants can not see each other, but each has a copy of the

Table 1: Characteristics of each communicative mode

	Tel-non-map	Tel-map	Face-to-face
Communication channel	speech	speech	speech pointing
Information available		map (not-shared)	map (shared)

same map, and face-to-face conversation with map sharing (face-to-face) where the participants can see each other and share the same map. Therefore, 48 conversation data were collected and analyzed. These experiment conditions were characterized by two factors: bandwidth of the communication channel and information available. **Table 1** shows the characteristics of each condition. In tel-non-map, conversation participants communicate only by speech. The recipient has no map of the locality. In tel-map, the participants communicate only by speech, but each participant has a copy of the same map. In face-to-face, they look at the same map and can use speech and pointing for communication. Therefore, the map is regarded as shared information.

3 RESULTS

3.1 Confirmation by the Recipient

The recipient confirms the giver's utterances in different ways as shown in **Dialogue 1**. We classified the recipient's confirmations according to the strength of evidence of understanding [2, 3, 11]. The recipient's response to the giver's questions and requests were not taken into consideration because the purpose of this study is to analyze how the recipient actively collaborates in developing conversation. We discovered three different classes of utterance surface expressions.

Acknowledgment: Returns back-channel response such as Yes, Yeah, Uh ¹.

Literal confirmation: Confirms that the recipient accurately heard the giver's utterance and recognized the linguistic meaning of that utterance.

Repeat: verbally repeating whole or part of the giver's previous utterance.

Paraphrase: paraphrasing the linguistic meaning of the giver's previous utterance

Inferential confirmation: Confirms that the recipient understood the giver's explanation correctly by indicating what the recipient has inferred from the knowledge gained up to that point.

Logical inference: Indicates inferences based on giver supplied information.

¹This type of utterance has some function in discourse [10]. It could be used to show that the recipient fully understands the giver's utterance and agrees with him/her. It, however, only assures that the recipients have acknowledged the giver's utterance

Table 2: The mean of contribution from the recipient in each communicative mode

	Types of Confirmation Utterances		
	Acknowledgement	Literal Confirmation	Inferential Confirmation
Tel-non-map	41.56	6.75	5.69
Tel-map	38.00	5.19	7.94
Face-to-face	31.81	2.88	3.56

Cognitive inference: Indicates inferences based on information gained from sources other than the giver (ex. reference to the map or what the recipient already knows).

Acknowledgment is the weakest contribution because it only assures that the recipient is paying attention to the giver. Inferential confirmation is the strongest contribution because it provides the recipient's literal understanding of the giver's utterance and more positive commitment to the conversation. In **Dialogue 1**, B1 is an example of acknowledgment, B2 is Literal confirmation, B3 is Inferential confirmation.

Average number of recipient confirmations per conversation for each mode is shown in **Table 2**. As the results of ANOVA with repeated measures, the total numbers of contributions are not statistically different between the three modes ($F(2,30)=2.822$, NS). The numbers of acknowledgment also showed no statistically significant difference between the modes ($F(2,30)=1.363$, NS). Literal confirmation was most frequent in the tel-non-map mode (6.75) and the difference in numbers of literal confirmation among the three modes was statistically significant ($F(2,30)=4.081$, $p<.03$). Inferential contribution was most frequent in the tel-map mode (7.94), and the difference in numbers among the three modes was also significant ($F(2,30)=3.969$, $p<.03$).

These results indicate that the total number of confirmations issued by the recipient does not depend on the communicative mode: the type of confirmation does however. The recipient issued the same number of acknowledgments regardless of the communication mode. Linguistic information was frequently confirmed in the tel-non-map mode. Inferential confirmation was most frequently observed in the tel-map mode. Therefore, it is suggested that the recipient most positively committed to the conversation in this mode. These empirical results verify that the recipient's contribution to the conversation depends on the communicative mode.

3.2 The Giver's Response to the Recipient's Confirmation Utterance

The previous section confirmed that recipients change their way of confirming the giver's explanation depending on the communicative mode. It is also important to analyze the giver's response to the recipient's confirmation because the communication act established by the recipient's confirmation is completed by the giver. It is expected that the giver's response depends on the type of the recipient's confirmation utterance and the communicative mode. An example of the giver's response is shown below.

Dialogue 2

A: If you turn right, you quickly find a big intersection.

B: The first one?

A: Yes, It's a big intersection, TATSUMI-BRIDGE.

In this example, B confirms the place of the intersection by issuing an inferential confirmation. The giver then makes a positive response and provides information about the name of the intersection.

In this section, we examine how the giver's response depends on the confirmation type and the communicative mode. As observed data, we collected responses to literal confirmations and inferential confirmations because the giver usually does not respond to acknowledgments. Moreover, we did not analyze negative responses because it is obvious that the giver usually corrects erroneous confirmations. We analyzed the responses to correct confirmations. The responses were found to consist of four classes.

No-response: No response in anyway.

Acknowledgment: Returning a back-channel response or saying only "Yes" to the recipient's confirmation.

Literal response: Repeating and paraphrasing the recipient's previous utterance.

Inferential response: Elaborating the recipient's utterance by giving additional information or re-explaining something from other point of view.

No-response is the weakest contribution by the giver and inferential response is the strongest because it adds supplemental information. The recipient's response in **Dialogue 2** is an example of an inferential response.

We compared the rate of acknowledgment as the simplest response and the rate of inferential response as the most complex response. The frequency and proportion of these two kinds of responses are shown in **Table 3**. As the response to literal confirmation, in all modes, simple acknowledgment was more frequent than inferential response. On the other hand, the response to inferential confirmation depended on the communicative mode. In tel-non-map, the rate of acknowledgment equaled that of inferential response (40.7%). In the other modes, the rate of the acknowledgment was higher than that of inferential response, especially in the tel-map mode (54.5% for acknowledgment, 19.3% for inferential response).

These results indicate that the giver simply acknowledges literal confirmations regardless of the mode while the response to inferential confirmation changes according to the communicative mode. In the tel-non-map mode, the recipient issues complimentary and cooperative responses; simple acknowledgment is preferred in the tel-map modality.

4 DISCUSSION

The previous sections analyzed the confirmation utterances performed by the recipient and the responses issued by the giver. In this section, we discuss the empirical results for each communicative mode.

Table 3: The number and proportion of the giver's response

Response to Literal Confirmation			
	Acknowledgement	Inferential Confirmation	Total
Tel-non-map	62(58.5%)	16(15.1%)	106
Tel-map	47(56.6%)	11(13.3%)	83
Face-to-face	25(54.3%)	7(15.2%)	46

Response to Inferential Confirmation			
	Acknowledgement	Inferential Confirmation	Total
Tel-non-map	22(40.7%)	22(40.7%)	54
Tel-map	48(54.5%)	17(19.3%)	88
Face-to-face	15(38.5%)	11(28.2%)	39

In [7], mental models are constructed in the comprehension of discourse. Mental models do not have to be imagistic. They consist of multiple representations of discourse. [6] discussed that in comprehending illustrated texts, the mental model integrates two sources of information: descriptive propositional information derived from the text and perceptual information from the picture. Based on these ideas, understanding of the route, in the route explanation task, could be regarded as a process of constructing a mental model by integrating information provided by the giver and that known by the recipient. That is,

$$P \wedge R \Rightarrow Q$$

P: Information provided by the giver through the communication channel.

R: Information available to the participants such as domain knowledge and map information.

Q: Mental model that is constructed based on **P** and **R**.

In this experiment, the amount of information provided by the giver is decided by the communication bandwidth (face-to-face vs. not-face-to-face) and the amount of information available to the recipient depends on whether the map can be used or not (with-map vs. without-map)².

As shown in **Table 1**, in the tel-non-map mode, the recipient has to construct the mental model **Q** based only on the information contained within the giver's speech **P**. In this relatively poor communicative mode, it is important for the recipient to catch the giver's words accurately in order to get the necessary information. That is why, in the tel-non-map mode, literal confirmation was most frequently used by the recipient. The giver, on the other hand, usually responded to literal confirmation with a simple acknowledgment in all modes. It is considered that literal confirmation is used to confirm the literal content of the giver's utterances, and the giver is not required to make inferences in order to compare the confirmation contents with his mental model; a simple acknowledgment suffices for the recipient if the confirmation is correct. With inferential confirmations in the

²The information available to the recipient is decided by map availability. It is assumed that all subjects had the same amount of domain knowledge because none of them knew about the route being explained

tel-non-map mode, however, it is difficult for the giver to be sure that the recipient's inferential confirmation is correct because the participants do not have information that can be used as common background. The giver has to check the inferential confirmation by comparing it with his own mental model. Therefore, the giver is not sure that they have realized common ground even if the recipient's utterance is correct. That is why the giver responded more cooperatively to inferential confirmations. He often added supplemental information and re-explained what he had mentioned before (inferential response). These cooperative responses are effective in offsetting the poor communication bandwidth and limited information sources.

In the tel-map mode, the recipient has to integrate the information from the giver's speech **P** and the information on the map **R** into a mental model. In order to integrate **P** and **R**, the recipient has to confirm information that is on the map but that was not mentioned by the giver. That is why the number of inferential confirmations was most frequent in this mode. The giver's response, however, was quite simple. In this mode the participants can use the map as common information and the mental model of each participant should be based on the map. Therefore, it is considered that they realize common ground by checking some places on the map. That is why simple responses were made to inferential confirmations in the tel-map mode.

Finally, in the face-to-face mode, the participants can use information from the speech channel as well as other communication channels such as pointing. Moreover, the map is shared. In this mode, the giver traces the route on the map while explaining it to the recipient. Therefore, it is not necessary for the recipient to integrate **P** and **R** because the giver explains the route by coordinating speech and map information; the recipient can understand the route by following the giver's explanation. That is why there were fewer literal and inferential confirmations than seen in the telephone modes. The recipient simply acknowledged the giver's utterances, the same as in the telephone modes. In the face-to-face, the giver is very sure that the recipient understands the explanation. So, it is reasonable to assume that the giver responds to a confirmation in a simple way, but this was not clear in the data observed. We have to analyze face to face communication in more detail.

5 CONCLUSION

In this paper, we focused on the confirmation issued by the recipient as his/her contribution to the conversation. The results of an empirical study indicated that the types of confirmations, and the giver's response to them, depended on the communicative mode. We also explained these results using two factors: the bandwidth of the communication channel and the information available to the participants. As future works, in order to utilize this empirical model in a multimodal dialogue system, it is necessary to establish a prediction model that takes into account the recipient's level's of understanding. It is also necessary to investigate communication using computer interfaces such as pointing devices, zoom up, and highlighting because this study analyzed completely natural human-to-human conversation.

References

- [1] Arita, H., Kogure, K., Nogaito, I., Maeda, H., & Iida, H. Media-Dependent Conversation Manners -Comparison of Telephone and Keyboard Conversations-, *SIG-NL-61-5, Information Processing Society of Japan*. (in Japanese). 1987.
- [2] Clark, H. H., & Schaefer, E. F. Contributing to Discourse. *Cognitive Science*, Vol. 13, pp. 259-294, 1989.
- [3] Clark, H. H. Managing Problems in Speaking. In *Proceedings of International Symposium on Spoken Dialogue*, pp. 181-184, 1993.
- [4] Cohen, P. R. The Pragmatics of Referring and the Modality of Communication. *Computational Linguistics*, Vol.10, No.2, pp.97-146, 1984.
- [5] Chapanis, A., Parrish, R. N., Ochsman, R. B. & Weeks, G. D. Studies in Interactive Communication: II. The Effects of Four Communication Modes on the Linguistic Performance of Teams during Cooperative Problem Solving. *Human Factors*, Vol. 19, pp. 101-126, 1977.
- [6] Glenberg, A. M., & Langston, W. E. Comprehension of Illustrated Text: Pictures Help to Build Mental Models. *Journal of Memory and Language*, Vol. 31, pp. 129-151, 1992.
- [7] Johnson-Laird, P. N. *Mental models*. Cambridge, MA: Harvard University Press, 1983.
- [8] Kowtko, J. C., Isard, S. D., & Doherty, G. M. Conversation Games within Dialogue. Research Paper HCRC/RP-31, Edinburgh, U. K.; HCRC, 1992.
- [9] Oviatt, S. L. & Cohen, P. R. The Contributing Influence of Speech and Interaction on Human Discourse Patterns. In Sullivan, J. W., & Tyler, S. W. (Eds.) *Intelligent User Interfaces*, pp. 69-83, Addison Wesley, 1991.
- [10] Shimazu, A., Kawamori, M., & Kogure, K. Analysis of Interjectory Responses in Dialogue. *SIG-NLC-95-9, Institute of Electronics Information and Communication Engineers*. (In Japanese). 1993.
- [11] Walker, M. A. Redundancy in Collaborative Dialogue. In *COLING 92*, pp.345-351, 1992.