TR-IT-0091

Analysis of Gesture Behavior in a Multimedia/multimodal Interpreting Experiment ;Human vs Wizard of Oz Interpretation Method

Young-Duk Park, Kyung-Ho Loken-Kim,
Larel Fais and Suguru Mizunashi

March , 1995

Abstract

To investigate the user's linguistic and paralinguistic behaviors when communicating with a multimedia/multimodal interpreting system, we performed two experiments based on the human and the Wizard of Oz interpretation method. In this report, we mainly describe the results of subjects' gesture behaviors in the experiments. We analyzed the collected gestures from the usage ratio, the shapes, and their intention. The experimental results shows that the gestures are very useful to the not only to the users but also to the machine in the interpreting telecommunication. Therefore, we believe that the future machine interpreting cannot be completed without analyzing the characteristics of the multimedia/multimodal data and researching the recognition methodology which contains other modalities as well as speech. To take the research more concrete the more data collection for gestures in various interpreting situations will be necessary.

1. Introduction

Gestures are significant in multimedia/multimodal user interface. It eliminates the need for a lengthy definite description and simplifies the dialogues. [1][2]These also have a favorable influence upon interpreting machine because of reduction in the recognition workload. [3] Therefore, we performed an experiment to investigate the user's linguistic and paralinguistic behavior when dealing with a multimedia/multimodal system in an interpreting dialogue. In this report, we mainly describe the results of the subjects' gesture that is one of the paralinguistic behaviors in the human and the Wizard of Oz interpretation experiment. The major questions for the gesture through these experiment are

1) How often gestures are used in a session,

2) What kind of gestures are used in each interpreting mode,

3) For what purpose did subject used gestures,

4) Are there any common characteristics in subjects' gesture,

5) How verbal and gesture behaviors are interrelated,

6) What the implications of our findings are for a multimedia/multimodal interpreting system.

To investigate the above questions, we measured the number of turns that included gesture, the shapes of gesture, the usage patterns and intentions of gesture, and the timing relation between speech and gesture in the human and the machine interpreting case.

Our study in these experiments showed that the gesture usage of the machine interpreting was greater than that of the human interpreting case, and it is clear that the gestures are significant and very useful to the users and the interpreting machine. In this report, we describe the method of the experiment, the results of subjects' gestures, and our future directions.

2. Method

2.1 Tasks

Subjects did two tasks in the experiments: asking directions and making hotel reservation. In the asking directions task, North American native speakers of English, acting as clients, were asked to imagine that they were arriving for the first time in Kyoto Station and that they wanted to know the way to the International Conference Center. They called on the "Interpreting Service Center" and

talked to an Japanese-speaking conference agent who gave directions to the foreign visitors through an interpreter. In making hotel reservation, clients wanted to reserve a hotel room, and the agent gave information about hotel reservations through the interpreter. Each client performed above tasks in two communication environments: the telephone and the multimedia/multimodal, using the EMMI (Environment for MultiModal Interaction).

2.2 Subjects and equipment

2.2.1 Subjects

A total of 39 subjects (18 Japanese and 21 North American native speaker of English) took part in the experiments. Below table show the number of participants in each experiment.

Subjects	Human	Wizard	Total
Agents	5 *	10 *	15 *
Interpreters	2 *	1 * 1 #	3 * 1 #
Clients	10 #	10 #	20 #
Total	7 * 10 #	11 * 11 #	18 * 21 #

* : Japanese, # : North American native speaker of English

2.2.2 Equipment

1) Three telephone for voice communication
2) Two NeXT Workstations for multimedia multimodal communication (client and agent)
3) Two Sun Workstation for multimedia and multimodal communication (interpreters)
4) Four Headset for recording the subjects' speech
5) Two Digital Audio Tape Recorders for recording the subjects' speech
6) Three Amplifier for the speech amplification
7) A BOSS Super Effects Processor for the distortion of the interpreters' speech
8) A Down converter for converting the video signal of the interpreter's screen
9) One VCR for recording the interpreter's screen

2.3 Experimental Configuration

Fig. 1 illustrates the experimental configuration for the data collection. The interpreter was physically separated from the agent and the client, and sound-absorbing partitions were placed between the agent and the client places to prevent transmission of the participants' voices. For details of the system

configuration in human interpreting experiment see [4]. We selected the Wizard of Oz technique as the simulation method for the machine interpreting experiment. The basic idea behind the Wizard of Oz method is very simple. A human plays the role of the machine in the simulation, but subject think they are communicating with machine [5]. In this report, we mainly describe system configuration in the machine interpreting experiment, which is different from that of the previous human interpreting experiment.

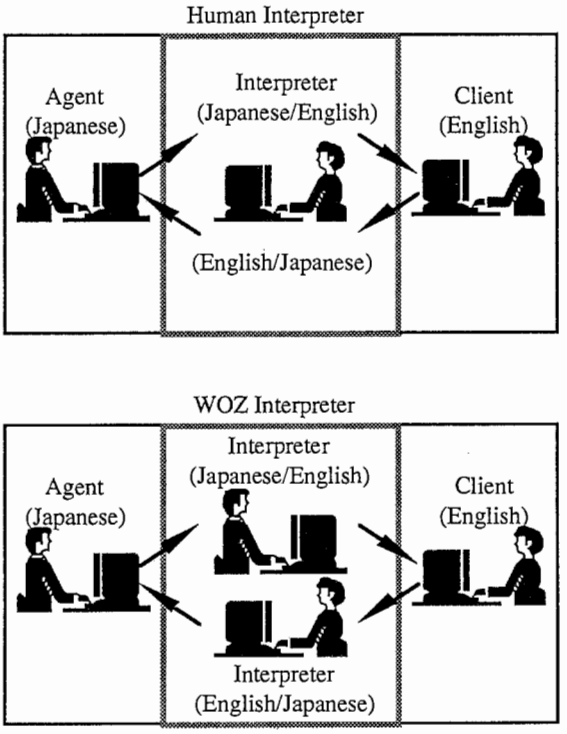

Fig.1 Experimental Configuration

First, we used two native speakers as the wizard to interpret with native utterance. One native speaker of Japanese acted as a wizard, translating the English into Japanese, while another native speaker of North American English translated the Japanese into English. Therefore, we added on more SUN sparc station to support two wizard interpreters at the same time. These wizards modulated their speech to be as monotonic and syllable-timed as possible, simulating the machine's speech. Also the wizards' voice were distorted by voice effector to make the subjects believe that they were communicating with machine interpreting system.

Second, we changed interpreters' gesture handling scheme. If the client (agent) dragged some point, that position transmitted to the English/Japanese wizard interpreter (Japanese/English wizard interpreter) only. After that step, the wizard interpreter clicked automatic transmission button, and the client's or agent's dragging position with the wizards' interpreted speech was transmitted to other communication partner. However, the interpreter's dragging was only displayed on the agent's or client's computer screen with the interpreted speech. Also the interpreters could transmit his/her dragging information to the agent and the client by clicking the manual transmission button.

Lastly, the arbitrator controls the subjects' speech problem. If the agent or the client spoke too long sentence, too many interjection words in a sentence, too fast, and too many false starts, to recognize by machine, arbitrator instructed the wizard interpreters to ask the subjects to repeat an utterance. These utterances by the wizards were speaks "Please repeat", "Please speak slowly". Example of these procedures for the gesture included dialogues are attached at the appendix.

2.4 Collected data

1) Client's speech
2) Agent's speech
3) Interpreter's speech
4) Computer screen of the English/Japanese WOZ interpreter
5) Computer screen of the agents
6) Subject's suggestions in the pre and post experiment interview
7) Subject's conversational behavior during experiment

2.5 Experiment sequences

Clients were divided into two groups. "A" group began with the telephone and "B" group began in the reverse sequence to decrease the client's learning effect on the experiment. The participants were given instructions about the background of the experiment, the tasks of the experiment, and the operation of the system. In the Wizard of Oz experiment, all subjects were told that their speech would be interpreted by machine which had been developed at ATR. After the introduction, participants practiced operating the system. Then the participants performed the actual experiment with a one-page brochure about the conference, and clients asked the two tasks in random order. We did not give any prepared scenarios to help the subjects' conversation. Also, we asked them to talk naturally, and not to be constrained when they had a problem during the actual experiment. After the

experiment, we asked some questions to learn the subjects' impressions and opinions about the experiment. In the Wizard of Oz experiment, no subject indicated any doubt that his or her speech was being interpreted by machine.

3. Results

From the transcription data and the videotape that recorded the experiment, we observed the subjects' speech and gesture behaviors in each interpretation mode (9 dialogue sessions in human interpreting mode and 10 dialogue sessions in wizard interpreting mode). Then we measured the number of turns, the number of gesture included turns, and gesture shapes and their intention that each participant used in a session. On the basis of the measured data, we analyzed the common characteristics of the subjects' gesture behaviors in the multimedia and multimodal interpreting conversations.

3.1 Gesture Usage

To analyze the subjects' gesture usage, first, we measured the total number of turns in each interpreting mode, the total number of turns in each task, and the total number of turns that included gestures in each task.

Table 1 and 2 show the gesture usage ratio of the collected gestures in each interpreting mode. We computed the usage ratio of the gesture by dividing the total number of collected gestures by the total number of turns. Some major results of the gesture usage are as follows.

Table 1. Gesture Usage in Human Interpreting Mode

Human	Agent	Client	Total
No. of Direction_Turn	274	288	562
No. of Direction_Gesture	36	10	46
Direction_Gesture %	13%	3%	8%
No. of Hotel_Turn	372	358	730
No. of Hotel_Gesture	6	0	6
Hotel_Gesture %	2%	0%	1%
No. of Total_Turn	646	646	1292
No. of Total_Gesture	42	10	52
Total_Gesture %	7%	2%	4%

Table 2. Gesture Usage in Wizard Interpreting Mode

Human	Agent	Client	Total
No. of Direction_Turn	196	178	374
No. of Direction_Gesture	47	22	69
Direction_Gesture %	24%	12%	18%
No. of Hotel_Turn	323	304	627
No. of Hotel_Gesture	8	5	13
Hotel_Gesture %	2%	2%	2%
No. of Total_Turn	519	482	1001
No. of Total_Gesture	55	27	82
Total_Gesture %	11%	6%	8%

1) The average number of turns for human interpreter experiment (agent: 71.8 turns, client: 71.8 turns) was greater than the wizard case (agent: 51.9 turns, client: 48.2 turns).

2) The usage ratio of the gesture for wizard (8%) was greater than the human interpreter case (4%). The reason for this is the system discredit of the subjects. Therefore, they used more gestures in the WOZ case to upgrade other communication partner's understanding.

3) The gestures occurred mainly in direction finding task (direction finding task: 8%, hotel reservation: 1% in human interpreting mode). The difference was quite significant in the wizard interpreting mode (direction finding task: 18%, hotel reservation: 2% in wizard interpreting mode), because the major modality in hotel reservation task was not a gesture but a typing to fill in the reservation form. In some case, agent used gesture in the hotel reservation task, when client asked location of the hotel.

4) The agent used more gestures than client, because the agent had to explain direction or location on the map in detail with gestures (agent: 24%, client: 12% in the wizard interpreting mode). The clients, on the other hand, preferred the speech to the gestures, because they thought that the former provides a more convenient access.

3.2 Gesture Classification and Intention

3.2.1 Gesture Classification

We observed experiments' videotapes, and classified subject's gestures into five classes. These were

1) circling; draw a circle-like shape in the area of map,

2) dragging; draw a line-like shape in the map,

3) pointing; click a point in the map,

4) spiraling; draw a spiral-like shape in the area of map, and

5) checking; check a point in the map as shown in Fig. 2.

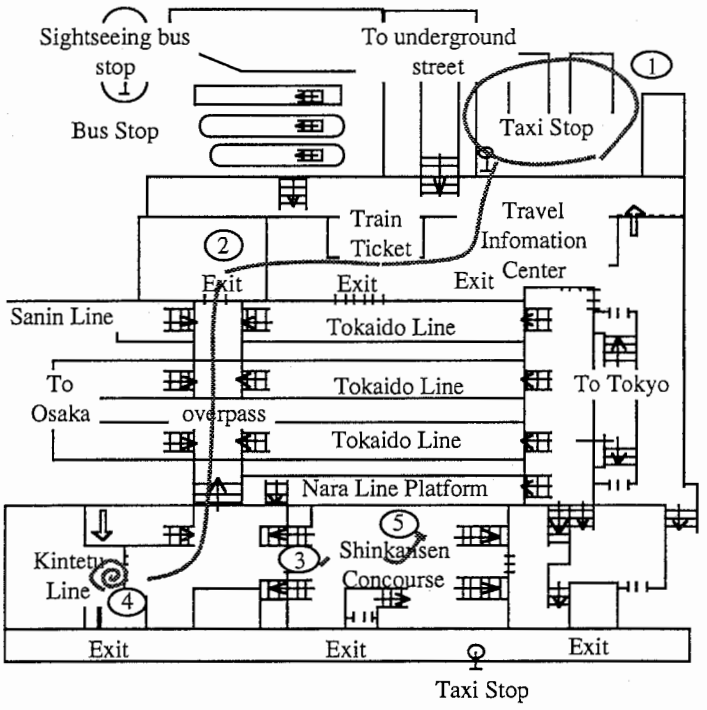

Fig. 2 Example Shapes of the Collected Gestures

Table 2 and 4 show the gesture classification of the collected gestures. Some major results of the gesture classification are as follows.

1) There was no significant difference between human and wizard interpreting mode in the gesture classification. Table 2 and 4 show the gesture classification of the collected gestures.

2) The subjects' main gestures were the circling and the dragging gesture (90% and 84% of gesture in human and wizard interpreting mode respectively).

3) There was some difference between agents' and the clients' gesture usage. The agents' used mainly the circling and the dragging gesture (96% and 99% of gesture in human and wizard interpreting

mode respectively). However, client used more various gestures than agent. These results may be occurred not because of the nationality but because of the agent's and the client's role in the task.

Table 3 Gesture Classification in Human Interpreting Mode

Human	Circling	Dragging	Pointing	Spiraling	Checking	Total
Agent	18 40%	25 56%	0 0%	1 2%	1 2%	45 100%
Client	4 33%	4 33%	3 25%	1 8%	0 0%	12 100%
Total	22 39%	29 51%	3 5%	2 4%	1 2%	57 100%

Table 4 Gesture Classification in Wizard Interpreting Mode

Wizard	Circling	Dragging	Pointing	Spiraling	Checking	Total
Agent	42 46%	48 53%	0 0%	0 0%	1 1%	91 100%
Client	3 9%	13 37%	11 31%	5 14%	3 9%	35 100%
Total	45 36%	61 48%	11 9%	5 4%	4 3%	126 100%

3.2.2 Gesture Intention and their Characteristics

The circling gestures were mainly used for location-indications (100% of circling gesture are in this case). Table 5 and 6 show the intention of gesture in two interpreting mode. Therefore, there were many deictic words and proper nouns in the duration of the circling gesture. The examples of deictic word and proper noun are kochira (here), kono (this), koko (here), name of hotel and subway station and so on. As shown in the below e.g. 1 and e.g. 2 sentences, the circling gestures occurred in the middle of the sentence where the deictic words and proper nouns were located. The brackets in the example sentences denote the start and end time of gesture. Also the circling gestures started with deictic words or proper noun, and included the back words of the deictic words because of the drawing time of circling gestures .

e.g. 1; [このあたり] ですね。

e.g. 2; 新幹線は[ここです]。

There were many kinds of parts in a sentence with dragging gesture, because the main usage of the dragging gesture was direction expression (86% in human interpreting mode, and 97% in wizard interpreting mode). Also, some subjects used dragging gestures as a referent-identification while they dragged under an object. The dragging gesture and the scrambling gesture included many words in contrast to the circling gestures (e.g. 3 and e.g. 4), and their start time was random. Also the drawing time of dragging gesture was longer than circling gesture.

e.g. 3; [まっすぐ歩いてすぐ右のところ]です。

e.g. 4; [この地下鉄乗り場てなっていますよね、こちら]。

The pointing gestures, the spiraling gestures and the checking gestures were mainly used by the clients for referent-identification and their location answer.

Table 5 Intention of Gesture in Human Interpreting Mode

Human	Circling	Dragging	Pointing	Spiraling	Checking	Total
Location-indication	22 100%	4 14%	3 100%	2 100%	1 100%	32 56%
Show the way	0 0%	25 86%	0 0%	0 0%	0 0%	25 44%
Total	22	29	3	2	1	57

Table 6 Intention of Gesture in Wizard Interpreting Mode

Wizard	Circling	Dragging	Pointing	Spiraling	Checking	Total
Location-indication	45 100%	2 3%	9 82%	5 100%	3 75%	64 51%
Show the way	0 0%	59 97%	2 18%	0 0%	1 25%	62 49%
Total	45	61	11	5	4	126

3.3 Gesture Behavior after "Please Repeat"

If the agent or the client spoke too long sentence, too many interjection words in a sentence, too fast, and too many false starts, to recognize by machine, WOZ interpreter spoke "Please repeat", "Please speak slowly" to the agent or the client. After that message, there were six kinds of subject's responses. In the below case examples, <C@ ..... @C>, <D@ ..... @D>, <P@ ..... @P>, <M@ ..... @M>, <S@ ..... @S> denotes circling, dragging, pointing, marking, spiraling gesture respectively.

Examples of these dialogues are attached at the appendix. These gesture behaviors after "Please repeat" were occurred randomly based on the dialogue situations and the subject's personality. Therefore, it is very difficult to find out the common rule that explain the gesture behaviors after "Please repeat".

1) Case 1: No gesture after "Please repeat"

In this case subjects used no gesture after "Please repeat" even though they used gestures before "Please repeat".

A: 二つ目の、<C@蹴上駅で@C>、降りてください。
I: もう一度、お願いします。
C: 二つ目の駅、蹴上駅で、降りてください。

2) Case 2: Same gestures as before

In this case, subjects used same shape of gestures that was drawn before "Please repeat". However, there was some difference in the timing and word in the gestures.

A: I'm at thi Kintetsu Line. <S@I'm putting a mark where I'm standing@S>
I: Please repeat
C: I'm standing at <S@thi mark@S> near thi Kintetsu Line

A: <C@あなたは今ここにいます@C>。
I: ゆっくり、お願いします。
C: <C@あなたは今ここですね@C>。

3) Case 3: Simplified gesture

In this case, subjects used simplified gestures that are compare with before "Please repeat".

A: /ls/ Yes, I see <D@the map, so you want me to follow thi arrow @D>into <D@(thi) this taxi stop?@D>
I: Please repeat
C: I can follow <D@this line @D>to the taxi stop?

4) Case 4: Different gestures from before

In this case, subjects used different kinds of gestures after "Please repeat".

A: タクシー乗り場がすぐ見えると思いますから<C@ここです@C>。
I: ゆっくり、お願いします。
C: タクシー乗り場がありますから<S@ここから乗ってください@S>

5) Case 5: New gestures after "Please repeat"

In this case, subjects used new gestures even though they did not use gesture before "please repeat".

A: この、赤い、(あ、じゃなかった)ここです。
I: もう一度、お願いします。
C: <C@ここです@C>。

6) Case 6: Changed the topic

In this case, subjects changed the topic, and moved to other question .

A: ここにタクシー乗り場があります。<D@ @D>
I: ゆっくり、お願いします。
C: 地図は見えますか。

4. Conclusion and Future Direction

In multimedia and multimodal environment, gestures facilitates the understanding of the communication partner. From the major results of this experiment, we found the basic characteristics and handling scheme of the subject's gestures in an interpreting dialogue.

First, the gesture usage ratio between two interpreting conditions was different. The usage ratio of the gesture in wizard was greater than the human interpreting case. The reason for this is the system discredit of the subject. Therefore, they used more gestures in the wizard case to upgrade other communication partner's understanding.

Second, the gesture usage ratio between the agent and the client. The agent used more gestures than the client because of the their role in the given task. These results may be occurred not because of the nationality but because of the agent's and the client's role in the task.

Third, there was no significant difference between two interpreting conditions in the subject's gesture shape and their intention. The subject's major gestures were the circling and the dragging gesture to express their intention in direction finding task. In case of the circling gestures, it is easy to find out what the gesture means, because it was mainly used to express the location-indication. Also the timing between referent-identification words and circling gestures was clear. However, the dragging gestures were used for various direction expressions, and the timing between speech and gestures was very random. This makes it very hard to recognize the exact and common meaning of the dragging gesture.

From the experiment results, it is clear that multimedia and multimodal communication is helpful not only to the users but to the interpreting machine. The results shows that the effect of multimedia/multimodal in machine interpreting is higher than that of human interpreting cases. Especially, the importance and benefit of the gesture in machine interpreting are not only the reduction of the speech recognition workload in system but also the upgrade of other communication partner's understanding. These also provide a positive effect on the dialogue management of the system. So, in machine interpreting system, it is important to examine the optimal gesture handling scheme to improve the system's speech processing capability.

As a conclusion, there's no need to discuss whether multimedia and multimodal effects on the machine interpreting field or not any more. We believe that the future machine interpreting can not be completed without analyzing the characteristics of the multimedia and multimodal data and researching the recognition methodology which contains other modalities as well as speech. Also, it is important to continue the following research items to concrete our finding about the gesture behavior in the interpreting field.

1) Data collection on the various kind of task

2) Investigation of the task- and map- dependence of the gesture behavior

3) Relation between the gesture and the nationality

4) Synchronization and representation scheme of the gesture behavior after "Please repeat"

References

1. Jeannette G. Neal and Stuart C. Shapiro, "Intelligent Multimedia Technology", Addison Wesley, 1991.

2. Meera M. Blattner and Roger B. Dannenberg, "Multimedia Interface Design", ACM Press, 1992.

3. Fumihiro Yato, Tsuyoshi Morimoto, Yasuhiro Yamazaki, and Akira Kurematsu, "Important Issues for Automatic Interpreting Telephone Technologies", Proc. of EuroSpeech-93, Berlin, Sep., 1993.

4. Kyung-ho Loken-Kim, Fumihiro Yato, and Tsuyoshi Morimoto, "A Simulation Environment for Multi-Modal Interpreting Telecommunication", Proc. of SIG-AVM, 94-AVM-4, 1994.

5. Norman M. Fraser and G. Nigel Gilbert, "Simulating Speech Systems", Academic Press Limited, 1991.

6. Young-Duk Park, Kyung-ho Loken-Kim, Fumihiro Yato, Laurel Fais and Tsuyoshi Morimoto, " Analysis of Telephone and Multimedia/Multimodal Conversations in an Interpreting Experiment: Linguistic and Paralinguistic Behaviors", Proc. of JW-MMC '94, Oct., 1994.

7. Young-Duk Park, Kyung-ho Loken-Kim, Fumihiro Yato, "Analysis of Gestures in a Multimedia/Multimodal Interpreting Experiment", IEICE Conference, March 27-30, Fukuoka, Japan.

8. Laurel Fais, Kyung-ho Loken-Kim, Young-Duk Park, " Speakers' Response to Requests for Repetition in a Multimedia Language Processing Environment", International Conference Cooperative Multimodal Communication, Theory and Applications CMC/95, Eindhoven, May 24-26, 1995.

Appendix

Gesture Behavior after "Please Repear" in WOZ experiment

1. Transcription Method of Speech Part

1) Interjection speech

[...] : denotes interjection speech

2) Simultaneous speech

- Speech uttered by one person simultaneously with speech uttered by another

I: {C: ... :C} : denotes simultaneous speech between I and C.

C: {I: ... :I}

- Two pieces of simultaneous speech occur together

A: {I: .... {C: .... :C} :I}

I: {A: .... {C: ... :C} ... :A}

C: {A,I: ... :A,I}

3) False start

(..) : denotes false start

4) Non-language sound

/ ... / : denotes non-language sound

5) False dialogue

! .. ! : denotes false dialogue

2. Transcription Method of Gesture Part

1) Circling gesture

<C@ ..... @C> : denotes circling gesture

2) Dragging gesture

<D@ ..... @D> : denotes dragging gesture

3) Pointing gesture

<P@ ..... @P> : denotes pointing gesture

4) Checking gesture

<M@ ..... @M> : denotes checking gesture

5) Spiraling gesture

<S@ ..... @S> : denotes spiraling gesture

Dialogue 1;

A : <D@バス乗り場に@D@>行かれたら、<D@そのまま@D>[えー]五番系統のバスで、(す)三条京阪の、停留所<teiryuujo>まで行ってください。

I : もう一度、お願いします。

A : バス乗り場で、五番のバスに乗ってもらって、三条京阪まで、行ってください。

Dialogue 2;

A : [あ]それから、[えー]京津<keishin>線の、浜大津駅行き<iki>の電車に乗ってもらって、二つ目の、<C@蹴上駅で@C>、降りてください。

I : もう一度、お願いします。

A : 京津<keishin>線の、浜大津行き<iki>の電車に乗ってもらって、二つ目の駅、蹴上駅で、降りてください。

Dialogue 3;

C : I'm at thi Kintetsu <S@Line. I'm putting a mark where I'm standing@S>

IE : Please repeat

C : I'm standing at <S@thi mark@S> near thi Kintetsu Line

Dialogue 4;

A : はい、タクシーでしたら、<D@近鉄線から@D> <D@真っ直ぐ行っていただいて、JRの{I1;チケット売場がここ@D>に;I1}。

I : {A1;ゆっくり;A1}お願いします。

A : JRのチケット売場の<D@ところを、通り過ぎ@D>{I2;て;I2}。

I : {A2;認;A2}識できません。¦ゆっくりお願いします。

A : はい。<C@近鉄の乗り場から@C>。

IE : From Kintetsu platform

A : <D@ @D>わたしが今書いた線のとおりに歩いてください。ここにタクシー乗り場があります。<D@ @D>

I : ゆっくりお願いします。

A : 地図は見えますか。

Dialogue 5;

C : /ls/ Yes, I see <D@the map, so you want me to follow thi arrow @D>into <D@(thi) this taxi stop?@D>

IE : Please repeat

C : I can follow <D@this line @D>to the taxi stop?

Dialogue 6;

A : <C@ @C>この丸を書いたところがタクシー乗り場です。

I : ゆっくりお願いします。

A : <C@この場所がタクシー乗り場です@C>。

Dialogue 7;

A : [えー]その前に、タクシー乗り場がありますから、そこで、タクシーに乗ってください。<D@ @D>

I : ゆっくり、お願いします。

A : あなたの前に、タクシー乗り場がありますから、そこでタクシーに、乗ってください。

Dialogue 8;

C : [hum] /ls/ OK, I see <D@which @D>taxi stand. [um] (what) Which building do I say to the taxi driver?

IE : Please repeat

C : /laugh/ [um] /ls/ When I'm in the taxi, (how) [ah] where does the taxi go to?

Dialogue 9;
A : <C@あなたは今ここにいます@C>。

I : ゆっくり、お願いします。

A : <C@あなたは今ここですね@C>。

Dialogue 10;
A : 分かりました。<D@それでは、出口を、八条口に出てください@D>。<C@タクシー乗り場は、ここにあります@C>。

I : もう一度、お願いします。

A : 分かりました。<D@それでは、八条口に出てください。タクシー乗り場は@D>、<C@ここにあります@C>。

Dialogue 11;
C : [ah] I'm at thi [um] <P@here @P> /laugh/

IE : Please repeat

C : <P@Here@P>

Dialogue 12;
A : (それでは、え、今いる、ところ、を真っ直ぐ)[あ]今いるところから、タクシー乗り場がすぐ見えると思いますから、{I;<C@ここです@C>;I}。

I : {A;ゆっくり;A}、お願いします。

A : すぐ前に、タクシー乗り場がありますから、<S@[え]ここから、タクシーに乗ってください@S>。

Dialogue 13;

A : では、近鉄線から、<D@正面の改札を出て@D>、そして、<D@橋を渡って@D>、<D@正面の@D>、<D@出口、から、真っ直ぐ@D>、行ってください。そうすると、<D@右手に@D>、タクシー乗り場、<D@あるいは、左にバス乗り場が@D>、あります。よろしいか。

I : もう一度、お願いします。

A : [お]<D@前の改札を、出て@D>、そして、橋を渡って、正面出口に出てください。

Dialogue 14;

A : この、赤い、(あ、じゃなかった)ここです。

I : もう一度、お願いします。

A : <C@ここです@C>。

Dialogue 15;

A : <D@そして、京{I2;津<keishin>線;I2}@D>、<D@京津<keishin>線に@D>、<D@乗り換えて@D>、今度は、<D@ここまで、行ってください@D>。そうすると、<C@目の前に@C>、あります。

IE : {A2;Then;A2}

I : もう一度、お願いします。

A : <D@ここで乗り換えて、そして、一つ、二つ目の、駅で、降りて、ください@D>。<D@すぐ前に、あります@D>。