Cooperative Interactions Generated by Incompleteness in Robots' Utterance

(ロボットの言葉足らずな発話が生み出す協調的インタラクションについて)

January, 2021

Doctor of Philosophy(Engineering)

Yusaku Nishiwaki
西脇 裕作

Toyohashi University of Technology

別紙４－１（課程博士（英文））

Date of Submission（month day, year）: January 8th, 2021

Department of Computer Science and Engineering	Student ID Number	D143364	
Applicant's name	Yusaku Nishiwaki		

Supervisors	Michio Okada Tetsuto Minami

Abstract（Doctor）

Title of Thesis	Cooperative Interactions Generated by Incompleteness in Robots' Utterance

Approx. 800 words

The conversation has always got attention as a way for people to interact with system s naturally. Improvements in machine learning have made it possible for systems to u nderstand human speech and human intentions. Therefore, systems such as smart spe akers have become commonplace. It is important to clarify the role of people and syst ems. The basis of the interaction is that the person speaks commands, and the syste m understands them. The system has focused on specializing its functions on an indiv idual, such as recognizing speech and understanding intention of one person. However, Mikhail Bakhtin mentioned that incompleteness is crucial in interacting with others, beyond conveying information. The incompleteness is a factor that allows the speaker to give the other person a new interpretation of part of the text. He said that incomp leteness is a chance to create new meanings with them. According to Lotman, the info rmation function of conversation and relationship-building function are not two separat e things. For Human-Computer conversation development, it is necessary to evolve the system conversation discussing the relationship-building function that involves a perso n, aiming at conversation created with others. This study tested a method to prepare for ambiguity, which has not been discussed and examined its effects.

First, we investigated the structure of speech in asymmetrical relationships between p eople and the practice of eliciting others' participation in conversations. We could see the practices as other-initiated repairs and fishing-devices that encourage others' parti cipation in our daily conversations. For example, sociolinguistic science described tellin g part of the story as a trigger to elicit participation from those who have more respo nsibility. In this study, we developed an "incomplete utterance strategy" in which the robot's speech is made lacking by removing additional pieces of information from its s peech. The possibilities and effects of this speech strategy were investigated through t hree experiments.

Then, to investigate the validity of the incomplete utterance method, two strategies were tested. We compared human behaviors between two conditions, one semantically incomplete utterance and structurally incomplete utterance in human-robot conversation. The results showed that people increased their responses to the semantically incomplete robots through active participation in the conversation (questions, the introduction of new content related to the conversation). Next, we examined what impressions were obtained from conversations between humans and incomplete utterance robots. We used two of the most average interactions from the last experiment. We asked for new participants to evaluate the two conversations. The results showed that participants perceived conversations with humans and semantically incomplete utterance robots as cooperative. Finally, I tackled the question: Can an incomplete utterance elicit human involvement? We analyzed whether incomplete utterances elicit human participation in terms of changes in participation attitudes. A

multi-party conversation was set up to analyze the changes in participation attitude. Den reported it is easier to change participation in a multi-party conversation than in a one-to-one conversation. The experiment set up a multi-party conversation between a speaker robot, a listener robot, and a participant. The participants could choose to explain with the speaker robot or listen with the listener robot. We compared impressions and behaviors between the semantically incomplete utterance robot and the fully explaining robot. Results showed that people increased their participation as the speaker with incomplete utterance robot. Furthermore, they decreased their participation as the listener also. Besides, it was confirmed if the robot spoke incompletely, but participants joined the conversation, the rate of information transfer is estimated the same as the fully explained robot. However, this study found that this incomplete utterance method had limitations. It requires the speaker robot to select a person as the next speaker directly. The impression showed that although participants increased their speech amount with incomplete utterance robot, they did not feel like they were explaining together, which is a subject for future research.

We tested the incomplete utterance robot through three experiments. It was confirmed that the robot's incompleteness changes human behavior and participation attitude w hen the robot tries to engage directly with a person. The result also showed that com munication efficiency does not change if humans participate, even if the robot falls sh ort of words. So, it is possible to use this technology as an interaction technique to el icit human participation. However, this study's implementation was not enough to disc uss collaborative actions. It will be necessary to examine the model for long-term impl ementation in the future.
In recent years, robots that do not speak Japanese at all on purpose have been commercialized and are gaining popularity in Japan. A robot that conveys information accurately and has space for people to interact will continue to attract attention in the future. This study proposed a new conversational design, which is speech incompleteness. We investigated its effects and limitations. This result may contribute to the engineering field's development as a basis for the next generation of interaction technology for human-interactive systems.

Contents

List of Figures

List of Tables

1 Introduction

"Hello XX, what is the weather today?" "The current weather in Toyohashi is cloudy, and today's forecast is sunny, with a high of 24 degrees and a low of 18 degrees." "Thank you." Information exchange between people and systems has become familiar with the appearance of smart speakers and voice assistants. On the other hand, if we look at everyday Japanese conversations between people, such as "kyo wa yoka ttane (Today is so good.)" "ne. tenki ga yokute yokatta (Yes, the weather is so good.)" the conversation is going on without the speaker clearly telling what was good about it at the beginning. The conversation became a place to put on each other's thoughts. People seem to be able to communicate with deficiency and excess in a way that is not possible between people and systems.

According to Saeki [1] and Reddy [2, 3], developmental and cultural psychologists, and Tannen [4], a sociologist, there are two main types of communication. In one of the communications, the participants' roles in the communication are separated. On the other hand, the roles are not clearly separated. It could be said that the conversation between people and systems have been focused on the former type of communication. The conversation of systems has been developing remarkably in recent years. For example, question-and-answering, guidance, reception works are all about repeating understanding and answering. If a system could understand a user's intentions correctly, they can provide information and services. The users need to say their request correctly, and the system needs to understand it correctly, so the roles of the conversation were divided. Therefore, researchers focused on individual systems' skills, improving speech recognition, and natural language understanding. These abilities have been greatly improved with machine learning, and conversation between people and systems now spreads.

Porcheron points out that the “conversational” interface is named as an interface that uses spoken dialogue to interact with people. However, the “Human-System conversation” is different from the conversation people have with each other [5]. Conversations are not just about exchanging information, but it is more diverse and more prosperous. Humans can build relationships with others and creating new insights in conversation. Holmes has also pointed out that the transactional aspects of conversation, such as information transfer, and the phatic aspect, such as relationship-building, cannot be separated [6]. According to Lotman, humans change the ratio of the aspects depending on the situation [7]. Today’s conversations between people and systems have been developed by improving the transactional aspects of information transfer and intention understanding. However, if the conversation functions cannot be separated, it is necessary to consider interaction design with the phatic aspect and relating to others.

One of the principles of our conversation is that we are talking to someone else. Conversations are created not only by one’s thoughts and intentions but also by others’ intentions who may be different from one. For this reason, the system often loses control of conversations with unclear objectives, such as small talk. Even if the systems try to grasp the person’s thoughts and intentions entirely, it is not easy. Then, why don’t we stop trying to grasp the conversation entirely all the time and reduce the separated role a little? As the conversations between people, it may be essential for the systems to drop perfection and use deficiency and excess as humans [8].

This study attempted to examine the system’s conversation design to realize a jointly constructed conversation instead of dividing the roles into individuals such as “speakers and listeners.” This study is a challenge to shift the conventional design from information exchange to co-constructed conversations. This study hopefully contributes to new interaction technology between people and systems.

2 Research Backgrounds

2.1 Classification of Social Interaction

We meet and relate to various people and objects in our daily lives. These people include friends, teachers, juniors, and strangers we meet on the street. Recently, ticket vending machine systems and information robots have also appeared. It has become familiar that people interact with these systems daily. How does a person relate to those people and objects? This study relies on Mikhail Bakhtin's dialogism to categorize those interactions.

Mikhail Bakhtin studied social interaction, such as how the words of a novel related to the reader and how the words of teachers and parents interact with students and children, focusing on the "voices" exchanged by participants in the interaction [9, 10]. As Wertsch mentioned [11, 12], for Bakhtin, the notion of voice cannot be reduced to an account of vocal-auditory signals. Voice affects the speaking personality that involves someone who takes a certain perspective or belongs to particular cultural and social categories. Bakhtin said that the accumulation of "voices constructs the dialogue."

Moreover, "voices" could take on different aspects depending on those voices' ability to interact with other voices. Bakhtin classified "authoritative words" that always seek only approval and desire from the other person, such that they do not intersect with other "voices." The word is considered "a word whose speech and its meaning is fixed and does not change when it meets other new voices. [11, 12]" When such "authoritative words" are used in a dialogue, the listener can only receive them, which will result in a conversation with a clear division of roles between speaker and listener.

On the other hand, He categorized words that can commune with other voices and create meaning with other voices as "Internally Persuasive words." The word is considered a state in which the self and others are mixed. Others' participation creates meaning and words that could not be created by the self alone. It is said that "the semantic structure of the internally persuasive words is not complete but open." It allows the listener to interpret new meanings.

How do these words manifest themselves in today's conversation? Bakhtin gave examples of "authoritative words" being used in conversations between teachers and students and between fathers and children. Suppose one is not allowed to question or express one's own opinion. In that case, it is one of the situations in which "authoritative words" are used.

In contrast, examples of "internally persuasive words" are meetings or conversations with friends. People are often allowed to express their opinions and thoughts to each other. Situations in which people exchange opinions, create new ideas, and build relationships would be situations in which "internally persuasive words" are used. Thus, this study suggests one of the critical points in categorizing social interaction is whether there is openness for the other and whether the action is not complete to the self.

2.2 Conversations with Clear Purposes

Today's most used conversations between people and systems are smart speakers such as Siri [1], Google Assistant [2], and Alexa [3]. With the development of research in speech recognition technology and dialogue understanding, today, they are al-

[1] "Siri," accessed December 20, 2020, https://www.apple.com/siri/.
[2] "Google Assistant," accessed December 20, 2020, https://assistant.google.com/.
[3] "Amazon Alexa," accessed December 20, 2020, https://developer.amazon.com/ja-JP/alexa.

ready widely implemented in society, not just in the research world. Bakhtin's classification of social interaction, which of these "words" do these smart speakers use?

The conversational design commonly used in smart speakers, called Request / Response [5]. It requires people to have a clear purpose of what they want from the system. As in the example given at the beginning of Chapter 1, the system provides information that people seek. Speaker and human sometimes repeat the exchange of answers and questions according to that purpose. People can find such conversations in social situations when people buy tickets at the reception desk in a train station or answer customer support questions. It is said that these conversations require participants to say their objectives clearly and understand the words. Since machine learning has made it possible to establish this ability to some extent, the Conversational Interface field has made remarkable progress in recent years. A variety of products have been developed as businesses.

The current mainstream conversations seem to use conversations with a predetermined protocol. They do not allow others to intervene in the other person's words. The system and humans always held the own objective or information: "What do you want to know?" "Where do you want to go?" In other words, the current widespread conversation between people and systems has the aspect of "authoritative words."

2.3 Conversations Aimed Not Only at Transactional Function

Conversations intended to engage are commonly referred to as "non-task-oriented conversations." This classification of conversations is used in natural language processing, artificial intelligence and linguistics, and sociolinguistic sciences. "Small

Talk" is often mentioned as a conversation that refers to this non-task oriented conversation. Murata and Ide summarized the form of chit-chatting in Japanese [13]. The purpose of a conversation is not only to convey information or building consensus. The speaker and the listener also recognize each other and create a sense of togetherness and intimacy in conversation.

However, Is it possible to completely separate the function of relating to others from communicating information? Holmes points out that it is impossible to divide conversation into two categories completely [6]. According to Lotman, the ratio of information transfer and relationship-building functions determines the conversation aspect [7].

Much research on non-task-oriented conversations aims to construct conversations called "chats," They try to develop to introduce topics and allow for speaker turnover naturally [14]. However, from the stance of this study, as mentioned above, while conversations require information transfer, they simultaneously require engagement. Then, the system conversation needs to move forward with a discussion of relationships embedded in information transfer. It is essential to develop a system with more aspects of "internally persuasive words" in order to have a design that is both task-oriented and human connection-oriented.

2.4 Research Considering the Incompleteness

Considering how systems can have the aspect of "internally persuasive words," this study will point out the recent human-agent interaction (HAI) research. Several studies have not utilized linguistic methods, but they could generate interactions with human interpretation and motivation.

2.4.1 Non-Verbal Voice Communication

First of all, some studies in HAI research do not use Japanese or English languages. They use methods not used by adults, such as baby utters or animal barks, beeps. Beeps are among the most common sounds used in systems such as machines and robots. However, according to Komatsu et al., by changing the beep's fundamental frequency and length, people can estimate the state of the computer [15].

Humans can also understand the voices from the anime "Pingu" and Pokemon's sounds, based on their situation. Some researchers investigate the interaction between humans and those agents. Suzuki et al.' studied "non-syllabic sound." The sound was generated by using the echo-mimicry of human voices. It was found that people feel empathy to the voice [16]. Kiyomaru et al. confirmed that the Japanese word chain game between a person and an agent that uses only the "Do" and "La" sounds is possible [17].

The research has been conducted on how people interpret the actions of non-verbal agents to create interactions. Recently, there has been a movement in Japan toward non-verbal voice communication as a product (Figure 2.1), such as LOVOT by GROOVE X, Inc. [1] and BOCCO emo by YUKAI Engineering Co. [2]

Through these studies and products, it has been shown that an explicit language is not necessarily to produce interaction because interactions could be generated by human interpretation.

2.4.2 Design to Elicit Human Participation

Next, there is a study of design that elicits human participation without speech or language.

[1] "LOVOT," accessed December 20, 2020, https://lovot.life/en/
[2] "BOCCO emo," accessed December 20, 2020, https://www.bocco.me/en/emo/

Figure 2.1: The Figures of LOVOT (left) and BOCCO emo (right)

Yamaji et al.'s "Sociable Trash Box(Figure 2.2)" is a robot that does not have a mechanism to pick up trash on its own [18]. However, the trash box robots have motions toward trashes. The feature makes the interaction between humans and robots through a common ground of "trash." The result leads to the human putting trash into the robot. The robots were able to achieve to pick up trashes without their equipment. This phenomenon is told as an example of the "human-dependent robot" approach. Another approach eliciting human participation is "shi-ka-ke." Matsumura developed the trash can attaching a basket goal. Matsumura's trash can (Figure 2.3) was reported to increase trash collected. People have fun putting the trash in like a shot on goal [19, 20]. The studies of "shi-ka-ke" mentioned people's past experiences, such as culture, conventions, and pleasurable acts, become triggers to change people's behavior.

Figure 2.2: The Figure of Three Sociable Trash Boxes.

Figure 2.3: The Figure of the Trash Can Attaching a Basketball Goal

If trying to solve the problem of littering, it is possible to create a robot that can pick up trash by attaching an arm to a robot. However, the previous study showed that some purposes could be accomplishing things due to the robots' design, behavior, and mechanisms. The simple factors create human interaction without utilizing fancy algorithms, systems, and sensors.

This study considered that the same could be said for conversation. So far, the construction of dialogue systems has often involved the consideration of machine learning and dialogue processing algorithms. However, focusing on simple aspects such as the method of speech production and the conversation design has the chance to attract humans and the advantage of being easy to use in other systems and cost-effective.

2.5 Positioning of This Study

The general behavior of today's robots is to speak firmly and explain. However, as daily conversations between people, we could realize that they do not always tell everything. According to Mikhail Bakhtin's dialogism, if a robot always explains entirely to people, it becomes using the "authoritative word" that requires the other person only to accept it.

Baxter developed Bakhtin's dialogic theory into a communication theory of intimate relationships [21, 22]. According to her, discrepancy or uncertainty is considered one crucial element in the continuity of a relationship. Also, referring to constructing interaction in HAI research, it is possible to change the relationship designing interaction other than natural language processing algorithms in conversations.

This study tried dealing with the conversation design that does not convey too much information. It explores the co-construction of conversation instead of transferring information individually. Concerning examples of human conversations and methods of eliciting others' participation, the ways to make an opportunity for the system's utterances were examined. This paper will explore the incompleteness in one-on-one conversations and multi-party conversations. Incorporating incomplete elements into conversations can contribute to the development of conversation and interface technology between people and systems.

3 Incomplete Utterance Approach

Chapter 2 mentioned that including incompleteness in an utterance is one possibility to achieve co-constructed conversation and have a place to build relationships. This chapter starts with illustrating the relationships and conversations between multiple persons who show incompleteness to others. Furthermore, the practices of daily conversation are also described. After all, those phenomena are summarized as How do people elicit the participation of others. Finally, this study named the approach to have incompleteness in the Japanese robot's speech as "Incomplete Utterance" and explains its speech strategy.

3.1 Findings in a Child-Caregiver Conversation

This section will describe conversations between infants and their caregivers. The relationship between them is one of the most exciting relationships. Children are still in the process of language development. So they may not be able to tell a complete story by themselves or may not be able to explain things in standard terms due to the small number of words they know. However, in everyday conversations between infants and their caregivers, the infant's undeveloped language skills are not apparent.

Firstly, this section mentions one video. It is available on YouTube [1]. The

[1] "230_2歳11ヶ月子供『会話が膨らむようになってきた』conversation. 2year old child," accessed December 20, 2020, https://youtu.be/caeRbVA7Sxs.

transcript of the video is shown below. In the video, a two year and 11-month-old child named Towa has a conversation with her parents. The transcript symbols for describing behaviors are some of the symbols used in conversation analysis [23]. The details of the symbols are also shown in the appendix. It should be noted that there is no record of what kind of talks or events happen before the conversation. However, it can be seen that Towa and her parents were talking about other children. The main topic of conversation seems to be what Towa was doing with her grandparents when she was away from her home and parents.

```
 1     CHI: She pooped in the toilet.
 2          (0.7)
 3     MOT: hhhhhhhh
 4     FAT: hhh
 5     MOT: (hhh)Poo in the toilet
 6     CHI: Yeah.
 7 →   FAT: Hmm?(.)Which child?
 8          (1.4)
 9 →   MOT: e
10          (0.7)
11 →   MOT: She lives near your grandma's house?
12     CHI: Yes.
13          (0.6)
14     MOT: oh yeah.
15          (0.4)
16     MOT: Where has she been?
17          (0.4)
18     MOT: Where did you play today?(.)With that Kawasaki Midori-chan.
19          (0.7)
20     CHI: there
21          (0.4)
22     MOT: h(.)there(.)(hh)hhhhh
23          (0.3)
24     FAT: What's her name?
25          (0.8)
26     CHI: Kawasaki Minori-chan
27     MOT: hh(.)hhhhhhhhh(h)hhhh(.)(hhh)You're right.
28          (2.1)
29 →   MOT: What would she be like?
30          (1.3)
31     CHI: She will play with me.
32          (0.2)
33     MOT: I see.
34     CHI: yeah.
35          (0.1)
36 →   MOT: Midori-chan is:::Is she coming with her father or
```

```
37          mother or someone?
38     CHI: yes.
39          (0.6)
40 →   MOT: Who was coming with her?
41          (1.2)
42     CHI: Dad and Mom
43          (0.3)
44     MOT: Ah. Both
45          (0.5)
46     CHI: Both
47          (0.4)
48     MOT: really:::
```

Since the child might be at the grandparents' house alone, it might say that only the child had the experience of how she played. Alternatively, it may be said the mother knew how the child played because she had heard from the grandparents or the mother had seen it before. However, the phenomena that could be seen in the transcript are that the parents asked what their child did. As seen in the above transcript, Towa began by talking about other children. Then the father and mother questioned, "Where did you play today?" and asked where she was playing with her friend in the park. The conversations in the transcript generally run as child talk a little and her parent asked about that. The caregivers sometimes asked questions that can be answered with yes and no so that Towa can answer about what she did.

Children are limited in what they can explain because their language skills are developing. However, the developing skills usually are not shown because they appeared only after comparing their language ability with adults' one. It is also difficult to imagine that children have trouble communicating what they want to say because they laugh and enjoy the conversation. It is natural to think that both parents and child think they are communicating sufficiently.

Ordinal children are not able to have a presentation or cannot thoroughly explain something. If so, by whose ability does the conversation happen? The reason may be that children are limited in the number of people they talk to. They do not have to meet new people and explain things to them as adults do. They often speak in relationships with family members, relatives, teachers, neighborhood friends, and other people who live in the same place and have many shared events. Children know how to relate to such people. In conversations that occur in intimate relationships, they may not need to talk to strangers as adults do.

The situation of immature speaking can be evaluated as an unfortunate inability to have a conversation. It has been treated as a negative thing, at least for the system. However, let us look at it from a different perspective. It can be said that it is a conversation design that allows others to participate and create a conversation. Alternatively, there is a place or opportunity in this communication because the child cannot communicate alone. Also, the caregiver participates in the incomplete part of the conversation and creates it together.

This conversation design between child and caregiver could also be seen in a conversation between language learners and native speakers, although the situation is limited. First of all, one of the factors which limit the situation is the expectation of the conversation. For example, let us consider the situation of the presentation. It is a situation where the speaker needs to convey information and not expect help from the listener. If language learners have a presentation, they need to speak only with the capability they have. Usually, listeners are expected to inform, and the listener expects them to provide information. This situation cannot be said to be the same as a conversation between child and caregiver.

On the other hand, think about when we talk with friends from overseas. Even if ESL speakers cannot accurately convey information to the native speakers, they can still communicate through small talk. Fan et al. and Iwata reported that even if the speaker's linguistic ability is limited, the conversation can be controlled and continued by the listener's linguistic ability [24, 25]. In some cases, the speaker is not the only person who contributes to the conversation. The listener also plays a significant contribution. If the listener is allowed to assist, conversations like those between children and caregivers can occur even between adults.

3.2 Practices in Conversation

Section 3.1 mentioned relationships in which participants of a conversation are asymmetrical in their abilities and positions. It was noted that the listener contributes as a member of the conversation. However, such conversations are not limited to the relationship between children and their caregivers. Such conversation can also be found among ordinary adults. This section describes conversation practices that even if there is no difference in basic conversation abilities, a listener contributes to the conversation.

"Repair" is one phenomenon in which the listener participates in the conversation. The "repair" has been discussed mainly in the conversation analysis research field [26, 27, 28, 23]. The phenomenon is known as a way of repairing the problems that occurred by comprehension in forego utterances. There is a wide range of issues regarding the "repair." Hayano summarized the repair practices in Japanese conversations [29]. When speakers repeat or rephrase after their own utterances, it is one form of repair called self-initiated repair [26]. Repair initiated by others is also discussed. The "other-initiated repair" is well-known as repeat requesting, questioning, and understanding check, and more. Suzuki reported two types of other-initiated repair in Japanese conversation [30]. One of them is "Nani-ga? (What was that?)," which combines a wh- interrogative and a particle. The function of the form is to specify the problem in the forego utterance by a listener. The other is "Nani? (What?)," which forms only interrogative. This form is used for asking the whole forego utterance.

The conversation is an ongoing process of understanding among the participants. The listener's utterances have an essential role in structuring the speaker's utterances, as illustrated by the "repair" practices.

In addition, not only the speech but also the subtle behavior of the listener is one of the most important factors. Kushida and Mizukami have reported that stammering and pauses in speech (silence) cause others to participate in the conversation as if they were offering a helping [31, 32]. In these repair or helping, the speaker and listener try to solve the understanding of the conversation.

So far, this paper has discussed ways of participation from the listener that relate to issues of comprehension. On the other hand, it is known that the speaker leaves opportunities for participation in the utterance so that the speaker elicits the listener's participation. The practices are called "fishing device" shed light by Pomerantz [33], and "explanation prompting," as discussed by Toe [34].

Below transcript is an example used in the description of Pomerantz's "Fishing Device" [33], where speaker A says what he knows, which elicits speaker B, who knows more about the event. In line 1, speaker A stated that when he called, the line was busy. Speaker A describes what he or she knows, and speaker B, who knows why it happened, makes a statement to add content.

```
      [NB:II:2.-1 ]
1     A:   Yer line's been busy.
2 →   B:   Yeuh my fu(hh)! 'hh my father's wife called me...
```

Next, "explanatory prompting" is discussed in a conversation between a parent, child, and another adult in the below example. It is mentioned that the mother's participation is elicited by embedding topics related to her child in the conversation. The below transcript is from Toe's study translated by the author [34]. In line 1, a person described what the baby is doing now. The mother then describes the behavior as her child's mother, adding the meaning and reason for the behavior. This phenomenon provides a place to display what she knows about her child as the child's mother.

```
      Move with hips
      ((A is a staff member of a child care support circle, B is a mother))
1     A:   He was moving with his hips.
2     B:   oh, yes!
3 →   B:   He travel a short distance with it.
4     A:   Oh, really?
5     B:   yes
```

These phenomena have been observed to elicit participation in conversations by telling parts of the phenomenon to people who have detailed information about the conversation or are responsible for it.

3.3 Incomplete Utterance of This Study

This study will discuss a new conversation design that focuses on incompleteness, discussed in Mikhail Bakhtin's dialogue theory [9, 10]. The conversation design was named "Incomplete Utterance." So far, this chapter has given examples of incompleteness in conversations, and it works as an opportunity for participation. However, since children and language learners are composed of various factors, it is difficult to imagine that a robot can be equipped with the exact same incompleteness. Because robots have creators as their programming and design, it is difficult with current technology to perfectly mimic the numerous factors such as children and people. Furthermore, the relationships between children and caregivers have been built up over the years. It is also hard to reproduce the same thing in a laboratory. So, The relationship between the child and the caregiver is unlikely to be copied.

Therefore, this study uses the constructivist approach. The constructivist approach is known as a way of understanding things through making. Based on people's communication, some factors were selected for the new conversation design for robots. The effect of incompleteness can be examined by looking at how the selected elements work in the conversation. Since the researchers must select the factors, this approach cannot say that they have the same incompleteness as children or language learners. However, by dealing with incompleteness that has not been discussed before, it will be possible to discuss them sufficiently to create a new conversation design.

This research considered how can be implemented incompleteness into robots with these limitations. The conversation practices and conversations between infants and their caregivers can be considered "wordless" as a factor of incompleteness. The implementation can be grammar restrictions and some missing parts of speech to make robots' utterances wordless. This study proceeded with an implementation based on mimicking child-like grammar and conversation practices that elicit others' participation.

On the other hand, it should be mentioned that there were other ways to implement the information other than linguistic information. As described earlier in this chapter, stammering and silence are known to be opportunities for others to become involved. The nonverbal information is already starting to investigate. Matsushita et al. found that when a robot makes stutters depending on the listener's gaze, the experiment participants feel an increased sense of participation in conversation [35]. There are other reasons why this study did not use nonverbal methods. It requires large sensors to capture gaze information now, which increases the cost of robot manufacturing and limits the robot's appearance. The linguistic method was selected, which has not been studied so far and does not have any limitations, such as the robot's appearance.

3.4 Implementation

In order to apply incompleteness, this study employed grammatical features that mimic child-like features and conversation practices called "repair" and "fishing devices." First, this section illustrates the child-like grammar, which implemented to Incomplete Utterance approach.

3.4.1 Syntax as a Feature of Child's Speech

One characteristic of conversations between young children and their caregivers is children's underdeveloped language skills. They use a limited number of words or still have difficulty in explaining themselves to the adults. As one of the features of this child-like speech, the study of children's grammar was utilized. Amano reported on Japanese children's grammar [36]. There is a database called CHILDESD that records actual conversations between actual children and adults [37].

First, Amano compared children in a 4 to 5-year-olds class with children in a 5 to 6-year-old class. He investigated what percentage of children were able to reproduce various Japanese basic syntax. While there was no significant difference in the children's ability to produce syntactic constructions in 3-Phrase sentences, a significant difference was found among 4-Phrase sentences. If the form was like subject + object + verb, the children could reproduce it even at a very young age. This study shows that children have grammatical limitations before they can speak firmly.

Next, here discuss the sentence length of children's speech. For the sentence length of speech, there is a measure called Mean Length of Utterances (MLU) used by developmental linguists [38]. Furthermore, there are two types of MLU, morphological MLU and the independent MLU. This study chose the independent MLU because of its compatibility with the sentence clause analyzer. Also, from the Okayama corpus, a colloquial speech database [39, 40, 41], the author calculated independent MLU using 30 four-year-olds children's data. The results showed that the average utterance length was about 2.86, which means that each utterance is equivalent to three-Phrase sentences. Besides, Inaba reported the independent MLU of native Japanese children and adults. It is said that three-year-old children used 2.84 independent MLU, although adults used 5.84 independent MLU [42]. Indeed, the child seems to be speaking in shorter sentences. It would be possible to implement not making the speech too long to make people think that robots speak non-fluently.

3.4.2 Implementation Based on Conversation Practices

The next step is to examine the implementation regarding the "repair [26]" and "fishing device [33]" methods mentioned in Chapter 3. In the "repair" sequence, the utterance has elements lacking somehow, and participants try to fix it. Also, "fishing device" is a way of triggering participation from responsible for the topic by saying a part of the event that they happen to know. This responsibility includes being familiar with the conversation contents. Overall, these factors suggest that lacking content works as an opportunity for others to participate. It would be possible to implement the lack of some element from utterances for an opportunity to participate.

3.4.3 Procedure

The following is how Incomplete Utterance was implemented. First, a robot needs topics to talk about. For the topic, many commercial robots use news of current affairs or weather, and explanations from encyclopedias, and more. These contents are written mostly for informational purposes.

News and explanations are written to convey information and phenomena and organized to summarize the necessary information [43, 44]. Generating the "Incomplete Utterance," this study implemented limitations such as a limited number of sentences, missing object words, and additions such modality to make utterance spoken language. The conversation design's general goal is to give people the impression of a lack of stories that would make the listener feel missing.

In order to take the methods mentioned, natural language resources need to be analyzed. MeCab [1] and Juman [2]. were used for morphological analysis, and

[1] "MeCab: Yet Another Part-of-Speech and Morphological Analyzer," accessed December 20, 2020, https://taku910.github.io/mecab/

[2] "JUMAN (a User-Extensible Morphological Analyzer for Japanese) ," accessed December 20, 2020, http://nlp.ist.i.kyoto-u.ac.jp/EN/?JUMAN

CaboCha [3] and KNP [4] for dependency parser. In the case of news sources, mecab-ipadic-Neologd [45] was used as a dictionary file to improve MeCab's word analysis accuracy.

At first, the following figure 3.1 illustrates the overview of the implementation.

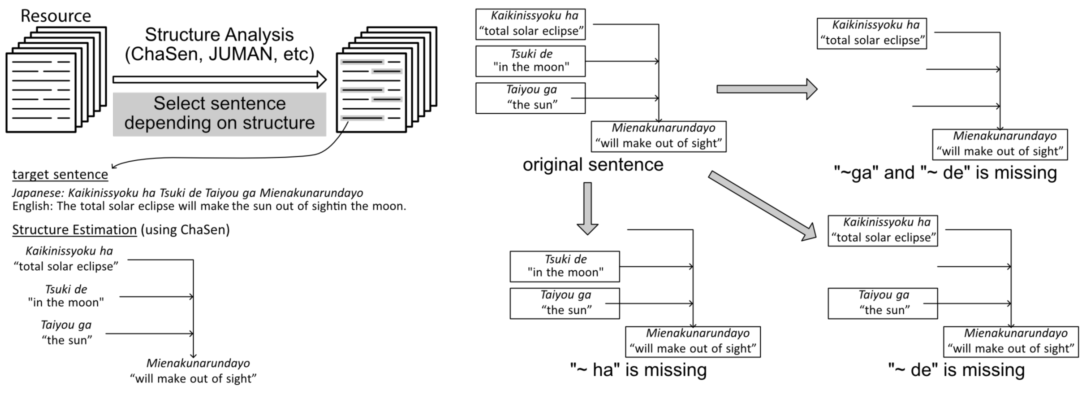


Figure 3.1: Overview of Incomplete Utterance Generation

(1) How to exclude target words

The part of the sentence will be excluded to implement the lack of content. There is a way to split Japanese sentences called Wakachigaki. A simple explanation of Wakachigaki is that it is a way to split a word and a preposition into a single set. Every target for excluding could be a phrase that consists of "word + preposition." The sentence analyzers can be used to get the connection between each of these sets. Japanese grammar is often structured as follows: a subject part, other elements, and a predicate part. It is common to have the subject of the action at the beginning and the verb at the end, as shown in Figure 3.1: subject + some element + predicate. Removing the middle element does not mean losing the subject and predicate, which is the sentence's primary meaning. Therefore,

[3] "https://taku910.github.io/cabocha/", accessed December 20, 2020, https://taku910.github.io/cabocha/

[4] "Japanese Dependency and Case Structure Analyzer KNP,", accessed December 20, 2020, http://nlp.ist.i.kyoto-u.ac.jp/EN/?KNP

this study employed this method in order to create the appearance that something is lacking.

(2) Syntax and number of phrases

It was considered that sufficiently long utterances might not allow people to sense the opportunity to participate. Limiting the length of the phrase of sentences was one additional factor in this study. Written text is often used as a resource for the robot's speech. However, Kaji and Hayashi have mentioned that written language, a speech resource, has a more complex vocabulary and longer phrases [43, 44]. More content in one utterance means more content is well-explained. Avoiding well-explained, this implementation use around three phrases, referring to the number of sentences in section 3.4.1.

(3) Modality and filler

Sentences such as news articles and stories are meant to be read by people and need to be changed as spoken words. To display that the robot is talking to the other person, modalities such as "ne" and "yo" were added. These modalities emphasize talking to others [46, 47]. To add the modality to verbs, MeCab is used to select the verb form according to the modality. Also, fillers such as "anone" and "ettone" were added to express the beginning of each speech turns.

4 Incomplete Utterance in One-to-One Conversations (1)

This chapter's primary purpose is to examine the incompleteness of "Incomplete Utterance" defined in Chapter 3. To analyze how people speak and behave to the robot using "Incomplete Utterance," one-on-one conversations are prepared. This chapter was composed based on Nishiwaki's previous papers [48].

4.1 One-to-One Conversation Between Human and System

Vlahos and Kawano summarize recent one-to-one conversational systems [49, 50]. They say that recent conversations between people and systems are low context conversations. In other words, they point out that the conversations do not need common knowledge with individuals. However, Hinds says that Japanese is a high-context culture where the listener also has a role in fulfilling the communication [51]. Therefore, a low-context conversation in Japanese would not always be natural. Although it is difficult for a system to implement a person's prerequisite knowledge, the system tries to have a good conversation with many questions and information to avoid misunderstanding. The systems talk in a way that people can clearly understand or do not mistakenly understand. When the system asks a question, speak so that it knows it is a question. It is seen that the system also speaks politely to ensure no lack of information that might require additional questions. However, if people have a conversation, "repair" practice is naturally happening. The misunderstanding often occurs in a daily conversation, so it may also not need for systems to be avoided excessively. Instead, there are ways to resolve the lack of common ground as humans do. This study supposes that the process of resolving incompleteness also be a process of communication. It is known as proof that Seaman [1], a video game released in the 1990s, provided good communication value in the way that the system was helped by people [50]. What would be the effect of prioritizing high-context methods in conversation as the Incomplete Utterance approach?

4.2 "Repair" to Build Understanding

The conversation between a person and a system that uses Incomplete Utterance is expected based on conversation practices mentioned in Chapter 3. In this study, the robot speaks Incomplete Utterance, so people need to deal with the problem of understanding. It is expected that people will begin the process of "repair" for the speaker robot utterances. As this study consider that repair is part of communication, robots should not entirely avoid repair.

It is known that repair does not always happen just because there is a problem of understanding [30]. A human may decide that there is no problem even if the robot's utterance includes an understanding problem. Also, there is no guarantee that a person will actively participate in the process as an attitude to participation. Using incompleteness has this problem of indeterminacy. However, voluntary participation is also an essential aspect of this "repair" as communication for systems. Free participation has the potential to measure one's attitude toward participation in the conversation. Therefore, it is valuable to investigate

[1] "Seaman (video game)," accessed December 20, 2020, https://en.wikipedia.org/wiki/Seaman_(video_game)

the effects of conversations with incompleteness between systems and people from the aspect of building communication.

4.3 Purpose

In conversations between people and systems, understanding problems have been thought of as an obstructive factor [52]. On the other hand, today's systems are always having low-context conversations that do not create understanding problems. It can be said that this is a situation that removes any room for human intervention. Based on practices of people's conversations, this study supposes that it would be a worth opportunity to elicit questions and participation from the participants in conversations.

Suppose the robots can have a conversation using incompleteness with human assistance. In that case, it will become a new possibility to have conversations in other ways than having the system speak firmly. However, until now, no research has been conducted to reduce the robot's ability to speak for generating incompleteness. This section describes the first investigation of how "Incomplete Utterance" changes people's behaviors and responses. This experiment was conducted with two different conditions of the Incomplete Utterance strategy and investigated the human responses and behaviors.

4.4 Procedure

This study compares two different conditions. The aim is to verify the selected incompleteness that can be used in conversation to elicit participation. For that purpose, the one-to-one conversation was used. It is the most general situation when humans have a conversation with a system.

4.4.1 Conditions and Resources

Since there is no conventional way to make a robot's utterance incomplete, this study will compare two possible incompleteness. This experiment set up two Incomplete Utterance conditions. The semantic incompleteness condition is the first one. This condition is the same as the utterance strategy described in Chapter 3. The reasons for calling it Semantic incomplete are that it takes the primary meaning out of a sentence and lacks the explicit sentence content. Next, the second condition is structural incompleteness. This condition is generated by pausing and stopping in the middle of the sentence. It is named this way because it is an utterance before it is completed as a sentence structure.

To explain the two conditions, firstly, it is necessary to look at the resource that is the robot's speech source. This experiment prepared six articles from a news site called NEWS WEB EASY [1]. The contents of each news item are shown in Table 4.1. This news site is written based on "Yasashii Nihongo(Easy Japanese) [53, 54]." The syntax is simple, and the subjects and predicates are exact; It is written to avoid complex expressions. In Japanese conversation, it is known that the written language tends to use more complex vocabulary than the spoken language and that sentences tend to be longer [43, 44]. However, the sentences are written in "Yasashii Nihongo" excludes complexity from the written word. Its features can be used for reducing the process of making the spoken word.

Semantically Incomplete Utterance (SemICU)

The semantically Incomplete Utterance is a condition for considering a semantic deficiency such as undefined references and lack of information. After analyzing the original sentences' structure, segments of sentences are merged with the original sentence's verb. Chapter 3 illustrated how this study analyzes the sentence and separates the sentence into segments. This condition creates two sub-sentences from one original sentence (Fig. 4.1).

[1] "NEWS WEB EASY," accessed December 20, 2020, https://www3.nhk.or.jp/news/easy/

Table 4.1: Six News Article

No.	Content
1	Toyota corp. and Matsuda corp. develop together
2	The richest person in the world
3	The production quantity of wine in French
4	Panda in Ueno Zoo.
5	Jet aircraft by Honda
6	Total solar eclipse in the U.S.

The first sub-sentence reduced the segments connected to verbs and mainly consists of segments containing the subject and the verb. Because the first sub-sentence drops the part of the content, it has some vague references and a lack of information. Then the second sub-sentence consists of the remaining part of the original sentence and the verb. For example, Figure 4.1 shows the first sub-sentence is (a) Segment 1 + Verb phrase, second sub-sentence is (b) Segments2, 3 + Verb phrase (Fig. 4.1).

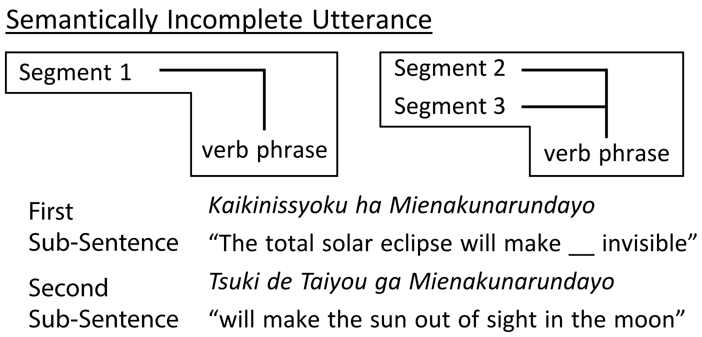


Figure 4.1: Segmentation for Semantically Incomplete Utterance

Depending on the sentence analysis results, there may be no segment that is connecting to the verb. If the sentence structure does not resemble Figure 4.1, extracting the sentence's primary meaning will not be simple. For example, if subjects are parallel, it is not simple to determine which is the main subject of a sentence. For such sentences, this experiment applied the method of Structurally Incomplete Utterance described below.

Structurally Incomplete Utterance (StrICU)

This experiment also defined Structurally Incomplete Utterances (StrICU). The StrICU is a condition for considering a structural deficiency such as stopping speech prematurely. StrICU divides an original sentence simply into two sub-sentences, the first sub-sentence is (a) Segment 1 + Segment 2, and the second sub-sentence is (b) Segment 3 + Verb phrase (Fig. 4.2). By the StrICU method, the first sub-sentence is seen as interrupted in the middle of the original sentence. And the second sub-sentence restart is seen as to speak the sentence. This phenomenon is one of the problems related to "repair" practices. It is known that if the speaker stutters in mid-speech or has a problem in speech production, listeners participate in the conversation as they offer help [55, 31].

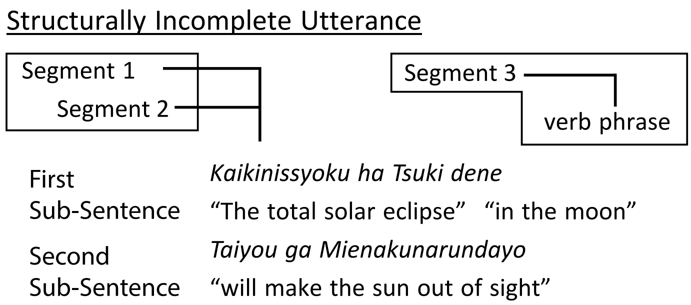


Figure 4.2: Segmentation for Structurally Incomplete Utterance

Adding colloquial expression (Common to Both Condition)

The process of segmentation only divided one original sentence into two sub-sentences. Because news articles are not written in a spoken language style, the sub-sentences are needed to some components to change the style. Therefore, the end particle of the utterance was changed. The "ne" particle and "yo" particle were added. The Japanese particles are seen in spoken language style. The particles act as a sign to invite the conversation partner [56]. Moreover, the filler “Anone” or “Ettone” was added before the first sub-sentence. Fillers do not contain the meaning themselves. However, Mizukami and Yamashita [32] reported that a filler has the function of maintaining the speaker’s right to speak in a conversation.

4.4.2 Robot and Experiment Field

4.4.2.1 Platform “Muu”

This study used the platform “Muu,” a robot developed at the Toyohashi University of Technology. Muu is used as a platform for social interfaces based on multi-party conversations. Its appearance is shown in figure 4.3.

Figure 4.3: The Figure of Platform Muu

It is designed based on a minimal design. A design that does not evoke a person, a dog, has no metaphor to estimate the robot’s ability. While animals and humans have two eyes, Muu has one. This design aims to reduce the factors caused by the adaptation gap [57]. The placement of the body and eyes is designed based on the baby schema. Moreover, the upper part of the body is designed to make it easier for humans to perceive the Muu’s direction.

This kind of robot has the advantage that it can be easily applied to other robots due to its small number of parts and simple mechanism. Also, since the simple construction, the number of contact areas with humans can be reduced. Robots with a single eye attached to a hemispherical body have recently been used in human-computer interaction for children, as in the work of Zaga et al. [58].

Hardware

The Muu consists of the main body and an upper part. The upper part is connected to the soft urethane outer cover. The main body has the PC and other processing devices.

As Figure 4.4, The upper part has an acrylic dome representing Muu's eyes. A web camera is installed inside it. Two servo motors are used to connect the upper part to the main body. The motors connect the upper part to the main body. As shown in the figure, the head moves in two axes, yaw and pitch. Therefore, Muu can express nodding or shaking its head left and right. The servo motor in the yaw direction is connected to the main body via a spring. When Muu moves his head, the movement is not mechanical and stable, but soft.

The main body contains the PC (GIGABYTE BRIX) and the circuitry and wiring to run the servo motors. The speaker is placed in the center of the main body. When Muu makes a sound, the sound can be heard from inside the body. The PC is capable of processing speech recognition, speech synthesis, and face recognition.

Software

Muu uses RealSense SDK [1]. by processing the video images acquired by Muu's upper part webcam, Muu can recognize the person in front of it. Also, by combining with the servo motor control, it is possible to direct Muu's gaze to the person. Furthermore, parameters such as the direction of the person's face can be used. For speech recognition, JSGF format recognition and free dictionary recognition

[1] Intel(R) RealSense Technology, accessed December 20, 2020, https://www.intelrealsense.com/.

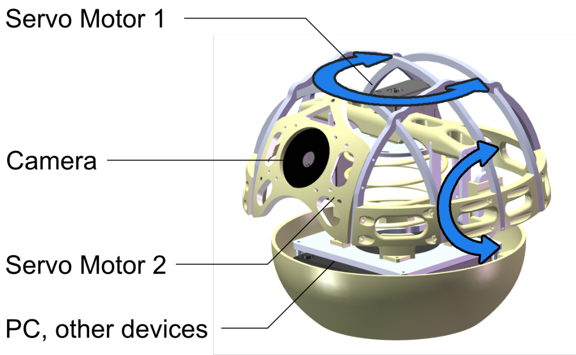


Figure 4.4: The Figure of Hardware of Muu

were used.

Muu can speak in Japanese. Muu uses the Wizard Voice SDK, a speech synthesizer software developed by ATR-Promotions. The software synthesizes a speech base on the voice of a five-year-old child in the Kansai region. That achieves to generate a voice close to that of a child.

For this experiment, the robot was programmed to utter the same sentences every time. It produces utterances every three seconds if the participants do not speak (e.g., turn pass). If participants continue to speak, the robot waits to speak its next utterance.

4.4.2.2 Experiment Field

This experiment used one of the platforms, Muu. The experiment room layout is as shown in Figure 4.5. When a participant interacts with the robot, participants could see a camera on their right side.

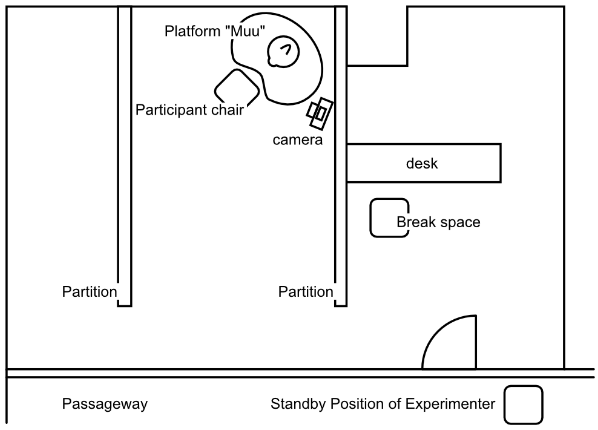


Figure 4.5: Experiment Field

4.4.3 Process

Participants who consented to participation in this study interacted with two types of robots, the conditions using the SemICU and the StrICU. One experiment included six interactions. The participants interacted three times with the SemICU robot and three times with the StrICU robot.

To balance topics and speech conditions, half of the participants interacted with the SemICU robot in the first, third, and fifth interactions. The other half interacted with the SemICU robot in the second, fourth, sixth interactions.

One interaction follows. (a) Participants come in front of the robot. (b) The experimenter explains before each interaction, "Please try to have a conversation with the robot while the robot talks to you." (c) The experimenter leaves the room. (d) After the robot has started to speak, they interacted around for 2-3 minutes, depending on the conversation topic. (e) The interaction ends with a signal by the experimenter. (f) In order to change the condition of the robot, the participant leaves the robot's presence. (g) Return to (a).

In the end, the participants were asked if they were interested in the six topics that the robot had spoken.

4.4.4 Analysis

This experiment explored how participants interacted with the robots and participated in the interaction. The categorization of human responses was based on Linnell's IR analysis [59, 25]. Table 4.2 shows all categories of human response in this research. Original IR analysis used categories 1, 2, 5, 6, 7, 8, and this study added categories 3, 4, 5, 10 considering the previous study using the platform Muu [60]. By using these categories, the participant's speeches were annotated.

The IR analysis clarifies how participants share, experience, and execute conversation together within a discussion on the conversational co-construction axis. Linell focused on whether each of the other speakers participated in the conversation [59]. This study defines cooperativeness as the participants' cooperative actions to complete the meaning of the robot's speech. In a situation where the robot is talking to the human and taking the initiative, this experiment discussed how much participants support a robot and increase conversation participation.

The annotation of human responses was annotated by two researchers, the author and one student in the same laboratory, using the annotation tool ELAN.

Table 4.2: Human Response Category

ID	Category	Sample Response (*ja* "en")
1	Turn pass	""
2	Minimum Response	*un* "ah"
3	Repeat partial Robot Utterances	*wain* "wine"
4	Encourage Robot Speech	*sorede?* "and?"
5	New information about topics	*shiranakatta* "I don't know it"
6	Question about topics	*dokode?* "to where?"
7	New information about not related topics	*Kyou ha tenki ga iine.* "This weather is good."
8	Question about not related topics	*Supotsu suki?* "Do you like sports?"
9	Present unhearable	*Nante itta?* "What did you say?"
10	Laugh	""

4.4.5 Evaluation Scope

Since a robot starts the conversation, the human response is the main target of analysis. In order to compare between SemICU and StrICU, the scope of the assessment was set in the interaction. As explained in Section 4.4.1, the transformed sentences are limited in this study. The scope of analysis is the part of original sentences utilized by both SemICU and SemICU. This study analyzed the human responses and behaviors to the first sub-sentences and second sub-sentences in both conditions.

4.5 Result

4.5.1 Participants

A total of 16 participants participated in this experiment. However, three participants were excluded from the analysis for homogeneity. One male and one female had a deficient number of utterances and continued silence, making the annotation impossible. They are excluded before undertaking detailed annotation. Also, the one female excluded participant spoke significantly more than other participants. The participant had a few Turn pass (category 1) and a minimum response (category 2); moreover, she overreacted to the robot. This participant's average interaction time was 215 seconds, while the overall average was 144 seconds. The interaction time of this participant was much longer and exceeded an upper inner fence. Thus, a total of 13 participants (8 males and five females, mean age 27.7, standard deviation 11.2) are analyzed.

4.5.2 Result 1: Response Category

Including all the participants, the number of human verbal responses are 205 times to the first sub-sentences (sub-s) of SemICU, 170 times to the first sub-s of StrICU, 210 times to the second sub-s of SemICU, and 193 times to the second sub-s of StrICU.

Figures 4.6, 4.7 show the result of the annotations of 13 participants. The vertical axis represents the average number of times the human responded. The horizontal axis represents the human response categories. Because the target scopes were defined, if participants increase one category response (e.g., Question), other categories' responses decrease (e.g., Minimum response).

Figure 4.6 depicts the result of a Wilcoxon's signed rank test between SemICU and StrICU. If participants interacted with the SemICU robot, they decreased

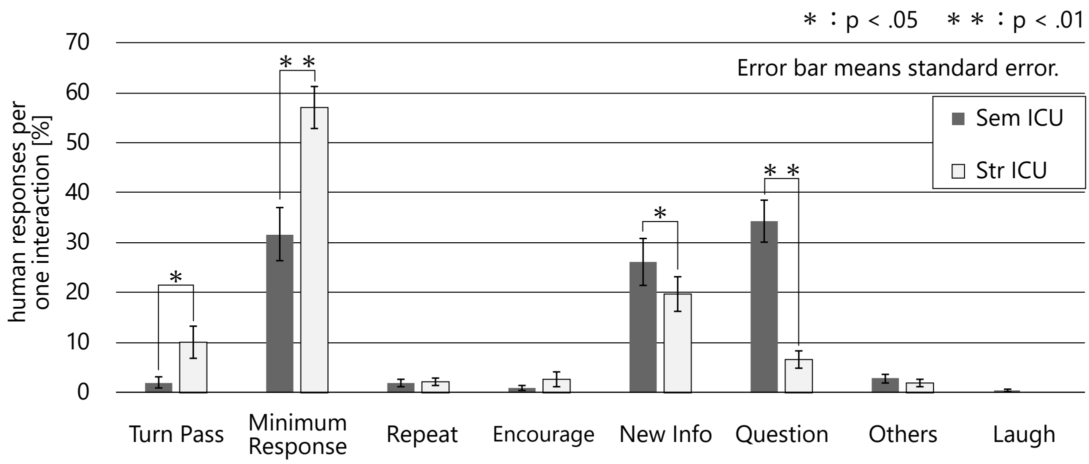


Figure 4.6: The Average Number of Human Response Time Compared to Semantically ICU and Structurally ICU

turn pass (t(12) = 1.5, p = .015) and minimum responses (t(12)= 0, p = .002), moreover increased new info (t(12) = 14, p = .030) and questions (t(12) = 0, p = .002).

Because the sentence was split into two parts in the process of speech generation, there are two types of human responses: responses directed to the first sub-sentence and responses directed to utterance second sub-sentence.

	First sub-sentence	Second sub-sentence
SemICU	A	B
StrICU	C	D

Table 4.3: Four Cases into Two Robot Speech Conditions and Two Utterance Timings

Additionally, four response categories were analyzed to find out more about items that showed significant differences. The results of Wilcoxon's signed-rank test using Bonferroni's adjustment method are shown in Figure 4.7. The symbols "A, B, C, D" in figure 4.7 are the combination of the two conditions and two

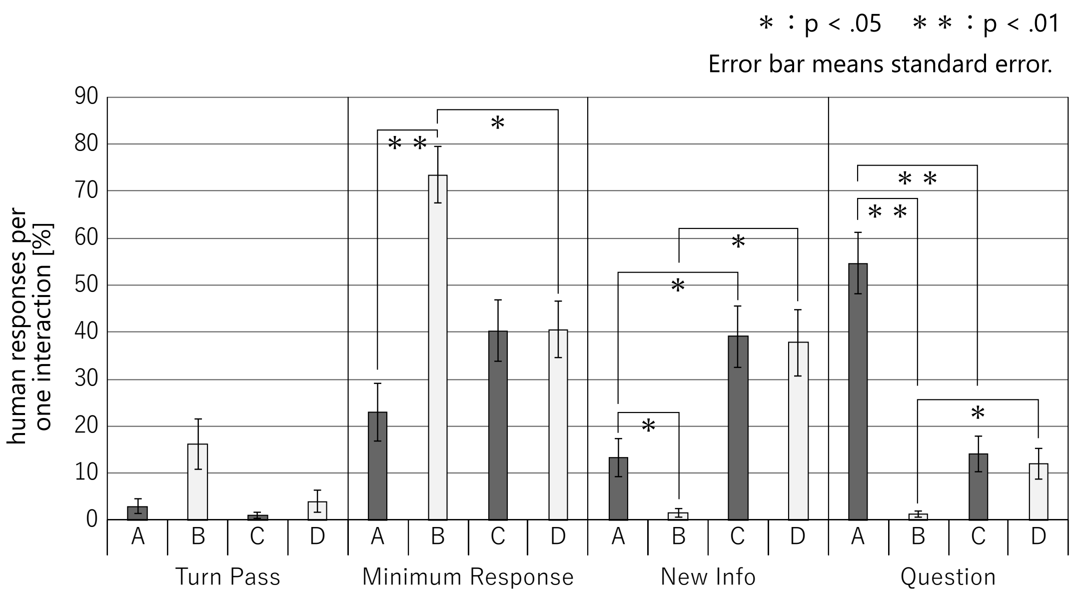


Figure 4.7: Average Speech Frequency Organized into Two Robot Speech Conditions and Two Utterance Timings

sub-sentences. The detail is shown in Table 4.3. Figure 4.7 shows the percentage of human responses addressed to each A, B, C, and D utterances.

The results showed that the B's minimum response was significantly higher than the responses to the other utterances. Furthermore, for category 5, "adding information," the responses increased in the order B $<$ A $<$ C, D. And the questions directed to the A utterance were significantly more than the questions directed to the other utterances.

4.5.3 Result 2: Response Time

Next, this study analyzed the response time. It is the time between the end of the robot's speech and the beginning of the human's speech. Figure 4.8 shows the result. The graph depicts the start time of the human responses. When the human speech overlaps the robot's speech, it is negative.

The response time was tested with Friedman's test between four conditions: the robot's utterance condition and the first sub-sentence and second sub-sentence. The response turns with no speech were considered as missing values. Therefore, the number of evaluation data was 166 for A, B, C, and D. The test results were $\chi\ 2(3) = 45.874$, $p = .000$, indicating a significant difference between the four conditions. As a post hoc test, Wilcoxon's signed-rank test was conducted for each condition using the Bonferroni method to adjust the confidence intervals. As a result, a significant difference was detected between A-B, A-C, and C-D with $p < .000$.

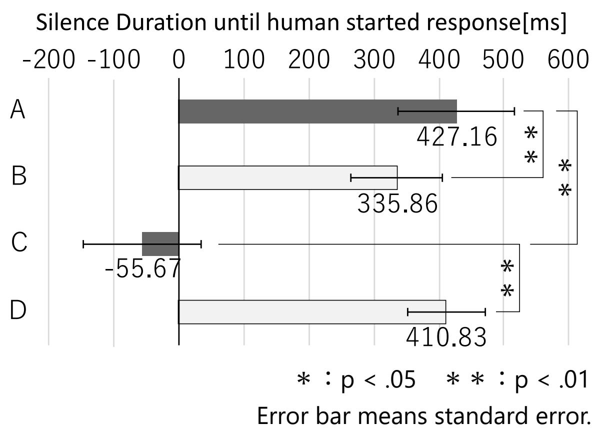


Figure 4.8: The Silence Duration until Human Started to Respond. Zero line means the end of robot utterances. The number next to the stick is the time of silence until speaking.

4.6 Discussion

As Figure 4.6 shows, the human response to Semantically Incomplete Utterance increased the amount of information added and questions. Semantically Incomplete Utterances increased the number of responses from people with a higher initiative character. The categories indicate that the semantically incomplete condition allowed people and robots to construct a more balanced conversation than the Structurally Incomplete condition.

Semantically Incomplete Utterance increased the questions. The results in Figure 4.7 show that significantly more questions were asked after utterance A (the first sub-sentence of Semantically Incomplete Utterance). Most of the questions asked by people were asked in the form of “interrogative word” + “particle.” According to Suzuki’s study, this indicated that participants wanted the robot to repair the problem because any elements were not present in the preceding robot’s utterance [30].

Figure 4.7 showed no difference in the number of questions found in utterances C (the second sub-sentence of SemICU) and D (the second sub-sentence of StrICU). There are no differences between the second sub-sentences. The reason for this may be in the news that originated the story. It is expected that there will be no excesses or deficiencies at the end of every sentence in the news. After this result, additional analysis was done. There were 32 places where people questioned in utterance D. And out of those 32 places, 29 places were also questioned in utterance C. As a supporting result, additional analysis showed that Ninety percent of the second sub-sentence questions were asked in the same place in utterances C and D.

The increase in “Adding Information” in the Semantically Incomplete Utterance condition caused by the first sub-sentences had verbs. In the Structurally Incomplete Utterance condition, the robot expresses incompleteness by interrupting an utterance in mid-sentence. People perceived it as the robot’s continuation of the utterance and did not actively engage in it.

As evidence, Figure 4.7 shows that in first sub-sentence in the Structurally Incomplete Utterance condition, the number of “minimal response” utterances increased, and people tried to make the robot continue its utterance. It is seen that an utterance is interrupted in conversations between people; another person takes over the utterance and completes it. However, in conversations between humans and robots, "taking over" was not observed simply by interrupting an utterance.

The response time results in Figure 4.8 show that the response time to first subsentence in the Semantically Incomplete Utterance condition is longer. The reason for the long response to utterance A may be the preference for "other-initiated repair," as pointed out by Schegloff [26]. According to Schegloff, speakers tend to repair the missing parts by themselves (self-initiated repair) before the listener asks them. However, the robot was not able to repair itself like a human. So, the participants initiated responses when they recognized the end of the robot's utterances and that it would not repair itself. It could be the reason why there was a delay in response to utterance A.

Besides, in response to utterance C (Semantically Incomplete Utterance 2), the person spoke in a way that overlapped the robot's utterance. The human's characteristic behavior is to make a response that indicates understanding during the preceding robot's speech. This overlapping suggests that the robot may have answered the person's Question in utterance C. The person was expressing that the Question had been answered when the word containing the answer was presented.

Study 1 was suggested to increase active participation from people to include predicates and end them as turns while lacking content semantically.

5 Incomplete Utterance in One-to-One Conversations (2)

Study 1 examined how Incomplete Utterance affects human responses. Since the previous study focused on analyzing the responses and behaviors, it did not have questionnaires for asking the participants. Although people increased their assistance to the robot by Incomplete Utterance, it left with a discussion about their impressions. Therefore, study 2 conducted a third-person evaluation of the robot's conversation and the participant recorded in study 1. New participants watched the video and answered their impressions of the conversation. The answers were included the attitude of the participants and whether the conversation was successful. This chapter is composed of part of Nishiwaki's previous paper [61].

5.1 Third-Person Evaluation of Conversation

Ogawa conducted studies in which observers evaluate conversations [62, 63]. Ogawa described the advantage of an observer's evaluation as a smaller cognitive load than that of a speaker. Participants can evaluate conversations calmly. Also, the third-person evaluation worth for robots and systems development. Because robots and systems have become more common today, every human can see someone uses the systems. It means that anyone could be seen using the tool by someone else. Therefore, it is crucial to know how a third person will view a conversation between a person and a robot. However, it is also known that an observer and a speaker's information processing are different due to cognitive load difference [64]. While third-person evaluation has merit for considering social applicability, it is also necessary to understand its limitations.

5.2 Purpose

The conversation in study 1 was the conversation that the robot spoke as the speaker and humans participated as the listener. In this situation, the robot started to talk every time. It remained whether the interaction was evaluated as a conversation even if humans responded to the robots. The previous study results also found that people's responses to semantically Incomplete Utterance were significantly different from people's responses to Structurally Incomplete Utterance. However, people still participated in conversations with robots that used Structurally Incomplete Utterance by using minimal responses. They also made remarks about what they knew about the topics and how they felt about them. Thus, the conversation with the StrICU robot could also be evaluated as a successful conversation even with few questions. From those concerns, the purpose here was to verify whether the observer felt the conversation established the conversation and how the speaker's participation attitude and the robot are evaluated by third-person.

5.3 Procedure

5.3.1 Condition and Video

This experiment used some of the recorded videos of study 1. The participants in that video have consented to the use of their videos in the experiment. The conditions are one with Semantically Incomplete Utterance and the other with

Structurally Incomplete Utterance.

This section explains how the videos are chosen to be shown to new participants. As for the video to be shown, it needs to be the average conversation from study 1. Therefore, the ten response categories and their ratio were attention. The videos were selected primarily by identifying participants who were having average conversations. Due to select the average human-robot conversation and its participant, the following three factors were used.

First, the ten categories of human responses were used in previous study 1. The categories and example responses were shown in Table 4.2. The previous study counted the frequency of categories for every participant. This study generated feature vectors from the frequencies. And then, the average feature vectors of both conditions were calculated. The conversation closest to the average was defined as being close to the two feature vectors of Semantically Incomplete Utterance conversation and Structurally Incomplete Utterance conversation. Therefore, the cosine similarity between the feature vectors and each participant's conversation was calculated.

Second, It is about the exclusion of conversations. In some conversations, the exact same questions were asked several times in multiple turns. There were also some behaviors in which people could not catch the robot's speech. In such cases where the nature of the speech was likely to be different from the speech in the category, it was considered inappropriate for video-based evaluation.

Third, Considering the situation in which the robot is used, if the participant is not interested in a topic, it may be skipped. Therefore, this video-based study selected conversations in which the participants were interested.

As the result of three factors, two conversations of a 23-year-old female participant were selected. The Semantically Incomplete Utterance conversation was about "the richest person in the world" with a cos similarity of 0.91, and the Structurally Incomplete Utterance conversation was about "Honda jet airplane" with a cos similarity of 0.97. The full conversations are shown in the following transcript. The transcript symbols for describing behaviors are some of the symbols used in conversation analysis [23]. The details of the symbols are also shown in the appendix.

Finally, as a video to be used in the experiment, the videos were anonymized. The videos were applying a mosaic so that the person speaking in the video could not be identified.

```
(1) Conversation in the video (Semantically Incomplete Utterance)
The first line is the alphabetical representation of the original
Japanese text. (X: ) means who is speaking.
The second line is the translation by the author
R:    untone
      so...
      (1.1)
H:    un(0.7)dou shita(3.0)h
      yeah(0.7)What happened(3.0)h
      (0.3)
R:    anone(0.3)america no zasshi no forbusu wa happyo shitan dayo
      well,(0.3)Forbes, an American magazine, announced
H:    n↑wani o?
      n↑What announced?
      (2.0)
R:    sekai no okanemochi ni [tsui]te happyo shitanda[yo
      It announced about the richest people in the world.
H:                           [un  ]                 [.hh(0.4)kyoumi ga
      arune sore wa
                              Ah                    I am interested in
      that.
      (4.0)
R:    untone(0.3)Jefu Bezosu san ga natta soudayo
      so... I heard that Jeff Bezos became one.
H:    he:::(1.4)dorekurai okanemochi nano kana?
      I see.    How rich is he?
      (1.7)
R:    sekai de ichiban no okanemochi ni natta soudayo
      I hear he's become the richest man in the world.
      (0.4)
H:    un:
      yes.
      (1.8)
R:    unto ne(0.3)bezosu san wa tsukuttan dayo
      well        Bezos made
      (0.5)
H:    wani o?
      What did he made?
      (1,8)
R:    amazon dotto comu o [tsukuttan dayo
      He created Amazon.com
H:                        [hh .hh
H:    watashi mo itsumo riyou shiteiru yo(1.1).hh
      I use it regularly too.
      (2.5)
R:    Bezosu san wa 901 okudoru motteiru n datte
      Mr. Bezos has $90.1 billion.
H:    hh(0.8).hh sugoi ne(0.2)issyo asonde kuraseru ne(0,9).hh
                 Unbelievable. He can play for the rest of his life.
      (1.1)
```

R: ano ne(0.3)Bezosu san wa hazimete natta n dayo
you know. Mr. Bezos has never become before.
H: n? nani ni: natta no?
uh? What did he become?
(3.6)
R: sekai de ichiban no okanemochi ni nattan dayo
He's now the richest man in the world.
H: un(.)sokka
yes.(.)That's right.
(3.8)
M: untone(0.3)amazon wa inta-netto de utte iru yo
so(0.3) Amazon sells on the Internet.
H: un
yes.
(2.1)
R: ironna mono o utte iruyo
They sell all kinds of stuff.
(0.,3)
H: souda ne zibun mo iroiro katte iruyo:
Yeah, I'm buying a lot of stuff too:
(3.1)
R: e:tto ne(0.4)bezosu san wa(0.2)motte iru yo
Um, yeah.   Mr. Bezos has.
(0.4)
H: un:(1.4)wani o motte ru?
Okay. What does he get?
(3.0)
R: amazon no kabu o 17% [kurai motte iru yo
He owns about 17% of Amazon's stock.
H: [a(h).hh
H: zibun de(h)motte run dane(0.2).hh
He's got it with him.
(1,3)
R: ima made ichiban datta nowa ne
The person who's always been number one.
H: un:
yes.
(2.8)
R: biru geitsu san dattan [dayo
It was Mr. Bill Gates.
H: [a
H: un shitteru yo
Yeah, I know.
(1.7)
R: geitsu san wa(0.25)900 oku doru [motte irun dayo
Mr. Gates has $90 billion.
H: [hh
H: .hh sukoshi gurai wakete hoshine
I hope he will share a little.

(2) Conversation in the video (Structurally Incomplete Utterance)
The first line is the alphabetical representation of the original Japanese text. (X:) means who is speaking.
The second line is the translation by the author

R: untone
so...
(0.6)
H: un
yes.
(1.8)
R: e:to ne(0.4)jidousya no honda wane
uh the Honda of the automotive company.
H: un
yes.
(2.9)
R: kogatahikouki o tsukutte irun dayo
build a small plane.
H: a(.)sounanda(.)jidousya dake zya naindane
Oh, really? So it’s not just automobiles.
(3.3)
R: untone(0.5)honda no hikouki wane
well Honda’s airplanes are
H: un
yes.
(2.8)
R: nana nin made norerun dayo
It can carry up to seven people.
H: e::sugoine(.)noritai ne issyo ni hh
Wow, that’s great. I want to ride with you.
(2.3)
R: honda ni yoruto ne
According to Honda.
H: un
yeah
(2.3)
R: kotoshi no rokugatsu made ni ne
By June of this year.
(2.7)
H: un
Mm.
R: anone(0.3)honda no hikouki wa okyaku san nine
You know what? Honda’s airplanes are for customers.
H: un
uh-huh
(2.3)
R: nijuyon ki todoke raretan datte
delivered to 24 customers.
H: e(0.2)sou nanda h
Oh, yeah.
(3.9)

R: america ya ro-roppa dewane
In the US and Europe.
H: un
uh
(2.1)
R: kogata hikouki o ne
A small plane.
H: un
yeah.
(2.4)
R: kau kaisya ga fuete irundayo
More and more customers are buying them.
H: he:::(0.2)minna okane mochi dane hhh(0.3).hh
Wow, they're all rich.
(1.9)
R: untone(0.3)honda wa sukunai gasu dene
Honda uses less gas.
H: un:
yes.
(2.7)
R: tobu koto ga dekirun dayo
It can fly.
H: ↑un(0.9)sokka
Yeah. I see.
(2.7)
R: e:ttone(0.4)ima wa ichinen ni ne
Now in a year.
H: un
yes.
(2.6)
R: gozyu tsukutte rundayo
They're making 50.
(0.6)
H: ha::(0.5)zyaa tyotto sukunai n [dane
So it's a little less then.
R: [kore kara hachizyu gura[i]ni
fuyasun dayo
They're going to increase it to
about 80 now.
H: [un]
uh-huh
H: un(.)sokka
I see.

Table 5.1: Questionnaire of Study 2

Questions about the mood of the conversation between humans and robots

1	They had a conversation.	Formation of Conversation
2	The conversation was natural.	
3	The conversation was cooperative.	Commonality
4	The conversation was one-sided.	
5	Human and robot, trying to talk together.	

Attitudes of people and robots toward participating in conversations

6	The person was talking to a robot.	Attitude of Participation
7	The person was listening to the robot.	
8	Person was actively trying to participate in the conversation.	
9	The person was trying to get a story out of the robot.	Assistance
10	The person was trying to maintain a conversation.	
11	The robot seemed to want the person to listen to it.	

5.3.2 Impression Evaluation

The eleven questions were prepared to evaluate impressions of the conversations. The questionnaires used a 5-point scale: 5: Agree, 4: Slightly agree, 3: Neither agree nor disagree, 2: Not really agree, 1: Disagree. The sentences of questionnaire items are shown in Table 5.1. The questions were designed to evaluate whether the conversation was successful, the conversation's collaboration, and the attitude of participation in the conversation. In the experiment, the order of these questions was presented randomly.

5.3.3 Process

This experiment was conducted as a video-based evaluation. The participants were asked to watch two videos. While watching the video, the participants were asked to wear headphones or earphones to listen to the conversation correctly. This experiment also includes a step to adjust the volume. The participants adjusted the volume to hear the people and robots' voices in the video and then watched the condition video. After watching the video, the participants were asked to answer the questions shown in the previous section.

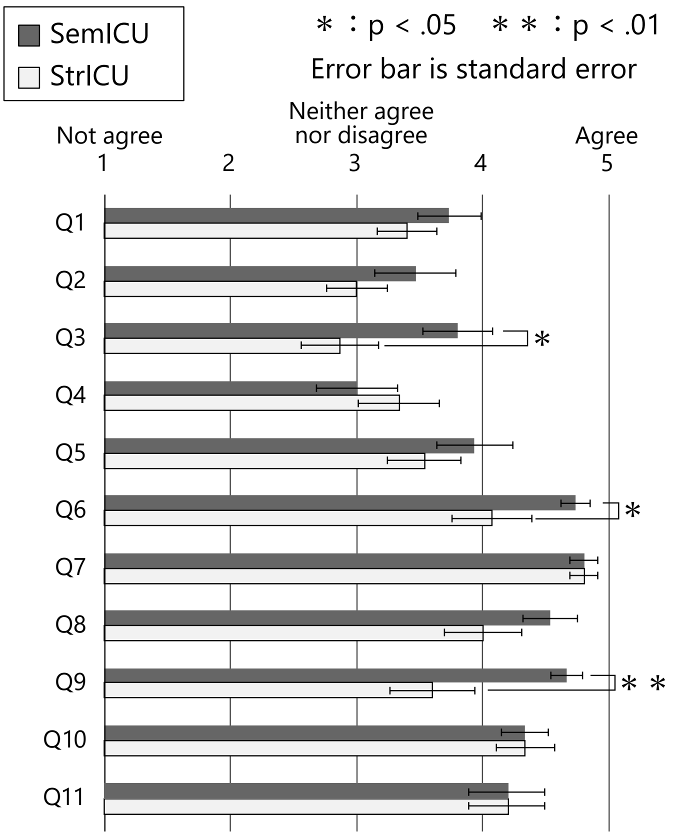
5.3.4 Participants

Fifteen participants (mean age: 28.4 years, standard deviation: 8.5 years), seven males and eight females, participated in the video evaluation experiment.

5.4 Result

5.4.1 Questionnaire

The results of the Wilcoxon's signed-rank test for each question are shown in Figure 5.1. The result shows that Q6(The person was talking to the robot) and Q9(The person was trying to get a story out of the robot) were scored significantly higher in the conversations in the Semantically Incomplete Utterance condition.



	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
mean(Up)	3.73	3.47	3.80	3.00	3.93	4.73	4.80	4.53	4.67	4.33	4.20
mean(Low)	3.40	3.00	2.87	3.33	3.53	4.07	4.80	4.00	3.60	4.33	4.20
T(14)	12	32	10	15	14	0	1.5	3	4.5	18	10.5
p-value	.211	.193	.018	.401	.333	.019	1	.066	.006	1	1

Figure 5.1: The Result of Third-Person Evaluation for the Videos (SemICU = Semantically Incomplete Utterance, StrICU = Structurally Incomplete Utterance)

5.4.2 Additional Comments

In addition to answering the questions, this study prepared free descriptions asking why they answered as they did. The following lists show the free descriptions of the reasons why they answered as they chose.

5.4.2.1 Conversation of Semantically Incomplete Utterance

Summary 1: The robot did not respond to people's responses. The robot did not seem to be answering people's questions.

It did not seem like the robot responded to what people said because it did not respond to the person's talking.
The robot did not seem to be responding to people.
Did the robot respond to the person's questions in a few parts?

Summary 2: The human was responding to the robot's speech. The conversation was made possible by the coordination of the human.

I felt that the conversation continued as the person responded to the robot's speech. However, I also felt that the conversation continued similarly.
Firstly, the robot showed the results, the person was drawing out the content, and the robot answered.
I answered that I thought the conversation was a success a little bit because the robot seemed to be answering the questions that people asked. However, I also got the impression that the human chose questions that the robot could easily answer. For these reasons, I chose "Neither agree nor disagree" for the question of whether the conversation was natural.

Summary 3: It was a conversation. The robot was answering people's questions appropriately.

I think it was a conversation.
The robot was answering questions from people appropriately.
The conversation looked like an adult listening to a preschooler.

Others

Because people's emotional expressions were natural, I felt that people were actively trying to participate in the conversation, that the conversation was cooperative, and that the conversation was conducted naturally.
The robot side always presented the topic, so it felt a little one-sided.
It seemed less one-sided than before.
I got the impression that the conversation was more natural than the last time between a robot and a person. The person was asked to answer the "wo (Part of Japanese particle)" part of the conversation.

5.4.2.2 Conversation of Structurally Incomplete Utterance

Summary 1: Conversation with the child

It sounded like she was talking to a small child.
Conversations from people sounded like they were talking to a child.

Summary 2: No response from the robot

I felt uncomfortable about the conversation because the robot did not ask any questions or give any feedback to the human.
I got the impression that the human was responding to the robot that was speaking the predetermined content.

Summary 3: Discomfort in the robot's response

There were some parts where the robot's response was not quite right.
The flow of the conversation was a little uncomfortable in some parts.

Others

It was not easy to hear proper nouns in the robot's speech.
The gap between the robot's speech and the concrete description was fascinating.
The conversations between people and robots were natural, with almost no pauses or interruptions. The tempo was as if an infant and an adult were talking.
I felt a little uncomfortable that the robot sounded like a child.
The way the robot talked made me interested in the content of the conversation.

5.5 Discussion

The third-person evaluation showed the result was compatible with the results of Study 1. It was confirmed that more questions and information were added in the "Semantically Incomplete Utterance" condition. The third-person evaluation also evaluated the fact that people were trying to talk to the robot and get a story out of the robot.

Why Semantically Incomplete Utterance scored significantly higher for Q3, "The conversation was cooperative"? According to the free descriptions, there was a difference in whether the person made adjustments or not. In Structurally Incomplete Utterance, the person was only able to give minimal responses. However, in Semantically Incomplete Utterance, the person could ask questions, and she did. So the conversation may have been rated as cooperative because the person worked to move the conversation forward.

This experiment discovered that cooperation could have been created by human participation. However, the evaluation of the actual speaker's impressions was remaining. The next experiment will explore the subjective evaluation and how to utilize robots other than one-on-one conversations.

6 Incomplete Utterance in Multi-Party Conversations

So far, this paper has investigated one-to-one conversations between a person and a robot. The previous studies remained the actual participant's evaluation. In considering the interactor's evaluations, it is necessary to resolve the limitations of one-on-one conversations. In the one-to-one conversation, people need to be a listener while the robot is talking. The people can speak among the times the robot was not talking. If it is viewed from the opposite side, the one-to-one conversation forced people to talk in the time the robot was no talking. The one-to-one conversation consisted of only the robot and one participant. If either of them were to stop talking, the conversation was broken. People were not able to choose the role to participate in their conversations. Therefore, the other conversational situation was needed to investigate people's participation attitudes and methodology further. This study was designed to use multi-party conversations. Based on Nishiwaki's paper [65], this chapter describes the results of an analysis of the effects of Incomplete Utterance and people's attitudes in a three-party conversation.

6.1 Multi-Party Conversation

This study will first discuss the usefulness of multi-party conversations.

6.1.1 Multi-Party Conversation between People

When spouses or couples participate in a conversation, they may share the episodes told by either one. Lerner and Mandelbaum found the typical conversations with a spouse or couple inside [66, 67]. They found that the partners do not participate in the conversation as just listeners while either telling a story. They make statements confirming what is correct about the episode and adding related episodes. Partners who know about the same event conduct a conversation with each other.

Den also reports on the composition of a three-party conversation [68]. In contrast to the one-to-one speaker-listener configuration, in a three-party conversation, one more person joins in it. Therefore, another person who is not a speaker or listener can participate in the speaker's explanation. Also, two people can listen together e in the conversation, and the other person can choose not to participate in the conversation as bystanders. Compared to one-on-one conversations, multi-party conversations offer various ways to participate and allow multiple people to form a single role.

6.1.2 Multi-party Conversations with Systems

Active Listening is the most effective way to elicit people's participation as speakers. Shimooka et al. investigated Active Listening using a comprehensive method and described the limitations of speech recognition accuracy [69]. In recent years, there has been much focus on multi-party conversations to avoid technical limitations and allow people to participate in conversations.

For example, Matsuyama et al. studied using a facilitation robot to help switching speakers between participants [70]. Iio et al. developed a method of preventing the collapse of conversations caused by speech recognition errors by preparing two robots as listeners [71]. These studies considered a single robot taking the role of a speaker and listener without the need for help from others or multiple robots performing a single role. The previous section mentioned that people who have the same knowledge could collaborate to be speakers and listeners. This phenomenon will be one of the methods to elicit people's participation in the conversation. However, such acts between robots and people have not yet been studied, and many works remain.

This study focused on the interaction “human and robot explaining together” instead of robot explaining alone. This study investigated the human's participation attitudes using Incomplete Utterance, taking into account their interactions with others in a multi-party conversation.

6.2 Purpose

Studies 1 and 2 reported that semantically incomplete robots increase the number of responses that indicate active participation. However, there is no control of the knowledge of the conversational content. Besides, it is yet unknown that the participants' attitude and their impressions of the robot themselves.

This study focused on investigating the impression and the human's participation attitude. Therefore the robot using Enough Utterance and Incomplete Utterance was compared in a multi-party conversation where it is easy to change how to participate. Moreover, it was discussed a method for co-constructing a conversation between a human and a system.

6.3 Hypothesis

Concerning the previous study, Mikhail Bakhtin's dialogue theory [9], and phenomena found in conversations between people, the following two hypotheses were set for this study.

1. When the robot explains clearly, people do not feel a sense of participation and behave to accept the information, increasing their participation as listeners.

2. When the robot uses Incomplete Utterance, people feel a sense of participation, add additional content to the conversation, and increase their participation as speakers.

6.4 Experiment

6.4.1 Participants

All participants received an explanation that the Ethical Review Committee had approved this experiment of the Toyohashi University of Technology. A total of 40 undergraduate and graduate students and staff members (26 males and 14 females, mean age 26.1 years, standard deviation 9.22 years) who consented to the study participated in the experiment.

6.4.2 "Explaining Together" Interaction

This experiment set up an "explaining together" interaction, referring to the conversations in which spouses talk about events in a conversation reported by Lerner and Mandelbaum [66]. This experiment also referred to the study by Inoue et al. on people explaining together [72]. Based on these previous studies, this experiment was set up a three-party conversation between the speaker robot, the listener robot, and the participants shown in Fig 6.1. The participants joined as "co-explainers" who described the same contents with the speaker robot. The speaker robot was the green robot in Figure 6.1, and the listener robot was the orange robot in Figure 6.1. The listener robot was set as the participant who always listens in the conversation.

By the way, studies 1 and 2 were one-on-one between a speaker robot and a listener robot, as shown in the dotted line in Fig. 6.1.

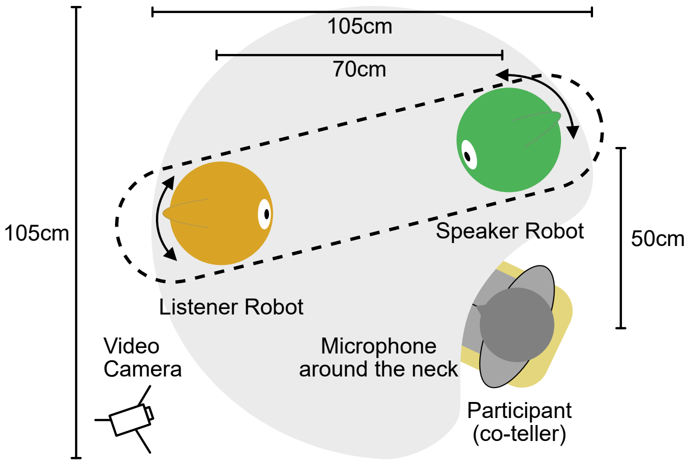


Figure 6.1: Experiment Field

6.4.3 Speech Content

Referring to the previous study by Inoue et al. [72], the participant and the speaker robot did the task of explaining the video content together. This study used "The Black Hole [1]" for the video explanation task. The video is a 3-minute silent short film with a single male actor. The speaker robot's explanation contents were created based on the explanations given by three students in the laboratory. The three students watched the video twice and then gave oral explanations about the contents. The experimenter recorded the explanation and created a scenario sheet. The sentences of the scenario sheet are shown in the left column of Table 6.1.

[1] "The Black Hole," IMDb: https://www.imdb.com/title/tt1430144/

6.4.4 Conditions

This study set two conditions: (1) the "Full" condition, in which the robot explains as described in the scenario, and (2) the "Lack" condition, in which the robot excludes some words from the scenario as described in Chapter 3. The Lack condition is the named "Semantically Incomplete Utterance" in the previous study. The robot's explanation was divided into ten utterances. The contents of the utterances were shown in Table 6.1. The middle column is for the Full condition, and the right column is for the Lack condition. Each content of the table is before adding the endings to make the content to spoken language.

6.4.5 Situations

This study set the two conditions for proving the hypothesis. Furthermore, the two situations were also set up to investigate the generality and limitation.

Figure 6.2, 6.3 shows the same flow of the participant joining the explanation. However, there are two situations: one where the speaker robot is directly involved with the participant and the speaker robot does not try to engage. It was necessary to change the speaker robot's speech pattern and behavior and the listener robot. Therefore, the situation was set as a situation, not a condition. No comparison between situations is assumed.

Table 6.1: Scenario Sheet Sentences and Speech Content for Each Speech Condition

No.	Scenario Sheet	No.	Full Condition	Lack Condition
		1	You know... It was a strange story.	You know... It was a strange story.
1	A man is standing in an office, looking tired.	2	A man was standing in his office, looking tired.	A man was standing.
2	He pushed the button on the copier too many times.	3	He pushed the button on the copier too many times.	He pushed the button.
3	A paper with a black circle on it came out of the copier.	4	A paper with a black circle on it came out of the copier.	A Black circle came out.
4	The man was surprised when the cup was vacuumed into a black circle.	5	The man was surprised when the cup was vacuumed into a black circle.	The man was surprised.
5	He put his hand carefully into the black circle,	6	He put his hand in the black circle and pulled out a cup.	He put his hand in the black circle and took it out.
6	and took out the cup			
7	The man came up with something.			
8	The man moved to the front of the vending machine.	7	The man moved to the front of the vending machine.	The man moved.
9	He put the paper on the vending machine	8	He put the paper on the vending machine and took snacks.	He took snacks.
10	and took snacks.			
11	He unlocked the door and entered the room.	9	Finally, he unlocked the door and entered a room.	Finally, he entered a room.
12	The man moved to the front of the vault.			
13	He took a lot of money from the vault.	10	He took a lot of money from the vault.	He took money.

Co-Teller Situation

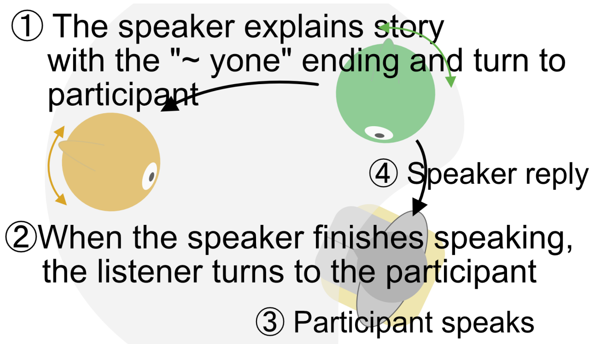


Figure 6.2: The Steps of How the Speaker Robot and the Listener Robot Behave in Co-Teller Situation

The first situation is the Co-Teller situation shown in Fig. 6.2. The speaker robot asks people to participate in the conversation. Compared to the conversation in Study 1, this situation is different in terms of knowledge control and multi-party conversation. The same thing was the speaker robot speaks directly to the participants. Before the speaker robot finished speaking, it turned to the person and said, "[⋯] yone". The ending of "ne" means "to confirm" or "to ask for cooperation" [46, 47]. The listener robot also turned toward the person after the speaker robot's speech. It took a stance to listen to the participant's speech. When the participant spoke, the speaker robot responded, and the listener robot made a nodding motion. When the speaker robot responded to the participant, the robot changed its response according to the participant's speech. If a participant told to speaker robot what the speaker robot already explained, then the speaker robot responded with "sou dane. (Yes, I see)."

On the other hand, If a participant told to speaker robot what the speaker robot not explained yet, then the speaker robot responded with "sounan dane. (Oh, I see)."

Single Situation

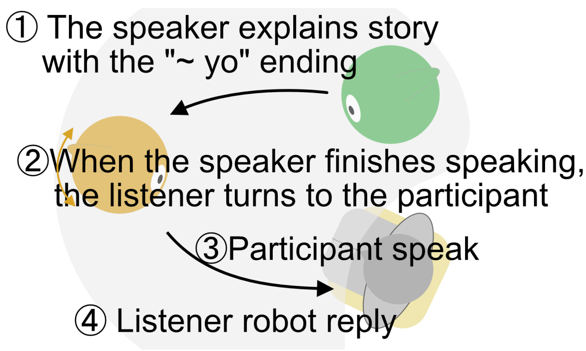


Figure 6.3: The Steps of How the Speaker Robot and the Listener Robot Behave in Single Situation

The second situation is shown in Figure 6.3, where the speaker robot explains without asking for human participation. However, suppose the speaker robot completed the explanation by itself. In that case, the conversation is closed between the speaker robot and the listener robot. There is no opportunity for a human to explain together.

Therefore, this experiment referred to the phenomena that appear in people's daily lives. For example, let us consider a multi-party conversation consisting of native speakers and a non-native speaker, and the non-native speaker and a native speaker are friends. It is known that a native speaker sometimes responds to the other native speaker despite the non-native speaker asked to him/her. Ostheimer found that people tend to make responses to the bystander instead of the person directly asked due to the preconception of ability [73]. Referring to this phenomenon, the listener (robot) may ask the participant to join the conversation instead of the speaker robot who is explaining. In the Single situation, the speaker robot said "[⋯] yo" without looking back at the person. This "yo" ending implies an assertive attitude in addition to emphasizing the address [46]. After the speaker robot's speech, the listener robot turned to the person's speech and listened to the participant's speech. When the participant spoke, the listener robot responded in the same way as in the co-teller situation.

6.4.6 Impression Evaluation

This study examined the effects and limitations of incompleteness by comparing the Lack condition and the Full condition in two different situations. The evaluation items were impression evaluation and behavior evaluation. The participants were asked to answer a pre-questionnaire and a post-questionnaire. The pre-questionnaire was the “Anxiety toward Communication Capability of Robots (S1)” and “Anxiety toward Discourse with Robots (S3)” of Nomura et al.'s Robot Anxiety Scale (RAS) [74]. The post-questionnaires were also set the following five-question groups as shown in Table 6.2.

(a) Perception of participation attitude and behavior (P1-6 in the table)

(b) Impression of the interaction (I1-5 in the table)

(c) Comparison between the speaker and listener robots (C1, C2, and liking in the table)

(d) Assessment of explanation achievement and own explanation percentage

(e) Part of the social skills test (KisXX in the table)

As for the questionnaire items, (c) uses the adjectives of Likeability from the GodSpeed questionnaire [75]. RAS-S1 and RAS-S3 in the table are the same as in the reference [74].

The answer choices of the RAS and the social skills test were the same as the references [74, 76]. Other answer choices were the following. The questions (a, b) were rated on a 5-point scale from “agree” to “disagree.” Question (c) compares the speaker robot and the listener robot with a 7-point scale. It was from “absolute green Muu" to "absolute orange Muu." The details of the 7-point scale were "absolute green or orange Muu," "green or orange Muu," "Somewhat green or orange Muu," and "Neither green Muu nor orange Muu." The 7-point scale was referring to Onodera's choice expressions [77]. Furthermore, the question content (d) was implemented with a sliding bar from 0 to 100 in increments of 1. This impression evaluation avoids specifying the roles of "speaker robot" and "listener robot." Therefore the answer choices used"green muu" and "orange muu," which refer to the speaker robot and listener robot in the text, respectively.

6.4.7 Behavior Evaluation

The total length of the speech interval was used to calculate the amount of human speech. This experiment recorded participants' speech with a microphone worn around their necks and took video with a video camera. The experimenter annotated the segments of the speech utterances using the audio resources. The results of this annotation were also used to analyze the characteristic behaviors of the participants. The transcript symbols for describing behaviors are some of the symbols used in conversation analysis [23]. The details of the symbols are also shown in the appendix. The annotations were also used for investigating the change of participant's speech. This experiment analyzed how participants respond to the robot's ten utterances shown in Table 6.1.

6.4.8 Process for Participating in the Experiment

This experiment had two experimental conditions and two situations of multi-party conversation. Therefore, there were four interactions as a combination. Participants were randomly assigned to one of the situations. After agreeing to the study, the participants answered a pre-questionnaire. They were given a scenario

Table 6.2: List of Questions Used for Impression Evaluation

Question ID	text of a questionnaire	Scale
RAS-S1	Anxiety toward Communication Capability of Robots	3 - 18
RAS-S3	Anxiety toward Discourse with Robots	3 - 18
P1	You were explaining with the green muu	1 - 5
P2	You were listening to the explanation together with the orange muu	1 - 5
P3	You participated in the conversation.	1 - 5
P4	You were adding information and scenes that were lacking in the description of the green muu.	1 - 5
P5	You added your thoughts and comments to the description of the green muu	1 - 5
P6	You agreed to the green muu's explanation.	1 - 5
I1	It was difficult to explain with the green muu	1 - 5
I2	The green muu explanation needed to be followed up.	1 - 5
I3	It was fun to explain with the green muu	1 - 5
I4	By explaining with the green muu, you were able to convey a lot more information than if you had explained alone.	1 - 5
I5	I think I may explain with green muu again.	1 - 5
C1	With which muu do you feel you were involved?	1 - 7
C2	Which muu did you talk to more?	1 - 7
Likeability	Which muu did you find XXX (adjective)?	1 - 7
Achievement	What percentage of the video content did you convey with the green muu?	0 - 100
HRate	Please indicate the percentage of what you explained based on the 100 percent of the content explained by you and Green Muu.	0 - 100
Kis01	Are you a person who does not stop talking much when you talk with others?	1 - 5
Kis03	Can you help others in a good way?	1 - 5
Kis05	Can you start a conversation with a stranger easily?	1 - 5
Kis10	Do you feel free to join in where others are talking?	1 - 5
Kis13	Can you express your feelings and emotions honestly?	1 - 5
Kis15	Can you introduce yourself well to people you meet for the first time?	1 - 5
Kis17	Can you get along with the people around you even if they think differently than you do?	1 - 5

sheet with the sentences shown in the left column of Table 1. With the scenario sheet, the participants watched the video twice. After that, the participants were given the following three instructions before proceeding to the interaction. (1) Your task is to partner with the green robot and explain the video content to the orange robot. (2) Explain the contents of the video together with the green robot following what the green robot says. (3) The content of the green robot's explanation is based on the content of the scenario sheet. After the interaction, the participants answered the post-interaction questionnaire shown in Table 6.2.

6.5 Analysis

6.5.1 Procedure

Each evaluation item was compared between the robot's Lack and Full conditions. There was no normality among the questions. Therefore, the Brunner-Munzel test (R package brunnermunzel.test), a nonparametric testing method, was used for statistical hypothesis testing. Box plots were used for the graphs.

The test results between the Lack and Full conditions in the Co-Teller and Single situations are presented in Table 6.4 and 6.5. W_N^{BF} in the table is a statistic, and MW is the superior probability of the estimated effect size. The effect size of the superior probability is small (0.44, 0.56), medium (0.36, 0.64), and large (0.29, 0.71), with 0.5 being the condition with the smallest difference [78, 79]. The two tables are included at the end of this chapter.

Firstly, the participants' attributes were compared to see if there were any differences between the participants to be compared. As shown in Fig. 6.4, there was no significant difference in the dialogue anxiety scale and the social skill test between the two conditions. Secondly, questionnaire item I2 was compared to see whether the experimental condition's incompleteness was satisfied. As shown in

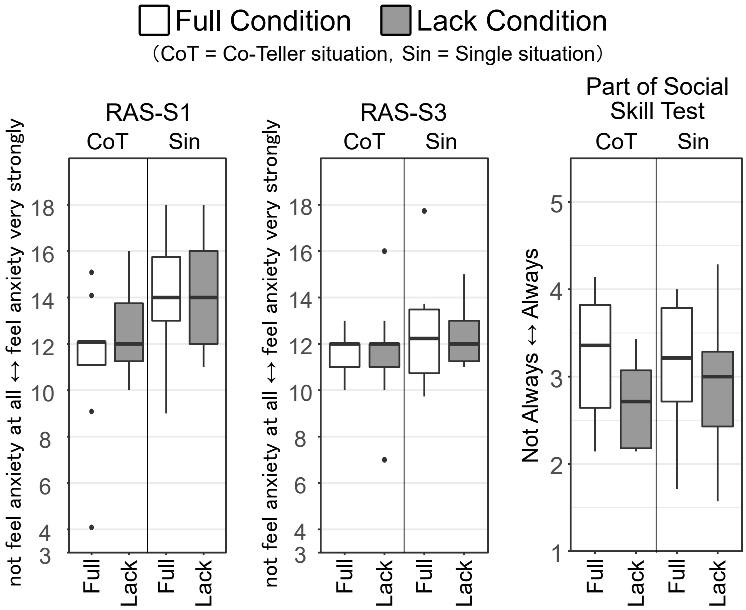


Figure 6.4: The Attributes of the Participants Based on the Post-Questionnaire

Fig. 6.5, there was a significant difference between the conditions. This study's participants perceived that the Lack condition's robot was less able for explaining. These comparisons indicate no significant differences in the participants' attributes in the experiment and that the participants recognized the incompleteness of the experimental conditions.

6.6 Result

6.6.1 Impression Evaluation

Figure 6.6 shows the results of the comparison for question item P2 related to participation as a listener. When the robot spoke in the Full condition in the CoTeller situation, the participants felt significantly listening together with the listener robot. However, there was no difference in the Single situation. Hypothesis 1, whether the attitude as a listener would increase, was limited by the situation.

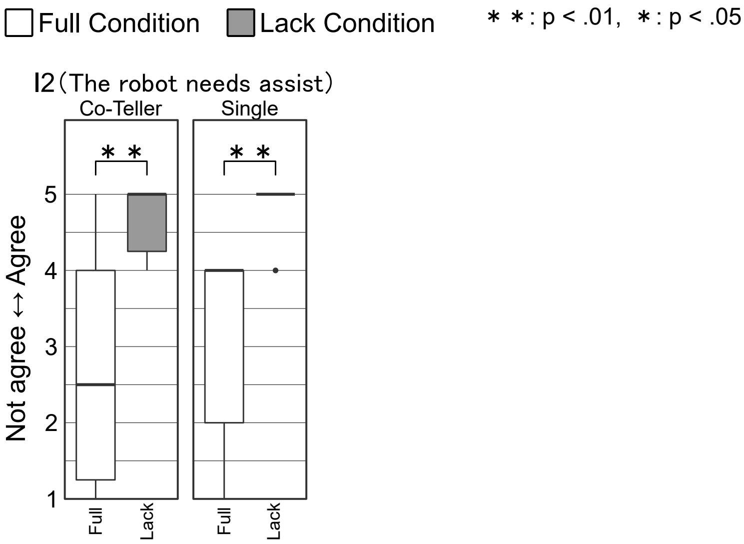


Figure 6.5: Establishment of Incompleteness Based on Post-Questionnaire

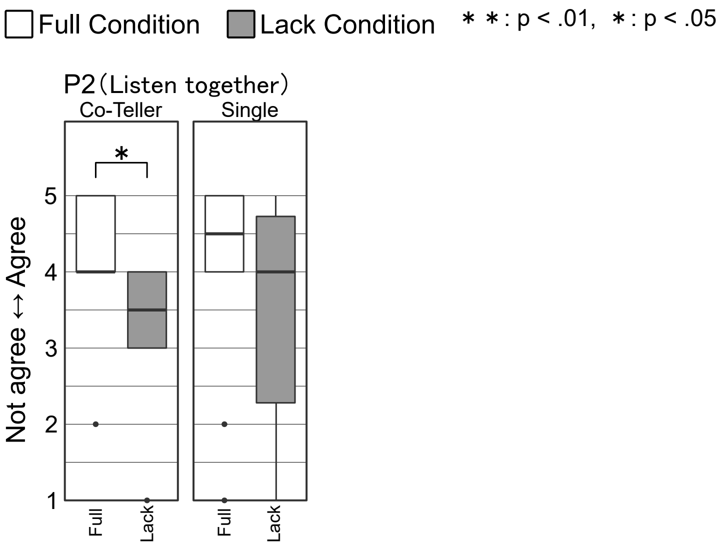


Figure 6.6: Participation Attitudes as a Listener Based on Post-Questionnaires

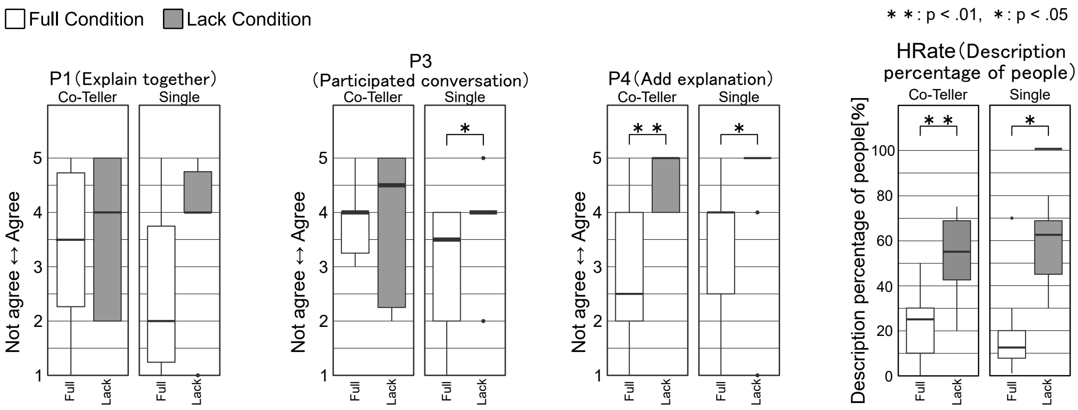


Figure 6.7: Participation Attitudes as a Speaker Based on Post-Questionnaires

Next, Figure 6.7 shows the results of the comparison of P1, P3, P4, and the percentage of human explanations related to participation as a speaker. As illustrated in the graphs, the Lack condition participants felt that they made a significantly more supplementary speech when explaining with the robot. The percentage of human explanations also increased. However, there was no difference between the conditions in whether the participants were explaining together with the speaker robot. Concerning hypothesis 2, it was confirmed that there was an increase in the attitude of participation as a speaker. However, there was no evidence that they were explaining "together."

Finally, Figure 6.8 shows the results of a comparison to evaluate if there was a difference in the achievement of the explanation when the Lack Condition robot participated in the explanation.

6.6.2 Behavior Evaluation

This section report the behavior of people observed in the experiment. The first is the amount of speech. Figure 6.9 shows the results of the comparison of human speech amount. The graphs showed that people increased their speech significantly in the Lack condition of the Co-Teller situation. However, in the Single situation,

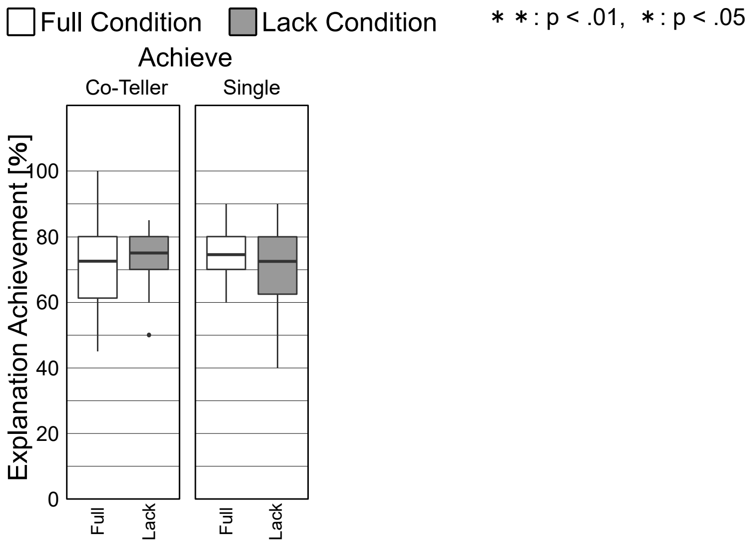


Figure 6.8: Performance of the Conversation Based on the Post-Questionnaire.

there was no difference between the conditions.

As seen in the impression evaluation, the participants added information to the speaker robot's explanation (6.7). Transcript (1) shows an example of a participant explaining.

```
    (1) Example of Participants add explanations.
    The first line is the alphabetical representation of the original
    Japanese text. (X: ) means who is speaking.
    (SP: Speaker Robot, LI: Listener Robot, HU: Participant)
    The second line is the translation by the author
    SP:  soredene
         so...
         (3.0)
    SP:  okashi o tottann dayone
         He took snacks.
         (0.8)
 →  HU:  soudane(.)jidouhanbaiki ni kami o haritsukete
         yes.      He put the paper on the vending machine
         (0.8)
    HU:  ne
         (0.4)
 →  HU:  sononaka no okashi o tottan dayone
         He took snacks from it.
         (2.0)
    SP:  sounan dane
         Oh, yeah.
```

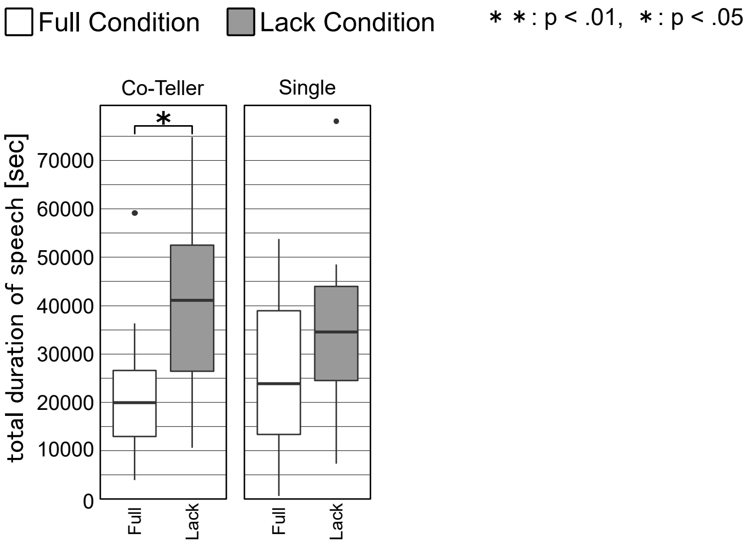


Figure 6.9: The Amount of Speech as a Behavior.

Also, Fig. 6.10 illustrates the shift of speech. It indicates that some participants responded with a minimal response or a single word. Transcript (2) shows how the participants gave minimal responses to the speaker robot's explanation.

```
     (2) Example of Participants add minimal responses.
     The first line is the alphabetical representation of the original
     Japanese text. (X: ) means who is speaking.
     (SP: Speaker Robot, LI: Listener Robot, HU: Participant)
     The second line is the translation by the author
     SP:  soushitara ne
          then,
          (2.7)
     SP:  jidouhanbaiki ni kami o haritsukete okashi o totta n dayo
          He put the paper on the vending machine and took snacks.
          (3.5)
  →  HU:  sou dane
          That's right.
          (2.1)
     LI:  un(2.2)sounanda
          oh     I see.
```

Finally, this behavior evaluation reports the differences in the participants' speech endings as a characteristic behavior. The participants were divided into three groups. One group added "ne" to their explanations, as shown in the transcript (1). The second group added the modality of "yo," as shown in the transcript

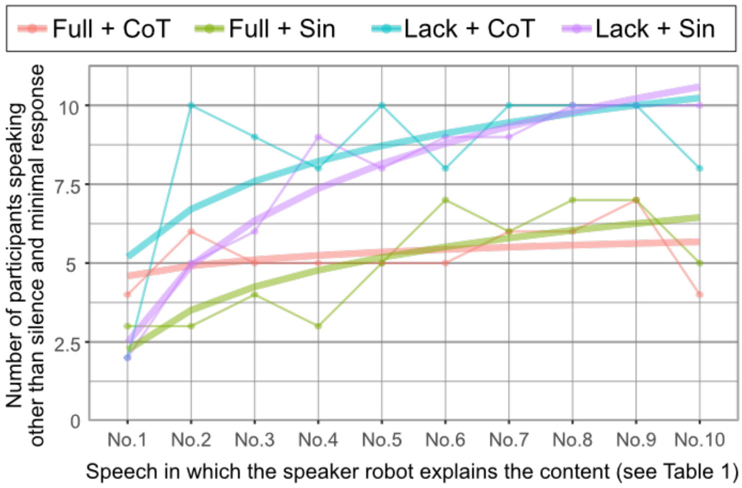


Figure 6.10: Changes in the Number of Participants Using Speech Other than Minimal Responses. (CoT = Co-Teller situation, Sin = Single situation)

(3). The third group used "ta" for all their speech, as shown in the transcript (4). In the case of using "ne," the participants turned their faces to the speaker robot while speaking. In using "yo," they turned their faces to the listener robot while speaking. As shown in Table 6.3, bias was found in the utterances using "ne" in the Co-Teller situation. However, no feature of bias toward "ne" and "yo" was found in the Single situation.

Four participants spoke with the "ta" ending, and they all experimented with a Single situation. Nakamura says that it is unusual to see this form of expression in Japanese conversation [80], and those who participated using this "ta" ending did not perform minimal responses such as "un (I see.)", "soudane (That is right.)"

```
(3) Example of Participants speaking with the ‘‘yo’’ ending.
The first line is the alphabetical representation of the original
Japanese text. (X: ) means who is speaking.
(SP: Speaker Robot, LI: Listener Robot, HU: Participant)
The second line is the translation by the author
SP:   sore dene
      so...
      (3.2)
SP:   otoko no hito wa idou shita n dayo
      The man moved.
```

Table 6.3: Classification of the Number of Participants According to the Modality ("none" means "yes" or other minimal response or silence)

		ne	yo	ne, yo mixed	ta	none
Co-Teller	Full	9	0	1	0	0
	Lack	7	0	3	0	0
Single	Full	3	1	3	2	1
	Lack	3	2	3	2	0

```
       (3.4)
→ HU:  otoko no hito wa nanika o(0.3)omoi tsuita mitai(0.9)datta yo
       The man seemed to have come up with something.
       (2.0)
  LI:  un
       Uh-huh.

  (4) Example of Participants speaking with the ‘‘ta’’ ending.
  The first line is the alphabetical representation of the original
  Japanese text. (X: ) means who is speaking.
  (SP: Speaker Robot, LI: Listener Robot, HU: Participant)
  The second line is the translation by the author
  SP:  sore dene
       so...
       (3.1)
  SP:  okane o tottan dayo
       He took money.
       (3.8)
  HU:  sono kuroi kami o tsukatte(0.5)kinko no naka kara okane o dashi
       mashita
       He used that black paper to get the money out of the vault.
       (1.7)
  H :  sou nan dane
       Ah, I see.
```

6.6.3 Extra Impression Evaluation (Excluding the four people with the "ta" ending)

The participants in the "ta" endings might be interacting differently from the other participants in the Single situation. Therefore, the characteristics of the

Single situation may have been lost in the previous analysis. The two participants in the Lack condition and two participants in the Full condition were excluded from the single situation and compared again. As shown in Fig. 6.11, there was a significant difference between the Lack and Full conditions in item C2, which compares whether the participant was talking more to the speaker or listener robot (p=0.012, df=7.98, WBFN=-0.74, MW=0.86). No significant differences from the previous test were identified for the other items, except for the participants with the "ta" ending.

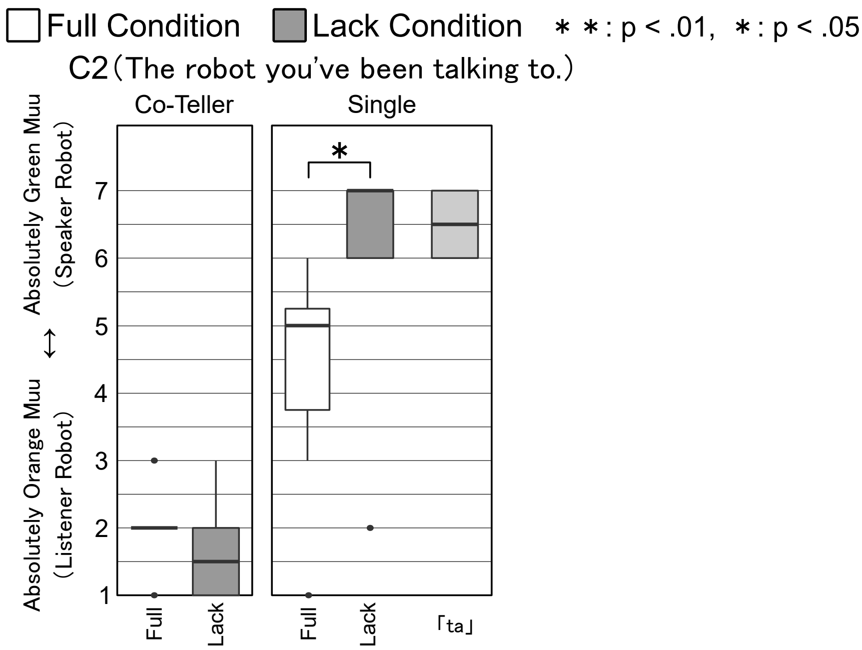


Figure 6.11: Comparative Results: Participants Felt Spoke to Which Robot with Separating Participants Who Used the "ta" Ending for Speech.

6.7 Discussion

6.7.1 Discussion 1: Attitude of Participation as a Listener

From the results of P2 "Were you listening with the listening robot?" in Fig. 6.6. Comparing the amount of human speech in Fig. 6.9 confirmed that Hypothesis 1 was limited by the situation. The shift of human speech during the conversation in Fig. 6.10 shows that in the Full condition, a certain number of people became listeners by interrupting and participated in conversation silently. In the Co-Teller situation, the person became a listener when the robot gave a clear explanation, as in Hypothesis 1. However, there was no difference in the P2 evaluation and the amount of speech between the conditions of the Single situation. Here discuss the situation and the listener's attitude.

Figure 6.6 shows a uniform distribution of P2 "listening to the explanation together." The uniform distribution means that there was no bias in the impressions of the participants. The unbias in the Single situation's lack condition is one reason why no significant differences were found. This study supposed the reason might be that the listener robot was between the speaker robot and the participant. Participants' speeches may not have been as stable as in the Co-Teller situation.

Since this study's primary purpose was to verify the effectiveness of the speaker robot's speech, the listener robot's intervention was set to a minimum, so the listener robot only turned toward the person silently. When participants considered how to participate in achieving the goal of explaining together, they were not able to use the listener robot's prior utterances as a resource. Therefore, in the Single situation, the method of participation was mostly freely entrusted to the participants. Table 6.3 shows that participants participated with the "ta" ending in both the Lack and Full conditions only in the Single situation. As explained, the "ta" ending is not seen in everyday conversation. It may indicate that the participation method was not normal in the Single situation and that the speech condition's effect was not observed.

6.7.2 Discussion 2: Attitude of Participation as a Speaker

It was confirmed that the person significantly increased their participation as a speaker when the robot became incomplete. It is evidenced by the results of P4 (whether or not the robot was adding supplementary information), the percentage of explanation by the person. These results suggest that the speaker robot's Incomplete Utterance increased the participants' attitude as speakers, and Hypothesis 2 was accepted. However, no significant difference was found in the impression of whether the participants were explaining together (P1). Why did the participants not feel that they were "explaining together"?

The task was to explain the content of the video together. It is known that tasks are divided into two types: tasks that have a single answer (well-defined tasks) and tasks that have no single answer (ill-defined tasks) [81]. This joint explanation would be categorized as the former because it has an answer as the video content. Suzuki et al. reported that "presentation of ideas" and "presentation of questions" are the behaviors that affect the satisfaction of a well-defined task. It is also reported that the satisfaction is affected by the participants performing and receiving these behaviors each other. In this experiment, the speaker robot was only able to "explain" as speaking and "show agreement" as listening, and there was no diversity in behavior as reported by Suzuki et al. [81]. The interaction could not say that was "each other." Therefore, it is possible that the commonality of "together" was not evaluated, and further study is needed.

6.7.3 Discussion 3: Using Incomplete Utterance

Incomplete Utterance was used in the Lack condition. It was confirmed that the method increased the amount of human participation as in Study 1. Also, no significant difference was found in the degree of accomplishment of the explanation in Figure 6.8. This comparison means that even if the robot's speech became incomplete, the explanation's degree of accomplishment did not change with human participation.

This study mainly examined the effect of "incompleteness" in a multi-party conversation. Considering both Study 1 and the present study together, the robot's incomplete state affects people's participation attitude and the amount of speech in the situation where the speaker robot talks directly to people.

Additionally, Figure 6.11 showed the evaluation of "Were you participating in the conversation?" According to the result, participants felt that they were not participating in the conversation when the robot was in the Single situation's Full condition. When humans and robots describe an event together, robots should not speak firmly as the Single situation in Japanese. It may take away the human's role as a speaker and make them feel that they were not participating in the conversation.

Since this experiment took a constructivist approach, it was not possible to show a clear causal relationship between incompleteness and the amount of human speech. Although this study has examined the incomplete elements to be added to the robot, further analysis is needed.

6.7.4 Discussion 4: Limitations and Prospects of This Study

This study was an in-laboratory experiment, and the participants were instructed. Conversations in daily life are not limited in time beforehand. There is no limit to the number of interactions with the same person outside the laboratory. If this study would be used outside the laboratory, the current strategy might be not enough. It is just talking incompletely. It is monotonous and may cause boredom problems. In order to implement the system in society, further verification will be necessary.

Future study should be how to create incompleteness in a conversation. In conversations between people, the listener points out what the speaker's speech lacks, as in Suzuki [30]. The "listeners point out behavior" could be considered one way to avoid monotonous conversations with predetermined rules. On the other hand, Mimura et al. explored the methods of "Estimating the other person's comprehension using human facial expressions." It will be possible to discuss the incompleteness between a person and a speaker robot [82].

As one of the other considerations, future studies need to discuss the interaction where incompleteness can be enjoyed. The results of Study 1 to 3 can say that the robot's incompleteness changes people's behavior to participate in the Japanese conversation when the speaker robot speak toward a human directly. However, if a person asks the robot for directions and the robot gives an incomplete explanation, that does not make sense. The interaction will be different from what the people expected, and the people will feel uncomfortable. It is necessary to study further the interactions that are suitable for incompleteness that can elicit human participation.

Table 6.4: Statistical Test Results for Co-Teller Situation

	Full			Lack			*: p <0.05, **: p <0.01, †: p <0.1				
Question item	1st Qu.	median	3rd Qu.	1st Qu.	median	3rd Qu.	W_N^{BF}	df	p	MW	
RAS-S1	11.0	12.0	12.0	11.3	12.0	13.8	0.94	18.0	0.357	0.62	N.S.
RAS-S3	11.0	12.0	12.0	11.0	12.0	12.0	0.19	16.0	0.850	0.53	N.S.
P1	2.25	3.50	4.75	2.00	4.00	5.00	0.30	17.8	0.765	0.54	
P2	4.00	4.00	5.00	3.00	3.50	4.00	-3.22	15.4	0.006	0.21	**
P3	3.25	4.00	4.00	2.25	4.50	5.00	0.67	16.6	0.513	0.59	
P4	2.00	2.50	4.00	4.00	5.00	5.00	4.79	12.7	0.000	0.87	**
P5	2.25	4.00	4.00	3.00	4.00	5.00	0.77	18.0	0.450	0.60	
P6	4.25	5.00	5.00	4.00	4.50	5.00	-0.88	17.9	0.388	0.40	
I1	2.00	3.00	4.75	2.00	4.00	5.00	0.58	17.9	0.567	0.58	
I2	1.25	2.50	4.00	4.25	5.00	5.00	3.81	11.3	0.003	0.84	**
I3	2.00	4.00	4.00	4.00	4.00	4.00	1.25	16.2	0.229	0.65	
I4	3.00	3.50	4.00	3.25	4.00	4.00	0.42	18.0	0.678	0.56	
I5	2.25	3.50	4.00	4.00	4.00	4.00	1.42	11.7	0.180	0.68	
C1	1.00	1.50	2.00	1.00	2.00	2.00	0.65	17.0	0.524	0.58	
C2	2.00	2.00	2.00	1.00	1.50	2.00	-0.74	14.1	0.470	0.41	
Likeability	2.00	2.50	3.45	2.05	2.40	3.00	-0.11	17.2	0.914	0.49	
Part of Social Skill test	2.64	3.36	3.82	2.18	2.71	3.07	-1.88	14.5	0.080	0.28	†
Achievement	61.25	72.50	80.00	70.00	75.00	80.00	0.33	15.1	0.747	0.55	N.S.
HRate	10.00	25.00	30.00	42.50	55.00	68.75	5.09	16.5	0.000	0.88	**
Speech Amount	12924	19919	26576	26415	41079	52484	2.24	17.7	0.038	0.76	*

Table 6.5: Statistical Test Results for Single Situation

	Full			Lack							
Question Items	1st Qu.	median	3rd Qu.	1st Qu.	median	3rd Qu.	W_N^{BF}	df	p	MW	
RAS-S1	13.0	14.0	15.8	12.0	14.0	16.0	-0.25	16.1	0.803	0.47	N.S.
RAS-S3	11.0	12.5	13.8	11.3	12.0	13.0	0.07	16.0	0.942	0.51	N.S.
P1	1.25	2.00	3.75	4.00	4.00	4.75	1.76	16.9	0.097	0.72	†
P2	4.00	4.50	5.00	2.25	4.00	4.75	-0.91	17.9	0.375	0.39	
P3	2.00	3.50	4.00	4.00	4.00	4.00	2.36	16.9	0.030	0.73	*
P4	2.50	4.00	4.00	5.00	5.00	5.00	2.68	17.0	0.016	0.79	*
P5	1.25	2.50	3.75	1.00	1.50	4.75	0.00	12.3	1.000	0.50	
P6	3.25	4.50	5.00	3.25	4.50	5.00	0.16	17.3	0.876	0.52	
I1	3.25	4.00	4.75	1.25	2.00	3.50	-1.65	16.3	0.119	0.30	
I2	2.00	4.00	4.00	5.00	5.00	5.00	4.01	13.8	0.001	0.84	**
I3	3.25	4.00	4.00	4.00	4.00	4.00	0.86	17.4	0.400	0.60	
I4	3.00	4.00	4.75	2.00	3.00	3.75	-1.56	14.8	0.140	0.31	
I5	2.25	4.00	5.00	4.00	4.00	4.00	0.07	13.0	0.943	0.51	
C1	2.00	4.00	5.75	4.00	5.50	6.00	1.30	15.5	0.214	0.66	
C2	4.25	5.00	6.00	6.00	6.50	7.00	2.06	17.8	0.054	0.74	†
Likeability	2.80	3.60	4.40	2.65	3.10	3.65	-0.93	16.6	0.365	0.38	
Part of Social Skill test	2.71	3.21	3.79	2.43	3.00	3.29	-0.59	18.0	0.565	0.42	N.S.
Achievement	70.00	74.50	80.00	62.50	72.50	80.00	-0.41	16.0	0.689	0.45	N.S.
HRate	7.75	12.50	20.00	45.00	62.50	68.75	2.90	16.0	0.010	0.83	*
Speech Amount	13346	23844	38923	24495	34533	43956	1.07	16.8	0.301	0.64	

7 Conclusion

This paper investigates the incompleteness in a speaker robot's speech. This study discussed the opportunity for human participation based on Mikhail Bakhtin's dialogue theory. The three studies were conducted with interactions in one-to-one conversational situations and two-to-one cooperative explaining situations. The two interaction scenes shed light on the possibility of constructing collaborative conversations based on human participation, as opposed to traditional information transfer oriented conversation designs.

Chapters 1 and 2 considered conversations as social interaction and classified conversations between people and systems. First, this paper described imperfections of human-to-human conversations in contrast to human-to-system conversations. Although the information-transfer is mainstream for conversational systems, this study stated the necessity to reveal the collaboration to create conversation. This chapter summarized this study's theoretical background and mentioned the position of this study.

Chapter 3 defined the "Incomplete Utterance" that was an approach discussed across this paper. This chapter described how to implement incompleteness to a robot. This approach was referring the characteristics of children and practices of daily conversations. The points of implementation were how to create opportunities to elicit people's participation.

Chapter 4 used a single robot to set up a one-to-one conversation. This chapter analyzed human responses to two Incomplete Utterance strategies for testing the appropriateness. The results suggest that the Semantically Incomplete Utterance increased human responses that indicated active participation in the conversation, such as "questions" and "introduction of new information relevant to the current topics."

Chapter 5 evaluated the impressions of conversations using the videos recorded in the Chapter 4 experiment. This third-person evaluation confirmed that the robot's conversation had used semantically Incomplete Utterance was significantly evaluated as cooperative. However, the results of Chapters 4 and 5 did not reflect the impressions of actual interactors. Therefore this paper conducted the next study.

In Chapter 6, multi-party conversations were set up using two robots. This chapter focused on evaluating the interactors' impressions. This chapter described why the third-party conversation is satisfied for this purpose. This study used a speaker robot and a listener robot, and the setting was "A human and a robot explain together." Two types of speaker robots were compared between fully explain robot and incompletely explain robot. The result indicated that if the speaker robot used Incomplete Utterance, the explanation achievement was not changed with human assistance. It was also suggested that the Incomplete Utterance needs to be used as if robots talk directly to a person. The limitation was needed to engage participants and change their participation attitude. Togetherness is one of the remaining challenges of this study. Although Incomplete Utterance can trigger people to participate in the robot's action, continuous use alone may not produce collaborative action. In the future, it is necessary to examine and construct a model that focuses on the chain of actions between humans and the system.

This study discovered the effects and limitations of incompleteness in speaker robot's speech. As a result, this work was able to be summarized as a patent (Japanese Patent No. 3752522). It could be said that this research aims to contribute to the basis for the next generation of interactive technology, has been achieved. In recent years, robots that do not speak Japanese at all on purpose have been commercialized and are gaining popularity in Japan. As robots that can accurately convey information expands due to their convenience, there is a possibility that robots equipped with opportunities for human participation will also expand. However, the current situation is limited to Japan. This study also leaves an issue of application to other languages than Japanese. The background of this study, fortunately, is not limited to the Japanese language. There must be methods to create opportunities for people to participate according to their language and culture. I could only do so much in my doctoral program's three years, but I look forward to future research.

Bibliography

[1] Yutaka Saeki. *Kyokan - Sodachiau Hoiku no Nakade*. 2007. (Japanese).

[2] Vasudevi Reddy. *How infants know minds*. Harvard University Press, 2010.

[3] Vasudevi Reddy and Yutaka Saeki(translator). *Odorokubeki Nyuyouji no Sekai: "Nininsho apuro-chi" kara miete kurukoto*. Mineruvua Shobou, 2015. (Japanese).

[4] Deborah Tannen. *You just don't understand: Women and men in conversation*. Morrow New York, 1990.

[5] Martin Porcheron, Joel E Fischer, Stuart Reeves, and Sarah Sharples. Voice interfaces in everyday life. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, pp. 1–12, 2018.

[6] Janet Holmes. Doing collegiality and keeping control at work: Small talk in government departments. *Small talk*, Vol. 32, p. 61, 2000.

[7] Yu. M. Lotman. Text within a text. *Soviet Psychology*, Vol. 26, No. 3, pp. 32–51, 1988.

[8] Michio Okada. On indeterminancy in talk-in-interaction. *Journal of Japanese Society for Artificial Intelligence*, Vol. 16, No. 6, pp. 819–825, 2001. (Japanese).

[9] Mikhail M. Bakhtin. *The Dialogic Imagination: Four Essays.* University of Texas Press, 1983.

[10] Mikhail M. Bakhtin and Ichiro Ito (translator). *Shosetsu no kotoba.* Heibon sya, 1996. (Japanese).

[11] James V Wertsch and James V Wertsch. *Voices of the mind: Sociocultural approach to mediated action.* Harvard University Press, 2009.

[12] James V Wertsch, Nobumoto Tajima (translator), Koji Sato (translator), Yuji Moro (translator), and Kayoko Uemura (translator). *Kokoro no Koe– Baikaisareta Koui eno Syakaibunkateki apuro-chi.* Fukumura Shuppan, 2004. (Japanese).

[13] Kazuyo Murata and Risako Ide. *The kaleidoscope of small talk : a linguistic approach.* Hitsuji Shobou, 2016. (Japanese).

[14] Ryuichiro Higashinaka, Kotaro Funakoshi, Michimasa Inaba, Yuiko Tsunomori, Tetsuro Takahashi, and Reina Akama. The dialogue system live competition. *SIG-SLUD*, Vol. B5, No. 02, pp. 106–111, 2018. (Japanese).

[15] Takanori Komatsu, Seiji Yamada, Kazuki Kobayashi, Kotaro Funakoshi, and Mikio Nakano. Artificial subtle expressions: Intuitive notification methodology of artifacts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, p. 1941–1944. Association for Computing Machinery, 2010.

[16] Noriko Suzuki, Yugo Takeuchi, Kazuo Ishii, and Michio Okada. Psychological effects derived from echoic mimicry using inarticulate sounds. *Transactions of Information Processing Society of Japan*, Vol. 41, No. 5, pp. 1328–1338, 2000. (Japanese).

[17] Hirokazu Kiyomaru, Masahiko Osawa, and Michita Imai. Predictive recognition-based linguistic communication without using natural language. *The Japanese Society for Artificial Intelligence SIG-AGI-006-03*, 2017. (Japanese).

[18] Y. Yamaji, T. Miyake, Y. Yoshiike, P. R. De Silva, and M. Okada. Stb: Human-dependent sociable trash box. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 197–198, 2010.

[19] Naohiro Matsumura. *Shikake: The Japanese Art of Shaping Behavior Through Design*. Liveright, 2020.

[20] Naohiro Matsumura. *Shikake gaku: Hito wo ugokasu aidea no tsukurikata*. TOYO KEIZAI INC., 2016. (Japanese).

[21] Leslie A Baxter. Dialectical contradictions in relationship development. *Journal of Social and Personal relationships*, Vol. 7, No. 1, pp. 69–88, 1990.

[22] Leslie A Baxter and Barbara M Montgomery. *Relating: Dialogues and dialectics*. Guilford Press, 1996.

[23] Tomoko Takagi, Yuri Hosoda, and Emi Morita. *Basics of Conversation Analysis*. Hitsuji Shobou, 2016. (Japanese).

[24] Sau Kuen Clare Fan. Contact situations and language management. *Multilingua-Journal of Cross-Cultural and Interlanguage Communication*, Vol. 13, No. 3, pp. 237–252, 1994.

[25] Natsuho Iwata. The change of features of participation in a conversation between a japanese language learner and a japanese native speaker:from asymmetrical participation to symmetrical participation. *Japanese-language education around the globe*, pp. 135–151, 2005. (Japanese).

[26] Emanuel A Schegloff, Gail Jefferson, and Harvey Sacks. The preference for self-correction in the organization of repair in conversation. *Language*, Vol. 53, No. 2, pp. 361–382, 1977.

[27] Emanuel A Schegloff. Practices and actions: Boundary cases of other-initiated repair. *Discourse processes*, Vol. 23, No. 3, pp. 499–545, 1997.

[28] Gene H Lerner. *Conversation analysis: Studies from the first generation*, Vol. 125. John Benjamins Publishing, 2004.

[29] Kaoru Hayano. Shufuku no soshiki (special issue the science of interaction). *Japanese Linguistics*, Vol. 36, No. 4, pp. 82–92, 2017. (Japanese).

[30] Kana Suzuki. Other-initiated repair on unexpressed utterance elements : Rethinking "argument ellipsis" as a conversational phenomenon. *The Japanese Journal of Language in Society*, Vol. 10, No. 2, pp. 70–82, 2008. (Japanese).

[31] Shuya Kushida. Recipient's methods for facilitating progress of telling: Sustaining, prompting and promoting continuation. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society*, Vol. 16, No. 1, pp. 12–23, 2009. (Japanese).

[32] Etsuo Mizukami and Koji Yamashita. An examination of a function of fillers as maintaining the speaker's right to speak. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society*, Vol. 14, No. 4, pp. 588–603, 2007. (Japanese).

[33] Anita Pomerantz. Telling my side: "limited access" as a "fishing" device. *Sociological inquiry*, Vol. 50, No. 3-4, pp. 186–198, 1980.

[34] Tetsuri Toe. Parent identity in action. *Japanese Sociological Review*, Vol. 62, No. 4, pp. 536–553, 2012. (Japanese).

[35] Hitomi Matsushita, Yohei Kurata, P Ravindra S De Silva, and Michio Okada. Talking-ally: What is the future of robot's utterance generation? In *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 291–296. IEEE, 2015.

[36] Kiyoshi Amano. *Youji no Bunpou nouryoku.* Tokyo Shoseki, 1977. (Japanese).

[37] Brian MacWhinney. *The CHILDES Project: Tools for analyzing talk. transcription format and programs*, Vol. 1. Psychology Press, 2000.

[38] Susanne Miyata. Guideline for japanese mlu : How to compute mluw and mlum. *Journal of health and medical science*, No. 2, pp. 1–17, 2012.

[39] Yoshiko Okayama. Haha to ko no taiwa no shuurooku [a collection of mother-child conversations], 1970. (Japanese).

[40] Yoshiko Okayama. Haha to ko no taiwa no shuurooku [a collection of mother-child conversations], 1973. (Japanese).

[41] Yasuhiro Shirai, Susanne Miyata, Norio Naka, and Yoshiko Sakazaki. The acquisition of causative morphology in japanese: A prototype account. *East Asian Language Acquisition*, 2001.

[42] Midori Inaba. Developmental changes of mlu in japanese- and english-language narratives. *Bulletin of Aichi Univ. of Education (Humanities and Social Sciences)*, Vol. 67, No. 1, pp. 47–55, 2018. (Japanese).

[43] Yukiko Hayashi and Shigeki Matsubara. Sentence style conversion for text-to-speech application. *IPSJ Natural Language Processing*, No. 47(2007-NL-179), 2007. (Japanese).

[44] Nomuhiro Kaji, Masashi Okamoto, and Sadao Kurohashi. Paraphrasing predicates from written language specific vocabulary into spoken language vocab-

ulary using the world wide web. *Journal of Natural Language Processing*, Vol. 11, No. 5, pp. 19–37, 2004. (Japanese).

[45] Sato Toshinori. Neologism dictionary based on the language resources on the web for mecab, 2015.

[46] Senko K Maynard. *Discourse modality: Subjectivity, emotion and voice in the Japanese language*. John Benjamins Publishing, 1993.

[47] Emi Morita. Functions of the japanese interactional particle ne and its socio-pragmatic implications. *The Japanese Journal of Language in Society*, Vol. 10, No. 2, pp. 42–54, 2008. (Japanese).

[48] Yusaku Nishiwaki, Sho Itashiki, Nihan Karatas, and Michio Okada. Cooperative interactions generated by incompleteness in robots' utterance. In *Proceedings of the 6th International Conference on Human-Agent Interaction*, HAI '18, p. 76–83, New York, NY, USA, 2018. Association for Computing Machinery.

[49] James Vlahos. *Talk to Me: How Voice Computing Will Transform the Way We Live, Work, and Think*. Eamon Dolan / Houghton Mifflin Harcourt, 2019.

[50] Michinari Kohno. *Onsei ni Mirai ha Arunoka?* Nikkei Business Publications, Inc., 2018. (Japanese).

[51] John Hinds. Reader versus writer responsibility: A new typology. *Landmark essays on ESL writing*, Vol. 17, p. 63, 2001.

[52] Ryuichiro Higashinaka, Kotaro Funakoshi, Masahiro Araki, Hiroshi Tsukahara, Yuka Kobayashi, and Masahiro Mizukami. Towards taxonomy of errors in chat-oriented dialogue systems. In *Proceedings of the 16th annual meeting of the special interest group on discourse and dialogue*, pp. 87–95, 2015.

[53] Isao Iori. From the viewpoint of “easy japanese”: what can the pedagogical grammar of japanese say to “grassroots” japanese-language classrooms? *Hitotsubashi review of arts and sciences*, No. 3, pp. 126–141, 2009. (Japanese).

[54] Hideki Tanaka, Hideya Mino, Shinji Ochi, and Motoya Shibata. Trial service of easy japanese news on the we. *NHK Science & Technology Research Laboratories R&D*, No. 139, pp. 20–29, 2013. (Japanese).

[55] Shuya Kushida. *Tasukebune to Osekkai - Kaiwa ni okeru Sanyo to Kategorika niokeru ichi kousatsu*, pp. 124–147. Sekai Shisou sya, 1999. (Japanese).

[56] Duck-Young Lee. Involvement and the japanese interactive particles ne and yo. *Journal of Pragmatics*, Vol. 39, No. 2, pp. 363–388, feb 2007.

[57] Takanori Komatsu and Seiji Yamada. Adaptation gap hypothesis: How differences between users’ expected and perceived agent functions affect their subjective impression. *Journal of Systemics, Cybernetics and Informatics*, Vol. 9, No. 1, pp. 67–74, 2011.

[58] Cristina Zaga, Roelof A.J. de Vries, Jamy Li, Khiet P. Truong, and Vanessa Evers. A simple nod of the head: The effect of minimal robot movements on children’s perception of a low-anthropomorphic robot. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 336–341, 2017.

[59] Per Linell, Lennart Gustavsson, and Päivi Juvonen. Interactional dominance in dyadic communication: a presentation of initiative-response analysis. *Linguistics*, Vol. 26, No. 3, pp. 415–442, 1988.

[60] Yusaku Nishiwaki, Kenta Yosimi, and Michio Okada. Why incomplete utterance attract people’s minds? In *Human Agent Interaction Symposium*,

pp. D–1. Human-Agent Interaction Symposium Program Committee, 2016. (Japanese).

[61] Yusaku Nishiwaki, Sho Itashiki, and Michio Okada. Cooperative interactions generated by incompleteness in robot's utterance. *The Transactions of Human Interface Society*, Vol. 21, No. 1, pp. 1–12, 2019. (Japanese).

[62] Kazumi Ogawa. The effects of utterance balance in a conversation on a neutral observer's impressions of the conversants and the conversational process. *The Japanese Journal of Experimental Social Psychology*, Vol. 43, No. 1, pp. 63–74, 2003. (Japanese).

[63] Kazumi Ogawa. The influence of cue information on impression formation of dyadic conversations (special issue: Investigating verbal behaviour with experiments). *The Japanese Journal of Language in Society*, Vol. 9, No. 1, pp. 27–36, 2006. (Japanese).

[64] Jennifer L Monahan. Information processing differences of conversational participants and observers: The effects of self-presentation concerns and cognitive load. *Communications Monographs*, Vol. 62, No. 3, pp. 265–281, 1995.

[65] Yusaku Nishiwaki, Naoki Ohshima, and Michio Okada. Effects of incomplete utterance on conversational engagement in a multi-party conversation. *Transactions of the Japanese Society for Artificial Intelligence*, Vol. 36, No. 2, p. 12 pages, 2021. (Japanese).

[66] Gene H. Lerner. Assisted storytelling: Deploying shared knowledge as a practical matter. *Qualitative Sociology*, Vol. 15, No. 3, pp. 247–271, 1992.

[67] Jennifer Mandelbaum. Couples sharing stories. *Communication Quarterly*, Vol. 35, No. 2, pp. 144–170, 1987.

[68] Yasuharu Den. Dynamics of three-party conversation. *Japanese linguistics*, Vol. 32, No. 1, pp. 4–13, 2013. (Japanese).

[69] Kazuya Shitaoka, Ryoko Tokuhisa, Takayoshi Yoshimura, Hiroyuki Hoshino, and Narimasa Watanabe. Active listening system for a conversation robot. *Journal of Natural Language Processing*, Vol. 24, No. 1, pp. 3–47, 2017. (Japanese).

[70] Yoichi Matsuyama, Iwao Akiba, Shinya Fujie, and Tetsunori Kobayashi. Four-participant group conversation: A facilitation robot controlling engagement density as the fourth participant. *Computer Speech & Language*, Vol. 33, No. 1, pp. 1–24, 2015.

[71] Takamasa Iio, Yuichiro Yoshikawa, and Hiroshi Ishiguro. Pre-scheduled turn-taking between robots to make conversation coherent. In *Proceedings of the Fourth International Conference on Human Agent Interaction*, pp. 19–25, 2016.

[72] Suguru Inoue, Yasuyuki Sumi, and Katsuya Takanashi. Collaborative utterances with gestures in descriptive conversation. In *Information Processing Society of Japan, Interaction*, Vol. 3, pp. 457–460, 2011. (Japanese).

[73] Teja OSTHEIDER. "3rd person response" as a form of overaccommodation : An example of inappropiate communication behaviour towards foreign and physically handicapped people. *The Japanese Journal of Language in Society*, Vol. 7, No. 2, pp. 39–49, 2005. (Japanese).

[74] Tatsuya Nomura, Takayuki Kanda, Tomohiro Suzuki, and Kensuke Kato. Prediction of human behavior in human-robot interaction using psychological scales for anxiety and negative attitudes toward robots. *IEEE transactions on robotics*, Vol. 24, No. 2, pp. 442–451, 2008.

[75] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, Vol. 1, No. 1, pp. 71–81, 2009.

[76] Hiromichi Hori. *Psychological Measurement Scales II Interpersonal Relationships and Values: Understanding the Connection between Humans and Society.* SAIENSU-SHA Co.,Ltd., 2011. (Japanese).

[77] Noriko Onodera. Different answers for “hijou ni” and “kanarini”. a study of alternative expressions in international comparative surveys . The NHK monthly report on broadcast research, 2001. (Japanese).

[78] Robert J Grissom. Probability of the superior outcome of one treatment over another. *Journal of Applied Psychology*, Vol. 79, No. 2, p. 314, 1994.

[79] Volker W Rahlfs, Helmuth Zimmermann, and Kennedy R Lees. Effect size measures and their relationships in stroke studies. *Stroke*, Vol. 45, No. 2, pp. 627–633, 2014.

[80] Kanae Nakamura. Interactive negotiation of perspectives and adjustment of assertiveness in conversation. *The Japanese Journal of Language in Society*, Vol. 14, No. 1, pp. 33–47, 2011. (Japanese).

[81] Shuntaro Suzuki and Toshiaki Muramoto. Facilitating mutual interactions and objective evaluations increase member satisfaction with collaborative problem solving. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society*, Vol. 16, No. 1, pp. 39–50, 2009. (Japanese).

[82] Atsushi Mumura and Masafumi Hagiwara. Understanding presumption system from facial images. *IEEJ Transactions on Electronics, Information and Systems*, Vol. 120, No. 2, pp. 273–278, 2000. (Japanese).

Appendix

Glossary of Transcript Symbols

This paper transcribes the interactions obtained from the experiments. The symbols used in the transcription have the following meanings. These symbols are reprints of the contents of the following books [28, 23].

[	A left bracket indicates the point of overlap onset.
]	A right bracket indicates the point at which two overlapping utterances end, if they end simultaneously, or the point at which one of them ends in the course of the other. It also is used to parse out segments of overlapping utterances.
(0.0)	Numbers in parentheses indicate elapsed time by tenths of seconds.
(.)	A dot in parentheses indicates a brief interval ($\pm$ a tenth of a second) within or between utterances.
:	Colons indicate prolongation of the immediately prior sound. The longer the colon row, the longer the prolongation.

Acknowledgements

Here, I want to express my sincere gratitude for many people's cooperation and guidance in compiling this paper. I would like to appreciate them sincerely. I am deeply grateful to Professor Michio Okada, Department of Computer Science and Engineering, Toyohashi University of Technology. He helps to conduct this research and offer his kind guidance in preparing this study. He had supported me from the time I was a master's student, not only in the philosophy and ideology of research but also in various writing and presentation skills.

In conducting this research and writing the thesis, I received support from the Leading Program for Doctoral Education at the Toyohashi University of Technology. I would also like to express my sincerest gratitude to Professor Yugo Takeuchi of Shizuoka University, Professor Tetsuto Minami of the Toyohashi University of Technology, and Assistant Professor Marlena Fraune of New Mexico State University for their guidance on the philosophy, ideas, and methods of my research, and to have received their valuable opinions and important advice despite their busy schedules.

Regarding the examination of my doctoral degree, I appreciate the precious comments and helpful and essential suggestions from my group supervisor Dr. Tetsuhito Minami, Professor Shigeru Kuriyama of the same department, and Project Professor Hiromu Ishii of Research Administration Center.

Also, I wish to express my special thanks to Dr. Naoki Oshima of Senior Lecturer of Electronics-Inspired Interdisciplinary Research Institute, Toyohashi University of Technology, and assistant professor Koumei Hasegawa of the Department of Computer Science and Engineering, Toyohashi University of Technology, for their excellent advice in logically constructing this research.

Finally, I want to express my special thankfulness to the members of Prof. Okada's laboratory, Department of Computer Science and Engineering, Toyohashi University of Technology, for their cooperation in conducting this research and preparing this paper. It was an invaluable experience for me to discuss and work with them until late and prepare for the external presentation. I would also like to express my gratitude to the Toyohashi University of Technology staff and the Leading Program Graduate School for their essential support in carrying out my research. Furthermore, I sincerely thank my parents and family who have supported me so far.

This research was supported by the Ministry of Education, Culture, Sports, Science and Technology's Doctoral Education Program "Training Brain Information Architects." I want to express my gratitude to all of them.

January 2021 Yusaku Nishiwaki

Publication List

List of Papers with Referee's Review

[1] Yusaku Nishiwaki, Naoki Ohshima, Michio Okada, "Effects of Incomplete Utterance on Conversational Engagement in a Multiparty Conversation," Transactions of the Japanese Society for Artificial Intelligence, Volume 36, Issue 2, p. B-K44_1-12, 2021 (Japanese).

[2] Yusaku Nishiwaki, Sho Itashiki, Michio Okada, "Cooporative interactions generated by incompleteness in robot's utterance," The Transactions of Human Interface Society, Volume 21, Issue 1, pp. 1–12, 2019 (Japanese).

[3] Sho Itashiki, Yusaku Nishiwaki, Naoki Ohshima, Michio Okada, "Why we feel distant from smart speakers? The role of the Japanese pronoun that creates intimacy with robots," The Transactions of Human Interface Society, Volume 22, Issue 2, pp. 65–76, 2020 (Japanese).

List of Papers at International Conference with Referee's Review

[1] Yusaku Nishiwaki, Sho Itashiki, Nihan Karatas, and Michio Okada, "Cooperative Interactions Generated by Incompleteness in Robots' Utterance," In Proceedings of the 6th International Conference on Human-Agent Interac-

tion (HAI ’18), pp. 76–83, 2018

[2] Shinpei Onoda, Yusaku Nishiwaki, and Michio Okada, “Interaction Design and Field Study of a Forgetful Social Robot, “Talking-Bones”,” In Proceedings of the 7th International Conference on Human-Agent Interaction (HAI ’19), pp. 259–261, 2019

List of Papers at Domestic Conference

[1] Yusaku Nishiwaki, Michio Okada: “Potential of Incomplete Utterance: Attracting Hearer in Human-Robot Interaction,” SIG-SLUD-B509-06, pp. 31–36, 2018

[2] Yusaku Nishiwaki, Michio Okada: “The interaction and incompleteness of Incomplete Utterance Robots,” In Proceedings of the Society of Instrument and Control Engineers, p. SS13-13, 2018

[3] Yusaku Nishiwaki, Michio Okada: “Conversation Robot designed room for interaction,” In Proceedings of the 34th Annual Meeting of the Japanese Congnitive Science Society, pp.435–441, 2017

[4] Sho Itashiki, Yusaku Nishiwaki, Komei Hasegawa, Michio Okada: “Why we feel distant from smart speakers? The role of the Japanese pronoun that creates intimacy with robots,” Human Interface Symposium 2018 DVD-ROM Proceedings, pp. 84–89, 2018

[5] Shinpei Onoda, Yuji Yamamura, Masaki Ishikawa, Yusaku Nishiwaki, Michio Okada: “Umm, what was it?: A forgetful social robot “Talking-Bones”,” In Proceedings of the Entertainment Computing 2018, pp. 137–139, 2018

[6] Yuji Yamamura, Yusaku Nishiwaki, Shohei Hoshino, Michio Okada: “Talking-Bones: Interaction Design Inspired by Theory of Relational Development,”

Human Interface Symposium 2018 DVD-ROM Proceedings, pp.641-644, 2017

List of Awards

[1] Human Interface Society (HIS), Best Paper Award 2020,
Cooperative interactions generated by incompleteness in robot's utterance

[2] JSAI Incentive Award 2017,
Potential of Incomplete Utterance: Attracting Hearer in Human-Robot Interaction, March 2018.

[3] Unity Award, Entertainment Computing 2018,
Umm, what was it?: A forgetful social robot "Talking-Bones"