EXPERIENCES WITH AN ON-LINE TRANSLATING DIALOGUE SYSTEM

Seiji MIIKE, Koichi HASEBE, Harold SOMERS*, Shin-ya AMANO
Research and Development Center
Toshiba Corporation
1, Komukai Toshiba-cho, Saiwai-ku
Kawasaki-City, Kanagawa, 210 Japan

ABSTRACT

An English-Japanese bi-directional machine translation system was connected to a keyboard conversation function on a workstation, and tested via a satellite link with users in Japan and Switzerland. The set-up is described, and some informal observations on the nature of the bilingual dialogues reported.

INTRODUCTION

We have been developing an English-Japanese bi-directional machine translation system implemented on a workstation (Amano 1986, Amano *et al.* 1987). The system, which is interactive and designed for use by a translator, normally runs in an interactive mode, and includes a number of special bilingual editing functions. We recently realized a real-time on-line communication system with an automatic translation function by combining a non-interactive version of our Machine Translation system with the keyboard conversation function just like *talk* in UNIX**. Using this system, bilingual conversations were held between members of our laboratory in Japan and visitors to the 5th World Telecommunications Exhibition *Telecom 87*, organized by the International Telecommunication Union, held in Geneva from 20th to 27th October 1987.

In the first part of this paper, we discuss in detail the configuration of this system, and give some indications of its performance. In the second part, we report informally on what for us was an interesting aspect of the experiment, namely the nature of the dialogues held using the system. In particular we were struck by the amount of **metadialogue**, i.e. dialogue discussing the previous interchanges: since contributions to the conversation were being translated, this metadialogue posed certain problems which we think are of general interest. In future systems of a similar nature, we feel there is a need for users to be briefly trained in certain conventions regarding metadialogue, and typical system translation errors. Furthermore, an environment which minimizes such errors is desirable and the system must be 'tuned' to make translations appropriate to conversation.

SYSTEM CONFIGURATION

A general idea of the system is illustrated in Figure 1. Workstations were situated in Japan and Switzerland, and linked by a conventional satellite telephone connection. The workstations at either end were AS3260C machines. Running UNIX, they support the Toshiba Machine Translation system AS-TRANSAC. On this occasion, the Machine Translation capability was installed only at the Japanese end, though in practice both terminals could run AS-TRANSAC.

The workstation screens are divided into three windows, as shown in Figure 2, not unlike in the normal version of UNIX's *talk*. The top window shows the user's dialogue, the middle window the correspondent's replies. The important difference is that both sides of the dialogue are displayed in the language appropriate to the location of the terminal. However, in a third small window, a workspace at the bottom of the screen, the raw input is also displayed. (This access to the English input at the Japanese end is significant in the case of Japanese users having some knowledge of English, and of course *vice versa* if appropriate.) The bottom window also served the purpose of indicating to the users that their conversation partners were transmitting.

*the Centre for Computational Linguistics, University of Manchester Institute of Science and Technology, England

**UNIX is a trademark of AT&T Bell Laboratories.

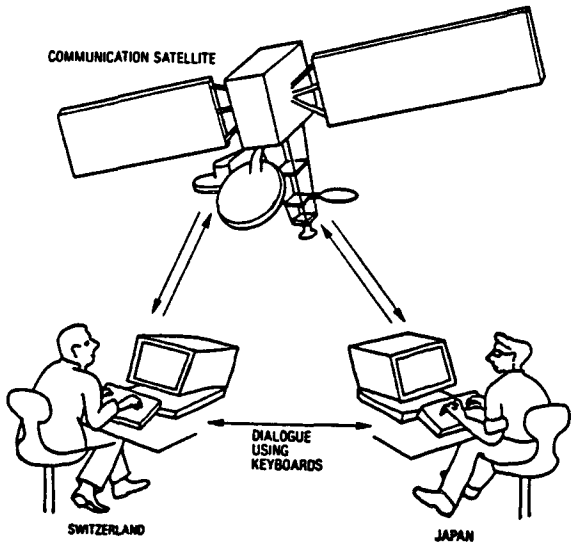
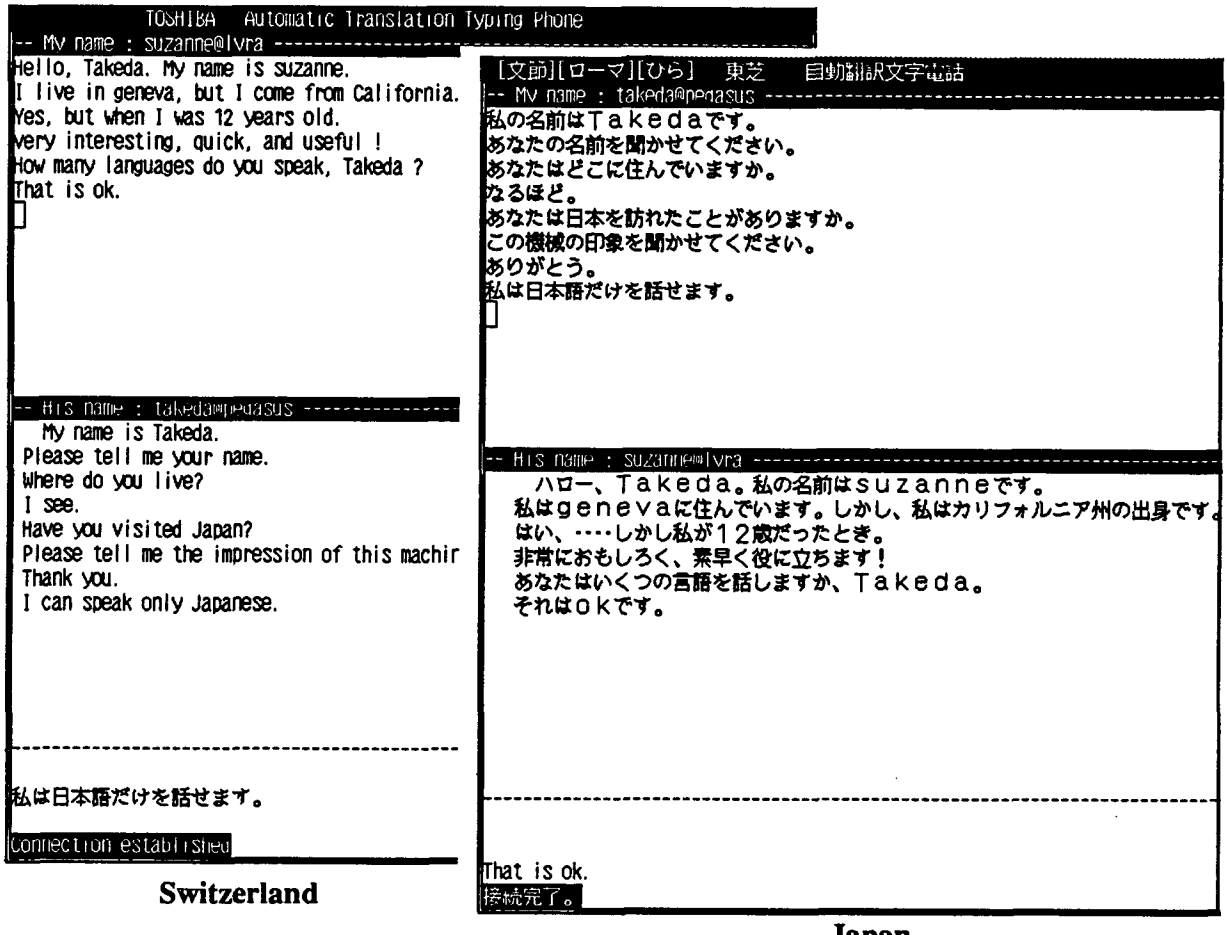


Figure 1. General Set-up



Switzerland **Japan**

Figure 2. Screen Display

Figure 3 shows the set-up in more detail. At the Japanese end, the user inputs Japanese at the keyboard, which is displayed in the upper window of the workstation screen. The input is passed to the translation system and the English output, along with the original input is then transmitted via telecommunications links (KDD's Venus-P and the Swiss PTT's Telepac in this case) to Switzerland. There it is processed by the keyboard conversation function, which displays the original input in the workspace at the bottom of the screen, and the translated message in the middle window on the screen. The set-up at the Swiss end is similar to that at the Japanese end, with the important exception that only the original input message is transmitted, since the translation will take place at the receiving end.

TRANSLATION METHOD

An input sentence is translated by morphological analyzer, dictionary look-up module, parser, semantic analyzer, and target sentence generator. Introducing a full-fledged semantic analyzer conflicts with avoiding increases in processing time and memory use. To resolve this conflict, a Lexical Transition Network Grammar (LTNG) has been developed for this system.

LTNG provides a semantic framework for an MT system, at the same time satisfying processing time and memory requirements. Its main role is to separate parsing from semantic analysis, i.e., to make these processes independent of each other. In LTNG, parsing includes no semantic analysis. Any ambiguities in an input sentence remain in the syntactic structure of the sentence until processed by the semantic analyzer. Semantic analysis proceeds according to a lexical grammar consisting of rules for converting syntactic structures into semantic structures. These rules are specific to words in a pre-compiled lexicon. The lexicon consists of one hundred thousand entries for both English and Japanese.

SYSTEM PERFORMANCE

Once the connection has been established, conversation proceeds as in UNIX's *talk*. An important feature of the function is that conversers do not have to take turns or wait for each other to finish typing before replying, unlike with *write*. This has a significant effect on conversational strategy, and occasionally leads to disjointed conversations, both in monolingual and bilingual dialogues. For example, a user might start to reply to a message the content of which can be predicted after the first few words are typed in; or one user might start to change the topic of conversation while the other is still typing a reply.

Transmission of input via the satellite was generally fast enough not to be a problem: the real bottle-neck was the physical act of input. Novice users do not attain high speed or accuracy, a problem exacerbated at the Swiss end by a slow screen echo. But the problem is even greater for Japanese input: users typed either in *romaji* (i.e. using a standard transcription into the Roman alphabet) or in *hiragana* (i.e. using Japanese-syllable values for the keys). In either case, conversion into *kanji* (Chinese characters) is necessary (see Kawada *et al.* 1979 and Mori *et al.* 1983 on *kana-to-kanji* conversion); and this conversion is needed for between a third and a half of the input, on average (cf. Hayashi 1982:211). Because of the large num-

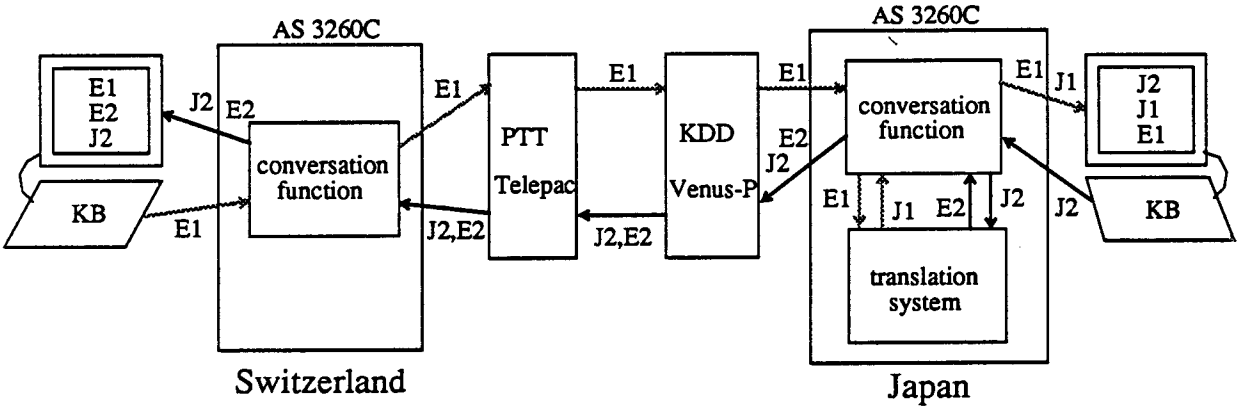


Figure 3. Configuration

ber of homophones in Japanese, this can slow down the speed of input considerably. For example, even for professional typists, an input speed of 100 characters (including conversions) per minute is considered reasonable (compare expected speeds of up to 100 words/minute for English typing). It is of interest to note that this *kana-to-kanji* conversion, which is accepted as a normal part of Japanese word-processor usage, is in fact a natural form of pre-editing, given that it serves as a partial disambiguation of the input.

On the other hand, slow typing speeds are also encountered for English input, one side-effect of which is the use of abbreviations and shorthand. In fact, we did not encounter this phenomenon in Geneva, though in practice sessions (with native English speakers) in Japan, this had been quite common. Examples included contractions (e.g. `pls` for `please`, `u` for `you`, `cn` for `can`), omissions of apostrophes (e.g. `cant`, `wont`, `dont`) and non-capitalization (e.g. `i`, `tokyo`, `jal`).

The translation time itself did not cause significant delays compared to the input time, thanks to a very fast parsing algorithm, which is described elsewhere (Nogami *et al.* 1988). Input sentences were typically rather short (English five to ten words, Japanese around 20 characters), and translation was generally about 0.7 seconds per word (5000 words/hour). Given users' typing speed and the knowledge that the dialogue was being transmitted half way around the world, what would, under other circumstances, be an unacceptably long delay of about 15 seconds (for translation and transmission) was generally quite tolerable, because users could observe in the third window that the correspondent was inputting something, even if it could not read.

TRANSLATION QUALITY

This environment was a good practical test of our Machine Translation system, given that many of the users had little or no knowledge of the target language: the effectiveness of the translation could be judged by the extent to which communication was possible. Having said this, it should also be remarked that the Japanese-English half of the bilingual translation system is still in the experimental stage and so translations in this direction were not always of a quality comparable to those in the other direction. To offset this, the users at the Japanese end, who were mainly researchers at our laboratory and therefore familiar with some of the problems of Machine Translation, generally tried to avoid using difficult constructions, and tried to 'assist' the system in some other ways, notably by including subject and object pronouns which might otherwise have been omitted in more natural language.

We recognized that the translation of certain phrases in the context of a dialogue might be different from their translation under normal circumstances. For example, English *I see* should be translated as *naruhodo* rather than *watashi ga miru*, Japanese *wakarimashita* should be *I understand* rather than *I have understood*, and so on. Nevertheless, the variety of such conversational fillers is so wide that we inevitably could not foresee them all.

The English-Japanese translation was of a high quality, except of course where the users - being inexperienced and often non-native speakers of English - made typing mistakes, e.g. (1). (In these and subsequent examples, E: indicates English input, J: Japanese input, and T: translation. Translations into Japanese are not shown. Typing errors and mistranslations are of course reproduced from the original transmission.)

(1a) E: this morning i came fro st. galle to vizite the exosition.
E: it is vwery inyteresti ng to see so many apparates here.

(1b) E: I arderd "today's menu".

(1c) E: i would go tolike a girl.
J: わかりません。
T: I don't understand.
J: t o l i k e は何ですか。
T: What is tolike?

These were sometimes compounded by the delay in screen echo of input characters, as in example (2).

(2) E: Sometimes, I chanteh the topic, suddenly.
E: I change teh topic.
J: わかります。
T: I understand.
E: I had many mistakes.
J: あなたは疲れていますか。
T: Are you tired?
E: A little.
E: But the main reason is the delay fo dispaying.
E: But the main reason is the delay of display.

Failure to identify proper names or acronyms often led to errors (by the system) or misunderstandings (by the correspondent), as in (3a), especially when the form not to be translated happens to be identical to a known word, as in (3b). In (3b), '*go men na sai*' means in Japanese that I'm sorry.

(3a) E: lars engvall.
J: l a r s e n g v a l l は何ですか。
T: What is lars engvall?
E: this is my name.

(3b) [having been asked if he knows Japanese]
E: How about go men na sai?
T: どのように_について_行く_人_n a_ s a i。

This was avoided on the Japanese-English side where proper names were typed in *romaji* (4).

(4) J: 私の名前はN o g a m i です。
T: My name is Nogami.

As with any system, there were a number of occasions when the translation was too literal, though even these were often successfully understood (5).

(5) E: Do you want something to drink?
J: はい。
T: Yes.
E: What drink do you want?
J: 温かいコーヒーを私は飲みたい。
T: I want to drink a warm coffee.
E: warm coffee?
E: Not a hot one?
J: 熱いコーヒーです。
T: It is a hot coffee.

One problem was that the system must always give some output, even when it cannot analyse the input correctly: in this environment failure to give some result is simply unacceptable. However, this is difficult when the input contains an unknown word, especially when the source language is Japanese and the unknown word is transmitted as a *kanji*. Our example (6) nevertheless shows how a cooperative user will make the most of the output. Here, the non-translation of *tsuki mae* (月前) is compounded by its mis-translation as a prepositional object. The first Japanese sentence said that I married two months ago. But the English correspondent imagines the untranslated *kanji* might mean 'wives'!

(6) J: 私は2月前に結婚した。
T: I married to 2 月前.
E: are married to 2 what???.
J: 今年の6月に結婚した。
T: I married in this year June.
E: now i understand.
E: i thought you married 2 women.

In the reverse direction, the problem is less acute, since most Japanese users can at least read Roman characters, even if they do not understand them (7): this led in this case to an interesting metadialogue. Again, the English user was cooperative, and rephrased what he wanted to say in a way that the system could translate correctly.

(7) E: can you give me a crash course in japanese?.
J: c r a s h c o u r s e は何ですか。
T: What is crash course?
E: it means learn much in a very short time.

Mistranslations were a major source of metadialogue, to be discussed below, though see particularly example (10).

THE NATURE OF THE DIALOGUES

There has been some interesting research recently (at ATR in Osaka) into the nature of keyboard dialogues (Arita *et al.* 1987; Iida 1987) mainly aimed at comparing telephone and keyboard conversions. They have concluded that keyboard has the same fundamental conversational features as telephone conversation, notwithstanding the differences between written and spoken language. No mention is made of what we are calling here metadialogue, though it should be remembered that our dialogues are quite different from those reported by the ATR researchers in that we had a translation system as an intermediary. No comparable experiment is known to us, so it is difficult to find a yardstick against which to assess our findings.

Regarding the subject matter of our dialogues, this was of a very general nature, often about the local situation (time, weather), the dialogue partner (name, marital status, interests) or about recent news. A lot of the dialogue actually concerned the system itself, or the conversation. An

obvious example of this would be a request to rephrase in the case of mistranslation, as we have seen in (6) above, though not all users seemed to understand the necessity of this tactic (8).

(8) E: how does your sistem work please .
J: 私はその文の意味がわかりません。
T: I don't understand a meaning of the sentence.
E: how does your sistem work?

Often, a user would seek clarification of a mis- or un-translated word as in (9), or (3) above.

(9) E: I could have riz in the dinner.
J: r i z はフランス語ですか。
T: Is riz French?
E: May be. I'm not sure.
J: 米ですか。
T: Is it rice?
E: In my guess, you are right.
J: 当然です。
T: It is natural.
E: What is natural?
J: 私はフランス語がわかります。
T: I understand French.
J: r i z は米です。
T: Riz is rice.

The most interesting metadialogues however occurred when users failed to distinguish cited words - a problem linguists are familiar with - for example by quotation marks: these would then be re-translated, sometimes leading to further confusion (10).

(10) J1: 日本の印象を話してください。
T: Please speak a Japanese impression.
E1: ichibana.
J2: わかりません。
J3: i c h i b a n a は何ですか。
T:What is ichibana?
E2: i thought it means number one.
J4: 何が一番ですか。
T:What is the first?
E3: the translation to you was incorrect.

This example may need explanation. First the translation of the Japanese question (J1) has been misunderstood: the translation should have been 'Please give me your impressions of Japan', but the English user (E-user) has understood *Japanese* to mean 'Japanese language'. That is, E-user has understood J1 to be saying 'Please speak an impressive Japanese word.' Then E-user confused **ichiban** ('number 1' or 'the first') and **ikebana** ('flower arranging'). The word ***ichibana*** (E1) does not exist in Japanese. His explanation 'number one' was correctly translated (not shown here) as **ichiban**. But not realizing of course that the meaning of his first sentence (J1) was incorrectly understood, the Japanese user (J-user) could not understand E1 (J2) and asked for its sense (J3). So E-user tried to explain the meaning of ***ichibana***, which in fact was **ichiban**. By the answer, J-user has identified what E-user ment, but since J-user still did not realized that his first sentence was incorrectly understood and hence J-user has understood E2 to be saying that something was 'number 1', he tried to ask what was 'number 1' (J4).

But in the translation of this question, ***ichiban*** (一番) was translated as 'the first'. At this point, it is not clear which comment E-user is referring to in E3, but anyway, not realizing what answer J-user have expected and not knowing enough Japanese to realize what has happened - i.e. the connection between 'number one' and 'the first' - E-user gives up and changes the subject. If E-user had intended to *speak* **ikebana** and explained its meaning, J-user could have realized J1 had been misunderstood. Because it is meaningless in a sentence saying someone's impression that something is **ikebana**.

On the other hand, where the user knew a little of the foreign language (typically the Japanese user knowing English rather than vice versa), such a misunderstanding could be quickly dealt with (11).

(11) E: How is the weathere in Tokyo?
J: w e a t h e r e はw e a t h e r ですか。
T: Is weathere weather?

CONCLUSIONS

There are a number of things to be learnt from this experiment, even if it was not in fact set up to elicit information of this kind. Clearly, typing errors are a huge source of problems, so an environment which minimizes these is highly desirable. Two obvious features of such an environment are fast screen echo, and delete and back-space keys which work in a predictable manner in relation to what is seen on the screen. For the correction of typing errors, the system should have a spelling-

check function which works word-by-word as users are typing in. The main reasons for syntax errors are ellipsis and unknown words. Therefore, the system should have a rapid syntax-check function which can work before transmission or translation and can indicate to users that there is a syntax error so that users can edit the input sentence or reenter the correct sentence. These are in themselves not new ideas, of course (e.g. Kay 1980 and others).

Conventions for citing forms not to be translated, especially in metadialogue, must be developed, and the Machine Translation system must be sensitive to these. The system must be 'tuned' in other ways to make translations appropriate to conversation, in particular in the translation of conversational fillers like *I see* and *wakarimashita*.

Finally, it seems to be desirable that users be trained briefly, not only to learn these conventions, but also so that they understand the limits of the system, and the kind of errors that get produced, especially since these are rather different from the errors occasionally produced by human translators or people conversing in a foreign language that they know only partially.

REFERENCES

AMANO Shin-ya, Hiroyasu NOGAMI & Seiji MIIKE (1988). A Step towards Telecommunication with Machine Interpreter. Second International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages, June 12-14, 1988 (CMU).

NOGAMI Hiroyasu, Yumiko YOSHIMURA & Shin-ya AMANO (1988). Parsing with look-ahead in a real time on-line translation system. To be appeared in COLING 88, Budapest.

AMANO shin-ya, Kimihito TAKEDA, Koichi HASEBE & Hideki HIRAKAWA (1988). Experiment of Automatic Translation Typing Phone (in Japanese). Information Processing Society of Japan (1988.3.16-18)

TAKEDA Kimihito, Koichi HASEBE & Shin-ya AMANO (1988). System Configuration of Automatic Translation Typing Phone (in Japanese). Information Processing Society of Japan (1988.3.16-18)

ASAHIOKA Yoshimi, Yumiko YOSHIMURA, Seiji MIIKE & Hiroyasu NOGAMI (1988). Analysis of the Translated Dialogue by Automatic Translation Typing Phone (in Japanese). Information Processing Society of Japan (1988.3.16-18)

ARITA Hidekazu, Kiyoshi KOGURE, Izuru NOGAITO, Hiroyuki MAEDA & Hitoshi IIDA (1987). Media-dependent conversation manners: Comparison of telephone and keyboard conversations (in Japanese). Information Processing Society of Japan 87-34 (1987.5.22)

IIDA Hitoshi (1987). Distinctive features of conversations and inter-keyboard interpretation. Workshop on Natural Language Dialogue Interpretation, November 27-28, 1987, Advanced Telecommunications Research Institute (ATR), Osaka.

AMANO Shin-ya (1986). The Toshiba Machine Translation system. Japan Computer Quarterly 64 "Machine Translation - Threat or Tool" (Japan Information Processing Development Center, Tokyo), 32-35.

AMANO Shin-ya, Hideki HIRAKAWA & Yoshinao TSUTSUMI (1987). TAURAS: The Toshiba Machine Translation system. MT Machine Translation Summit, Manuscripts & Program, Tokyo: Japan Electronic Industry Development Association (JEIDA), 15-23.

KAY Martin (1980). The proper place of men and machines in language translation. Report CSL-80-11, Xerox-PARC, Palo Alto, CA.

HAYASHI Ooki (kanshuu) (1982). 林大(監修). 「図説日本語」角川小辞典9. 東京(昭和57, 1982). 角川書店.

KAWADA Tsutomu, Shin-ya AMANO, Ken-ichi MORI & Koji KODAMA (1979). Japanese word processor JW-10. Compcon 79 (Nineteenth IEEE Computer Society International Conference, Washington DC), 238-242.

MORI Ken-ichi, Tsutomu KAWADA & Shin-ya AMANO (1983). Japanese word processor. In T. KITAGAWA (Ed.) Japan Annual Reviews in Electronics, Computers & Telecommunications Volume 7: Computer Sciences & Technologies, Tokyo: Ohmsha, 119-128.

APPENDIX A.
Overall Performance Data

sessions	78	times
utterances	1429	times (100%)
	18.3	time/session
utterances that were successfully translated	1289	times (90%)
utterances that were mis-translated	140	times (10%)
metadialogues	31	times
	0.4	time/session

APPENDIX B.
Subject Matter in Utterances

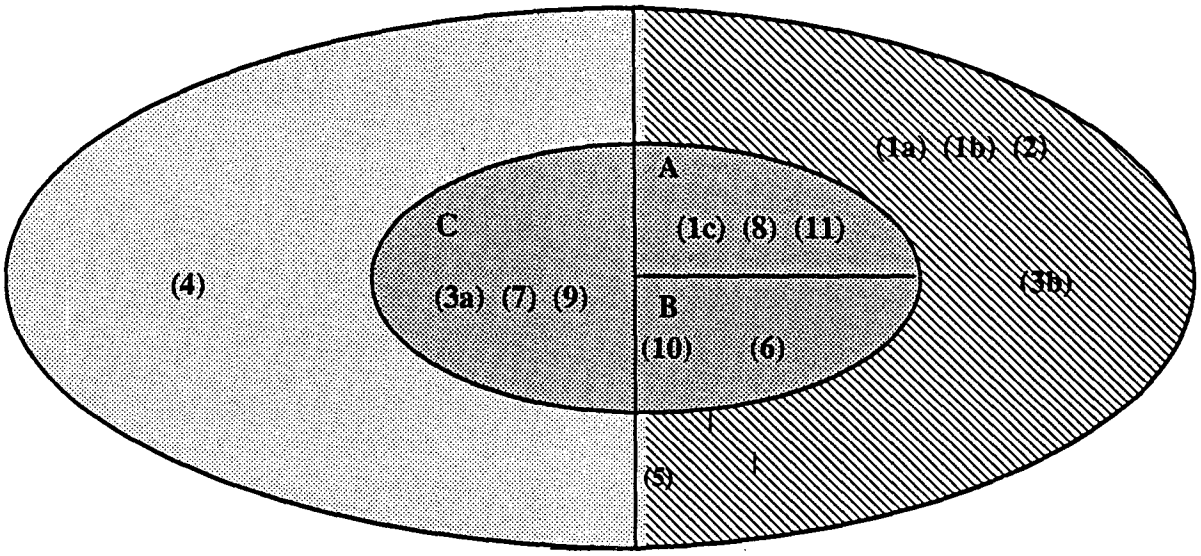
total utterances	1429 times (100%)
greeting and self introduction	470 times (33%)
response signals	154 times (11%)
about weather	92 times (6%)
about time	56 times (4%)
others	657 times (46%)

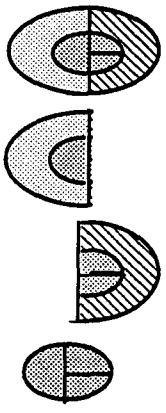
APPENDIX C.
Type of Expressions in Metadialogue

total metadialogues	31 times
repetition of a part of partner's utterances (e.g. What is ichibana?)	22 times
(English users' are 2 and Japanese users' are 20)	
telling typing errors or mistranslations (e.g. Error in Translation.)	9 times
(English users' are 6 and Japanese users' are 3)	

APPENDIX D. Distribution of Utterances

(1a), (1b), (2) and so on are corresponding to examples in the text. Those numbers are put in the area in which main utterances in the examples are involved.





: total 1429 utterances

: 1289 utterances that were successfully translated

: 140 utterances that were mis-translated

: 31 utterances that caused metadialogues
A: by typing errors (7times)
B: by mistranslations (5times)
C: by unknown words to the partner and so on (19times)