

TR-IT-0177

Pilot Experiment on the Understandability of Interactive Disambiguation Dialogues

Hervé BLANCHON
Laurel FAIS

1996.7.17

We report on a pilot experiment that was carried out at ATR-ITL on the topic of interactive disambiguation and more precisely, on the understandability of interactive disambiguation questions.

Two classes of questions (human-like and machine-like) were proposed using a textual modality. The human-like questions were worded in natural, conversational language, and the machine-like questions were the ones that we would have been able to produce using the disambiguation methodology proposed in [1-3]. The answers to the questions were of two kinds, "easy" (corresponding to the most frequent or most natural interpretation of a given sequence of words) and "hard" (corresponding to a very infrequent or unnatural interpretation of a given sequence of words). We had two groups of subjects, each subject participating in only one setting.

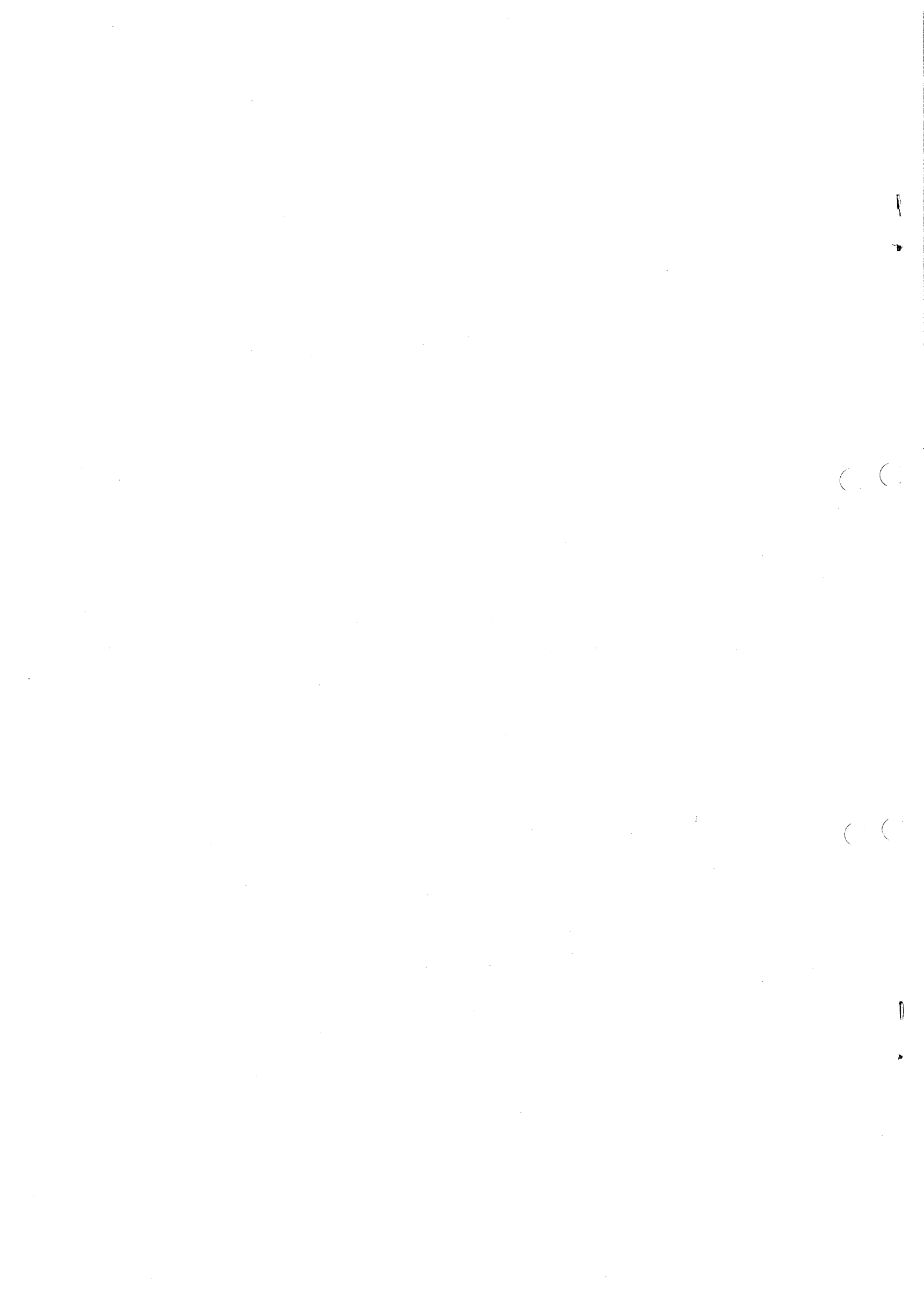
A text containing 35 questions (21 hard, and 14 easy) was read aloud by the subjects, who then answered the questions. The analysis of the results shows that the experiment was not well designed. The text was too difficult and unnatural to enable the subjects to focus their attention correctly; they performed poorly with each one of the two sets of questions.

ATR – Interpreting Telecommunications Research Laboratories
2-2 Hikari-dai, Seika-cho, Soraku-gun, Kyoto 619-02, Japan

© 1995 by ATR Interpreting Telecommunications Research Laboratories

Contents

INTRODUCTION	1
I. SETTING	1
I.1. EXPERIMENTAL CONDITIONS	1
I.1.1 Presentation of the questions by the experimenter	2
I.1.2 Presentation of the questions to the subjects	3
I.2. EXPERIMENTAL SETTINGS	3
I.2.1 Text	3
I.2.2 Questions	4
II. RESULTS	5
III. ANALYSIS	5
III.1. STATISTICAL ANALYSIS	5
III.1.1. General Analysis: Two Approaches	5
III.1.2. Separated Statistical Approach	6
III.2. BEHAVIORS	8
III.3. QUESTIONNAIRES	8
III.3.1. About each kind of ambiguity	8
III.3.2. About general feelings & proposals	10
CONCLUSION	11
Lessons learned	11
Implications for the design of the next experiment	11
REFERENCES	12
APPENDIX I: TEXT PRESENTED TO THE SUBJECTS	13
APPENDIX II: QUESTIONS PRESENTED TO THE SUBJECTS	15
APPENDIX III: RESULTS IN THE HUMAN-LIKE SETTING	17
APPENDIX VI: RESULTS IN THE MACHINE-LIKE SETTING	18



Figures & Tables

Figure 1. Experimental setting.....	1
Figure 2. The warm-up text with 4 questions	2
Figure 3. A control card for the experimenter	2
Figure 4. The subject screen with a dialogue box.....	3
Table 1. Classification & definition of the ambiguity classes	4
Figure 5. Average number of correct answers per ambiguity category,.....	6
Figure 6. Average number of correct answers per ambiguity category,.....	7
Figure 7. Average number of incorrect answers per ambiguity category,.....	7
Figure 8. Average number of “no answer” responses per ambiguity category,	8
Table 2. Subjects’ evaluation of ease of solving each class of ambiguity	9
Table 3. Subjects’ feelings about the experiment	10
Figure 9. The text read by the subjects.	14
Table 4. Questions asked to the subjects in both settings	16
Table 5. Subjects’ answers to the human-like dialogues	17
Table 6. Subjects’ answers to the machine-like dialogues	18

Introduction

A methodology to produce disambiguation questions automatically and to present them to a user of a system using natural language (spoken or written) as an input modality has been proposed [2]. It is important to evaluate the understandability of the questions to be produced with this methodology. During summer 1995, a pilot experiment was carried out at ATR-ITL to test the understandability of the disambiguation questions that may be produced using this methodology. The results of this experiment are reported in this technical report.

We designed a text that contained a set of 35 ambiguous sentences. Among those 35 sentences 14 had easy interpretations and 21 had hard interpretations. By "easy" interpretation we mean the sense which pops up first in someone's mind for a given sequence of words; the "hard" one is the less immediate interpretation. For example, the easy interpretation of the sentence "I would like to check in to the hotel" is "I would like to register at the hotel;" the hard one is "I would like to investigate the hotel."

The subjects were given a hard copy of the text beforehand. They were asked to read it for themselves trying to get as clear an understanding as possible. Afterward, they were invited to read the text aloud slowly and carefully, trying to understand and answer the questions whenever they were asked.

We report the results of this experiment in three parts. The first part describes the actual setting we used for this experiment. The second part gives the results as we got them. Finally, in the third part we analyze the results. In the conclusion we give some summing up comments and draw some implications.

I. Setting

I.1. Experimental conditions

For this experiment, the workstations of the subject and the experimenter were back to back. They were not communicating.

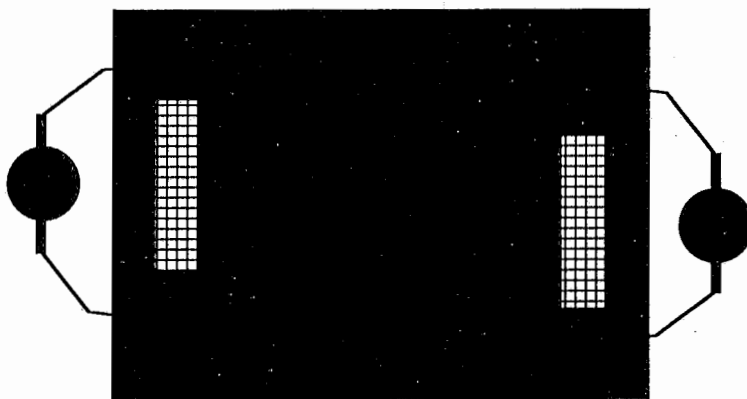


Figure 1. Experimental setting.

The subjects were first asked to read beforehand, to themselves, the text they were to answer questions about. The text was then displayed inside a text window on the subject's workstation screen. The subjects were asked to read aloud slowly and carefully and pause between the sentences. The scrolling of the text window was controlled by the subject.

For the subjects to get used to the setting, a set of warm-up sentences was given to them as shown in figure 2.

I bought some apples yesterday. **They are cooking apples.** That hospital is run by old, conservative doctors who don't know the latest hospital techniques. **They need some new blood.** There were two girls walking near the pond. One girl had binoculars. The other had a camera. **I saw the girl with the binoculars.** I lost a leg in a car accident, so I called a company to find out about getting an artificial leg. Boy, were they fast. When I opened the door five minutes later, **there stood a salesman with three artificial legs.**

Figure 2. The warm-up text with 4 questions

The experimenter controlled the presentation of the questions to the subjects.

1.1.1 Presentation of the questions by the experimenter

The presentation of the questions was controlled through an HyperCard stack. There was one stack per subject. All the cards in this stack had the same presentation (Fig. 3). On the top part of each card was displayed a part of the text with the ambiguous sentences bolded. On the lower part was:

- a set of buttons (e.g., **traveling**, **exitout**, **trainticket**), each one activating the display of a dialogue box on the subject screen.
- a set of textual fields, used to record automatically the answer of the subject to the related question,
- a set of two buttons allowing the experimenter to go to the previous or next card (the arrows).

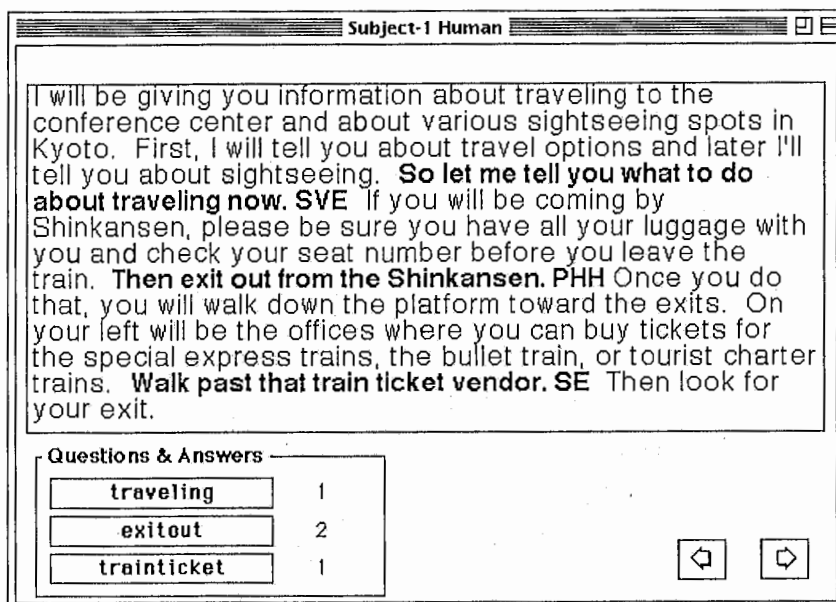


Figure 3. A control card for the experimenter

On this card, which corresponds to the beginning of the text for subject-1 with the human-like set of questions, we can see that the subject has chosen: answer 1 for the question about **traveling**, answer 2 for the question about **exitout** and answer 1 for the question about **trainticket**.

On the first card was a prologue button used to play a speech synthesized message asking the subject to start reading slowly and carefully. On the last card was an epilogue button used to play an acknowledgment message.

1.1.2 Presentation of the questions to the subjects

The questions were displayed in a dialogue box presenting an ambiguous utterance and two possible interpretations (Fig. 4). The dialogue box appeared in the middle of the textual window. The special feature of this dialogue box was that there was no choice selected by default. We chose to do that so as not to lead the subjects toward a particular item.

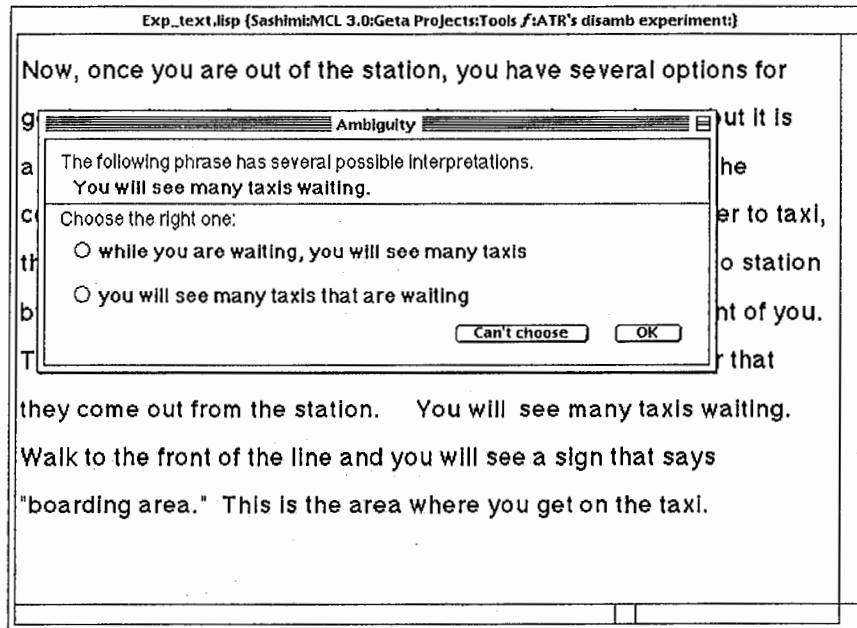


Figure 4. The subject screen with a dialogue box

1.2. Experimental settings

Two classes of dialogues were used. This required two groups of subjects. The text and the ambiguities to be solved were the same for each group.

1.2.1 Text

The text to be read was made up of two different stories (Appendix 1). It contained 35 ambiguous sentences, five in each category of ambiguities described in the table below. The ambiguities were selected from the corpus of spontaneous speech collected in EMMI [4-7]. Each ambiguity is in boldface. In each category there were two sentences with the easy interpretation and three sentences with the hard one.

The codes found in the text (Table 1) give the abbreviation for the ambiguity type, followed by H for "hard" interpretation or E for "easy" interpretation. The abbreviations can be described as follows:

D	Decoration	Problem with the function of a phrase as in "please exit by Exit 14." – exit through Exit 14 – exit in proximity to Exit 14.
PH	Phrasal Verb	Interpretation of a phrasal verb as in "Go over it carefully." – pass over, – review.
CO	Coordination	Interpretation of conjoined phrases as in "Western bed and bath." – western bath, – not western bath.
S	Subordination without the verb	Subordination of phrases not involving a verbal phrase as in "a train ticket vendor." – train ticket, – ticket vendor.
P	Polysemy	A word with several senses such as "story." – narrative, – floor.
SV	Subordination with the Verb	The complementation of the verb is not certain as in "You will see many taxis waiting." – the taxis are waiting, – You are waiting.
SC	Syntactic class	Ambiguity of syntactic class as in "You can take a bus or taxi." – You can taxi, – You can take a taxi.

Table 1. Classification & definition of the ambiguity classes

1.2.2 Questions

We designed two sets of questions: human-like ones and machine-like ones.

The human-like ones were wordy and gave a plain and natural paraphrasing of the meaning of the ambiguous utterances.

The machine-like ones were those which could be produced by a system using the automatic disambiguation method proposed in [2]. Thus they are more compact, somewhat more abstract and less descriptive.

In Appendix II we give the labeling of the questions in the human and the machine like settings. The ambiguous utterances are in bold type, and precede the two proposed interpretations.

II. Results

The Tables giving the answers we got for each setting and for each subject are given in Appendix III (in the human-like setting) and Appendix IV (in the machine-like setting).

III. Analysis

The analysis of the collected information is divided into three parts: statistical analysis, behavioral analysis, and the post-experiment questionnaire analysis.

III.1. Statistical analysis

We conducted two analyses of the results. In the first one we went across every response to every question. In the second one, we grouped responses for each class.

III.1.1. General Analysis: Two Approaches

Within each category, analyzing all of the answers together required us to assign values to those answers. If the only possible answers had been "right" and "wrong," that is, if the experiment had had a forced choice design, we could have given "right" the value of "1" and "wrong" the value of "0" and the subsequent analysis would have been fairly clear. But we were faced with the problem of giving "no answer" some value.

In order to be able to analyze the results over all the answers given, then, we calculated the results in two ways: once with "0.5" as the value for "none" and once with no value for "none." The two sets of results differed only slightly. We interpret the results of the analysis with no value for "no answer" as focusing on how well subjects could succeed in the task, and thus, how well they could deal with the dialogue boxes. The results from assigning 0.5 as the value for "no answer" gives subjects credit for understanding the text, but doesn't reveal much about how they dealt with the boxes.

1.1.1. Results for Dialogue Type

The only ambiguity class for which there was a clear significant difference due to dialogue type was Syntactic Class. There was a weak significant difference for Decoration also in the case where we disregarded the non-answers. This was not significant where we counted the non-answers as 0.5.

In both approaches, Polysemy was strongly *not* significant. Where non-answers were given 0.5, Coordination and Subordination were also strongly *not* significant.

1.1.2. Results for Questions

There was also a significant difference in the answers depending upon which question was being answered. It was significant for all categories except Subordination in the case where "no answer" had no value. In the case where "no answer" had 0.5 value, all were significant except Decoration.

1.1.3. Results for Questions depending upon Dialogue Type

There was an interaction between the questions and the dialogue type as well. That is, there was a significant difference in the way that questions were answered depending upon the dialogue type, for some classes. This was significant for Syntactic Class and Polysemy (in both approaches to analysis).

1.1.4. Summary of General Approach

We observed a real difference in dialogue type for Syntactic Class, and a weak difference for Decoration. For Polysemy, especially, but also possibly for Coordination and Subordination as well, there was strong *non*-significance. So the dialogue type really made no difference for these ambiguity classes.

As for the effect of the questions, it is clear that the nature of each question had an effect on how it was answered. This was true for virtually every category. This difference interacted with the dialogue difference for Syntactic Class and Polysemy. In the case of Syntactic class, the effects of dialogue and question were both so strong that the combined effect was also significant. The case of Polysemy is different; there was a strong *non*-effect of dialogue type for Polysemy. But the effect of question was strong enough to make the interaction of dialogue type and question significant.

III.1.2. Separated Statistical Approach

Instead of analyzing the results all together, as we did in our general approach, in this second approach, we separated the results for correct answers, incorrect answers and no answer. While this makes it more difficult to understand the results for each category overall, it does allow us to look at how accurately subjects answered questions (right and wrong answers) and how confident they felt about answering the questions (no answers).

Recall that each class of ambiguity was represented in the experiment by five examples each. These examples had both easy and difficult interpretations. Figure 5 illustrates the difference between the two types of questions; subjects gave significantly more correct answers in the case of the easy examples.

The main point of Figure 5 is that we were, in fact, correct in our assessment of how difficult or easy the interpretations were. Our intuitions that subjects would be able to answer the easy interpretations more accurately were borne out.

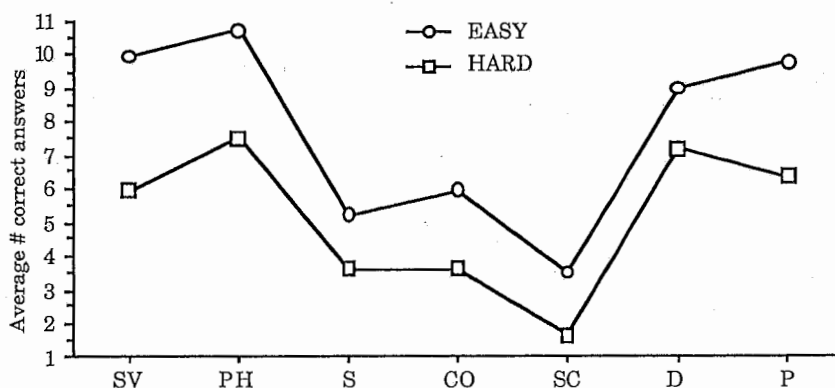


Figure 5. Average number of correct answers per ambiguity category, for easy and hard interpretations.

Since subjects had the options both to make a definite choice or to choose "no answer," we looked at the results for all three possible outcomes: the subject chose the correct response; the subject chose the incorrect response ; and the subject did not choose any response ("no answer"). The results for each case are shown below.

Figures 6 and 7 give us some insight into how subjects reacted to the two different types of disambiguation dialogues. First of all, the only ambiguity class for which there was a clear difference between the two types of wording was Syntactic Class. These ambiguities were more difficult to resolve given the system or machine-like dialogue. We can see this not only in the fewer correct answers given for Syntactic Class ambiguities (Figure 6), but also in the greater number of incorrect answers (Figure 7).

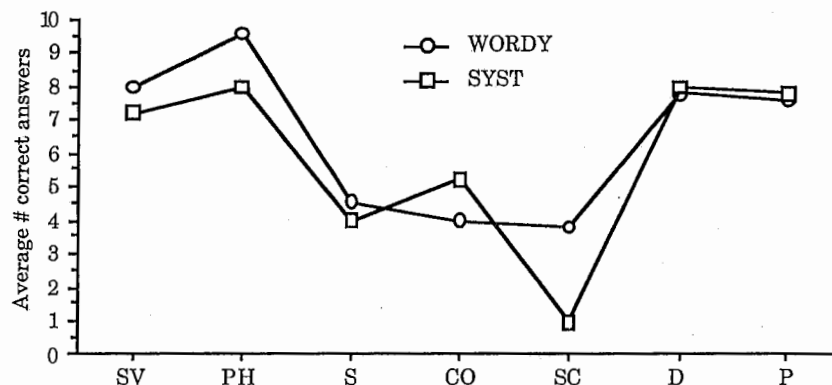


Figure 6. Average number of *correct* answers per ambiguity category, for "wordy" and "system" choices.

The ambiguity classes seemed to fall into two categories depending upon how accurately subjects were able to interpret the intended senses. Subordination involving the Verb (SV), Phrasal Verb (PH), Decoration (D), and Polysemy (P) ambiguities seemed to group together, receiving a relatively high number of correct (and low number of incorrect) answers. On the other hand, subjects seemed to be less accurate in resolving Subordination (S), Coordination (CO), and Syntactic Class (SC). These received a relatively lower number of correct answers and higher number of incorrect answers. These results give us a rough indication of which types of ambiguities are most accurately resolved by subjects.

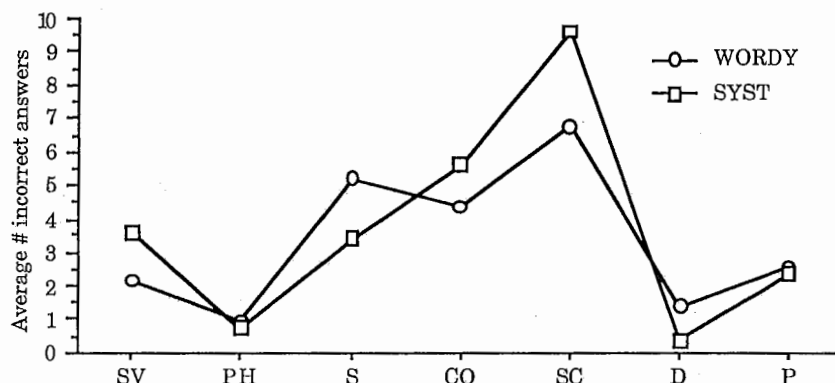


Figure 7. Average number of *incorrect* answers per ambiguity category, for "wordy" and "system" choices.

The “no answer” case is also interesting. Figure 8, which illustrates these results, shows a difference between the “wordy” and “system” dialogues for the Phrasal Verb, Subordination, and Coordination categories of ambiguity. In the first two cases, the “system” dialogues were more difficult to resolve; subjects had a greater number of “no answers” in those cases. For Coordination ambiguities, however, the “wordy” dialogue was more difficult to respond to; subjects had greater number of “no answers” in that case.

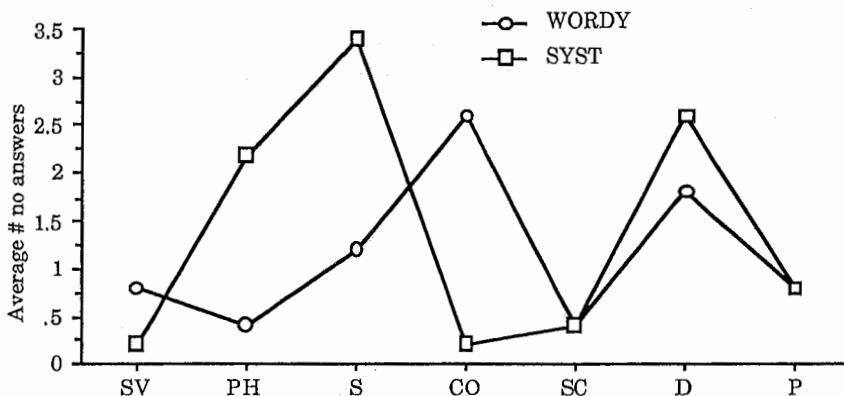


Figure 8. Average number of “no answer” responses per ambiguity category, for “wordy” and “system” choices.

In summary, then, only in the case of Subordination ambiguities was there a difference in subject performance depending upon dialogue type; subjects responded to the human-like dialogues more accurately. Overall, subjects were able to respond relatively more accurately to Subordination involving the verb, Phrasal verb, Decoration and Polysemy ambiguities, and less accurately to Subordination, Coordination and Syntactic class ambiguities. Subjects seemed somewhat more confident in their ability to respond to the human-like or “wordy” phrasings (as measured by the number of “no responses” given) with the notable exception of Coordination ambiguities, which exhibited the opposite trend.

III.2. Behaviors

Subjects had difficulty concentrating at times since they were not isolated. Subjects were puzzled by the text (one of them refused to believe that the text was written by a native English speaker) and the questions. But we do think they did their best to carry out the task as well as possible.

III.3. Questionnaires

Each subject was asked to answer a post-experiment questionnaire (cf. Appendix 2 & 3). The questionnaire had two parts; one part was specifically about the different classes of ambiguity solved, and the second one was about their general feelings and proposals. Here are some of the results we got for each part of the questionnaire, with short comments.

III.3.1. About each kind of ambiguity

In the first part of the questionnaire there was a question about each kind of ambiguity to be solved in the given text. For each class, subject were asked to say if the questions were:

- 1 easy to answer
- 2 had to think, but then the answer was clear

- 3 had to think about it and was still unsure
- 4 really doubtful about the answer
- 5 completely impossible to answer

After each one rated 3 or over, they were also asked to comment on why they thought they had some difficulty.

- A couldn't tell from the text which meaning was intended
- B couldn't understand the meaning of the choices in the dialogue box
- C both choices in the dialogue box seemed the same
- D could understand the dialogue box, but couldn't see how it related to the example
- E other (please explain)

We got the following results:

Sub Nb	Q set	Polysemy		Ph Verb		Synt Class		Coord		Decoration		S wout Vb		S with Vb	
		Rat	Com	Rat	Com	Rat	Com	Rat	Com	Rat	Com	Rat	Com	Rat	Com
2	S	1		1		1		1		1&3	B	3	C	2	
3	S	1		5	C	2		1		3	A	5	B&C	1&4	B
5	S	1		3	B	5	B&C	3	A	2		4	B	4	A
7	S	1		3	B	2&3		1		1		1		2	
9	S	1		1&45	C	3	B	2		2		2&3	C	2	
11	S	1&3	A	2&3	B	3	C	1&3	C	1		2		2	B
13	S	1		2		3	A	3	A	2		1		5	A
15	S	2		2		1		1		4	B	5	B	2	
16	S	1		4	C	3	B	2		4	B	3	C	3	B
19	S	1		2		2		1		2		5	C	1&4	A
21	S	1		2		3	B&C	2		4	E	5	C	1	
1	W	2		1		3	A	1		1		1		1	
4	W	1		1		1&3		1		1		1		1&5	A
6	W	1	B	1	B	2	C	1		2	B	2		2	(B≈)
8	W	1		1&3	A	2		3	C	2		1&3	B	3	A
10	W	1		2		2		1		3	A	2		2	
12	W	1		2		2		3	A	3	A	3	A	2	
14	W	3	A	2		1		1&3	A	2		1&4	A	1&4	D
17	W	1		1&3	E	2		1&3	E	1		2		2	
18	W	1		1		2		1		1		1		1	
20	W	1		1		1	(E)	1&2		3	B	1&4	C	1	
22	W	1		1		2		4&5	E	2		2		2&3	E

Table 2. Subjects' evaluation of ease of solving each class of ambiguity

Subjects almost always responded to the particular examples we gave them (it is shown by the different answers given by the subjects for each class of ambiguity); they could not generalize to the class of ambiguities. In a way, this is natural; there were often big differences among the ambiguous sentences in terms of how difficult they were.

III.3.2. About general feelings & proposals

In the second part of the questionnaire, they were asked about what they generally felt about:

- The disambiguation dialogues,
- Being interrupted to disambiguate,
- Resuming the task after disambiguation.

Subject Number	Question set	Feelings about Dialogues	Interruption	Resuming
2	M	Hard&Easy	Fine	Fine
3	M	Not Wonderful	Too much	Fine
5	M	Stilted	Fine	Fine
7	M	Fine	Fine	Fine
9	M	Confused		Need Pointer
11	M	Hard to understand	Fine	Need Pointer
13	M	Fine	Fine	Fine
15	M	Hard to understand	Fine	Fine
16	M	Fine	Fine	Fine
19	M	Fine	Bothered	Fine
21	M	Bothersome&Fine	Fine	Need Pointer
1	H	Confusing	Fine	Fine
4	H	Mostly ok	Fine	Need Pointer
6	H	Confusing	Fine	Need Pointer
8	H	Confused	Fine	Need Pointer
10	H	Text Confusing	Expected	Fine
12	H	Irritated	Irritated/Surprised	Need Pointer
14	H	Verbose/Unnatural	Fine	Fine
17	H	Not Clear	Anticipating	Fine
18	H	Fine	Fine	Fine
20	H	Fine	Fine	Need Pointer
22	H	Irritated	Fine	Fine

Table 3. Subjects' feeling about the experiment

Subjects often responded that being interrupted was not a problem. We suspect this was because being interrupted was the point of the experiment. They had agreed to do the experiment, so they weren't bothered by the interruptions. This doesn't tell us much about what their reactions would be in a "real life" situation.

They were also asked about any more comments or suggestions about the design of a disambiguation module. The majority of subjects pointed out the need for the whole sentence to appear in the dialogue box with the ambiguity underlined. They also asked for "no click to confirm", a "selection button on the right hand side of the choices", a "default choice (the most plausible one)", a "voice interface and use of prosody".

Conclusions

Lessons learned

In the course of conducting the pilot experiment and discussing their impressions with subjects afterwards, we learned a number of things that affected the design of the following experiment.

Subjects frequently commented on how unnatural the text seemed to be. There were three reasons why the text sounded unnatural. First, it included actual spoken English examples in written form, surrounded by (made up) written context. Transcriptions of spoken English often sound unnatural, especially embedded in written text. Second, some of the “hard” interpretations were ones that, in real life, only a computer would have trouble understanding. In trying to motivate these difficult interpretations, unnatural text was produced. And third, there was a much higher density of ambiguous sentences in the text than would be found in real text or speech.

The questionnaire did not allow us to draw any conclusions about the easiness or hardness of the questions according to the kind of ambiguity to be solved. The confidence level of the subjects was quite good, even if in the reality the actual results were poor. The global feeling about the dialogue was that they were hard to understand or confusing. We think that the feeling of confusion came from the unnaturalness of the interpretations we were trying to impose for most of the ambiguous sentences.

Implications for the design of the next experiment

We made a number of changes to the experimental format for the second experiment based on our experience in the pilot experiment. We realized that using hard interpretations in the pilot experiment was a mistake. It made the text sound unnatural and made the task more difficult for the subjects, clouding the real issue: how well they could respond to the different wordings of the dialogues.

We also changed the arrangement of the screen so that subjects could check back to the text to confirm their understanding of the ambiguity involved. Subjects complained that they could not do this in the pilot; this was also an unnecessary obstacle to the accomplishment of the task.

While the “no answer” option gave some interesting results, it also made it difficult to see trends in answering the questions clearly. For that reason, we designed the subsequent experiment to be forced choice.

References

- [1] **Blanchon H. (1994)** *Pattern-based approach to interactive disambiguation: first definition and implementation*. Technical Report, TR-IT-0073, ATR-Interpreting Telecommunications Research Laboratories, Sept. 8, 1994, 91 p.
- [2] **Blanchon H. (1995)** *An Interactive Disambiguation Module for English Input: an Engine and the Associated Lingware*. Technical Report, TR-IT-0129, ATR-ITL, August, 1995, 110 p.
- [3] **Blanchon H., Fais L., Loken-Kim K.-h. & Morimoto T. (1995)** *A Pattern-Based Approach for Interactive Clarification of Natural Language Utterances*. Bulletin of the Information Processing Society of Japan (95-NL-107), 95/52, 11-18.
- [4] **Fais L. & Loken-Kim K.-h. (1994)** *Effects of communicative mode on spontaneous English speech*. Technical Report, NLC94-22, Institute of Electronics, Information and Communication Engineers, Oct. 94, 6 p.
- [5] **Fais L. & Loken-Kim K.-h. (1995)** *Lexical Accommodation in Human-interpreted and Machine-interpreted Dual Language Interactions*. Proc. ESCA Workshop on Spoken Dialogue Systems, Vigsø, Denmark, May 30-June 2, 1995, vol. 1/1, 69-72.
- [6] **Fais L., Loken-Kim K.-h. & Park Y.-d. (1995)** *Speakers' responses to requests for repetition in a multimedia cooperative dialogue*. Proc. International Conference on Cooperative Multimodal Communication, Eindhoven, The Netherlands, May 24-26, 1995, vol. 1/1, 129-144.
- [7] **Loken-Kim K.-h., Yato F., Kurihara K., Fais L. & Furukawa R. (1993)** *EMMI-ATR environment for multi-modal interactions*. Technical Report, TR-IT-0018, ATR-ITL, Sept. 30, 1993, 28 p.

Appendix I: Text presented to the subjects

I will be giving you information about traveling to the conference center and about various sightseeing spots in Kyoto. First, I will tell you about travel options and later I'll tell you about sightseeing. So let me **tell you what to do about traveling now.** **SVE** If you will be coming by Shinkansen, please be sure you have all your luggage with you and check your seat number before you leave the train. Then **exit out from the Shinkansen.** **PHH** Once you do that, you will walk down the platform toward the exits. On your left will be the offices where you can buy tickets for the special express trains, the bullet train, or tourist charter trains. Walk past that **train ticket vendor.** **SE** Then look for your exit. The exits have either names, such as South Central, or numbers, such as Exit 3. They are clearly marked above the exit door. So you should be able to see the sign saying the **exit with the number, or the name.** **COH** Follow the signs for Exit 14. At this point, you are on the lowest level of the station. Please take the escalators up. As you go up, you will reach one level that is built like a set of bleachers, in steps, and then two platforms that are level. You can **catch a taxi at the second level platform.** **SCH** When you arrive at this level, there will be additional signs for Exit 14. The exit you want to take is the South Central, right next to Exit 14. So, please **exit by Exit 14,** through South Central. **DH** Once you are at this platform, please continue through the main doors of the station which should be directly in front of you. At this point, you must be careful of your luggage. If you have a large suitcase, **it may be difficult to get out of Kyoto Station.** **PHH** If you have trouble getting through the door with any large luggage, there is a special luggage exit, one door to your left. You should have no trouble with your luggage if you just **take that one door to your left.** **SCE** You will have to go over a drainage conduit cover embedded in the floor of the station just past the exit. It is a little bumpy, and your luggage wheels may get caught, so you may want to be sure to **go over it carefully.** **PHH** Now, once you are out of the station, you have several options for getting to the conference center. You can take a subway, but it is a bit complicated. Most of our customers prefer to taxi to the conference center. I will assume that you would also prefer to taxi, though **you could take a bus or taxi.** **SCH** Once you have left Kyoto station by Exit 14, you should see the taxi stand directly in front of you. The taxis wait here in line to take each customer in the order that they come out from the station. **You will see many taxis waiting.** **SVE** Walk to the front of the line and you will see a sign that says "boarding area." This is the area **where you get on the taxi.** **PHE.** There are a number of different kinds of taxis and they will be different colors. The red ones are taxis from the conference center. Just tell the driver of a red cab where you want to go, and the trip should take you five minutes **by taxi from the conference center.** **SH** Be sure to take a red cab. Once taxis get to the conference center, all other taxis must turn left and stop in a lot that is a little far from the center entrance. However, the red taxis don't have to turn left, they come **right,** to the front of the center. **PH** Because these are special conference center taxis, the conference center will pay part of the cost of the taxi. Please inquire at the conference information desk to find out the cost of the taxi to you and **the cost of the taxi to the conference center.** **DH** The conference information desk can also help you with information about your return. Please consult them to help you make reservations to **take the taxi back.** **PHE**

The area around the conference center is rich with cultural sights. There are numerous shrines and temples and the famous castle, Nijo Castle. The conference information desk can also help you with transportation to **these temples or castles like Nijo.** **COH** Most conference goers will not leave immediately after the conference. If you are free on the day after the conference, perhaps you would like to join the sightseeing tour arranged especially for conference participants at that time. Please notify the desk if you would like to take part in this **tour the day after the conference before you leave.** **SE** Or you might like to arrange your own sightseeing. Although many taxi drivers in Japan speak English, not all of them do. Therefore, you may require help with communicating with your taxi driver. Please just ask anyone at the conference information desk. They know what to say and can help you with **where you should tell the taxi to go.** **DE**

If you do not wish, for some reason, to come to the center by taxi, please refer to the conference packet you received when you sent in your pre-registration. In it you will find a brochure. Consult it if you want to **look at** another method of travel. **PH** There are trains from the eastern terminal of the subway, or buses from Kyoto Station itself which can get you to the conference center. Please read the brochure concerning taking **a train or bus from the Station.** **COH**

I represent a travel service company. We specialize in helping visitors to Kyoto make arrangements for their stay here. We understand that you are coming to Kyoto for a conference. And **we'd also like to make such arrangements for you.** **DE** We have a number of agents, both human and mechanical. Our mechanical agents, or robots, can only display information for you graphically. Among our speaking agents, we have some who are French, some who are German and some who are English. I am an **English speaking agent.** **SCH**

To begin with, you have several options for making your arrangements. You can make a reservation through a foreign travel agency, a Japanese travel agency or by calling the hotel itself. However, out of all these options, **the most efficient one will be calling us to make your arrangements.** **SCE** Of course you may be calling some other places to help you as well. We would like to know who else you are calling after you call us so that we can know who our competitors are. For that reason, we'd like to know **where you are calling, from now.** **SVH** The information I will give you may be rather complex. We want to be sure that you fully understand it. So, please feel free to ask any questions you might have so that you can be sure to **get it down correctly.** **PE** Now I'd like to give you some information about making reservations for your hotel. At most Japanese hotels, you can make a reservation for a night including breakfast or for a night not including breakfast. The price is thirteen thousand yen for a night not including breakfast and **fifteen thousand for a night including breakfast.** **SH** Of course, hotels have rooms in

both Japanese and Western styles. You can reserve a Japanese room with either a Japanese-style or a Western-style bath. However, the western-style rooms only have western-style baths. If you reserve a western-style room, it will have a regular western bed and bath. **COE** There are deluxe Japanese rooms available as well. These look out on their own private gardens. Often these gardens have ponds with decorative carp. You could be eating your Japanese style breakfast with fish outside your window. **SVH** We have a number of sizes of rooms to choose from also. You can arrange to have a suite with one bedroom or a suite with two bedrooms. Just let us know if you'd like a single bedroom suite or a double bedroom. **SH** Some hotels have extensive recreational equipment. The Miyako, for instance, has an Olympic size swimming pool where many guests swim laps for exercise, and a gym complete with weights and exercise bicycles. You might enjoy staying at such a hotel where there is a swimming pool and also a gym where you can train. **COE**

Once you have chosen the hotel you'd like to stay in, we will handle your reservation, but we require a deposit. The deposit must be in yen, and it *must* be made by postal money order or by bank transfer. No cash or credit cards, please. Although you cannot make the deposit with a credit card, you can pay the final bill with a credit card. If you would like to pay the final bill with your credit card, we will need to see your credit card before you register. For that reason, please let us have the deposit with your credit card. **DH** Because we are particular about our clients, we assume our clients are particular, too. For that reason, you will have an opportunity to check out the hotel and be sure that it is satisfactory before you register. Please let us know when you would like to check out your hotel. Most guests plan to check out the day before they register. **PH**

Or, if you would rather make your own reservation, you can do that, too. Right next to the conference center where your conference is being held is a Tourist Information and Hotel Reservation Office with a number of stories. **PE** They can help you make reservations for hotels anywhere in the city. So it is very easy for you to make a reservation for a hotel near the conference center. **SVH**

Figure 9. The text read by the subjects.

Appendix II: Questions presented to the subjects

Human	Machine
let me tell you what to do about traveling now let me tell you now about traveling let me tell you about your forthcoming travel	let me tell you what to do about traveling now now, let me tell you what to do about traveling let me tell you what to do about (traveling now)
exit out from the Shinkansen to out of the Shinkansen area to exit the Shinkansen itself	exit out from the Shinkansen to leave from the Shinkansen to leave the shinkansen
that train ticket vendor someone selling train ticket someone selling tickets to buy a train	that train ticket vendor that vendor for train ticket that ticket vendor for train
the exit with the number, or the name the name the exit with the name	the exit with the number, or the name (the exit with the number) or (the name) (the exit with the number) or (the exit with the name)
level the platform is on the second level the platform itself is level	level a platform that is (second level) a (second platform) that is level
exit by Exit 14, through South Central get out near Exit 14 get out through Exit 14	exit by Exit 14, through South Central exit next to Exit 14 exit through Exit 14
to get out of Kyoto Station to get yourself out of Kyoto Station to take something else out of Kyoto Station	to get out of Kyoto Station to leave Kyoto Station to get (out of Kyoto Station)
that go through that door to your left take that through the door to your left	that take (that one door) to your left take (that) one door to your left
go over it carefully look it over and consider it carefully walk or move above and over carefully	go over it carefully to examine it carefully to go (over it) carefully
taxi you could take a bus or a taxi you could take a bus or you could taxi	taxi you could take a taxi you could taxi
you will see many taxis waiting while you are waiting, you will see many taxis you will see many taxis that are waiting	you will see many taxis waiting waiting, you will see many taxis you will see many (taxis waiting)
get on the taxi actually climb into the taxi arrive at somewhere on the taxi	get on the taxi board the taxi get (on the taxi)
by taxi from the conference center the trip is five minutes from the conference center the taxi is the conference center's property	by taxi form the conference center from the conference center, five minutes by taxi by (taxi from the conference center); five minutes
right on the right side, not on the left straight there; no detour	right opposite to the left directly
the cost of the taxi to the conference center the taxi is going to the conference center the conference center will pay for the taxi	the cost of the taxi to the conference center the cost of the taxi towards the conference center the cost of the taxi in case of the conference center
take the taxi back take the taxi and return it take the taxi and go back	take the taxi back return the taxi take the taxi (back)
these temples or castles like Nijo these temples these temples like Nijo	these temples or castles like Nijo (these temples) or (these castles like Nijo) (these temples like Nijo) and (these castles like Nijo)
the day after the conference before you leave it the day before you leave it the day after the conference which is right before you leave	the day after the conference before you leave (the day before you leave) after the conference the day after (the conference before you leave)
where you should tell the taxi to go where you should speak where the taxi should go	where you should tell the taxi to go at what place you should tell the taxi to go to what place you should tell the taxi to go

look at actually looking with your eyes think about or think over	look at to see to consider
a train or bus from the station the bus only is from the station the train and the bus are from the station	a train or bus from the station (a train) or (a bus from the station) (an train from the station) or (a bus from the station)
arrangements for you the arrangements are for you the arrangement are made on your behalf, possibly by someone else	arrangements for you arrangements appropriate to you arrangements instead of you
speaking someone who speaks English someone who is English and who is speaking	speaking an agent that is (English speaking) A (speaking agent) that is English
calling the most efficient option will be to call us to make the arrangement the most efficient person will be calling us to make the arrangement	calling the most efficient one (will call) us to make the arrangement the most efficient one (will be) calling us to make the arrangement
where are you calling from now where are you calling from at this point of time from now, where are you calling to.	where are you calling from now now, where are you calling from from now, where are you calling
get down to make sure something is put down to get a complete understanding	get down to lower to understand
fifteen thousand for a night including breakfast fifteen thousand for a night which happens to also include breakfast fifteen thousand for a particular kind of night which includes breakfast	fifteen thousand for a night including breakfast including breakfast, fifteen thousand for a night for (a night including breakfast), fifteen thousand
a western bed and bath a bath a western bath	a western bed and bath (a western bed) and (a bath) (a western bed) and (a western bath)
You could be eating you Japanese style breakfast with fish outside your window your Japanese breakfast includes fish the fish are outside the window	You could be eating you Japanese style breakfast with fish outside your window outside your window, you could be eating you Japanese style breakfast with fish with fish outside your window, you could be eating you Japanese style breakfast
a double bed-room two bedrooms a room with a double bed	a double bed-room a double (bedroom) a (double bed) room
a swimming pool and also a gym where you can train a swimming pool a swimming pool where you can train	a swimming pool and also a gym where you can train (a swimming pool) and (a gym where you can train) (a swimming pool where you can train) and (a gym where you can train)
the deposit with your credit card use your credit card for the deposit let us have your deposit and your credit card	the deposit with your credit card the deposit using your credit card the deposit and your credit card
check out to settle the bill and leave the hotel to look into or investigate something	check out to pay and leave the hotel to verify
story like fairy tales or legends different levels or floors	story narrative floor
it is very easy to make a reservation for a hotel near the conference center the reservation takes place near the conference center the hotel is near the conference center	it is very easy to make a reservation for a hotel near the conference center near the conference center, it is very easy to make a reservation for a hotel it is very easy to make a reservation for (a hotel near the conference center)

Table 4. Questions asked to the subjects in both settings

Appendix III: Results in the human-like setting

Question	Class	Diff	Ans	1	4	6	8	10	12	14	16	18	20	22	GA
traveling	SV	E	1	2	1	2	1	1	2	1	1	1	1	2	7
exitout	PH	H	2	0	2	1	2	2	0	2	2	2	2	2	8
trainticket	S	E	1	1	1	1	1	1	0	1	0	1	1	1	9
sign	⊙	H	1	2	0	2	2	2	2	2	1	1	0	0	2
platform	SC	H	2	1	1	1	2	2	1	1	1	1	1	1	2
exit14	D	H	1	1	0	1	1	1	0	2	1	2	1	1	7
getout	PH	H	2	2	1	1	2	2	1	2	0	2	0	2	6
door	SC	E	2	1	1	1	1	1	2	1	1	1	1	1	1
goover	PH	H	2	2	2	2	2	2	2	2	2	2	2	2	11
taxi	SC	H	2	2	1	1	1	0	1	1	1	1	0	1	1
waiting	SV	E	2	2	2	2	2	2	2	2	2	2	2	2	11
geton	PH	E	1	1	1	1	1	1	1	1	1	1	1	1	11
taxifrom	S	H	2	1	0	1	1	1	2	0	0	1	1	1	1
right	P	H	1	2	1	2	1	1	1	1	1	1	1	1	9
taxicost	D	H	2	0	2	2	2	2	2	2	0	2	0	2	8
takeback	PH	E	2	2	2	2	2	2	2	2	2	2	2	2	11
temples	⊙	H	1	1	1	1	1	1	0	0	1	0	0	0	6
tour	S	E	1	2	2	2	2	2	2	0	1	0	0	2	1
telltaxi	D	E	2	0	2	2	1	2	2	2	2	2	2	0	8
lookat	P	H	1	2	1	2	2	2	0	2	2	2	1	2	2
trainorbus	⊙	H	1	2	2	2	1	2	2	2	2	2	2	2	1
arrangement	D	E	1	1	1	2	1	1	1	1	0	1	2	1	7
speaking	SC	H	2	1	1	1	2	1	1	1	1	1	2	1	2
callingus	SC	E	1	1	1	1	1	1	1	1	2	1	1	1	10
callingfrom	SV	H	2	2	0	1	1	2	0	0	1	1	2	2	4
getdown	P	E	2	2	0	2	2	2	1	1	2	0	2	2	7
breakfast	S	H	2	1	2	2	1	0	2	2	1	1	1	1	4
bath	⊙	E	2	2	2	2	2	2	2	2	2	2	2	0	10
fish	SV	H	2	2	2	2	2	2	2	2	1	2	2	2	10
bedroom	S	H	1	2	1	2	1	1	2	1	2	1	1	2	6
pool	⊙	E	2	2	1	1	2	1	1	1	1	0	0	0	2
deposit	D	H	2	2	0	2	2	0	2	2	2	1	2	0	7
checkout	P	H	2	2	2	2	2	2	2	0	2	2	2	1	9
story	P	E	2	2	2	2	2	2	2	2	2	2	2	2	11
reservation	SV	H	1	2	1	1	1	2	2	2	1	2	2	0	4

Table 5. Subjects' answers with the human-like dialogues

Legend

Question: head word of the ambiguous phrase,

Class: class of the ambiguity,

Diff: difficulty of the question; Easy or Hard,

Ans: expected answer,

GA: total of Good Answers.

Appendix VI: Results in the machine-like setting

Question	Class	Diff	A	2	3	5	7	9	11	13	15	17	19	21	GA
traveling	SV	E	1	1	1	1	1	1	1	1	1	1	1	1	11
exitout	PH	H	2	0	0	0	0	0	2	2	1	2	0	1	3
trainticket	S	E	1	0	1	1	0	1	2	1	0	1	1	1	7
sign	⊙	H	1	2	2	2	1	2	2	2	2	0	2	2	1
platform	SC	H	2	1	1	2	1	1	2	1	1	1	1	1	2
exit14	D	H	1	1	0	2	1	0	2	1	1	2	1	1	6
getout	PH	H	2	0	0	2	1	0	1	2	1	2	1	1	3
door	SC	E	2	1	1	1	1	2	1	1	1	2	1	1	2
goover	PH	H	2	2	2	2	2	2	2	2	2	2	2	2	11
taxi	SC	H	2	1	1	0	1	0	1	1	1	2	1	1	1
waiting	SV	E	2	2	2	2	2	2	2	2	2	2	2	2	11
geton	PH	E	1	1	0	1	1	1	1	1	1	1	1	1	10
taxifrom	S	H	2	x	0	1	1	2	1	2	1	1	0	1	2
right	P	H	1	1	0	1	1	1	1	1	0	1	1	1	9
taxicost	D	H	2	0	0	2	0	2	2	2	0	2	2	2	7
takeback	PH	E	2	2	2	2	2	2	2	2	2	2	2	2	11
temples	⊙	H	1	1	1	1	1	1	1	1	1	1	1	1	11
tour	S	E	1	0	0	1	1	2	1	2	2	2	1	0	4
telltaxi	D	E	2	2	2	2	2	2	2	2	2	2	2	2	11
lookat	P	H	1	2	2	2	0	2	2	2	2	2	2	2	0
trainorbus	⊙	H	1	2	2	2	1	2	2	2	2	2	2	2	1
arrangement	D	E	1	1	0	1	1	1	1	1	1	1	1	0	9
speaking	SC	H	2	1	1	2	2	1	1	1	1	1	1	1	2
callingus	SC	E	1	2	2	2	2	2	2	2	2	1	2	2	1
callingfrom	SV	H	2	1	1	1	2	2	1	1	1	2	2	0	4
getdown	P	E	2	0	2	2	2	2	2	2	2	2	2	2	10
breakfast	S	H	2	0	0	2	1	0	1	2	0	2	0	2	4
bath	⊙	E	2	2	2	2	2	2	2	2	2	2	2	2	11
fish	SV	H	2	2	2	2	1	2	2	2	2	2	2	2	10
bedroom	S	H	1	0	1	0	2	2	2	2	1	1	1	1	5
pool	⊙	E	2	2	1	1	1	1	0	1	1	1	1	1	1
deposit	D	H	2	0	2	2	2	2	2	2	0	2	2	0	8
checkout	P	H	2	2	2	1	1	2	2	2	2	2	2	2	9
story	P	E	2	2	2	2	2	2	2	2	2	2	2	2	11
reservation	SV	H	1	2	2	2	2	2	2	2	2	1	1	2	2

Table 6. Subjects' answers with the machine-like dialogues

Legend

Question: head word of the ambiguous phrase,

Class: class of the ambiguity,

Diff: difficulty of the question; Easy or Hard,

Ans: expected answer,

GA: total of Good Answers.