

概要

近年の機械翻訳では，対訳データから自動的に機械翻訳システムを構築する統計翻訳が主流となっている．しかし，統計翻訳において，様々な問題点がある．その問題点の一つとして，主語が省略されている日本語文の日英翻訳が挙げられる．日本語において，主語が省略されている日本語文が多く存在する．猪澤ら [1] は，この主語が省略されている日本語文の日英翻訳は，翻訳精度が低いということを報告している．

そこで本研究では，統計翻訳を行う前に，主語が省略されている日本語文に対して主語を付与し，どのように翻訳精度が変化するか調査した．結果として，自動評価と人手評価の両方において翻訳精度が向上し，統計翻訳において，主語が省略されている文に対し主語を付与する手法の有効性を確認できた．

目次

1	はじめに	1
2	日英統計翻訳システム	2
2.1	統計翻訳の概要	2
2.2	翻訳モデル	3
2.2.1	フレーズテーブル	3
2.2.2	IBM モデル	3
2.2.3	フレーズテーブルの作成法	9
2.3	言語モデル	13
2.4	デコーダ	13
2.5	パラメータチューニング	14
3	関連研究	15
3.1	統計翻訳の問題点	15
3.2	主語補完における先行研究	16
4	“主語省略文”	17
4.1	主語の定義	17
4.2	命令文	17
4.3	“主語省略文”の判断条件	18
5	主語補完	20
5.1	日本語補完	20
5.2	対訳補完	24
5.3	train 補完	26
6	実験方法	27
6.1	実験手順	27
6.2	実験の種類	28
7	実験環境	29
7.1	デコーダのパラメータ	29
7.2	翻訳モデルの学習	29

7.3	言語モデルの学習	29
7.4	実験データ	29
8	評価方法	31
8.1	自動評価	31
8.2	人手評価	33
9	実験結果	34
9.1	主語補完結果	34
9.2	自動評価結果	34
9.3	人手評価結果	36
9.3.1	対比較評価結果	36
9.3.2	対比較評価例文	36
10	考察	39
10.1	英語文の文頭調査	39
10.2	日本語補完についての考察	40
10.2.1	“主語省略文”抽出の精度	40
10.2.2	日本語補完における主語補完の精度	41
10.2.3	日本語補完における自動評価の解析	43
10.2.4	日本語補完における人手評価の解析	43
10.3	対訳補完についての考察	46
10.3.1	対訳補完における主語補完の精度	46
10.3.2	対訳補完における自動評価の解析	47
10.4	train 補完についての考察	47
11	今後の研究	48
12	おわりに	49

目 次

1	日英統計翻訳手順	2
2	デコーダ動作例	14
3	主語補完手順	22
4	文頭単語抽出例 1	24
5	文頭単語抽出例 2	25
6	翻訳手順	27

表 目 次

1	フレーズテーブルの例	3
2	日英方向の単語対応	9
3	英日方向の単語対応	9
4	intersection の例	10
5	union の例	10
6	grow の例	11
7	grow-diag の例	11
8	grow-diag-final の例	12
9	grow-diag-final-and	12
10	言語モデルの例	13
11	翻訳精度が低下した例	15
12	出力文が用いたフレーズ対 (単語区切りフレーズ対の問題)	15
13	文内照応の例	16
14	“主語省略文”の例	17
15	命令文の例	17
16	“主語省略文”の判断条件	18
17	“主語省略文”の判断条件を満たす文の例	18
18	“主語省略文”の判断条件を満たさない文の例	19
19	形態素解析結果の例	20
20	抽出される文の例	21
21	抽出されない文の例 1	21
22	抽出されない文の例 2	21
23	主語補完例	22
24	主語補完例	23
25	対訳補完 “主語省略文”例	24
26	文頭単語の日本語への変換規則	25
27	主語補完例	26
28	主語補完例	26
29	単文コーパス例文	30
30	正解文例	31

31	正解文例	33
32	主語補完文数	34
33	10,000 文での自動評価結果	34
34	1,526 文での自動評価結果	35
35	1,109 文での自動評価結果	35
36	対比較評価結果	36
37	日本語補完が優れていると判断した例 1	37
38	日本語補完が優れていると判断した例 2	37
39	ベースラインが優れていると判断した例 1	37
40	ベースラインが優れていると判断した例 2	38
41	差なしと判断した例 1	38
42	差なしと判断した例 2	38
43	対訳英語文の文頭単語	39
44	“主語省略文”抽出評価結果	40
45	“主語省略文”抽出不適切文例	40
46	“主語省略文”抽出評価結果	41
47	“主語省略文”抽出不適切文例	41
48	主語補完評価結果	42
49	主語補完誤り文例	42
50	主語補完による主語出力改善の例	43
51	主語補完によるフレーズ対応改善の例	44
52	フレーズ対応改善の例	44
53	主語補完の誤りによる翻訳精度の低下	44
54	主語補完の誤りによる翻訳精度の低下	45
55	主語補完評価結果	46
56	主語補完誤り文例	46

1 はじめに

機械翻訳は、人手によって翻訳ルールを記述し、翻訳を行うルールベース翻訳が用いられていた。しかし、ルールベース翻訳は、人手で翻訳ルールを記述することに大きなコストがかかる。また、多言語間で翻訳を行う場合、言語間ごとに翻訳ルールを記述する必要があり、多言語化が困難であるという問題点があった。そこで、近年の機械翻訳では、対訳データから自動的に機械翻訳システムを構築する統計翻訳が主流となっている。統計翻訳は、対訳データがあれば、自動的に翻訳規則を作成できるため、コストが小さく、また多言語化も容易に行うことができる。

この統計翻訳において、様々な問題点がある。その問題点の一つとして、主語が省略されている日本語文の日英翻訳が挙げられる。日本語において、主語が省略されている日本語文が多く存在する。猪澤ら [1] は、この主語が省略されている日本語文の日英翻訳は、翻訳精度が低いということを報告している。しかし、統計翻訳において、主語が省略されている日本語文に主語を付与し、解析を行っている実験は見られない。

そこで本研究では、統計翻訳を行う前に、主語が省略されている日本語文に対して主語を付与する(主語補完と呼ぶ)。この主語補完による翻訳精度を調査し、主語補完の有効性を確認する。調査の結果、主語補完を行わない翻訳と比較し、主語補完を行った翻訳では、単語対応の改善が見られ、翻訳精度が向上した。この調査結果より、主語補完の有効性が確認できた。

2章において、日英統計翻訳システムについての説明を行う。3章において、主語が省略されている文において述べている研究と、主語補完を行っている研究について述べる。4章において、主語が省略されている文の定義を行う。5章において、主語補完方法を説明する。6章において、行った実験の種類と、実験の流れを説明する。7章において、実験環境についての説明を行う。8章において、実験の評価方法について説明する。9章において、主語補完の有効性を述べる。10章において、解析および考察を行う。12章において、結論を述べる。

2 日英統計翻訳システム

2.1 統計翻訳の概要

統計翻訳とは、大量の対訳コーパス(原言語と目的言語の対訳文)より、自動的に翻訳規則を作成し翻訳を行う翻訳システムである。そのため、統計翻訳では、翻訳精度が対訳コーパスの量に大きく依存する。しかし、単言語コーパス(単言語のみのコーパス)は大量に入手しやすいが、対訳コーパスは作成に膨大なコストがかかるため、入手しにくい。図1に日英統計翻訳の手順を示す。

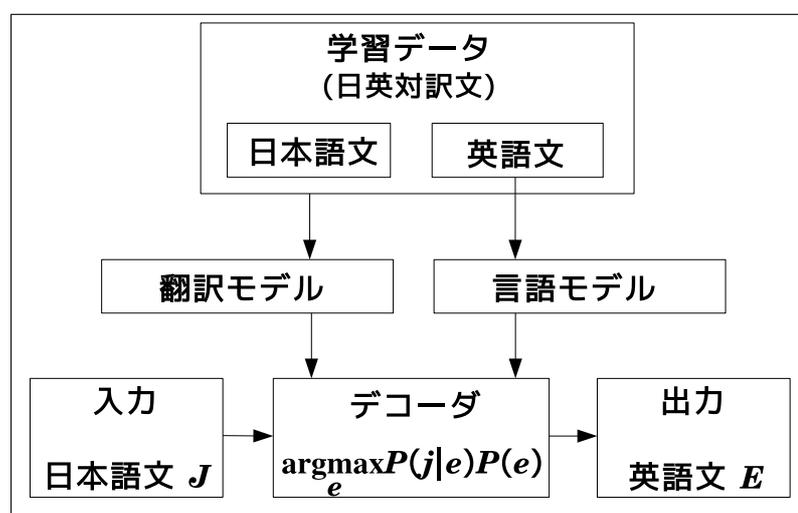


図 1: 日英統計翻訳手順

日英統計翻訳は、入力文として日本語文 J が与えられた場合に、翻訳モデルと言語モデルの組み合わせの中から確率が最大となる英語文 E を探索することで翻訳を行う。以下にその基本モデルを示す。

$$E = \arg \max_e P(e|j) \quad (1)$$

$$\simeq \arg \max_e P(j|e)P(e) \quad (2)$$

ここで、 $P(j|e)$ は翻訳モデルであり、 $P(e)$ は言語モデルである。翻訳モデルは、対訳コーパスから学習し、言語モデルは目的言語の単言語コーパスから学習する。また、デコーダを用いて、 $P(j|e)P(e)$ が最大となる E を探索する。デコーダの動作については、2.4 節で説明する。

2.2 翻訳モデル

翻訳モデルには，大きくわけて単語ベースの翻訳モデルと句ベースの翻訳モデル [3] がある．初期の統計翻訳は，単語ベースの翻訳モデルを用いていた．しかし，単語ベースの翻訳モデルと比較し，句ベースの翻訳モデルは翻訳精度が高いため，現在は句ベースの翻訳モデルが主流となっている．句ベースの翻訳モデルは，日本語の単語列から英語の単語列へ確率的に翻訳を行うためのモデルであり，対訳コーパスより学習される．以下に，句ベースの翻訳モデルの作成手順を示す．

手順 1 GIZA++[2] を用いて，IBM モデルを推定する．IBM モデルについては，詳細を 2.2.2 節で説明する．

手順 2 IBM モデルから，ヒューリスティックなルールに基づき，句ベースの翻訳モデルを作成する．作成方法については，2.2.3 節で説明する．

この句ベースの翻訳モデルについて，詳細を以下に示す．

2.2.1 フレーズテーブル

句ベースの翻訳モデルは，フレーズテーブルと呼ばれる表で管理されている．表 1 にフレーズテーブルの例を示す．

表 1: フレーズテーブルの例

あの歌		That song		1	0.20	1	0.16
おもちゃ		toy		0.51	0.57	0.30	0.31
彼は		He is		0.23	0.05	0.04	0.01

左から，日本語フレーズ，英語フレーズ，フレーズの英日方向の翻訳確率 $P(j|e)$ ，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となっている．

2.2.2 IBM モデル

統計翻訳の手法として代表的なものに，IBM の仏英翻訳モデルがある．この IBM モデルは model1 から model5 までの 5 つのモデルからなっている．IBM 翻訳モデルではフランス語から英語への翻訳を想定しているため，以下の説明は仏英翻訳を前提とする．IBM

モデルでは英語文 E , フランス語文 F の翻訳モデル $P(F|E)$ を計算するため, ある仏単語 f と英単語 e の対応関係であるアライメント a を用い, 以下の式を用いる.

$$P(F|E) = \sum_a P(F, a|E) \quad (3)$$

アライメント a は, IBM モデルでは, 英単語は仏単語と 1:n の対応を持ち, 仏単語は一つの英単語のみと対応すると仮定する. また, 英語文頭に特殊文字 e_0 があるとし, 仏単語 f が適切な英単語と対応しない場合は, この e_0 と対応付ける.

モデル 1

式 (3) は以下の式に分解することが出来る. m はフランス語文の長さ, a_1^{j-1} はフランス語文における, 1 番目から $j-1$ 番目までのアライメント, f_1^{j-1} はフランス語文における, 1 番目から $j-1$ 番目まで単語を表している.

$$P(F, a|E) = P(m|E) \prod_{j=1}^m P(a_j | a_1^{j-1}, f_1^{j-1}, m, E) P(f_j | a_1^j, f_1^{j-1}, m, E) \quad (4)$$

このままでは複雑であるため計算が困難である. そこで, モデル 1 では以下のように仮定し, 簡略化を行う.

- フランス語文の長さの確率は m, E に依存しないとし, 定数 ϵ で置く

$$P(m|E) = \epsilon$$

- アライメントの確率は英語文の長さ l のみに依存する

$$P(a_j | a_1^{j-1}, f_1^{j-1}, m, E) = (l+1)^{-1}$$

- フランス語の翻訳確率 $t(f_j | e_{a_j})$ は, 仏単語 f_j に対応付けられる英単語 e_{a_j} のみに依存する

$$P(f_j | a_1^j, f_1^{j-1}, m, e) = t(f_j | e_{a_j})$$

以上の簡略化を行うことで、 $P(F, a|E)$ と $P(F, E)$ は以下の式で表される。

$$P(F, a|E) = \frac{\epsilon}{(l+1)^m} \prod_{j=1}^m t(f_j|e_{a_j}) \quad (5)$$

$$P(F|E) = \frac{\epsilon}{(l+1)^m} \sum_{a_1=0}^l \cdots \sum_{a_m=0}^l \prod_{j=1}^m t(f_j|e_{a_j}) \quad (6)$$

$$= \frac{\epsilon}{(l+1)^m} \prod_{j=1}^m \sum_{i=0}^l t(f_j|e_{a_j}) \quad (7)$$

モデル1では翻訳確率 $t(f|e)$ の初期値が0以外の場合、以下のようなEMアルゴリズムを繰り返し行うことで得られる期待値を用いて最適解を推定する。

1. 翻訳確率 $t(f|e)$ に適当な初期値を設定する。
2. 仏英対訳対 $(F^{(s)}, E^{(s)})$ (但し, $1 \leq s \leq S$) において、仏単語 f と英単語 e が対応する回数の期待値を以下の式により計算する。

$$c(f|e; F, E) = \frac{t(f|e)}{t(f|e_0) + \cdots + t(f|e_l)} \sum_{j=1}^m \delta(f, f_j) \sum_{i=0}^l \delta(e, e_i) \quad (8)$$

$\delta(f, f_j)$ はフランス語文 F 中で仏単語 f が出現する回数、 $\delta(e, e_i)$ は英語文 E 中で英単語 e が出現する回数を表している。

3. 英語文 $E^{(s)}$ の中で1回以上出現する英単語 e に対して、翻訳確率 $t(f|e)$ を計算する。
 1. 定数 λ_e を以下の式により計算する。

$$\lambda_e = \sum_f \sum_{s=1}^S c(f|e; F^{(s)}, E^{(s)}) \quad (9)$$

2. (9) 式より求めた λ_e を用いて、翻訳確率 $t(f|e)$ を再計算する。

$$\begin{aligned} t(f|e) &= \lambda_e^{-1} \sum_{s=1}^S c(f|e; F^{(s)}, E^{(s)}) \\ &= \frac{\sum_{s=1}^S c(f|e; F^{(s)}, E^{(s)})}{\sum_f \sum_{s=1}^S c(f|e; F^{(s)}, E^{(s)})} \end{aligned} \quad (10)$$

4. 翻訳確率 $t(f|e)$ が収束するまで手順2と手順3を繰り返す。

モデル2

モデル1では、全ての単語の対応に対して、単語対応の確率を一定としているが、これは現実的ではない。そこでモデル2では、 j 番目の仏単語 f_j と対応する英単語の位置 a_j は英語文の長さ l に加えて、 j と、フランス語文の長さ m に依存するとし、以下のよう仮定する。

$$a(a_j|j, m, l) \equiv P(a_j|a_1^{j-1}, f_1^{j-1}, m, l) \quad (11)$$

この仮定より、モデル1における (6) 式は、以下の式に変換できる。

$$P(F|E) = \epsilon \sum_{a_1=0}^l \cdots \sum_{a_m=0}^l \prod_{j=1}^m t(f_j|e_{a_j}) a(a_j|j, m, l) \quad (12)$$

$$= \epsilon \prod_{j=1}^m \sum_{i=0}^l t(f_j|e_{a_j}) a(a_j|j, m, l) \quad (13)$$

モデル2では、期待値は $c(f|e; F, E)$ と $c(i|j, m, l; F, E)$ の2つが存在する。以下の式から求められる。

$$c(f|e; F, E) = \frac{t(f|e)}{t(f|e_0) + \cdots + t(f|e_l)} \sum_{j=1}^m \delta(f, f_j) \sum_{i=1}^l \delta(e, e_i) \quad (14)$$

$$= \sum_{j=1}^m \sum_{i=0}^l \frac{t(f|e) a(i|j, m, l) \delta(f, f_j) \delta(e, e_i)}{t(f|e_0) a(0|j, m, l) + \cdots + t(f|e_l) a(l|j, m, l)} \quad (15)$$

$$c(i|j, m, l; F, E) = \sum_a P(a|E, F) \delta(i, a_j) \quad (16)$$

$$= \frac{t(f_j|e_i) a(i|j, m, l)}{t(f_j|e_0) a(0|j, m, l) + \cdots + t(f_j|e_l) a(l|j, m, l)} \quad (17)$$

$c(f|e; F, E)$ は対訳文中の英単語 e と仏単語 f が対応付けされる回数の期待値を表し、 $c(i|j, m, l; F, E)$ は英単語の位置 i が仏単語の位置 j に対応付けされる回数の期待値を表している。

モデル2は、複数の極大値を持つため、最適解が得られない可能性がある。モデル1では $a(i|j, m, l) = l + 1^{-1}$ となるモデル2の特殊な場合であると考えられる。モデル1は最適解に必ず収束するため、モデル1を用いることで最適解を得ることができる。

モデル3

モデル3は、英語単語と仏単語の対応は1対1の場合のみを想定していたモデル1、モデル2とは異なり、1つの単語が複数対応する単語の繁殖数や単語の翻訳位置の歪みについて考慮する。またモデル3では単語の位置を絶対位置として考える。モデル3では以下の3つのパラメータを用いる。

- 翻訳確率 $P(f|e)$
英単語 e が仏単語 f に翻訳される確率
- 繁殖確率 $n(\phi|e)$
英単語 e が ϕ 個の仏単語と対応する確率
- 歪み確率 $d(j|i, m, l)$
英語文の長さ l 、フランス語文の長さ m のとき、 i 番目の英単語 e_i が j 番目の仏単語 f_j に翻訳される確率

さらに、英単語が仏単語に翻訳されない個数を ϕ_0 とし、その確率 p_0 を以下の式で求める。このとき、歪み確率は $\frac{1}{\phi_0!}$ で、 $p_0 + p_1 = 1$ で p_0, p_1 は0より大きいとする。

$$P(\phi_0|\phi_1^l, E) = \binom{\phi_1 + \dots + \phi_l}{\phi_0} p_0^{\phi_1 + \dots + \phi_l - \phi_0} p_1^{\phi_0} \quad (18)$$

以上により、モデル3は以下の式で求められる。

$$\begin{aligned} P(F|E) &= \sum_{a_1=0}^l \dots \sum_{a_m=0}^l P(F, a|E) \\ &= \sum_{a_1=0}^l \dots \sum_{a_m=0}^l \binom{m - \phi_0}{\phi_0} p_0^{m - 2\phi_0} p_1^{\phi_0} \prod_{i=1}^l \phi_i! n(\phi_i|e_i) \\ &\quad \times \prod_{j=1}^m t(f_j|e_{a_j}) d(j|a_j, m, l) \end{aligned} \quad (19)$$

$$\quad (20)$$

モデル3では、全てのアライメントを計算するため、計算量が膨大となる。そのため、期待値を近似により求める。

モデル4

モデル4では、モデル3と異なり、単語の位置を絶対位置ではなく、相対位置で考える。またモデル3では各単語の位置が考慮されていない。例えば、各言語間での形容詞と名詞の関係である。このような場合を考慮し、モデル4では歪み確率 $d(j|i,m,l)$ を2つの場合で考える。

- 繁殖数が1以上である英単語に対応する仏単語の中で、最も文頭に近い場合

$$P(\Pi_{[i]1} = j | \pi_1^{[i]-1}, \tau_0^l, \phi_0^l, E) = d_1(j - \odot_{i-1} | \mathcal{A}(e_{[i-1]}), \mathcal{B}(f_j)) \quad (21)$$

\odot_{i-1} は $i-1$ 番目の英単語に対応する仏単語の位置を表している。

- それ以外の場合

$$P(\Pi_{[i]k} = j | \pi_{[i]1}^{k-1}, \pi_1^{[i]-1}, \tau_0^l, \phi_0^l, E) = d_{>1}(j - \pi_{[i]k-1} | \mathcal{B}(f_j)) \quad (22)$$

$\pi_{[i]k-1}$ は同じ英単語に対応している直前の仏単語を表している。

モデル5

モデル4では、単語の位置に関して直前の単語以外は考慮されていない。したがって、複数の単語が同じ位置に生じたり、単語の存在しない位置が生成される。このような問題を避けるために、モデル5では単語を空白部分に配置するよう改善が施されている。

2.2.3 フレーズテーブルの作成法

単語対応の作成

日本語と英語の単語対応の作成には，2.2.2 節で説明した IBM モデルを用いる．IBM モデルは，GIZA++[2] を用いて推定する．まず，英日方向，日英方向の両方向に対し，単語対応を作成する．表 2 に日英方向の単語対応を示す．また，表 3 に英日方向の単語対応を示す．

表 2: 日英方向の単語対応

	I	go	to	see	a	movie
私	■					
は					■	
映画						■
を			■			
見				■		
に				■		
行く		■				

表 3: 英日方向の単語対応

	I	go	to	see	a	movie
私	■					
は						
映画						■
を					■	
見				■		
に						
行く		■	■			

intersection と union

“intersection”(積集合)は，日英方向と英日方向の両方向に対応がある場合，両方向の対応点とする．表 4 に intersection の例を示す．

表 4: intersection の例

	I	go	to	see	a	movie
私	■					
は						
映画						■
を						
見				■		
に						
行く		■				

また，“union”(和集合)は，日英方向と英日方向の両方向の対応点全てを対応点とする．表 5 に union の例を示す．

表 5: union の例

	I	go	to	see	a	movie
私	■					
は					■	
映画						■
を			■		■	
見				■		
に				■		
行く		■	■			

grow

“grow”は，intersection の対応点の縦横方向に union の対応点があれば，その対応点を intersection の対応点に追加させていく．grow の例を表 6 に示す．

表 6: grow の例

	I	go	to	see	a	movie
私	■					
は						
映画						■
を						
見				■		
に				■		
行く		■	■			

grow-diag

“grow-diag”は，intersection の対応点の縦横方向，または対角方向に union の対応点があれば，その対応点を intersection の対応点に追加させていく．grow-diag の例を表 7 に示す．

表 7: grow-diag の例

	I	go	to	see	a	movie
私	■					
は					■	
映画						■
を			■		■	
見				■		
に				■		
行く		■	■			

grow-diag-final

“grow-diag-final”は，grow-diag で，日英両方向のうち少なくとも片方の単語対応がない場合，union に対応点があれば追加させていく．grow-diag-final の例を表 8 に示す．

表 8: grow-diag-final の例

	I	go	to	see	a	movie
私	■					
は					■	
映画						■
を			■		■	
見				■		
に				■		
行く		■	■			

grow-diag-final-and

“grow-diag-final”は，grow-diag で，日英両方向の単語対応がない場合，union に対応点があれば追加させていく．grow-diag-final-and の例を表 9 に示す．

表 9: grow-diag-final-and

	I	go	to	see	a	movie
私	■					
は					■	
映画						■
を			■		■	
見				■		
に				■		
行く		■	■			

フレーズ対応の抽出

両方向の単語対応から，フレーズ対応が矛盾しないようにフレーズ対応を抽出する．そして，抽出したフレーズ対応に対し，翻訳確率を計算し付与することで，フレーズテーブルを作成する．翻訳確率について，日本語フレーズ J_{phrase} と英語フレーズ E_{phrase} からなるフレーズ対応の確率を求める式を以下に示す．

$$P(J_{phrase} | E_{phrase}) = \frac{\text{学習データ中で } J_{phrase} \text{ と } E_{phrase} \text{ が同時に出現した数}}{\text{学習データ中で } E_{phrase} \text{ が出現した数}} \quad (23)$$

$$P(E_{phrase} | J_{phrase}) = \frac{\text{学習データ中で } J_{phrase} \text{ と } E_{phrase} \text{ が同時に出現した数}}{\text{学習データ中で } J_{phrase} \text{ が出現した数}} \quad (24)$$

2.3 言語モデル

言語モデルは、単語列が生じる確率を付与するモデルである。翻訳モデルで生成された翻訳候補から英語として自然な文を選出する。言語モデルは、単言語コーパスから学習される。統計翻訳において、言語モデルは通常、 N -gram モデルが用いられる。 N -gram モデルは、“単語列 $P(W_1^n) = w_1, w_2, \dots, w_n$ の i 番目の単語 w_i の生起確率 $P(w_i)$ は直前の単語列 $w_{i-(N-1)}, w_{i-(N-2)}, \dots, w_{i-1}$ に依存する”，という仮説に基づくモデルである。以下に計算式を示す。

$$P(W_1^n) = \prod_{i=1}^n P(w_i | w_{i-(N-1)}^{i-1}) \quad (25)$$

また、 $P(w_i | w_{i-(N-1)}^{i-1})$ の計算には、以下の式を用いる。なお、 $C()$ は単語列の出現数である。

$$P(w_i | w_{i-(N-1)}^{i-1}) = \frac{C(w_{i-(N-1)}^i)}{C(w_{i-(N-1)}^{i-1})} \quad (26)$$

表 10 に、言語モデルの例を示す。表 10 の言語モデルは、2-gram(2 単語間)のものである。

表 10: 言語モデルの例

-2.041488	He is	-0.000405
-2.804095	He had	-0.006078
-2.895682	He made	0.001770

一番上の例に関して、左から、“He”のあとに“is”がくる確率を常用対数で表した値“ $\log_{10}(P(is | He)) = -2.041488$ ”，2-gram で表された単語列“He is”，バックオフスムージングで推定された，“He”の後に“is”がくる確率を常用対数で表した値“ $\log_{10}(P(is | He)) = -0.000405$ ”である。ここで、バックオフスムージングとは、高次の N -gram が存在しない場合、低次の N -gram を用いる手法である。この低次の確率を改良したスムージングの手法が、Kneser-Ney スムージングである。言語モデルにおける N -gram 作成には、一般的に Kneser-Ney スムージングが用いられている。

2.4 デコーダ

デコーダは、翻訳モデルと言語モデルを用い、確率が最大となる出力文を探索し、翻訳を行う。図 2 に、例を示す。

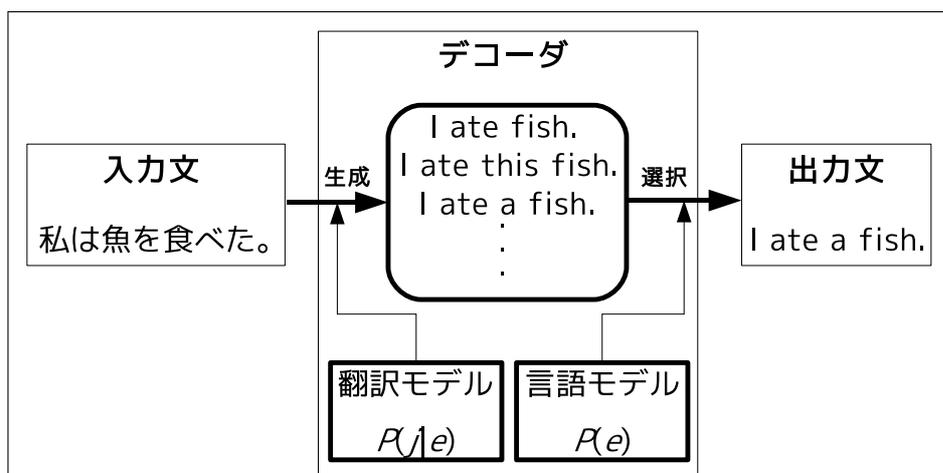


図 2: デコーダ動作例

図2は，入力文として“私は魚を食べた。”を与えられた場合の例である．この場合では， $\operatorname{argmax}_e P(j|e)P(e)$ の確率が最大となる “I ate a fish .” が出力される．統計翻訳において，出力文を決定するためには，莫大な計算量が必要となり，莫大な時間が必要となる．そこで，ビームサーチ法を用いることによって，計算量を削減する．

2.5 パラメータチューニング

パラメータチューニングは，デコーダで用いるパラメータを最適化し，評価関数（一般的には BLEU）を最大にする翻訳結果が選ばれるように，パラメータ調整を行うことである．パラメータチューニングには，Minimum Error Rate Training(MERT)[4] という手法を用いる．パラメータチューニングを行う際，ディベロップメントデータと呼ばれる，試し翻訳を行うデータを与え，各文に対して上位 100 個程度の翻訳候補を出力し，よりよい翻訳候補が上位にくるようにパラメータを調整する．

しかし，パラメータチューニングは，ディベロップメントデータに依存してしまう．また，パラメータチューニングの初期値に乱数を用いている．そのため，パラメータチューニングを行った場合，チューニングごとに値が異なるため，比較が困難である．

3 関連研究

3.1 統計翻訳の問題点

猪澤らは、文節区切りとした学習データからフレーズテーブルを作成し、翻訳精度を向上させることに成功した [1]。しかし、文節区切りとした学習データを用いて作成したフレーズテーブルを用いた翻訳では、主語が省略されている文においてうまく主語を生成することができず、翻訳精度が低下する原因であると述べている。表 11 に主語が省略されていることが原因で翻訳精度が低下した例を示す。なお、表 11 の従来手法では、単語区切りの学習データを用いてフレーズテーブルを作成している。また、表 11 の先行研究では、文節区切りの学習データを用いてフレーズテーブルを作成している。

表 11: 翻訳精度が低下した例

入力文	どこかの図書館で数か月懸命に勉強することが必要だ。
従来手法の出力文	I need some months in the library to study hard.
先行研究の出力文	Some few months in the library to study hard necessary.

ここで、従来手法のフレーズ対と、先行研究のフレーズ対を表 12 に示す。

表 12: 出力文が用いたフレーズ対 (単語区切りフレーズ対の問題)

従来手法のフレーズ対	先行研究のフレーズ対
どこ I	どこかの Some
かの some	図書館で in the library
図書館で in the library	数ヶ月 few months
数ヶ月 months	懸命に勉強する to study hard
懸命に勉強する to study	ことが必要だ。 necessary .
ことが必要だ need	
。 .	

従来の学習データを用いたフレーズ対において、フレーズ対”どこ ||| I”は不適切である。しかし、入力文には主語がなく、”どこ ||| I”を用いて主語を作ることで、出力文の翻訳精度が向上している。一方で、文節区切りの学習データを用いたフレーズ対では、”どこかの ||| Some”を用いている。このフレーズ対は適切ではあるが、他のフレーズ対でも主語が生成されず、翻訳精度が低下する原因となっている。

このような主語が省略されている文に対し主語補完を行うことで、主語が生成され、翻訳精度が向上するのではないかと考えられる。

3.2 主語補完における先行研究

中岩ら [5][6][7][8] は、ゼロ代名詞 (日本語において省略されている格要素) の解析を行い、ゼロ代名詞照応解析の方法を提案した。まずゼロ代名詞の解析について、日英機械翻訳システム評価用例文 3,718 文を解析した結果、照応解析を要するゼロ代名詞が 484 文、510 箇所存在することを示した。そして、このゼロ代名詞 510 件のうち、そこに補完すべき指示対象が文内に存在する場合は 139 件、文外に存在する場合は 371 件存在すると報告した。また、ゼロ代名詞の照応解析方法として、文内照応、文間照応、文章外照応の 3 タイプを提案した。文内照応は、ゼロ代名詞と、補完すべき指示対象が同じ文内に存在する場合、同じ文内の指示対象を補完する方法である。例を表 13 に示す。

表 13: 文内照応の例

ゼロ代名詞を含む文	彼は方程式を解いて答えを出した。
文内照応解析による補完結果	彼は方程式を解いて (彼が) 答えを出した。

この文内照応は、助詞の種類や、接続語・用言意味属性・様相表現による制約を用い、照応解析を行う。この文内照応解析では、再現率 98%、適合率 100% の精度で正しい指示対象を決定できたことが報告された。文間照応解析は、新聞記事文のような複数文からなる文章においては、補完すべき指示対象が、ゼロ代名詞が現れる文と異なる文中に存在する場合は多いことを利用し、照応解析を行う方法である。この文間照応解析では、ゼロ代名詞の 84% に対して照応解析が成功したことを報告した。文章外照応は、文章中に補完すべき指示対象が現れない場合の照応解析方法である。この文章外照応では、格への意味的制約・用言意味属性・様相表現・接続語などを考慮したルールを構築し、ゼロ代名詞の指示対象を推定する方法である。この文章外照応解析の結果、ルールを複雑にすれば、ほとんどの文において、省略各要素を復元が可能であることを報告した。

これらの方法は、機械翻訳に対して用いることが可能であるが、統計的機械翻訳における主語補完の効果は報告されていない。

4 “主語省略文”

日本語文において、主語が省略されている文が存在する。本研究では、主語が省略されている文を“主語省略文”とする。以下に、“主語省略文”の例を示す。

表 14: “主語省略文”の例

主語省略文	昼食をたっぷり取った。
対訳英語文	He had a big lunch .
主語省略文	サッカーをした。
対訳英語文	I played soccer .
主語省略文	せいぜい1時間しかかからなかった。
対訳英語文	It took little more than an hour .

4.1 主語の定義

日本語では、“これは”や“私が”のように、名詞の後に助詞“は”もしくは“が”もしくは“も”があれば主語となることが多い。そこで本研究では、名詞と助詞“は”もしくは“が”もしくは“も”の句を主語とする。以下に、主語の例を示す。

主語の例

私は 私が 私も 彼は 彼が 彼も
これは それが あれも クジラは 電車が 海も

4.2 命令文

日本語において、命令文には主語が無いことが多い。また英語においても、命令文には主語が無いことが多い。表 15 に命令文の例を示す。

表 15: 命令文の例

日本語の命令文	早くしろ。
英語の命令文	Hurry it up .
日本語の命令文	次の文を英訳せよ。
英語の命令文	Put the following sentences into English .
日本語の命令文	声を低くしなさい。
英語の命令文	Keep your voice down .

このような命令文には、主語が無いことが妥当である。よって本研究では、命令文に対し主語補完を行わない。

4.3 “主語省略文”の判断条件

4.1 節と 4.2 節を踏まえ、本研究では、“主語省略文”であるための条件を表 16 のように定義する。

表 16: “主語省略文”の判断条件

条件 1	助詞 “は” が含まれない文
条件 2	助詞 “が” が含まれない文
条件 3	助詞 “も” が含まれない文
条件 4	動詞が命令形ではない文

本研究では、この 4 つの条件を用いて“主語省略文”を抽出し、主語補完を行う。本研究では、主語補完方法として、日本語補完、対訳補完、train 補完の 3 つを用いる。日本語補完では、表 16 の条件 1～条件 4 を全て満たす場合、“主語省略文”として抽出する。また、対訳補完と train 補完では、表 16 の条件 1～条件 3 の 3 つの条件を満たす場合、“主語省略文”として抽出する。この 3 つの主語補完方法については、5 章で説明する。ここで、“主語省略文”の 4 つの条件をすべて満たす日本語文の例を表 17 に示す。また、“主語省略文”の 4 つの条件を 1 つ以上満たさない日本語文の例を表 18 に示す。

表 17: “主語省略文”の判断条件を満たす文の例

例 1	主語省略文例	昼食をたっぷり取った。
	対訳英語文例	He had a big lunch .
例 2	主語省略文例	サッカーをした。
	対訳英語文例	I played soccer .
例 3	主語省略文例	盗品を山分けにした。
	対訳英語文例	They divided the loot among themselves .

例 1 の場合、表 16 の条件 1～条件 4 を全て満たしているため、“主語省略文”であると判断する。例 2 の場合も同様に、条件 1～条件 4 を全て満たすため、“主語省略文”であると判断する。例 3 の場合も同様である。

表 18: “主語省略文”の判断条件を満たさない文の例

例 1	主語がある文例	私は野球をした。
	対訳英語文	I played baseball.
例 2	主語がある文例	彼が一番後にやってきた。
	対訳英語文	He was the last to come .
例 3	命令文例	窓を開けなさい。
	対訳英語文	Open the window .

表 18 の例 1 の場合，助詞 “は” が含まれている．よって，表 16 の条件 1 を満たさない
ので，“主語省略文”ではないと判断する．例 2 の場合，助詞 “が” が含まれている．よっ
て，表 16 の条件 2 を満たさないので，“主語省略文”ではないと判断する．また，例 3 の
場合，動詞 “開けなさい” が命令形である．よって，表 16 の条件 4 を満たさないので，“
主語省略文”ではないと判断する．ここで，対訳補完と train 補完では，対訳英語文を参
照して主語補完を行う．このため，命令文に対して主語補完が行われることがない．以
上の点を考慮し，対訳補完と train 補完において，例 3 のように，表 16 の条件 1～条件 3
は満たすが条件 4 のみを満たさない日本語文は，“主語省略文”として抽出する．詳細は
5.2 節に記述する．

5 主語補完

本研究では，テストデータの日本語データに対し，日本語補完と対訳補完の2種類の方法で主語補完を行う．また，学習データに対し，train 補完を用いて主語補完を行う．以下に，日本語補完，対訳補完，train 補完における主語補完の方法をそれぞれ示す．

5.1 日本語補完

日本語補完では，テストデータの日本語文に対して，対訳英語文を参照せずに主語補完を行う．日本語補完は，“主語省略文”に主語補完を行うことによる翻訳精度の変化を調査することを目的とする．そのため，日本語補完で補完する主語は，“私は”である必要はない．しかし本研究では，日本語データを調査した結果，“私は”がもっとも一般的な主語であると判断した(10章)ため，“主語省略文”の文頭に“私は”を補完することとする．以下に，主語補完の手順を示す．

手順1 日本語文に対し，形態素解析を行う．表19に，形態素解析結果の例を示す．

表 19: 形態素解析結果の例

サッカー	サッカー	サッカー	名詞-一般		
を	ヲ	を	助詞-格助詞-一般		
し	シ	する	動詞-自立	サ変・スル	連用形
た	タ	た	助動詞	特殊・タ	基本形
。	。	。	記号-句点		

表19は，入力文として“サッカーをした。”が入力された場合の形態素解析結果である．形態素解析を行うと，表19のように，縦方向に単語区切りで出力される．形態素解析の出力は，左から順に，形態素，読み，原型，品詞，品詞の細分類，活用形となっている．

手順2 形態素解析を行ったデータを参照し，表16の条件1～条件4を全て満たす文を“主

語省略文”として抽出する．表20に“主語省略文”として抽出される文の形態素解析結果の例を示す．また，表21と表22に，“主語省略文”として抽出されない文の形態素解析結果の例を示す．

表 20: 抽出される文の例

サッカー	サッカー	サッカー	名詞-一般		
を	ヲ	を	助詞-格助詞-一般		
し	シ	する	動詞-自立	サ変・スル	連用形
た	タ	た	助動詞	特殊・タ	基本形
。	。	。	記号-句点		

表 20 は，“サッカーをした。”に対する形態素解析結果である。この場合，助詞“は”および“が”および“も”を含まない。また，動詞“し”は，命令形ではない。よって，表 16 の条件を全てみたすので，“主語省略文”として抽出される。

表 21: 抽出されない文の例 1

私	ワタシ	私	名詞-代名詞-一般		
は	ハ	は	助詞-係助詞		
野球	ヤキュウ	野球	名詞-一般		
を	ヲ	を	助詞-格助詞-一般		
し	シ	する	動詞-自立	サ変・スル	連用形
た	タ	た	助動詞	特殊・タ	基本形
。	。	。	記号-句点		

表 21 は，“私は野球をした。”に対する形態素解析結果である。この場合，助詞“は”が含まれている。よって，表 16 の条件 1 を満たさないため，抽出されない。

表 22: 抽出されない文の例 2

窓	マド	窓	名詞-一般		
を	ヲ	を	助詞-格助詞-一般		
開け	アケ	開ける	動詞-自立	一段	連用形
なさい	ナサイ	なさる	動詞-非自立	五段・ラ行特殊	命令 i
。	。	。	記号-句点		

表 22 は，“窓を開けなさい。”に対する形態素解析結果である。この場合，動詞“なさい”が命令形である。よって，表 16 の条件 4 を満たさないため，抽出されない。

手順 3 手順 2 で抽出した“主語省略文”に対し，文頭に“私は”を補完する。以下に例を示す。

表 23: 主語補完例

主語補完前	サッカーをした。
主語補完後	私はサッカーをした。

また、図 3 に主語補完の手順を示す。

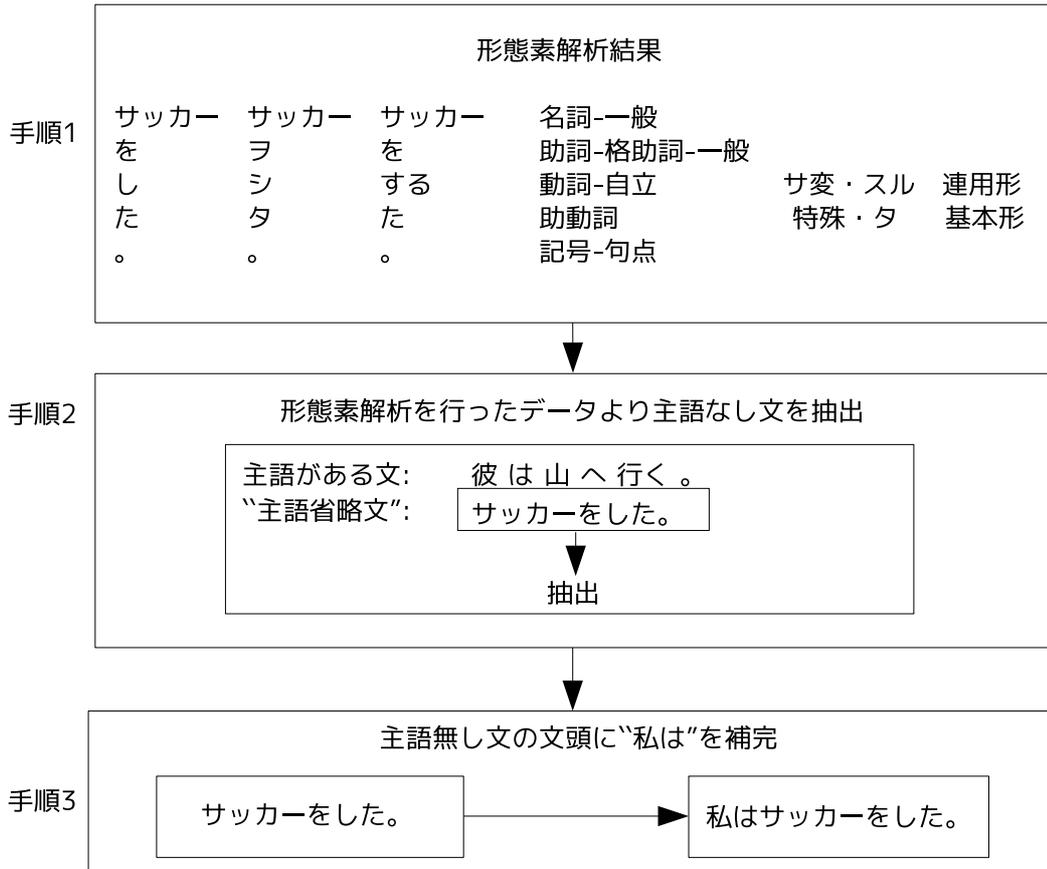


図 3: 主語補完手順

日本語補完例

表 24 に、日本語補完を用いて主語補完を行った例を示す。

表 24: 主語補完例

主語省略文	サッカーをした。
日本語補完した文	私はサッカーをした。
対訳英語文	I played soccer .
主語省略文	勝負に勝った。
日本語補完した文	私は勝負に勝った。
対訳英語文	We won game .
主語省略文	ブランデーを一息に飲んだ。
日本語補完した文	私はブランデーを一息に飲んだ。
対訳英語文	He swallowed the brandy in one swig .
主語省略文	その後まもなくナポリで結婚した。
日本語補完した文	私はその後まもなくナポリで結婚した。
対訳英語文	They got married soon afterward in Naples .
主語省略文	彼の手のひらからコインを取った。
日本語補完した文	私は彼の手のひらからコインを取った。
対訳英語文	She took the coin from his palm .

表 24 に示すように、対訳英語文の主語が “I” ではない場合も、“主語省略文”の文頭に “私は” を補完している。しかし、日本語補完を行う目的は、統計翻訳において主語補完が有効であるかを調査することである。よって、対訳英語文の主語が “I” ではない場合にも、“主語省略文”の文頭に “私は” を補完することは妥当であると考えられる。

5.2 対訳補完

対訳補完では，テストデータの日本語文に対して，対訳英語文を参照することにより主語補完を行う．この対訳補完は，本来存在しない対訳英語文を用いているため，実際の翻訳では対訳補完を行うことができない．しかし本研究では，理想的な主語補完を行った際の翻訳精度を調査するためにこの手法を用いて翻訳を行う．対訳補完における主語補完の手順を以下に示す．

手順1 日本語文に対し，形態素解析を行う．

手順2 形態素解析を行ったデータを参照し，表16の条件1～条件3を満たす文のみを“主語省略文”とし，抽出する．“主語省略文”として抽出される文の例を以下に示す．

表 25: 対訳補完 “主語省略文” 例

主語省略文	昼食をたっぷり取った。
対訳英語文	He had a big lunch .
主語省略文	声を低くしなさい。
対訳英語文	Keep your voice down .

手順3 手順2で抽出した“主語省略文”に対する対訳英語文の文頭単語を抽出する．文頭単語の抽出の例を図4と図5に示す．

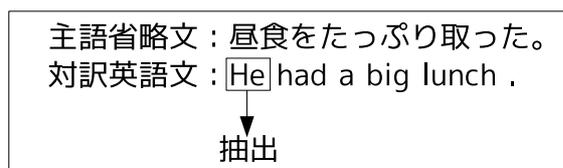


図 4: 文頭単語抽出例 1

図4の場合では，“昼食をたっぷり取った。”の対訳英語文である“He had a big lunch .”の文頭単語“He”が抽出される．



図 5: 文頭単語抽出例 2

図 5 の場合では，“声を低くしなさい。”の対訳英語文である “Keep your voice down .” の文頭単語 “Keep” が抽出される。

手順 4 抽出した対訳英語文の文頭単語を，変換規則に従って日本語に変換する。変換規則を表 26 に示す。図 4 の場合では，抽出した単語 “He” を変換規則に従って変換し，“彼は”を得る。しかし，図 5 のように，抽出した対訳英語文の文頭単語が表 26 の変換規則に適応しなかった場合は，主語補完は行わない。よって，命令文に対し，主語補完が行われることがないので，対訳補完において表 16 の条件 4 は意味をなさないため用いない。

表 26: 文頭単語の日本語への変換規則

正解文の文頭単語	変換する日本語
I	私は
He	彼は
She	彼女は
We	私たちは
It	それは
You	あなたは
They	彼らは
Someone	誰かが
Anyone	誰かが
Somebody	誰かが
Anybody	誰かが

手順 5 変換した日本語を，手順 2 で抽出した “主語省略文” の文頭に補完する。例を表 27 に示す。

表 27: 主語補完例

主語補完前	昼食をたっぷり取った。
主語補完後	彼は昼食をたっぷり取った。
対訳英語文	He had a big lunch .

対訳補完の例

表 28 に、対訳補完を用いて主語補完を行った例を示す。

表 28: 主語補完例

主語省略文	サッカーをした。
日本語補完した文	私はサッカーをした。
対訳英語文	I played soccer .
主語省略文	勝負に勝った。
日本語補完した文	私たちは勝負に勝った。
対訳英語文	We won game .
主語省略文	ブランデーを一息に飲んだ。
日本語補完した文	彼はブランデーを一息に飲んだ。
対訳英語文	He swallowed the brandy in one swig .
主語省略文	その後まもなくナポリで結婚した。
日本語補完した文	彼らはその後まもなくナポリで結婚した。
対訳英語文	They got married soon afterward in Naples .
主語省略文	彼の手のひらからコインを取った。
日本語補完した文	彼女は彼の手のひらからコインを取った。
対訳英語文	She took the coin from his palm .

表 28 に示すように、対訳補完では、対訳英語文を参照することで適切な主語補完を行っている。

5.3 train 補完

train 補完は、学習データに対し、対訳英語文を参照することで主語補完を行う。学習データには対訳文が存在するため、対訳英語文を参照することが可能である。train 補完は、対訳補完における主語補完の方法と同じ方法で主語補完を行う。

6 実験方法

本研究では，テストデータ 10,000 文に対して，5 章で述べた日本語補完と対訳補完の 2 種類の方法を用いて主語補完を行う．また，学習データ 100,000 文に対して，train 補完を用いて主語補完を行う．本研究では，これらの主語補完方法を用いて，日英統計翻訳における主語補完の有効性を調査する．本研究で行う実験を 6.2 節にまとめる．

6.1 実験手順

実験手順の例として，テストデータに対して，日本語補完を用いて主語補完を行う場合の翻訳手順を以下に示す．

手順 1 テストデータの“主語省略文”に対し，日本語補完を用いて主語補完を行う．

手順 2 主語補完を行ったテストデータと，主語補完を行わないテストデータに対して統計翻訳を行う．

また，翻訳手順を図 6 に示す．

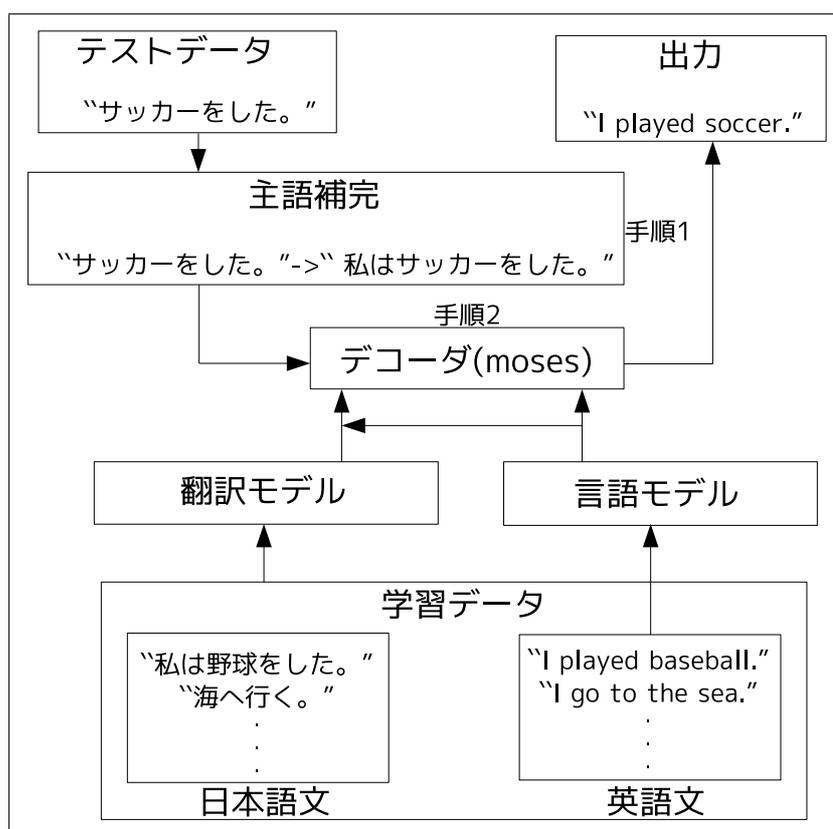


図 6: 翻訳手順

6.2 実験の種類

日本語補完

以下に、日本語補完を用いる実験を示す.

- 日本語補完のみ用いた実験
- 日本語補完と train 補完を併用した実験

対訳補完

以下に、対訳補完を用いる実験を示す.

- 対訳補完のみ用いた実験
- 対訳補完と train 補完を併用した実験

これらの4つの実験と比較を行うため、train 補完のみを用いた実験も行う.

7 実験環境

7.1 デコーダのパラメータ

デコーダには, Moses[9] を用いる. 過去の研究 [10] から, 翻訳モデルには日英翻訳確率と英日翻訳確率の共起確率を用いる. したがって, フレーズテーブルの各種の重み “weight-t” を “0.5 0.0 0.5 0.0 0.0” とする. また, “ttable-limit” の値は 60 とする.

また本研究では, 翻訳実験の比較を行うため, パラメータチューニング [4] を行わない.

7.2 翻訳モデルの学習

翻訳モデルの学習には, Moses に付属している “train-factored-phrase-model.perl” を用いる.

7.3 言語モデルの学習

言語モデルには N -gram モデルを用いる. 本研究では, SRILM[11] の “ngram-count” を用いて 5-gram の言語モデルを学習する. 尚, スムージングに “Kneser-Ney discount” を用いる.

7.4 実験データ

実験データには, 辞書の例文より抽出した単文コーパス 181,988 文から, 学習データとして 100,000 文, テストデータとして 10,000 文を用いる. 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[12] を用いて分かち書きを行う. さらに, 各コーパスの英語文に対して, tokenizer.perl[9] を用いて, 分かち書きを行う. 表 29 に日英対訳文の例を示す.

表 29: 単文コーパス例文

日本語文	昼食をたっぷり取った。
英語文	I had a big lunch .
日本語文	梅雨が始まった。
英語文	The rainy season has set in .
日本語文	私は猫を1匹飼っている。
英語文	I have a cat .

また、主語補完を行う際の形態素解析には、CaboCha[13]を用いる。

8 評価方法

本研究では、翻訳実験の評価として、自動評価を行う。また、日本語補完のみを用いた実験において人手評価を行うことで、日英統計翻訳における主語補完の効果を調査する。

8.1 自動評価

本研究で用いる自動評価は、あらかじめ用意した正解文(対訳英語文)と、翻訳システムが出力した文とを比較する評価方法である。以下に入力文とそれに対する正解文の例を示す。

表 30: 正解文例

入力文	私は映画を見に行く。
正解文	I go to see a movie .
入力文	1点差で敗北した。
正解文	We lost the game by one point .
入力文	机の上をきれいにしなさい。
正解文	Please clean off the desk .

自動評価法には多くの方法があるが、本研究では、BLEU[14], METEOR[15], NIST[16]を用いる。

BLEU

BLEU は語順 (4-gram) が正しい場合に高いスコアを出力し、0 から 1 までの値で評価を行う。BLEU の算出には、以下の式を用いる。

$$BLEU_{score} = BP \times \sqrt[N]{\prod_{i=1}^N P_n} \quad (27)$$

$$P_n = \frac{\sum_i \text{出力文 } i \text{ と正解文 } i \text{ で一致した } N\text{-gram の数}}{\sum_i \text{出力文 } i \text{ の } N\text{-gram の数}} \quad (28)$$

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 をとなり、出力文が正解文より短い場合は 1 未満の値となる

METEOR

METEORは、正解文と比較し、3人称単数や時制などの単語属性が正しい場合に高いスコアを出力し、0から1までの値で評価を行う。METEORは以下の式で計算される。

$$METEOR_{score} = F_{mean} \times (1 - Pen) \quad (29)$$

$$F_{mean} = \frac{P \times R}{\alpha \times P + (1 - \alpha) \times R} \quad (30)$$

$$Pen = \gamma \times \left(\frac{c}{m}\right)^\beta \quad (31)$$

METEORはF値(再現率 R と適合率 P の調和平均)と、単語の非連続性に対するペナルティ関数 Pen を用い計算される。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。 α , β , γ の値はパラメータである。

NIST

NISTは語順(5-gram)と文長(単語数)が正しい場合に高いスコアを出力し、0から ∞ までの値で評価を行う。NISTは、BLEUをベースとし、BLEUに改良を加えられている。以下にNISTの計算式を示す。

$$NIST = \sum_{n=1}^N \left\{ \frac{\sum_{all w_1 \dots w_n \text{ in sys output}} info(w_1 \dots w_n)}{\sum_{all w_1 \dots w_n \text{ in sys output}} 1} \right\} \times \exp \left\{ \beta \log^2 \left[\frac{L_{sys}}{\bar{L}_{ref}}, 1 \right] \right\} \quad (32)$$

ここで、 \bar{L}_{ref} は正解文内での平均単語数を表す。 L_{sys} および N は、評価対象に含まれる単語数および最大 n -gram長を表す。 β は以下の式で計算される。

$$\beta = \frac{\log(0.5)}{\left\{ \log(1.5) \right\}} \quad (33)$$

また、 $info(w_1 \dots w_n)$ は、評価に用いられるすべての正解文を用い、以下の式で計算される。

$$info(w_1 \dots w_n) = \log_2 \left(\frac{\text{the number of occurrence of } w_1 \dots w_{n-1}}{\text{the number of occurrence of } w_1 \dots w_n} \right) \quad (34)$$

8.2 人手評価

本研究では、主語補完を行わない翻訳をベースラインとする。人手評価として、日本語補完とベースラインの対比較評価を行う。対比較評価は、日本語補完を用いて主語補完した1,526文に対する出力文からランダムに100文抽出し、ベースラインの出力文と比較し、評価する。以下に評価基準を示す。

- a) 日本語補完○ 日本語補完の翻訳品質がベースラインの翻訳品質より優れている場合
- b) ベースライン○ ベースラインの翻訳品質が日本語補完の翻訳品質より優れている場合
- c) 差なし 日本語補完の翻訳品質とベースラインの翻訳品質に差がない場合
- d) 同一出力 日本語補完の翻訳結果とベースラインの翻訳結果が同じ場合

なお、日本語補完では、文頭に“私は”を補完する。この日本語補完の目的は、日英統計翻訳における主語補完の効果を調査することである。よって本研究では、出力文の主語が正解文の主語と異なる場合も、出力文は正しい翻訳であると評価することとする。表31に例を示す。

表 31: 正解文例

主語省略文	昼食をたっぷり取った。
日本語補完入力文	私は昼食をたっぷり取った。
正解文 (対訳英語文)	He had a big lunch .
ベースライン出力文	He had a big lunch .
日本語補完出力文	I had a big lunch .

この例の場合、翻訳結果として、ベースライン出力文では、主語として“He”が出力され、正解文の主語と一致している。これに対し、日本語補完出力文では、主語として“I”が出力されており、正解文の主語と不一致である。しかし、本研究では主語補完による翻訳精度の調査を目的としているため、日本語補完出力文は、適切な翻訳を行っていると判断する。よって、ベースライン出力文と日本語補完出力文の翻訳精度は差がないと判断する。

9 実験結果

9.1 主語補完結果

日本語補完と対訳補完において，テストデータ 10,000 文に対して主語補完を行った文数，また，train 補完において，学習データ 100,000 文に対して主語補完を行った文数を表 32 に示す．

表 32: 主語補完文数

	総文数	主語補完文数
日本語補完	10,000	1,526
対訳補完	10,000	1,109
train 補完	100,000	11,060

9.2 自動評価結果

テストデータ 10,000 文を用いた実験

ベースライン，日本語補完のみを用いて主語補完を行った実験，日本語補完と train 補完を併用した実験，対訳補完のみを用いて主語補完を行った実験，対訳補完と train 補完を併用した実験，train 補完のみを用いて主語補完を行った実験の 6 つの実験における自動評価の結果を表 33 に示す．6 つの実験は全て，テストデータとして 10,000 文，学習データとして 100,000 文用いる．

表 33: 10,000 文での自動評価結果

	BLEU	METEOR	NIST
ベースライン	0.1146	0.3983	4.0832
日本語補完	0.1195	0.4068	4.2807
日本語補完+train 補完	0.1187	0.4060	4.2645
対訳補完	0.1218	0.4153	4.3579
対訳補完+train 補完	0.1209	0.4135	4.3266
train 補完	0.1130	0.3997	4.1405

表 33 では，対訳補完は，ベースラインと比較し，BLEU 値が 0.72% 向上している．これより，統計翻訳において，主語補完が有効であることが分かる．また，日本語補完では，ベースラインと比較し，BLEU 値が 0.49% 向上している．これより，対訳補完には

劣っているが、日本語補完も有効であることが分かる。また、日本語補完+train 補完は、日本語補完と比較し、BLEU 値が 0.08%低下している。そして、対訳補完+train 補完は、対訳補完と比較し、BLEU 値が 0.09%低下している。さらに、train 補完は、ベースラインと比較し、BLEU 値が 0.16%低下している。これより、train 補完は有効ではないと考えられる。

b) 日本語補完において主語補完したデータのみを用いた実験

日本語補完を行った実験と、日本語補完と train 補完を併用した実験について、日本語補完で主語補完した 1,526 文のみをテストデータとして用いた翻訳実験を行った。評価結果を表 34 に示す。

表 34: 1,526 文での自動評価結果

	BLEU	METEOR	NIST
ベースライン	0.0958	0.3559	3.0051
日本語補完	0.1035	0.3690	3.4898
日本語補完+train 補完	0.1030	0.3619	3.4380

表 34 では、ベースラインと比較し、日本語補完は BLEU 値が 0.77%向上した。表 33 と比較し、BLEU 値の向上が大きいことが確認できる。また、日本語補完+train 補完は、日本語補完と比較し、0.03%低下した。これより、主語があるテストデータのみに対し翻訳を行った場合でも、train 補完が有効でないことが確認できる。

c) 対訳補完において主語補完したデータのみを用いた実験

対訳補完を行った実験と、対訳補完と train 補完を併用した実験について、対訳補完で主語補完した 1,109 文を用いた実験結果を表 35 に示す。

表 35: 1,109 文での自動評価結果

	BLEU	METEOR	NIST
ベースライン	0.1082	0.3723	3.0857
対訳補完	0.1328	0.4471	4.0784
対訳補完+train 補完	0.1304	0.4517	4.0971

表 35 では、対訳補完は、ベースラインと比較し、BLEU 値が 2.46%向上した。これより、主語補完した文の翻訳精度が大きく向上していることが確認できる。また、対訳補完+train 補完は、対訳補完と比較し、0.24%低下した。これより、多くの文について適

切な主語が補完されているテストデータのみに対し翻訳を行った場合でも，train 補完が有効でないと考えられる．

9.3 人手評価結果

日本語補完で主語補完を行った1,526文の出力文より，ランダムに100文抽出し，ベースラインの出力文との対比較評価を行った．

9.3.1 対比較評価結果

対比較評価結果を表36に示す．

表 36: 対比較評価結果

日本語補完○	ベースライン○	差なし	同一出力
21	9	65	5

対比較評価結果より，ベースラインと比較して日本語補完の方が翻訳品質が良いことが確認できる．

9.3.2 対比較評価例文

対比較評価を行った出力文の例を示す．

日本語補完が優れていると判断した例

対比較評価において，日本語補完が優れていると判断した例を，表37と表38に示す．

表37において，ベースラインの出力文は“うそ”に対応する単語が出力されていない．それに対し，日本語補完の出力文は“うそをつく”に対し，“tells lies”と出力されている．よって，日本語補完が優れていると判断した．

表 37: 日本語補完が優れていると判断した例 1

ベースライン入力文	平気ですをつく。
日本語補完入力文	私は平気ですをつく。
正解文	She doesn't scruple to tell a lie .
ベースライン出力文	It is a has .
日本語補完出力文	I tells lies without shame .

表 38 において、ベースラインは“酒を”の翻訳が行われていない。それに対し、日本語補完は、“酒を飲む”という意味である“drinking”が出力されている。よって、日本語補完が優れていると判断した。

表 38: 日本語補完が優れていると判断した例 2

ベースライン入力文	酒をやめている。
日本語補完入力文	私は酒をやめている。
正解文	I am off liquor .
ベースライン出力文	I gave up .
日本語補完出力文	I gave up drinking .

ベースラインが優れていると判断した例

対比較評価において、ベースラインが優れていると判断した例を表 39 と表 40 に示す。表 39 において、ベースラインは、“見つめた”が“looked at”と翻訳されている。それに対し、日本語補完は“見つめた”にベースラインの方が優れていると判断した。

表 39: ベースラインが優れていると判断した例 1

ベースライン入力文	じっと手を見つめた。
日本語補完入力文	私はじっと手を見つめた。
正解文	He stared down at his hands .
ベースライン出力文	She looked at hand .
日本語補完出力文	I washed my hands on me .

表 40 において、ベースラインは、入力文に対し、適切に翻訳されている。それに対し、日本語補完は、“私は”に対し、“of my life”と訳されており、不適切である。これより、ベースラインが優れていると判断した。

表 40: ベースラインが優れていると判断した例 2

ベースライン入力文	泥棒に入られた。
日本語補完入力文	私は泥棒に入られた。
正解文	I had a burglar break into my home .
ベースライン出力文	My house was robbed .
日本語補完出力文	My house was robbed of my life .

差なしと判断した例

対比較評価において、ベースラインと日本語補完の出力に差が無いと判断した例を表 41 と表 42 に示す。表 41 において、ベースラインと日本語補完は、主語が異なるだけなので、差なしと判断した。

表 41: 差なしと判断した例 1

ベースライン入力文	暴動を起こした。
日本語補完入力文	私は暴動を起こした。
正解文	They started a mutiny .
ベースライン出力文	He had a riot .
日本語補完出力文	I had a riot .

表 42 において、ベースラインの出力文と日本語補完の出力文の両方が、意味が入力文と異なる。よって、差なしと判断した。

表 42: 差なしと判断した例 2

ベースライン入力文	100セントで1ドルになる。
日本語補完入力文	私は100セントで1ドルになる。
正解文	One hundred cents make a dollar .
ベースライン出力文	There are a hundred dollars .
日本語補完出力文	I make a ten dollars .

10 考察

10.1 英語文の文頭調査

日本語補完では、テストデータの“主語省略文”の文頭に“私は”を補完した。この妥当性を示すため、テストデータの対訳英語文の文頭単語を調査した。命令文などの文頭単語が動詞になる文は、“その他”に分類した。表 43 に調査結果を示す。

表 43: 対訳英語文の文頭単語

I	He	She	We	They	固有名詞	その他
23	19	4	4	2	18	30

表 43 より、対訳英語文の文頭単語は、“I”が一番多いことが確認できる。よって、“主語省略文”の文頭に“私は”を補完することは妥当であると考えられる。また表 43 より、テストデータの対訳英語文の文頭が“I”ではない文が多く存在することが確認できた。これは、表 33 と表 34 の自動評価結果において、日本語補完の BLEU 値が大きく向上しなかったことの原因の一つであると考えられる。しかし、人手評価においては、8.2 節で述べたように、例えば“He”が出力された場合と“I”が出力された場合は翻訳品質に差がないと判断するため、“私は”を補完することは適切であると考えられる。

10.2 日本語補完についての考察

日本語補完では、4章で定義した条件1～条件4を全て満たす場合，“主語省略文”として抽出した。そして，抽出した全ての文の文頭に“私は”を補完した。この日本語補完についての考察を以下に記述する。

10.2.1 “主語省略文”抽出の精度

“主語省略文”1,526 に対する調査

日本語補完で抽出した“主語省略文”1,526文からランダムに100文抽出し，“主語省略文”抽出の精度調査を行った。抽出精度調査の評価基準を以下に示す。

抽出○ 抽出した文に主語が無い場合

抽出× 抽出した文に主語がある場合

表 44 に結果を示す。

表 44: “主語省略文”抽出評価結果

抽出○	抽出×
98	2

また，主語がある文を“主語省略文”と判断し，抽出した文の例を表 45 に示す。

表 45: “主語省略文”抽出不適切文例

文例 1	摩周湖の美しさに全員感嘆の声を上げた。
文例 2	これ、キアリーって作家の書いた本です。

表 45 の文例 1 の主語は“全員”である。このとき，“全員”のあとに，助詞“が”が省略されているため，“主語省略文”と判断した。また，表 45 の文例 2 の主語は“これ”である。このとき，“これ”のあとに，助詞“は”が省略されているため，“主語省略文”と判断した。しかし，表 45 のような文は少なく，表 44 から分かるように，“主語省略文”抽出の精度が高いことが分かる。

主語があると判断された 8,474 文に対する調査

日本語補完で“主語省略文”として抽出されなかった 8,474 文から，ランダムに 100 文抽出し，“主語省略文”抽出の精度を調査した。抽出精度調査の評価基準を以下に示す。

抽出○ 抽出した文に主語がある場合

抽出× 抽出した文に主語が無く、主語補完を行うべきであると判断される場合

表 47 に結果を示す。

表 46: “主語省略文”抽出評価結果

抽出○	抽出×
98	2

また、抽出×であった文を表 47 に示す。

表 47: “主語省略文”抽出不適切文例

文例 1	学生時代は良き友人に恵まれた。
文例 2	背が高く、ほとんど幽霊のようにやせていた。

文例 1 では、助詞 “は” が含まれているため、主語があると判断した。しかし、“学生時代は” は、主語ではない。よって、抽出が不適切であると判断した。この文例 1 の場合では、“私は” や “彼は” のような主語を補完を行うことが好ましいと考える。また文例 2 では、助詞 “が” が含まれているため、主語があると判断した。しかし、“背が” は、“やせていた” に対する主語ではない。よって、抽出が不適切であると判断した。この文例 2 の場合では、“私は” や “彼は” のような主語を補完を行うことが好ましいと考える。

表 47 のような文の主語の有無を判断するのは、本研究の主語補完アルゴリズムでは不可能である。しかし、このような文は、少数しか存在しない。また、本研究では、日英統計翻訳における主語補完の効果を調査することが目的であるので、今回は誤差であるとし、考慮しないこととする。しかし、主語補完を統計翻訳に実用的に組み込む場合、このような助詞 “は” もしくは “が” もしくは “も” が含まれている文に主語が含まれない場合においても、主語が省略されていると判断し抽出する必要があると考える。今後、新たな主語補完アルゴリズムを考案する必要があると考える。

10.2.2 日本語補完における主語補完の精度

日本語補完で主語補完を行った 1,526 文より、ランダムに 100 文抽出し、主語補完の精度の評価を行った。以下に評価基準を示す。

主語補完○ 主語補完した文が日本語として文法的に、また意味的に正しい場合

主語補完× 主語補完した文が日本語として文法的に、また意味的に正しくない場合

表 48 に評価結果を示す。

表 48: 主語補完評価結果

主語補完○	主語補完×
93	7

主語補完が正しくない例文を表 49 に示す。表 49 では、日本語補完により補完された主語は、“私は”である。

表 49: 主語補完誤り文例

文例 1	“私は”たばこの消し忘れで大火事になった。
文例 2	“私は”雨になりそうだ。
文例 3	“私は”みんな口をぽかんとあけた。

表 49 の文例 1 では、“大火事になった”に対する主語は、人ではなく、建物や“それは”などがふさわしい。よって、文例 1 の主語補完は誤りである。文例 2 では、主語は“天気は”などがふさわしいと考える。また、文例 3 では、“みんな”が主語であるが、“みんな”のあとに助詞が省略されている。そのため、“主語省略文”として抽出されたことにより、誤って主語補完されている。表 48 の評価結果より、わずかではあるが、主語補完に誤りがあることが分かる。主語補完の精度を向上させることにより、さらに翻訳精度が向上すると予想できる。

10.2.3 日本語補完における自動評価の解析

10.2 節では、テストデータの主語が“私は”である文は、全体のおよそ 20%であることが確認できた。さらに、10.2.2 節では、主語補完の精度はおよそ 90%であることが確認できた。以上の環境での翻訳実験であったが、テストデータ 10,000 文を用いた日本語補完における自動評価結果では、ベースラインと比較し、BLEU 値において 0.49%の向上を確認できた。また、主語補完を行った 1,526 文を用いた日本語補完における自動評価結果では、ベースラインと比較し、BLEU 値において 0.77%の向上を確認できた。これにより、本研究の目的である、日英統計翻訳における主語補完の有効性が確認できたといえる。今後、適切な主語補完を行うためのアルゴリズムを考案し、日英統計翻訳に組み込むことで、日英統計翻訳の翻訳精度は向上すると考える。

10.2.4 日本語補完における人手評価の解析

人手評価において、主語補完を行うことで、翻訳精度が向上することが確認できた。しかし、主語補完によって翻訳精度が低下した文が存在する。そこで、主語補完を行うことで翻訳精度が向上する場合と、翻訳精度が低下する場合において、解析を行った。

日本語補完による翻訳精度の向上

日本語補完を用いることで翻訳精度が向上した原因は、2つ挙げられる。まず1つ目は、入力文の主語が省略されていることで出力されなかった主語が、主語補完を行うことで出力されることである。表 50 に例を示す。

表 50: 主語補完による主語出力改善の例

ベースライン入力文	戸をぱっと開いた。
日本語補完入力文	私は戸をぱっと開いた。
正解文	I threw open the door .
ベースライン出力文	open the door .
日本語補完出力文	I open the door .

表 50 の場合、ベースライン入力文は、主語が省略されている。そのため、主語が出力されず、ベースライン出力文は命令文になっている。しかし、主語補完を行うことによって、主語を出力し、翻訳精度が向上したことが分かる。

また2つ目は、入力文の主語が省略されている場合、うまくフレーズ対応をとることができないことが原因で出力に誤りが生じた。これに対し、主語補完を行うことでうまくフレーズ対応をとることができた点である。表 51 に例を示す。

表 51: 主語補完によるフレーズ対応改善の例

主語省略文入力文	平気でうそをつく。
日本語補完入力文	私は平気でうそをつく。
正解文	She doesn't scruple to tell a lie .
ベースライン出力文	It is a has .
日本語補完出力文	I tells lies without shame .

また、フレーズ対応を表 52 に示す。

表 52: フレーズ対応改善の例

“主語省略文”のフレーズ対応	主語補完した文のフレーズ対応
平気で has うそを It is a つく。 .	私 I は 平気 で うそ を つく。 tells lies without shame .

表 52 では、ベースラインでは、フレーズ対応がうまくとれていない。しかし、主語補完を行うことで、フレーズ対応が改善され、翻訳精度が向上したことが確認できる。

日本語補完による翻訳精度の低下

日本語補完を用いることで翻訳精度が低下した原因は、2つ挙げられる。まず1つ目は、主語補完の誤りによる翻訳精度の低下である。表 53 に例を示す。

表 53: 主語補完の誤りによる翻訳精度の低下

ベースライン入力文	雨になりそうだ。
日本語補完入力文	私は雨になりそうだ。
正解文	It is likely to rain .
ベースライン出力文	It looks like rain .
日本語補完出力文	I is likely to rain .

“雨になりそうだ。”の主語としては、“天気は”のようなものが好ましい。しかし、誤って文頭に“私は”を補完したため、出力文が誤っている。このような主語補完の誤りによる翻訳精度の低下は、適切な主語補完アルゴリズムの考案によって改善される。例えば、動詞や形容詞より、主語が人かどうかを判断し、人であると判断した場合のみ“私は”を補完するというような方法で改善が可能であると考えられる。

2つ目に、“主語省略文”翻訳を翻訳した際に、すでに出力文にうまく主語が翻訳されている場合、主語を補完することで、誤ったフレーズが出力されてしまうことである。表 54 に例を示す。

表 54: 主語補完の誤りによる翻訳精度の低下

ベースライン入力文	泥棒に入られた。
日本語補完入力文	私は泥棒に入られた。
正解文	I had a burglar break into my home .
ベースライン出力文	My house was robbed .
日本語補完出力文	My house was robbed of my life .

表 54 では、入力文に主語が省略されている場合でも、主語として “My house” が出力されている。この場合に入力文に対し主語補完を行うと、“私は” がうまく翻訳されず、“of my life” のように、誤ったフレーズが出力されてしまう。この翻訳精度の低下は、学習データには “主語省略文” が含まれているので、“主語省略文” が入力された場合でも主語を出力し、適切な翻訳を行う場合があることが原因であると考えられる。このような場合、“主語省略文” に主語補完を行うことで、誤ったフレーズを出力する。

このような原因で翻訳精度の低下が見られる文はわずかである。しかし、統計翻訳を実用する際に主語補完を行う場合、この翻訳精度の低下を無視することはできないと考える。この問題に対しては、学習データに主語補完を行うことにより改善できると考えるが、train 補完の実験結果で翻訳精度の低下を確認したため、今後新たな対策の考案が必要であると考えられる。

10.3 対訳補完についての考察

対訳補完では、4章で定義した条件1～条件3の、3つの条件を満たす場合、主語が省略されている文として抽出した。そして、抽出した文の対訳英語文の文頭を参照し、主語を補完した。この対訳補完についての考察を以下に記述する。

10.3.1 対訳補完における主語補完の精度

対訳補完で主語補完を行った1,109文より、ランダムに100文抽出し、主語補完の精度評価を行った。評価は、10.2.2節と同じ評価基準で行う。表55に結果を示す。

表 55: 主語補完評価結果

主語補完○	主語補完×
96	4

また、主語補完が正しくない例文を表56に示す。

表 56: 主語補完誤り文例

対訳補完日本語文1	“それは”天気になる。
対訳英語文1	It clears up .
対訳補完日本語文2	“それは”犬にコートを着せるなんて過保護だ。
対訳英語文2	It is overprotective to put a coat on a dog .
対訳補完日本語文3	“誰かが”誰かに足を踏まれた。
対訳英語文3	Somebody stepped on my foot .

表56の対訳補完日本語文1では、“それは”を補完している。しかしこの文では、日本語としては、“それは”を補完すべきではない。対訳補完日本語文2についても同様に、“それは”を補完すべきではない。また、対訳補完日本語文3では、“誰かが”を補完している。ここで、対訳英語文2を日本語に変換すると、“誰かが私の足を踏んだ。”となる。しかし、テストデータの日本語は、“誰かに足を踏まれた”と、受身の構造となっている。これが原因で、主語補完が誤っている。

表55より、対訳補完でも、主語補完に誤りがある文が存在することが確認できた。しかし、本研究で提案した対訳補完アルゴリズムでは、これらの主語補完の誤りを改善するのは難しい。これより、より主語補完の精度を向上させるためには、新たな主語補完アルゴリズムを考案しなければならないと考えている。

10.3.2 対訳補完における自動評価の解析

対訳補完を用いて主語補完を行った1,109文での自動評価では、ベースラインと比較し、BLEUにおいて、2.46%向上した。これは、表34で示した、日本語補完を行った1,526文のみを用いた実験の自動評価結果を大きく上回っている。これより、適切な主語補完を行うことができれば、翻訳精度が大きく向上すると考える。

10.4 train 補完についての考察

日本語補完と比較し、train 補完では、BLEU 値の向上が見られなかった。学習データに主語補完を行うと、学習データの“主語省略文”が大幅に減少する。そのとき、テストデータに残っている“主語省略文”の翻訳精度が下がり、train 補完の効果が見られない原因の一つとして考えられる。また、主語補完に誤りがあることが原因の一つとして考えられる。主語補完の精度をさらに上げることで、train 補完が有効になる可能性があると考えている。これについては、さらなる解析が必要である。

11 今後の研究

以下に，今後行う研究をまとめる．

実験 1 適切な主語補完を行うアルゴリズムを考案し，統計翻訳の精度を向上を目指す．

実験 2 本研究では単文のみの実験を行ったが，重文複文についての実験を考えている

実験 3 原言語と目的言語の文法構造が近いほど，翻訳精度が向上することが分かっている [17]．これより，主語補完を行った後，主語を動詞の前に移動させることで翻訳精度が向上するのではないかと考えられる．このような，主語補完と文法構造変換の手法を併用した実験を行いたいと考えている．

12 おわりに

本研究では、主語が省略されている日本語文への主語補完による翻訳精度の変化を調査した。まず、主語が省略されている文の文頭に“私は”を補完した。その結果、翻訳精度の向上を確認することができた。また、主語が省略されている文に対して、対訳英語文を参照することにより、精度の高い主語補完を行った。その結果、自動評価において、大きな翻訳精度の向上を確認することができた。さらに、テストデータの主語が省略されている文の文頭に“私は”を補完した翻訳の解析を行った。その結果、主語補完を行うことで翻訳精度が向上する原因として、主語出力の改善と、フレーズ対応の改善が挙げられることが分かった。結論として、主語補完することにより翻訳精度が向上し、主語補完の有効性が確認できた。

謝辞

最後に、一年間に渡り、本研究の御指導をいただきました鳥取大学工学部知能情報工学科計算機講座C研究室の村田真樹教授，村上仁一准教授，徳久雅人講師に深く感謝するとともに厚くお礼を申し上げます。また，計算機工学講座C研究室の皆様にも厚くお礼を申し上げます。また，参考にさせて頂いた論文の著者の方々に対して，深く感謝します。

参考文献

- [1] 猪澤雅史, 村上仁一, 徳久雅人, 池原悟, “文節区切りの学習データを用いた, 日英統計翻訳の検討”, 言語処理学会年次大会発表論文集, pp1022-1025,2010.
- [2] GIZA++
<http://www.fjoch.com/GIZA++.html>
- [3] Philipp Koehn, Franz J. Och, and Daniel Marcu, “Statistical phrase-based translation”, In Proceedings of HLT-NAACL 2003, pp.127-133, 2003.
- [4] Franz Josef Och, “Minimum error rate training for statistical machine translation”, Proceedings of the ACL, 2003.
- [5] 中岩浩巳, 池原悟, “日英翻訳システムにおける用言意味属性を用いたゼロ代名詞照応解析”, 情報処理学会論文誌 34(8), pp1705-1715, 1993.
- [6] 中岩浩巳, 池原悟, “語用論的・意味論的制約を用いた日本語ゼロ代名詞の文内照応解析”, 自然言語処理, Vol.3, No.4, pp50-65, 1996.
- [7] 中岩浩巳, “日英対訳コーパスを用いたゼロ代名詞とその指示対象の自動認定” 自然言語処理研究会 (NL-123), pp33-40, 1998.
- [8] 中岩浩巳 “日英機械翻訳におけるゼロ代名詞照応解析に関する研究”, 2002.
- [9] Philipp Koehn, Marcello Federico, Brooke Cowan, Richard Zens, Chris Dyer, Ondřej Bojar, Alexandra Constantin, Evan Herbst, “Moses: Open Source Toolkit for Statistical Machine Translation”, Proceedings of the ACL 2007 Demo and Poster Sessions, pages 177-180, June 2007.
- [10] Jin'ichi Murakami, Masato Tokuhisa, Satoru Ikehara, “Statistical Machine Translation using Large J/E Parallel Corpus and Long Phrase Tables”, International Workshop on Spoken Language Translation 2007, pp.151-155, 2007
- [11] SRILM, The SRI Language Modeling Toolkit
<http://www.speech.sri.com/projects/srilm/>
- [12] MeCab <http://mecab.sourceforge.net/>
- [13] CaboCha : Yet Another Japanese Dependency Structure Analyzer
<http://chasen.org/~taku/software/cabocha>
- [14] BLEU, NIST Open MT Scoring
<http://www.itl.nist.gov/iad/894.01/tests/mt/2008/scoring.html>

- [15] METEOR, The METEOR Automatic Machine Translation Evaluation System
<http://www-2.cs.cmu.edu/~alavie/METEOR/>
- [16] NIST, Automatic Evaluation of Machine Translation Quality Using N-gram Co-Occurrence Statistics
<http://www.itl.nist.gov/iad/mig/tests/mt/2008/scoring.html>
- [17] Sara Stymne, “Definite Noun Phrases in Statistical Machine Translation into Danish”, Workshop on Extracting and Using Constructions in NLP, pp4-9, 2009.