MULTI-LINGUAL SPOKEN DIALOG TRANSLATION SYSTEM USING TRANSFER-DRIVEN MACHINE TRANSLATION

Hideki Mima, Osamu Furuse*, Yumi Wakita, and Hitoshi Iida

ATR Interpreting Telecommunications Research Laboratories
2-2 Hikaridai Seika-cho Soraku-gun Kyoto 619-02, Japan
{mima, yumi, iida}@itl.atr.co.jp, furuse@cslab.kecl.ntt.co.jp

Abstract. This paper describes a Transfer-Driven Machine Translation (TDMT) system as a prototype for efficient multi-lingual spoken-dialog translation. Currently, the TDMT system deals with dialogues in the travel domain, such as travel scheduling, hotel reservation, and trouble-shooting, and covers almost all expressions presented in commercially-available travel conversation guides. In addition, to put a speech dialog translation system into practical use, it is necessary to develop a mechanism that can handle the speech recognition errors. In TDMT, robust translation can be achieved by using an example-based correct parts extraction (CPE) technique to translate the plausible parts from speech recognition results even if the results have several recognition errors. We have applied TDMT to three language pairs, i.e., Japanese-English, Japanese-Korean, Japanese-German. Simulations of dialog communication between different language speakers can be provided via a TCP/IP network. In our performance evaluation for the translation of TDMT utilizing 69-87 unseen dialogs, we achieved about 70% acceptability in the JE, KJ translations, almost 60% acceptability in the EJ and JG translations, and about 90% acceptability in the JK translations. In the case of handling erroneous sentences caused by speech recognition errors, although almost all translation results end up as unacceptable translation in conventional methods, 69% of the speech translation results are improved by the CPE technique.

1 Introduction

Systems that deal with spoken dialogs generally require different techniques than systems that deal with written languages. The main requirements for the former are techniques to handle 1) spoken languages containing ungrammatical expressions, and 2) real-time translation to avoid interrupting smooth communication. TDMT [Furuse et al., 1995] has proven to be an efficient method for spoken-dialog translation. In TDMT, constituent boundary patterns [Furuse and Iida, 1996] are applied to an input incrementally; this contrasts with the linguistic manner of applying grammar rules. The result provides for robust parsing that can even handle ungrammatical phenomena such as derivation in metonymical relationships. Additionally, by dealing with best-only substructures utilizing translation examples, the explosion of structural ambiguities is significantly constrained. Accordingly, robust and efficient translation of a spoken-language input can be achieved.

Furthermore, despite recent efforts to improve speech-recognition accuracy, how to handle misrecognized sentences still remains one of the most crucial problems for speech-to-speech systems. Several methods have already been proposed to parse ill-formed sentences or phrases using global linguistic constraints based on a context-free-grammar (CFG) framework, and their effectiveness against misrecognized speech sentences have been confirmed [Mellish, 1989] [Saitou et al., 1988]. These parsings are also used for translation [Lavie et al., 1996]. In these studies, even if the parsing was unsuccessful for erroneous parts, the parsing could be continued by deleting or recovering the erroneous parts. The parsing is assumption-based and uses such techniques as applying a sort of script-based default reasoning for error correction, supplementing missing information [Mayfield et al., 1995], and using stochastic estimation executed by corpus statistics.

On the other hand, this paper addresses the necessity of 'partial translation', which translates misrecognized sentences by utilizing a correct parts extraction (CPE) technique [Wakita et al., 1997]. The CPE uses global linguistic and semantic constraints by an example-based approach in order to extract correct parts from the sentences. Although those translated parts, i.e., subtrees, are often fragmental, the subtrees are not sufficient in extracting suitable meaningful candidate structures because these linguistic constraints are based on the grammatical constraint without semantics.

From the preliminary experiments of the partial translation that utilizes CPE for both the largest part of correct recognition phrases and all possible translation phrases, we can state that CPE can be expected to improve the performance of speech translation systems.

In the next section of this paper, we give an overview of TDMT. Section 3 presents advantages of TDMT in comparison with related research efforts. Section 4 mentions our performance evaluation of the

* Current affiliation: NTT Communication Science Laboratories, 1-1 Hikarinooka Yokosuka-Shi Kanagawa 239 Japan.

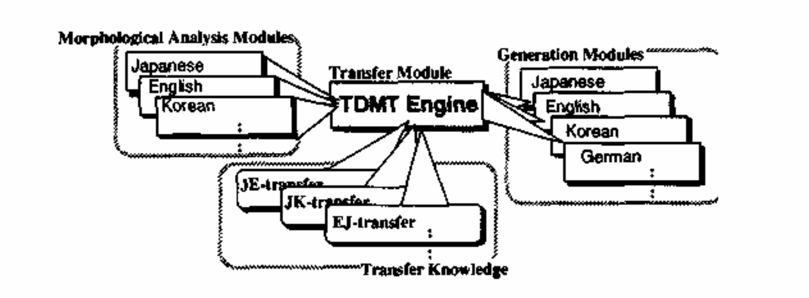

Figure 1. Structure of Multi-lingual TDMT System.

TDMT prototype system using 69-87 unseen dialogs (about 1,300 unseen sentences). Section 5 describes the CPE method for improving the system's performance together with evaluation results of CPE applied to Japanese-to-English speech translation experiments. In section 6, we state our conclusions.

2 Overview of TDMT

In TDMT, translation is performed mainly by a 'transfer module' that applies 'transfer knowledge' to an input sentence. Figure 1 shows the brief structure of a multi-lingual TDMT system. The transfer module, which is an essential component, is a common part of the translation system for every language pair, whereas in language-oriented modules, such as morphological analysis, sentence generation is provided for each and every source/target language. In the following subsections, we will briefly explain 'transfer knowledge' and 'transfer processing'.

2.1 Transfer Knowledge

Transfer knowledge describes the correspondence between source-language expressions and target-language expressions at various linguistic levels.

Source and target-language expressions are expressed in terms of patterns representing meaningful units of linguistic structure to be interpreted by TDMT. A pattern is defined as a sequence that consists of variables and constituent boundary markers such as surface functional words [Furuse and Iida, 1996]. A variable is substituted for a linguistic constituent, and is expressed with a capital letter, such as an *X*.

Transfer knowledge is compiled from such actual translation examples in every source pattern. For example, the Japanese pattern *X no Y,* including the particle *no,* is a very frequently used Japanese expression. We can derive the following Japanese-to-English transfer knowledge about *X no Y* by compiling such translation examples as the source-target pair of *hoteru no juusho → the address of the hotel, eigo no panfuretto → the pamphlet in English,* etc.

X no Y=>
Y' of X' *((hoteru , juusho),* ...),
'hotel' 'address',
Y' in X' *((eigo , panfuretto),* ...),
'English' 'pamphlet'
Y' for X' ((asu , tenkou), ...),
'tomorrow' 'weather'

Within this pattern, *X'* is the target word of *X,* and a corresponding English word is written under each Japanese word. For example, *hoteru* means 'hotel', and *juusho* means 'address'.

This transfer knowledge expression indicates that the Japanese pattern *X no Y* corresponds to many-possible English expressions, *(hoteru, juusho)* are merely sample bindings for *X no Y,* where *X = hoteru,* and *Y = juusho.*

TDMT makes the most of such an example-based framework, which produces an output sentence by mimicking the closest translation example to an input sentence. The semantic distance from the input is calculated for all examples. Then, the example closest to the input is chosen, and the target expression of that example is extracted. The distance between an input and a translation example is measured in terms of a thesaurus hierarchy [Sumita and Iida, 1995]. The distance is calculated quickly because of the simplicity of the mechanism employed.

Suppose that the input is *nihongo no panfuretto,* where *nihongo* means 'Japanese', and the input is closest to *(eigo, panfuretto); the pamphlet in Japanese* can be gained by choosing *Y' in X'* as the best target expression.

2.2 Transfer

The flow of transfer involves the derivation of possible source structures by Constituent Boundary Parsing (CB-parsing) [Furuse and Iida, 1996] and the mapping to target structures based on the results of distance calculations (Fig. 2). When structural ambiguities occur, the best structure is determined by computing the totals for all possible combinations of partial distance values.

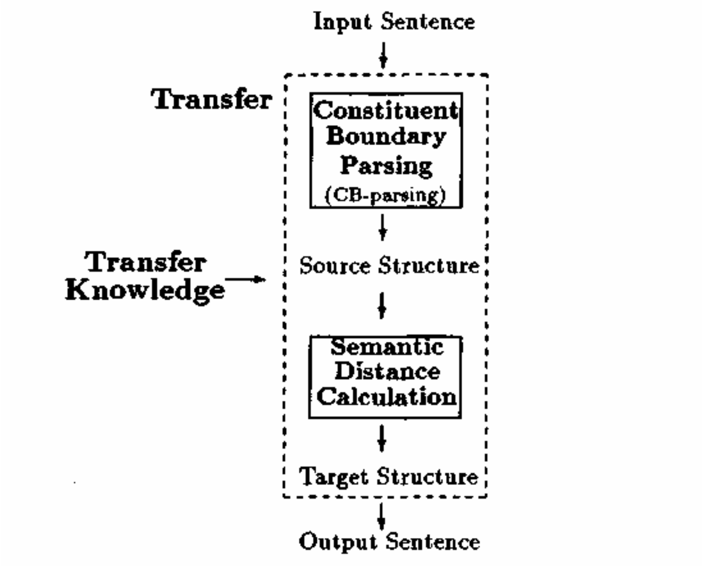

Figure 2. Transfer in TDMT

In general, the EBMT method is particularly effective when the structure of an input expression is short or well-defined and its boundaries have been recognized. When applying it in translation of longer utterances, the input must first be chunked to determine potential patterns by analyzing its phrases after adding part-of-speech tags. In TDMT, as we have already mentioned, translation is performed by stored translation examples, which are represented by "constituent boundary patterns". These are built using limited word-tag information, derived from morphological analysis, in the following sequence [Furuse and Iida, 1996]: (a) insertion of constituent boundary markers, (b) derivation of possible structures by pattern matching, and (c) structural disambiguation using similarity calculation [Sumita and Iida, 1995].

We will show the flow of transfer using the Japanese sentence *Kyouto ni ki te kudasai.* The transfer module derives source structures by combining such source parts of the transfer knowledge as *X te kudasai, X ni Y, Kyouto* and *ki.* Then, based on the result of distance calculations, partial source expressions in the source structure are transferred to *please X', Y' to X', Kyoto* and *come.* By combining these target expressions, the target structure is obtained. From this target structure, the translation output *Please come to Kyoto* can be generated.

If the similarity calculations for candidate phrase patterns were executed top-down & breadth-first, then the calculation cost would be too expensive and the decision on the best phrase would have to be postponed. The translation cost is reduced in TDMT and phrases or partial sentences are analyzed because the current TDMT instead uses an incremental method to determine the best structure locally in a bottom-up & best-only way; this is done to constrain the number of competing structures. Accordingly, even though TDMT fails at whole sentence analysis, partially analyzed substructures can still be obtained.

3 Related Research on Multi-Lingual Translation

In attempts to develop a multi-lingual translation system [Lavie et al., 1996] [Wahlster et al., 1993], an interlingua-based approach has been studied as a theoretically efficient mechanism. Several schemes have been recently proposed based on this approach. The semantic pattern-based parsing in JANUS [Lavie et al., 1996] uses frame-based semantics [Goddeau et al., 1994] with semantic phrase grammar. In this scheme, a recognized speech input is paraphrased into a concrete and simple expression that conforms to one of the system's internal representations. This is performed to make the utterance's meaning easier for the system to understand. Although these inference schemes are powerful in explaining a speaker's intention and the propositional content of the utterance even from a keyword or phrase, plausible default values have to be prepared for achieving heuristic inferences. For example, it is impossible to accept a sentence that includes a metonymical expression like:

(1) "An inexpensive and clean election campaign is nice."

without preparing features able to bridge the semantic gap between inexpensive (clean) and election campaign *a priori.* Therefore, such an approach may work well within a certain domain, but less scalability may be available when extending a prototype system in practice. Additionally, even if a semantic concept unit for at least three languages can be defined, designing a basic generation mechanism that can bridge the semantic gap between languages still remains one of the more difficult problems.

On the other hand, it is generally accepted that people learn source-to-target expression pairs when they learn a foreign language. Therefore, it would seem practical to design reasonable pairs of expressions for each pair of languages based on a common MT mechanism. In other words, it is reasonable that a source language should be analyzed to suit each target language in achieving a practical multi-lingual translation. The motivation for developing TDMT based on this observation.

The main principle of TDMT is to produce translations based on the synchronization of the source and target language structures [Abeillé et al., 1990]. In addition, TDMT has the advantages of: 1) an efficient parsing algorithm; and 2) the capability to handle various linguistic phenomena by utilizing translation examples with a simple best-first mechanism. TDMT's CB-parsing can be implemented according to bottom-up left-to-right chart parsing algorithms. Using best-first syntactic/semantic similarity makes TDMT more flexible [Sumita et al., 1991]. For example, TDMT can accept the previous example sentence (1) by utilizing a similar example like "An inexpensive and clean room would be good"[2]. Therefore, TDMT can achieve both scalability and efficiency in multi-lingual spoken dialog translation.

4 Evaluation of TDMT

Currently, the TDMT system is implemented in LISP and running on UNIX-based machines and deals with dialogues in the travel domain, such as travel scheduling, hotel reservation, and trouble-shooting, and covers almost all expressions presented in commercially-available travel conversation guides. We have applied TDMT to three language pairs, i.e., Japanese-English, Japanese-Korean, Japanese-German. Simulations of dialog communication with free keyboard inputs between different language speakers can be provided via a TCP/IP network.

A system dealing with spoken-dialog requires a quick and accurate response, rather than a grammatical response, in order to provide smooth communication. Moreover, since every process, including speech recognition, translation and speech synthesis, runs automatically from start to finish, there is no room for manual pre/post-editing of input/output sentences in order to make the sentence easier for either the translation process or the user to read. In other words, assuring both efficiency and acceptability in spoken-language translation are the most crucial tasks in devising such a system. Therefore, we evaluated TDMT for both speed and acceptability of translation and analyzed the evaluation results from this point of view.

4.1 The Evaluation Procedure

We evaluated TDMT's translation quality by separately using morphological analysis, a translation module (including a generation module) and a parsing scheme (CB-parsing).

Manually checked morpheme sequences were used to avoid errors and unknown words in testing the translation module itself. This allowed us to assess how well the TDMT would function individually. Details of the evaluation on the morphological analyzer are not described here[3].

[2] Although 'room' and 'election campaign' are semantically rather distant, the syntactic similarity enables TDMT to translate the first example sentence into acceptable Japanese while preserving the metonymical nuance of the meaning.

[3] The success rate in perceiving morphemes was more than 99%, and that in assigning linguistic categories was more than 98% while the analysis speed was less than 0.2 second for each source language measured using a SPARC Station 10 workstation.

Table 1. Experimental conditions.

	JE	JK	JG	EJ	KJ
Vocabulary size	10000			6000	3000
No. of training sentences(diff.)	2602	1195	1553	2431	493
average morphemes/sentence	10.1	9.0	9.3	8.4	7.5
No. of patterns for transfer	887	624	787	1194	320
No. of examples for transfer	10227	3605	2941	7008	1701

Table 2. Evaluation results.

(a) Evaluation results of TDMT

	JE	JK	JG	EJ	KJ
No. of test dialogs	69 (1247 sentences)			73 (1323)	87 (1169)
average morphemes/sentence	9.4			7.1	8.0
(A) (%)	30.1	46.6	27.6	23.7	34.4
(B) (%)	18.1	29.1	11.4	17.2	17.5
(C) (%)	20.6	14.4	10.6	18.5	21.2
(D) (%)	31.2	9.9	50.4	40.5	26.9

(b) Total evaluation results

	JE	JK	JG	EJ	KJ
i. Translation quality					
Acceptability ((A)+(B)+(C)) (%)	68.8	90.1	49.6	59.5	73.1
ii. Parsing quality					
Success ratio of CB-parsing (%)	70.9	60.2	58.7	63.9	43.3
iii. Translation time (sec.)					
Average	0.4	0.3	0.3	0.3	0.1
max (No. of morphemes)	4.4(26)	3.7(31)	3.1(31)	3.3(16)	9.2(15)

Table 1 shows our experimental conditions and evaluation. The reader should note that as the JG and KJ translation project have just started, the transfer knowledge will not always be of the same quality for all language pairs. In addition, though the quality of the thesaurus for each language is an important topic for example-based frameworks, according to the experimental results obtained by applying various kinds of thesauri into TDMT, no remarkable differences in translation quality were observed except for the number of translation outputs.

Translations for 69-87 unseen dialogs (about 1,300 unseen sentences) were manually evaluated by assigning a grade. Two or three native speakers of each of the target languages performed the assessments; all of the examiners were also familiar with their respective source language in order to judge the correctness of the information.

We used the same dialogs for all of the translations whose source language was Japanese, i.e., JE, JK, and JG translations, in order to compare such aspects as differences in the transfer knowledge quality (patterns of expression, examples, etc) and the linguistic distance between languages.

Each sentence was assigned one of four grades for translation quality: (A) Perfect - accurate translation; (B) Fair - a translation that makes it easy to understand the expressions but grammatically flawed; (C) Acceptable - an acceptable translation; (D) Nonsense, - an unacceptable translation or where important information has been translated incorrectly. The translation speed was measured on a SPARC Station 10 workstation with 256MB of memory.

4.2 Results

Table 2 shows the evaluation results for the TDMT, where "acceptability" is the sum of the (A), (B) and (C) grade sentences. The translation speed does not include the time needed for morphological analysis. All ratios are taken from the average of two or three examiners. As a result, about 70% acceptability was achieved in the JE and KJ translations, and almost 60% acceptability was achieved in the EJ and JG translations; remarkably, more than 90% acceptability was achieved in the JK translation, although JK translations needed less transfer knowledge than the others. This observation can be explained from the viewpoint of linguistic similarity; while the Japanese-English (German) language pair is linguistically distant, the Japanese-Korean pair is rather close.

The main problem in the translations involved insufficient examples for CB-parsing. However, an increase in the ratio with the number of examples can be observed in the results. Thus, total accuracy and acceptability should improve in proportion to an increase in transfer knowledge [4]. This trend was observed in the parsing quality results. However, in spite of the rate of false parsing, the quality of both the JK and KJ translations were very high. We consider this finding to be due not only to the linguistic distance between the languages but also to the TDMT pattern-based approach having the robustness to translate phrases while preserving the nuance even for ungrammatical input sentences.

[4] This is accepted in general for the example-based framework since the exact match ratio is certain to increase in proportion to the increase in translation examples. In fact, a total accuracy of more than 93% was achieved in our closed test evaluation for over 1,000 sentences for EJ TDMT translation.However, we have to ascertain the practical satiation limit, or how much the transfer knowledge can be expanded.

Although speed depends on the amount of knowledge and sentence length, average translation times were less than 0.4 sec; thus, TDMT can be considered as an efficient multi-lingual spoken dialog translation mechanism.

5 Handling Ill-formedness in Speech Translation

5.1 Major Trends in Speech Recognition Results

Almost all current continuous speech recognition systems produce word sequences as their output. Generally, they do not make errors in spelling words because wrong phoneme sequences are handled by modifying them into certain word sequences. Thus, a method to eliminate incorrect word sequences, including unnecessary or unrelated words, that is different from an error correction method is expected to avoid errors for speech-to-speech systems. At least three types of 'ill-natured' speech recognition outputs can be observed: 1) a word sequence including an incorrect subsequence that can be analyzed by a kind of syntactic parser, 2) an incorrect subsequence that can be interpreted as a correct meaning by a kind of semantic analysis, and 3) a word sequence difficult to divide into sentences or clauses determined by syntactic constraints.

In addition, a general speech recognition system for dialogue utterances produces no boundary marker between utterances nor punctuation in written texts. A Japanese sentence, in particular, requires a verb at the end of a sentence, and the sentence can be regarded as an embedded clause for modifying the following noun word that must be at the head of the next sentence if the two sentences are concatenated with no sentence boundary. Such a wrong analysis easily happens when a system is handling speech inputs that include a lot of fragmental utterances.

Consequently, we believe that the following problems must be resolved for the best candidate in one sentence lattice produced by the speech recognition process:

— how to identify grammatically correct phrases/clauses
— how to cut off wrong sub-phrases
— how to select meaningful phrases
— how to correct wrong sub-phrases

Despite recent efforts to achieve speech-to-speech systems, the above problems still remain the most to handle.

However, although these problems are crucial, both parsing techniques and utilizing memorized phrases as language usage examples have good effects on achieving understanding between speakers in dialogues. A method of handling ill-formed sentences in translation based on these approaches is shown in the next subsection.

5.2 Correct Parts Extraction for Partial Translations

A conventional syntactic parser normally handles a definite sentence or phrase. On the other hand, continuous speech is often fragmentary and stopped in the middle of utterances. Under such circumstances, a parser is required that can always analyze such incomplete sentences/phrases in spite of the forms of the sentences/phrases. At a minimum, a bottom-up parsing method is desirable because an input cannot be identified as a sentence/clause/phrase before it is analyzed; moreover, referring to phrase examples memorized as language usage is required so that a grammatical or semantical chunk can be easily identified as a sub-structure. In this latter process, the estimation of finding a suitable example corresponding to the chunk depends on a certain criterion judged against existing examples similar to the chunk in the example data. Here, for a sub-structure Example-based MT [Sumita et al., 1991] obtained a suitable example in the same language, and at the same time produced one appropriate translation result for the structure. This means that capturing a similar example to a phrase of the input makes it possible to translate the part, which can be regarded as a correct speech recognition phrase. On the other hand, if the translation fails, the phrase must be ungrammatical. Such a failure happens in some case - for example, an incorrect recognition word sequence, an ungrammatical one, an incorrect segmented one that indicates a concatenation of two phrases/sentences as mentioned above, and so on.

The partial translation method using the correct parts extraction (CPE) technique has been proposed based on these observations. CPE obtains correct parts from recognition results by using the following two factors in CB-parsing for the extraction:

— the semantic distance between the input expression and an example expression, and
— the structure selected by the shortest semantic distance.

The advantages of using CB-parsing are as follows.

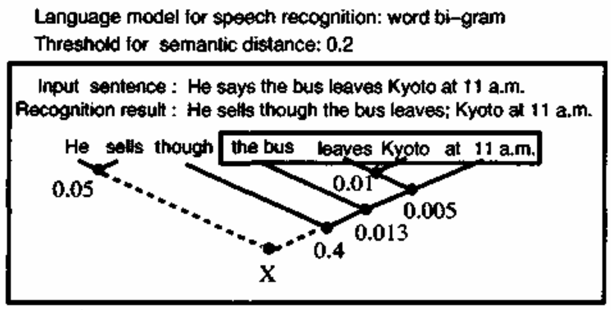

Figure 3. Example of correct part extraction

—CB-parsing can analyze spontaneous speech that cannot be analyzed by the CFG framework only if the example expressions are selected from a spontaneous speech corpus. With more expressions in spontaneous speech, there is an increased ability to distinguish between erroneous sentences and correct ones.
—CB-parsing can deal with patterns including over N words that can not be dealt with during speech recognition.
—CB-parsing can extract partial structures independently from results of parsing, even if the parsing fails for a whole sentence.

Correct parts are extracted under the following conditions:

—When expressions including erroneous words show large distance values to the examples. When the distances are over the distance threshold, the parts are defined as "erroneous parts".
- Correct parts are extracted only from global parts consisting of over N words. If local parts including less than N words cannot have a relation to other parts, the parts are defined as "erroneous parts", even if the semantic distances are under the threshold.

Figure 3 shows an example of CPE. The input sentence /He says the bus leaves Kyoto at 11 a.m./ is recognized as /He sells though the bus leaves Kyoto at 11 a.m./ by continuous speech recognition using a word bi-gram. The solid lines in Figure 3 indicate partial structures and the number for each structure denotes the corresponding semantic distance value. The dotted line indicates the failure analysis result. In this example, the analysis for the whole sentence is unsuccessful because the part /He says/ is misrecognized as /He sell though/. First, the distance value of the longest part, /though the bus leaves Kyoto at 11 a.m./, is compared with the threshold value . The part is considered to include erroneous words because the distance value 0.4 is larger than the threshold value 0.2. Then, the next longest part /the bus leaves Kyoto at 11 a.m./ is evaluated. This part is extracted as a correct part because the distance 0.005 is under the threshold value. Finally, the remaining part /He sells/ is evaluated. The distance of the part /He sells/ is under the threshold value, but the part includes only two words, which are fewer than N, so the part /He sells/ is regarded as an erroneous part.

5.3 Evaluation of CPE for Effectiveness to Speech Translation

We conducted preliminary evaluation of CPE using the Japanese-to-English TDMT system mentioned in the previous section. The obtained recognition results were first analyzed and then partial structures and their semantic distances were output. Next, the correct parts were extracted and only the extracted parts were translated into target sentences.

For the evaluations, we used 70 erroneous results output by a speech recognition experiment using the ATR spoken language database on travel arrangement. The threshold values of CPE in the semantic distance and word length were determined as 0.2 and 3, respectively, based on the observations of recall and precision rates in terms of the number of correct words in extracted parts [Wakita et al., 1997].

The translation results were evaluated manually by five Japanese. They assigned one of the following five levels (L1)-(L5) to each misrecognition result after extraction by comparing the results with the corresponding correct sentence before speech recognition. The five levels were:

(L1) Able to understand the same meaning as the correct sentence
(L2) Able to understand, but the expression is slightly awkward
(L3) Unable to understand, but the result is helpful in imagining the correct sentence
(L4) Understanding of the wrong meaning: CPE is not helpful
(L5) Output of the message "Cannot translate"

Table 3. Effect of CPE on translating misrecognition results

Levels	(L1)	(L2)	(L3)	(L4)	(L5)
without CPE	11.9%	0%	0%	2.4%	85.7%
after CPE	25.7%	16.7%	26.6%	21.0%	10.0%

Each of the average rates of the three evaluators is shown in Table 3.

Without CPE, 85.7% of the recognition results could not be translated. It seems that CPE is good for (L1)-(L3) but poor for (L4); (L5) shows a negligible effect. The correctness rate for translation after CPE is more than double the rate before CPE (11.9% to 25.7%). The sum of (L1)-(L3) is 69%. This means that the proposed CPE is effective in improving translation performance. However, we cannot ignore the fact that 21% of the recognition results were translated into erroneous sentences.

6 Conclusion

TDMT has been presented as a general and efficient mechanism for multi-lingual spoken dialog translation. Experimentation of the prototype system has shown that the TDMT can be expected to achieve efficient multi-lingual spoken dialog translation.

In addition, although handling ill-formedness is one of the most crucial problems for speech translation, we have examined what the speech recognition process needs for understanding the misrecognized sequences and have demonstrated the necessity of partial translation by utilizing the CPE method.

A preliminary experiment showing performance improvements in speech translation by using TDMT and CPE makes us believe that the proposed scheme can achieve robust and efficient spoken dialog translations.

In order to increase the generality of TDMT, we are planning to expand the system's transfer knowledge by continuing with training in linguistic knowledge, patterns of expression, and translation examples.

Additionally, we will try to feed the partial results extracted by CPE back into the speech recognition process for re-recognizing only the non-extracted parts to improve the performance of speech recognition. This kind of tight integration of the speech and translation processes to improve the robustness and acceptability of speech translation is also an interest of future work.

References

[Abeillé et al., 1990] Anne Abeillé, Yves Schabes and Aravind K. Joshi. 1990. Using Lexicalized Tags for Machine Translation. In *Proceedings of the 13th International Conference on Computational Linguistics,* pages 1-6.

[Furuse et al., 1995] Osamu Furuse, Jun Kawai, Hitoshi Iida, Susumu Akamine and Deok-Bong Kim. 1995. Multi-lingual Spoken-Language Translation Utilizing Translation Examples. In *Proceedings of NLPRS'95,* pages 544-549.

[Furuse and Iida, 1996] Osamu Furuse and Hitoshi Iida. 1996. Incremental Translation Utilizing Constituent Boundary Patterns. In *Proceedings of Coling '96,* pages 412-417.

[Goddeau et al., 1994] David Goddeau, Eric Brill, James Glass, Christine Pao, Michael Phillips, Joseph Polifroni, Stephanie Seneff and Victor Zue. 1994. Galaxy: A Human-Language Interface to On-Line Travel Information. In *Proceedings of ICSLP94,* pages 707-710.

[Lavie et al., 1996] Alon Lavie, Donna Gates, Marsal Gavalda, Laura Mayfield, Alex Waibel and Lori Leven. 1996. Multi-lingual Translation of Spontaneously Spoken Language in a Limited Domain. In *Proceedings of Coling '96,* pages 442-447.

[Mayfield et al., 1995] L. J. Mayfield, M. Gavalda, Y-H. Seo, B. Suhm, W. Ward, A. Waibel. 1995. Parsing Real Input in JANUS: a Concept-base Approach to Spoken Language Translation. In *Proceedings of TMI-95,* pages 196-205.

[Mellish, 1989] C. S. Mellish. 1989. Some chart-based techniques for parsing ill-formed input. In *Proceedings of the Annual Meeting of the ACL,* pages 102-109.

[Saitou et al., 1988] Hiroaki Saitou and Masaru Tomita. 1988. Parsing noisy sentences. In *Proceedings of COL-ING'88,* pages 561-566.

[Sumita et al., 1991] Eiichiro Sumita and Hitoshi Iida. 1991. Experiments and Prospects of Example-based Machine Translation. In *Proceedings of 29th ACL,* pages 185-192.

[Sumita and Iida, 1995] Eiichiro Sumita, and Hitoshi Iida. 1995. An Example-Based Disambiguation of English Prepositional Phrase attachment. In *Proc. Syst. and Com. in Japan,* Vol.26 , No.4, pages 30-41.

[Wahlster et al., 1993] Wolfgang Wahlster. 1993. Verbmobil: Translation of Face-To-Face Dialogs. In *Proceedings of MT-Summit IV,* pages 127-135.

[Wakita et al., 1997] Yumi Wakita, Jun Kawai, and Hitoshi Iida. 1997. Correct Parts Extraction from Speech Recognition Results using Semantic Distance Calculation, and Its Application to Speech Translation. In *Proceedings of an ACL Workshop on Spoken Language Translation,* pages 24-31.