TR-IT-0167

A Second Experiment on the Understandability of Interactive Disambiguation Dialogues

Hervé BLANCHON†
Laurel FAIS

1996.04.19

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

A first pilot experiment was carried out during the summer 1995 on the same topic [Blanchon & Fais 1995]. This second experiment is the fruit of what we learned and the results are far better.

Two classes of questions (human-like and machine-like) were proposed using two different modalities (text and speech). The human-like questions were the easy ones and the machine-like ones were the difficult ones. We had four groups of subjects, each subject participating in one setting.

A text was read aloud by the subjects, who then answered 35 questions. The analysis of the results shows that there is basically no significant difference among the four settings. This tends to prove that the machine style questions, which could be produced automatically, are not that hard to answer.

† The work report in this report has been Carried out while H. Blanchon was invited, from GETA-CLIPS (Grenoble, France), for three weeks at ATR_ITL in the framework of the MIDDIM project.

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

TABLE OF CONTENTS

LIST OF FIGURES

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 [Blanchon 1994 ; Blanchon, *et al.* 1995 ; Blanchon 1995a]. It is now important to evaluate the understandability of the questions to be produced with this methodology. During the summer 1995 a first 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 were not so good [Blanchon & Fais 1995], but we learned a lot.

For this second experiment, we designed a text that was easy to understand, with no difficult interpretations of any of the sentences. We recruited enough subjects to test four settings: two settings presented human-like disambiguation questions, and two settings, machine-like disambiguation questions. The settings differed also by the modality used to present the questions to the subjects: two settings presented spoken questions, and two settings, written questions. We used spoken settings because speech is reported to be an interesting modality for disambiguation [Lehiste 1973 ; Lehiste, *et al.* 1976 ; O'Malley, *et al.* 1973 ; O'Shaughnessy 1989 ; Streeter 1978]. The subjects 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 wizard and the subjects were separated, each of them on both sides of a partition. They communicated through head sets (microphone, headphones).

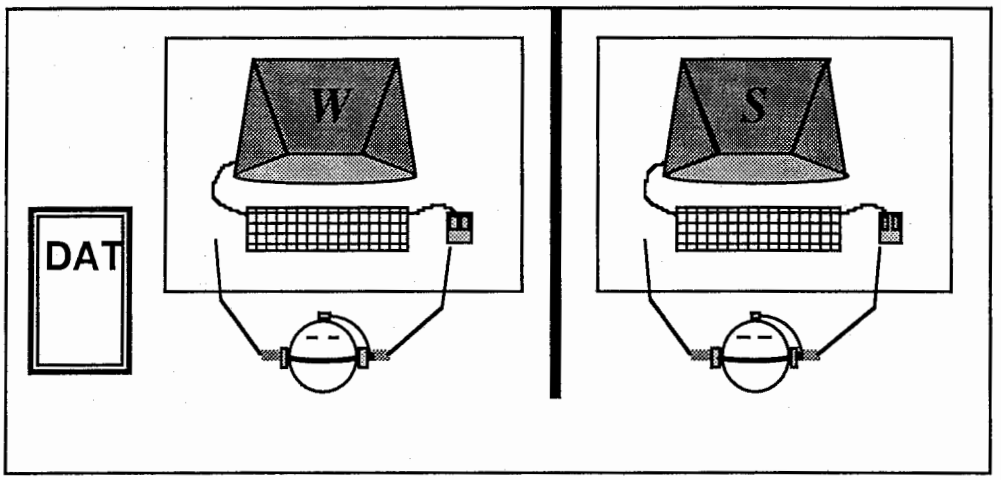


Figure 1: Experimental setting.

The subject was asked to read aloud, slowly and carefully, a text displayed inside a text window (Fig. 5) and pause between each sentence. The scrolling of the text window was

controlled by the wizard (i.e. the text windows of the subject and the wizard were synchronized).

The wizard controlled

- the recording, on a DAT tape, of the subjects' reading,
- the scrolling of the text windows for himself and the subject,
- the setting the subject participated in:
  - Speech Human (spoken, human-like dialogues),
  - Speech Machine (spoken dialogues that could have been generated automatically),
  - Text Human (human-like text dialogues),
  - Text Machine (text dialogues that could have been generated automatically).

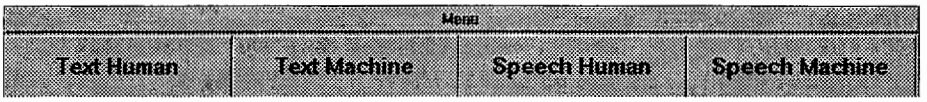


Figure 2: Menu for the selection of the dialogues set to be proposed to the subject.

- the presentation of the questions, written or spoken.

I.1.1. Presentation of the questions by the wizard

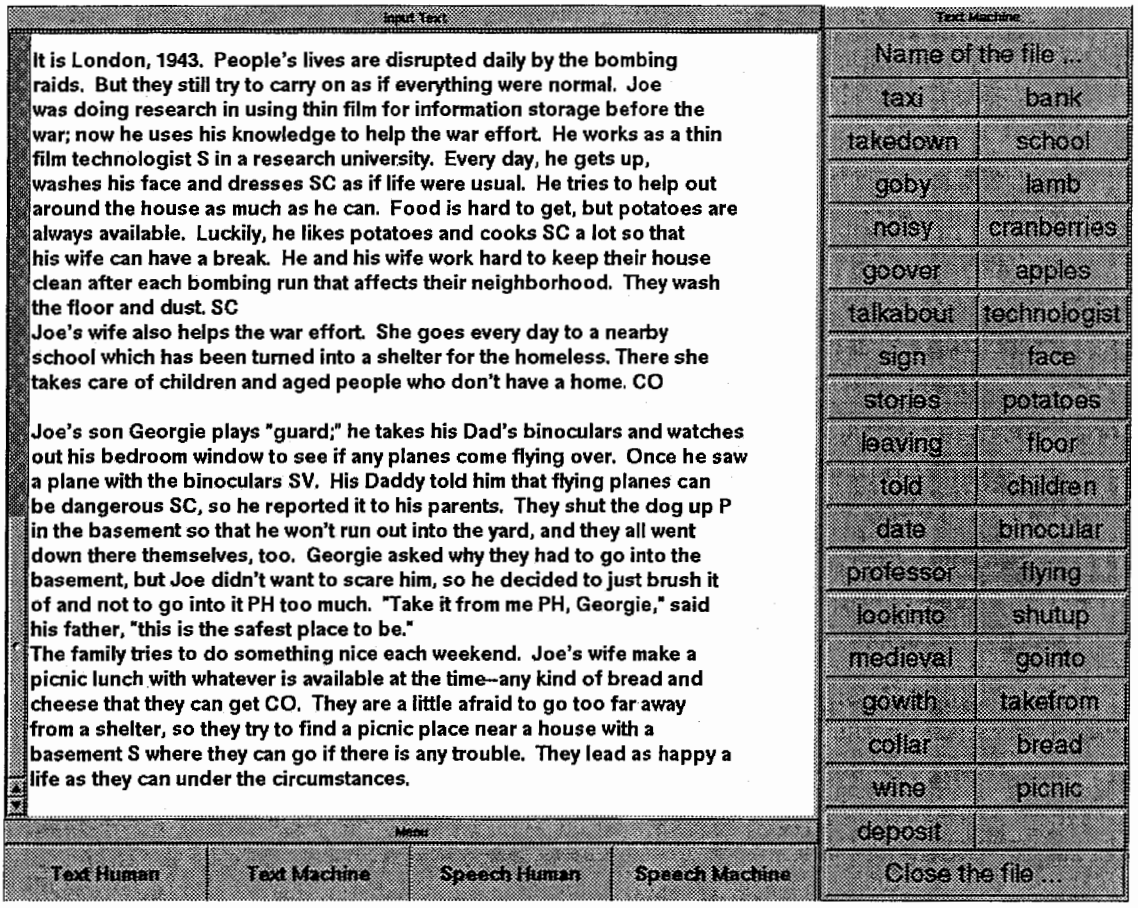


Figure 3: The wizard's screen in the text situation.

The presentation of the question was controlled through a palette made of buttons. Once a button was released the corresponding question was asked (spoken or displayed in text on the subject screen).

With the text questions (Fig. 5 & 7), the answers were recorded in a text file. The button at the top of the button palette was used to create and open this file (one per subject). The button at the bottom of the palette was used to close the file.

With the spoken questions (Fig. 4) the answers were recorded on a tape. Thus the wizard only had to present the questions.

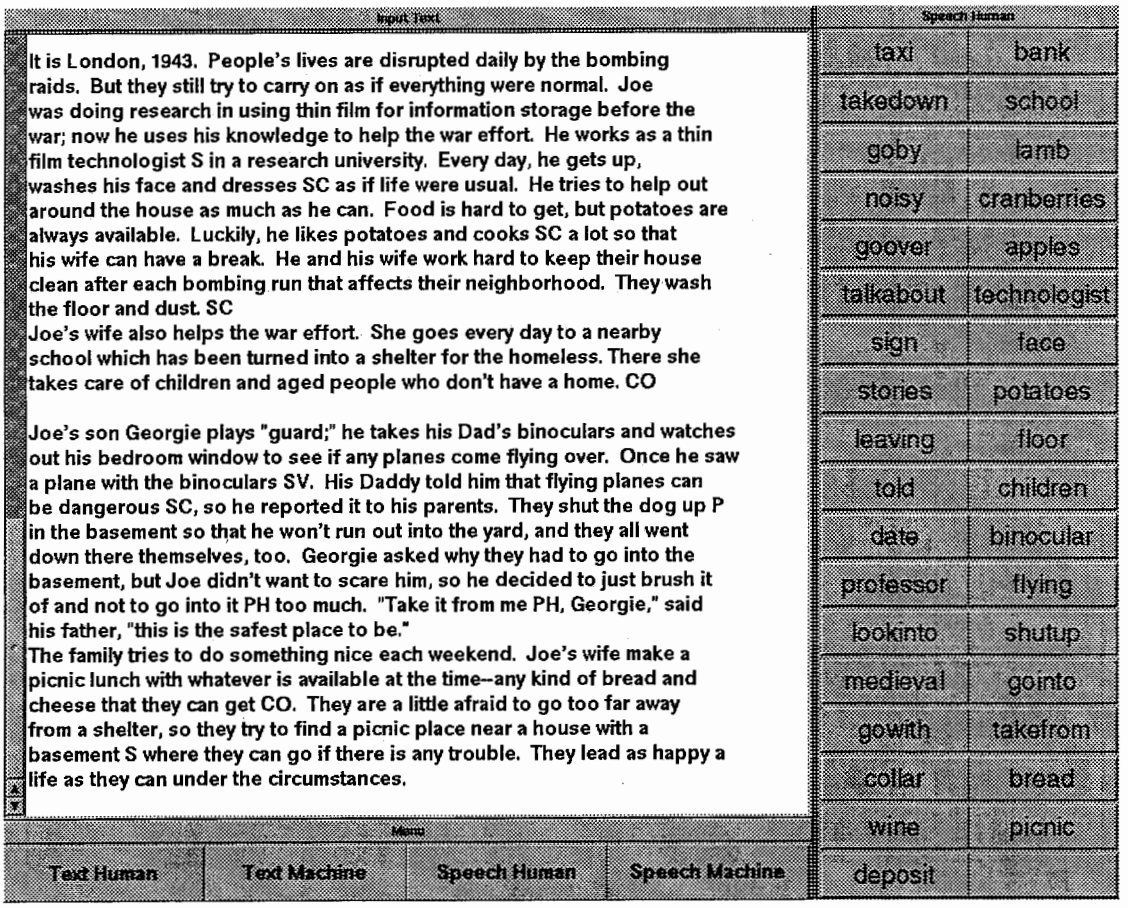


Figure 4: The wizard's screen in the spoken situation.

I.1.2. Presentation of the questions to the subject

The written questions were displayed in a dialogue box presenting an ambiguous utterance and two possible interpretations (Fig. 5 & 7). When a question was presented to be answered, the back ground color of the text turned from white to gray. The appearance of the subject's screen while a question is pending is shown in Figure 5.

The spoken questions were prerecorded on the subject workstation. When a question was to be asked, the corresponding sound file was played by the wizard.

I.2. Experimental settings

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

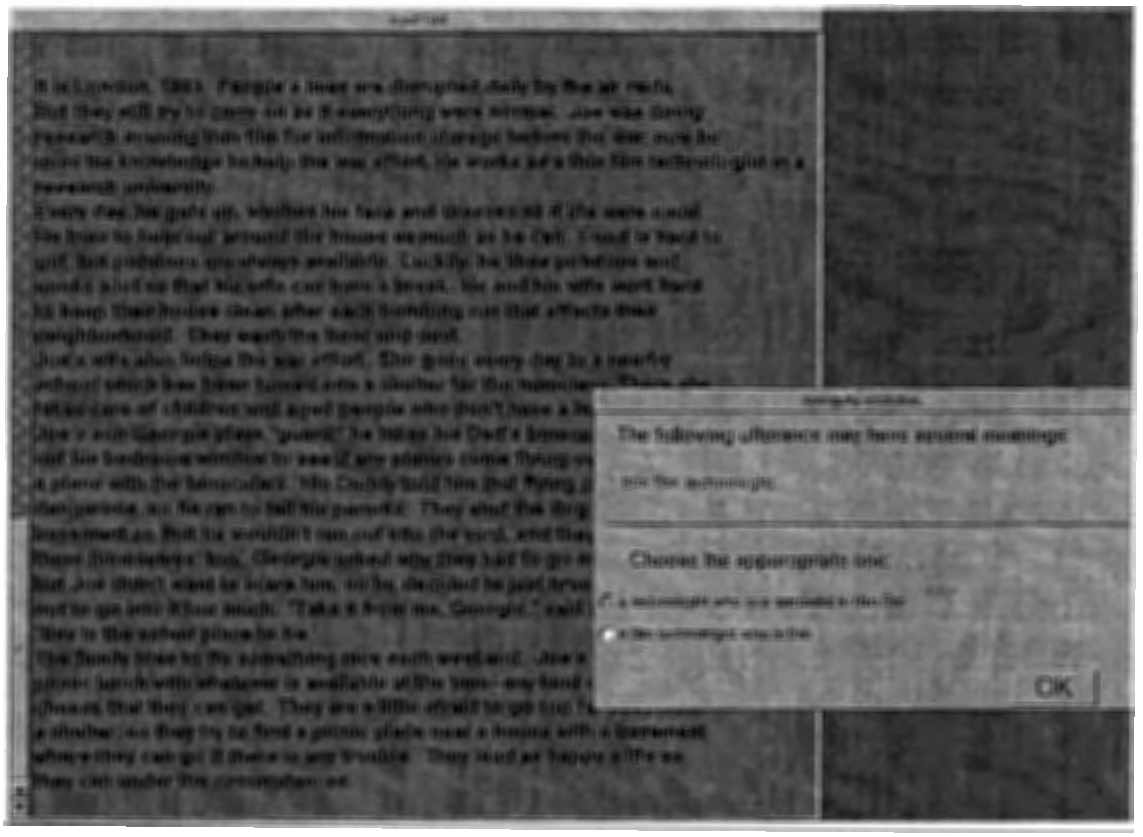


Figure 5: the subject screen when a textual question is pending

I.2.1. Text

The text to be read was made up of three different stories (Figure 6). It contained 35 ambiguous sentences, five in each category of ambiguities described in the table below. Each ambiguity is in boldface in Figure 6.

The codes found in the text (Fig. 6) give the ambiguity type described as follows:

D	Decoration	Problem with the function of a phrase as in "The lamb did look ready to eat." – the lamb will eat, – the lamb will be eaten.
PH	Phrasal Verb	Interpretation of a phrasal verb as in "Take it from me." – remove it from me, – believe me.
CO	Coordination	Interpretation of conjoined phrases as in "Medieval food and drink." – medieval drink, – not medieval drink.
S	Subordination without the verb	Subordination of phrases not involving a verbal phrase as in "He is a thin film technologist." – thin film, – thin technologist

P	Polysemy	A word with several senses such as "bank." – river side, – financial institution
SV	Subordination with the Verb	The complementation of the verb is not certain as in "He brought a bottle of wine from France." – a bottle of French wine, – a bottle, brought from France, of wine
SC	Syntactic class	Ambiguity of syntactic class as in "She washes her face and dresses." – She washes her dresses, – She dresses.

Bill had to go to the library to get a book. But he didn't want to have to pay **the cost of the taxi to the library.D** So he decided to drive himself. However, he wasn't sure how to get there, so he called his friend Tom. Tom gave him the directions over the telephone and told Bill to **take them down PH** carefully. Tom told Bill to **go by the highway D,** and then get off at the Charles Street exit. But Bill told Tom that he hated taking the highway; the cars and the trucks drove so quickly and made so much noise. He especially didn't like all the **noisy cars and trucks CO**. So Tom gave him different directions, using a much longer route so that he could avoid the highway. The directions were so complicated that they had to **go over them PH** twice before Bill understood. They **talked about an hour D** before they were finished.
The next day, Bill drove to the library. He found the sign Tom had told him about–the sign **in front of the gas station with the red roof. S** He looked to the left, as Tom had told him to, and there he saw the library building with a number of **stories P**. On the third floor, he found the book that he wanted, but the librarian was nowhere to be seen. Bill decided to go home, but he **met the librarian as he was leaving SV.** He signed out the book and took it home with him. The trip took him all day, and he was exhausted when he got home. He called Tom and **told him he had had a terrible time that day SV.**

There was going to be a big Medieval Arts festival that night in the nearby city of Newton, and Susan really wanted to go. She didn't **have a date P** to go to the festival, though, so she wondered what to do. Finally she thought she might call up her **Old English professor S** and see if he wanted to go with her. He hadn't heard about the festival and told her that he would **look into it PH**. He called back an hour later to say that he had talked to a friend about it and that it sounded quite interesting. He heard that everything there would be medieval: medieval entertainment and **medieval food and drink CO.** He agreed to pick her up at 8:00, so she started getting ready.
She searched through her closet for something appropriate to wear. She found a cape, but she also wanted a long dress that would match it. Finally she found a dress that would **go with P** the cape. They both were made from velvet and the cape had a feather collar. She put on **the dress and the cape with the feather collar CO** and waited for her professor. He arrived with a **bottle of wine from France SV,** so they had a glass of wine before they left.
The festival was easy to find–the area it was in was brightly lit and the music could be heard from a long way off. It was a warm night, so she **deposited her cape with the coat check clerk D** and she and the professor went to see what there was to eat. The food stalls were lined up along the river under a string of white lights, so they headed for the **bank P**. There they found a wide variety of foods available. It was early, so some of the stalls weren't quite open, but they found one that advertised a special French style of mutton. The professor admitted that he was quite fond of **the French school of cooking S** and so they went to see if they could get some mutton stew. The **lamb did look ready to eat D,** so they ordered a dish to share between the two of them. The stew also contained chestnuts and cranberries, and was delicious. They enjoyed walking along the river **eating their mutton stew with cranberries. SV**
They were still hungry, so they continued to wander among the food stalls. Other people were also strolling along the river. Some were eating cakes and **some were eating apples SC.** But soon they got tired, so they walked back to the entrance, picked up her cape, and went home.

It is London, 1943. People's lives are disrupted daily by the air raids. But they still try to carry on as if everything were normal. Joe was doing research in using thin film for information storage before the war; now he uses his knowledge to help the war effort. He work as a **thin film technologist S** in a research university.
Every day, he gets up, **washes his face and dresses SC** as if life were usual. He tries to help out around the house as much as he can. Food is hard to get, but potatoes are always available. Luckily, he **likes potatoes and cooks SC** a lot so that his wife can have a break. He and his wife work hard to

keep their house clean after each bombing run that affects their neighborhood. They **wash the floor and dust. SC**
Joe's wife also helps the war effort. She goes every day to a nearby school which has been turned into a shelter for the homeless. There she takes care of **children and aged people who don't have a home. CO**
Joe's son Georgie plays "guard;" he takes his Dad's binoculars and watches out his bedroom window to see if any planes come flying over. Once he saw a **plane with the binoculars SV**. His Daddy told him that **flying planes can be dangerous SC**, so he ran to tell his parents. They **shut the dog up P** in the basement so that he wouldn't run out into the yard, and they all went down there themselves, too. Georgie asked why they had to go into the basement, but Joe didn't want to scare him, so he decided to just brush it off and not to **go into it PH** too much. **"Take it from me PH**, Georgie," said his father, "this is the safest place to be."
The family tries to do something nice each weekend. Joe's wife makes a picnic lunch with whatever is available at the time--any kind of **bread and cheese that they can get CO**. They are a little afraid to go too far away from a shelter, so they try to find a picnic **place near a house with a basement S** where they can go if there is any trouble. They lead as happy a life as they can under the circumstances.

Figure 6: The text read by the subjects

I.2.2. Questions

We designed four sets of questions using two modalities: spoken and written. For each modality we had two styles of questions: human-like ones, and machine-like ones.

The human-like ones were wordy and gave a plain and accurate 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 [Blanchon 1995b]. Thus they are more compact, somewhat more abstract and less descriptive.

I.2.2.1. Text dialogues

The following table gives the labeling of the text questions in the human and the machine styles. The ambiguous utterance is in bold type, and precedes the two proposed interpretations.

Human	Machine
the cost of the taxi to the library the cost of the taxi that is going to the library the cost of the taxi that will be paid by the library	**the cost of the taxi to the library** the cost of the taxi towards the library the cost of the taxi for the library
take down record in writing move something to a lower level	**take down** record take downwards
go by the highway take the highway to go there go by the side of the highway	**go by the highway** go via the highway go next to the highway
noisy cars and trucks noisy cars and not necessarily noisy trucks noisy cars and noisy trucks	**noisy cars and trucks** (noisy cars) and trucks noisy (cars and trucks)
go over pass or move over do again for practice	**go over** pass over review
talked about an hour the discussion concerned an hour the discussion took an hour	**talked about an hour** (talked about) an hour talked (about an hour)
the sign in front of the gas station with the red roof the sign has a red roof on it the gas station has a red roof on it	**the sign in front of the gas station with the red roof** the sign with the red roof the gas station with the red roof

stories floors in a building tales or narratives in books	**stories** floors narratives
he met the librarian as he was leaving Bill was leaving when he met the librarian the librarian was leaving when Bill met him	**he met the librarian as he was leaving** Bill was leaving the librarian was leaving
he told him he had had a terrible time that day he told him on the same day, that he had had a terrible time he told him that that day had been terrible	**he told him he had had a terrible time that day** that day, he told him he had had a terrible time he told him (he had had a terrible time that day)
date day on the calendar social engagement with someone	**date** day engagement
Old English professor a professor who teaches Old English an English professor who is old	**Old English professor** (Old English) professor old (English professor)
look into try to find out about look into the inside of	**look into** investigate look inside
medieval food and drink medieval food and not necessarily medieval drink medieval food and medieval drink	**medieval food and drink** (medieval food) and drink medieval (food and drink)
go with match or suit color or style go somewhere in the company of	**go with** match accompany
the dress and the cape with the feather collar the dress with the feather collar and the cape with the feather collar the dress with no feather collar and the cape with the feather collar	**the dress and the cape with the feather collar** the dress and (the cape with the feather collar) (the dress and the cape) with the feather collar
he brought a bottle of wine from France he brought a bottle of French wine he himself brought some kind of wine from France	**He brought a bottle of wine from France** He brought a bottle of (wine from France) from France, he brought a bottle of wine
she deposited her cape with the coat check clerk she deposited her cape and the coat check clerk together somewhere she gave her cape to the coat check clerk to take care of	**she deposited her cape with the coat check clerk** she deposited her cape and the coat check clerk she deposited her cape to the coat check clerk
bank the land along the side of a river a place where people deposit and borrow money	**bank** river side financial institution
French school of cooking cooking in the French style the school of cooking located in France	**French school of cooking** (French school) of cooking French (school of cooking)
The lamb did look ready to eat The lamb is going to eat The lamb is going to be eaten	**The lamb did look ready to eat** The lamb eats The lamb is to eat
eating their mutton stew with cranberries eating their mutton stew that has cranberries in it eating their mutton stew in the company of cranberries	**eating their mutton stew with cranberries** eating (their mutton stew with cranberries) with cranberries, eating their mutton stew
some are eating apples some are consuming apples some are the kind of apples that can be eaten	**some are eating apples** some eat apples some are (eating apples)
thin film technologist a technologist who is a specialist in thin film a film technologist who is thin	**thin film technologist** (thin film) technologist thin (film technologist)
he washes his face and dresses he washes his face and he washes his dresses he washes his face and he dresses	**washes his face and dresses** washes (his face and dresses) (washes his face) and dresses
he likes potatoes and cooks he likes potatoes and he likes cooks he likes potatoes and he cooks	**he likes potatoes and cooks** he likes (potatoes and cooks) he (likes potatoes) and (cooks)
They wash the floor and dust They wash the floor and they wash the dust They wash the floor and they dust	**They wash the floor and dust** They wash (the floor and dust) They (wash the floor) and (dust)

children and aged people who don't have a home children who don't have a home and aged people who don't have a home any kind of children, and aged people who don't have a home	**children and aged people who don't have a home** (children and aged people) who don't have a home (children) and (aged people who don't have a home)
he saw a plane with the binoculars using binoculars, he saw a plane he saw a plane that had binoculars on it	**he saw a plane with the binoculars** with the binoculars, he saw a plane he saw (a plane with the binoculars)
flying planes can be dangerous it can be dangerous to fly on or operate planes planes that are flying can be dangerous	**flying planes can be dangerous** using or operating planes can be dangerous (flying planes) can be dangerous
shut up to close something or someone into a room to make someone stop making noise	**shut up** close up make quiet
go into it explain it in more detail go inside of something	**go into it** explain it enter it
Take it from me take it away from me take my advice	**Take it from me** remove it from me believe me
bread and cheese that they can get any kind of bread but only cheese that they can get only bread they can get and cheese they can get	**bread and cheese that they can get** bread and (cheese that they can get) (bread and cheese) that they can get
a picnic place near a house with a basement the house itself has a basement the picnic place itself has a basement	**a picnic place near a house with a basement** a picnic place near (a house with a basement) (a picnic place near a house) with a basement

The text questions were displayed in a dialogue box as shown in the next figure.

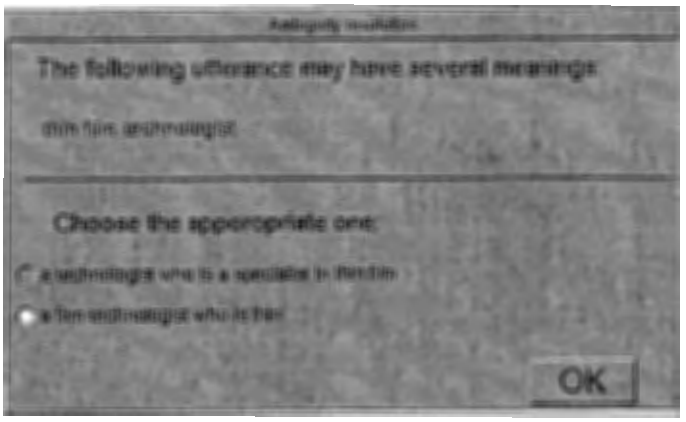


Figure 7: A dialogue box as presented to the subjects.

I.2.2.2. Spoken dialogues

For both of the spoken situations (Machine & Human), the proposed rephrasings were the same as in the text situations. they were played from a sound file.

In the spoken machine situation, the bracketing was replaced by pauses. The bracketed words were said in one phrase and there was a long pause between the bracketed phrase and the rest of the utterance.

The spoken questions were presented as follows:

"XX may have two meanings, please choose the appropriate meaning:
A, YY ;
B, ZZ"

Where XX is the ambiguous utterance, and YY & ZZ the proposed rephrasings.

II. Results

The following tables give the answers we got for each setting and for each subject.

II.1. Text Human

Question	Class	A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
taxi	D	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
takedown	PH	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
goby	D	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
noisy	C	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
goover	PH	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
talkabout	D	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
sign	S	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
stories	P	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
leaving	SV	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
told	SV	2	2	2	2	2	2	2	2	*1*	2	2	2	2	2	2	2
date	P	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
professor	S	1	1	*2*	1	1	1	1	1	1	1	1	1	1	*2*	1	*2*
lookinto	PH	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
medieval	CO	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
gowith	P	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
collar	CO	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
wine	SV	1	1	*2*	1	*2*	1	1	1	1	*2*	*2*	1	1	1	1	1
deposit	D	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
bank	P	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
school	S	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
lamb	D	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
cranberries	SV	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
apples	SC	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
technologist	S	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
face	SC	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
potatoes	SC	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
floor	SC	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
children	CO	1	*2*	*2*	*2*	*2*	1	1	1	*2*	*2*	*2*	1	1	1	1	*2*
binoculars	SV	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
flying	SC	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
shutup	P	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
gointo	PH	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
takefrom	PH	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
bread	CO	2	2	*1*	2	2	2	2	2	2	2	2	2	2	2	2	2
picnic	S	1	1	*2*	1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 8: Results in the text human setting

Comment

We observe a high rate of success for the subjects, except for the question about the coordination phrase containing the word "children;" this received seven correct answers and eight wrong answers. It seems we have a case here in which subjects did not use the text to choose an interpretation but based their decisions on their own convictions. One subject said "of course you will be helping all the children." Moreover, the wrong answers are not affected by the gender of the subject (five males & three females).

II.2. Spoken Human

Question	Class	A	16	18	19	20	21	22	23	24	25	26	27	28	29	30	31
taxi	D	1	1	1	1	1	1	2	2	1	1	2	1	1	1	1	1
takedown	PH	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
goby	D	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
noisy	C	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
goover	PH	2	2	2	2	2	1	2	2	2	2	2	1	2	2	2	2
talkabout	D	2	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2
sign	S	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
stories	P	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
leaving	SV	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
told	SV	2	2	1	1	2	2	2	2	1	2	2	1	1	1	2	1
date	P	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
professor	S	1	1	1	1	2	1	1	1	1	1	1	2	1	2	1	1
lookinto	PH	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
medieval	CO	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
gowith	P	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
collar	CO	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
wine	SV	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
deposit	D	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
bank	P	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
school	S	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
lamb	D	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
cranberries	SV	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
apples	SC	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
technologist	S	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
face	SC	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
potatoes	SC	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
floor	SC	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
children	CO	1	1	1	1	2	1	1	1	1	1	1	2	1	1	1	1
binoculars	SV	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
flying	SC	2	2	2	2	2	2	2	2	2	1	2	2	2	2	2	2
shutup	P	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
gointo	PH	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
takefrom	PH	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
bread	CO	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
picnic	S	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 9: Results in the spoken human setting

Comment

The answers to the coordination ambiguity containing "children" are very good. It may be because of the modality which more or less obliges the subject to choose the right interpretation on the fly. Once the first interpretation is chosen, the subject is less receptive to the second one.

The result for the question about "Tom telling Bill he had had a terrible time that day" is poor (eight correct answers and seven wrong). The result is excellent in the other settings. May be it was difficult to understand the difference between the rephrasings.

II.3. Text Machine

Question	Class	A	32	33	34	35	36	44	45	46	47	48	56	57	58	59	60
taxi	D	1	1	2	1	1	2	1	1	1	2	2	2	1	1	1	1
takedown	PH	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
goby	D	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
noisy	C	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
goover	PH	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
talkabout	D	2	2	2	2	1	2	2	2	2	2	2	2	2	2	2	2
sign	S	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2	2
stories	P	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
leaving	SV	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
told	SV	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
date	P	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
professor	S	1	1	2	1	2	1	1	1	2	2	2	1	1	1	1	1
lookinto	PH	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
medieval	CO	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
gowith	P	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2
collar	CO	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
wine	SV	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
deposit	D	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
bank	P	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
school	S	1	1	2	1	1	2	2	2	1	2	1	2	1	2	2	2
lamb	D	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
cranberries	SV	1	1	1	1	1	1	1	1	1	1	2	1	1	1	2	1
apples	SC	1	2	1	1	2	2	2	2	2	2	2	1	2	2	2	2
technologist	S	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
face	SC	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
potatoes	SC	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
floor	SC	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
children	CO	1	1	1	1	1	1	1	1	2	1	1	1	1	1	2	1
binoculars	SV	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
flying	SC	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2	2
shutup	P	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1
gointo	PH	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
takefrom	PH	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
bread	CO	2	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2
picnic	S	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 10: Results in the text machine setting

Comment

The ambiguities containing "taxi" and "professor" may have been a little tricky. On the other hand, we do have a real problem with the ambiguities containing "school" and "apples". Almost all the subjects were asked by the experimenter about their understanding of the sentences and they all understood the right meaning. The choices for the rephrasings were thus not labeled correctly.

II.4. Spoken Machine

Question	Class	A	17	37	38	39	40	41	42	43	49	50	51	52	53	54	55
taxi	D	1	1	1	1	2	1	1	1	1	1	1	2	1	1	2	1
takedown	PH	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1
goby	D	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
noisy	C	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2	2
goover	PH	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
talkabout	D	2	2	2	2	1	2	2	2	2	2	2	2	2	2	2	2
sign	S	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
stories	P	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
leaving	SV	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
told	SV	2	2	2	2	1	2	2	2	2	2	2	2	2	2	2	2
date	P	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
professor	S	1	2	2	1	1	1	1	1	1	1	1	1	2	2	1	1
lookinto	PH	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
medieval	CO	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
gowith	P	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
collar	CO	2	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1
wine	SV	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1
deposit	D	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
bank	P	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
school	S	1	2	2	2	1	1	2	2	2	2	2	2	2	2	2	2
lamb	D	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
cranberries	SV	1	1	1	1	1	1	1	2	1	1	1	1	1	2	1	1
apples	SC	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
technologist	S	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
face	SC	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
potatoes	SC	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
floor	SC	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
children	CO	1	1	1	1	2	1	1	1	1	2	1	1	1	1	1	1
binoculars	SV	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
flying	SC	2	2	2	2	2	2	2	2	2	1	2	2	2	2	2	2
shutup	P	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
gointo	PH	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1
takefrom	PH	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
bread	CO	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
picnic	S	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 11: Results in the spoken machine setting

Comment

The ambiguities containing "taxi" and "professor" seemed less tricky in the machine spoken than in the machine text condition.

The responses to the ambiguities containing "school" and "apples" are more accurate in this setting. For both questions almost all the subjects (except two) are in agreement with the answers.

For the resolution of the ambiguity containing "school," the technique is the same as the one used for the other ambiguities of this type. The problem of interpretation of the wording used here may be linked with the representation people have of the French style of cooking.

For the resolution of the ambiguity containing "apples" we used a slightly different technique than for the other ambiguity of the same kind, "Flying planes can be dangerous." For the "apples" ambiguity, the interpretations were labeled as follows: A – some eat apples, B – some are (eating apples). To be coherent with the interpretations for "planes," the labeling should have been: A – some are masticating apples, B – some are (eating apples). With this revised labeling the right interpretation may have been easier to understand. It is also to be said that the pronounciation of the second rephrasing for "apples" was following the normal way of saying that people were eating apples. This is why this choice has been unanimously chosen by the subjects.

III. Analysis

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

III.1. Statistical analysis

III.1.1. Actually collected data analysis

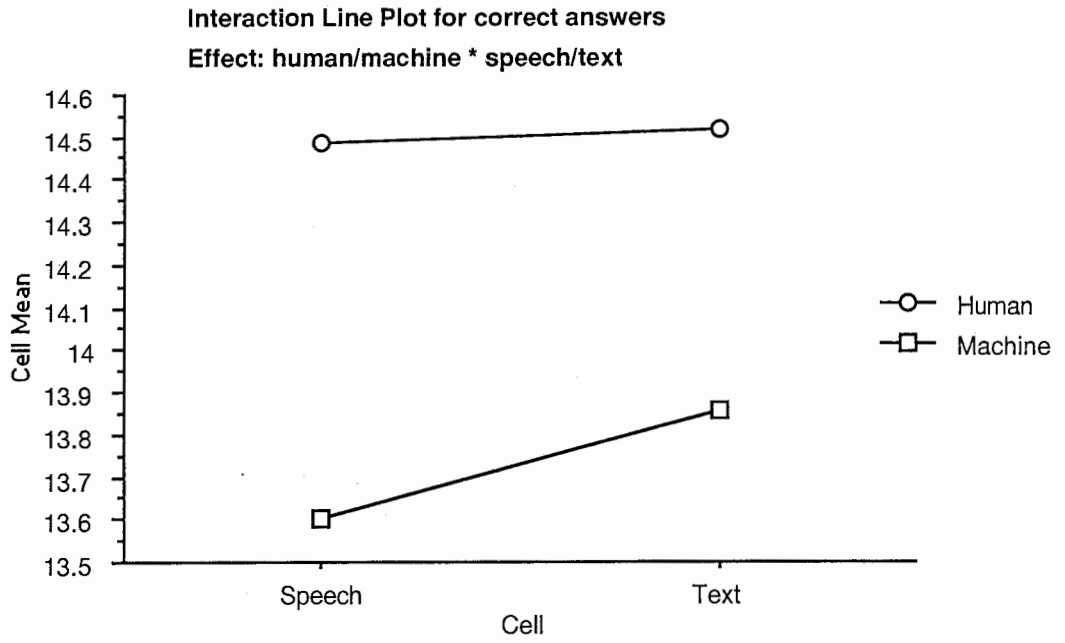


Figure 12: (Actual data) analysis

If we use the actual answers collected, the basic result is that the difference between the human setting and the machine setting is significant ($p<.05$). This is caused by the two

problematic questions (those with "school" and with "apples") we had in the machine setting. We will see in a different analysis below that the difference is not significant anymore.

It appears that in both machine-like and human-like phrasings, the performance of the subjects was better with text questions, but we can't draw any definitive conclusion since the differences between spoken and text dialogues seen in Figure 12 are not significant.

III.1.2. Filtering: ambiguities containing "school" and "apples" excluded from machine-like dialogue results

Because of the questions raised above about the appropriateness of the dialogues for the ambiguities containing "speech" and "apples," we also analyzed our results excluding the questions about those ambiguities from the results for the machine-like dialogues. In this case there is no significant difference between the subjects' performance in the machine-like dialogue settings and the human-like dialogue settings.

Subjects seem to show better performance for text dialogues over spoken dialogues in the human settings; however, there is no difference at all between text and speech in the machine settings. Again, the differences are not significant so no definitive conclusion can be drawn.

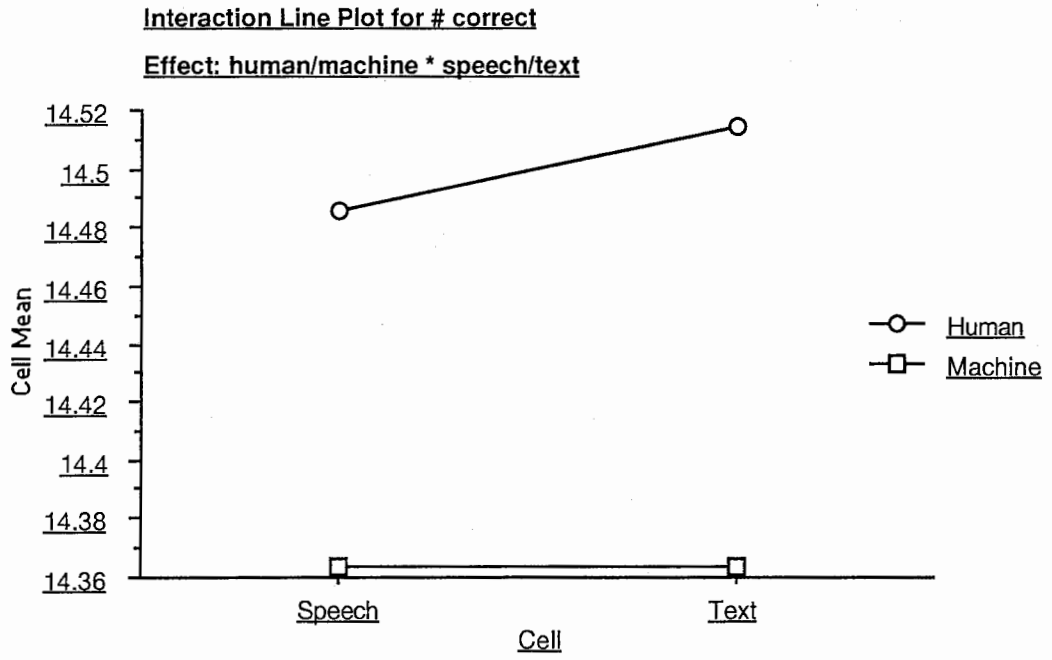


Figure 13: (Filtered Data) analysis

III.1.3. Projection: ambiguity containing "school" excluded from machine-like dialogue results and results for ambiguity containing "apples" corrected

In this third way of looking at the data, we excluded results for the ambiguity containing "school" from the machine setting results and considered that the answers to the dialogues about "apples" may have been different from the interpretation labeled as in the question about "flying planes". In this third case, there is again no significant difference between the machine-like and human-like dialogue settings.

Here the performances of the modalities are different from the machine-like and the human-like dialogue settings. The difference is again not significant; thus no definitive conclusion can be drawn.

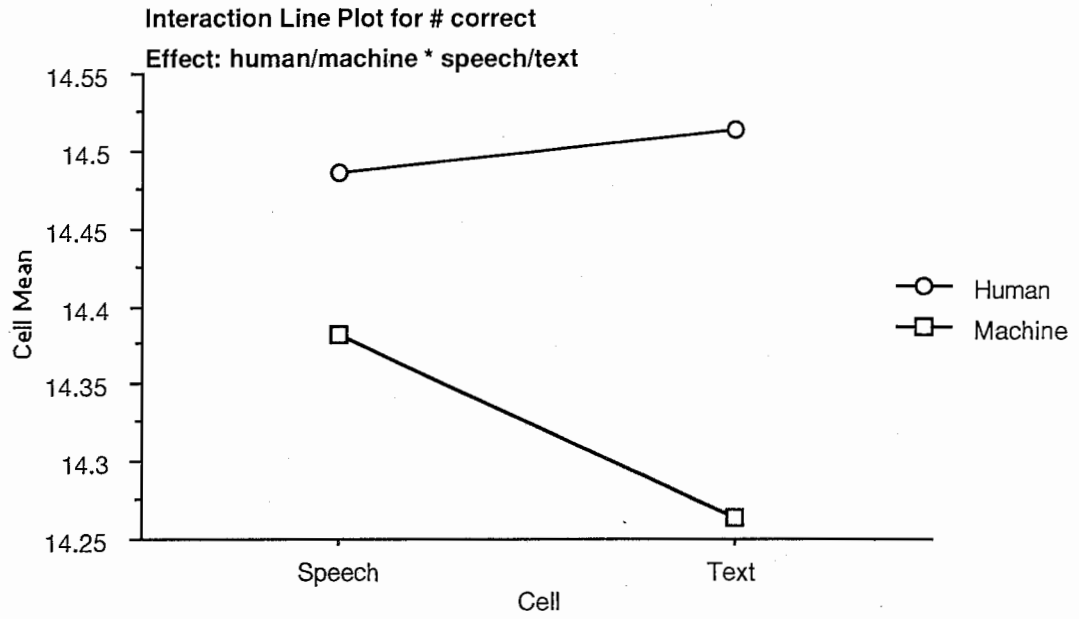


Figure 14: (Projected data) analysis

III.2. Behaviors

The behavior of the subjects can be evaluated according several parameters that were pertinent to the task.

III.2.1. Lecture speed and pause length

The lecture speed and the pause duration between the sentences were very different from one subject to another.

The wizard did not explicitly ask the fast subjects to slow down but rather let them realize that they were too fast for the system by presenting the disambiguation question after the beginning of the sentence following the ambiguity. Subjects eventually accommodated slightly by reducing their speed somewhat.

There was real accommodation and slowing down of the speech of the subject in the spoken settings. The slow speed of the speech used in the spoken disambiguation questions may have influenced the subjects.

III.2.2. Use of the "repeat" option

In the spoken settings, subjects could ask to have the disambiguation dialogue repeated. However, the repeat option was very infrequently used:

- 17 times in the spoken human setting over 525 questions, namely 3.24%;
- 19 times in the spoken machine setting over 525 question, namely 6.6%.

III.2.3. Answer time

The answer time was very different from one subject to another. The tapes have not been analyzed yet so no specific data are available at this time.

III.2.4. Comfort with the task

The subjects were very comfortable with the task and self confident about their answers. The task was easy to carry out. That is what the experimenters intended.

III.3. Questionnaires

Each subject was asked to answer a post-experiment questionnaire (cf. Appendix 3). Here are some of the results, with short comments.

III.3.1. About the ease of answering

	Very Easy	Quite Easy	Easy	Fairly Easy	Confusing
Text Human	2	2	4	8	1
Spoken Human	6	3		7	
Text Machine	4	3	4	3	6
Spoken Machine	3	3	3	8	2

Comment

The questions were thus easy to answer. Only 15% of the subjects felt that the answers were difficult.

III.3.2. Recommendations to make the questions easier to understand

	Proposal	Number
Text Human	– More context in the question	1
	– Better distinction between the items	2
Spoken Human	– Less monotonous voice	1
	– Shorter questions	5
	– Written choices	1
	– Be asked about my own text	1
Text Machine	– More context	3
	– More verbose choices	2
	– Brackets + bold the emphasis	1
	– User-proposed meaning	1
	– Grammar markers	1

Spoken Machine	– Example sentences instead of intonation & pauses – Shorter questions – Use of intonation – More detailed explanation available – Choice on the screen with brackets	1 3 1 2 2

Comment

In the spoken situations, several subjects (three) felt that they would have preferred to be asked questions in a written setting. That is a small number, but it shows that several settings could be available for the user to choose among.

In the spoken situation the meta-labeling to present the question should be as short as possible and intonation may be used more.

In the text machine situation, typography may be used as a complement to the bracketing to show the emphasis.

III.3.3. Strategies

	Strategy	Number
Text Human	– Use of the context – Substitution in the whole sentence	8 4
Spoken Human	– Use of the context – Substitution in the whole sentence	7 2
Text Machine	– Use of the context – Substitution in the whole sentence	9 3
Spoken Machine	– Use of the context – Substitution in the whole sentence	4 2

Comment

Substitution and use of the context are the most used techniques in all settings.

III.3.4. Predictability & recognition of patterns

	Predictability & recognition
Text Human	3
Spoken Human	3
Text Machine	2
Spoken Machine	5

Comment

Some of the subjects were able to tell us how predictable the question was, according to several syntactic patterns. We would have liked their number to be bigger but this does indicate that there are patterns in the disambiguation process.

Conclusion & perspectives

If we allow for the problematic questions concerning items with "school" and "apples," we see that there were no significant differences according to the style (machine, human) of the presentation of the disambiguation dialogues, and no significant differences according to the modality (spoken or text). The former result is essential to the success of an automatic interactive disambiguation program. We have seen that subjects are able to interpret the dialogues when presented in human-like, i.e., natural, phrasing, but it is not likely that automatically generated dialogues can be so natural. Therefore, it is critical that users be able to interpret the type of dialogues that machines are likely to be able to generate. The results reported here show that this is indeed the case.

We also investigated whether spoken or text dialogues would be easier to understand. This is a design question; it affects how an automatic system will be designed, but is not crucial to the system. The results found here, as well as comments made by some of the participants about wanting to have text instead of speech, suggest that one design feature for an interactive disambiguation system should be the option for users to choose in which modality they would like to have the dialogues presented. According to our results, both modalities are understandable.

Although the "repeat" option was not extensively used in the spoken setting, it is still necessary to include it for cases where users cannot understand the dialogue after the first hearing. Other suggestions made by the subjects can be easily implemented. For example, more of the context of the ambiguity can be included in the dialogue; this would also support the most frequent strategy used by the subjects in determining their responses. In addition, spoken utterances can be made shorter and faster. How best to use intonation in the spoken presentation of disambiguation dialogues is an open and interesting question.

It will be also necessary to run an experiment using as a textual support a text provided by each subject himself. This may be the only way to have a better analysis of the interactive disambiguation methodology we proposed.

Acknowledgment

The authors would like to acknowledge Mr Yamazaki and Mr Morimoto for their kind invitation of H. Blanchon in the framework of the joint MIDDIM project between ATR-ITL and the GETA-CLIP laboratories. They are very grateful to Mr Loken-Kim who have them access to the experiment room equipment. They finally want to thank Mr Kurihara who programmed the actual setting and set up the experiement.

References

Blanchon H. (1994). *Pattern-based approach to interactive disambiguation: first definition and implementation.* Rap. ATR-Interpreting Telecommunications Research Laboratories. Technical Report. n° TR-IT-0073. Sept. 8, 1994. 91 p.

Blanchon H. (1995a). *An Interactive Disambiguation Module for English Input: an Engine and the Associated Lingware.* Rap. ATR-ITL. Technical Report. n° TR-IT-0129. August, 1995. 110 p.

Blanchon H. (1995b). *An Interactive Disambiguation Module for English Natural Language Utterances.* Proc. NLPRS'95. Seoul. Dec 4-7, 1995, vol. **2/2** : pp. 550-555.

Blanchon H. & Fais L. (1995). *Pilot Experiment on the Understandability of Interactive Disambiguation Dialogues.* Rap. ATR-ITL. Technical repport. n° TR-IT-0130. September, 1995.

Blanchon H., Fais L., Loken-Kim K. H. & Morimoto T. (1995). *A Pattern-Based Approach for Interactive Clarification of Natural Language Utterances.* in Bulletin of the Information Processing Society of Japan (95-NL-107). vol. **95**(52) : pp. 11-18.

Lehiste I. (1973). *Phonetic disambiguation of syntactic ambiguity.* in Glossa. vol. **7**(2) : pp. 107-122.

Lehiste I., Olive J. P. & Streeter L. A. (1976). *Role of duration in disambiguating syntactically ambiguous sentences.* in Journal of the Accoustical Society of America. vol. **60**(5) : pp. 1199-1202.

O'Malley M. H., Kloker D. R. & Dara-Abrams B. (1973). *Recovering Parentheses From Spoken Algebraic Expressions.* in IEEE Transactions on Audio and Electroacoustics. vol. **AU-21**(3) : pp. 217-220.

O'Shaughnessy D. (1989). *Specifying accent marks in French text for teletext and speech synthesis.* in International Journal of Man-Machine Studies. vol. **31**(4) : pp. 405-414.

Streeter L. A. (1978). *Acoustic determinants of phrase boundary perception.* in Journal of the Accoustical Society of America. vol. **60**(6) : pp. 1582-1592.

Appendix

A.1. Instructions (textual setting)

The purpose of this experiment is to test a language processing system. Specifically, we are testing a module that helps the system "understand" the correct meaning of sentences. This system will be monitored by someone who may interact with you in case of a problem.

You are going to read a text presented on a computer screen, and the system is going to ask questions about several sentences. You should answer those questions as accurately as you can.

Reading

- Read aloud **slowly and carefully trying to understand what you are reading as much as possible** since you may be asked questions about the sentences you are reading.
- Pause at the end of each sentence. If you do not pause, the system will have difficulty understanding your speech.

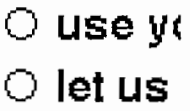
Scrolling

- Scroll down the text **page by page** with **one** click at the bottom of the scroll bar as shown to the right.

Questions

Selection

- When a question is presented, select the appropriate meaning by clicking on the corresponding circle. If you change your mind, you may change your selection at this point.



Validation

- Once you are sure about your choice, click on the button "OK".

Continuation

- after you click "OK," continue reading.

A.2. Instructions (spoken setting)

The purpose of this experiment is to test a language processing system. Specifically, we are testing a module that helps the system "understand" the correct meaning of sentences. This system will be monitored by someone who may interact with you in case of a problem.

You are going to read a text presented on a computer screen, and the system is going to ask questions about several sentences. You should answer those questions as accurately as you can.

Reading

- Read aloud **slowly and carefully trying to understand what you are reading as much as possible** since you may be asked questions about the sentences you are reading.
- **Pause** at the end of each sentence. If you do not pause, the system will have difficulty understanding your speech.

Scrolling

- When you are ready to go on to the next page, please say "next page." The system will automatically change to the next page.

Questions

Selection

- After some sentences, the system will interrupt you to ask you a question orally. You will be given two choices for the meaning of the sentence you just read. Choose the appropriate meaning by saying either "A" or "B." If you want to **change** your answer, you may. If, for example, you chose "A" and you want to choose "B" instead, please say "not A, B."

Repeat

- If you would like to hear the choices again, please say "Repeat."

Validation

- Once you are sure about your choice, say "OK."

Continuation

- after you say "OK," continue reading.

A.3. Post-experiment questionnaire

Please answer the following questions in as much detail as you can.

1. How hard or how easy was it to answer the questions?

2. Can you think of anything that would have made it easier to answer the questions?

3. What kinds of strategies did you use to answer the questions?

4. Any other comment:

Thank you for your help.