

TR-326

Generating Natural Language Responses
Appropriate to Conversational Situations
—On the Case of Japanese—

by

N. Noguchi, M. Takahashi
& H. Yasukawa

November, 1987

©1987. ICOT

ICOT

Mita Kokusai Bldg. 21F
4-28 Mita 1-Chome
Minato-ku Tokyo 108 Japan

(03) 456-3191~5
Telex ICOT J32964

Institute for New Generation Computer Technology

Generating natural language responses appropriate to conversational situations – On the case of Japanese –

Naohiko Noguchi Masanori Takahashi Hideki Yasukawa

Tokyo Systems Research Department
Matsushita Electric Industrial Co., Ltd.

17-15, Shinbashi 6-chome,
Minato-ku, Tokyo 105, Japan

Abstract

The goal of this work is to make human-machine interaction in natural language smooth and efficient. To realize this, the machine has to generate natural language sentences which are understood by users most easily. And it is important for this understandability that each sentence is appropriate to the conversational situation at the time of its utterance. In this paper, we will shed light on this appropriateness and consider how we can formalize the method to generate such appropriate sentences. First, we will define such appropriateness, second we will discuss some surface features of Japanese sentences and consider how people decide to add such features to his sentences depending on situational factors. Then we will show brief methods of deciding to add suitable surface features to sentences in suitable situation. Finally, we will show an implementation of a conversational system which adopts these methods.

1 Introduction

On a conversational system which has some purposes and communicates with people(users) cooperatively in some natural language to achieve that purposes, it is important that the system makes the conversation with users as smooth as possible. And this enables the system to run effectively. To realize this aspect, such system has to:

- (1) interpret user's input sentences correctly, and grasp his true intention;

In order that the system can deal with conversations similar to which people make, the system has to grasp their true intention. Both the formalization and recognition of the speaker's intention are one of the major problems in the fields of linguistics and AI. The problem about formalizing the speaker's intention has been dealt with by the speech act theories [Searl 1969] and the problem about recognizing it has been dealt with by the plan-goal approaches [Allen 1983] based on the speech act theories.

- (2) decide the system's own response in accordance with its purposes and user's true intention;

After interpreting user's input sentence, the system has to decide its own response. One of the criteria for deciding it is to choose one response which will realize its purposes as fast as possible. And one of the necessity conditions for deciding it is to have some accordance with user's true intention which have been interpreted just before.

- (3) show the response in a natural language expression which users can understand most easily.

Finally, the system has to show its response in some natural language form. And it is a desirable aspect that users can easily understand the system's expression from the standpoint of continuation of the conversation. If this aspect is not realized, users may misunderstand the system's sentence, the system and users must have further conversation in order to remove their misunderstanding, which decreases efficiency of whole conversation.

The system can do these three things only when it can refer to conversational situations at each point of the conversation. That is, to recognize user's true intention from their sentence, it is necessary to take account of the things which have been understood from the previous conversation. And the system has to decide its own response based on the facts which had been recognized in the previous conversation. And to show its response in a natural language expression users can easily understand, it is necessary to consider what had been communicated and what kind of expression had been used in the previous conversation.

In this paper, we will argue (3) of above points. In other words, we will consider, when the system has recognized user's true intention correctly ((1) has been done) and has decided its response((2) has been done), how the system should express its response based on conversational situations at that time. To say more concretely, we are going to discuss

- (a) what kind of factors in conversational situations should be parameterized for deciding the expression of its response and
- (b) how are they used for deciding the expression.

Section 2 of this paper sheds light on the appropriateness of the expression which people generate in a conversation. The appropriateness is defined from the viewpoint how easily people can understand that expression. In section 3, we consider some surface features of Japanese sentences and how can people decide to add these features to sentences appropriately. This appropriateness is evaluated in the meaning we defined at section 2 and the methods are formalized in the form we stated above — some situational factors are used for deciding some surface features of Japanese sentences. Section 4 is related to implementation of these methods. Finally, section 5 summarizes this paper.

2 What is appropriateness?

What kind of sentences are easily understood by people? Can we define concrete conditions for such sentences? Grice [Grice 1975] stated a principle and some maxims for cooperative conversations. These were described as rules which must be obeyed by all participants in a cooperative conversation. But these can be seen as the conditions for sentences in a conversation to be understood easily. In this paper, we define some conditions below which are a little bit concrete version of Grice's and which refer to levels of surface expressions.

The surface expressions of sentences in a conversation should :

- (1) be consistent with what has been said before (corresponds to the Maxim of quality),
Do not use expressions which are inconsistent with the things which has been remarked and understood each other so far.
- (2) offer as much information as to describe what one really wants to say (corresponds to the Maxim of quantity),
Do not use expressions which offer more or less information than is required.
- (3) have some relations to what has been said before (corresponds to the Maxim of relation), and
Do use expressions which clearly shows some relations to the conversation so far.
- (4) be clear (corresponds to the Maxim of manner).
Do not use ambiguous expressions and do use as clearer expressions as possible.

We don't argue whether Grice's principle and maxims are valid or not. Through this paper to the end, we think of those sentences which suit these conditions above as "appropriate" sentences concerning to conversational situations.

3 Considerations to some linguistic phenomena on Japanese

In this section, we will think about some surface features of Japanese sentences and consider some relations between these features and conversational situations through the analysis of the related linguistic phenomena. We will also consider in what situation, what kind of surface features sentences should have in order to satisfy the "appropriateness condition" stated in section 2. We deal with three categories of surface features of Japanese sentences here. They are

- styles, topicalizations and word-order variations,
- connectives, and
- some referential expressions and ellipsis.

3.1 Styles, topicalizations and word-order variations

When we want to say something, there are many surface expressions which describe what we want to say correctly. For example, when we want to ask someone where he manufactures some certain products, we can choose any sentences written below to express what we want to ask.

- (1) *seihin-wo doko-de seizoushimasu-ka?*
*products-ACC where-LOC manufacture-INT*¹
Where do you manufacture the products?
- (2) *seihin-wo seizousuru basho-wa doko-desu-ka?*
products-ACC manufacture place-TOP where-COP-INT
Where is it that you manufacture the products?
- (3) *seihin-wa doko-de seizoushimasu-ka?*
products-TOP where-LOC manufacture-INT
As for the products, where do you manufacture these?
- (4) *doko-de seihin-wo seizoushimasu-ka?*
where-LOC products-ACC manufacture-INT
Where do you manufacture the products?

The sentence (1) and (2) differ in their styles. For convenience, we call the style the sentence (1) has "doushi-bun" (which corresponds to normal Japanese sentences in which the main verb is located at the end of the sentence), and the style the sentence (2) has "meishi-bun" (which corresponds to the cleft sentences of English). In the sentence (3), accusative case "products" is topicalized and located at the beginning of the sentence, and in the sentence (1), it is also located at the beginning of the sentence but is not topicalized. And between the sentence (1) and (4), their word-orders are different. We will think of these differences and how these differences are related to the "appropriateness conditions" stated in section 2.

Let's see the examples demonstrated below one by one in order to clarify the points.

¹ Through this paper, some abbreviations to express Japanese postpositions are used. They are,

NOM:nominative,

ACC:accusative,

GEN:genitive,

PART:postposition which represents part-of relation,

LOC:locative,

TOP:postposition which marks topicalization,

INT:postposition which expresses interrogative mood,

COP:copula,

CON:coordinating conjunction, and

COMP:complimentizer.

- (5) keisanki-no memori-wo amerika-de seizoushimasu.
 (wa) (de-wa)
computer-PART memory-ACC America-LOC manufacture
 (TOP) (LOC-TOP)
I will manufacture memories of computers in America.

- (6) keisanki-no memori-wo seizousuru basho-wa amerika-desu.
computer-PART memory-ACC manufacture place-TOP America-COP
The place where I will manufacture memories of computers is America.
(It is America where I will manufacture memories of computers.)

- (7) keisanki-no memori-wo amerika-de seizoushimasu-ka?
 (wa) (de-wa)
computer-PART memory-ACC America-LOC manufacture-INT
 (TOP) (LOC-TOP)
Will you manufacture memories of computers in America?

- (8) keisanki-no memori-wo seizousuru basho-wa amerika-desuka?
computer-PART memory-ACC manufacture place-TOP America-COP-INT
Is the place where you will manufacture memories of computers America?
(Is it America where you will manufacture memories of computers?)

- (9) keisanki-no memori-wo doko-de seizoushimasu-ka?
 (wa)
computer-PART memory-ACC where-LOC manufacture-INT
 (TOP)
Where will you manufacture memories of computers?

- (10) keisanki-no memori-wo seizousuru basho-wa doko-desu-ka?
computer-PART memory-ACC manufacture place-TOP where-COP-INT
Where is the place in which you will manufacture memories of computers?
(Where is it in which you will manufacture memories of computers?)

Sentences (5),(7),(9) are "doushi-bun"s and sentences (6),(8),(10) are "meishi-bun"s. First, we consider the difference between "doushi-bun" and "meishi-bun" on declarative sentences. (5) and (6) seem to mean exactly the same thing, but their true meaning are slightly different because of the difference of styles. (5) insists that the specific event in which the speaker manufactures the products is true. While (6) insists that the place where the speaker manufactures the products and America are the same. That kind of difference is clarified if you imagine the situation in which one wants to manufacture the products in America and England. In such situation, it is possible to use such expressions as below.

keisanki-no memori-wo amerika-de seizoushimasu.
computer-PART memory-ACC America-LOC manufacture
I will manufacture memories of computers in America.

keisanki-no memori-wo amerika-to igirisu-de seizoushimasu.
computer-PART memory-ACC America CON England-LOC manufacture
I will manufacture memories of computers in America and England.

keisanki-no memori-wo seizousuru basho-wa amerika-to igirisu-desu.
computer-PART memory-ACC manufacture place-TOP America-CON England-COP
The places where I will manufacture memories of computers are America and England.
(These are America and England where I will manufacture memories of computers.)

But (6) is not suitable for this situation. Because, "The place where I will manufacture memories of computers" and "America" are not exactly the same. And in (6), "keisanki-no memori-wo seizousuru

basho(the place where I will manufacture memories of computers)" is topicalized, so we can recognize that the place is old information, i.e. someone has already referred to such place in the previous conversation. To say more clearly, only in the situation in which someone has already referred to the place where someone would manufacture memories of computers, he can use the sentence (6) as an appropriate expression. On the other hand, (5) does not have such limitation.

Next, let's think about (7) and (8). On such yes/no-questions, there is an essential ambiguity — what does the speaker wants to ask? For example, there are several possible interpretations of (7):

- Will you manufacture memories of computers or will you manufacture other things?
- Will you manufacture memories of computers in America or will you manufacture them in other places?
- Will you manufacture memories of computers in America or Won't you manufacture them in America?

Ordinary, one can change the order of postpositional phrases(PPs) or topicalize some PP to make it clearly that what he really wants to ask. Then, what is the relationship between this kind of ambiguities and these linguistic methods(topicalizations, word-order variations)? Since (7) has two PPs, we can make those sentences below from (7) by changing the order of PPs or topicalizing these PPs.

- (11) *keisanki-no memori-wa amerika-de seizoushimasu-ka?*
computer-PART memory-TOP America-LOC manufacture-INT
 As for memories of computers, will you manufacture them in America?

This means that "will you manufacture memories of computers in America or will you manufacture them in other place?".

- (12) *keisanki-no memori-wo amerika-de-wa seizoushimasu-ka?*
computer-PART memory-ACC America-LOC-TOP manufacture-INT
 As for America, will you manufacture memories of computers there?

This sentence is appropriate only in the situation in which the speaker and the hearer know that the hearer will manufacture memories of computers at least in one country except for America, or that he won't manufacture them at least in one country except for America.(latter is more ordinary)

It means that "although you won't manufacture in that country, will you manufacture in America?".(this "wa" should be recognized as "contrastive-"wa")

- (13) *amerika-de-wa keisanki-no memori-wo seizoushimasu-ka?*
America-LOC-TOP computer-PART memory-ACC manufacture-INT
 As for America, will you manufacture memories of computers there?

This means "will you manufacture memories of computers in America or will you manufacture the other products in America?"

It is presupposed that the hearer will manufacture something in America.

- (14) *amerika-de keisanki-no memori-wa seizoushimasu-ka?*
America-LOC computer-PART memory-TOP manufacture-INT
 As for memories of computers, will you manufacture them in America?

The meaning of this sentence is the same as (11).

Or, this sentence is appropriate only in the situation in which the speaker and the hearer know that the hearer will manufacture at least one product except for memories of computers in America, or that he won't manufacture at least one product except for memories of computers in America.(latter is more ordinary) In this case, this sentence means that "although you won't manufacture the other products, will you manufacture memories of computers?".

(15) *keisanki-no memori-wa amerika-de-wa seizoushimasu-ka?*

computer-PART memory-TOP America-LOC-TOP manufacture-INT

As for memories of computers and as for America, will you manufacture them there?

As for memories of computers, will you manufacture them in America in contrast with any other countries?

This sentence is appropriate only in the situation in which the speaker and the hearer know that the hearer will manufacture memories of computers at least in one country except for America, or that he won't manufacture them at least in one country except for America. (former is more ordinary)

It means that "although you will manufacture in other countries, will you manufacture in America?".

It is presupposed that the hearer will manufacture memories of computers somewhere.

(16) *amerika-de-wa keisanki-no memori-wa seizoushimasu-ka?*

America-LOC-TOP computer-PART memory-TOP manufacture-INT

As for America and as for memories of computers, will you manufacture them there?

As for America, will you manufacture memories of computers in contrast with any other products there?

This sentence is appropriate only in the situation in which the speaker and the hearer know that the hearer will manufacture at least one product except for memories of computers in America, or that he won't manufacture at least one product except for memories of computers in America. (former is more ordinary) It means that "although you will manufacture the other product, will you manufacture memories of computers?".

It is presupposed that the hearer will manufacture something in America.

According to above examples, we can understand that the PP which is topicalized and moved to the beginning of the sentence takes a role in the presupposed event. And what the speaker wants to ask is clear in the sentence some PP of which is topicalized, because its presupposed facts become more obvious. On the other hand, the PP located in front of main verb is the sign of what the speaker wants to ask, but if the PP that is located in front of main verb is topicalized one, the situation may be changed. Although we can recognize it as such sign by interpreting "wa" which marks that PP as "contrastive-*wa*". (Especially in (12), we can only interpret this way) To sum up, when one utters yes/no-question in "doushi-bun" style, he can hold the sentence appropriate by topicalizing some PP or changing the order of PPs in relation to the presupposed facts at the time of the utterance. On the other hand, in (8), the speaker verifies the equality between "the place where you will manufacture memories of computers" and "America", so he presupposes that "there is a place where you will manufacture memories of computers" and also that "you will manufacture memories of computers". Therefore we can't interpret this sentence as the speaker asking someone whether he will manufacture or not, and there is no ambiguities at all that what the speaker wants to ask. That is to say, in (7), the speaker verifies whether a specific event is true or not, or whether a specific role is valid to let a specific event be true, but in (8), he verifies the equality between two objects. We can understand this difference well by considering the situation in which the hearer will manufacture memories of computers in America and England. In such situation, each pair of question-response described below is appropriate.

Q: *keisanki-no memori-wo amerika-de seizoushimasu-ka?*

computer-PART memory-ACC America-LOC manufacture-INT

Will you manufacture memories of computers in America?

A: *hai, amerika-to igirisu-de seizoushimasu.*

yes, America-CON England-LOC manufacture

Yes, I will manufacture them in America and England.

Q: *keisanki-no memori-wo seizousuru-basho-wa amerika-desu-ka?*

computer-PART memory-ACC manufacture-place-TOP America-COP-INT

Is the place where you will manufacture memories of computers America?

(Is it America where you will manufacture memories of computers?)

A: *ie, amerika-to igirisu-desu.*
no, America-CON England-COP
No, these are America and England.

Finally, let's think about differences between (9) and (10) which are both *wh*-questions. There is also a difference between (9) and (10) which we have already considered in the case of *yes/no*-questions. That is to say, in (9), the speaker wants to ask where it is which makes the event that the hearer will manufacture memories of computers true, but in (10), he wants to ask where it is which equals to the place where the hearer will manufacture memories of computers. This difference will be more clear when you look some examples below.

Q1: *keisanki-no memori-wo doko-de seizoushimasu-ka?*
computer-PART memory-ACC where-LOC manufacture-INT
Where will you manufacture memories of computers?

A1: *amerika-de seizoushimasu.*
America-LOC manufacture
I will manufacture in America.

Q2: *hoka-no basho-de-mo sore-wo seizoushimasu-ka?*
other place-LOC-also these-ACC manufacture-INT
Will you manufacture them in any other country?

A2: *hai, igirisu-de-mo seizoushimasu.*
yes, England-LOC-also manufacture
Yes, I will manufacture in England, too.

Q: *keisanki-no memori-wo seizousuru-basho-wa doko-desu-ka?*
computer-PART memory-ACC manufacture-place-TOP where-COP-INT
Where is the place in which you will manufacture memories of computers?

A: *amerika-to igirisu-desu.*
America-CON England-COP
These are America and England.

Let's consider the sentence (17) together.

(17) *keisanki-no memori-wo dokoka-de seizoushimasu-ka?*
computer-PART memory-ACC somewhere-LOC manufacture-INT
Will you manufacture memories of computers somewhere and where is it?

We can interpret (17) as either *yes/no*-question or *wh*-question. If we interpret it as *yes/no*-question, what the speaker wants to ask is whether the event one will manufacture memories of computers is true or not. And if we interpret it as *wh*-question, its meaning is the same as that of (9). Because of this ambiguity, if the speaker uses the sentence like (17), he can ask whether the event described in the sentence becomes true or not and if it's true, he can also ask where does it become true, when does it become true, what role is required for the event to become true, and so on. To say more, in (17), of course it is not presupposed that the hearer will manufacture memories of computers.

So far, we have considered the differences between "doushi-bun" style sentences and "meishi-bun" style sentences. And we have also considered the relationship between such differences and surface features of the sentences - word-order variations and topicalizations. We conclude that when the speaker decide the expressions, his decision depends on those situational factors which are described below:

- 1-a presupposed facts relating to the event which the speaker wants to describe, and
- 1-b the point which the speaker wants to ask or insist on in the event which he wants to describe.

If one does not decide them depending on these factors correctly, he may make sentences which are inappropriate to the conversational situation. For example, he may say "keisanki-no memori-wo seizousuru

basho-wa doko-desu-ka?(Where is the place in which you will manufacture memories of computers?)" in the situation where the hearer has no intention of manufacturing memories of computers.

3.2 Connectives

Connectives, which are located at the beginning of the sentences, are used for expressing the relationship between the previous conversation and the present utterance. For example, connectives like "tokorode(by the way)", "sate (well)" are used for changing formerly discussed topics to the new one, connectives like "soredewa(if so)", "sorede(then)" are used for maintaining formerly discussed topics or entering into details of these topics. For making the relationship between the previous conversation and the present utterance clear, it is necessary to use suitable connectives in the utterance in suitable situations. Then, what kind of situational information motivates the making of "connective-added" sentences? There are two aspect of the relationship which the present utterance has in the conversation. They are:

- A. relationship between the speaker's(one's own) utterance just before and the present utterance and
- B. relationship between the hearer's(the other's) utterance just before and the present utterance.

The relationship which belongs to the category A comes to the surface when one participant(the speaker) has continued to take the leadership of the conversation. In ordinary question-answering conversation, the questioner has continued to take the leadership and repeats the pair of "question-and-answer" utterances. In such situation, the questioner chooses a question from his prepared questions one by one depending only upon his purpose. So, the relationship between the question he made just before and the present question is projected to the relationship between the utterance he made just before and the present utterance. On the other hand, the relationship belongs to the category B comes to the surface when the conversation go in the direction which the speaker doesn't expect. When the other's answer is not sufficient for his question, or when it is incorrect, or when the other says that he wants to take the leadership of the conversation, the speaker should make some untypical conversation. For the first case, the speaker should say that "please describe in details.", for the second case, he should say that "you are wrong, because... and please answer it correctly.", and for the last case, he should reply to the other's request. In such situations, the speaker should understand the other's true intention and the relationship between the previous conversation and the present utterance depends on how the speaker decides his intention and content of response according to the other's true intention. Let's think about an example below:

Q1: keisanki-no memori-nikansuru tokkyo-wo tsukatteiru-no-desu-ka?
 computer-PART memory-concerning patent-ACC use-COP-INT
 Have you used the patent concerning memories of computers?

A1: hai.
 yes
 Yes.

Q2: sorede, sono-tokkyo-wa zenbude nanken-desu-ka?
 then the-patent-TOP all how many-COP-INT
 Then, how many are the patents in all?

A2: ikken-desu.
 one-COP
 one.

Q3: soredewa, sono-tokkyobangou-wa irete-kudasai.
 then the-number of the patent-TOP input-please
 Then, please input the number of the patent.

A3: nihontokkyo2312203.
 Japan patent 2312203
 Japan patent 2312203.

Q4: tokorode, keisanki-no memori-wo seizoushiteiru-basho-wa doko-desu-ka?
by the way computer-PART memory-ACC manufacture-place-TOP where-COP-INT
By the way, where is it in which you have manufactured memories of computers?

A4: nihon-de seizoushi, nihon-to amerika-de hanbaishiteiru.
Japan-LOC manufacture Japan-CON America-LOC have sold
we have manufactured in Japan and sold in Japan and America.

Q5: sousuruto, amerika-ni yushutsushiteiru-no-desu-ne?
so America-LOC have exported-COMP-COP-INT
So, is it true that you have exported it into America?

A5: hai.
yes
Yes.

In above example, from Q1 to Q4, the relationship in the category A came to the surface of the conversation and connectives like "sorede", "soredewa", "tokorode" were appeared. That is, the questioner repeated questions in accordance with his purpose, and he maintained topics("sorede"), or he entered into the details("soredewa"), or he changed topics("tokorode"). But at Q5, because the answerer had said more than had been required, the questioner confirmed the result he inferred from the answerer's remarks(if one manufacture some products in Japan and sells them in Japan and America, probably it is true that he exports it from Japan to America). In such situation, the category B's relationship came to and connective "sousuruto" was added to the sentence.

To sum up, when the speaker decides whether he add connectives to his utterance or not and decides what kind of connectives he add, he depends on two situational factors:

2-a the structures of his prepared topics, and

2-b how does he decide his intention and contents of his utterance in accordance with the other's true intention just before.

3.3 Some referential expressions and ellipsis

When people make a conversation, they use many referential expressions each of which is appropriate to conversational situations at the time of their utterance and they use ellipsis which is also appropriate at that time. In this section, we will discuss how we decide referential expressions to some referents, how we decide to use ellipsis. And we will also discuss what kind of situational factors are used for such decision. First, let's think about ellipsis. In real conversations which people make, the participants hardly offer more information than is required. This is quite so when one is asked and he replies. For example, when you are asked,

"Where are you going to see the movie?"

it's more ordinary to reply

"Shibuya."

or

"Shinjuku."

than to reply

"I'm going to Shinjuku."

or

"I'm going to Shibuya to see a movie."

But if you know that the questioner doesn't know what movie you are going to see and you think that he wants to know the movie, you ordinary reply,

"I'm going to Shinjuku to see "Rocky-4"."

which has more information than is asked. That is to say, when one replies to some question, it's ordinary to say only what he is asked. And he offers more information if the information has a enough reason to be uttered in the situation at that time. This is so not only in this case(question-and-answer case) but also in any other case. (It corresponds to Grice's "maxim of quantity") Therefore, deciding the use of ellipsis to some referent is reduced to examine whether it has enough reason to be uttered. And if there is no reason to be uttered, one can realize an appropriate utterance by using ellipsis to that referent. There are many reasons for some referents to be uttered such as:

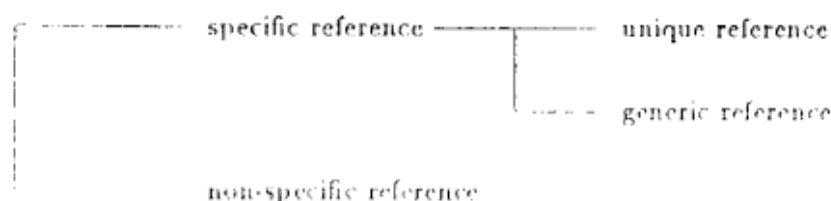
- referent which is a central topic in the conversation,
- referent which has never been uttered in the previous conversation,
- referent which is central role in the situation corresponds to the speaker's utterance.

If a referent is not omitted from one's utterance, it can be referred in several types of expressions. These types are:

- (1) pronoun,
- (2) specifier-added expression("sono-tokkyo"("the patent" or "that patent"), "kono-hito"("the man" or "this man"), etc.),
- (3) proper noun,
- (4) common noun,
- (5) expression with relative clause.

Probably, the speaker chooses one of these types of expressions for each referent, which is the most suitable to the situation at that time. Then, what is the relationship between that choice and conversational situations?

At first, we will consider type (1) and (2) expressions. Generally, there are several types of the way to refer by noun phrases.



The sentences below are examples of these types of referring.

- (18) sono-tokkyo-wa boku-ga kakimashita.
the-patent-TOP I-NOM wrote
 The patent, I wrote.

- (19) tokkyo-wa aru-ittai-no kikan-wo sugiru-to mukou-ni-naru.
patent-TOP some-fixed duration-ACC pass-COMP non effective become
 Patents will be non-effective after some period.

- (20) wareware-wa ichinen-ni ikken tokkyo-wo kakanakereba-naranai.
we-TOP one year-per one patent-ACC write-must
 We must write one patent per year.

In (18), "sono-tokkyo(the patent)" refers to some specific patent which the speaker wrote, in (19), "tokkyo(patents)" refers to generic class of patents. And in (20), "tokkyo(patent)" refers to one non-specific patent. That difference comes from the difference of the way how the speaker grasp these referents. And the specifier-added expression is only used for unique reference.

Moreover, all referents should be identified by the hearer, there are several types of context which the hearer can use for that identification process. They are:



Referential expression which refers to a referent in linguistical context is ordinarily called "anaphora", and which refers to a referent in non-linguistical direct context is ordinarily called "deixis". Specifier-added expression can not be used for any referents in non-linguistical indirect context. And when the hearer identifies the referent and he chooses a candidate from these contexts, he will choose one which has been introduced into the context most recently. To say reversely, the speaker uses referential expressions considering the hearer's processes of identification of the referent. It concludes that when the speaker uses referential expression, he decides the expression depending on how recently its referent has been introduced into the context.

Next, let's consider type (3),(4) and (5) expressions. The way to decide which type of these expressions the speaker uses mainly depends on how he grasps the referent and whether there is an indication act or not. It is generally accepted that all proper nouns indicate referents specifically which are in non-linguistical indirect context. So when the speaker uses proper noun to refer to the referent, there is an obvious indication act by him and he grasps the referent as an unique and specific object. On the other hand, when he uses common noun to refer to the referent, it may not be so. It is not decided from the referential expression that how the speaker grasps the referent and whether there is an indication act or not. It depends on the conversational situation. And noun phrases with relative clauses guarantees that there is some identification act by the speaker if the relative clauses are used restrictively. But it is not decided from only referential expression that whether these are used restrictively or not. It needs some information contained in the conversational situations. To say reversely, from the standpoint of the speaker, if he wants to indicate some object by a referential expression(i.e. if there is an indication act), he wants the hearer to identify the object correctly. Therefore, to communicate such intention to the hearer and to let the hearer recognize that, the speaker should use such referential expressions as express the existence of the indication act clearly. If it is realized, the speaker can make more clearer sentences in the sense of Grice's maxim of manner. As a result, we have a general criterion, which defines the way to choose referential expressions, that when one wants to indicate some object, he should choose such referential expression as expresses the existence of the indication act more clearly.

We briefly summarize the points discussed here. In deciding which type of referential expression the speaker should use, he has to consider some situational factors described below:

- 3-a whether there is an indication act by the speaker or not,
- 3-b how does the speaker grasp the referent,
- 3-c what type of context does the referent are in, and
- 3-d how recently has the referent been introduced into the context.

4 Implementation and some example

We are now making researches into a conversational system, which is named "ToR". Its entire view is a kind of computer-aided documentation system which supports people to make a contract. This system

makes a conversation with human(hereinafter called USER), extracts USER's intention and request, interprets them as contracting conditions and finally makes a form of contract. Now we have developed a prototype of the system(hereinafter called ToR) and what we have argued in the previous sections are applied to the interface of ToR which makes a conversation with USER in Japanese. To say more concretely, when ToR offers its response on the CRT-display to USER, ToR considers the conversational situation at that point of time to make sentences which USER can understand most easily. In this section, we will state how ToR makes sentences relating to what we have discussed in section 3. In 4.1 we will describe how to represent the information about conversational situations. In 4.2, we will describe the framework and the methods to decide three surface features, which are discussed in section 3, based on that information. And we will briefly look over the implementation of ToR in 4.3. Finally, in 4.4, we will demonstrate some example and explain how ToR runs according to the example.

4.1 Representation of the information about conversational situations

In ToR, the contents of USER's utterances, the concepts about contracts and the contents of ToR's utterances are all represented in LAST(Language based on Situations and Types), which is a kind of knowledge representation language based on Situation theory [Barwise 1987] and it has major functions requisite for making a conversational system. [Yasukawa 1987] On ToR, the facts which are recognized through the conversation with USER are registered into LAST database. For example, when USER says,

uchi-ga ICOT-no tokkyo-wo tsukatteiru.

we-NOM ICOT-GEN patent-ACC have used

we have used the ICOT's patents.

the instance of a situation type, where USER has used patents which are owned by ICOT in order to manufacture and sell something, is registered into the database as a fact. Figure 1 shows such instance of a situation type.

$X = \text{use}(\text{agt:S1,obj:O,purpose:P}),$

$Y = \text{export}(\text{agt:S2,obj:O}),$

$S1 = \text{USER},$

$S2 = \text{ICOT},$

$O \text{ typeof } \text{patent},$

$P \text{ typeof } \text{product}.$

registered objects = $[X,Y,S1,S2,O,P]$

Figure 1: LAST expressions registered into the database

And the contents of ToR's response to USER are also represented in a set of intermediate expressions of LAST(i.e. expressions not registered in the database). For instance, the content of an utterance which is described below

ICOT-no tokkyo-wo tsukatte, keisanki-no memori-wo seizoushiteiru.

ICOT-GEN patent-ACC use computer-PART memory-ACC have manufactured

we have manufactured memories of computers using ICOT's patents.

is represented in a set of intermediate expressions of LAST described in Figure 2.

So, the information about the conversational situation at the point of each utterance is basically contained in LAST database. And referring to the conversational situation is reduced to observing the

```

X = manufacture(agt:S1,obj:O1),
Y = use(agt:S1,obj:O2,purpose:O1),
Z = possess(agt:S2,obj:O2),
W = partof(obj1:O1,obj2:O3),
S1 = USER,
S2 = ICOT,
O1 typeof memory,
O2 typeof patent,
O3 typeof computer.

utterance content = [X,Y,Z,W,S1,S2,O1,O2,O3]

```

Figure 2: LAST expressions corresponding to the utterance content

relationship between the intermediate LAST expressions which represents the utterance content and LAST database. To observe such relationship, LAST provides some built-in predicates such as:

```

hold(X):
    examines whether the partially instantiated type X holds in the database or not.
hold_same_rel(Rel,Y):
    examines whether the specific relation Rel holds in the instantiated situation type Y in the
    database or not.

```

Using these predicates, we can examine the condition 1-a discussed in 3.1. And there are several sorts among the objects of LAST. They are:

- instance(unique object)
- generic(generic object)
- set

The sorts of LAST objects represents the difference between the way for ToR to grasp objects. Therefore this represents the information 3-b discussed in 3.3. we can get the sort of an object by the built-in predicate "sortof".

```

sortof(X,Sort):
    returns sort Sort of the object X.

```

But LAST database contains only recognized facts flatly, so there is no distinction among the objects whether one object was referred in the real utterance or was deduced by inference. And we can't understand how recently each object has been introduced into the conversation. To represent these aspect, we construct some structured data such as:

- (1) the list of LAST objects which are referred in each utterance:
 - it stores LAST objects which are referred explicitly by either USER or ToR
- (2) the list of LAST objects which are introduced into the conversation:

it stores USSR objects which are introduced into the conversation relating to each utterances, and also it categorizes these objects according to the time they has been introduced;

- 1 objects which were introduced by USER's utterance just before,
- 2 objects which were introduced by ToR's utterance just before, and
- 3 objects which had been introduced by some utterances before.

We can consider the objects in (1) as the objects in linguistical context. And also we can consider the objects categorized in (2) as introduced into the attentional state of the participants by some utterances, so there also exists objects which have not been referred explicitly. Categories 1,2,3 represent the time sequence of these objects.(This represents 3-c,3-d discussed in 3.3)

We also construct some structured data relating to the continuity of the conversation such as:

- (3) the retroactive records of topics which have been discussed in the conversation:

ToR has a set of topics to ask which is about making contract. At the time of each utterance, ToR registers the central topic of the utterance.

- (4) the retroactive records of intentions which each participant have at the time of each utterance:

ToR registers intentions of USER and ToR at the time of each utterance. Categories of intentions are defined provisionally as described below. Categories of USER's intentions are:

- 1 answering to ToR's question,
- 2 requesting some explanation,
- 3 requesting to know some information about the contract which has been partially decided by the conversation so far, and
- 4 requesting to know how USER answers to ToR's question.

Categories of ToR's intentions are:

- 1 informing USER of some information,
- 2 explaining,
- 3 asking, and
- 4 confirming.

These data represent 2-1,2-b discussed in 3.2.

4.2 How to decide surface features of ToR's responses

ToR generates Japanese sentences corresponding to the utterance contents, which are represented in a set of LAST expressions, examining the relationship between the utterance contents and the information about the conversational situations described in 4.1. Figure 3 is the framework of this process.

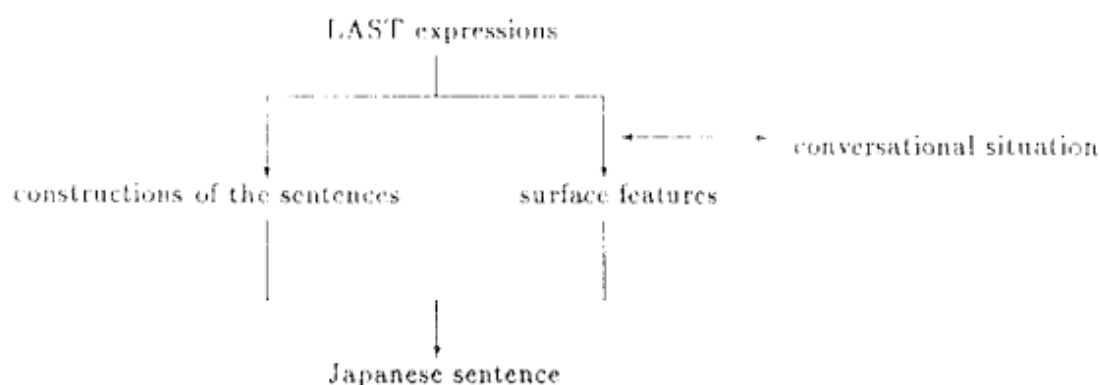


Figure 3: The framework of the process

Now, we will describe the methods which decide surface features of ToR's sentences from a set of LAST expressions.

4.2.1 Method to decide styles, topicalized element, word-order

According to the discussion in 3.1, these surface features such as styles, topicalizations and word-orders are decided based on conversational situations. The method to do this is described below:

1. In the case where there is some presupposed fact of the utterance content (this can be examined by the predicate "hold"),
 - 1-1. In the case where there is a point ToR wants to ask or insist on especially,
 - 1-1-1. choose "meishi-bun" style and put the expression of the object, which corresponds to what ToR wants to ask or insist on, on the end of the sentence,
 - 1-1-2. if we can't choose "meishi-bun" style because of some reason,
 - choose "doushi-bun" style,
 - topicalize the object which corresponds to the presupposed fact and
 - put the expression of the object, which corresponds to what ToR wants to ask or insist on, on the place just before the sentence predicate.
 - 1-2. In the case where there is no point ToR wants to ask or insist on especially,
 - choose "doushi-bun" style and
 - topicalize the object which corresponds to the presupposed fact.
2. In the case where there is no presupposed fact of the utterance content,
 - 2-1. In the case where there is a point ToR wants to ask or insist on especially,
 - choose "doushi-bun" style,
 - put the expression of the object, which corresponds to what ToR wants to ask or insist on, on the place just before the sentence predicate and
 - if the utterance content corresponds to the wh-question, choose the expression such as (17) — use "some-wh" in the sentence.
 - 2-2. In the case where there is no point ToR wants to ask or insist on especially, choose "doushi-bun" style.

4.2.2 Method to add suitable connectives

According to the discussion in 3.2, we can summarize how to decide to add suitable connectives to the sentences.

1. In the case where ToR has been taking the leadership of the conversation, ToR should refer to the retroactive records of topics and should add an appropriate connective in accordance with the relationship of the current topic and the topic just before. There are several kind of such relationships described below:
 - changing the topic,
 - maintaining the topic and
 - entering into the details of the topic just before.
2. In the case where USER answers what ToR didn't expect to be answered, ToR should refer to the retroactive records of utterance intentions and should add an appropriate connective in accordance with the relationship between the current ToR's intention and intentions so far.

4.2.3 Method to use ellipsis, to choose referential expressions

We can describe how to use ellipsis and how to choose referential expressions below.

1. If there is some reason for a object not to omit from the sentence in the conversational situation at that time, choose referential expression for it according to the method described in 2(below). Otherwise, use ellipsis for the object. The following example represent such reasons:
 - the object which ToR wants to ask or insist on especially.
 - the object which is a kind of new information that is not introduced in the conversation so far.
 - the object which ToR wants to confirm.
2. For the objects which are decided not to use ellipsis,
 - 2-1. an object, which is introduced by both ToR's utterance just before and USER's utterance just before, and which is expressed explicitly in at least one of these utterances, should be referred by a pronoun.
 - 2-2. an object, which has a proper name, should be referred by the proper name.
 - 2-3. an object, which is introduced by either ToR's utterance just before or USER's utterance just before and, which is expressed explicitly in at least one of these utterances, should be referred by a specifier-added expression.
 - 2-4. an object, which was expressed explicitly in the former utterances, should be referred by the same expression.
 - 2-5. an object other than described above should be referred by a common noun which corresponds to the class of the object.

4.3 Implementation

Now, ToR has been implemented on Quintus-PROLOG, which is running on Micro VAX 2. And ToR can make a simple conversation in Japanese with USER and then write out a draft of contract in English. In this section, we will briefly explain the implementation, especially of the process which decides surface features of ToR's response, which has been discussed in this paper so far. That process receives a set of LAST expressions which corresponds to ToR's sentence as an input and yields a rough construction of ToR's sentence and surface features which applies to the construction as an output. It consists of a group of production rules, a rule applier which selects suitable rules and applies these rules, and a information manager which holds and manages information about conversational situation. (It is depicted in Figure 4)



Figure 4: Process configuration

The format of the production rules is described below.

```
RULENAME::PRIORITY,
PATTERNS OF LAST EXPRESSION,
PATTERNS OF WORKING MEMORIES,
CONDITIONS OF CONVERSATIONAL SITUATION
⇒ SURFACE FEATURES.
```

Figure 5 shows an example of that rule.

```
style_concept_completion::15,  
[_,X,Sit],  
working([[[utterance_type,[[question_type,concept_completion]]]]],  
discourse_condition([hold(Sit,Sit1)])  
⇒ [additional_information,[[style,meishibun]]].
```

Figure 5: An example of the rule

The meaning of each items of the rule is:

- (1) RULE_NAME: identifier of the rule,
- (2) PRIORITY: priority of the rule,
- (3) PATTERNS OF LAST EXPRESSION:
patterns of a set of LAST expressions corresponding to ToR's utterance,
- (4) PATTERNS OF WORKING MEMORIES:
patterns of a set of surface features which are decided by application of other rules,
- (5) CONDITIONS OF CONVERSATIONAL SITUATION:
conditions on the conversational situation at each point of time to be able to apply the rule,
- (6) SURFACE FEATURES:
surface features of ToR's sentence which is concluded by the application of this rule.

For example, the meaning of the rule depicted in Figure 5 is,

"If the set of LAST expressions, which corresponds to ToR's utterance content, represents that the utterance is wh-question, and if the event, which corresponds to ToR's utterance content, is a presupposed fact, then choose "meishi-bun" style."

The rule applier selects the applicable rules, and apply them one by one to decide surface features of ToR's sentence. If more than two rules are applicable at the same time, it choose one rule whose priority is the highest among these rules. When there is no applicable rule, surface features which have decided so far are the result.

4.4 How does ToR work?

Figure 6 shows an example of a conversation between ToR and USER. *Sns* are ToR's utterances, and *Uns* are USER's utterances.

S1: *konkai-wa donna goyou-desu-ka?*
What can I do for you?

U1: *ICOT-no tokkyo-wo uchi-ga tsukatteiru.*
ICOT-GEN patent-ACC we-NOM have used
I have used ICOT's patents.

S2: *sono-tokkyo-wa nanken-desu-ka?*
that-patent-TOP how many-COP-INT
How many are the patents?

U2: *nihontokkyo2312203-to nihontokkyo2312201-no 2ken desu.*
Japanese patent 2312203-CON Japanese patent 2312201 two COP
They are two, Japanese patent 2312203 and Japanese patent 2312201.

S3: sore-wa donoyouna seihin-ni tsukatteimasu-ka?
Them-TOP what kind product-ACC have used-INT
 What kind of products have you used them for ?

U3: keisanki-no memori-desu.
computer-PART memory-COP
 Memories of computers.

S4: tokorode, sono-memori-wa dokoka-de seizoushiteimasu-ka?
by the way the-memory-TOP somewhere-LOC have manufactured-INT
 By the way, Have you manufactured the memories somewhere?

U4: nihon-de seizoushite, nihon-to amerika-de hanbaishiteiru.
Japan-LOC manufacture Japan-CON America-LOC have sold
 We have manufactured (them) in Japan and sold (them) in Japan and America.

S5: amerika-ni yushutsushiteimasu-ne?
America-LOC have exported-INT
 Is it true that you have exported (them) to America?

U5: hai.
yes
 Yes.

Figure 6: An example of a conversation between ToR and USER

Now we will explain how ToR decides surface features of its response in above example.

In S2, ToR asked the number of patents which USER had used. Because the patents which ToR referred in S2 were the same as the patents which USER referred in U1, ToR chose specifier-added referential expression “sono-tokkyo(the patents)” in S2.

In S3, since the patents had been specified by the conversation before, ToR asked what the products USER had manufactured using these patents. This time, ToR chose pronoun to refer to ICOT’s patent in place of specifier-added expression. That is because the patents had been referred in both S2 and U2. And at that point of conversation, ToR had recognized that USER had used ICOT’s patents based on the interpretation of U1, so it’s appropriate to topicalize the patents in S3(“sore-wa(them-TOP)”).

In S4, since both the patents and the products had been specified by the conversation so far, ToR changed topics and asked about manufacturing and selling the products. At that point, there had been a situation that ToR had continued to take leadership of conversation and there was change of topics, therefore ToR decided to add a connective “tokorode(by the way)” to S4. And also at that point, ToR had recognized that USER had used the patents relating to memories of computers but it had not recognized that USER had manufactured and sold these products using these patents. Because of such situations, ToR chose the sentence like S4(word “somewhere” was added in the end of the sentence). That is to say, the meaning of S4 is, ToR asked whether USER had manufactured or not and if USER had manufactured, also asked where USER had manufactured. Moreover, memories of computers was topicalized(“sono-memori(the memories)”) because of the fact that the memories were referred in U3 by the expression “keisanki-no memori(memories of computers)”.

5 Summary and concluding remarks

So far, we have analyzed the method which makes responses of the system more understandable for user in order for the system to make efficient conversation with people. And we have considered these method from a standpoint of performance of language.

we have argued about the information in conversational situations needed for decision of surface features of sentences. The correspondence between the conditions to be appropriate sentence and surface features we discussed in section 3 is described below:

- (1) To decide style, topicalized object and word-order appropriately makes the sentence consistent and clear.
- (2) To insert appropriate connectives makes the relationship between the sentence and the previous conversation clear.
- (3) To use ellipsis appropriately makes the information conveyed by the sentence necessary and sufficient.
- (4) To decide referential expressions appropriately makes the sentence clear.

And we have realized that method in the prototype of the making-contract system(ToR) which we had already developed.

It seems that surface features of sentences are decided by many factors in conversational situations, so the method discussed in this paper is only one of these methods and is restricted to only three surface features. But the framework that surface features of sentences are decided based on such situational factors is general, and this framework can be applied to a future conversational system.

Acknowledgements

We would like to thank Masato Yamazaki and Katsuhiko Komorida for their giving us a chance to do this work and their helpful comments and suggestions through this work. We would also like to thank members of our laboratory for their attendance to our discussions.

This work was done as one of the sub-projects of the Fifth Generation Computer System(FGCS) project.

References

- [Allen 1983] J. Allen: *Recognizing intentions from natural language utterances*, in M. Brady and R.C. Berwick Eds., *Computational Models of Discourse*, MIT Press, 1983, pp.107-166.
- [Barwise 1987] J. Barwise: *Notes on a model of a theory of situations, sets, types and propositions*, informal manuscript, 1987.
- [Grice 1975] H. P. Grice: *Logic and Conversation*, in P. Cole and J.L. Morgan Eds., *Syntax and Semantics: Vol.3, Speech Acts*, Academic Press, 1975, pp.41-58.
- [Kaplan 1983] J. Kaplan: *Cooperative responses from a portable natural language database query system*, in M. Brady and R.C. Berwick Eds., *Computational Models of Discourse*, MIT Press, 1983, pp.167-208.
- [Searl 1969] J. R. Searl: *Speech Acts: an Essay in the Philosophy of Language*, Cambridge Univ. Press, 1969.
- [Yasukawa 1987] H. Yasukawa and H. Suzuki: *Knowledge Representation Language based on Situation Theory*, France-Japan Artificial Intelligence and Computer Science Symposium '87(to be appeared).