A Reordering Model Using a Source-Side Parse-Tree for Statistical Machine Translation

SUMMARY This paper presents a reordering model using a source-side parse-tree for phrase-based statistical machine translation. The proposed model is an extension of IST-ITG (imposing source tree on inversion transduction grammar) constraints. In the proposed method, the target-side word order is obtained by rotating nodes of the source-side parse-tree. We modeled the node rotation, monotone or swap, using word alignments based on a training parallel corpus and source-side parse-trees. The model efficiently suppresses erroneous target word orderings, especially global orderings. Furthermore, the proposed method conducts a probabilistic evaluation of target word reorderings. In English-to-Japanese and Japanese-to-Chinese translation experiments, the proposed method resulted in a 0.49-point improvement (29.31 to 29.80) and a 0.33-point improvement (18.60 to 18.93) in word BLEU-4 compared with IST-ITG constraints, respectively. This indicates the validity of the proposed reordering model.

key words: phrase-based statistical machine translation, reordering model, parse-tree, syntactic information

1. Introduction

Statistical machine translation has been widely applied in many state-of-the-art translation systems. A popular statistical machine translation paradigm is the phrase-based model [1], [2]. In phrase-based statistical machine translation, errors in word reordering, especially global reordering, are one of the most serious problems. To resolve this problem, many word-reordering constraint techniques have been proposed. These techniques are categorized into two types. The first type is linguistically syntax-based. In this approach, tree structures for the source [3], [4], target [5], [6], or both [7] are used for model training. The second type is formal constraints on word permutations. IBM constraints [8], the lexical word reordering model [9], and inversion transduction grammar (ITG) constraints [10], [11] belong to this type of approach. For ITG constraints, the target-side word order is obtained by rotating nodes of the source-side binary tree. In these node rotations, the source binary tree instance is not considered. Imposing a source tree on ITG (IST-ITG) constraints [12] is an extension of ITG constraints and a hybrid of the first and second type of approach. IST-ITG constraints directly introduce a source sentence tree structure. Therefore, IST-ITG can obtain stronger constraints for word reordering than the original ITG constraints. For example, IST-ITG constraints allow only eight word orderings for a four-word sentence, even though twenty-two word orderings are possible with respect to the original ITG constraints. Although IST-ITG constraints efficiently suppress erroneous target word orderings, the method cannot assign the probability to the target word orderings.

This paper presents a reordering model using a source-side parse-tree for phrase-based statistical machine translation. The proposed reordering model is an extension of IST-ITG constraints. In the proposed method, the target-side word order is obtained by rotating nodes of a source-side parse-tree in a similar fashion to IST-ITG constraints. We modeled the rotating positions, monotone or swap, from word alignments of a training parallel corpus and source-side parse-trees. The proposed method conducts a probabilistic evaluation of target word orderings using the source-side parse-tree.

The rest of this paper is organized as follows. Section 2 describes the previous approach to resolving erroneous word reordering. In Sect. 3, the reordering model using a source-side parse-tree is presented. Section 4 shows experimental results. Finally, Sect. 5 presents the summary and some concluding remarks and future works.

2. Previous Works

First, we introduce two previous studies on related word reordering constraints, ITG and IST-ITG constraints.

2.1 ITG Constraints

In one-to-one word-alignment, the source word $f_i$ is translated into the target word $e_i$. The source sentence $[f_1, f_2, \ldots, f_N]$ is translated into the target sentence which is the reordered target word sequence $[e_1, e_2, \ldots, e_N]$. Then, the number of reorderings is $N!$. 
Stochastic synchronous grammars provide a generative process to produce a sentence and its translation simultaneously. An inversion transduction grammar (ITG) [10], [11] is a well-studied synchronous grammar formalism. To allow for movement during translation, non-terminal productions can be either straight (monotone) or inverted. Straight productions are output in the given order in both sentences. Inverted productions are output in the reverse order in the foreign sentence only. ITG cannot represent all possible permutations of concepts that many occur during translation, because some permutations will require discontinuous constituents. When these ITG constraints are introduced, the number of reorderings \( N! \) can be reduced in accordance with the following constraints.

- All possible source-side binary tree structures are generated from the source word sequence.
- The target sentence is obtained by rotating any node of the generated source-side binary trees.

When \( N = 4 \), the ITG constraints can reduce the number of reorderings from \( 4! = 24 \) to 22 by rejecting the orders \([e_3, e_1, e_4, e_2]\) and \([e_2, e_4, e_1, e_3]\) that cannot be represented by ITG. Such target word orders are called inside-out alignments [11]. For a four-word sentence, the search space is reduced to 92% (22/24), but for a 10-word sentence, the search space is only 6% (206,098/3,628,800) of the original full space.

### 2.2 IST-ITG Constraints

In ITG constraints, the source-side binary tree instance is not considered. Therefore, if a source sentence tree structure is utilized, stronger constraints than the original ITG constraints can be created. IST-ITG constraints [12] directly introduce a source sentence tree structure. The target sentence is obtained with the following constraints.

- A source sentence tree structure is generated from the source sentence.
- The target sentence is obtained by rotating any node of the source sentence tree structure.

By parsing the source sentence, the source-side parse-tree is obtained. After parsing the source sentence, a bracketed sentence is obtained by removing the node syntactic labels; this bracketed sentence can then be converted into a tree structure. For example, the source-side parse-tree “(S1 (S (NP (DT This)) (VP (AUX is) (NP (DT a) (NN pen))))))” is obtained from the source sentence “This is a pen” which consists of four words. By removing the node syntactic labels, the bracketed sentence “((This (is) (a) (pen)))” is obtained. Such a bracketed sentence can be used to produce constraints. If IST-ITG constraints are applied, the number of target word orders in \( N = 4 \) is reduced to 8, down from 22 with ITG constraints. For example, for the source-side bracketed tree “((f_1 f_2) (f_3 f_4)),” the eight target sequences \([e_1, e_2, e_3, e_4, e_2, e_1, e_3, e_4]\), \([e_1, e_2, e_4, e_3]\), \([e_2, e_1, e_4, e_3]\), \([e_3, e_4, e_1, e_2]\), \([e_3, e_4, e_2, e_1]\), \([e_4, e_3, e_1, e_2]\), and \([e_4, e_3, e_2, e_1]\) are accepted. For the source-side bracketed tree “(((f_1 f_2) f_3) f_4),” the eight sequences \([e_1, e_2, e_3, e_4]\), \([e_2, e_1, e_3, e_4]\), \([e_3, e_2, e_1, e_4]\), \([e_2, e_1, e_3, e_4]\), \([e_3, e_2, e_1, e_4]\), \([e_2, e_1, e_3, e_4]\), \([e_3, e_2, e_1, e_4]\), and \([e_4, e_3, e_2, e_1]\) are accepted. When the source sentence tree structure is a binary tree, the number of word orderings is reduced to \( 2^{N-1} \). However, the parsing results sometimes do not produce binary trees. In this case, some subtrees have more than two child nodes. For a non-binary subtree, any reordering of child nodes is allowed. If a subtree has three child nodes, six reorderings of the nodes are accepted.

In phrase-based statistical machine translation, a source “phrase” is translated into a target “phrase.” However, with IST-ITG constraints, “word” must be used for the constraint unit since the parse unit is a “word.” To absorb different units between translation models and IST-ITG constraints, a new limitation for word reordering is applied.

- Word ordering that destroys a phrase is not allowed.

When this limitation is applied, the translated word ordering is obtained from the bracketed source sentence tree by reordering the nodes in the tree, which is the same as for one-to-one word-alignment.

### 3. Reordering Model Using the Source-Side Parse-Tree

In this section, we present a new reordering model using syntactic information of a source-side parse-tree.

#### 3.1 Abstract of Proposed Method

The IST-ITG constraints method efficiently suppresses erroneous target word orderings. However, IST-ITG constraints cannot evaluate the accuracy of the target word orderings; i.e., IST-ITG constraints assign an equal probability to all target word orderings. This paper proposes a reordering model using the source-side parse-tree as an extension of IST-ITG constraints. The proposed reordering model conducts a probabilistic evaluation of target word orderings using syntactic information of the source-side parse-tree.

In the proposed method, the target-side word order is obtained by rotating nodes of the source-side parse-tree in a similar fashion to IST-ITG constraints. Reordering probabilities are assigned to each subtree of source-side parse-tree S by reordering the positions into two types: monotone (straight) and swap. If the subtree has more than two child nodes, the number of child node order is more than two. However, we assume the child node order other than monotone to be swap.

The source-side parse-tree S consists of subtrees \( s_1, s_2, \cdots, s_k \), where \( K \) is the number of subtrees included in the source-side parse-tree. The subtree \( s_k \) is represented by the parent node’s syntactic label and the order, from sentence head to sentence tail, of the child node’s syntactic labels. For example, Fig. 1 shows a source-side parse-tree for a four-word source sentence consisting of three subtrees. In Fig. 1, the subtrees \( s_1, s_2, s_3 \) are represented...
by S+NP+VP, VP+AUX+NP, and NP+DT+NN, respectively. Each subtree has a probability \( P(t \mid s) \), where \( t \) is monotone (m) or swap (s). The probability of the target word reordering is calculated as follows.

\[
P_r = \prod_{k=1}^{\kappa} P(t \mid s_k)
\]

By Eq. (1), each target candidate is assigned the different reordering probability. The proposed reordering probabilities of higher-level subtrees are effective for global word reordering, and ones of lower-level subtrees are effective for local word reordering.

### 3.2 Training of the Proposed Model

We modeled monotone or swap node rotating automatically from word alignments of a training parallel corpus and source-side parse-trees. The training algorithm for the proposed reordering model is as follows.

1. The training process begins with a word-aligned corpus. We obtained the word alignments using Koehn et al.’s method (2003), which is based on Och and Ney’s work (2004). This involves running GIZA++ [13] on the corpus in both directions, and applying refinement rules (the variant they designate is “final-and”) to obtain a single many-to-many word alignment for each sentence.

2. Source-side parse-trees are created using a source language phrase structure parser, which annotates each node with a syntactic label. A source-side parse-tree consists of several subtrees with syntactic labels. For example, the parse-tree “(S1 (S (NP (DT This)) (VP (AUX is) (NP (DT a) (NN pen))))” is obtained from the source sentence “This is a pen” which consists of four words.

3. Word alignments and source-side parse-trees are combined. Leaf nodes are assigned target word positions obtained from word alignments. Via the bottom-up process, target word positions are assigned to all nodes. For example, in Fig. 2, the left-side (sentence head) child node of subtree \( s_2 \) is assigned the target word position “4,” and the right-side (sentence tail) child node is assigned the target word positions “2” and “3,” which are assigned to the child nodes of subtree \( s_1 \).

4. The monotone and swap reordering positions are checked and counted for each subtree. By comparing the target word positions, which are assigned in the above step, the reordering position is determined. If the target word position of the left-side child node is smaller than one of the right-side child node, the reordering position determined as monotone. For example, in Fig. 2, the subtrees \( s_1, s_2 \) and \( s_3 \) are monotone, swap, and monotone, respectively.

5. The reordering probability of the subtree can be directly estimated by counting the reordering positions in the training data.

\[
P(t \mid s) = \frac{c_t(s)}{\sum_t c_t(s)}
\]

where \( c_t(s) \) is the count of reordering position \( t \) included all training samples for the subtree \( s \).

The parsing results sometimes do not produce binary trees. For a non-binary subtree, any reordering of child nodes is allowed. However, the proposed reordering model assumes that reordering positions are only two, monotone and swap. That is, the reordering position which the order of child nodes do not change is monotone, and the other positions are swap. Therefore, the probability of swap \( P(s \mid s_k) \) is derived from the probability of monotone \( P(m \mid s_k) \) as follows.

\[
P(s \mid s_k) = 1.0 - P(m \mid s_k)
\]

Table 1 shows the example of proposed reordering models.

If a subtree is represented by a binary-tree, there are \( L^3 \) possible subtrees, where \( L \) is the number of syntactic labels. However, in the possible subtrees, there are subtrees observed only a few times in training sentences, especially
some problematic cases. In Fig. 3, the subtree 

$$s$$

sons, errors of automatic word alignments, syntactic anal-

verb, modal and so on. Others are due to non-linguistic rea-

such as negation (French “ne...pas” and English “not”), ad-

In the accepted candidate, the reordering positions for

all subtrees included the source side parse-tree are

checked by comparing the source phrase $$f$$ with the

source phrase sequence used before.

Subtrees checked reordering positions are assigned a

probability–monotone or swap–by the proposed reordering

model, and the target word order is evaluated by Eq (1).

Phrase-based statistical machine translation uses a

“phrase” as the translation unit. However, the proposed

reordering model needs a “word” order. Because “word”

alignments from the source phrase to target phrase are not

clear, we cannot determine the reordering position of sub-

tree included in a phrase. Therefore, in the decoding pro-

cess using the proposed reordering model, we define that

higher probability, monotone or swap, are assigned to sub-

trees included in a source phrase. For example, in Fig. 4,

the source sentence $$[f_1, f_2, f_3, f_4]$$ is translated into the tar-

get sentence $$[e_1, e_2, e_4, e_3]$$, where $$[f_1, f_2]$$ and $$[e_1, e_2]$$ are

used as phrases. Then, the source phrase $$[f_1, f_2]$$ includes

the subtree $$s_2$$. If the monotone probabilities of subtrees $$s_1$$,

$$s_2$$, and $$s_3$$ are 0.8, 0.4 and 0.7, the proposed reordering prob-

ability is $$0.8 \times 0.6 \times 0.3 = 0.144$$. If a source phrase is

$$[f_1, f_2, f_3, f_4]$$ and a source-side parse-tree has the same tree

structure used in Fig. 4, the subtrees $$s_1$$, $$s_2$$, and $$s_3$$ are

assigned higher reordering probabilities. If the source phrase

$$[f_1, f_2, f_3, f_4]$$ used in Fig. 4, the subtrees $$s_1$$, $$s_2$$, and $$s_3$$ are

assigned higher reordering probabilities.

Non-binary subtrees are often observed in the source-

side parse-tree. When a source phrase $$f$$ is included in a non-

binary subtree and does not include a non-binary subtree, we

cannot determine the reordering position. For example, the

phrase is selected to extend a target candidate. The checking

algorithm is as follows.

1. For old translation candidates, the subtree $$s$$, which in-

cludes both translated and untranslated words, and its

untranslated part $$u$$ are calculated.

2. When a new target phrase $$\bar{e}$$ is generated, the source

phrase $$\bar{f}$$ and the untranslated part $$u$$ calculated in the

above step are compared. If the source phrase $$\bar{f}$$ does

not include the untranslated part $$u$$ and is not included

$$u$$, the new candidate is rejected.

3. In the accepted candidate, the reordering positions for

all subtrees included the source side parse-tree are

checked by comparing the source phrase $$\bar{f}$$ with the

source phrase sequence used before.

when the subtree consists of more than three child nodes.

Although a large number of subtree models can capture var-

iations in the training samples, too many models lead to the

over-fitting problem. Therefore, subtrees where the number

of training samples is less than a heuristic threshold and un-

seen subtrees are clustered to deal with the data sparseness

problem for robust model estimations.

After creating word alignments of a training parallel

corpus, there are target word orders which are not derived

from rotating nodes of source-side parse-trees. Figure 3

shows a sample which is not derived from rotating nodes.

Some are due to linguistic reasons, structural differences

such as negation (French “ne...pas” and English “not”), ad-

verb, modal and so on. Others are due to non-linguistic rea-

sons, errors of automatic word alignments, syntactic anal-

ysis, or human translation [14]. The proposed method discards

such problematic cases. In Fig. 3, the subtree $$s_1$$ is then

removed from training samples, and the subtrees $$s_2$$ and $$s_3$$

are used as training samples.

### 3.3 Decoding Using the Proposed Reordering Model

In this section, we describe a one-pass phrase-based decoding

algorithm that uses the proposed reordering model in the
decoder. The translation target sentence is sequentially gen-
erated from left (sentence head) to right (sentence tail), and

all reordering is conducted on the source side. To introduce

the proposed reordering model into the decoder, the target

candidate must be checked for whether the reordering posi-
tion of a subtree is either monotone or swap whenever a new

| Subtree type       | Monotone probability |
|--------------------|----------------------|
| S+PP+NP+VP+NP+NP   | 0.764                |
| PP+NP              | 0.816                |
| NP+DT+NP+NN+NN     | 0.664                |
| VP+VBN+NP          | 0.837                |
| NP+NP              | 0.805                |
| NP+DT+JJ+NN        | 0.653                |
| NP+DT+JJ+VBP+NN    | 0.412                |
| NP+DT+NN+CC+VB     | 0.357                |

Fig. 3  Example of a target word order which is not derived from rotating

the nodes of source-side parse trees.

Fig. 4  Example of a target candidate including a phrase.
reordering position of subtree $s_2$ in Fig. 5, which includes the phrase $[f_3, f_4]$, can not be determined. In this case, we define that such subtrees are also to be assigned a higher probability.

4. Experiments

To evaluate the proposed model, we conducted two experiments: English-to-Japanese and English-to-Chinese translation.

4.1 English-to-Japanese Paper Abstract Translation Experiments

The first experiment was the English-to-Japanese (E-J) translation. Table 2 shows the training, development and test corpus statistics. JST Japanese-English paper abstract corpus consists of 1.0 M parallel sentences were used for model training. This corpus was constructed from 2.0 M Japanese-English paper abstract corpus belongs to JST [15] by NICT using the method of Uchiyama and Isahara [16]. For phrase-based translation model training, we used the GIZA++ toolkit [13], and 1.0 M bilingual sentences. For language model training, we used the SRI language model toolkit[17], and 1.0 M sentences for the translation model training. The language model type was word 5-gram smoothed by Kneser-Ney discounting [18]. To tune the decoder parameters, we conducted minimum error rate training [19] with respect to the four word BLEU score [20] using 2.0 K development sentence pairs. The test set with 2.0 K sentences is used. In the evaluation and development sets, a single reference was used. For the creation of English sentence parse trees and segmentation of the English, we used the Charniak parser [21]. We used Chasen [22] for segmentation of the Japanese sentences. We used Cleopatra made at ATR for the decoding, which is compatible with Moses [23]. The performance of this decoder was configured to be the same as Moses. Other conditions were the same as the default conditions of the Moses decoder.

In this experiment, the following three methods were compared.

- Baseline: The IBM constraints and the lexical reordering model were used for target word reordering.
- IST-ITG: The IST-ITG constraints, the IBM constraints, and the lexical reordering model were used for target word reordering.
- Proposed: The proposed reordering model, the IBM constraints, and the lexical reordering model were used for target word reordering.

During minimum error training, each method used each reordering model and reordering constraint.

The proposed reordering model are trained from 1.0 M bilingual sentences which are used for the translation model training. The amount of available training samples represented by subtrees was 9.8 M. In the available training samples, there were 54 K subtree types. The heuristic threshold was 10, and subtrees with training samples of less than 10 were clustered. The proposed reordering model consisted of 5,960 subtrees types and one clustered model. The models not including the clustered model covered 99.29% of all training samples.

The BLEU and WER are presented in Table 3. In comparing "Baseline" method with "IST-ITG" method, the improvement in BLEU was a 1.44-point and improvement in WER was 4.76%. Furthermore, in comparing "IST-ITG" method with "Proposed" method, the improvement in BLEU was a 0.49-point and improvement in WER was 0.65%. Table 4 shows the number of outputs that improved or got worse in BLEU score from "IST-ITG" for E-J translation.

Table 2: Statistics of training, development and test corpus for E-J translation.

|          | English | Japanese |
|----------|---------|----------|
| Train    | 1.0 M   |          |
| Dev      | 24.6 M  | 28.8 M   |
| Test     | 30.1 K  | 58.7 K   |

Table 3: BLEU score results for E-J translation. (1-reference)

| Method   | BLEU  | WER  |
|----------|-------|------|
| Baseline | 27.87 | 77.20|
| IST-ITG  | 29.31 | 72.44|
| Proposed | 29.80 | 71.79|

Table 4: The number of output that “Proposed” improved and got worse in BLEU score from “IST-ITG” for E-J translation.

|        | positive | negative | equal |
|--------|----------|----------|-------|
| # of outputs | 605      | 539      | 831   |
ever, when the source sentence consists many words and the source sentence structure is complex, the results using the proposed reordering model is better than one using the IST-ITG constraints. In this experiment, when the number of source words was more than 30, 45% of test sentences were improved by the proposed reordering model. Therefore, “Proposed” method resulted in a better BLEU and WER. The improvement could clearly be seen from visual inspection of the output, a few examples of which are presented in the Appendix.

4.2 NIST MT08 English-to-Chinese Translation Experiments

Next, we conducted English-to-Chinese (E-C) newspaper translation experiments for different language pairs. The NIST MT08 evaluation campaign English-to-Chinese translation task was used for the training and evaluation corpora. Table 5 shows the training, development and test corpus statistics. For the translation model training, we used 4.6 M bilingual sentences. For the language model training, we used 4.6 M sentences which are used for the translation model training. The language model type was word 3-gram smoothed by Kneser-Ney discounting. A development set with 1.6 K sentences was used as evaluation data in the Chinese-to-English translation track for the NIST MT07 evaluation campaign. A single reference was used in the development set. The evaluation set with 1.9 K sentences is the same as the MT08 evaluation data, with 4 references. In this experiment, the compared methods were the same as in the E-J experiment.

The proposed reordering model are trained from 4.6 M bilingual sentences which are used for the translation model training. The amount of available training samples represented by subtrees was 39.6 M. In the available training samples, there were 193 K subtree types. As in the E-J experiments, the heuristic threshold was 10. The proposed reordering model consisted of 18,955 subtree types and one clustered model. The models not including the clustered model covered 99.45% of all training samples.

The BLEU and WER are presented in Table 6. In comparing “Baseline” method with “IST-ITG” method, the proposed method resulted in a better BLEU and WER. The improvement could clearly be seen from visual inspection of the output, a few examples of which are presented in the Appendix.

| Table 5 | Statistics of training, development and test corpus for E-C translation. |
|---------|-----------------------------|
|         | Train Sentences | 4.6 M |
|         | Dev Sentences | 46.4 K |
|         | Test Sentences | 45.7 K |
|         | Words | 79.6 M |
|         | Words | 1.6 K |
|         | Words | 1.9 K |

| Table 6 | BLEU score results for E-C translation. (4-reference) |
|---------|-----------------------------|
|         | BLEU | Baseline | IST-ITG | Proposed |
| BLEU    | 17.54 | 18.60 | 18.93 |
| WER     | 78.07 | 75.43 | 75.57 |

improvement in BLEU was a 1.06-point. Furthermore, in comparing “IST-ITG” method with “Proposed” method, the improvement in BLEU was a 0.33-point. As in the E-J experiments, “Proposed” method performed the highest BLEU. Consequently, we demonstrated that the proposed method is effective for multiple language pairs. However, the improvement of BLEU and WER in E-C translation is smaller than the improvement in E-J translation. Table 7 shows the number of outputs that improved or got worse in BLEU after comparing “Proposed” method with “IST-ITG” method. These results cannot indicate a statistically significant difference at 95% confidence level between “Proposed” method and “IST-ITG” method. That is because English and Chinese are similar sentence tree structures, such as SVO-languages (Japanese is SOV-language). When the sentence tree structures are different, the proposed reordering model is effective.

5. Conclusion

This paper proposed a new word reordering model using a source-side parse-tree for phrase-based statistical machine translation. The proposed model is an extension of the IST-ITG constraints. In both IST-ITG constraints and the proposed method, the target-side word order is obtained by rotating nodes of the source-side tree structure. Both the IST-ITG constraints and the proposed reordering model fix the phrase position for the global reorderings. However, the proposed method can conduct a probabilistic evaluation of target word reorderings which the IST-ITG constraints cannot. In E-J and E-C translation experiments, the proposed method resulted in a 0.49-point improvement (29.31 to 29.80) and a 0.33-point improvement (18.60 to 18.93) in word BLEU-4 compared with IST-ITG constraints, respectively. This indicates the validity of the proposed reordering model.

Future work will focus on a simultaneous training of translation and reordering models. Moreover, we will deal with difference between source and target tree structures in multi level like in [24].

References

[1] P. Koehn, F.J. Och, and D. Marcu, “Statistical phrase-based translation.” Proc. HLT-NAACL 2003, pp.127–133, 2003.
[2] F.J. Och and H. Ney, “The alignment template approach to statistical machine translation,” Computational Linguistics, vol.30, no.4, pp.417–449, 2004.
[3] C. Quirk, A. Menezes, and C. Cherry, “Dependency treelet translation: Syntactically informed phrasal SMT,” Proc. ACL, pp.271–279, 2005.
[4] L. Huang, K. Knight, and A. Joshi, “Statistical syntax-directed translation with extended domain of locality,” Proc. AMTA, 2006.
A.2 Sentence 2

Source: From result of the consideration, it was pointed that radiation from the loop elements was weak.

Baseline: 考察の結果からすることを指摘し、ループ素子からの放射は弱かった。

Reference: 考察結果より、ループ素子からの放射が弱いことを指摘する。

Proposed: 考察の結果から、ことを指摘し、ループの要素からの放射は弱かった。

Proposed: 考察の結果から、ループ素子からの放射は弱いことを示した。

Kei Hashimoto was born in 1984. He received the B.E., and M.E. degrees in computer science, and computer science and Engineering from the Nagoya Institute of technology. Nagoya, Japan in 2004, and 2006, respectively. He is currently a Doctor’s candidate at the Nagoya Institute of technology. His research interests include statistical speech recognition, synthesis and machine translation. He is a student member of the Acoustical Society of Japan.

Hirofumi Yamamoto received the M.S. degree in agriculture from the University of Tokyo 1981 and the Ph.D. degree in global information and telecommunication from the Waseda University in 2004. He is currently a professor at Kinki University School of Science and Engineering Dept. Informatics, short-term researcher at National Institute of Communications Technology. His research interests include speech recognition and machine translation. He is a member of the IEEE, the ASJ and the ANLP.

Appendix: Samples from the English to Japanese Translation

A.1 Sentence 1

Source: The free-running period (τ) in constant darkness (DD) at 20°C became shorter than that at 25°C, suggesting that the phase advance of locomotor activity in LD cycles at 20°C was caused by the decrease in τ.

Baseline: 一定の暗間で20°Cで25°Cでのそれに比べ短くなっていることから、20°CでLD周期で歩行活動の位相前進（τ）のフリーランニング期（DD）の減少が原因であった。

Reference: 20°Cでの一定の暗環境（DD）における自由継続期（τ）は25°Cよりも短くなり、20°CでのLD周期における運動活動の位相前進がτの減少に起因することを示唆した。

IST-ITG: 20°Cで一定の暗間（DD）でフリーランニング期間（τ）よりも短くなり、25°Cで20°CでLD周期で歩行活動の位相前進であることが示唆され、τの減少が原因であった。

Proposed: 20°Cで一定の暗間（DD）でフリーランニング期間（τ）25°Cでのそれに比べ短くなっていることから、20°CでLD周期で歩行活動の位相前進 τの減少が原因であった。

A.2 Sentence 2

Source: From result of the consideration, it was pointed that radiation from the loop elements was weak.

Baseline: 考察の結果からすることを指摘し、ループ素子からの放射は弱かった。

Reference: 考察結果より、ループ素子からの放射が弱いことを指摘する。

Proposed: 考察の結果から、ことを指摘し、ループの要素からの放射は弱かった。

Proposed: 考察の結果から、ループ素子からの放射は弱いことを示した。
Hideo Okuma received the B.S. degree in mathematics from Tokyo University of Science in 1987. Mr. Okuma is currently a researcher at the National Institute of Communications Technology. His research include machine translation.

Eiichiro Sumita received the M.S. degree in computer science from the University of Electro-Communications in 1982 and the Ph.D. degree in engineering from Kyoto University in 1999. Dr. Sumita is the head of ATR/SLC/NLP, the leader of NICT/KCCRC/LTG, visiting professor of Kobe University and vice-president of ATR-Langue. His research interests include machine translation and e-Learning. He is a member of the IEEE, the ACL, the IPSJ, the ASJ and the ANLP. He serves Associate Editor of ACM/TSLP.

Keiichi Tokuda received the B.E. degree in electrical and electronic engineering from the Nagoya Institute of Technology, Nagoya, Japan, in 1984 and the M.E. and Dr.Eng. degrees in 8 IEICE TRANS. FUNDAMENTALS, VOL.E86.A, NO.5 MAY 2003 information processing from the Tokyo Institute of Technology, Tokyo, Japan, in 1986 and 1989, respectively. From 1989 to 1996, he was a Research Associate in the Department of Electronic and Electric Engineering, Tokyo Institute of Technology. From 1996 to 2004, he was an Associate Professor in the Department of Computer Science, Nagoya Institute of Technology, where he is currently a Professor. He is also an Invited Researcher at the ATR Spoken Language Communication Research Laboratories, Kyoto, Japan, and was a Visiting Researcher at Carnegie Mellon University, Pittsburgh, PA, from 2001 to 2002. His research interests include speech coding, speech synthesis and recognition, and statistical machine learning. Prof. Tokuda is a coreipient of the Paper Award and the Inose Award, both from the IEICE in 2001, and the TELECOM System Technology Prize from the Telecommunications Advancement Foundation Award, Japan, in 2001 and 2008. He was a member of the Speech Technical Committee of the IEEE Signal Processing Society. He is a member the ISCA, IPSJ, ASJ, and JSAI.