

# 概要

近年，機械翻訳の分野で原言語から目的言語に翻訳する統計翻訳が注目されている．統計翻訳は対訳文を用いてフレーズごとの翻訳確率翻訳モデルや単語確率である言語モデルを学習する．そして，文が入力された時に翻訳モデルと言語モデルの組合せから確率が最大となる文を出力文として得る．統計翻訳において，対訳文数が多ければ翻訳精度は高くなるが，対訳文数の量には限界がある．

この問題に対し，対訳コーパスに対訳句コーパスを追加する手法の研究が行われ，自動評価結果が向上したとの報告がある．この研究は，セルビア語英語間，スペイン語英語間 [1] そしてブルトン語フランス語間 [2] における句に基づく統計翻訳 (PSMT) において行われている．また日野らの研究においても，日本語英語間の PSMT における対訳コーパスへの対訳句コーパスの追加の効果が報告されている [3] ．

ところで，階層型統計翻訳 (HSMT) は文の構造を考慮して翻訳を行うため，PSMT よりも翻訳精度が高いことが一般的に知られている．

そこで，本研究では対訳コーパスに対訳句コーパスを追加する手法を日本語英語間の HSMT において行う．辞書のデータから抽出した対訳句コーパスを用いる．日英対訳コーパスを用いて言語モデルの学習する．日英対訳コーパスに対訳句コーパスを加えて翻訳モデルを学習し，HSMT を行う方法を提案手法とし，翻訳精度の調査を行う．

日英対訳コーパスは単文コーパスと重文複文コーパス [4] を用いる．対訳句コーパスとして鳥バンク [5] と英辞郎 [6] を用いる．

その結果，自動評価結果と人手評価結果のどちらにおいても提案手法の翻訳精度は向上した．よって，日本語英語間の HSMT でも PSMT と同様に，対訳句コーパスを用いた提案手法は有効であると確認できた．

# 目次

第1章	はじめに	1
第2章	統計翻訳システム	3
2.1	基本概念	3
2.1.1	翻訳モデル	4
2.1.1.1	GIZA++	4
2.1.2	言語モデル	4
2.1.2.1	$N$ -gram モデル	4
2.1.3	デコーダ	6
2.1.4	句に基づく統計翻訳	6
2.1.4.1	句に基づく統計翻訳の概要	6
2.1.4.2	翻訳モデル	6
2.1.5	階層型統計翻訳	7
2.1.5.1	階層型統計翻訳の概要	7
2.1.5.2	翻訳モデル	8
2.2	評価方法	8
2.2.1	自動評価法	9
2.2.1.1	BLEU	9
2.2.1.2	NIST	9
2.2.1.3	METEOR	10
2.2.1.4	RIBES	10
2.2.2	人手評価	11
第3章	提案手法	12
第4章	実験環境	14
4.1	日英対訳コーパス	14

4.2	翻訳モデルの学習	14
4.3	言語モデルの学習	14
4.4	デコーダ	14
4.5	対訳句コーパス	15
4.5.1	鳥バンク	15
4.5.2	英辞郎	16
4.6	実験内容	16
4.6.1	評価方法	16
<b>第5章</b>	<b>実験結果</b>	<b>17</b>
5.1	自動評価	17
5.2	人手評価	19
5.2.1	実験結果のまとめ	20
5.2.2	対比較の出力例	20
<b>第6章</b>	<b>考察</b>	<b>26</b>
6.1	HSMT と PSMT の比較	26
6.1.1	HSMT と PSMT の比較のまとめ	27
6.1.2	対比較の出力例	27
6.2	鳥バンクと英辞郎の比較	30
6.2.1	鳥バンクと英辞郎の比較のまとめ	30
6.2.2	対比較の出力例	31
6.3	語彙の増加が生じない対訳句追加	33
<b>第7章</b>	<b>分野外実験</b>	<b>34</b>
7.1	分野外実験の内容	34
7.2	分野外実験のデータ	34
7.3	分野外実験結果	34
7.3.1	自動評価結果	34
7.3.2	人手評価	35
7.3.3	分野外実験結果のまとめ	36
7.3.4	対比較の出力例	36



# 目 次

2.1	日英統計翻訳の流れ . . . . .	3
2.2	句に基づく統計翻訳による翻訳の例 . . . . .	6
2.3	階層型統計翻訳による翻訳の例 . . . . .	8
3.1	日英統計翻訳の場合の提案手法の流れ . . . . .	12

# 表 目 次

2.1	日英フレーズテーブルの例 . . . . .	7
2.2	日英ルールテーブルの例 . . . . .	8
2.3	対比較例評価出力例 . . . . .	11
2.4	対比較評価基準例 . . . . .	11
2.5	対比較評価結果例 . . . . .	11
4.1	日英対訳文対数 . . . . .	14
4.2	対訳句対数 . . . . .	15
4.3	鳥バンクから抽出した対訳句の例 . . . . .	15
4.4	英辞郎から抽出した対訳句の例 . . . . .	16
5.1	単文 日英翻訳 . . . . .	18
5.2	単文 英日翻訳 . . . . .	18
5.3	重文複文 日英翻訳 . . . . .	18
5.4	重文複文 英日翻訳 . . . . .	18
5.5	単文 日英翻訳 . . . . .	19
5.6	単文 英日翻訳 . . . . .	19
5.7	重文複文 日英翻訳 . . . . .	20
5.8	重文複文 英日翻訳 . . . . .	20
5.9	“HSMT”と“HSMT+鳥”の対比較の出力例 (単文 日英) . . . . .	22
5.10	“HSMT”と“HSMT+英”の対比較の出力例 (単文 日英) . . . . .	22
5.11	“HSMT”と“HSMT+鳥”の対比較の出力例 (単文 英日) . . . . .	23
5.12	“HSMT”と“HSMT+英”の対比較の出力例 (単文 英日) . . . . .	23
5.13	“HSMT”と“HSMT+鳥”の対比較の出力例 (重文複文 日英) . . . . .	24
5.14	“HSMT”と“HSMT+英”の対比較の出力例 (重文複文 日英) . . . . .	24
5.15	“HSMT”と“HSMT+鳥”の対比較の出力例 (重文複文 英日) . . . . .	25
5.16	“HSMT”と“HSMT+英”の対比較の出力例 (重文複文 英日) . . . . .	25

6.1	単文 日英翻訳	26
6.2	単文 英日翻訳	26
6.3	重文複文 日英翻訳	26
6.4	重文複文 英日翻訳	26
6.5	“PSMT+鳥”と“HSMT+鳥”の対比較の出力例 (単文 日英)	28
6.6	“PSMT+鳥”と“HSMT+鳥”の対比較の出力例 (単文 英日)	28
6.7	“PSMT+鳥”と“HSMT+鳥”の対比較の出力例 (重文複文 日英)	29
6.8	“PSMT+鳥”と“HSMT+鳥”の対比較の出力例 (重文複文 英日)	29
6.9	単文 日英翻訳	30
6.10	単文 英日翻訳	30
6.11	重文複文 日英翻訳	30
6.12	重文複文 英日翻訳	30
6.13	“HSMT+鳥”と“HSMT+英”の対比較の出力例 (単文 日英)	31
6.14	“HSMT+鳥”と“HSMT+英”の対比較の出力例 (単文 英日)	32
6.15	“HSMT+鳥”と“HSMT+英”の対比較の出力例 (重文複文 日英)	32
6.16	“HSMT+鳥”と“HSMT+英”の対比較の出力例 (重文複文 英日)	33
7.1	重文複文 日英翻訳	35
7.2	重文複文 英日翻訳	35
7.3	重文複文 日英翻訳	35
7.4	重文複文 英日翻訳	36
7.5	“HSMT”と“HSMT+鳥”の対比較の出力例 (分野外の重文複文 日英)	37
7.6	“HSMT”と“HSMT+鳥”の対比較の出力例 (分野外の重文複文 英日)	37

# 第1章 はじめに

近年，機械翻訳の分野で原言語から目的言語に翻訳する統計翻訳が注目されている．統計翻訳は対訳文を用いてフレーズごとの翻訳確率翻訳モデルや単語確率である言語モデルを学習する．そして，文が入力された時に翻訳モデルと言語モデルの組合せから確率が最大となる文を出力文として得る．統計翻訳において，対訳文数が多ければ翻訳精度は高くなるが，対訳文数には限りがある．

この問題に対し，対訳コーパスに対訳句コーパスを追加する手法の研究が行われ，自動評価結果が向上したとの報告がある．この研究は，セルビア語英語間，スペイン語英語間 [1]，ブルトン語フランス語間 [2] そして，日本語英語間 [3] における句に基づく統計翻訳 (PSMT) において行われている．

Maja Popovićらはセルビア語英語間，スペイン語英語間において，対訳コーパスに対訳句コーパスを追加し句に基づく統計翻訳を行った [1]．セルビア語英語間の翻訳で，対訳句コーパスとして 351 対訳句を使用した．セルビア語英語方向の翻訳において，対訳コーパス 0.2k に対訳句コーパスを追加することによって，BLEU スコアで 0.020 の向上を報告した．また，スペイン語英語間の翻訳で，対訳句コーパスとして 52,566 対訳句を使用した．スペイン語英語方向の翻訳において，対訳コーパス 1k に対訳句コーパスを追加することによって，BLEU スコアで 0.06 の向上を報告した．

Francis M Tyers はブルトン語フランス語間の翻訳において対訳句コーパスを用いた [2]．その結果，対訳コーパス 27,987 文に語彙変化しない対訳句コーパスとして 7k 対訳句対を追加することによって，BLEU スコアで 0.01 の向上を報告した．また，対訳コーパスに語彙変化させた対訳句コーパスとして 43k 対訳句対を追加することによって BLEU スコアで 0.07 の向上を報告した．

日野らは日本語英語間の翻訳において対訳句コーパスを用いた [3]．その結果，対訳コーパス 100k に対訳句コーパスとして 698,472 対訳句を追加することによって，BLEU スコアで 0.02 の向上を報告した．

ところで，階層型統計翻訳 (HSMT) は文の構造を考慮して翻訳を行うため，PSMT よりも翻訳精度が高いことが一般的に知られている．



そこで、本研究では対訳コーパスに対訳句コーパスを追加する手法を日本語英語間の HSMT において行う。辞書のデータから抽出した対訳句コーパスを用いる。日英対訳コーパスを用いて言語モデルの学習する。日英対訳コーパスに対訳句コーパスを追加して翻訳モデルを学習し、HSMT を行う方法を提案手法とし、翻訳精度の調査を行う。

日英対訳コーパスは単文コーパスと重文複文コーパス [4] を用いる。対訳句コーパスとして鳥バンク [5] から抽出した対訳句 と英辞郎 [6] から抽出した対訳句 を用いる。

その結果、自動評価結果と人手評価結果のどちらにおいても提案手法の翻訳精度は向上した。よって、日本語英語間の HSMT でも PSMT と同様に、対訳句コーパスを用いた提案手法は有効であると確認できた。

本論文の構成は以下の通りである。まず、2 章で統計翻訳システムの概要を示し、各モデルの学習、評価方法について述べる。3 章では、本研究の提案手法について述べる。4 章では、実験に用いるデータやツールといった実験環境について述べる。5 章では、提案手法の結果を示す。そして、6 章で考察を行い、7 章で分野外実験を行う。最後に 8 章で結論を述べ、まとめる。

## 第2章 統計翻訳システム

統計翻訳システムを，原言語（翻訳の対象となる入力された言語）を日本語，目的言語（翻訳された後に出力される言語）を英語とする日英統計翻訳の場合を例として説明する．

### 2.1 基本概念

日英統計翻訳システムは，日本語入力文  $j$  が与えられたとき，全ての組み合わせの中から確率が最大となる英語文  $\hat{e}$  を探索することで翻訳を行う．以下に基本モデルを示す．

$$\hat{e} = \arg \max_e P(e|j) \quad (2.1)$$

$$\approx \arg \max_e P(j|e)P(e) \quad (2.2)$$

$P(j|e)$  は翻訳モデル， $P(e)$  は言語モデルと呼ぶ． $\hat{e}$  を探索する翻訳システムをデコーダと呼ぶ．図 2.1 に日英統計翻訳の流れを示す．

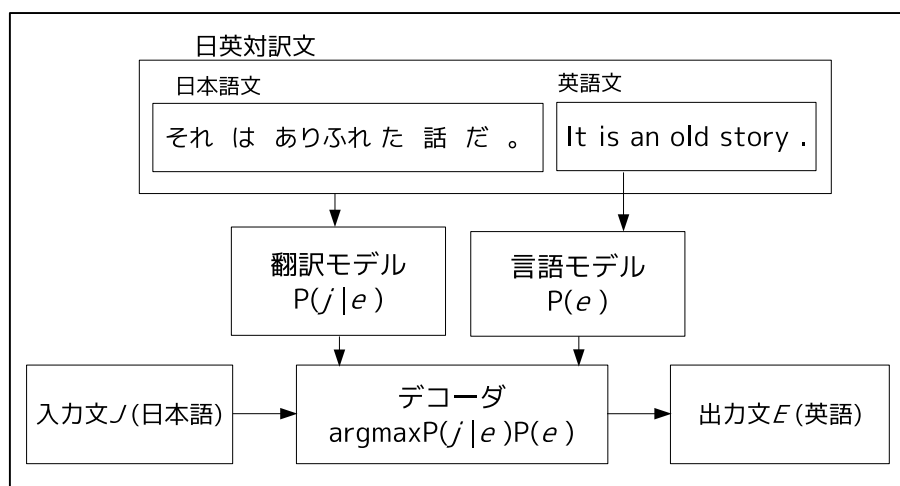


図 2.1: 日英統計翻訳の流れ

### 2.1.1 翻訳モデル

翻訳モデルの代表例として IBM 翻訳モデル [7] がある。IBM モデルは英語文  $e$  , 日本語文  $j$  の翻訳モデル  $P(j|e)$  を計算するためにアラインメント  $a$  と呼ばれる概念を導入し、以下のような式を考える。なお、アラインメントとは、ある日本語単語  $j$  と英単語  $e$  の対応関係のことを示す。

$$P(j|e) = \sum_a P(j, a|e)$$

IBM モデルでは、日英統計翻訳の場合、英単語は 1 対  $n$  の対応を持ち、日本語の単語は 1 つの英単語のみと対応すると仮定する。また、日本語の単語の対応関係として適切な英単語がなかった場合、英語文の文頭の特殊文字  $e$  と対応付けを行う。

#### 2.1.1.1 GIZA++

GIZA++ [8] とは、統計翻訳を用いることを前提に作られた単語対応のアラインメントを行うツールである。IBM モデルを学習し、単語の対応関係の確率値を計算する。

### 2.1.2 言語モデル

言語モデルは翻訳候補の文に対して目的言語の文らしさの指標を与えるモデルである。翻訳モデルでは、訳語の選択や訳語の位置の選択に対する評価を与えることはできるが、作られた翻訳候補が目的言語の文としてふさわしいかどうかを判断する評価を与えることはできない。そのため、言語モデルでは日英統計翻訳の場合、より英語らしい文に対して、高い確率を与えることで、翻訳モデルで翻訳された訳文候補の中から英語として自然な文を選出する。

#### 2.1.2.1 $N$ -gram モデル

統計翻訳では一般的に、 $N$ -gram モデルを用いる。 $N$ -gram モデルとは“単語列  $w_1^n = w_1, w_2, w_3, \dots, w_n$  の  $i$  番目の単語  $w_i$  の生起確率  $P(w_i)$  は直前の  $(N-1)$  の単語列  $w_{i-(N-1)}, w_{i-(N-2)}, w_{i-(N-3)}, \dots, w_{i-1}$  に依存する” という仮説に基づくものである。また、 $N = 1$  のモデルを uni-gram ,  $N = 2$  のモデルを bi-gram ,  $N = 3$  のモデルを tri-gram と特有の呼び

かたをする． $N = 4$  以上は 4-gram など数値を用いて呼ぶ． $N$ -gram は以下の式で表現される．ここで， $w_i^j$  は  $i$  から  $j$  番目までの単語列を表す．

$$P(w_1^n) = P(w_1) \times P(w_2|w_1) \times P(w_3|w_1^2) \dots P(w_n|w_1^{n-1}) \quad (2.3)$$

$$\approx P(w_1) \times P(w_2|w_1) \times P(w_3|w_1^2) \dots P(w_n|w_{n-(N-1)}^{n-1}) \quad (2.4)$$

$$= \prod_{i=1}^n P(w_i|w_{i-(N-1)}^{i-1}) \quad (2.5)$$

また， $P(w_i|w_{i-(N-1)}^{i-1})$  は以下の式で計算される．ここで  $C(w_1^i)$  は単語列  $w_1^i$  が出現する頻度を表す．

$$P(w_i|w_{i-(N-1)}^{i-1}) = \frac{C(w_{i-(N-1)}^i)}{C(w_{i-(N-1)}^{i-1})} \quad (2.6)$$

たとえば，“I have dogs .” という単語列に対して  $N = 2$  とした bi-gram モデルの言語モデルを適応した場合，単語列が生成される確率は以下の式で計算される．

$$P(\text{“I have dogs .”}) \simeq P(I) \times P(\text{have}|I) \times P(\text{dogs|have}) \dots P(.|\text{dogs}) \quad (2.7)$$

tri-gram モデルであれば， $P(\text{dogs}|I \text{ have})$ ，4-gram モデルであれば  $P(.|I \text{ have dogs})$  となる．

(2.6) 式から信頼性の高い値を推定するためには，単語列  $w_1^n$  が多く出現している必要がある．しかし，実際には多くの単語列は出現数が 0 となることが多いため信頼できる値を推定できない場合が多い．低頻度な語彙の場合， $C(w_{i-(N-1)}^i)$ ， $C(w_{i-(N-1)}^{i-1})$  の値が小さく，信頼性が低い．また，学習データ中に単語列  $w_1^i$  が存在しない場合，この単語列の出現確率は 0 と推定される．そのため，(2.6) 式から信頼できる値を算出するためには，大規模なコーパスを用いて，各単語列の出現数を高める必要がある．そこで，出現頻度の少ない単語列をモデルの学習から削除(カットオフ)する方法や，確率が 0 となるのを防ぐために，大きい確率を小さく，小さい確率を大きくするスムージング手法が提案されている．スムージングの代表的な手法にバックオフ・スムージングがある．バックオフ・スムージングは学習データに出現しない  $N$ -gram の値をより低い次数の  $(N-1)$ -gram の値から推定する．

### 2.1.3 デコーダ

デコーダは翻訳モデルと言語モデルの全ての組み合わせの中から確率が最大となる出力文を探索して翻訳を行う。代表的なデコーダに Moses[9] がある。

### 2.1.4 句に基づく統計翻訳

#### 2.1.4.1 句に基づく統計翻訳の概要

2000年の初めに「句に基づく統計翻訳」が提案された。句に基づく統計翻訳は、句の対応を翻訳モデルに用いる。句を構成する単語の数が、翻訳する文の句と目的文の句で一致する必要がないため、単語に基づく統計翻訳の単語対応の問題を解決した。また、並び替えにおいても、単語に基づく統計翻訳よりも優れている。そのため、近年では句に基づく統計翻訳が主流となっている。なお、本論文では、句に基づく統計翻訳を「PSMT」と表記する。

句に基づく統計翻訳による翻訳の例を図 2.2 に示す。

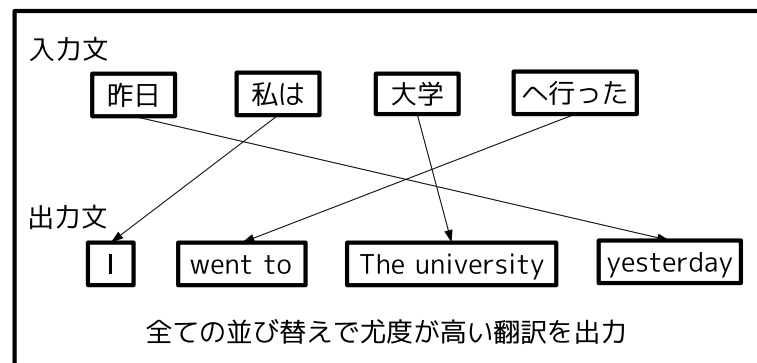


図 2.2: 句に基づく統計翻訳による翻訳の例

#### 2.1.4.2 翻訳モデル

句に基づく統計翻訳の翻訳モデルはフレーズテーブルによって管理されている。フレーズテーブルの例を表 2.1 に示す。

表 2.1: 日英フレーズテーブルの例

これは	This is a	0.525424 0.0311118 0.118774 0.000752339
これは	This is an	0.666667 0.0179866 0.0153257 0.000128339
これは	This is	0.127148 0.0102929 0.141762 0.0110294
これは私の	This is my	0.333333 0.000404608 0.5 0.000884628
これは私の	of this	0.00877193 4.32631e-06 0.5 0.0256013

左から，日本語フレーズ，英語フレーズ，フレーズの英日翻訳確率  $P(j|e)$ ，英日方向の単語翻訳確率の積，フレーズの日英方向の翻訳確率  $P(e|j)$ ，日英方向の単語翻訳確率の積である．

## 2.1.5 階層型統計翻訳

### 2.1.5.1 階層型統計翻訳の概要

2005年に「階層型統計翻訳」が提案された．階層型統計翻訳は，句に基づく統計翻訳を拡張した手法である．階層型統計翻訳は，階層句の対応で，木構造を用いて翻訳を行うため，文法構造を考慮する．階層句とは，非終端記号を含む句である．そのため，文法構造の大きく異なる日本語-英語間での翻訳において，句に基づく統計翻訳よりも翻訳精度が高い．

階層型統計翻訳による翻訳の例を図 2.3 に示す．手順として，まず，学習文から階層句の対応を学習する．そして，入力文が与えられたとき，句の中に句が含まれる階層に分解し，それぞれの階層に階層句を当てはめて空白を埋めていくことで目的言語の出力文を生成する．なお，本論文では，階層型統計翻訳を「HSMT」と表記する．

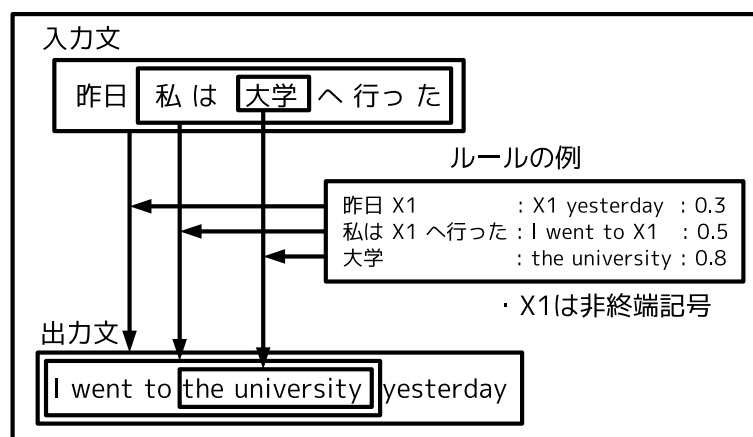


図 2.3: 階層型統計翻訳による翻訳の例

### 2.1.5.2 翻訳モデル

階層型統計翻訳における翻訳モデルには，ルールテーブルを用いる．ルールテーブルの例を表 2.2 に示す．

表 2.2: 日英ルールテーブルの例

[X]		これは		This is		0.138021	0.0102929	0.140684	0.0110294			
[X]		これは	[X,1][X,2]		This is a [X,1][X,2]		0.5879	0.03111	0.1734	0.00075		
[X]		これは	[X,1][X,2]	です。		This is a [X,1][X,2]	.		0.038	0.00018	0.7	0.0076
[X]		新しい		a new		0.165997	0.39984	0.029316	0.0253706			
[X]		新しい		new		0.460302	0.39984	0.79316	0.884752			

左から，非終端記号，日本語ルール，英語ルール，ルールの英日翻訳確率  $P(j|e)$ ，英日方向の単語翻訳確率の積，ルールの日英方向の翻訳確率  $P(e|j)$ ，日英方向の単語翻訳確率の積である．

## 2.2 評価方法

本研究では，翻訳システムによって出力した文の評価に自動評価法と人手評価法を用いる．

## 2.2.1 自動評価法

機械翻訳システムの翻訳精度を自動的に評価する手法として、あらかじめ用意した正解文と、翻訳システムで出力した文とを比較する手法が一般的である。自動評価法には多くの手法がある。本研究では、BLEU[10]、NIST[11]、METEOR[12]、RIBES[13]を用いる。

### 2.2.1.1 BLEU

BLEUは語順(4-gram)が正しい場合に高いスコアを出す。BLEUは以下の式で計算される。

$$BLEU_{score} = BP \times \exp\left(\sum_{n=1}^N \frac{1}{N} \log p_n\right) \quad (2.8)$$

$$p_n = \frac{\sum_i \text{出力文 } i \text{ と正解文 } i \text{ で一致した } N\text{-gram の数}}{\sum_i \text{出力文 } i \text{ の } N\text{-gram の数}} \quad (2.9)$$

ここで、 $P_n$ は出力文と正解文の $N$ -gramの一致率を表している。BLEUはこの一致率を1-gramから4-gramまで計算し、その幾何平均をとる。また、出力文が正解文より短い場合、“ $\sum_i \text{出力文 } i \text{ の } N\text{-gram の数}$ ”が小さくなり、不当にスコアが高くなる可能性がある。そこで、正解文より短い文に対するペナルティとして、 $BP$ を用いる。 $BP$ は出力文が正解文より長い場合は1となり、出力文が正解文より短い場合は1未満の値となる。

### 2.2.1.2 NIST

NISTではBLEUと同様に語順の正しさで比較を行うが、5-gramを用いる。NISTは以下の式で計算される。

$$NIST_{score} = BP \times \sum_{N=1}^N \frac{\sum_i \left( \sum_{\substack{\text{出力文 } i \text{ と正解文 } i \text{ に} \\ \text{共通する } w_1 \dots w_n}} \text{Info}_i(w_1 \dots w_n) \right)}{\sum_i \text{出力文 } i \text{ と正解文 } i \text{ で一致した } N\text{-gram の数}} \quad (2.10)$$

$$\text{Info}_i(w_1 \dots w_n) = \log_2 \frac{\text{評価コーパス中 } (w_1 \dots w_{n-1}) \text{ の数}}{\text{評価コーパス中の } (w_1 \dots w_n) \text{ の数}} \quad (2.11)$$



### 2.2.1.3 METEOR

METEOR は単語属性 (3 人称単数など) が正しい場合に高いスコアを出す。METEOR は以下の式で計算される。

$$METEOR_{score} = F_{mean} \times (1 - Pen) \quad (2.12)$$

$$F_{mean} = \frac{P \times R}{\alpha \times P + (1 - \alpha) \times R} \quad (2.13)$$

$$Pen = \gamma \times \left(\frac{c}{m}\right)^\beta \quad (2.14)$$

METEOR はまず再現率  $R$  と適合率  $P$  に基づく F 値を求め、次に、単語の非連続性に対するペナルティとして関数  $Pen$  を与える。ペナルティ関数  $Pen$  において、 $m$  は出力文と正解文の単語の一致率を表す。そして、 $c$  は一致した単語を対象に、正解文と語順が同じものを 1 つのまとまりとして統合した場合の、まとまりの数を表す。そのため、出力文と正解文が同じ文であるとき  $c=1$  となる。また、一致率の計算において、WordNet による類義語を用いて、似た意味を持つ単語は同一であると判断される。 $\alpha, \beta, \gamma$  の値はパラメータである。

BLEU と METEOR と RIBES では 0 から 1 までの間で評価され、NIST では 0 から 1 までの間で評価される。いずれの評価方法でも、評価方法が高いほど翻訳精度が高いことを表す。なお、本研究では入力文 1 文に対して正解文 1 文を用いて評価を行う。

### 2.2.1.4 RIBES

RIBES は、正解文と出力文の間で、共通単語の出現順序を順位相関係数で評価を行う評価法である。RIBES は以下の式で計算される。

$$RIBES(S) = NSR \times P^\alpha \quad (2.15)$$

$$RIBES(K) = NKT \times P^\alpha \quad (2.16)$$

$$P = \frac{n}{h} \quad (2.17)$$

ここで、 $NSR$  はスピアマンの順位相関係数であり、 $NKT$  はケンドールの順位相関係数である。また  $P$  はペナルティであり、 $n$  は出力文と正解文との間で共通な単語の数、 $h$  は出力文の単語数で計算する。 $\alpha$  はペナルティに対する主として使用され、 $0 < \alpha < 1$  の値である。単語の出現順を順位相関係数を用いて評価することで、文全体の語順に着目することができる。

## 2.2.2 人手評価

人手評価として、対比較評価がある。対比較評価は、二つの文を相対的に比較して、どちらがより正しい文であるかを人手で選択する評価方法である。二つの翻訳システムの出力で優劣を判断する場合に有効である。

対比較を行う例を表 2.3 に示す。表 2.3 より、“手法 A”と“手法 B”を比較すると、正しい翻訳をしている文は“手法 B”である。よって、この場合は評価者によって、表 2.4 の選択肢から“手法 B”が選択される。また、“手法 A”と“手法 B”の対比較評価を 100 文に対して行った場合の結果の例を表 2.5 に示す。表 2.5 では“手法 B”が優れていると見て取ることができる。

また、人手評価は評価者によって評価に差が出てくる可能性がある。しかし、表 2.3 のように正解文を用意しておき、参照して対比較評価を行うことで、信頼性を高める。

表 2.3: 対比較例評価出力例

例 1	
入力文	これは卒業論文です。
正解文	This is a graduation thesis.
手法 A	This is a student yearbook.
手法 B	This is a thesis of graduation.

表 2.4: 対比較評価基準例

手法 A	手法 A の方が良い
手法 B	手法 B の方が良い
差なし	手法 A と手法 B の出力文の単語対応や伝わる意味に差がない
一致	手法 A と手法 B の出力文が完全一致

表 2.5: 対比較評価結果例

手法 A	手法 B	差なし	一致
3	12	78	7

### 第3章 提案手法

本研究では辞書のデータから抽出した対訳句コーパスを日英対訳コーパスに追加し，HSMT を行う手法を提案手法とする．

図 3.1 に日英統計翻訳の場合の提案手法の流れを示す．

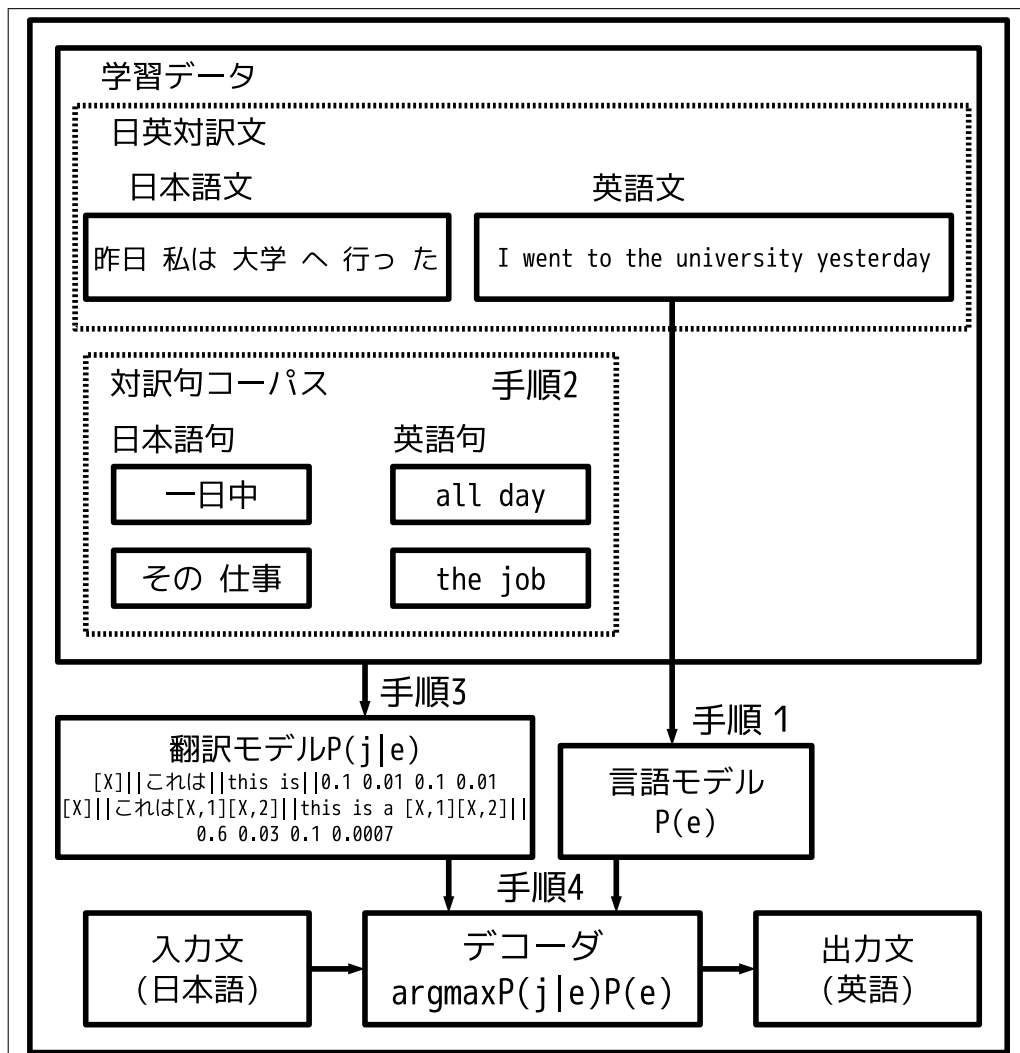


図 3.1: 日英統計翻訳の場合の提案手法の流れ

提案手法の手順を以下に示す。

手順 1 日英対訳コーパスを学習データとして言語モデルを作成する

手順 2 日英対訳コーパスに対訳句コーパスを追加する

手順 3 手順 2 で作成したコーパスを学習データとして翻訳モデルを作成する

手順 4 手順 1 と手順 3 で作成したモデルを用いて統計翻訳を行う

## 第4章 実験環境

### 4.1 日英対訳コーパス

本研究では日英対訳コーパスとして，単文コーパスと重文複文コーパス [4] を用いる．統計翻訳の前処理として，各コーパスの日本語文に対して，MeCab[14] を用いて形態素解析を行う．また，英語文に対して “tokenizer.sed [15]” を用いて正規化を行う．本研究で用いる単文コーパス [4]，重文複文コーパス [4] を表 4.1 の内訳で用いる．

表 4.1: 日英対訳文対数

	単文	重文複文
学習データ	100,000 文	100,000 文
ディベロップメントデータ	1,000 文	1,000 文
テストデータ	10,000 文	10,000 文

### 4.2 翻訳モデルの学習

翻訳モデルの学習には，“train-model.perl[9]” を用いる．

### 4.3 言語モデルの学習

言語モデルの学習には，“SRILM[16]” の “ngram-count” を用いる．本研究では， $N$ -gram モデルは 5-gram とする．

### 4.4 デコーダ

本実験では PSMT のデコーダとして，“moses[9]” を用いる．また，HSMT のデコーダとしても “moses” を用いる．

moses のパラメータは, “mert-moses.pl[9]” を用いてチューニングを行う。また, 統計翻訳においては, moses の設定ファイル “moses.ini” の “distortion-limit” の値を “-1” に変更する。distortion-limit はフレーズの並び替えを制限する値である。-1 にすることで, 無制限とする。

## 4.5 対訳句コーパス

本研究では対訳句コーパスとして鳥バンク [5] と英辞郎 [6] を用いる。本研究で対訳句コーパスとして用いる鳥バンク, 英辞郎の対訳句数を表 4.2 に示す。

表 4.2: 対訳句対数

鳥バンク	548,694 句
英辞郎	1,366,458 句

### 4.5.1 鳥バンク

鳥バンクは自然言語処理のための言語知識ベースを収録したデータバンクであり, 日本語の重文と複文を対象とする「意味類型パターン辞書」が収録されている。本研究では, 鳥バンクの対訳句として表 4.1 中の重文複文の学習データから抽出した 548,694 対訳句 [?] を用いる。対訳句の例を表 4.3 に示す。

表 4.3: 鳥バンクから抽出した対訳句の例

<u>彼のお母さん</u>
his mother
<u>息子の話</u>
son 's story
<u>社会通念</u>
commonly accepted ideas

## 4.5.2 英辞郎

英辞郎は、EDP(Electronic Dictionary Project) がアップデートし続けている英和・和英辞書である。英辞郎のデータには対訳句の他に翻訳例や注釈、本来の文に出てこない“~”等の記号が含まれる。表 4.4 に英辞郎の対訳句の例を示す。

表 4.4: 英辞郎から抽出した対訳句の例

<u>に急落する</u>
<u>plummet to</u>
<u>から出てくる</u>
<u>come out from</u>
<u>の結果として生じる</u>
<u>come out from</u>

## 4.6 実験内容

本研究では、以下に示す合計 16 種類の翻訳実験を行う。

- 日英対訳文として、単文と重文複文を用いる。
- 翻訳実験として、日英統計翻訳と英日統計翻訳を行う。
- 統計翻訳の手法として、HSMT と PSMT を用いる。
- 追加する対訳句として、鳥バンクと英辞郎を用いる。

### 4.6.1 評価方法

本研究では、出力文の評価として自動評価と人手評価を行う。自動評価は自動評価法“BLEU[10]”，“NIST[11]”，“METEOR[12]”，“RIBES[13]”を用いる。人手評価は出力文からランダムに 100 文抽出し、対比較評価を行う。

## 第5章 実験結果

表記方法を以下に示す．

- 対訳句の追加を行っていない HSMT の翻訳結果を“HSMT”と表記する．
- 対訳句の追加を行っていない PSMT の翻訳結果を“PSMT”と表記する．
- 学習データに対して鳥バンクの対訳句の追加を行った HSMT の翻訳結果を“HSMT+鳥”と表記する．
- 学習データに対して鳥バンクの対訳句の追加を行った PSMT の翻訳結果を“PSMT+鳥”と表記する．
- 学習データに対して鳥バンクの対訳句の追加を行った HSMT の翻訳結果を“HSMT+英”と表記する．
- 学習データに対して鳥バンクの対訳句の追加を行った PSMT の翻訳結果を“PSMT+英”と表記する．
- 対比較評価において翻訳後の出力が完全に同一な場合は，“一致”と表記する．
- 対比較評価において翻訳精度にかかわらず同程度である場合は，“差なし”と表記する．

### 5.1 自動評価

自動評価における，単文の日英翻訳の結果を表 5.1，単文の英日翻訳の結果を表 5.2，重文複文の日英翻訳の結果を表 5.3，重文複文の英日翻訳の結果を表 5.4 に示す．



表 5.1: 単文 日英翻訳

	BLEU	NIST	METEOR	RIBES
HSMT	0.1397	4.8430	0.4590	0.7143
HSMT+鳥	0.1544	5.2274	0.4909	0.7268
HSMT+英	0.1527	5.2628	0.4919	0.7315
PSMT	0.1388	4.9161	0.4608	0.7152
PSMT+鳥	0.1561	5.2203	0.4897	0.7308
PSMT+英	<b>0.1595</b>	<b>5.3405</b>	<b>0.4964</b>	<b>0.7355</b>

表 5.2: 単文 英日翻訳

	BLEU	NIST	METEOR	RIBES
HSMT	0.1827	4.7136	0.4581	0.7718
HSMT+鳥	0.1998	5.0070	0.4824	0.7848
HSMT+英	<b>0.2064</b>	<b>5.1298</b>	<b>0.4868</b>	<b>0.7871</b>
PSMT	0.1811	4.6653	0.4563	0.7699
PSMT+鳥	0.1904	4.9788	0.4665	0.7765
PSMT+英	0.1963	4.9880	0.4762	0.7760

表 5.3: 重文複文 日英翻訳

	BLEU	NIST	METEOR	RIBES
HSMT	0.1263	4.5297	0.4183	0.6770
HSMT+鳥	0.1400	4.9040	0.4486	0.6915
HSMT+英	0.1412	<b>4.9925</b>	0.4546	<b>0.6998</b>
PSMT	0.1223	4.3717	0.4051	0.6694
PSMT+鳥	<b>0.1429</b>	4.9310	<b>0.4569</b>	0.6925
PSMT+英	0.1426	4.9136	0.4525	0.6961

表 5.4: 重文複文 英日翻訳

	BLEU	NIST	METEOR	RIBES
HSMT	0.1502	4.4037	0.4050	0.7220
HSMT+鳥	0.1580	4.4772	0.4154	0.7234
HSMT+英	<b>0.1648</b>	4.6284	<b>0.4271</b>	<b>0.7311</b>
PSMT	0.1498	4.2698	0.3983	0.7140
PSMT+鳥	0.1573	4.5137	0.4116	0.7155
PSMT+英	0.1635	<b>4.6480</b>	0.4257	0.7230

## 5.2 人手評価

人手評価における，単文の日英翻訳の結果を表 5.5，単文の英日翻訳の結果を表 5.6，重文複文の日英翻訳の結果を表 5.7，重文複文の英日翻訳の結果を表 5.8 に示す．

具体的には，表 11 の一番上の表は，対訳句を追加していない HSMT における翻訳結果と，学習データに鳥バンクの対訳句を追加した HSMT の翻訳結果との対比較を行い，HSMT における翻訳結果が優れていると判断したのが 6 文，HSMT において学習データに鳥バンクの対訳句を追加して行った翻訳結果が優れていると判断したのが 11 文，差なしと判断したのが 79 文，一致と判断したのが 4 文であることを表している。

表 5.5: 単文 日英翻訳

HSMT	HSMT+鳥	差なし	一致
6	11	79	4
HSMT	HSMT+英	差なし	一致
4	10	79	7
PSMT	PSMT+鳥	差なし	一致
4	16	73	7
PSMT	PSMT+英	差なし	一致
4	14	72	10

表 5.6: 単文 英日翻訳

HSMT	HSMT+鳥	差なし	一致
2	10	83	5
HSMT	HSMT+英	差なし	一致
1	12	74	13
PSMT	PSMT+鳥	差なし	一致
4	9	80	7
PSMT	PSMT+英	差なし	一致
5	14	72	9

表 5.7: 重文複文 日英翻訳

HSMT	HSMT+鳥	差なし	一致
6	11	78	5
HSMT	HSMT+英	差なし	一致
5	7	84	4
PSMT	PSMT+鳥	差なし	一致
3	14	78	5
PSMT	PSMT+英	差なし	一致
3	16	81	0

表 5.8: 重文複文 英日翻訳

HSMT	HSMT+鳥	差なし	一致
4	7	83	6
HSMT	HSMT+英	差なし	一致
4	9	84	3
PSMT	PSMT+鳥	差なし	一致
2	11	83	4
PSMT	PSMT+英	差なし	一致
4	17	77	2

### 5.2.1 実験結果のまとめ

- HSMT においても全ての実験結果において、対訳句の追加における翻訳精度の向上が見られた。これは PSMT の場合と同様の結果である。
- HSMT と PSMT のそれぞれにおける対訳句の追加の効果を比べた場合、PSMT における対訳句の追加の効果の方が高かった。ただし、単文の英日翻訳においては HSMT における対訳句の追加の効果の方が高かった。
- 鳥バンクと英辞郎の対訳句の追加の効果の差は見られなかった。

### 5.2.2 対比較の出力例

単文の日英翻訳における“HSMT”と“HSMT+鳥”の対比較の出力例を表 5.9, 単文の日英翻訳における“HSMT”と“HSMT+英”の対比較の出力例を表 5.10, 単文の英日翻訳におけ

る“HSMT”と“HSMT+鳥”の対比較の出力例を表 5.11, 単文の英日翻訳における“HSMT”と“HSMT+英”の対比較の出力例を表 5.12, 重文複文の日英翻訳における“HSMT”と“HSMT+鳥”の対比較の出力例を表 5.13, 重文複文の日英翻訳における“HSMT”と“HSMT+英”の対比較の出力例を表 5.14, 重文複文の英日翻訳における“HSMT”と“HSMT+鳥”の対比較の出力例を表 5.15, 重文複文の英日翻訳における“HSMT”と“HSMT+英”の対比較の出力例を表 5.16 に示す .

表 5.9: “HSMT”と“HSMT+鳥”の対比較の出力例 (単文 日英)

HSMT の出力例	
入力文	我が社には優秀な技術者が居ります。
正解文	Our company has some talented technicians .
HSMT	There is excellent engineers in our company .
HSMT+鳥	Our company has excellent engineers is 居り .
HSMT+鳥 の出力例	
入力文	彼女はいつもクラスの首席を占めている。
正解文	She is always at the top of her class .
HSMT	She always has the 首席 in the class .
HSMT+鳥	She is always the top of his class .
差無しの出力例	
入力文	その試合はテレビで見ました。
正解文	I watched the game on television .
HSMT	I saw the game by a television .
HSMT+鳥	I saw the game on television .

表 5.10: “HSMT”と“HSMT+英”の対比較の出力例 (単文 日英)

HSMT の出力例	
入力文	ぼくは純粋な動機からそう言っているのだ。
正解文	I am speaking from a disinterested motive .
HSMT	I say that from the neat motives .
HSMT+英	I is say that out of pure motives .
HSMT+英 の出力例	
入力文	彼の声は落ち着き払っていた。
正解文	His voice was even .
HSMT	His voice was 落ち着き払っ .
HSMT+英	His voice was composure .
差無しの出力例	
入力文	自動車が動かない。
正解文	The car will not start .
HSMT	The car won't work .
HSMT+英	The car doesn't work .

表 5.11: “HSMT” と “HSMT+鳥” の対比較の出力例 (単文 英日)

HSMT の出力例	
入力文	The strike originated in a demand for higher wages .
正解文	ストは賃上げ要求から起こった。
HSMT	ストライキは賃金値上げの要求で始まった。
HSMT+鳥	賃上げ要求ストに源を発している。
HSMT+鳥 の出力例	
入力文	We were totally crushed by the terrible news .
正解文	そのひどい知らせにすっかり打ちひしがれた。
HSMT	私たちはひどい知らせでいた。
HSMT+鳥	私たちはその恐ろしい知らせに完全に打ちひしがれた。
差無しの出力例	
入力文	The debris from the accident fell over a wide area .
正解文	その事故の残骸は広い範囲に落下した。
HSMT	その事故の残骸は広い地域に落ちた。
HSMT+鳥	その事故の残骸からは広い地域に落ちた。

表 5.12: “HSMT” と “HSMT+英” の対比較の出力例 (単文 英日)

HSMT の出力例	
入力文	I have recently come to like Latin music .
正解文	このごろラテン音楽が好きになった。
HSMT	私は最近ラテン語音楽が好きになってきた。
HSMT+英	私は音楽が最近ラテン語が好きになる。
HSMT+英 の出力例	
入力文	He has a keen intuition .
正解文	彼は鋭い勘をもっています。
HSMT	彼は鋭いintuitionしている。
HSMT+英	彼は鋭い洞察をしている。
差無しの出力例	
入力文	This industry had its beginning in 1864 .
正解文	この産業の始まりは1864年である。
HSMT	この業界の1864できた。
HSMT+英	この業界で1864始めた。

表 5.13: “HSMT”と“HSMT+鳥”の対比較の出力例 (重文複文 日英)

HSMT の出力例	
入力文	どうも そう 思え て しょうが ない。
正解文	I am strongly inclined to think so .
HSMT	I am strongly inclined to think so .
HSMT+鳥	I think so .
HSMT+鳥 の出力例	
入力文	彼 は 愉快 な 人 だ 。
正解文	He is a lot of fun .
HSMT	He is a very good time .
HSMT+鳥	He is a cheerful person .
差無し の出力例	
入力文	この 章 を 読み 終える まで 待っ て ください 。
正解文	Wait till I have finished this chapter .
HSMT	Please wait until I finish read this chapter .
HSMT+鳥	Please wait until I had finished reading this chapter .

表 5.14: “HSMT”と“HSMT+英”の対比較の出力例 (重文複文 日英)

HSMT の出力例	
入力文	生き物 を いじめ る の は 悪い 。
正解文	It is wrong to torment any creature .
HSMT	It is wrong teasing the living thing .
HSMT+英	Cruelty life is bad .
HSMT+英 の出力例	
入力文	警官 は 私 に 止まれ と 合図 した 。
正解文	The policeman gave me a sign to stop .
HSMT	The policeman me to stop .
HSMT+英	The policeman gave me a sign to stop .
差無し の出力例	
入力文	道路 を 横断 する とき は 車 に 注意 し なさい 。
正解文	Watch out for the traffic when you cross the street .
HSMT	Watch out for cars when you cross the street .
HSMT+英	Be careful for cars when you cross the street .

表 5.15: “HSMT”と“HSMT+鳥”の対比較の出力例 (重文複文 英日)

HSMT の出力例	
入力文	I think he made a good choice .
正解文	彼はよい選び方をしたと思う。
HSMT	彼がよい選択したと思う。
HSMT+鳥	彼を信用したのがいいと思います。
HSMT+鳥 の出力例	
入力文	They demolished the old factory and built a condominium .
正解文	彼らは古い工場を取り壊してマンションを建てた。
HSMT	彼らはその古い工場 demolished ているマンションを建てた。
HSMT+鳥	彼らは古い工場を壊してマンションを建てた。
差無し of 出力例	
入力文	I reflected upon what he had said .
正解文	彼が言ったことについてよく考えてみた。
HMST	私が言ったことを反映している。
HSMT+鳥	私は彼が言ったことを反映している。

表 5.16: “HSMT”と“HSMT+英”の対比較の出力例 (重文複文 英日)

HSMT の出力例	
入力文	I looked everywhere , but I could not find it .
正解文	方々探したが見つからなかった。
HSMT	あちこち捜したが、見つからなかった。
HSMT+英	あちこち捜したが、私は見つからなかった。
HSMT+英 の出力例	
入力文	We made the suggestion that she should resign .
正解文	彼女に辞めてはどうかと言った。
HSMT	彼女が辞職を提案した。
HSMT+英	彼女が辞職することを提案した。
差無し of 出力例	
入力文	I stepped cautiously on the fresh snow .
正解文	私は新雪を踏みしめて歩いた。
HSMT	雪の上に用心して止めた。
HSMT+英	雪を新たに用心した。



## 第6章 考察

### 6.1 HSMT と PSMT の比較

HSMT と PSMT の翻訳精度の比較を行うため，以下に示す人手評価を行った．単文の日英翻訳における結果を表 6.1, 単文の英日翻訳における結果を表 6.2, 重文複文の日英翻訳における結果を表 6.3, 重文複文の英日翻訳における結果を表 6.4 に示す．

表 6.1: 単文 日英翻訳

PSMT	HSMT	差なし	一致
3	5	70	22
PSMT+鳥	HSMT+鳥	差なし	一致
6	5	64	25

表 6.2: 単文 英日翻訳

PSMT	HSMT	差なし	一致
2	6	71	21
PSMT+鳥	HSMT+鳥	差なし	一致
3	8	72	17

表 6.3: 重文複文 日英翻訳

PSMT	HSMT	差なし	一致
5	8	77	10
PSMT+鳥	HSMT+鳥	差なし	一致
8	5	71	16

表 6.4: 重文複文 英日翻訳

PSMT	HSMT	差なし	一致
5	7	75	13
PSMT+鳥	HSMT+鳥	差なし	一致
7	3	82	8

### 6.1.1 HSMT と PSMT の比較のまとめ

表 6.1 から表 6.4 の結果より，対訳句の追加を行った場合は，PSMT が HSMT よりも翻訳精度が高くなった．ただし，単文の英日翻訳においては，HSMT に対訳句を追加した翻訳の方が高くなった．この結果より以下のことが言える．

- HSMT における学習データへの対訳句の追加の効果よりも，PSMT における学習データへの対訳句の追加の効果の方が高い．
- PSMT において学習データに対訳句を追加して行った翻訳の結果が，すべての翻訳結果の中で一番精度が高い．

以上に示した 2 点については，5.1 節の自動評価でも同様の傾向が見られた．

### 6.1.2 対比較の出力例

単文の日英翻訳における“PSMT+鳥”と“HSMT+鳥”の対比較の出力例を表 6.5, 単文の英日翻訳における“PSMT+鳥”と“HSMT+鳥”の対比較の出力例を表 6.6,, 重文複文の日英翻訳における“PSMT+鳥”と“HSMT+鳥”の対比較の出力例を表 6.7, 重文複文の英日翻訳における“PSMT+鳥”と“HSMT+鳥”の対比較の出力例を表 6.8 に示す．

表 6.5: “PSMT+鳥”と“HSMT+鳥”の対比較の出力例 (単文 日英)

PSMT+鳥 の出力例	
入力文	ほお に キス して。
正解文	Kiss me on the cheek .
PSMT+鳥	Give me a kiss on the cheek .
HSMT+鳥	She kissed her cheek .
HSMT+鳥 の出力例	
入力文	その 試合 は テレビ で 見 ました。
正解文	I watched the game on television .
PSMT+鳥	I saw TV in the game .
HSMT+鳥	I saw the game on television .
差無し of 出力例	
入力文	あ の 人 は 日 本 語 が ち ょ っ と で き ま す。
正解文	That person can speak a little Japanese .
PSMT+鳥	He is a bit of a Japanese .
HSMT+鳥	That person is a Japanese .

表 6.6: “PSMT+鳥”と“HSMT+鳥”の対比較の出力例 (単文 英日)

PSMT+鳥 の出力例	
入力文	The bill failed to reach the Senate floor .
正解文	その 法案 は 上 院 の 審 議 に 回 さ れ な か っ た。
PSMT+鳥	その 法案 は 上 院 届 か な か っ た。
HSMT+鳥	その 法案 は 上 院 を 床 に 失 敗 し た。
HSMT+鳥 の出力例	
入力文	This cold medicine is effective .
正解文	こ の 風 邪 薬 は よ く 効 く。
PSMT+鳥	こ の 寒 さ に 効 く 薬 は な い。
HSMT+鳥	こ の 薬 は 風 邪 に 効 果 的 だ。
差無し of 出力例	
入力文	We grouped ourselves round our teacher .
正解文	私 たち は 先 生 の ま わ り に 集 ま っ た。
PSMT+鳥	私 たち は 先 生 の 周 り に ま と も に 群 が っ た。
HSMT+鳥	私 たち は 先 生 の 自 分 の ま わ り に 集 ま っ た。

表 6.7: “PSMT+鳥”と“HSMT+鳥”の対比較の出力例 (重文複文 日英)

PSMT+鳥 の出力例	
入力文	我々は彼は無罪だと推定した。
正解文	We presumed his innocence .
PSMT+鳥	We presumed that he is innocent .
HSMT+鳥	We are that he is innocent .
HSMT+鳥 の出力例	
入力文	女が二人寄ればきっと着物の品定めだ。
正解文	When you see two women together , you may be sure they are discussing their own clothes .
PSMT+鳥	When two women get together , they will be in kimono .
HSMT+鳥	When two women get together , I am sure they discuss their clothes .
差無し of 出力例	
入力文	会議は予定より30分遅れて始められた。
正解文	The meeting was thirty minutes late in starting .
PSMT+鳥	The conference began thirty minutes late than expected .
HSMT+鳥	The meeting started 30 minutes behind schedule .

表 6.8: “PSMT+鳥”と“HSMT+鳥”の対比較の出力例 (重文複文 英日)

PSMT+鳥 の出力例	
入力文	It was good to see you again .
正解文	再びお会いできてうれしく存じます。
PSMT+鳥	また会えて嬉しく思います。
HSMT+鳥	またお会いしてよいことだった。
HSMT+鳥 の出力例	
入力文	They demolished the old factory and built a condominium .
正解文	彼らは古い工場を取りこわしてマンションを建てた。
PSMT+鳥	彼らはその古い工場はマンションを壊して建てた。
HSMT+鳥	彼らは古い工場を壊してマンションを建てた。
差無し of 出力例	
入力文	I was under the apprehension that you were John 's wife .
正解文	あなたはジョンの奥さんだと思いこんでいた。
PSMT+鳥	彼の奥さんれていると理解していました。
HSMT+鳥	君が彼の妻と理解していなかった。

## 6.2 鳥バンクと英辞郎の比較

鳥バンクと英辞郎の対訳句の追加の効果と比較するために以下に示す人手評価を行った。単文の日英翻訳における結果を表 6.9, 単文の英日翻訳における結果を表 6.10, 重文複文の日英翻訳における結果を表 6.11, 重文複文の英日翻訳における結果を表 6.12 に示す。

表 6.9: 単文 日英翻訳

HSMT+鳥	HSMT+英	差なし	一致
9	4	82	5
PSMT+鳥	PSMT+英	差なし	一致
7	10	73	10

表 6.10: 単文 英日翻訳

HSMT+鳥	HSMT+英	差なし	一致
8	8	81	3
PSMT+鳥	PSMT+英	差なし	一致
7	7	74	12

表 6.11: 重文複文 日英翻訳

HSMT+鳥	HSMT+英	差なし	一致
6	6	85	3
PSMT+鳥	PSMT+英	差なし	一致
10	8	80	2

表 6.12: 重文複文 英日翻訳

HSMT+鳥	HSMT+英	差なし	一致
6	6	86	2
PSMT+鳥	PSMT+英	差なし	一致
7	9	82	2

### 6.2.1 鳥バンクと英辞郎の比較のまとめ

表 6.9 から表 6.12 より, 英辞郎と鳥バンクの対訳句の追加における翻訳精度の差は見られなかった。

しかし, 対訳句数に着目すると, 英辞郎の対訳句数が 1,366,458 句に対して鳥バンクの対訳句数は 548,694 句であるので, 英辞郎の対訳句数の方が 2 倍以上も多い。それにも関わらず, 翻訳精度に差が生じなかった理由として, 鳥バンクの対訳句が重文複文コーパスから抽出されたものであるため, 実験で用いたテスト文との親和性が高かったことが要因のひとつではないかと考える。

## 6.2.2 対比較の出力例

単文の日英翻訳における“HSMT+鳥”と“HSMT+英”の対比較の出力例を表 6.13, 単文の英日翻訳における“HSMT+鳥”と“HSMT+英”の対比較の出力例を表 6.14,, 重文複文の日英翻訳における“HSMT+鳥”と“HSMT+英”の対比較の出力例を表 6.15, 重文複文の英日翻訳における“HSMT+鳥”と“HSMT+英”の対比較の出力例を表 6.16 に示す .

表 6.13: “HSMT+鳥”と“HSMT+英”の対比較の出力例 (単文 日英)

HSMT+鳥 の出力例	
入力文	住宅投資にも先細りの兆しが見える。
正解文	There are also signs that housing investments are tapering off .
HSMT+鳥	There are signs of decline in housing investments .
HSMT+英	Housing investment in some signs shrinking .
HSMT+英 の出力例	
入力文	彼はマラソンで脱落した。
正解文	He dropped away in a marathon race .
HSMT+鳥	He is in the marathon race .
HSMT+英	He made a hiatus in the marathon .
差無し of 出力例	
入力文	努力は水のあわとなった。
正解文	All my efforts came to nothing .
HSMT+鳥	I took the efforts of water .
HSMT+英	The efforts and took of water .

表 6.14: “HSMT+鳥”と“HSMT+英”の対比較の出力例 (単文 英日)

HSMT+鳥 の出力例	
入力文	His house confronts mine .
正解文	彼の家は私の家と向き合っている。
HSMT+鳥	彼の家は私のと向かい合っている。
HSMT+英	彼の家は私 confronts 。
HSMT+英 の出力例	
入力文	His temptation is irresistible .
正解文	彼の誘惑には抵抗しがたい。
HSMT+鳥	彼の誘惑は破竹の勢いであった。
HSMT+英	彼は誘惑には耐えられない。
差無し of 出力例	
入力文	There are various ranks in the military .
正解文	軍隊にはいろいろの階級がある。
HSMT+鳥	その軍事ではいろいろながあった。
HSMT+英	軍事にはいろいろの列がある。

表 6.15: “HSMT+鳥”と“HSMT+英”の対比較の出力例 (重文複文 日英)

HSMT+鳥 の出力例	
入力文	方々探したが見つからなかった。
正解文	I looked everywhere , but I could not find it .
HSMT+英	I looked everywhere , but I could not find it .
HSMT+鳥	I could not find the freewheeled hunting .
HSMT+英 の出力例	
入力文	世間の口はうるさいものだ。
正解文	People will talk .
HSMT+鳥	in the world is noisy .
HSMT+英	It is noisy mouth of the world .
差無し of 出力例	
入力文	彼は足を引きずってやっとのことで床についた。
正解文	He dragged himself to bed .
HSMT+鳥	He trudged and on the floor .
HSMT+英	He shuffled I tried to the floor .

表 6.16: “HSMT+鳥”と“HSMT+英”の対比較の出力例 (重文複文 英日)

HSMT+鳥 の出力例	
入力文	Such action admits of no excuse .
正解文	そんな事をしては弁解の余地がない。
HSMT+鳥	そんな事をしては言い訳が立たない。
HSMT+英	そのような行動は言い訳にはならない。
HSMT+英 の出力例	
入力文	They were on strike demanding higher wages .
正解文	彼らは賃銀の増額を要求してストライキを行なっていた。
HSMT+鳥	彼らは賃上げを要求している。
HSMT+英	彼らは賃上げを要求してストにしていた。
差無しの出力例	
入力文	A horse whisks off the flies .
正解文	馬が尾を振って蠅を払う。
HSMT+鳥	馬が尾を振って蠅をした。
HSMT+英	馬が whisks ハエを追い払った。

### 6.3 語彙の増加が生じない対訳句追加

本研究で用いた，鳥バンクの対訳句 548,694 句は本実験で用いた重文複文の日英対訳文 100,000 文から抽出した対訳句である．よって，5 章の実験の重文複文における鳥バンクの対訳句を追加して行った日英翻訳と英日翻訳の実験では，語彙の増加が生じない対訳句追加となる．しかし，5 章の自動評価，人手評価ともに，翻訳精度が向上している．この結果より，学習データにおいて語彙の増加が生じない場合であっても対訳句の追加が有効であると考えられる．



## 第7章 分野外実験

鳥バンクは重文複文コーパスから抽出した対訳句であるので，重文複文コーパスとの親和性が高い．そこで，重文複文の実験において，新たに分野外のオープンなデータである 3,950 文の重文複文コーパスをテスト文として用いて，分野外の実験を行った．

### 7.1 分野外実験の内容

分野外実験では，4.6 節の実験内容で示した項目の日英対訳文として重文複文のみを扱い実験を行う．したがって，合計 8 種類の翻訳実験を行う．

### 7.2 分野外実験のデータ

学習データとディベロップメントデータとして 4.1 節で示したものと同じものを使用し，テストデータとして新たに用意した重文複文 3,950 文を用いる．

### 7.3 分野外実験結果

#### 7.3.1 自動評価結果

自動評価における，日英翻訳の結果を表 7.1，英日翻訳の結果を表 7.2 に示す．

表 7.1: 重文複文 日英翻訳

	BLEU	NIST	METEOR	RIBES
HSMT	0.0834	3.8391	0.3759	0.6314
HSMT+鳥	0.0941	4.1222	0.4016	<b>0.6453</b>
HSMT+英	0.0861	4.0479	0.3893	0.6386
PSMT	0.0812	3.8039	0.3687	0.6182
PSMT+鳥	0.0928	4.1638	<b>0.4034</b>	0.6359
PSMT+英	<b>0.0942</b>	<b>4.1460</b>	0.3981	0.6378

表 7.2: 重文複文 英日翻訳

	BLEU	NIST	METEOR	RIBES
HSMT	0.1007	3.5336	0.3556	0.6683
HSMT+鳥	0.0996	3.5086	0.3532	0.6608
HSMT+英	<b>0.1050</b>	<b>3.5956</b>	<b>0.3613</b>	<b>0.6718</b>
PSMT	0.0924	3.3615	0.3377	0.6458
PSMT+鳥	0.0979	3.4974	0.3484	0.6428
PSMT+英	0.0980	3.5486	0.3535	0.6497

### 7.3.2 人手評価

実験結果として人手評価を示す。日英翻訳の結果を表 7.3，英日翻訳の結果を表 7.4 に示す。

表 7.3: 重文複文 日英翻訳

HSMT	HSMT+鳥	差なし	一致
4	6	88	2
HSMT	HSMT+英	差なし	一致
4	8	88	0
PSMT	PSMT+鳥	差なし	一致
5	10	81	4
PSMT	PSMT+英	差なし	一致
5	9	84	2

表 7.4: 重文複文 英日翻訳

HSMT	HSMT+鳥	差なし	一致
4	12	82	2
HSMT	HSMT+英	差なし	一致
6	12	79	3
PSMT	PSMT+鳥	差なし	一致
2	11	87	0
PSMT	PSMT+英	差なし	一致
1	7	91	1

### 7.3.3 分野外実験結果のまとめ

実験結果より，HSMT においても対訳句の追加における翻訳精度の向上が見られた．また，HSMT と PSMT のそれぞれにおける学習データへの対訳句の追加の効果を比べた場合，PSMT における学習データへの対訳句の追加の効果の方が高かった．これは，6.1 節の結果と同様の傾向である．

### 7.3.4 対比較の出力例

日英翻訳における“HSMT”と“HSMT+鳥”の対比較の出力例を表 7.5, 英日翻訳における“HSMT”と“HSMT+鳥”の対比較の出力例を表 7.6 に示す．

表 7.5: “HSMT” と “HSMT+鳥” の対比較の出力例 (分野外の重文複文 日英)

HSMT の出力例	
入力文	残念ですが、回復の見込みはほとんどありません。
正解文	I am afraid that there is almost no chance of it recovering .
HSMT	I regret that I have little hope of recovery .
HSMT+鳥	I am sorry , but I do not have the little hope of recovery .
HSMT+鳥 の出力例	
入力文	どういうふうに説明していいのかわからない。
正解文	I don 't know how to explain it .
HSMT	I do not know what you explain in this way .
HSMT+鳥	I do not know how to explain the word .
差無しの出力例	
入力文	温泉に入るのがなによりの楽しみだ。
正解文	Getting into a hot spa is my greatest pleasure , above everything else .
HSMT	We look forward to enter the hot spring bonus to me .
HSMT+鳥	I got into a hot spring than anything else .

表 7.6: “HSMT” と “HSMT+鳥” の対比較の出力例 (分野外の重文複文 英日)

HSMT の出力例	
入力文	She was stricken with sorrow after she lost her family .
正解文	家族を失って、彼女は悲しみにうちひしがれている。
HSMT	彼女は家族を失って悲しみに胸を裂かれるような思いがした。
HSMT+鳥	悲しみに胸を裂かれるような思いがして彼女は彼女の家を失った。
HSMT+鳥 の出力例	
入力文	The summary of this paper 's conclusion is as follows .
正解文	本稿の結論をまとめれば、次のようになる。
HSMT	この紙は以下の結論を約言すれば次のとおりである。
HSMT+鳥	この論文の結論はの概要を約言すれば次のとおりである。 。
差無しの出力例	
入力文	The sun does not rise before 7 o 'clock .
正解文	7 時頃になってようやく陽が昇る。
HSMT	太陽はあ 7 お 'clock 前にふえない。
HSMT+鳥	お 'clock 7 日前にしてはならない。

## 第8章 おわりに

本研究では，提案手法として対訳句コーパスを日英対訳コーパスに追加したコーパスを学習データとして使用し，HSMTを行うことによって，PSMTの場合と同様に対訳句コーパスの効果が得られるかを調査した．翻訳手法として，HSMTとPSMTを用い，対訳句コーパスとして鳥バンクと英辞郎を用い，単文コーパスと重文複文コーパスに対して日英統計翻訳と英日統計翻訳をそれぞれ行った．したがって合計16種類の翻訳実験を行った．

その結果，全ての提案手法の自動評価結果が向上した．また，人手評価法においても提案手法での翻訳結果の精度が高いことを示すことができ，提案手法の有効性を示すことができた．また，PSMTへの対訳句の追加がHSMTへの対訳句の追加よりも，高い効果を示していた．さらに，PSMTにおいて学習データに対して対訳句を追加した翻訳の結果が，すべての実験結果の中で一番高い精度を示した．また，鳥バンクと英辞郎の対訳句の追加における翻訳精度の差は見られなかった．

今後は，今回の実験で得られた結果の誤り分析を行い，さらなる精度の向上を目指し調査を行いたい．

# 謝辞

最後に，1年間に渡りご指導いただきました鳥取大学工学部知能情報工学科計算機工学講座C研究室の村田真樹教授，村上仁一准教授，徳久雅人講師そして計算機工学講座C研究室の方々に厚く御礼申し上げます．

また，参考にさせていただいた論文の著者の方々に対して深く感謝申し上げます．

## 参考文献

- [1] Popović Maja, and Ney Hermann “Statistical Machine Translation with a small amount of bilingual training data”, 5th LREC SALT MIL Workshop on Minority Languages. 2006.
- [2] Francis M Tyers “Rule-based augmentation of training data in Breton-French statistical machine translation”, 13th Annual Conference of the European Association for Machine Translation, pp.213-217. 2009.
- [3] 日野聡子：“日英統計翻訳における対訳句コーパスの効果”，2013年修論
- [4] 村上仁一，藤波進 “日本語と英語の対訳文対の収集と著作権の考察”，第一回コーパス日本語学ワークショップ, pp.119-130. 2012.
- [5] 鳥バンク：<http://unicorn.ike.tottori-u.ac.jp/toribank/>
- [6] 英辞郎：<http://www.alc.co.jp/>
- [7] Peter F.Brown, Stephen A.Della Pietra, Vincent J.Della Pietra, Robert L.Mercer, “ The mathematics of statistical machine translation:Parameter Estimation ”, Computational Linguistics, 1993.
- [8] GIZA++：<http://www.fjoch.com/GIZA++>
- [9] Moses: Philipp Koehn, Marcello Federico, Brooke Cowan, Richard Zens, Chris Dyer, Ondrej Bojar, Alexandra Constantin, Evan Herbst, “Moses: Open Source Toolkit for Statistical Machine Translation”, Proceedings of the ACL 2007 Demo and Poster Sessions, pp.177-180, 2007.
- [10] BLEU: Papineni Kishore, Salim Roukos, Todd Ward, Wei-Jing Zhu, “BLEU: a method for automatic evaluation of machine translation”, 40th Annual meeting of the Association for Computational Linguistics, pp.311-318, 2002.

- [11] NIST: “Automatic Evaluation of Machine Translation Quality Using n-gram Co-Occurrence Statistics” Proceedings of the Human Language Technology Conference (HLT), pp.128-132. 2002.
- [12] METEOR: Banerjee Satanjeev, Lavie Alon, “METEOR: An Automatic Metric for MT Evaluation with Improved Correlation with Human Judgments”, Proceedings of Workshop on Intrinsic and Extrinsic Evaluation Measures for MT and/or Summarization at the 43th Annual Meeting of the Association of Computational Linguistics (ACL-2005), pp.65-72, 2005.
- [13] RIBES: 平尾努, 磯崎秀樹, Kevin Duh, 須藤克仁, 塚田元, 永田昌明, “RIBES: 順位相関に基づく翻訳の自動評価法”, 言語処理学会第 17 年次大会発表論文集, pp.1111-1114, 2011.
- [14] MeCab: Taku Kudo, Kaoru Yamamoto, Yuji Matsumoto, “Applying Conditional Random Fields to Japanese Morphological Analysis”, Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP-2004), pp.230-237, 2004.
- [15] tokenizer.sed  
<http://www.cis.upenn.edu/treebank/tokenizer.sed>
- [16] SRILM: Andreas Stolcke, “SRILM - an Extensible Language Modeling Toolkit”, 7th International Conference on Spoken Language Processing, pp.901-904, 2002.
- [17] Mert: Franz Josef Och: “Minimum Error Rate Training in Statistical Machine Translation”, In Proceeding of the 41st Annual Meeting of the Association for Computational Linguistics, pp.160-167, 2003.