

TR-IT-0064

Ambiguity Analysis  
and  
MIDDIM-DB

Mutsuko Tomokiyo K.H.Loken-Kim

August 1994

Acknowledgement

We would like to heartily thank Dr.Pr.F.Peccoud (director of GETA), Dr.Pr.Ch.Boitet, Dr.Y.Yamazaki (president of ATR-ITL), T.Morimoto (department head), Dr.H.Blanchon for their constant support to this research and also researchers at GETA and ATR for interesting discussions and their technical support.  
Special thanks should go to J.Winship who programmed the database system.

## Abstract

This is the report on a cooperative research at GETA from 9 May 1994 to 8 July 1994 in the framework of a three year contract, signed last year between CNRS-IMAG-GETA and ATR-ITL. In fact, this research was performed as the third business trip in this contract, following H.Blanchon stay at ATR-ITL and K.H.Loken-Kim stay at GETA. This collaboration project, called "MIDDIM", aims at studying multimodal interactive disambiguation methods for spontaneously spoken dialogues as well as for written texts, for the Japanese-French-English languages.

The research also aims at building an ambiguity database, which will be used to help building a speech dialogue system as well as a machine translation system in the near future.

## Keywords

ambiguity, ambiguity classification, ambiguity database, ambiguity, representation formalism, disambiguation, multimodal interactive disambiguation, spoken language

# Contents

<b>1</b>	<b>Research activities</b>	<b>3</b>
1.1	Research team . . . . .	4
<b>2</b>	<b>Ambiguity analysis</b>	<b>5</b>
2.1	Chosen texts . . . . .	5
2.2	Ambiguity representation formalism . . . . .	7
2.3	Ambiguity classification . . . . .	9
2.4	Statistics of ambiguities . . . . .	17
2.5	Study toward a speech dialogue system . . . . .	18
<b>3</b>	<b>The MIDDIM-DB</b>	<b>22</b>
<b>Appendix</b>		
	<b>Ambiguity analysis results</b>	<b>23</b>
A.1	Japanese dialogue text-2a . . . . .	23
A.2	Japanese dialogue text-2b . . . . .	29
A.3	English dialogue text-5a . . . . .	37
A.4	English dialogue text-5b . . . . .	41
A.5	French written text-introduction of a scientific article . . . . .	50
A.6	English written text-introduction of a scientific article . . . . .	56
A.7	French transparency . . . . .	62

# 1 Research activities

Our collaboration research aims at studying multimodal interactive disambiguation methods for spontaneously spoken dialogue as well as for written texts for the Japanese, French, and English. For these aims, we started with ambiguity analysis. Recognition and analysis ambiguities in machine translation systems or dialogue interpretation systems are introduced by:

- the recognizer : *e.g.* bad recognition or misrecognition
- the clients : *e.g.* speech variability, incoherent input, use of out-of-vocabulary words, spontaneous input, change of mind
- the languages : *e.g.* homophony, polysemy, uncompleted sentences, ellipsis, absence of reference, etc.
- the communication channel quality : *e.g.* bandwidth or noise

Within the context of our overall research plan and according to "Program for Mrs. Tomokiyo" (1 & 3 March, 1994) by Ch.Boitet and K.H.Loken-Kim, we have concentrated on the analysis of ambiguities which are introduced by the languages. The analysis of the ambiguities due to the languages consists of :

- dialogue corpus analysis (in English and Japanese) → Appendix 1
- written text analysis (in English and French) → Appendix 2
- transparencies analysis (in English and French) → Appendix 3
- a classification of ambiguities → Section 2.3
- an ambiguity representation formalism → Section 2.2
- a contribution to the building of the MIDDIM database. → Section 3

Each item is detailed below.

## 1.1 Research team

We worked with Dr.Pr.Ch.Boitet, M.Axtmeyer, N.Nédobejkine and J.Winship

## 2 Ambiguity analysis

### 2.1 Chosen texts

It was decided to define a corpus for the base consisting of different types of documents in order to be able to study ambiguities in different contexts. The chosen types were articles, transparencies and transcribed dialogues, notably because of the fact that there is often call for their translation. As MIDDIM-DB was just being built, we started with a small subset of the corpus which becomes bigger at the GETA and at ATR. More precisely, for the transcribed dialogues, four are taken from the base of guidance dialogues <sup>1</sup> produced by ATR, and for the articles, two introductions, two abstracts and for the transparencies, 2 multimodal-concerning transparencies have been chosen from the GETA corpus.

corpus name	num. of Utts.	num. of words	num. of pause units	language
2a	23 utterances	163 words	81 pause	Japanese
2b	58 utterances	306 words	155 pause	Japanese
5a	21 utterances	328 words	69 pause	English
5b	30 utterances	586 words	119 pause	English
Total	132 utterances	1383 words	424 pause	

Example 2a

---

#### <sup>1</sup>Notes

We used transcription versions of the dialogue texts to analyze. They contain no pause units, pronunciations, punctuations, nor capital letters, with a few exceptions. Actually the transcription method should be studied according to its usage. New concept of segmentation unit in an utterance might be researched on in stead of *word* or *sentence* for the transcription of spoken texts, although it depends on speech recognition results.

*e.g.*

Good morning conference office can I help you  
[gú d mɔniŋkánfərənsɔfiskənai / hélp / ju]

When we listen to the cassette tape, the above sentence is segmented into 3 by pause units : after *I*, *help* and *you*. So, the segmentation is different from meaning units, because the utterance should be segmented into 3 units according to the meaning : *Good morning/ conference office/ can I help you*.

A : はい / (こ) / 京都国際交流センターです  
 TM : [あ / あの /] / わたくし三浦と申します

Example 5a

A : OK you're at Kyoto Station how would you like to get thi Interna-  
 tional Conference Center we have [ uh ] three methods of travel We have  
 the subway the bus or (th') a taxi  
 BL : [ um ][ um ] which way is the easiest /laugh/

We have chosen the following written texts to analyze ambiguities.

corpus name	articles	num. of sentences	num. of words	language
CB5	abstract	3 sentences	86 words	English
CB6	introduction	13 sentences	443 words	English
ML7	abstract	9 sentences	162 words	French
CB8	introduction	11 sentences	452 words	French
GF9	transparency	12 lines	50 words	French
GF10	transparency	10 lines	59 words	French
Total		36 sentences	1252 words	

Example CB6

Opinions on Machine Translation (MT) are often extreme. Some consider it only as a testbed for the scientific experimentation of their favorite theories or formalisms, while others see it as a purely technological and utilitarian enterprise. In most cases, each defends a particular paradigm, as one would defend an ideology.

Example CBS

Les avis sur la traduction automatique (TA) sont souvent extrêmes. Les uns ne la conçoivent que comme l'expérimentation scientifique de leurs théorie ou de leurs formalismes favoris, tandis que les autres y voient une entreprise purement technologique et utilitaire. Le plus souvent, chacun défend un "paradigme" particulier, comme on défendrait une idéologie.

## 2.2 Ambiguity representation formalism

Ambiguities are investigated at the morpho-syntactic, semantic and pragmatic levels and are represented by an attribute-value pair formalization. The attribute/value pairs appear recurrently in a representation.

Each representation has identification number, ambiguity classification number as the attributes and contains "object" and "status" as necessary values. The value of the "object" is words, phrases, clauses, sentences or fragments in texts to be analyzed. The value of the "status" is "accidental" or "fatal". "Accidental" is used for structural ambiguities which would not be regarded as ambiguous by the human reader, but for which the system lacks the information required to disambiguate. "Accidental" structural ambiguities occur due to an accidental combination of words having category ambiguities, due to alternative grammatical uses for syntactic constituents, or due to different possible combination of syntactic constituents. "fatal" is used for structural ambiguities which would be regarded as ambiguous even by the human reader. The following is well known as an example of "fatal" ambiguity : The man saw the girl with the telescope.

The other attributes appear according to ambiguity variations : *e.g.* an attribute "reference" is used for ambiguities of the anaphora, and an attribute "information known" is used for ambiguities due to ellipsis.

*e.g.*

```
(ambiguity42-5.2 ((object "yes")
                  (status fatal)
                  (IFT (agreement discourse-indication
                       information-offer))))
```

In this example, "42" is ambiguity identification number, and "5.2" is ambiguity classification number (→see 2.3 Ambiguity classification), which would be related to ambiguity resolution methods<sup>2</sup> For example, "yes" is here analyzed as an object of the ambiguity due to speaker's intention, so the speaker's intention is represented by the IFT ( *illocutionary force type* ) attribute and illocutionary force labels as its values. All attributes and values are shown below in "Ambiguity classification".

---

<sup>2</sup>We have started to investigate ambiguity resolution methods. However, it's difficult to finish for 2 month.



## Notes

We can also represent ambiguities by following alternatives. However, all ambiguities should be represented in the same way, the formalism we mentioned above is better than the followings:

```
(defambiguity cases noun
  (syn ((cat noun)
        (num plurie)))
  (sem ((LEX1 "cas")
        (LEX2 "caises"))))
```

```
(ambiguity "favorite" in "favorite théories or formalism"
  favorite -> théories
  favorite -> théories or formalism)
```

```
((NP -> ADJ N or N)
 (NP -> ADJ N or ADJ N)
 (ADJ -> favorite)
 (N -> théories)
 (N -> formalism))
```

## 2.3 Ambiguity classification

Ambiguities are classified from the point of view of a tricotomy of morpho-syntax, semantics and pragmatics. The method of classification would share a lot of opinions. However the important point is that all ambiguous phenomena should be analyzed by the same criterion for spoken and written texts, and it should be also represented by the same formalism. Here is an ambiguity classification list where all attributes and all values used in our ambiguity formalism are defined.

### 1. Categorical ambiguity

#### 1.1 morphological ambiguity :

attribute : category

value : verb, noun, pronoun, relative noun, adjective, adverb, conjunction, interjection, determiner

*e.g.*

```
(ambiguity17-1.1 ((object "that")
                  (status accidental)
                  (category (relative-noun pronoun))))
```

#### 1.2 syntactic ambiguity

attribute: function

value : gerund, participle present, noun, padv, padj, smodifier

*e.g.*

```
(ambiguity27-1.2 ((object "by building systems")
                  (status accidental)
                  (function (gerund participle-present noun))))
```

### 2. Structural ambiguity

#### 2.1 words, expression, segmentation

attribute : segmentation

value : texts

*e.g.*

```
(ambiguity28-2.1 ((object "right")
                  (status accidental)
                  (segmentation ("right" "right now"))))
```

## 2.2 grammatical constituent structure

### 2.2.1 coordinate or apposition<sup>3</sup>

attribute : coordinate : apposition

value : text or x..x<sub>n</sub>, y, y1, y2

#### Notes

x=previous utterance, x1..x<sub>n</sub>=previous utterance of same speaker,  
y=current utterance, y1=next utterance, y2=next utterance of same  
speaker)

### 2.2.2 subordinate

attribute : structure

value : text or x, x1..x<sub>n</sub>, y, y1, y2

*e.g.*

```
(ambiguity950-2.2.2 ((object "their favorite theories or formalisms")
                    (status accidental)
                    (structure (<their favorite theories
                               or their favorite formalism>
                               <<their favorite theories>
                               <or formalisms>>
                               <<their favorite theories>
                               <or their formalism>>))))))
```

---

<sup>3</sup>"coordinate" is used for the coordinate with conjunction of noun phrases and clauses.

### 2.2.3 specific or generic

attribute : specific / generic

value : only one, 1/2, possessive, modifier, certain

#### Notes

*only one* is used for the object of noun phrases existing as the unique existence in the world. *1/2* is for noun phrase meaning one of two that exists. *possessive* is for noun phrase with possessive pronoun. *modifier* is for noun phrase with adjectival phrase. *certain* is for noun phrase which is specified with nothing.)

*e.g.*

```
(ambiguity 901-2.2.3 ((object "some")
                      (status accidental)
                      (specific 1/2)
                      (generic certain)))
```

### 2.2.4 aspect, modality and tense

attribute : aspect, tense, modal

value :tense-future, present, past, future-in-past

:aspect- progressive, accomplishment, state

:modal-politeness, euphemism, duty, likelihood, opinion,

desire, presupposition, possibility

*e.g.*

```
(ambiguity 901-2.2.4 ((object "would")
                      (status accidental)
                      (modal (euphemism politeness)
                              (tense future-in-past)))
```

### 2.2.5 quantifier scope

attribute : qscope  
value : text

## 2.2.6 negation scope

attribute : nscope  
value : text

## 2.2.7 anaphora

### 2.2.7.1 pronoun

attribute : reference or deictic  
value : text or grammatical category  
*e.g.*

```
(ambiguity 18-2.2.7.1 ((object "some")  
                        (status accidental)  
                        (reference (to-inf "the main ticket  
                                     office")))))
```

#### **Notes**

The ambiguity is whether *specific* or *generic* is it.

### 2.2.7.2 adjective

attribute : reference or deictic  
value : text

### 2.2.7.3 adverb

attribute : deictic  
value : text, grammatical category or explanation  
*e.g.*

(ambiguity 44-2.2.7.3 ((object "here")  
(status fatal)  
(deictic ("on the screen" speaker's place  
"Kyoto station"))))

2.2.8 ellipsis  
2.2.8.1 simple

attribute : ellipsis  
value : text

2.2.8.2 ellipsis due to information known

attribute : information known  
value : text  
*e.g.*

(ambiguity69-2.2.8.2 ((object "each")  
(status accidental)  
(specific 1/2)  
(information known "some and others"))))

2.2.8.3 implied

attribute : implied  
value : text  
*e.g.*

(ambiguity814-2.2.8.3 ((object "d'autres approches")  
(status accidental)  
(implied ("construire des  
systèmes multilingues"  
"de passer par un interlingua"))))

2.2.9 inversion

attribute : inversion  
value : text

#### 2.2.10 repeat

attribute : repetition  
value : text

#### 2.2.11 insertion or revision

attribute : insertion or revision  
value : text

#### 2.2.12 ungrammatical utterance

attribute : ungra  
value : text

### 3. Semantic ambiguity

#### 3.1 polysemy

##### 3.1.1 internal polysemy

attribute : lex  
value : text

*e.g.*

```
(ambiguity 41-3.1.1 ((object "cases")  
                    (status fatal)  
                    (lex ("box" "example"))))
```

##### 3.1.2 external polysemy

attribute : Japanese, French, English  
value : text

*e.g.*

(ambiguity 901-3.1.2 ((object "calling")  
(status fatal)  
(Japanese ("電話する " "呼ぶ")))))

### 3.2 substitution

attribute : substitution  
value : text

## 4. Logico-semantic ambiguity

attribute : logicosem  
value : purpose, intention, result, cause, suppose, suggestion

*e.g.*

(ambiguity 901-4 ((object "in order to")  
(status accidental)  
(logicosem (purpose intention))))

## 5. Discourse ambiguity

### 5.1 speaker's intention

attribute : IFT  
value : discourse-indication, conjunctive, deixis, information-offer,  
information-request, attention-request, attention-offer, suggest, in-  
vitation, repetition-request, repetition, human-relation, agreement-  
offer, agreement-request, intention

*e.g.*



```
(ambiguity 901-2.2.3 ((object "yes")
                      (status accidental)
                      (IFT (agreement discourse-indication
                            information-offer))))
```

## 5.2 real world

attribute : situation  
value : text  
*e.g.*

```
(ambiguity 15-5-5.2 ((object "the second floor")
                     (status fatal)
                     (situation (rez-de-chaussé-0 re
                                  z-de-chaussé-1))))
```

## 2.4 Statistics of ambiguities

Among three languages, French, English and Japanese, there are some ambiguity tendency. For example, structural ambiguities are found in English text rather than Japanese or French texts. Ambiguities due to the omission of subjects or objects are found typically in Japanese text, contrastively with ambiguities due to the omission of implied meaning representation which are found in all texts.

Anaphoric ambiguity by pronouns is found notably in English texts. Deictic ambiguity appears typically in spoken texts. *Fatal* ambiguities are in general due to polysemy and *accidental* ambiguities to structural specification problems.

The followings are statistics of ambiguities according to the ambiguity classification.

Ambiguity	Japanese	English	French	total	%
morphological	1	1	0	2	0.74
syntactic	0	1	0	1	0.27
word, expression, segmentation	5	8	0	13	4.85
coordinate or apposition	1	3	0	4	1.49
subordinate	2	11	1	14	5.22
specific or generic	0	2	1	3	1.12
aspect, modality and tense	20	4	0	24	8.96
quantifier scope	0	0	0	0	0.00
negation scope	0	0	0	0	0.00
pronoun anaphora	7	26	8	42	16.0
adjective anaphora	1	5	4	10	3.73
adverb anaphora	7	13	2	22	8.21
simple ellipsis	29	1	0	30	11.0
information known	15	2	9	26	9.7
implied	4	3	3	10	3.73
inversion	0	0	0	0	0.00
repeat	1	0	0	1	0.37
insertion or revision	9	19	2	20	7.46
ungrammatical	0	0	0	0	0
internal polysemy	1	3	0	4	1.49
external polysemy	4	8	5	17	6.34
substitution	0	0	0	0	0.00
logico-semantic	0	2	3	5	1.87
speaker's intention	14	4	0	18	6.72
real world	5	16	5	26	9.7

## 2.5 Study toward a speech dialogue system

We tried to show entirely ambiguities due to the language in natural language processing (NLP) through our ambiguity analysis.

Let's take up here a speech interpreting system, which is one of NLP and a matter of our current concerns.

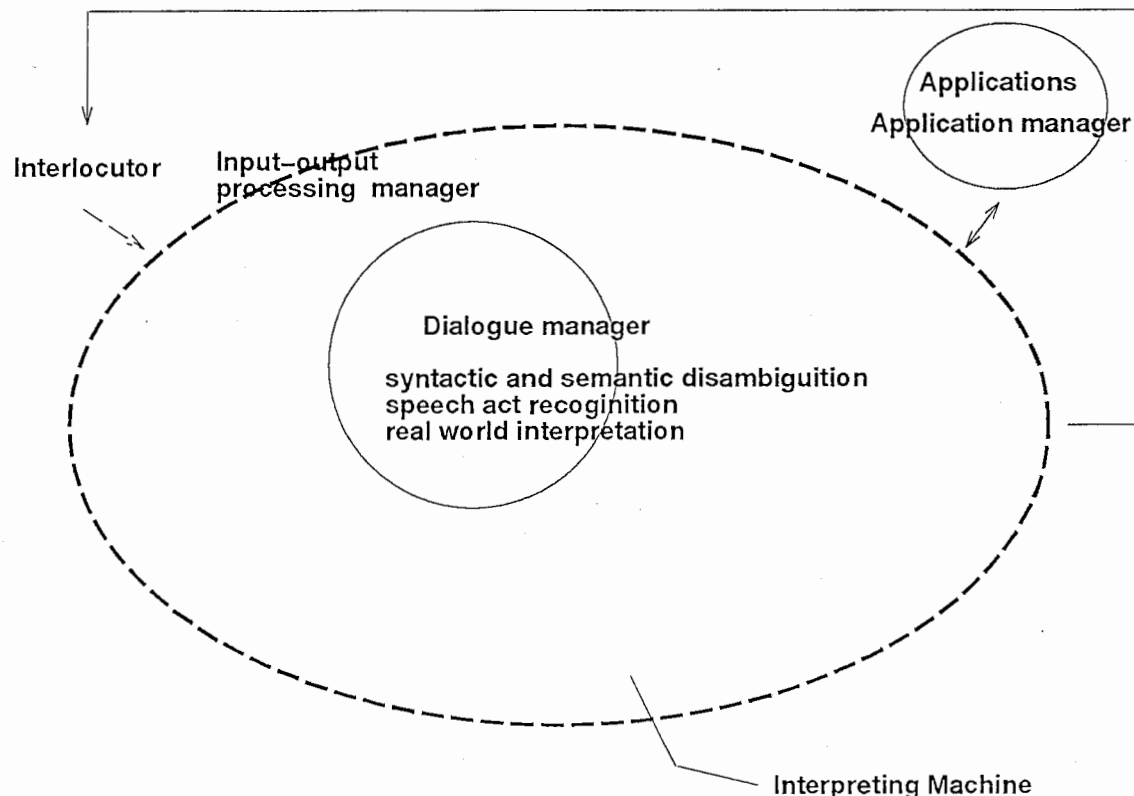


Figure 1: speech dialogue system

Monolingual speech interpreting system is quite different from multilingual speech dialogue system<sup>4</sup>. There are translations no longer in the monolingual speech interpreting system, but correct interpretation of speech acts in utterances should be explicitly made. We presume that a interactive method is taken in order to make

<sup>4</sup>A *Human Aided Machine Translation*(HAMT) system for multilanguages has been proposed by Christian Boitet(1993), where the human interpreter plays a role as supervisor, so the human interpreter is an expert of the system and he should *supervise the system, advice its clients, and adjust its parameters and/or take over if necessary*. When disambiguation is done interactively with the interlocutor, interactive way will be different between monolingual dialogue system and HAMT.

the interpretation unique for utterances, where disambiguation of interpretations relies on the assistance of interlocutors.

We have discovered communicative acts ambiguity, deictic anaphora ambiguity, polysemy and so forth through spoken dialogues analysis. If these ambiguities are not resolved, they cause misunderstanding or unnaturalness of resulting dialogue in monolingual or multilingual dialogue system, whatever.

Discussions focus on necessity and naturalness of interaction activity in monolingual interpretation system.

The system should be basically in charge of the dialogue activity, *viz.* input/output processing and syntactic/discourse analysis and in control of the applications. We presume that the architecture of the speech dialogue system consists of :

1. input/output manager, which accepts speech input/output and text input/output at least,
2. application manager, which controls the communication between the server and the application,
3. dialogue manager, which interprets the input coming from the input/output manager and drives the execution of the application and which lets maintain conversation flow by putting question or extending utterances in case of ambiguities, errors and so on.

What is asked by the system? To which aim? How to reduce the gap between the client and the system, or what is efficient interaction? On the one hand, it is important for the system to ensure an optimal interpretation; on the other hand, it is important to implement the continuous confirmation of hypothesis, ensuring at the same time a maximum naturalness of the resulting dialogue. The questions from the system are made to clarify, confirm, expand the client's utterances, to explain, to emphasize, to repeat important information and to give adequate information. To clarify and confirm is generally done by asking the interlocutor.

*example 1*

Tell me the train from Ulm to Hamburg around eight o'clock tomorrow.

There are two interpretations concerning the prepositional phrase attachment in *e.g.1*. In addition, there is an ambiguity of the time expression in this example. If the system puts a question to the client whether "around eight o'clock tomorrow" modifies "the train from Ulm to Hamburg" or "tell me", the client will not be happy, because the client would expect that such kind of ambiguities are automatically resolved. Contrastively to this, if the system puts a question to the client whether

"eight o'clock" is of p.m or a.m, the client would answer with pleasure. Because the answer to the later question is what only the client knows and it's an inevitable question to be put to the client. This means that what kind of interrogation from the system is optimal is one of important studies toward building a speech dialogue system.

In our ambiguity analysis, the ambiguity of the prepositional phrase attachment is analysed as an "accidental" ambiguity and it must be resolved by the system automatically. The ambiguity of the time expression is analysed as a "real world" ambiguity and it is the object of asking to the interlocutor.

The followings are examples concerning IFT ambiguity, adverb anaphora ambiguity, etc.

*example 2*

Agent : Where are you located now?

Client : I'm at Kyoto station

Agent : You are at Kyoto station. OK. Where are you located at Kyoto station?

(In the following examples, "Agent" is presumed to be the system and "Client" be the client.)

When a client inquires about itineraries, "Where are you located now?" from the agent is an important question. The agent also has to confirm by asking "You are at Kyoto station. OK". In addition, the agent might ask "Where are you located at the Kyoto station?" to get more details about client's place.

*example 3*

Client : I'm trying to find my way to the International Conference Center for tomorrow's symposium on interpreting telecommunications.

Agent : the International Conference Center for a symposium on Monday 20th of June?

The agent understands client's question and extends the deictic adverb "tomorrow" to a precise date, Monday the 20th of June in *e.g.3*. Because, a lot of symposiums will take place at the International Conference Center, the real date for tomorrow is needed in order to confirm.

*example 4*

Agent : Subway, bus, or taxi: what would you prefer?  
Client : Probably subway if it's easy.  
Agent : It's pretty easy you're already at Kyoto station.

The agent answers to the question which is put with if-clause in *e.g.4*.

*example 5*

Agent : Subway, bus or taxi what would you prefer?  
Client : Probably subway if it will rain.  
Agent You will take subway?

It's not necessary for the agent to answer to the if-clause, but the agent has to confirm whether the client wants to take the subway in *e.g.5*.

*example 6*

Agent : You will take the Keihan-Kyotsu Line to the Keage Station.  
Client : Kiage.  
Agent : Keage.  
Client : Keage Station.

The agent doesn't stop repeating until the client arrives at the correct pronunciation for "Keage Station " in *e.g.6*.

*example 7*

Client : OK. You said that 's two stops.  
Agent : Yes, Keage Station is the second stop.

The agent paraphrases "two stops", because the client doesn't understand whether it's the second stop or the third stop in *e.g.7*.

*example 8*

Agent : Walk to the first floor and out the north end of the building.

"the first floor" is used for the first floor and also the second floor, by counting floors of the building from the surface of a ground. French people count "rez-de-chaussée", "premier étage (first floor)" "deuxième étage"...., so a nice agent might explain for asked itineraries according to client's counting manner.

### 3 The MIDDIM-DB

The analysis results of ambiguities are put into the MIDDIM-DB(MIDDIM database) system. The MIDDIM-DB is created using Hypercard, which is an application for the Macintosh. The principal aim of building the MIDDIM-DB is :

- to use it for resolving analysis ambiguities on machine translation systems or (man-machine) dialogue systems
- to investigate and clarify what kind of ambiguities there are, when machine translation is concerned, which are ambiguities due to, how to resolve ambiguities.

The MIDDIM-DB contains morpho-syntactic, semantic and pragmatic information of ambiguous phenomena of the language in NLP. So it can also be considered as a core knowledge of discourse or semantic analyzer namely in speech dialogue system, because the ambiguities concerning the discourse are indicated as illocutionary force type ambiguity and the ambiguities concerning the general semantics are indicated as polysemy, anaphora, negation scope, ellipsis, etc. For speech dialogue system, the MIDDIM-DB serves to decide what kind of question should be put to the interlocutor as we mentioned above.

It contains concrete ambiguity examples and enables us to calculate statistics of ambiguity phenomena, so it can be useful to make a stochastic grammar.

## A Ambiguity analysis results

The analysis results of Japanese dialogues 2a,2b, English dialogues 5a,5b, French-English articles, and French transparency are as follows.

Texts are shown with identification number. The identification number is correspondent to the first number in the ambiguity representation. The second number is the ambiguity classification number.

### A.1 Japanese dialogue text-2a

A: はい (100) こ (101) 京都国際センターです (102,103,209)

```
(ambiguity100-5.1 ((object "はい")
                  (status accidental)
                  (IFT (human-relation discourse-indication
                       agreement-offer))))
```

```
(ambiguity102-2.2.11 ((object "こ")
                     (status accidental)
                     (insertion "こ")))
```

```
(ambiguity209-2.2.8.1 ((object "京都国際交流センターです")
                       (status accidental)
                       (ellipsis ("こちらは" "それは"))))
```

```
(ambiguity103-5.2 ((object "京都国際交流センターです")
                   (status accidental)
                   (situation (telephone something))))
```

TM: あ あの (104) わたくし (105) 三浦と申します

```
(ambiguity104-2.2.11 ((object "あ あの")
                      (status accidental)
                      (insertion ("あ あの"))))
```

```
(ambiguity105-2.2.8.1 ((object "わたくし")
                       (status accidental)
                       (ellipsis ("は" "が" ))))
```

A: はい (106)

```
(ambiguity106-5.2 ((object "はい")
```



(status accidental)  
(IFT (human-relation discourse-indication  
agreement-offer))))

TM: えっと (107) これから (108) ですね (109) えっと (107) 本日開催されます  
えーと (110) マルチメ (111) えっと (107) マシン エディット トランスステーショ  
ンにおけるマルチメディア文脈の役割についてという講演を聞きに行きたい (112) ん  
ですけれども (113)

(ambiguity107-2.2.11 ((object "えっと")  
(status accidental)  
(insertion ("えっと")))))

(ambiguity108-2.2.7.3 ((object "これから")  
(status fatal)  
(deictic at the time of day))))

(ambiguity109-5.1 ((object "はい")  
(status accidental)  
(IFT (discourse-indication attention-request  
information-request))))))

(ambiguity110-2.2.11 ((object "えっと")  
(status accidental)  
(insertion ("えーと")))))

(ambiguity111-2.2.11 ((object "マルチメ")  
(status accidental)  
(revision "マルチメ"))))

(ambiguity112-2.2.8.1 ((object "聞きに行きたい")  
(status accidental)  
(ellipsis "わたしは"))))

(ambiguity113-5.1 ((object "んですけれども")  
(status accidental)  
(IFT (attention-request information-request))))))

A: はい (109) TM: えーと (110) その (114) かいえ (115) そこ (116) まで行く (117)  
のに (118, 119) えーと (110) どのように行っ (120) たら よろしいでしょうか えっ  
と (107) 今 (121) ですね (109)

(ambiguity114-2.2.7.2 ((object "その")

(status accidental)  
 (reference ("講演" "役割について"  
 something))))  
 (ambiguity115-2.2.11 ((object "え")  
 (status accidental)  
 (insertion "え"))))  
 (ambiguity116-2.2.7.3 ((object "そこ")  
 (status accidental)  
 (deictic ("そのかい" "講演"のあるところ  
 somewhere))))  
 (ambiguity117-2.2.8.1 ((object "行く")  
 (status accidental)  
 (ellipsis ("わたしが" someone))))  
 (ambiguity118-3.1.1 ((object "のに")  
 (status accidental)  
 (LEX (対比 目的))))  
 (ambiguity119-3.1.2 ((object "のに")  
 (status accidental)  
 (Francais ("malgré tout" "pour"))  
 (English ("in spite of" "in order to"))))  
 (ambiguity120-2.2.8.1 ((object "行っ")  
 (status accidental)  
 (ellipsis ("わたしが" someone))))  
 (ambiguity121-2.2.7.3 ((object "今")  
 (status accidental)  
 (deictic "at the time of")))

A: はい (106)

TM: えっと (107) 京都駅 (122) にいる (123) んですよ (124)

(ambiguity122-5.2 ((object "京都駅")  
 (status fatal)  
 (situation (外 内 somewhere))))  
 (ambiguity123-2.2.8.1 ((object "いる")  
 (status accidental)  
 (ellipsis ("わたしが" someone))))  
 (ambiguity124-5.1 ((object "んですよ")

(status accidental)  
(IFT (attention-request  
information-offer))))

A: はい (106) えっと (107) ですね (109) タクシーに乗って (125) 頂けるとすぐな (126)  
んですけれども

(ambiguity125-2.2.8.1 ((object "乗って")  
(status accidental)  
(ellipsis ("あなたが" someone))))

(ambiguity126-2.2.8.2 ((object "すぐな")  
(status accidental)  
(information known "そこは"))))

TM: タクシーですか (127)

(ambiguity127-2.2.8.2 ((object "すぐな")  
(status accidental)  
(information known "わたしがのるのは"))))

A: はい (106) えっと (107)

TM: ええ (128)

(ambiguity128-5.1 ((object "ええ")  
(status accidental)  
(IFT (agreement-offer discourse-indication))))

A: 今 (121) 京都駅 (122) です (129) よね

(ambiguity129-2.2.8.2 ((object "京都駅です")  
(status accidental)  
(information known "あなたがいるのは"))))

TM: ええ (128) そうです (130)

(ambiguity130-2.2.7.1 ((object "そう")  
(status accidental)  
(reference ("京都駅"))))

A: あの (131) 烏丸中央口に あの (132) 正面に (132)(133) ですね (109) タクシー  
のりばがごさいます

(ambiguity131-2.2.11 ((object "あの")  
(status accidental)  
(insertion "あの"))))  
(ambiguity132-2.2.11 ((object "正面に")  
(status accidental)  
(revision "烏丸中央口に"))))  
(ambiguity133-2.2.8.3 ((object "正面")  
(status fatal)  
(implied "駅の"))))

TM: はい(106) A: それ(134)に乗って(135)いただきまして あの(136) 京都国際交流センターと言って(137,211)頂けましたら約二十分で到着(138,210)致しますけれども(139)

(ambiguity134-2.2.7.1 ((object "それ")  
(status accidental)  
(reference ("タクシー" something))))  
(ambiguity135-2.2.8.1 ((object "乗って")  
(status accidental)  
(ellipsis ("あなたが" someone))))  
(ambiguity136-2.2.11 ((object "あの")  
(status accidental)  
(insertion "あの"))))  
(ambiguity137-2.2.8.1 ((object "言って")  
(status accidental)  
(ellipsis ("あなたが" someone))))  
(ambiguity211-2.2.7.3 ((object "言って")  
(status accidental)  
(implied ("タクシーの運転手に" someone))))  
(ambiguity138-2.2.8.2 ((object "到着")  
(status accidental)  
(information known ("そこに" somewhere))))  
(ambiguity210-2.2.8.1 ((object "到着")  
(status accidental)  
(ellipsis ("あなたが" "タクシーが" someone))))  
(ambiguity139-5.1 ((object "ますけれども")  
(status accidental)  
(IFT (attention-request information-offer))))))

TM: あ(140) そう(141)ですか(142)

(ambiguity140-5.1 ((object "あ")  
(status accidental)  
(IFT (human-relation attention-offer))))

(ambiguity141-2.2.7.1 ((object "そう")  
(status accidental)  
(reference ("それに乗っていただきまして  
あの京都国際交流センター  
と言って頂けましたら  
約二十分で到着致しますけれども "  
"約二十分で到着致しますけれども " ))))

(ambiguity142-5.1 ((object "ですか")  
(status accidental)  
(IFT (information-request attention-offer))))

A: はい

TM: えっと 烏丸中央口でタクシーに乗る (142) わけですね

(ambiguity143-2.2.8.1 ((object "乗る")  
(status accidental)  
(ellipsis ("わたしが " "あなたが " someone))))

A: そう (144) です正面 (133) です

(ambiguity145-2.2.7.1 ((object "そう")  
(status accidental)  
(reference ("烏丸中央口でタクシーに乗る  
わけ " "烏丸中央口でタクシーに乗る "))))

TM: はい (106) わかりました (146,212)

(ambiguity146-2.2.8.1 ((object "わかり")  
(status accidental)  
(ellipsis ("わたしが " "あなたが " someone))))

(ambiguity212-2.2.8.2 ((object "わかり")  
(status accidental)  
(information known ("烏丸中央口でタクシーに乗る "))))

A: はい (106)

TM: ありがとうございます

A: はい (106) どうも失礼いたします

## A.2 Japanese dialogue text-2b

TM: あ あの(104) もしもし(150)

(ambiguity150-5.2 ((object "もしもし")  
(status accidental)  
(situation (電話会話 会話))))

A: あ(140) もしもし(150)

TM: あ(140)

A: 京都国際センターです(102,103)

TM: あ あの(104) わたくし(105)三浦と申します

A: はい(106)

TM: えっと(107) これから(108)ですね(109) えーと(110) そちら(151)の方に伺いたいんですけども

(ambiguity151-2.2.7.3 ((object "そちら")  
(status accidental)  
(deictic ("京都国際センター" somewhere))))

A: はい(106)

TM: えーと(110) ただいま(152)京都にいます(153) いるんですよ(154)

(ambiguity153-2.2.8.1 ((object "います")  
(status accidental)  
(ellipsis ("わたしが" someone))))

(ambiguity154-2.2.11 ((object "いるんですよ")  
(status accidental)  
(revision ("います"))))

(ambiguity154-5.1 ((object "んですよ")  
(status accidental)  
(IFT (information-offer attention-request))))

A: はい(106)

TM: きよじえ(155) JRの京都駅ですね(156, 157)

(ambiguity155-2.2.11 ((object "きよじえ")  
(status accidental)

(insertion "きよじえ"))  
(ambiguity156-2.2.8.2 ((object "JRの京都駅ですね")  
(status accidental)  
(information known "わたしがいるのは")))  
(ambiguity157-5.1 ((object "ですね")  
(status accidental)  
(IFT (information-offer attention-request))))

A: はい (106)

TM: それで (158) えーと (110) どのような交通機関を使ったら (159) えーと (110)  
早く着く (160) んでしょうか

(ambiguity110-2.2.1 ((object "それで")  
(status accidental)  
(segmentation ("<<それ>で>" "<それで>"))))  
(ambiguity159-2.2.8.1 ((object "使っ")  
(status accidental)  
(ellipsis ("わたしが" "あなたが" someone))))  
(ambiguity160-2.2.8.1 ((object "着く")  
(status accidental)  
(ellipsis ("わたしが" "あなたが" someone))))

A: えーと (110) 時間でしたらタクシーが一番 (161) 早いんですけども (113)

(ambiguity161-2.2.8.2 ((object "一番")  
(status accidental)  
(implied 他の交通機関と比べて)))

TM: ええ (162)

(ambiguity162-5.1 ((object "ええ")  
(status accidental)  
(IFT (agreement-offer attention-offer))))

A: え あとバスと電車を乗り継いで頂くと

(ambiguity163-2.2.8.1 ((object "乗り継い")  
(status accidental)  
(ellipsis ("あなたが" "わたしが" someone))))

TM: はー (164)

(ambiguity164-5.1 ((object " はー ")  
(status accidental)  
(IFT (attention-offer information-request))))

A: 四十五分ほどかかります (165)(166) けれども (113)

(ambiguity165-2.2.8.3 ((object "かかります ")  
(status accidental)  
(implied 時間が)))

(ambiguity166-2.2.8.2 ((object "かかります ")  
(status accidental)  
(information known 京都駅から国際会議センターまで)))

TM: タクシーでしたらどのくらいですか (167)

(ambiguity166-2.2.8.2 ((object "どのくらい ")  
(status accidental)  
(information known 京都駅から国際会議センターまで)))

(ambiguity167-3.1.2 ((object "どのくらい ")  
(status fatal)  
(Francais (combien cela coûte combien de temps il faut))  
(English (How much how long))))

A: 約二十分です (168)(169)

(ambiguity168-2.2.8.2 ((object " 約二十分です ")  
(status accidental)  
(information known 京都駅から国際会議センターまで)))

(ambiguity169-2.2.8.3 ((object " 約二十分です ")  
(status accidental)  
(implied 時間が)))

TM: あ (140) そう (170) ですか

(ambiguity170-2.2.7.1 ((object " そう ")  
(status accidental)  
(reference "約二十分です")))

A: はい (106) TM: ジャタクシーで行き (171) たいんですけれど

(ambiguity171-2.2.8.1 ((object " そう ")  
(status accidental)  
(ellipsis "わたしが")))



A: はい (106)

TM: えーと (110) タクシーのりば (171) 教えて (172) 頂けませんか (173)

(ambiguity171-2.2.8.1 ((object " タクシーのりば ")  
(status accidental)  
(ellipsis "を")))

(ambiguity172-2.2.8.1 ((object " 教えて ")  
(status accidental)  
(ellipsis ("あなたが " "わたしが " someone))))

(ambiguity173-5.1 ((object " 頂けませんか ")  
(status accidental)  
(IFT (suggest information-request))))

A: あ (140) はい (106) 少々お待ちください (174) え 今出てます あの (175) 地図がですね (109)

(ambiguity174-3.1.2 ((object " ください ")  
(status accidental)  
(Francais "s'il vous plaît donnez-moi...")  
(English "please give me")))

(ambiguity175-1.1 ((object "あの ")  
(status accidental)  
(cat (interjection determiner))))

TM: ええ (128)

A: JR 京都駅の内部の地図な (176) んですけれども (113)

(ambiguity176-2.2.8.1 ((object "地図な ")  
(status accidental)  
(ellipsis "今出てます地図が")))

TM: ええ (128) A: 現在 (177) 三浦様はどちら (178) にいらしゃいますでしょうか

(ambiguity177-2.2.7.3 ((object "現在 ")  
(status accidental)  
(deictic at the time of day)))

(ambiguity178-3.1.2 ((object "どちら ")  
(status accidental)  
(Francais (quel côté ou))  
(English (which side where))))

TM: えーと (110) これは (179) えーと (110) ですね (109)

(ambiguity179-2.2.7.1 ((object "これは ")  
(status accidental)  
(deictic (いまいるところ somewhere))))

A: はい (106)

TM: 今 (121) コンコースな (180) んですけど どこなんだろうな (181) ちょっとこ  
れ (194) だけじゃわかんない (182) んですけども (113)

(ambiguity180-2.2.8.2 ((object "今コンコースな ")  
(status accidental)  
(information known わたしがいるのは)))

(ambiguity181-2.2.8.2 ((object "どこなんだろうな ")  
(status accidental)  
(information known (わたしがいるのは ここ))))

(ambiguity182-2.2.8.2 ((object "わかんない ")  
(status accidental)  
(information known どこかが)))

(ambiguity194-2.2.7.1 ((object "これ ")  
(status accidental)  
(reference いまみているもの)))

A: わからないですか (183)

(ambiguity183-2.2.8.2 ((object "わからない ")  
(status accidental)  
(information known どこかが)))

(ambiguity184-2.2.8.1 ((object "わからない ")  
(status accidental)  
(ellipsis ("わたしが " "あなたが " someone))))

TM: ええ えーと (185) どの辺りの出口に出たら (186) タクシーのりば (187)(188)

(ambiguity185-2.1 ((object "ええ えーと ")  
(status accidental)  
(segmentation("<<ええ><えーと>> "  
"<えええ><ーと>>")))))

(ambiguity186-2.2.8.1 ((object "出たら ")  
(status accidental)  
(ellipsis (わたしが あなたが someone))))

(ambiguity187-2.2.2 ((object "タクシーのりば ")  
(status accidental)

(structure (<<出口に出たら>  
<タクシーのりば>> <<タクシー  
のりば><>>))))

(ambiguity188-2.2.8.1 ((object "タクシーのりば")  
(status accidental)  
(ellipsis (が は を に))))

A: あのー (189)

(ambiguity189-5.1 ((object "あのー")  
(status accidental)  
(IFT (discourse-indication attention-request))))

TM: 近い (190) んでしょうか

(ambiguity190-2.2.8.1 ((object "近い")  
(status accidental)  
(ellipsis (タクシーのりば))))

A: えっとですね烏丸町口っていう駅な (191) なんですけれども (113)

(ambiguity191-2.2.8.2 ((object "烏丸町口っていう駅な")  
(status accidental)  
(information known タクシーのりば))))

TM: ええ

A: あの (192) すいません

(ambiguity192-5.2 ((object "あの")  
(status accidental)  
(IFT (attention-request discourse-indication))))

TM: えああなんか

(ambiguity193-2.1 ((object "えああなんか")  
(status accidental)  
(segmentation (<<え><ああ>  
<なんか>> <えああなんか>))))

A: ここ (195) なんです

(ambiguity195-2.2.7.3 ((object "ここ")  
(status accidental)  
(deictic (画面上の一点 somewhere))))

TM: あええはいはあはあはあはあ (196)

(ambiguity196-2.1 ((object "あええはいはあはあはあはあ")  
(status accidental)  
(segmentation (<<あ><ええ><はい>  
<はあはあはあはあ>> <あ  
ええはいはあはあはあはあ>))))))

A: ここ (195) の真正面にあのこと (195) タクシースタンドってありまして (197)

(ambiguity197-2.2.8.3 ((object "タクシースタンドってありまして")  
(status accidental)  
(implied タクシースタンドって書いてありまして)))

TM: ええ A: すぐわかる (198) ようになっているんですけども

(ambiguity198-2.2.8.1 ((object "わかる")  
(status accidental)  
(ellipsis (あなたが 私が someone))))))

TM: はい

A: えっと上の駅の上の (199) (200) あの (201) 案内とか見て (202) 頂きましたら

(ambiguity199-2.2.10 ((object "駅の上の")  
(status accidental)  
(repetition 上の)))  
(ambiguity200-2.1 ((object "上の駅の上の")  
(status accidental)  
(segmentation (<上の駅の> <駅の上の>))))))  
(ambiguity201-2.1 ((object "あの案内")  
(status accidental)  
(segmentation (<あの案内> <<あの><案内>>))))))  
(ambiguity202-2.2.8.1 ((object "見て")  
(status accidental)  
(ellipsis (あなたが someone))))))

TM: ええ

A: 烏丸町口と出て (203) おりますので

(ambiguity203-2.2.8.2 ((object "出て")  
(status accidental)  
(information unknown (案内が 案内に))))))

TM: はい

A: あの出口まではあのわりと楽に出れる (204) と思う (205) んですけど

(ambiguity204-2.2.8.1 ((object "出れる ")  
(status accidental)  
(ellipsis (あなたが わたしが someone))))

(ambiguity205-2.2.8.1 ((object "思う ")  
(status accidental)  
(ellipsis (あなたが わたしが someone))))

TM: ああそう (206) ですか (207)

(ambiguity206-2.2.7.1 ((object "そう ")  
(status accidental)  
(reference 出口まではあのわりと楽に出れる)))

(ambiguity207-5.1 ((object "ですか ")  
(status accidental)  
(IFT (information-request attention-offer))))

A: はい

TM: はいわかり (208) ました

(ambiguity208-2.2.8.1 ((object "わかり ")  
(status accidental)  
(ellipsis (わたしが あなたが someone))))

A: はい

TM: どうもありがとうございました

A: あどうも失礼いたします

### A.3 English dialogue text-5a

A : Good morning Conference Office(1,2) can I help you(3)

(ambiguity1-2.2.8.1 ((object "Conference Office")  
(status accidental)  
(ellipsis "This is"))))

(ambiguity2-5.2 ((object "Conference Office")  
(status accidental)  
(situation (This-is are-you))))

(ambiguity3-5.2 ((object "can I help you")  
(status accidental)  
(situation (telephone conversation-by-sight))))

BL : /breath/ yes I'm trying to find my way(4) to thi International Conference  
Center for thi Symposium on Interpreting Telecommunications(6)

(ambiguity4-3.1.1 ((object "way")  
(status accidental)  
(LEX ("route" "method"))))

(ambiguity5-2.2.2 ((object "to thi International  
Conference Center")  
(status accidental)  
(structure (<<way><to thi International  
Conference Center>>  
<<find><to thi International  
Conference Center>>))))

(ambiguity6-2.2.2 ((object "for thi Symposium on Interpreting  
Telecommunications")  
(status accidental)  
(structure (<<to thi International Conference  
Center >  
<to thi International Conference  
Center>>  
<<find><to thi International Conference  
Center>>))))

A : o?'

BL : and (7)right now (8)I'm at [uh] Kyoto station(9)

(ambiguity7-2.2.1 ((object "and")  
                  (status accidental)  
                  (coordinate (x-and-y x<sub>1</sub>-and-y))))

### Notes

x=previous utterance y=current utterance x<sub>1</sub>=previous utterance of the same speaker

(ambiguity8-2.1 ((object "right now")  
                  (status accidental)  
                  (segmentation ("right now" "right/now"))))  
(ambiguity9-3.1.2 ((object "station")  
                  (status accidental)  
                  (Japonais ("駅" "場")  
                  (Francais ("gare" "station"))))

A : OK(10) you're at Kyoto station how would you like to get to(11) thi International Conference Center we have [uh] three methods of travel we have the subway the bus or (th') a taxi

(ambiguity10-5.1 ((object "OK")  
                  (status accidental)  
                  (IFT (agreement-offer discourse-indication))))  
(ambiguity11-2.1 ((object "get")  
                  (status accidental)  
                  (segmentation ("get" "get to"))))

BL : [um] [hm] which way is the easiest /laugh/

A : the easiest(12) is the taxi

(ambiguity12-2.2.8.2 ((object "the easiest")  
                  (status accidental)  
                  (information known way)))

BL : OK [um] (how would I)(13) how do I (14)get to(11) the taxi from where I am now(15)

(ambiguity13-2.2.11 ((object "how would I")  
(status accidental)  
(insertion "how would I"))))  
(ambiguity14-2.2.11 ((object "how do I")  
(status accidental)  
(revision "how would I"))))  
(ambiguity15-5.2 ((object "now")  
(status accidental)  
(situation at the time of day)))

A : (o?) OK(10) you're at Kyoto station (16) and you'll wanna walk down from the second floor(17) which is where you are at now(18) (19)(I) I believe you're on the second floor walk to the first floor and out the north end of the building [umm] that is where the main ticket office is and it's very easy to find (on the fir') on the first floor(20) located behind the ticket office across the street is the bus station and that(21)'s where you can pick up a taxi as well the taxi stand is located on thi east side of the bus station and there should be a line of taxis waiting all the time and it is the quickest way to get there(22) to (11)(12) thi International Conference Center

(ambiguity16-5.2 ((object " you're at Kyoto station")  
(status fatal)  
(situation (somewhere of Kyoto station))))  
(ambiguity17-5.2 ((object "second floor")  
(status fatal)  
(situation (一階 二階))))  
(ambiguity18-1.1 ((object " now")  
(status accidental)  
(category (conjunction adverb noun))))  
(ambiguity19-2.1 ((object "now")  
(status accidental)  
(segmentation ("now I believe" "where you are  
at now"))))  
(ambiguity20-2.1 ((object "on the first floor")  
(status accidental)  
(segmentation ("find on the first floor" "is  
located on the first floor"))))  
(ambiguity21-2.2.7.1 ((object "that")  
(status accidental)  
(reference ("the bus station"



"the street" "the ticket office"  
"on the first floor")))

(ambiguity22-2.2.7.1 ((object "there")  
(status accidental)  
(deictic ("where you are at" ))))

BL : OK [umm] what do I tell the taxi driver(23) when I get in

(ambiguity23-2.2.8.2 ((object "there")  
(status accidental)  
(information known ("in order to get to  
Conference Center" ))))

A : (the taxi driv') (24)/laugh/ the taxi drivers (25) speak English and you can just  
[uh] tell him that you're trying to get to (11)thi International Conference Center  
and it (26) should(27) be about a twenty minute ride

(ambiguity24-2.2.11 ((object "the taxi driv'")  
(status accidental)  
(insertion ("the taxi driv'"))))

(ambiguity25-2.2.11 ((object "the taxi drivers")  
(status accidental)  
(revision ("the taxi driv'" ))))

(ambiguity26-2.2.7.1 ((object "the taxi drivers")  
(status accidental)  
(reference ("to get to thi International  
Conference Center "  
"you're trying to get to thi  
International Conference  
Center"))))

(ambiguity27-2.2.4 ((object "should")  
(status accidental)  
(modal (duty likelihood opinion))))

BL : OK [umm] [hm] about how much does that(28) cost

(ambiguity28-2.2.7.1 ((object "that")  
(status accidental)  
(reference ("a twenty minute ride"  
expretif))))

BL: A : thi [uh] taxi costs ten thousand yen(29)

(ambiguity29-5.2 ((object "hi [uh] taxi costs ten thousand yen")  
(status accidental)  
(situation (taxi fee taxi-car))))

BL : OK

A : OK

BL : All right(30)

(ambiguity30-5.1 ((object "All right")  
(status accidental)  
(IFT (discourse-indication agreement  
attention-offer))))

A : is there anything else I can help you with today(31)

(ambiguity31-2.2.7.3 ((object "today")  
(status accidental)  
(deictic (date month year))))

BL : [umm] /ls/ no? I should(27) be able to find my way thank you very much

(ambiguity32-2.2.11 ((object "ls/ no? ")  
(status accidental)  
(insertion ("ls/ no? "))))

A : OK you're welcome goodbye

BL : bye

#### A.4 English dialogue text-5b

A : good morning conference office(1,2) can I help you(3)

BL : [um] yes I'm calling about thi [ah] International Symposium on Interpreting Telecommunications(40) right now(8) [um] I'm at Kyoto Station I'm trying to find my way (4) to thi [ah] International Conference +Center(5)+ so

(ambiguity40-2.2.2 ((object "on Interpreting Telecommunications")  
(status accidental)  
(structure (<< thi [ah] International Sympo

sium >  
< on Interpreting  
Telecommunications >>  
<< I'm calling >< on Interp  
reting Telecommunications >>)

A : /laugh/ +/sigh/+ OK you're calling from Kyoto Station

(ambiguity41-3.1.2 ((object "calling")  
(status accidental)  
(Japanese (電話する 呼ぶ))))

BL : yes

(ambiguity42-5.1 ((object "yes")  
(status accidental)  
(IFT (agreement discourse-indication  
information-offer))))

A : OK(43) I'm going to draw up a map [ah] of the station and we can [ah] show  
you where you are in relation to transportation OK lets see here(44) OK you're  
calling from Kyoto Station and you're located down here and you can get to thi  
International Conference Center by either subway bus(45) or a taxi how would you  
like to travel

(ambiguity43-5.1 ((object "OK")  
(status accidental)  
(IFT (agreement attention-offer  
discourse-indication))))

(ambiguity44-2.2.7.3 ((object "here")  
(status accidental)  
(deictic (on the screen speaker's place  
Kyoto station))))

(ambiguity45-2.1 ((object "which")  
(status accidental)  
(segmentation (< subway bus > < subway >< bus >))))

BL : [um] which (46)way is the least expensive /laugh/ or the quickest

(ambiguity46-2.2.7.2 ((object "which")  
(status accidental)  
(reference ("subway" bus" "taxi))))

A : (the least(47)) OK the least expensive [ah] method is not necessarily the cheapest but it(48)s the subway would you like me to [ah] tell ya(49) how to get to the International Conference Center on the subway(50)

(ambiguity47-2.2.11 ((object "the least")  
(status accidental)  
(insertion ("the least"))))

(ambiguity48-2.2.7.1 ((object "it")  
(status accidental)  
(reference ("the least expensive method"  
"the cheapest"))))

(ambiguity49-2.2.11 ((object "ya")  
(status accidental)  
(insertion ("ya"))))

(ambiguity50-2.2.2 ((object "on the subway")  
(status accidental)  
(structure ("<< the International Confe  
rence Center >< on the subway >> "  
"<< get >< on the subway >> "))))

BL : sure

A : OK [ah] are you familiar with the station

BL : [umm] no I'm not(51)

(ambiguity51-2.2.8.2 ((object "I'm not")  
(status accidental)  
(information known ("with the station"))))

A : OK I'm going to [ah] tell you roughly how you're gonna(52) go and then(53) I'll show you(54) the blow up of thi station so(55) you can [ah] catch the subway we(56)'re gonna go north on the subway you're going to take the (n') subway north so (57)let me go to thi station and show you(58) where you're located right now(59)(60) OK you're on the second floor platform(61) right here(62) and the subway only(63) goes one way so(57) you get back on the subway and take it(64) north

(ambiguity52-3.1.2 ((object "gonna")  
(status fatal)  
(Japanese ("たい" "べきだ")  
(Francais ("desirer" devoir))))

(ambiguity53-2.2.7.3 ((object "and then")  
(status accidental)

(deictic (after having done Z))))  
 (ambiguity54-5.2 ((object "show")  
 (status accidental)  
 (situation ("on the screen" "by sight"))))  
 (ambiguity55-2.2.7.3 ((object "so")  
 (status accidental)  
 (reference ("z and z1" "Z"))))

### Notes

Z=previous first sentence in an utterance, Z1= previous second sentence  
 in an utterance

(ambiguity56-2.2.7.1 ((object "we")  
 (status accidental)  
 (deictic ("speaker" "speaker and hearer"))))  
 (ambiguity57-2.2.7.3 ((object "so")  
 (status accidental)  
 (reference ("z and z1" "Z"))))  
 (ambiguity58-5.2 ((object "let me go to thi station and show you")  
 (status accidental)  
 (situation ("on the screen" "by sight"))))  
 (ambiguity59-2.2.7.3 ((object "right now")  
 (status accidental)  
 (deictic ("at the time of day"  
 "some-time"))))  
 (ambiguity60-2.1 ((object "right now")  
 (status accidental)  
 (segmentation ("right now" "right?/now"))))  
 (ambiguity61-5.2 ((object "second floor")  
 (status fatal)  
 (situation ("the second floor" "the  
 third floor"))))  
 (ambiguity62-2.2.7.3 ((object "here")  
 (status fatal)  
 (deictic ("at the point of the map"

```

                                somewhere))))
(ambiguity63-2.2.2 ((object "only")
                    (status fatal)
                    (structure ("only goes one way" "only
                                goes")))))
(ambiguity64-2.2.7.1 ((object "it")
                      (status accidental)
                      (reference ("the subway" "one way "))))

```

BL : OK

A : and now(65 ) we'll show(54) you where you're goinu go(66) OK you're going to take thi Kyoto Subway north(67) to Sanjo Station and(68) this (69) is Sanjo Station when you arrive at Sanjo Station you'll want to(70) get off the subway and(71) change subways to thi Keihen-Kyotsu Line so take the Kyoto subway north to Sanjo Station get off and get on the Keihan- Kyotsu Line

```

(ambiguity65-2.2.11 ((object "and no")
                    (status accidental)
                    (insertion ("and now"))))
(ambiguity66-2.2.11 ((object "you're goinu go")
                    (status accidental)
                    (insertion "you're goinu go"))))
(ambiguity67-2.2.1 ((object "north")
                    (status fatal)
                    (apposition ("to Sanjo Station")
                    (subordinate ("<< take thi Kyoto Subway >
                                < north >> " << take thi
                                Kyoto Subway >
                                < to Sanjo Station >> "))))))
(ambiguity68-2.2.11 ((object "and")
                    (status accidental)
                    (insertion ("z and Y" "Z1,Z and Y"))))
(ambiguity69-2.2.7.1 ((object "this")
                      (status accidental)
                      (deictic ("Sanjo Station" on the screen))))
(ambiguity70-3.1.2 ((object "want to")
                    (status accidental)
                    (Japanese ("たい" "べきだ")
                    (Francais ("desirer" "devoir"))))

```

(ambiguity71-2.2.1 ((object "and")  
                  (status accidental)  
                  (coordinate ("z,z1 and y" "Z and Y"))))

BL : /breath/ OK

A : And that(72) will take you down Sanjo dori to (kiage)(73) Keage station(74)  
that(75)'s the second stop you'll wanna get off the subway at Keage Station

(ambiguity72-2.2.7.1 ((object "that")  
                  (status accidental)  
                  (reference ("Keihan- Kyotsu Line"  
                              "the Kyoto subway"))))

(ambiguity73-2.2.11 ((object "(kiage)"  
                  (status accidental)  
                  (insertion "kiage"))))

(ambiguity74-2.2.2 ((object "to Keage station")  
                  (status accidental)  
                  (subordinate "<<Sanjo dori>  
                              <to Keage station>>" "  
                              <<take you down><to Keage station>>"))

(ambiguity75-2.2.7.1 ((object "that")  
                  (status accidental)  
                  (reference ("Keage Station"  
                              "Sanjo dori""that"))))

BL : all right

A : (ae') and you'll walk east on Sanjo Dori to the first street(76) at the first  
street(77) take a left and walk up to thi International Conference Center it(78)'s  
located on the left and it(79)'s a large gray building

(ambiguity76-2.2.2 ((object "to the first street")  
                  (status accidental)  
                  (subordinate ("<<Sanjo Dori>  
                              <to the first street>>" "<<walk>  
                              <to the first street>>"))))

(ambiguity77-2.2.2 ((object "at the first street")  
                  (status accidental)  
                  (subordinate ("<<the first street>  
                              <at the first street>>" "<< take a left>  
                              <at the first street>>"  
                              "<< take a left and walk up>

<at the first street>>"))))  
(ambiguity78-2.2.7.1 ((object "it")  
(status accidental)  
(reference International Conference)))

(ambiguity79-2.2.7.1 ((object "it")  
(status accidental)  
(reference ("International Conference"  
"it"))))

BL : OK

A : ('z)(80) do you [ah] wanna go over that(81) again or do you have that(82) down

(ambiguity80-2.2.11 ((object "z")  
(status accidental)  
(insertion "z")))

(ambiguity81-2.2.7.1 ((object "that")  
(status accidental)  
(reference previous utterances)))

(ambiguity82-2.2.7.1 ((object "that")  
(status accidental)  
(reference previous utterances)))

BL : [um] I(83)think I understand that [um] from where I am now at Kyoto station  
(I)(84) how do I catch the subway (can I just) (85)is that very close(87)

(ambiguity83-5.2 ((object "I")  
(status accidental)  
(situation previous speaker)))

(ambiguity84-2.2.11 ((object "I")  
(status accidental)  
(insertion "I")))

(ambiguity85-2.2.11 ((object "can I just")  
(status accidental)  
(insertion "can I just")))

(ambiguity86-2.2.7.1 ((object "that")  
(status accidental)  
(reference "the subway")))

(ambiguity87-2.2.8.3 ((object "close")  
(status accidental)



(implied "to this place"))

A : OK (at Kyoto stesh')(88) you're at Kyoto station now(88) so you'll(89) wanna walk up to the second floor let me go back(90) there(91) and we(92)'ll take a look at the station where you're at you're on the second floor right now(59)(60) and you just walk through these(93) steps up onto the platform and that(94) is the Kyoto subway line going north you're located about right here now and you'll just walk onto the platform and (that is thi)(95) [uh] that(96)'s where you catch the Kyoto train north Kyoto (su') subway

(ambiguity88-2.2.11 ((object "at Kyoto stesh")  
(status accidental)  
(insertion "at Kyoto stesh"))))  
(ambiguity88-2.2.7.3 ((object "now")  
(status accidental)  
(dectic "at the time"))))  
(ambiguity89-2.2.4 ((object "will")  
(status accidental)  
(modal politeness)  
(tense future)))  
(ambiguity90-5.2 ((object "go back")  
(status accidental)  
(situation move on the screen)))  
(ambiguity91-2.2.7.3 ((object "there")  
(status accidental)  
(dectic ("the second floor"  
"Kyoto station"))))  
(ambiguity92-5.2 ((object "we")  
(status accidental)  
(situation ("speaker"  
"speaker and hearer"))))  
(ambiguity93-2.2.7.2 ((object "these")  
(status accidental)  
(reference ("pointed out on the screen"  
))))  
(ambiguity94-2.2.7.1 ((object "that")  
(status accidental)  
(reference ("the platform"))))  
(ambiguity95-2.2.11 ((object "that is thi")  
(status accidental)

(insertion "that is thi"))  
(ambiguity96-2.2.7.1 ((object "that")  
(status accidental)  
(reference "onto the platform")))

BL : OK

A : OK

BL : I can understand that(97)

(ambiguity97-2.2.7.1 ((object "that")  
(status accidental)  
(reference previous utterance)))

A : (any) can I help you with any other things

BL : [um] I think that(98)'s it(99)

(ambiguity98-2.2.7.1 ((object "that")  
(status accidental)  
(reference vaguely)))

(ambiguity99-2.2.7.1 ((object "it")  
(status accidental)  
(reference vaguely)))

A : OK

BL : thank you very much

A : Thank you have a good day bye

BL : y' too(100) bye

(ambiguity100-2.2.11 ((object "y' too")  
(status accidental)  
(insertion "y' too")))

## A.5 French written text-introduction of a scientific article

Les avis sur la traduction automatique (TA) sont souvent extrêmes(800).

(ambiguity800-3.1.2 ((object "extrêmes")  
(status fatal)  
(Japonais ("最後の" "極端な"))))

Les uns(801) ne la(802) conçoivent que comme l'expérimentation scientifique de leurs théories ou de leurs formalismes favoris, tandis que les autres(803) y(804) voient une entreprise purement(805) technologique et utilitaire.

(ambiguity801-2.2.3 ((object "les uns")  
(status accidental)  
(generic certain)  
(specific 1/2)))

### Notes

For generic, there is "certain" as the value, for specific there are "only one", "1/2", "possessive", "modifier" as the values. "only one" is used for the object of noun phrases which exists as only one in the world. 1/2 is used for the object of noun phrases of which total quantity is limited. "possessive" and "modifier" are used for the object of nouns phrases is modified by the special modifier.

(ambiguity802-2.2.7.1 ((object "la")  
(status accidental)  
(reference "traduction automatique"))))

(ambiguity803-5.2 ((object "les autres")  
(status accidental)  
(situation "la moitié des gens qui ont leur avis sur la  
traduction automatique"))))

(ambiguity804-2.2.7.3 ((object "y")  
(status accidental)  
(reference "la traduction automatique"))))

```
(ambiguity805-2.2.2 ((object "purement")
                    (status accidental)
                    (structure (< purement technologique et
                                purement utilitaire > << purement
                                technologique > et < utilitaire >>)))
```

Le plus(806) souvent, chacun défend un "paradigme" particulier, comme(807) on défendrait une idéologie.

```
(ambiguity806-3.1.2 ((object "le plus")
                    (status accidental)
                    (Japonais (もっとも 非常に)))
(ambiguity807-4 ((object "comme")
                (status accidental)
                (logicosem (reason likelihood)))
```

Par exemple, on soutient que le progrès ne pourrait venir qu'en construisant des systèmes munis d'ontologie, qui "comprennent" explicitement, alors que c'(808)est rarement possible, et, quand ça(809) l'(810)est, rarement nécessaire.

```
(ambiguity808-2.2.7.1 ((object "ce")
                      (status accidental)
                      (reference (le progres en construisant des systèmes
                                  munis dontologie)))
(ambiguity809-2.2.7.1 ((object " ce")
                      (status accidental)
                      (reference (le progres en construisant des systeme
                                  munis d'ontologie)))
(ambiguity810-2.2.7.1 ((object "l'")
                      (status accidental)
                      (reference ("le progres" "en construisant
                                  des systèmes munis d'ontologie" "c'est
                                  rarement possible" "est rarement
                                  possible possible")))
```

Ou encore(811), on(829) affirme qu'il(812) est nécessaire de passer par un interlingua pour construire des systèmes multilingues, alors que cette(813) idée déjà ancienne est particulièrement difficile à mettre en oeuvre, et que d'autres(814) approches sont

également(815) possibles et nettement moins(816) coûteuses, sauf dans les rares situations très fortement multilingues (au moins 8 langues, avec flux équilibrés pour toutes les paires).

- (ambiguity811-3.1.2 ((object "encore")  
(status accidental)  
(Anglais ("still/yet" "more" "furthermore")  
(Japonais (もう まだ さらに))))
- (ambiguity829-5.2 ((object "on")  
(status accidental)  
(situation ('auteur et lecteur' 'les gens'))))
- (ambiguity812-2.2.7.1 ((object "il")  
(status accidental)  
(reference (de+inf something))))
- (ambiguity813-2.2.7.2 ((object "cette")  
(status accidental)  
(reference ("construire des systèmes multilingues"  
"de passer par un interlingua"))))
- (ambiguity814-2.2.8.3 ((object "d'autres approches")  
(status accidental)  
(implied ("construire des systèmes multilingues"  
"de passer par un interlingua"))))
- (ambiguity815-2.2.8.3 ((object "egalement")  
(status accidental)  
(implied "a ce qu'on passe passer par interlingua"  
)))
- (ambiguity816-2.2.8.3 ((object "moins")  
(status accidental)  
(implied "qu'on passe passer par un interlingua"))))

Notre(830) thèse est que la TAO n'est ni une science, ni une collection de recettes, mais plutôt une technologie scientifique, c'est-à-dire un ensemble de méthodes qui progresse autant par l'intégration épisodique d'idées théoriques que par l'amélioration incrémentale des savoir-faire.

- (ambiguity830-5.2 ((object "notre")  
(status accidental)  
(situation auteur)))
- (ambiguity817-2.2.7.3 ((object "c'est-à-dire")  
(status accidental)

(reference ("n'est ni une science,  
ni une collection de recettes,  
mais plutot une  
technologie scientifique" "plutot  
une technologie scientifique"))

D'autre part(818), les fameux "paradigmes" ne peuvent être placés sur une seule échelle de valeurs, mais sont à(819) comparer dans chaque type de situation traductionnelle.

(ambiguity818-3.1.2 ((object "D'autre part")  
(status accidental)  
(Japonais (他方で さらに))  
(Anglais ("the other hand" "morever"))))

(ambiguity819-4 ((object "sont a")  
(status accidental)  
(logicosem ("une relation de consequence"  
"un etat en cours" "une obligation"))))

Enfin, en ce qui concerne l'architecture linguistique des systèmes de TAO, il (820) reste de nombreuses voies intermédiaires à explorer.

(ambiguity820-4 ((object "il")  
(status accidental)  
(logicosem ("pronon personel" "expletif"))))

Cette thèse peut être illustrée par le nouveau paradigme de la TA fondée sur le dialogue (TAFL), ou TAO personnelle pour auteurs monolingues, que (821) nous proposons pour le cas de situations traductionnelles où d'autres approches, telles que la TA fondée sur la langue (TAFL), la TA fondée sur la connaissance (TAFC), et les aides informatisées au traducteur (THAM), sont inadéquates.

(ambiguity821-2.2.7.1 ((object "que")  
(status accidental)  
(reference ("la "TA fondée sur le dialogue"(TAFL),  
ou TAO personnelle pour auteur  
monolingue" "TAO personnelle pour  
auteur monolingue"))))

Dans cette approche, bien que la base de connaissances linguistiques du système reste cruciale, et doit même être la plus couvrante possible, et que connaissances extralinguistiques puissent éventuellement(822) être utilisées si elles(823) sont disponibles,

l'accent est mis sur une préédition indirecte, effectuée grâce à un dialogue de normalisation et de clarification avec l'auteur, visant à produire des traductions de haute qualité sans révision.

```
(ambiguity821-2.2.7.2 ((object "cette")
  (status accidental)
  (reference ("la "TA fondée sur le dialogue"(TAFL),
    devant ou TAO personnelle pour auteur
    monolingue" "TAO personnelle pour
    auteur monolingue" "la TA fondée sur
    la langue(TAFL), la TA fondée sur
    la connaissance(TAFC), et les
    aides informatisées
    au traducteur(THAM)"))))
(ambiguity822-3.1.2 ((object "éventuellement")
  (status accidental)
  (Japonais ("万一" "ひょっとして"))))
(ambiguity823-2.2.7.1 ((object "elles")
  (status accidental)
  (reference "des connaissances extralinguistiques")))
```

Nous(831) étudions diverses questions relatives à cette approche grâce à une maquette, LDIA-1(824), première étape vers la construction d'un système de TAFD pour une situation particulière(825), celle d'auteurs monolingues rédigeant en hypertexte de la documentation technique être diffusée en plusieurs langues(826).

```
(ambiguity831-5.2 ((object "nous")
  (status accidental)
  (situation auteur)))
(ambiguity824-2.2.11 ((object "LDIA-1")
  (status accidental)
  (insertion "LDIA-1")
  (apposition "une maquette")))
```

#### Notes

This formalization is used to represent that there is an ambiguity due to whether it's "insertion" or "apposition".

```

(ambiguity825-2.2.11 ((object "premiere étape....
                        une situation particulière")
                      (status accidental)
                      (insertion "premiere étape....
                                  une situation particulière")
                      (apposition ("LDIA-1(824)" "une maquette"))))
(ambiguity826-2.2.11 ((object "celle d'auteurs monolingues ...en
                              plusieurs langues")
                      (status accidental)
                      (insertion "celle d'auteurs monolingues ...en
                                  plusieurs langues")
                      (apposition ("LDIA-1(824)" "une maquette"
                                  "premiere étape....
                                  une situation particulière"))))

```

Nous terminons en évoquant quelques problèmes intéressants et nouveaux apparus durant ce(827) travail, et auxquels(828) il faudrait au moins trouver des solutions partielles pour passer ultérieurement à un prototype.

```

(ambiguity827-2.2.7.2 ((object "ce")
                      (status accidental)
                      (deictic "ce qui se passe dans cet article")
                      (reference something)))
(ambiguity828-2.2.7.2 ((object "ce")
                      (status accidental)
                      (reference "quelques problemes
                                  interessants" "nouveaux apparus"
                                  "ce travail"))))

```



## A.6 English written text-introduction of a scientific article

Opinions on Machine Translation (MT) are often extreme(900).

(ambiguity900-3.1.2 ((object "extreme")  
(status fatal)  
(Japanese ("最後の " "極端な")))))

Some(901) consider it(902) only as a testbed for the scientific experimentation of their favorite theories or formalisms(903), while others(904) see(907) it(905) as a purely technological and utilitarian enterprise(906)(908).

(ambiguity901-2.2.3 ((object "some")  
(status accidental)  
(generic certain)  
(specific 1/2)))

(ambiguity902-2.2.7.1 ((object "some")  
(status accidental)  
(reference ("machine translation" "Opinions"  
"Opinions on Machine Translation  
(MT) are often extreme."))))

(ambiguity903-2.2.2 ((object "favorite theories or formalism")  
(status accidental)  
(structure (< favorite theories or favorite  
formalism > << favorite theories >  
< or formalism >>))))

(ambiguity904-2.2.3 ((object "others")  
(status accidental)  
(generic certain)  
(specific 1/2)))

(ambiguity905-2.2.7.1 ((object "it")  
(status accidental)  
(reference ("it" "a testbed" "the  
scientific experimentation of  
their favorite theories or  
formalisms" "their favorite theories  
or formalisms"))))

(ambiguity906-2.2.2. ((object "a purely technological and  
utilitarian enterprise")  
(status accidental)

(structure (<< a purely technological >< and  
utilitarian > enterprise >>  
<< a purely technological and  
purely utilitarian ><  
enterprise >>)))

(ambiguity907-3.1.2 ((object "see")  
(status accidental)  
(Japanese (見る わかる))))  
(ambiguity908-3.1.2 ((object "enterprise")  
(status accidental)  
(Japanese (企業 積極性))))

In most cases(909), each defends a particular "paradigm", as one would(914) defend an ideology.

(ambiguity909-3.1.2 ((object "cases")  
(status accidental)  
(Japanese (問題 事実 立場 容体 もの・ひと  
訴訟 箱 覆い 枠 一組 格)  
(Francais (cas cause affaire etui coffre couverture))))  
(ambiguity910-3.1.1 ((object "cases")  
(status accidental)  
(LEX ("box" "example" "grammatical case"))))  
(ambiguity914-2.2.4 ((object "would")  
(status accidental)  
(tense (past present)  
(modal (politeness uncertanety))))

For example, it is asserted that progress could only(911) come by(916) building(915) systems equipped with an ontology, which "understand" explicitly, while this(912) is rarely possible, and, if it(913) is possible, rarely necessary.

(ambiguity911-2.2.2 ((object "only")  
(status accidental)  
(structure (< only come > < only come  
by building systems equipped with  
an ontology >))))  
(ambiguity912-2.2.7.1 ((object "this")

```

        (status accidental)
        (reference ("building systems equipped with
                    an ontology"
                    "progress" "understand explicitly"))))
(ambiguity913-2.2.7.1 ((object "it")
                      (status accidental)
                      (reference ("building systems equipped with
                                  an ontology"
                                  "progress" "understand explicitly"
                                  "this"))))
(ambiguity915-1.2 ((object "building")
                  (status accidental)
                  (function (gerund participle-present noun))))
(ambiguity916-4 ((object "building")
                 (status accidental)
                 (logicosem (place way))))

```

Or it is asserted that it is necessary to go through an interlingua in order to(917) build multilingual systems, while this(918) already old idea is particularly difficult to implement, and other approaches are equally(919) possible and considerably less(920) expensive, except in the rare(921) situations which are strongly multilingual (at least 8 languages, with balanced throughputs for all language pairs).

```

(ambiguity917-2.1 ((object "in order to")
                  (status accidental)
                  (segmentation (< in order to > << in order >
                                < to build multilingual systems))))
(ambiguity918-2.2.7.2 ((object "this")
                      (status accidental)
                      (reference ("to go through an interlingua"
                                  "to build multilingual systems"))))
(ambiguity919-2.2.8.3 ((object "equally")
                      (status accidental)
                      (implied "to going through an interlingua")))
(ambiguity920-2.2.8.3 ((object "less")
                      (status accidental)
                      (implied "than to going through an interlingua")))
(ambiguity921-3.1.1 ((object "rare")

```

(status accidental)  
(LEX ("lightly cooked" "uncommon"))))

Our(922) thesis is that MT is neither a science nor a collection of recipes, but rather a scientific technology, that is, a set of methods which progress as much through episodic integration of theoretical ideas as through incremental improvements in practical know-how.

(ambiguity922-5.2 ((object "our")  
(status accidental)  
(situation author))))

On the other hand(923), the famous "paradigms" cannot be ranked on a unique scale of values, but should be compared with reference to each type of translational situation.

(ambiguity923-2.2.7.3 ((object "On the other hand")  
(status fatal)  
(reference (previous-paragraph  
all-previously-mentioned))))))

Finally(924), as far as the linguistic architecture of MT systems is concerned, there remain many intermediate avenues to explore.

(ambiguity924-4 ((object "finally")  
(status fatal)  
(logicosem ("at last" "as the last of a number of things"))))

This thesis can be illustrated with the new paradigm of "dialogue-Based MT" or personal MT for monolingual authors, which we propose for the case of situations where other(925) approaches, such as Language-Based MT and Knowledge-Based MT, are inadequate.

(ambiguity925-2.2.7.2 ((object "other")  
(status accidental)  
(reference the new paradigm  
of "dialogue-Based"  
personal MT for monolingual authors))))

Although the linguistic knowledge base of the system remains crucial, and should even be as covering as possible, and although extralinguistic knowledge may possibly be used if it (926) is available, emphasis is on an indirect prediction, realized through a normalization and clarification dialogue with the author, making it (927) possible to produce high quality translations without postediting.

```
(ambiguity926-2.2.7.1 ((object "it")
                      (status accidental)
                      (reference "extralinguistic knowledge")))
```

```
(ambiguity927-2.2.7.1 ((object "it")
                      (status accidental)
                      (reference ("to produce high quality translations
                                without postediting" "emphasis"
                                "a normalization and clarification
                                dialogue with the author"))))
```

In the last part of the paper, we (928) present the specification and the current state of our (929) mockup, LIDIA-1 (930), first step toward building a DBMT system for a particular situation (931), that of monolingual authors writing hypertext technical documentation to be distributed in several languages (932). We also mention some interesting and new problems which have appeared during this (933) work, problems for which we should find at least partial solutions before we could (934) go further and build an operational prototype.

```
(ambiguity928-5.2 ((object "we")
                  (status accidental)
                  (situation author)))
```

```
(ambiguity929-5.2 ((object "our")
                  (status accidental)
                  (situation author)))
```

```
(ambiguity930-2.2.11 ((object "LDIA-1")
                     (status accidental)
                     (insertion "LDIA-1")
                     (apposition "our mockup")))
```

```

(ambiguity931-2.2.11 ((object " first step .....situation")
  (status accidental)
  (insertion " first step .....situation")
  (apposition ("LDIA-1" "our mockup"))))
(ambiguity932-2.2.11 ((object "celle d'auteurs monolingues ...en
  plusieurs langues")
  (status accidental)
  (insertion "that of monolingual...in several languages")
  (apposition ("LDIA-1" "our mockup"
    "first step .....situation"))))

(ambiguity933-2.2.7.2 ((object "this")
  (status accidental)
  (reference ("the specification and the
    current state of our mockup"
    "our mockup")))

(ambiguity934-2.2.4 ((object "could")
  (status accidental)
  (tense (past present)
  (modal (politeness uncertainty))))

```

## A.7 French transparency

Motivation générales(1000)

- En TAFD, accessibilité des Bases de Connaissances Linguistiques(1001)
  - rendre(1003) les connaissances linguistiques "découvrables"
  - apports de l'EAO(1004)
  - .....également(1005)
- Socle applicatif pour une modélisation(1006)
  - interfaces coopérants(1007)
  - utilisateurs-apprenants(1008)
  - .....dans les environnements d'apprentissage par découverte(1009)
- Utilisation effective, en situations d'apprentissage(1010)

(ambiguity1000-5.2 ((object "motivation générales")  
(status accidental)  
(situation (transparency article dialogue))))

(ambiguity1001-2.2.8.2 ((object "accessibilité des Bases de  
Linguistiques Connaissance")  
(status accidental)  
(information known (''il y a le problème de''  
''il y a l'''))))

(ambiguity1003-2.2.8.2 ((object "rendre")  
(status accidental)  
(information known (''TADF'' something))))

(ambiguity1004-2.2.8.2 ((object ''apports de l'EAO''  
(status accidental)  
(information known (''c'est des''  
''c'est les'' something))))

(ambiguity1005-2.2.8.2 ((object "également")  
(status accidental)  
(information known something)))

(ambiguity1006-2.2.8.2 ((object "Socle applicatif pour une modélisation")  
(status accidental)  
(information known something)))

(ambiguity1007-2.2.8.2 ((object "interfaces coopérants")  
(status accidental)  
(information Socle applicatif pour  
une modélisation)))

(ambiguity1008-2.2.8.2 ((object "utilisateurs-apprenants")  
(status accidental)  
(information Socle applicatif pour  
une modélisation)))

(ambiguity1009-2.2.8.2 ((object "dans les environnements  
d'apprentissage par découvert'')  
(status accidental)  
(information (''on a le socle applicatif pour  
une modélisation'' ''Il y a  
le socle applicatif pour une  
modélisation'')))

(ambiguity1010-2.2.8.2 ((object "Utilisation effective,  
en situations d'apprentissage")  
(status accidental)  
(information ''Il y a ...'' something)))



## References

- [1] Christian Boitet Human-Oriented Design and Human-Machine-Human Interactions in Machine Interpretation ATR Technical Report TR-IT-0013 1993
- [2] Eric Bilange A task independent oral Dialogue Model ACL proceedings Fifth European Conference Berlin 1991
- [3] G.N.Leech Principles of Pragmatics Longman London 1983
- [4] H.Blanchon désambiguïation interactive en TAO personnelle Avignon-92 Avignon juin 1992
- [5] H.Blanchon A Solution to the Problem of Interactive Disambiguation COLING-92 Nantes juillet 1992, ACL
- [6] H.Blanchon LIDIA-1 une première maquette vers la T A interactive "pour tous". Grenoble, 21/1/94, Univ. Grenoble 1 (UJF)
- [7] 柏崎雅世 日本語における行為指示型表現の機能 くろしお出版 Tokyo 1993
- [8] 国立国語研究所 日本語教育映画 基礎編 総合文型表 日本シニセル株式会社 1987
- [9] K.H.Loken-Kim F.Yato L.Fais K.Kurihara T.Furukawa Y.Kitagawa マルチモーダルシュミレータ EMMI を用いた道案内データベースのテキスト ATR-IT-0029 ATR Technical Report 1993
- [10] M.Lafourcade LEAF, ou comment garder l'origine de l'ambiguïté T-TA-TAO'93, Montréal, 28/9/93, Presses de l'Université de Mont
- [11] Mutsuko Tomokiyo Transfert de la langue parlée japonais-Anglais dans le système de traduction automatique ASURA T-TA-TAO Montréal, 28/9/93, Presses de l'Université de Mont
- [12] Mutsuko Tomokiyo and Tsuyoshi Morimoto Communicative Functions of Spoken Japanese and Its Meaning Interpretation on MT System TR-I-0260 ATR technical report ATR Kyoto 1992
- [13] Paul Heisterkamp AMBIGUITY AND UNCERTAINTY IN SPOKEN DIALOGUE Proceedings of ACL, 1993
- [14] Paul Larreya Enoncé performatifs Présupposition Université Nathhan information formation Fernand Nathan France 1979

- [15] Philippe Morin and Jean-claude Junqua HABITABLE INTERACTION IN GOAL-ORIENTED MULTIMODAL DIALOGUE SYSTEM Proceedings of ACL, 1993
- [16] Polly Szatrowski 日本語の談話の構造分析 くろしお出版 Tokyo 1993
- [17] 佐久間まゆみ 文段認定の一基準 筑波大学 Tukuba 1987
- [18] Stephen C. Levinson Pragmatics Cambridge University Press Cambridge 1983
- [19] 友清睦子 語用論的分析に基づく自然発話の長文分割 TR-IT-0019 ATR Technical Report Sept.1993
- [20] Willis Edmondson Spoken Discourse Longman London and New York 1981