Utilizing Extra-Grammatical Phenomena in Incremental English-Japanese Machine Translation *

Shigeki Matsubara and Yasuyoshi Inagaki

Department of Information Engineering
Nagoya University
Furo-cho, Chikusa-ku, Nagoya, 464-01, Japan
{matu,inagaki}@inagaki.nuie.nagoya-u.ac.jp

Abstract. Since spontaneous speech appears continuously, each stage of a spoken language machine translation system should work incrementally. This paper proposes an incremental machine translation system, which translates English spoken words into Japanese in accordance with the order of appearances of them. The system is composed of three modules: parsing, transfer and generation, which work incrementally and synchronously. The transfer module utilizes some of extra-grammatical phenomena characterizing Japanese spoken language: flexible word-order, ellipses, repetitions and so forth. This is influenced by the observational facts that such phenomena frequently appear in Japanese uttered by English-Japanese interpreters. Their frequent utilization is the key to success of the incremental translation between English and Japanese, which have different word-order. We have implemented a prototype system Sync/Trans. To evaluate Sync/Trans we have made an experiment with conversations consisting of 27 dialogues and 218 sentences. 190 of the sentences are correct, providing a success rate of 87.2%. This result shows our method to be an effective technique for spoken language translation with acceptable accuracy and high real-time nature.

1 Introduction

With the advance of communication technology in recent years, there have appeared new requirements of spoken language processing systems. In particular, on spoken language translation, which is a basic technology for advanced real-time machine translation systems such as automatic telephone interpretation and speech dialogue interpretation, a number of studies have been made. Spontaneously spoken language, different from written language with which most traditional machine translation systems have dealt, has the following features:
(1) ill-formed expressions appear frequently, and
(2) speech appears continuously.
It is thus important to consider the above two features for the purpose of the development of practical spoken language machine translation systems.

In order to cope with the first feature, that is, robustly translate ill-formed expressions including extra-grammatical phenomena, fragment expressions and so forth, many corpus-based approaches have been proposed recently. This framework, which includes statistics-based translation [Brown, 1990, Doi & Murakami, 1992], example-based translation [Sumita et al., 1990, Furuse & Iida, 1992], has achieved some degree of accuracy. Most machine translation systems

* This paper appeared in *Proceeding of TMI-97*. The authors wish to thank Mr. Yoshiaki Iwata for his valuable contributions on an earlier version of this study, Dr. Fransis Bond for providing us with the helpful material. The first author has been supported by Support Center for Advanced Telecommunications Technology Research, Foundation (SCAT).

which have been developed so far, however, do not cope with the continuous nature of speech. These systems, because of their sentence-by-sentence-based approaches, cannot start transferring it until the whole sentence is uttered, which inhibits real-time nature of processing. To satisfy the requirement of the second feature, it is necessary to develop a technique for incremental translation whose processing synchronizes with the input [Menzel, 1994]. Moreover, in spoken language machine translation, the high degree of incrementality is preferred to that of quality.

This paper proposes an incremental machine translation system which translates English input words into Japanese in accordance with their order of appearances [1]. The system is composed of three modules: incremental parsing, transfer and generation, which work synchronously. Since it translates English sentences in an exceedingly incremental way, the system can be expected to be available for an advanced real-time machine translation system with simultaneous interpretation [Inagaki & Matsubara, 1995].

The incremental transfer utilizes some extra-grammatical phenomena characterizing Japanese spoken language, that is, does Japanese translations including flexible word-order, ellipses, repetitions and so forth. The utilization of these characteristics is the key to success of incremental translation between English and Japanese, which have different word-order. The linguistic phenomena used in our system also appear frequently in human dialogue. Consequently, the user could understand the translations with ease.

We have implemented a prototype system Sync/Trans which can parse English dialogues on a word-by-word basis and generate Japanese speech immediately. To evaluate the effectiveness of Sync/Trans, we have made an experiment with all the dialogues which appear in an English textbook for the seventh grade in Japan. The results have shown our incremental method to be an effective technique for spoken language machine translation with acceptable accuracy and high real-time nature.

This paper is organized as follows: Section 2 discusses the issue of incremental translation. Section 3 explains the idea of utilizing extra-grammatical phenomena in the incremental translation. Section 4 introduces the technique for incremental English-Japanese machine translation shortly. Section 5 reports on experimental results on Sync/Trans.

2 Incremental Translation

As Menzel has pointed out, incrementality is an inevitable property of applications of spoken language processing such as the simultaneous interpretation [Menzel, 1994]. This view that the traditional sentence-by-sentence-based approaches might not be appropriate for spoken language processing systems already existed in about 1960. (For example, Hockett advocated such a viewpoint as "Grammar for the Hearer" [Hockett, 1961].) Such a view is strongly suggested by our intuition that the human does not start analyzing speech after hearing the whole sentence, but analyzes it as it is heard. Nonetheless, possible computational applications of incremental language processing to spoken language systems have received less attention so far [Milward & Cooper, 1994].

In this section, first we discuss the issue of incremental machine translation systems, and, next, refer to the related studies.

[1] The word "incremental" originally means to proceed bit-by-bit according to an order. Therefore, it has been used in some differing senses in the field of natural language processing [Wiren, 1989, Pereira & Pollack, 1991]. We will use "incremental translation" for the translation on a possibly word-by-word basis as the sentence is interpreted from left to right.

2.1 Fundamental Issue

Incremental translation generates target languages in a way of possibly word-by-word according to the order of appearances of source words. The important keys to succeed in such the translation are incrementality of the processing and accuracy of the results. However, there exists a trade-off between them.

Provided that incrementality is stressed, the appropriate translations might not be made up. For example, translating the following English sentence:

E1 I met Yumi in the park yesterday.

into Japanese from left to right, then we would obtain the translation J1-1 whose both quality and accuracy are very low.

J1-1 *watashi-wa* (I) *atta* (met) *Yumi-ni* (Yumi) *koen-de* (in the park) *kinoo* (yesterday).

Provided accuracy is stressed, in contrast, the order of appearances of English words cannot be taken into consideration. For example, the Japanese corresponding to E1 is as follows:

J1-2 *watashi-wa* (I), *kinoo* (yesterday), *koen-de* (in the park) *Yumi-ni* (Yumi) *atta* (met).

It is, however, impossible to generate "*kinoo koen-de Yumi-ni atta*" before the end of E1 because the word "yesterday" appears at the end of E1.

There exists an essential difference between English and Japanese in the word-order. This is a fatal issue to incremental translation.

2.2 Related Work

To develop a technique for translating incrementally and synchronously is important for advanced speech-to-speech translation systems. Kitano has proposed a method of incremental sentence production for modeling simultaneous English-Japanese interpretation [Kitano, 1990]. Amtrup has introduced a technique for chart-based incremental transfer for head switching in German-English machine translation [Amtrup, 1995]. However, they both adopt the method of generating the target of the source verb phrase after the source is completely uttered, namely that E1 is translated into J1-2. Therefore, we can say they did not achieve high degree of incrementality and synchronicity. On the other hand, our method generates the translations in an exceedingly incremental and synchronous way.

3 Utilizing Extra-Grammatical Phenomena

3.1 Characteristic Feature of Japanese Speech

To solve the issue pointed out in Section 2.1, it is an effective way to adopt an approach based on human interpretation processes. We have studied on features and phenomena characterizing Japanese utterances by English-Japanese interpreters [Matsubara & Inagaki, 1996a]. We used the bilingual dialogue data of ATR Dialogue Database (ADD) [Ehara et al., 1990]. Table 1 shows the features of the data. The results of their studies have clarified that extra-grammatical phenomena appear frequently: flexible word-order (26 utterances), repetitions (31 utterances) and ellipses (74 utterances). We have also observed these phenomena appear more frequently in Japanese speech by interpreters than in the usual one [2]. These facts are just considered to characterize incrementality of the interpretation process, and also suggest that humans can easily understand the utterances including the above mentioned phenomena.

[2] Compare with [Takezawa et al., 1995].

Table 1. Bilingual dialogue data

Number of dialogues	12
Dialogue domain	Inquiries about international conference, Applications of travels
Dialogue participants	1) English speaker, 2) E/J interpreter, 3) Japanese speaker, 4) J/E interpreter
Object of study	Utterances by 2) E/J interpreter
Number of utterances	512

Table 2. Example of incremental translation process

I	**met**	**Yumi**	**in the park**	**yesterday**	.
watashi wa atta.		*Yumi-ni*	*koen-de*	*kinoo*	*atta.*

3.2 Basic Idea

Let us now try to accurately and incrementally translate **E1** into the Japanese. Table 2 shows the process of the incremental translation, which we propose in this paper. Details of the process are as follows: For the first input word "**I**", the Japanese "*watashi*" is generated, and for second "**met**", a suffix "*wa*" and a verb "*met*". Since in Japanese a verb usually appears at the end of a sentence, here ends the first translation sentence once. That is, the translation is *divided* into two sentences. Although the translation "*watashi-wa atta* (I met $\cdots$. [3])" does not specify whom, where, and when I met, it includes the content to be translated at least.

Next, we turn to the second translation sentence. For an input word "**Yumi**", the target expression "*Yumi-ni* " is generated. Moreover, after the prepositional phrase "**in the park**", "*koen-de*", and for the last word "**yesterday**", "*kinoo*". Finally, in order to arrange the style of the translation the target verb "*atta*" is generated again. The word "*atta*" is *repeated* to add more detailed information to the first translation. The translation "*Yumi-ni koen-de kinoo atta.*" can be easily understood because of high *flexibility of word-order* in Japanese. Although the second translation *omits* the person who met her, it can be restored by the first translation.

4 An Incremental English-Japanese Machine Translation System

An incremental English-Japanese machine translation system is composed of three modules: incremental chart parsing, rule-based transfer and generation, which work synchronously. Figure 1 shows the configuration of the system. To incrementally translate English sentences, the system executes the below (I)-(III) for each word.

(I) **Incremental chart parser** makes the edge with the structure, whose category is S, as the label. From now on, we call such a structure, simply, *sentence structure*. The parsing procedure follows the bottom-up and top-down chart processing framework which we have proposed [Matsubara et al., 1996b] [4]. If no sentence structure can not be gained at all, the system fails in translation.

[3] This means the sentence whose objective case is omitted.

[4] Let us complement the small explanation of the procedure here. The altering point for the orthodox bottom-up chart parsing method is the syntactic operation for gradually embodying the sentence

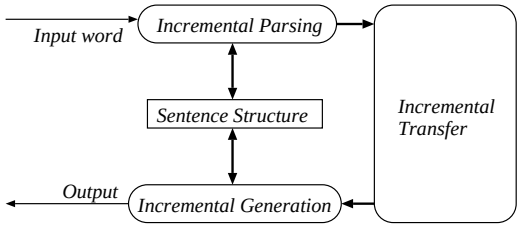

Figure 1. Configuration of an incremental English-Japanese machine translation system

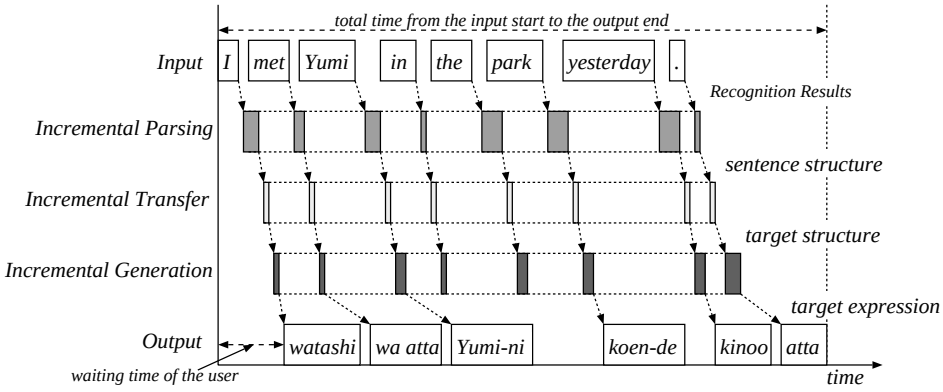

Figure 2. Trace of the translation process for E1

(II) **Incremental transfer** transforms the sentence structure which incremental chart parser has acquired, into the target structure by applying transfer rules in a top-down fashion. How it utilizes the extra-grammatical phenomena is described in the transfer rules. Obviously, the transfer module operates not only complete structures but also incomplete ones.

(III) **Incremental generation** transforms the target structure into the character string. In the spoken language translation, the source expressions transferred once does not have to be transferred any longer. Therefore, the incremental generation also operates on the sentence structure, and marks them as *transferred*.

Figure 2 shows the outline of the translation process for E1 [5] . It is clearly seen from the figure that the system processes the input sentence in an incremental and synchronous way, and therefore both the user's waiting time and the total translation time are getting shorter. Though we showed a translation example of a simple declarative sentence here, the system can cope with various type of English sentences by effectively utilizing extra-grammatical linguistic phenomena: such as *flexible word-order*, *repetitions* and *ellipses*.

---

structure, which is initialized as $[?]_S$. The operation is defined as the replacement of an undecided term in the sentence structure with the other undecided one. So, our parser behaves in the same way as combinatory categorial grammar [Ades & Steedman, 1982] which is famous as the framework for incremental parsing.

[5] To avoid conflicts between grammar rules or between transfer rules, our prototype system lists them according to the frequencies of use respectively.

Table 3. Translation results of 218 sentences

type	sentences	rate
(A) correct	171	78.5%
(B) correct but unnatural	19	8.7%
(C) failed	19	8.7%
(D) incorrect	9	4.1%

Table 4. Typical correct translations

(1) Declarative	(source)	I often eat tofu at dinner .
	(target)	*watashi wa yoku tabe-masu. tofu-wo yushoku-de tabe-masu.*
(2) Negative	(source)	But he doesn't cook tempura .
	(target)	*shikashi kare wa ryoori-shimasen. tempura-wo ryoori-shimasen.*
(3) Imperative	(source)	Please, come in .
	(target)	*onegai-shi-masu haitte-kudasai.*
(4) Interrogative	(source)	Do you play baseball ?
	(target)	*anata-wa shi-masu-ka ? yakyu-wo shi-masu-ka ?*
(5) Interrogative	(source)	What else can you make ?
	(target)	*nani-ka hoka-no-mono wo anata-wa tsukure-masu-ka ?*

5 Sync/Trans and Its Evaluation

To demonstrate the feasibility and effectiveness of the incremental machine translation system, we have developed a prototype system Sync/Trans in Common Lisp. The system has been implemented in the scale of English lexicon 500 words and 60 grammar rules, and the transfer rule has been established corresponding to each grammar rule. Sync/Trans executes the synchronous speech output with the input.

We have made an experiment on Sync/Trans. The conversations consisting of 27 dialogues and 218 sentences which appear in an English textbook for the seventh grade in Japan, have been used. As Table 3 shows, we have classified the English sentences according to the translation results.

English sentences classified into (A) and (B) are successful in translation. 190 of the sentences are semantically correct, providing a success rate of 87.2%. This reveals our technique to be effective for incremental translation with acceptable accuracy. Table 4 shows the typical correct translations. The translations on Sync/Trans are different, in the sense that they are ill-formed expressions, from those on a conventional system. Their translations, however, represent exactly the semantic contents.

Unsuccessful 28 sentences, which are classified into (C) or (D), can be also classified according to the causes as Table 5 shows. We explain their causes below.

- **Parallel phrase (6 sentences)**
  Sync/Trans cannot parse a parallel phrase. Because it is impossible to clarify the structure of the parallel phrase just when a conjunction is inputed.
  Example: I change clothes *and* eat dinner.
- **Structural ambiguity (13 sentences)**
  Sync/Trans might fail to parse an word with the multiple categories.

Table 5. Incorrect translations

type	level	cause	sentences	rate
(C)	Parser	parallel phrase	6	2.7%
		structural ambiguity	13	6.0%
(D)	Transfer	lexical ambiguity	5	2.3%
		idiomatic phrase	4	1.8%

Example: Are *these* books yours?

- **Lexical ambiguity (5 sentences)**
  Sync/Trans might incorrectly translate a word with the multiple senses.
  Example: Do many *Japanese* like origami ?
- **Idiomatic phrase (4 sentences)**
  Obviously, Sync/Trans cannot translate idiomatic phrases appropriately.
  Example: *I beg your pardon.*

6 Concluding Remarks

This paper has proposed an incremental English-Japanese machine translation system. The system is composed of three modules: parsing, transfer and generation, and they work incrementally and synchronously. Utilizing linguistic phenomena characterizing Japanese speech has been the key to success of the incremental machine translation between English and Japanese, which have different word-order.

To evaluate the feasibility and effectiveness of the system. we have implemented a prototype system Sync/Trans The experiment with Sync/Trans has shown our incremental method to be a promising technique for spoken language translation with acceptable accuracy and high real-time nature. The following points are left as future work.

- Section 5 has revealed the causes of the errors by Sync/Trans. These errors are certainly fatal to the incremental translation. However, even though the system generates the incorrect translations, it could correct the errors by utilizing self-repairs. In fact, a human speaker makes self-repairs very frequently in spontaneous speech [Sagawa et al., 1994]. We are studying the technique to naturally correct the errors by self-repairs, and will report about this at a later date.
- We have used very simple English spoken dialogues in the experiment with Sync/Trans. They all consist of only *simple sentences*, and the average length of them is 4.0 words (cf. The average length of English sentences in ADD is 4.9 words.) . In spontaneous dialogues, however, more complex sentences would also often appear. A study on an incremental transfer for such the sentences is required.

The achivement of this paper can be regarded as the first step toward simultaneous interpretation, which is just one of ambitious applications in the near future. As the other related application, we can consider multi-modal dialogue translation systems. In order that the user may simultaneously understand various information such as speech, text, pointing and gesture, the system would be required to output the translation in a syncronous way.

References

[Ades & Steedman, 1982] Ades, Anthony E. and Mark J. Steedman. 1982. On the Order of Words. *Linguistic and Philosophy* 4:517–558.

[Amtrup, 1995] Amtrup, Jan W. 1995. Chart-based Incremental Transfer in Machine Translation. In *Proceedings of the Sixth International Conference on Theoretical and Methodological Issues in Machine Translation, TMI '95*, Leuven, pp.188-195.

[Brown, 1990] Brown, P. F. 1990. A Statistical Approach to Machine Translation. *Computational Linguistics* 16:79–85.

[Doi & Murakami, 1992] Doi, S. and K. Murakami. 1992. Translation Ambiguity Resolution based on Text Corpora of Source and Target Languages. In *Proceedings of the Fifteenth International Conference on Computational Linguistics, COLING-92*, Nantes, pp.525–531.

[Ehara et al., 1990] Ehara, Terumasa, Naomi Inoue, Hideo Kohyama, Toshirou Hasegawa, Fumi Shohyama and Tsuyoshi Morimoto. 1990. Contents of the ATR Dialogue Database. *Technical report TR-I-0186*, ATR Interpreting Telephony Research Laboratories, Kyoto, Japan.

[Furuse & Iida, 1992] Furuse, Osamu and Hitoshi Iida. 1992. Cooperation between Transfer and Analysis in Example-Based Framework. In *Proceedings of the Fifteenth International Conference on Computational Linguistics, COLING-92*, Nantes, pp.645–651.

[Hockett, 1961] Hockett, Charles F. 1961. Grammar for the Hearer. In *Proceedings of Symposia in Applied Mathematics*, XII, American Mathematical Society.

[Inagaki & Matsubara, 1995] Inagaki, Yasuyoshi and Shigeki Matsubara. 1995. Models for Incremental Interpretation of Natural Language. In *Proceedings of the Second International Symposium on Natural Language Processing, SNLP '95*, Bangkok, Thailand, pp.51–60.

[Kitano, 1990] Kitano, Hiroaki. 1990. Incremental Sentence Production with a Parallel Marker-Passing Algorithm. In *Proceedings of the Thirteenth International Conference on Computational Linguistics, COLING-90*, Helsinki, pp.217–222.

[Matsubara & Inagaki, 1996a] Matsubara, Shigeki and Yasuyoshi Inagaki. 1996. A Study on Features and Phenomena of Japanese Spoken Language in Simultaneous Interpretation. In *Proceedings of 1996 Information and Systems Society Conference of the Institute of Electronics, Information and Communication Engineers*, Kanazawa, Japan, p.55. (In Japanese).

[Matsubara et al., 1996b] Matsubara, Shigeki, Yasuyoshi Inagaki and Katsuhiko Toyama. 1996. Soundness of The Principle of Incremental Compositionality. In *Proceedings of the Tenth Annual Conference of JSAI '96*, Tokyo, Japan, pp.335–338. (In Japanese).

[Menzel, 1994] Menzel, Wolfgang. 1994. Parsing of Spoken Language under Time Constraints. In *Proceedings of the Eleventh European Conference on Artificial Intelligence, ECAI '94*, Amsterdam, Netherlands, pp.560–564.

[Milward & Cooper, 1994] Milward, David and Robin Cooper. 1994. Incremental Interpretation: Applications, Theory, and Relationship to Dynamic Semantics. In *COLING-94: The 15th International Conference on Computational Linguistics*, Kyoto, Japan, pp.748–754.

[Pereira & Pollack, 1991] Pereira, Fernando, C.N. and Martha E. Pollack. 1991. Incremental Interpretation. *Artificial Intelligence* 50:37–82.

[Sagawa et al., 1994] Sagawa, Yuji, Noboru Ohnishi and Noboru Sugie. 1994. A Parser Coping with Self-Repaired Japanese Utterances and Large Corpus-based Evaluation. In *COLING '94: the 15th International Conference on Computational Linguistics*, Kyoto, Japan, pp.593–597.

[Sumita et al., 1990] Sumita, Eiji, Hitoshi Iida and Hideo Kohyama. 1990. Example-Based Approach in Machine Translation. In *The Third International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Language*, Austin, Texas, pp.203–212.

[Takezawa et al., 1995] Takezawa, Toshiyuki. Toshihisa Tashiro and Tsuyoshi Morimoto. 1995. A Study on Linguistic Phenomena of Spontaneous Speech in ATR Spokenf Language Database. *Research report SIG-SLUD-9403*, Japanese Society for Artificial Intelligence, Tokyo, Japan, pp.13–20. (In Japanese).

[Wiren, 1989] Wiren, Mats. 1989. Interactive Incremental Chart Parsing. In *Fourth Conference of the European Chapter of the Association for Computational Linguistics, EACL '89*, Manchester, England, pp.241–248.