Improvement of Translation Accuracy for the Outlines of Japanese Statutes by Splitting Parenthesized Expressions

OKADA Kouhei[1], OGAWA Yasuhiro[1,2,3], NAKAMURA Makoto[3,4],
OHNO Tomohiro[1,2], TOYAMA Katsuhiko[1,2,3]

[1] *Graduate School of Information Science*
[2] *Information Technology Center*
[3] *Japan Legal Information Institute*
[4] *Graduate School of Law*

Abstract. To globally share Japanese legal information, we translate the Outlines of Japanese statutes. These outlines are the official summaries of Japanese statutes and are useful to quickly understand their contents. In a previous statistical machine translation system for the outlines, we found that the training corpus consisted of both statutes and their outlines, including many long sentences that reduced the translation quality. To solve this problem, we shortened the length of sentences and focused on parenthesized expressions. In this paper, we propose a translation method that splits off parenthesized expressions from the sentences. Experimental result shows the effectiveness of our method.

Keywords: Statistical Machine Translation, The Outlines of Japanese Statutes, Legal Informatics

1. Introduction

Recently, the demand for providing Japanese legal information globally has increased for various reasons, such as promoting investment into Japan and facilitating international transactions. To respond to requests for legal information, the Japanese Ministry of Justice released the Japanese Law Translation Database System (JLT) in 2009 (Toyama el al., 2011), which provides Japanese legal information in English. However, JLT is not sufficient for the required amount and speed of translation due to the difficulty of legal translation.

To overcome this problem, Inagi et al. proposed to translate the Outlines of Japanese statutes (acts, cabinet orders, treaties) in place of the original statutes (Inagi el al., 2013). The Outlines, which are official summaries of Japanese statutes, are useful for easily and quickly understanding their contents. Since they are shorter and simpler than the original statutes, they are suitable for machine translation. The previous statistical machine translation (SMT) system designed for the outlines used sentences from both the statutes and the outlines (statutory sentences) to train its translation model. However, these

statutory sentences, most of which are long, reduced the translation quality.

After considering how to shorten the sentences, we focus on the parenthesized expressions in them. Parenthesized expressions are frequently used in statutory sentences to show a definition, an explanation, a supplementation, etc. They are usually independent of the original sentence, meaning the original sentence retains its own grammar even if the parenthesized expressions are removed from it. Thus, we split off such parenthesized expressions from a statutory sentence to shorten it and then translate them separately. In this paper, we propose an SMT method based on splitting statutory sentences and describe how to change the inputs and the training models.

Kawachi et al. are launching a website that disseminates daily news in English on Japanese legislation and embodies three concepts: *transparency*, *internationality*, and *promptness* (Kawachi et al., 2015). We plan to open the translation results of our system in this site so that our translation system will contribute to achieving the promptness. Therefore, our study promotes and supports the free access to law movement (Greenleaf, 2008).

This paper is organized as follows. In the next section, we introduce the Outlines of Japanese statutes. In Section 3, we describe the previous SMT system for the outlines. After explaining the classification of parenthesized expressions in Section 4, we propose an improved SMT method in Section 5. Our translation experiments are shown in Section 6. Finally, Section 7 summarizes this paper.

2. The Outlines of Japanese Statutes

The Japanese government has published the Outlines of Japanese statutes since 1973 for the general public's understanding of legislation (Ogawa et al., 2013). They are simultaneously published in official gazettes along with the promulgation of their original statutes.

Figure 1 is a translation example of an act's[1] outline, and Figure 2 is a translation of the original act by JLT (Inagi el al., 2013). In these figures, the upper broken-line boxes correspond to each other, while the lower broken-line boxes correspond to each other. Although Article 3 of the original act consists of six paragraphs, the outline contains only two sentences: The first corresponds to paragraphs (1) and (2) of the original, and the second corresponds to paragraphs (3) and (4). Paragraphs (5) and (6) of the original are omitted in the outline. Since

---

[1] An act is a statute enacted by the Diet.

---

3 Inspection

(i) If there are reasonable grounds to believe that a Ship in Japan's inland waters or territorial waters or the high seas is loaded with North Korea Specified Cargo, the Commandant of the Japan Coast Guard may, with the consent of the Captain, etc. of the ship if it is in Japan's territorial waters or the high seas, have a coast guard officer(s) take such measures as an inspection of the relevant cargo, documents and/or other articles. (The point of Article 3, Paragraphs 1 and 2)

(ii) If there are reasonable grounds to believe that a Ship, etc. located at a port in Japan is loaded with North Korea Specified Cargo or that any North Korea Specified Cargo is included in the cargo placed in a bonded area, the Director-General of the relevant Custom-House may have a customs official(s) take such measures as an inspection of the relevant cargo, documents and/or other articles. (The point of Article 3, Paragraphs 3 and 4)
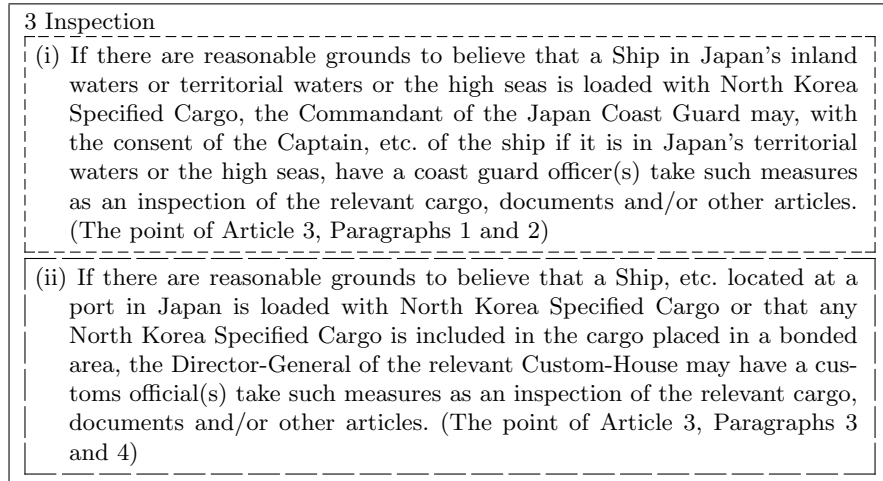

Figure 1. Translation of the Outline of Article 3 of Act on Special Measures concerning Cargo Inspections etc. Conducted by the Government Taking into Consideration United Nations Security Council Resolution 1874, etc. (Act No. 43 of 2010)

the outlines are shorter and simpler than their original acts, we can easily understand the content of the acts. From a translation viewpoint, since we expect lower translation costs, we can provide translations soon after the acts are promulgated.

The outlines are more than just simple summaries. For example, an amendment statute consists of amendment sentences that indicate how to revise the original statutory sentences, as shown in Figure 3. When we read the amendment statute, we cannot understand the sentence's meaning without referring to the statute that is being amended. By contrast, since the outlines provide an understanding of the substantive contents of the amendment, as shown in Figure 4, we can smoothly understand it. Therefore, the outlines are useful for understanding the changes made in statutes.

3. Translation of Outlines

3.1. STATISTICAL MACHINE TRANSLATION

SMT (Koehn, 2010), which is a common approach to automated translation systems, trains statistical models from a large amount of text data called a corpus. All we need is a training corpus to make a translation system at low cost in a short time. There are two kinds of models: a language model and a translation model. A language model, which is trained on a monolingual corpus, gives the probability of a sequence

(Inspection)

Article 3 (1) The Commandant of the Japan Coast Guard may order Coast Guard Officers to undertake the following measures, if there are reasonable grounds to believe that a vessel in the internal waters of Japan is carrying Specified North Korea Cargo:

   (i) to stop the vessel for inspection;

   (ii) to enter the vessel, inspect the cargo, documents, and other items, or question the crew and other relevant persons of the vessel;

   (iii) to collect samples, limited to the minimum amount necessary for inspection;

   (iv) to unload or reload the cargo or to instruct the captain of the vessel, etc. to unload or reload the cargo, within the limit necessary for inspection.

(2) The Commandant of the Japan Coast Guard may order Coast Guard Officers to undertake the following measures, if there are reasonable grounds to believe that a vessel in the territorial waters of Japan or on the high seas (including the Exclusive Economic Zone set forth in the United Nations Convention on the Law of the Sea; the same shall apply hereinafter) is carrying Specified North Korea Cargo:

   (i) to request the captain of the vessel, etc. to stop the vessel for inspection;

   (ii) to take the measures listed in item (ii) or item (iii) of the preceding paragraph, with the consent of the captain of the vessel, etc. ;

   (iii) to unload or reload the cargo with the consent of the captain of the vessel, etc., or to request the captain of the vessel, etc. to unload or reload the cargo, within the limit necessary for inspection.

(3) The Director-General of Customs may order Coast Guard Officers to undertake the following measures, if there are reasonable grounds to believe that a vessel or an aircraft (excluding military aircraft and aircraft owned or operated by a foreign government for non-commercial purposes; the same shall apply hereinafter) is carrying Specified North Korea Cargo at a port or an airport in Japan:

   (i) to enter the vessel or aircraft and inspect cargo, documents, and other items or question the crew and other relevant persons;

   (ii) to collect samples, limited to the minimum amount necessary for inspection;

   (iii) to unload or reload the cargo, or to order the captain of the vessel, etc., the captain of the aircraft, or the person acting on behalf of the captain (referred to in paragraph (2) of the following Article as the "captain of the aircraft, etc.") to unload or reload the cargo, within the limit necessary for inspection.

(4) The Director-General of Customs may order Coast Guard Officers to undertake an inspection of cargo, documents, and other items, to question the owner, possessor, manager, or other relevant persons, or to collect samples, limited to the minimum amount necessary for inspection, if there are reasonable grounds to believe that the cargo placed in a bonded area (meaning an area as prescribed in Article 29 of the Customs Act, including those designated by the Director-General of Customs pursuant to the provisions of Article 30, paragraph (1), item (ii) of the Customs Act; the same shall apply in paragraph (2) of the following Article).

(5) Pursuant to Ordinance of the Ministry of Land, Infrastructure, Transport and Tourism and Ordinance of the Ministry of Finance, Japan Coast Guard and Customs officials shall wear uniforms and carry identification cards when conducting the inspections prescribed in each of the preceding paragraphs.

(6) The authority pursuant to the provisions of paragraphs (1) through paragraph (4) may not be construed as having been granted for a criminal investigation.

Figure 2. Original of Article 3 of Act on Special Measures concerning Cargo Inspections etc. Conducted by the Government Taking into Consideration United Nations Security Council Resolution 1874, etc. (Act No. 43 of 2010)

> Part of the Act on Limitation on Number of Court Officials shall be revised as follows:
> The term "1,717" in the table of Article 1 shall be revised to "1,782," and "1,020" to "1,000.'

Figure 3. Article 1 of Partial Amendment Act for Act on Limitation on Number of Court Officials (Act No. 11 of 2010)

> Partial Amendment Act for Act on Limitation on Number of Court Officials (Act No. 11) (Ministry of Justice)
> 1 The number of judges shall be increased by 65. (The point of Article 1)
> 2 The number of assistant judges shall be reduced by 20. (The point of Article 1)

Figure 4. The Outline of Article 1 of Partial Amendment Act for Act on Limitation on Number of Court Officials (Act No. 11 of 2010)

of words in the target language. A translation model, which is trained on a parallel corpus, indicates how likely it is that the words in the target language are translations of words in the source language. Such words, which share a translation relation called word alignment, are estimated from parallel sentences during training to use in statistical models. The accuracy of word alignment is affected by the length of the parallel sentences: the longer the sentences, the more translation candidates, that is, lower alignment accuracy.

A decoder is translation machinery that performs an actual translation and finds the most probable sequence from the two models as a translation output. Both translation and alignment accuracy are affected by sentence length: the longer the inputs, the lower the translation quality. Accordingly, shorter sentences are desirable both in training and as inputs.

3.2. Baseline System for Outlines

In SMT, we can construct a translation system that specializes in a particular field for use as an in-domain corpus, since an outline corpus is usually used for translating the outlines. However, we cannot prepare a sufficiently large corpus, since only 1,742 translation sentences of the outlines are available, which are too few for training. We additionally used the translations of 166,977 Japanese statutes provided by JLT for the outlines, since Japanese statutes and their outlines share technical terms or phrases in the legal domain. Therefore, we compiled a training corpus that consists of 168,719 statutory sentences (1,742 + 166,977 = 168,719).

The current SMT system automatically changes the weights for the translation and language models to improve the translation accuracy (Och, 2003). This calculation is called parameter tuning. A small set of parallel sentences called development data and an automatic evaluation metric are used for determining the weights. Then we use part of the outline corpus as development data.

This system, however, has a problem: Most of the statutory sentences in the training corpus and the inputs are long, which reduces the translation accuracy, as mentioned in Section 3.1. To improve translation accuracy, we propose a new method that shortens sentences by focusing on the parenthesized expressions in statutory sentences.

4. Classification of Parenthesized Expressions

In this section, we investigate the use of parenthesized expressions in statutory sentences to consider the classes of parenthesized expressions processed in Section 5. The uses of parenthesized expressions in Japanese statutes have been investigated by (GYOSEI, 2013), and we confirmed these uses and found some additional ones in outlines.

4.1. Parenthesized Expressions in Japanese Statutes

There are two types of parenthesis characters in Japanese statutes: Japanese-style quotation marks[2] and round parentheses. In this paper, we focus on the latter.

Round parentheses in Japanese statutes are used in the following cases:

1. To indicate the range of articles included in a chapter or section in the table of contents.
2. To indicate the statute number of a statute cited in the provision.
3. To append a caption to an article.
4. To define the abbreviation of a term or phrase before the parentheses (see Example 1).
5. To define a term before the parentheses (see Example 2).
6. To complement conditions on the range of a term before parentheses, such as excluding or including something or limiting its range (see Example 3).
7. To replace a term or phrase before the parentheses with another one in certain cases (see Example 4).

---

[2] Japanese-style quotation marks correspond to double-quotation marks in English.

8. To append the summary of a quoted clause.
9. To indicate, in a table or a form, a related provision.
10. To show item numbers like "(1), (2), (3)" written horizontally[3].
11. To show sub-sub-item numbers.

We show some statutory sentences as examples below.

Example 1 The registered training institution must establish regulations concerning training (hereinafter referred to as "Training Regulations") and provide notification to the Minister of Land, Infrastructure, Transport and Tourism before the commencement of training.
Source: Article 26-10 Paragraph (1) of Construction Business Act (Act No. 100 of 1949)

Example 2 At a manufacturing facility, storage facility or handling facility, a person other than a hazardous materials engineer (meaning a person who has obtained a hazardous materials engineer's license; the same shall apply hereinafter) shall not handle hazardous materials without the attendance of a Class A hazardous materials engineer or Class B hazardous materials engineer.
Source: Article 13 Paragraph (3) of Fire Service Act (Act No. 186 of 1948)

Example 3 Compilations (not falling within the term "databases"; the same shall apply hereinafter) which, by reason of the selection or arrangement of their contents, constitute intellectual creations shall be protected as independent works.
Source: Article 12 Paragraph (1) of Copyright Act (Act No. 48 of 1970)

Example 4 Dismissal of Officers at Incorporation shall be determined by a majority of the votes of the incorporators (or by a majority of two thirds or more in case of dismissal of Auditors at Incorporation).
Source: Article 43 Paragraph (1) of Companies Act (Act No. 86 of 2005)

4.2. Classification of Parenthesized Expressions in Outlines

We investigated parenthesized expressions in outlines and found that the usages are the same as those of statutory sentences described in the previous subsection, except for three other usages in round parentheses as follows:

---

[3] Japanese statutes are usually written vertically in official gazettes; they are rarely written horizontally.

Table I. Parenthesized Expressions in Statutory Sentences

Length	Sentences	(with Parentheses)	Ratio (%)
1-20	71,049	1,112	1.56
21-40	36,925	4,205	11.39
41-60	20,770	4,807	23.14
61-80	10,404	4,044	38.87
81-	15,501	10,402	67.11
Total	154,649	24,570	15.89

12. To indicate the statute number and the name of the ministry having jurisdiction in a statute title.

13. To append relevant provisions.

14. To show numbers in itemized form, such as "(i), (ii), (iii)" and "(a), (b), (c)" (see Example 5).

Example 5 Restrictions or prohibitions described pursuant to (i) may only be put until September 11, 2011.
Source: The Outline of Article 1 Paragraph (2) of Act on Special Provisions of Building Restrictions in Urban Areas Severely Damaged by the Great East Japan Earthquake (Act No. 34 of 2011)

5. Proposed Method

We propose an SMT method that splits off parenthesized expressions from a sentence and translates them. We investigated how many parenthesized expressions are in statutory sentences, and the results is shown in Table I. Parenthesized expressions are frequently used in statutory sentences, especially in long sentences, which causes sentences to become long.

Since our approach can make sentences shorter, we expect the translation accuracy to improve. Note that this approach is effective not only for translation of sentences with parentheses but also for translation of sentences without parentheses, since we apply it to training a corpus as well as inputs.

An overview of this method is illustrated in Figure 5. The baseline system for outlines is described in (Inagi el al., 2013). We added a
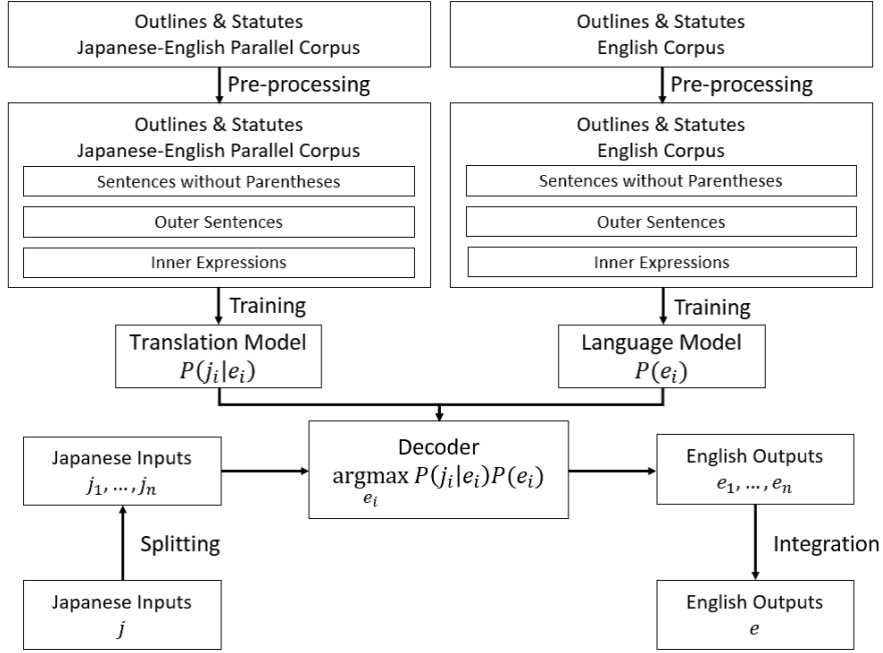
Figure 5. Overview of Proposed Method

pre-processing stage to the baseline system. This pre-processing stage consists of four steps:

1. splitting sentences

2. tokenization

3. lowercasing letters in English sentences

4. cleaning the training corpus for translation models

The details of splitting sentences in the training corpus are given in Section 5.2, and the details of splitting the inputs and integrating the outputs are presented in Section 5.3.

5.1. TARGET EXPRESSIONS

To shorten the statutory sentences, we split off the parenthesized expressions from the sentence when they are independent of each other, that is, when the sentence remains grammatically accurate even if the parenthesized expressions are removed. For the Japanese statutory sentences described in Section 3, the parenthesized expressions in usages 1. to 9., 12., and 13. are independent, so we split them off. On the other

hand, since the parenthesized expressions in usages 10., 11., 14., and Japanese-style quotation marks are not independent, we retain them. Hereafter, we focus on the former parenthesized expressions and call them "inner expressions" and the sentences in which inner expressions are removed are called "outer sentences." We can determine automatically whether parenthesized expressions are independent or not: If a parenthesized expression only consists of numerals or alphabets, it is certainly not independent.

If a statutory sentence has multiple inner expressions, we handle them as follows:

- We split all inner expressions off from the sentence.

- When the separated inner expression has another inner expression, that is, an original sentence includes a nested inner expression, we recursively split off the nested one.

5.2. Splitting Sentences in Training Corpus

The training corpus is a Japanese-English parallel corpus or an English monolingual corpus. After splitting, we identify the translation relations between the Japanese and English inner expressions. Next we compile a new corpus that consists of inner expressions, outer sentences, and sentences that did not originally have parentheses (hereinafter referred as to "sentences without parentheses").

Table II shows an example of splitting a parallel sentence. If a Japanese sentence has only one inner expression, as shown in Table II, the translation relation between Japanese and English inner expressions is clear. When a Japanese sentence has multiple inner expressions, we need to decide translation relations between Japanese and English. In order to decide the relation, we translate an inner expression into English by the baseline system and choose the most similar English inner expressions with the translation by calculating the Position independent word Error Rate (PER) (Och, 2003).

5.3. Splitting Inputs and Integrating Outputs

We split the input into an outer sentence and inner expressions, and translate them separately. For the output, we need to integrate the translation of the outer sentence and the inner expressions. First, the translated inner expressions are put in parentheses. After that, we insert each of them into the translated outer sentence. The insertion point is just after the translation of the word that appears just before the inner expression in the original sentence. This is because inner

Table II. Example of Splitting Parallel Sentences

Splitting		Japanese-English Parallel Sentences
Before		登録講習実施機関は、講習に関する規程（以下「講習規程」という。）を定め、講習の開始前に、国土交通大臣に届け出なければならない。
		The registered training institution must establish regulations concerning training **(hereinafter referred to as "Training Regulations")** and provide notification to the Minister of Land, Infrastructure, Transport and Tourism before the commencement of training.
After	Outer Sentence	登録講習実施機関は、講習に関する規程を定め、講習の開始前に、国土交通大臣に届け出なければならない。
		The registered training institution must establish regulations concerning training and provide notification to the Minister of Land, Infrastructure, Transport and Tourism before the commencement of training.
	Inner Expression	以下「講習規程」という。
		hereinafter referred to as "Training Regulations"

expressions often explain a phrase just before them, as described in Section 4.

6. Translation Experiments

In this section, we conduct translation experiments to verify the effectiveness of our proposed method. We compared two methods: the original SMT without splitting as a baseline and one with splitting as our proposed method.

6.1. EXPERIMENTAL PROCEDURE

We prepared a Japanese-English parallel corpus lowercased and tokenized by morphological analyzers: MeCab (Kudo et al., 2004) for

Table III. Number of Sentences Used in Training Translation Model

	Method	Before Cleaning	After Cleaning	Deleted
Baseline		168,719	150,472	18,247
Proposed	Sentence without Parentheses	142,199	133,847	8,352
	Outer Sentences	26,520	20,682	5,838
	Inner Expressions	32,127	31,771	356
	Total	200,846	186,300	14,546

Japanese and the tokenizer included with Moses (Koehn et al., 2007) for English. We used SRILM (Stolcke, 2002) to train a language model and GIZA++ (Och et al., 2005) to train a translation model on the training corpus, which consists of 166,977 statutory sentences in the JLT corpus and 1,742 outline sentences published in 2010. We used "-ukndiscount -interpolate" as the smoothing parameter for SRILM and "-grow-diag-final-and -msdbidirectional-fe" as the GIZA++ parameters. Since GIZA++ cannot handle long sentences, we cleaned up sentences longer than 80 words in the training corpus. We also cleaned up sentences that are empty or violate the 9-1 sentence ratio limit of GIZA++. Table III shows the number of sentences in the training corpus after cleaning for each method. We set the "distortion-limit" parameter in Moses to -1 (unlimited), since word order is usually changed in Japanese-English translation. We also conducted parameter tuning (Och, 2003), where we used BLEU (Papineni et al., 2002) as an automatic metric and 300 sentences randomly selected from the outlines published in 2010 as development data.

We used Moses as a decoder and translated 1,371 sentences of 52 outlines published in 2011. To evaluate the outputs of the translation systems, we used the automatic metrics BLEU and RIBES (Isozaki et al., 2010) with two reference data sets, that is, we prepared two human translations for each outline.

6.2. EXPERIMENTAL RESULTS AND DISCUSSION

Table IV shows the scores of each method as the results. The BLEU and RIBES scores for the proposed method were significantly higher than those of the baseline ($p < 0.05$). The accuracy of integrating the translation of an outer sentence and its inner expressions is 60.38%.

Table IV. BLEU and RIBES Scores for Each Method

Method	BLEU	RIBES
Baseline	37.37	71.05
Proposed	38.87	72.26

Table V. RIBES Scores Divided by Sentences With or Without Parentheses

Method	With	Without
Baseline	53.14	70.49
Proposed	55.76	72.05

We divided the outputs into 120 sentences with parentheses and 1,251 without them and evaluated these sentences by RIBES, which can evaluate each sentence[4]. The results are shown in Table V. The RIBES scores of the sentences with parentheses using the proposed method were 2.5 points higher than those using the baseline. Sentences without parentheses were 1.5 points higher. These results show that the translation accuracy improved not only with parentheses but also without them. The rise in the scores of sentences without parentheses is caused by the increased training data; fewer sentences were deleted in the cleaning up step since they are shorter than the original ones. In addition, shortening the sentences improved the alignment accuracy in the training step. These factors also improved the RIBES scores of the sentences with parentheses, and shortening the input sentences by splitting off the inner expressions improved the translation quality. The average lengths of input sentences in the baseline and in the proposed method are 34.64 and 31.04, respectively.

Table VI shows an example of the output sentences in this experiment. The inner expressions in each sentence are shown in bold. The baseline method translated the inner expressions into three parts, which are inaccurate translations, while the proposed method correctly translated the inner expressions. In this example, the RIBES scores by the baseline and the proposed methods are 52.98 and 67.04, respectively.

Based on the above results, we confirmed the effectiveness of our proposed method.

---

[4] BLEU is unsuitable for evaluating each sentence, so we only used RIBES.

Table VI. Example of Translations by Baseline and Proposed Methods

Japanese	地方 公共 団体 等 以外 の 者 は 、 港湾 運営 会社 の 株式 につい て 、 保有 基準 割合 （ 原則 として 総 株主 の 議決 権 の 一 〇 〇 分の 二 〇 ） 以上 の 数 の 議決 権 を 取得 し 、 又は 保有 し て は なら ない こと と した 。
Baseline	no person other than the holding ratio threshold **of the voting rights of all shareholders** or more **(20 / 100ths** , with respect to the shares of the company shall not acquire or hold a number of voting rights in the local government , etc. port operation **principle)** .
Proposed	local governments , etc. with regard to shares of the company shall not be acquired or held by a person other than the holding ratio **(100 20 percent of the voting rights of all shareholders in principle)** threshold of the voting rights of the port operation .
Reference 1	a person other than local governments or the like shall not obtain or hold a number of voting rights for the shares of the port operation company that is at least equal to the holding ratio threshold **(20 percent of the voting rights of all shareholders , in principle)** .
Reference 2	no person except for local public entities , etc. may acquire or hold voting rights of the port and harbor management company equivalent to or exceeding the holding ratio threshold **(twenty percent of the voting rights of all shareholders in principle)** .

7. Conclusion

We proposed a translation method that split off parenthesized expressions from statutory sentences in a training corpus and used them as inputs. Our experimental results show the effectiveness of this method. In this paper, we applied our proposed method to outlines, but it can also be adapted to statutes, since it uses features shared by outlines and statutes.

Future work will involve harvesting the correct order of words in translations. Even though the translation quality of technical terms is good in our translation system, the word order is often incorrect due to word-order differences in Japanese and English. A novel approach, such as (Hoshino et al., 2013), converts Japanese word order into English

word order in advance. We plan to introduce this approach to our system.

Acknowledgements

References

Greenleaf, G. *Legal Information Institutes and the Free Access to Law Movement*, GlobaLex website, 2008.

GYOSEI Corporation Legislation Seminar, *Zusetsu Houseishitsumu Nyumon*, GYO-SEI Corporation, 2013 (in Japanese) .

Hoshino, S., Miyao, Y., Sudoh, K., and Nagata, M. *Two-Stage Pre-ordering for Japanese-to-English Statistical Machine Translation*, Proc. of IJCNLP 2013, pp. 1062–1066, 2013. Inagi, D., Ogawa, Y., Nakamura, M., Ohno, T., and Toyama, K. *Statistical Machine Translation for Outlines of Japanese Statutes*, Proc. of JURISIN 2013, pp. 37–49, 2013.

Inagi, D., Ogawa, Y., Nakamura, M., Ohno, T., and Toyama, K. *Statistical Machine Translation for Outlines of Japanese Statutes*, Proc. of JURISIN 2013, pp. 37–49, 2013.

Isozaki, H., Hirao, T., Duh, K., Sudoh, K., and Tsukada, H. *Automatic Evaluation of Translation Quality for Distant Language Pairs*, Proc. of EMNLP 2010, pp. 944–952, 2010.

Kawachi, G., Ogawa, Y., Nakamura, M., Ohno, T., and Toyama, K. *Daily News on Japanese Legislation toward International Sharing of Japanese Legal Information*, Journal of Open Access to Law, Vol. 3, No. 1, 19 pages, 2015.

Koehn, P. *Statistical Machine Translation*, Cambridge University Press, 2010.

Koehn, P., Hoang, H., Birch, A., Callison-Burch, C., Federico, M., Bertoldi, N., Cowan, B., Shen, W., Moran, C., Zens, R., Dyer, C., Bojar, O., Constanin, A., and Herbst, E. *Moses: Open Source Toolkit for Statistical Machine Translation*, Proc. of ACL2007, pp. 177–180, 2007.

Kudo, T., Yamamoto, K., and Matsumoto, Y. *Applying Conditional Random Fields to Japanese Morphological Analysis*, Proc. of EMNLP 2004, pp. 230–237, 2004.

Och, F. J. *Minimum Error Rate Training in Statistical Machine Translation*, Proc. of ACL 2003, pp. 160–167, 2003.

Och, F. J. and Ney, H. *A Systematic Comparison of Various Statistical Alignment Models*, Computational Linguistics, Vol. 29, No. 1, pp. 19–51, 2005.

Ogawa, Y., Inagi, D., Nakamura, M., and Toyama, K. *Translation for Outlines of Japanese Acts*, Law via the Internet 2013, 12 pages, 2013.

Papineni, K., Roukos, S., Ward, T., and Zhu, W. *BLEU: A Method for Automatic Evaluation of Machine Translation*, Proc. of ACL 2002, pp. 138–145, 2002.

Stolcke, A. *SRILM - An Extensible Language Modeling Toolkit*, Proc. of ICSLP 2002, pp. 901–904, 2002.

Toyama, K., Saito, D., Sekine, Y., Ogawa, Y., Kakuta, T., Kimura, T., and Matsuura, Y. *Design and Development of Japanese Law Translation Database System*, Law via the Internet 2011, 12 pages, 2011.