

日英統計翻訳における主語補完の効果

古市将仁 村上仁一 徳久雅人 村田真樹
鳥取大学 工学部 知能情報工学科

{s072047, murakami, tokuhisa, murata} @ ike.tottori-u.ac.jp

1 はじめに

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。しかし、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 [1]。

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)。主語補完を行うことによって、どのように翻訳精度が変化するか調査する。

2 “主語省略文”の判断条件

日本語文において、主語が省略されている文が存在する。本論文では、主語が省略されている日本語文を“主語省略文”と呼ぶ。以下に“主語省略文”の例を示す。

表1 “主語省略文”例

昼食をたっぷり取った。
サッカーをした。

日本語では、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。また，“しろ”や，“しなさい”のような動詞を含む命令文には、主語が無いことが多い。以上を踏まえ、本論文において日本語文を“主語省略文”であると判断する条件を以下に示す。

判断条件 1 助詞“は”，もしくは“が”，もしくは“も”が含まれない文

判断条件 2 動詞が命令形ではない文

3 主語補完

本研究では、テストデータ 10,000 文に対し、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。以下に、日本語補完と対訳補完における主語補完の方法を示す。

3.1 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。以下に、主語補完の手順を示す。

手順 1 テストデータに対し、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、2 章で定義した判断条件 1 と判断条件 2 を満たす“主語省略文”を抽出する。以下に抽出される文と抽出されない文の例を示す。

表2 日本語補完抽出文例

抽出される文(主語省略文)	昼食をたっぷり取った。
抽出されない文	彼は山へ行く。

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3 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	私は昼食をたっぷり取った。

3.2 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。対訳補完における主語補完の手順を以下に示す。

手順 1 テストデータに対し、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、判断条件 1 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4 対訳補完“主語省略文”例

昼食をたっぷり取った。
声を低くしなさい。

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 1 と図 2 に示す。

主語無し文: 昼食をたっぷり取った。
対訳英語文: [He] had a big lunch .
↓ 抽出

図1 文頭単語抽出例 1

主語無し文: 声を低くしなさい。
対訳英語文: [Keep] your voice down .
↓ 抽出

図2 文頭単語抽出例 2

手順 4 抽出した正解文の文頭単語を、変換規則に従って日本語に変換する。変換規則を表 5 に示す。なお、図 2 のように、抽出した対訳英語文の文頭単語が表 5 の変換規則に適用しなかったときは、主語補完は行わない。

表5 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 6 に示す。

表6 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。

4 実験方法

4.1 翻訳手順

翻訳手順を以下に示す。

手順1 テストデータの日本語文に対し、主語補完を行う。

手順2 主語補完を行ったテストデータを用いて統計翻訳を行う。

また、翻訳手順を図3に示す。

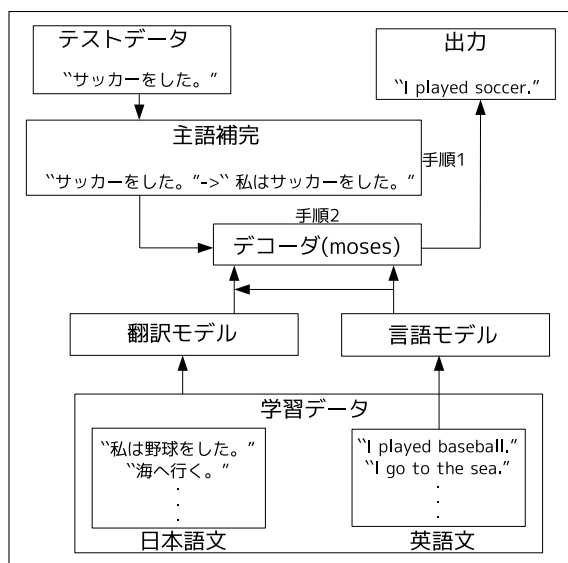


図3 翻訳手順

5 実験環境

5.1 デコーダのパラメータ

デコーダには、Moses[3]を用いる。また本研究では、パラメータチューニング[4]を行わない。

5.2 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[5]を用いて分かち書きを行う。また、各コーパスの英語文に対して、tokenizer.perl[3]を用いて、分かち書きを行う。表7に日英対訳文の例を示す。

表7 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	私は猫を1匹飼っている。
英語文	I have a cat .

5.3 評価方法

5.3.1 自動評価

本研究では、出力文の自動評価に、BLEU[6], METEOR[7], NIST[8]を使用する。

5.3.2 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文に対する出力文よりランダムに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

6 実験結果

6.1 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した文数を表8に示す。

表8 主語補完文数

	主語補完文数
日本語補完	1,526
対訳補完	1,109

6.2 自動評価結果

a) 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、対訳補完における自動評価の結果を表9に示す。テストデータは10,000文用いる。

表9 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
対訳補完	0.1218	0.4153	4.3579

表9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。

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の評価結果を表10に示す。

表10 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

表10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9と比較し、BLEU値の向上が大き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

対訳補完で主語補完した 1,109 文を用いた実験結果を表 11 に示す。

表 11 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

表 11 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。

6.3 対比較評価結果

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。結果を表 12 に示す。

表 12 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。

6.4 対比較評価例文

対比較評価を行った出力文の例を示す。

6.4.1 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 13 と表 14 に示す。

表 13 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	平気うそをつく。
日本語補完入力文	私は平気うそをつく。
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 .

表 14 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 .

6.4.2 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 15 と表 16 に示す。

表 15 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 16 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 .

6.4.3 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 17 と表 18 に示す。

表 17 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 18 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 .

7 考察

7.1 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 19 に調査結果を示す。

表 19 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 19 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 19 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 10 で、“私は”のみを補完した日本語補完において、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。

7.2 “主語省略文”抽出の精度

日本語補完で抽出した“主語省略文”1,526 文より、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。精度調査の評価基準を以下に示す。

- 抽出○ 抽出した“主語省略文”に主語が無い場合
- 抽出× 抽出した“主語省略文”に主語がある場合

表 20 に結果を示す。

表 20 “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 21 に示す。

表 21 “主語省略文”抽出誤り文例

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、キアラーって作家の書いた本です。

表 21 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 21 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 21 のような文は少なく、表 20 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。

7.3 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 22 に結果を示す。

主語補完○	主語補完×
93	7

主語補完が正しくない例文を表 23 に示す。表 23 では、日本語補完により補完された主語は、“私は”である。

文例 1	“私は”たばこの消し忘れて大火事になった。
文例 2	“私は”みんな口をぽかんとあけた。

表 23 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。また、表 23 の文例 2 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 22 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。

7.4 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、7.3 節と同じ評価基準で行う。表 24 に結果を示す。

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 25 に示す。

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	“誰かが”誰かに足を踏まれた。
対訳英語文 2	Somebody stepped on my foot .

表 25 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”を補完すべきでない。また、表 25 の対訳補完日本語文 2 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 24 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。

7.5 学習データへの主語補完

6 章の実験では、テストデータのみ主語補完を行った。そこで、学習データに対して主語補完を行うことによる翻訳精度の変化を調査した。

a) 学習データへの主語補完結果

テストデータと同様に、学習データに対して主語補完を行った。主語補完の方法は、3.2 節の対訳補完と同様に行った。表 26 に主語補完結果を示す。

学習データ文数	100,000
主語補完文数	10,600

b) 評価結果

テストデータに対して日本語補完で主語補完を行い、学習データに対しても主語補完を行う手法を、train 補完 (日本語補完) と呼ぶ。また、テストデータに対して対訳補完で主語補完を行い、学習データに対しても主語補完を行う手法を、train 補完 (対訳補完) と呼ぶ。

ベースライン、日本語補完、train 補完 (日本語補完)、対訳補完、train 補完 (対訳補完) に対し、自動評価を行った。表 27 に結果を示す。

	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
train 補完 (日本語補完)	0.1187	0.4060	4.2645
対訳補完	0.1218	0.4153	4.3579
train 補完 (対訳補完)	0.1209	0.4135	4.3266

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”の翻訳精度が下がると考えられ、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。

8 おわりに

本研究では、主語の無い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。結果として、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。

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] CaboCha : Yet Another Japanese Dependency Structure Analyzer
<http://chasen.org/~taku/software/cabocha>
- [3] Moses,moses.2007-05-29.tgz <http://www.statmt.org/moses/>
- [4] Franz Josef Och, “Minimum error rate training for statistical machine translation”, Proceedings of the ACL, 2003.
- [5] MeCab <http://mecab.sourceforge.net/>
- [6] BLEU, NIST Open MT Scoring
<http://www.itl.nist.gov/iad/894.01/tests/mt/2008/scoring.html>
- [7] METEOR, The METEOR Automatic Machine Translation Evaluation System
<http://www-2.cs.cmu.edu/~alavie/METEOR/>
- [8] NIST, Automatic Evaluation of Machine Translation Quality Ising N-gram Co-Occurrence Statistics
<http://www.itl.nist.gov/iad/mig/tests/mt/2008/scoring.html>
- [9] Sara Stymne, “Definite Noun Phrases in Statistical Machine Translation into Danish”, Workshop on Extracting and Using Constructions in NLP, pp4-9, 2009.