

平成 29 年度特許庁委託事業

拒絶理由通知等の日英機械翻訳に関する調査  
報告書

平成 30 年 3 月

一般財団法人 日本特許情報機構

# 目次

1. 概要 .....	3
1.1 目的 .....	3
1.2 調査概要 .....	3
2. 調査内容 .....	4
2.1 翻訳対象 .....	4
2.2 調査方法 .....	4
2.2.1 日英対訳コーパスの作成 .....	4
2.2.2 評価対象となる日本語文の選定 .....	4
2.2.3 学習コーパスの準備 .....	5
2.2.4 学習コーパスの分割 .....	6
2.3 評価対象となる機械翻訳文の取得 .....	6
2.3.1 統計機械翻訳 (SMT) .....	6
2.3.2 ニューラル機械翻訳 (NMT) .....	8
2.4 機械翻訳文の精度評価 .....	10
2.4.1 人手評価 .....	10
2.4.2 自動評価 .....	10
2.5 機械翻訳文の分析 .....	10
3. 調査結果 .....	11
3.1 機械翻訳文の精度評価 .....	11
3.1.1 人手評価 .....	11
3.1.2 自動評価 .....	13
3.2 機械翻訳文の分析 (統計機械翻訳) .....	16
3.2.1 機械翻訳の誤りの傾向 .....	16
3.2.2 学習コーパスの登録による改善点や改悪点 .....	23
3.2.3 同一の語の翻訳によるぶれの程度 .....	29
3.2.4 日英対訳コーパスの学習量に応じた機械翻訳精度の変化 .....	31
3.3 統計機械翻訳とニューラル機械翻訳の比較 .....	35
3.3.1 比較対象とした機械翻訳 .....	35
3.3.2 評価値の比較 .....	35
3.3.3 個別例の分析 .....	35
4. 精度向上のための検討 .....	43
4.1 係り受け誤りへの対応 .....	43
4.1.1 日本語構文解析の精度向上 .....	43
4.1.2 日本語の書き方の工夫 .....	43

4.2 学習コーパスの整備.....	43
4.2.1 文分割に誤りがある学習コーパスの除外.....	44
4.2.2 数字が対応しない学習コーパスの除外.....	44
4.3 未知語の取り扱いについて.....	45
5. 付録 .....	46
5.1 自動評価手法の概要.....	46
5.1.1 BLEU .....	46
5.1.2 Smoothed BLEU .....	47
5.1.3 NIST.....	48
5.1.4 RIBES .....	50
5.1.5 BLEU と RIBES の特徴.....	52
5.2 日英対訳コーパスの内訳（フォーマット A） .....	55

## 1. 概要

### 1.1 目的

特許庁は、外国特許庁と特許出願の手續や審査に関連する情報（ドシエ情報）を相互利用する取組みとして「高度産業財産ネットワーク」（AIPN）や「ワン・ポータル・ドシエ」（OPD）を運用し、外国特許庁に対して、拒絶理由通知書などのドシエ情報を、英語に機械翻訳して提供している。近年、海外庁からドシエ情報の機械翻訳文に対する精度向上の要望が上がっている他、平成 28 年 5 月に取りまとめられた報告書<sup>1</sup>において、海外庁の審査官が我が国の審査結果を参酌しやすいよう機械翻訳の取組を強化し、高精度な機械翻訳文を提供することにより、我が国の審査結果の利用を促進していくべきとの方針が示されるなど、ドシエ情報の翻訳精度のさらなる改善が求められている。

従来特許庁では、特許公報をもとに辞書・対訳コーパス等を作成する等して、特許の公報の機械翻訳の精度向上に寄与してきたが、ドシエ情報の翻訳精度の向上については十分検討がなされていない。拒絶理由通知等のドシエ情報は、特許公報と文体や使用される用語が異なる点が多いため、特許公報をもとに作成した辞書・対訳コーパスを使用するだけでは、十分な翻訳精度が得られない可能性が高い。

そこで、本調査では、ドシエ情報の中で重要な情報である拒絶理由通知について、日英機械翻訳をする際の課題を調査・分析することを目的とする。併せて、本事業を通じて拒絶理由通知をもとにした日英対訳コーパスを作成し、機械翻訳システムの精度向上に寄与させることを目的とする。

### 1.2 調査概要

特許庁が貸与する日本語の拒絶理由通知について英語の人手翻訳文を作成し、拒絶理由通知の機械翻訳文と人手翻訳文とを比較することにより、機械翻訳の課題を調査・分析する。また、本事業で作成した人手翻訳文をもとに、日英対訳コーパスを作成し、その日英対訳コーパスの有効性を検証する。

---

<sup>1</sup> 産業構造審議会 知的財産分科会 情報普及活用小委員会 報告書「特許情報のさらなる活用に向けて」（平成 28 年 5 月）

## 2. 調査内容

### 2.1 翻訳対象

本事業では、特許法第 29 条第 1 項第 3 号（新規性無し）または特許法第 29 条第 2 項（進歩性無し）の規定により特許を受けることができないことを通知した拒絶理由通知を翻訳の対象とし、特許庁から貸与された翻訳対象リストにて指定された 2,553 件の拒絶理由通知の翻訳を行った。

### 2.2 調査方法

#### 2.2.1 日英対訳コーパスの作成

2.1 において作成した翻訳文 2,553 件から、評価・分析用として 400 件を選定した。選定にあたり、拒絶理由通知データよりオフィスコードを抽出し、オフィスコードに偏りが無いようにした。選定した 400 件について特許庁の仕様に従い、「出願番号」、「書類番号」、「日本語文」、「英語文」、「原文（日本語）文字数」、「文の種類（認定・対比・判断）」の欄を設けた日英対訳コーパス 5,407 文対（フォーマット A）を作成した<sup>2</sup>。残る 2,153 件からは「文の種類（認定・対比・判断）」の欄を含まない日英対訳コーパス 30,363 文対（フォーマット B）を作成した。作成された日英対訳コーパスは、フォーマット A および B を合わせると 35,770 文対であった。

#### 2.2.2 評価対象となる日本語文の選定

2.2.1 で作成した日英対訳コーパス 5,407 文対（フォーマット A）から、評価対象とする文を 200 文選定した。選定にあたり文長<sup>3</sup>（短文、中文、長文）、審査部、文の種類（認定、対比、判断）が均等になるよう選定を行った。また、評価対象文のうち、100 文を人手評価の対象として選定した。

---

<sup>2</sup> 日英対訳コーパス（フォーマット A）の内訳は 5.2 を参照のこと。

<sup>3</sup> 評価対象の 1 文あたりの日本語文字数について集計を行い、1 文の文字数により文を短文（64 文字以下）、中文（65-116 文字）、長文（117 文字以上）の 3 区分とし、区分毎の文数が同じとなるようにした。

表 2.2-1 評価対象文 (200 文)

審査第一部	認定	対比	判断	合計
短文	6	5	6	17
中文	6	5	6	17
長文	6	5	6	17
小計	18	15	18	51
審査第二部	認定	対比	判断	合計
短文	6	4	6	16
中文	6	6	6	18
長文	6	6	5	17
小計	18	16	17	51
審査第三部	認定	対比	判断	合計
短文	6	5	5	16
中文	6	5	5	16
長文	5	5	5	15
小計	17	15	15	47
審査第四部	認定	対比	判断	合計
短文	6	5	6	17
中文	6	5	6	17
長文	6	5	6	17
小計	18	15	18	51
合計	71	61	68	200

### 2.2.3 学習コーパスの準備

機械翻訳文の取得にあたっては、特許庁から貸与された見解書コーパス<sup>4</sup>および 2.2.1 で作成した日英対訳コーパスを学習コーパスとして準備した。

学習コーパスの作成にあたり、本調査で使用する「みんなの自動翻訳@TexTra®<sup>5</sup> (以降、「みんなの自動翻訳」と記載)」において、オリジナル機械翻訳を作成する際に登録できるコーパスの最大長は 1 文あたり 1,000 バイトであるため、学習コーパスとして文長が 1,000 バイト以下のものを選定した。

また、日本語文・英語文がともに同じとなる重複データは削除した。さらに、見解書コーパスは日本語と英語が 1 文対 1 文で対応しているもののみ使用し、文の対応具合を示すスコア値がゼロのデータは除外するなど、コーパスとして不適切なものを除いた。

<sup>4</sup> 特許庁で作成した PCT 見解書の原文と、WIPO で作成した英語翻訳文から作成された対訳コーパス。

<sup>5</sup> <https://mt-auto-minhon-mlt.ucri.jgn-x.jp/content/menu/>

#### 2.2.4 学習コーパスの分割

本調査では学習コーパス量の増加に応じた機械翻訳精度の変化を調査するため、表 2.2-2 のように見解書コーパスを 3 分割した（学習コーパス A-C）。分割する際、文長や文の内容に偏りがないようにするため、文長の平均が均等となるようにした。

また、2.2.1 で作成した日英対訳コーパス（フォーマット A）から、2.2.2 で選定した評価対象文を除いたものを学習コーパス D、2.2.1 で作成した日英対訳コーパス（フォーマット B）から、2.2.2 で選定した評価対象文と日本語が一致した文対を除いたものを学習コーパス E とした。

表 2.2-2 学習コーパス一覧

学習コーパス名	コーパスの内容	コーパス文対数
学習コーパス A	見解書コーパスの 1/3 程度	571,446
学習コーパス B	見解書コーパスの 1/3 程度	571,448
学習コーパス C	見解書コーパスの 1/3 程度	571,443
学習コーパス D	2.2.1 にて作成した日英対訳コーパス（フォーマット A、但し 2.2.2 において自動評価の対象とした文は除く）	2,292
学習コーパス E	2.2.1 にて作成した日英対訳コーパス全て（フォーマット B、但し 2.2.2 において自動評価対象とした文と日本語が一致したものは除く）	20,395

### 2.3 評価対象となる機械翻訳文の取得

#### 2.3.1 統計機械翻訳（SMT）

##### (1) 機械翻訳環境

統計機械翻訳では、特許庁仕様書に基づき「みんなの自動翻訳」を使用した。表 2.2-2 に記載の見解書コーパス（学習コーパス A-C）について、2017 年 12 月に登録および学習を行い、表 2.3-1 に記載の統計機械翻訳 A-C を得た。表 2.2-2 に記載の本事業で作成した日英対訳コーパス（学習コーパス D-E）については、2018 年 1 月に登録・訓練を行い、表 2.3-1 に記載の統計機械翻訳 E を得た。

## (2) 学習・登録条件

表 2.3-1 に学習・登録条件<sup>6</sup>を示す。

表 2.3-1 機械翻訳一覧（みんなの自動翻訳）

機械翻訳名 <sup>7</sup>	登録・学習したデータ
統計機械翻訳 A ( SMT-A )	・ 学習コーパス A
統計機械翻訳 B ( SMT-B )	・ 学習コーパス A ・ 学習コーパス B
統計機械翻訳 C ( SMT-C )	・ 学習コーパス A ・ 学習コーパス B ・ 学習コーパス C
統計機械翻訳 E ( SMT-E )	・ 学習コーパス A ・ 学習コーパス B ・ 学習コーパス C ・ 学習コーパス D ・ 学習コーパス E

## (3) 機械翻訳文の取得

「みんなの自動翻訳」では、まず図 2.3-1 に示すような対訳データ登録画面から各学習コーパスを登録した。次に、図 2.3-2 に示すオリジナル機械翻訳作成画面にて学習を行い、表 2.3-1 に示した学習データを登録した統計機械翻訳 A-E を作成した。作成した統計機械翻訳を使用して、2.2.2 で選定された評価対象文の機械翻訳文を取得した。

<sup>6</sup> 「みんなの自動翻訳」の仕様変更に伴い、特許庁担当者と相談の上、記載の条件により機械翻訳文を取得することとした。

<sup>7</sup> カッコ内の表記はグラフ等で使用する機械翻訳の記号名称となる。





図 2.3-1 対訳データ登録画面

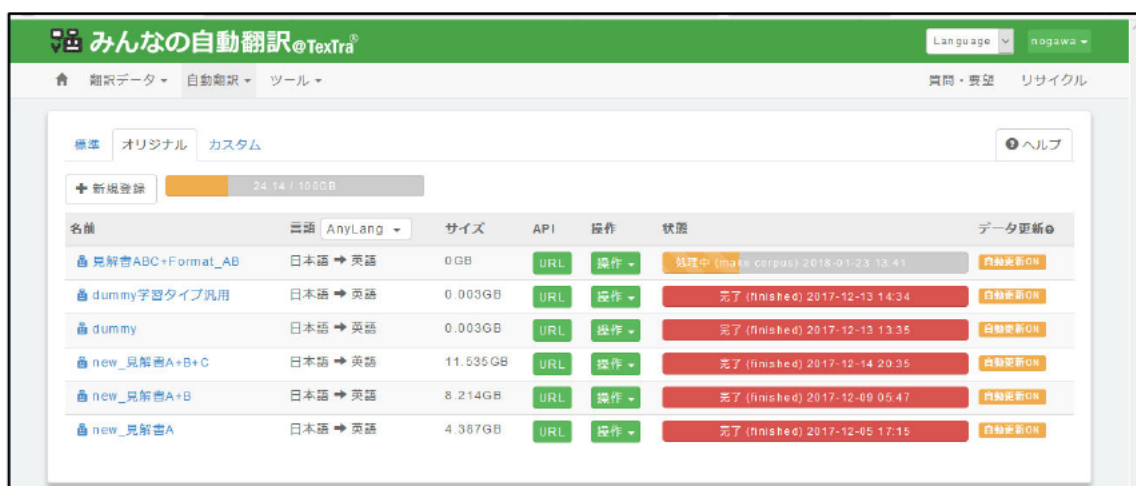


図 2.3-2 オリジナル機械翻訳作成画面

## 2.3.2 ニューラル機械翻訳 (NMT)

### (1) 機械翻訳環境

ニューラル機械翻訳では、オープンソースで利用可能な OpenNMT<sup>8</sup>を使用した。表 2.2-2 に記載の見解書コーパス（学習コーパス A-C）および本事業で作成した日英対訳コーパス（学習コーパス D-E）について、2018 年 1 月に登録および学習を行い、表 2.3-2 に記載のニューラル機械翻訳 A-E を得た。

<sup>8</sup> <http://opennmt.net/>（2017 年 5 月ダウンロード時点でのバージョンを使用。）

## (2) 学習・登録条件

表 2.3-2 に学習・登録条件 を示す。

表 2.3-2 機械翻訳一覧 (OpenNMT)

機械翻訳名 <sup>9</sup>	登録・学習したデータ
ニューラル機械翻訳 A ( NMT-A )	・ 学習コーパス A
ニューラル機械翻訳 B ( NMT-B )	・ 学習コーパス A ・ 学習コーパス B
ニューラル機械翻訳 C ( NMT-C )	・ 学習コーパス A ・ 学習コーパス B ・ 学習コーパス C
ニューラル機械翻訳 E ( NMT-E )	・ 学習コーパス A ・ 学習コーパス B ・ 学習コーパス C ・ 学習コーパス D ・ 学習コーパス E

## (3) 機械翻訳文の取得

OpenNMT では、原則として標準設定を使用して各学習コーパスの登録および学習を行い、表 2.3-2 に示した学習データを登録したニューラル機械翻訳 A-E を作成した。作成したニューラル機械翻訳を使用して、2.2.2 で選定された評価対象文の機械翻訳文を取得した。

OpenNMT の標準使用と異なる点について以下に示す。

オプション設定で ' -dbrrn ' を使用

Default の RNN ではなく、deep bidirectional RNN を使用。

未知語対応

未知語<sup>10</sup>があった場合に、対訳辞書を学習データとした文字単位のニューラル機械翻訳を使用して未知語を翻訳する処理を追加。

<sup>9</sup> カッコ内の表記はグラフ等で使用する機械翻訳の記号名称となる。

<sup>10</sup> ここでの未知語とは、翻訳結果中に原文の日本語がそのまま出力された部分を意味する。

## 2.4 機械翻訳文の精度評価

### 2.4.1 人手評価

2.3 で作成した統計機械翻訳文のうち、人手評価の対象とした各 100 文について、特許庁が公表している「特許文献機械翻訳の品質評価手順 ( ver1.0 )<sup>11</sup>」に従って、人手により「内容の伝達レベル」の観点での評価を行った。内容伝達レベルの人手評価結果は 5 段階評価とした ( 翻訳精度の最も高いものが評価値 5、最も低いものが評価値 1 )。また、評価は 2 名で行い、その平均値を評価結果とした。

### 2.4.2 自動評価

2.3 で作成した統計機械翻訳文およびニューラル機械翻訳文のうち、自動評価の対象とした各 200 文について、自動評価を行った。

自動評価にあたっては、代表的な自動評価基準である BLEU および RIBES<sup>12</sup>の 2 つの自動評価手法を利用した。

## 2.5 機械翻訳文の分析

2.4 で得られた結果をもとに、「機械翻訳の誤りの傾向」、「日英対訳コーパスの登録による改善点や改悪点」、「同一の語の翻訳によるぶれの程度」、「日英対訳コーパスの学習量に応じた機械翻訳精度の変化」について調査・分析を行い、拒絶理由通知を日英機械翻訳する際の課題とその解決策を検討した。

---

<sup>11</sup> [http://www.jpo.go.jp/shiryou/toushin/chousa/tokkyohonyaku\\_hyouka.htm](http://www.jpo.go.jp/shiryou/toushin/chousa/tokkyohonyaku_hyouka.htm)

<sup>12</sup> 5.1 参照。

### 3. 調査結果

#### 3.1 機械翻訳文の精度評価

##### 3.1.1 人手評価

###### (1) 全体

全体評価の結果を図 3.1-1 に示す。学習コーパス量が最も少ない統計機械翻訳 A と比較し、統計機械翻訳 B では評価値の上昇が見られたが、以降、学習コーパス量を増やしても、評価値の上昇は僅かなものであった。

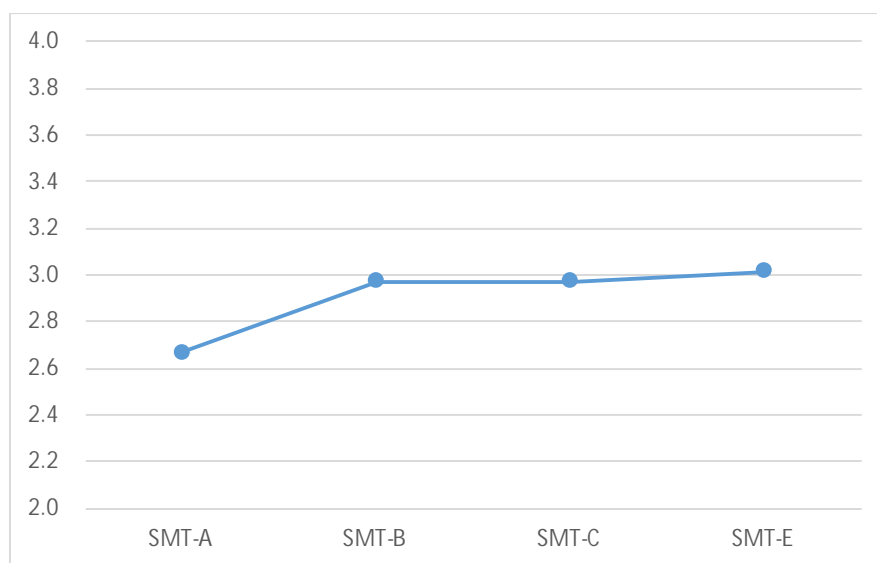


図 3.1-1 人手評価（全体）

###### (2) 審査部別

審査第二部は統計機械翻訳 B 以降、他の審査部に比べ評価値が高い傾向が見られた。それ以外の審査部はほぼ同様の評価結果となった。

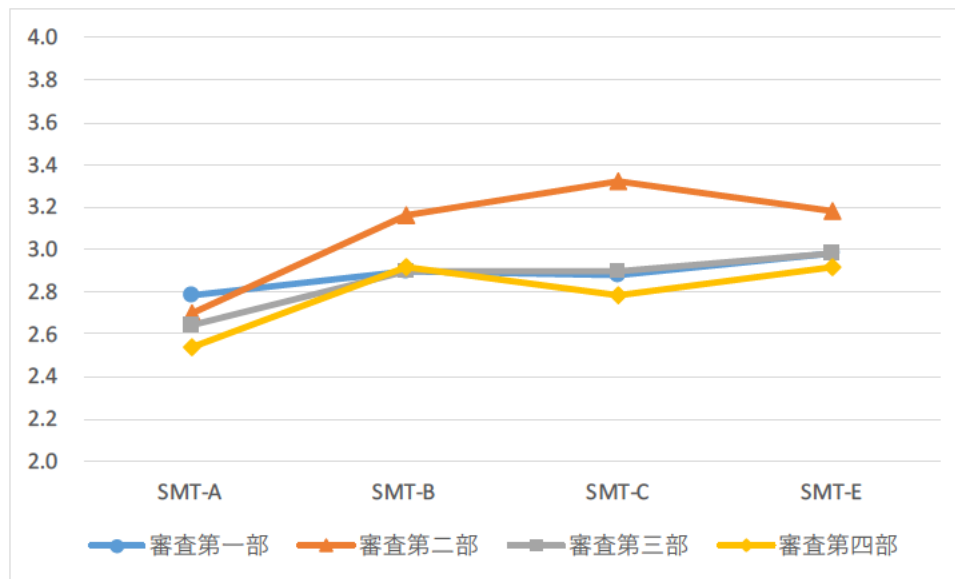


図 3.1-2 人手評価（審査部別）

### (3) 文長別

学習コーパスの量に関わらず、短文、中文、長文の順に評価値が高い結果となった。また、中文は学習コーパス量の増加に従い、評価値の上昇が見られた。一方、短文および長文は統計機械翻訳 B 以降、学習コーパスを増やしても評価値の上昇は見られなかった。

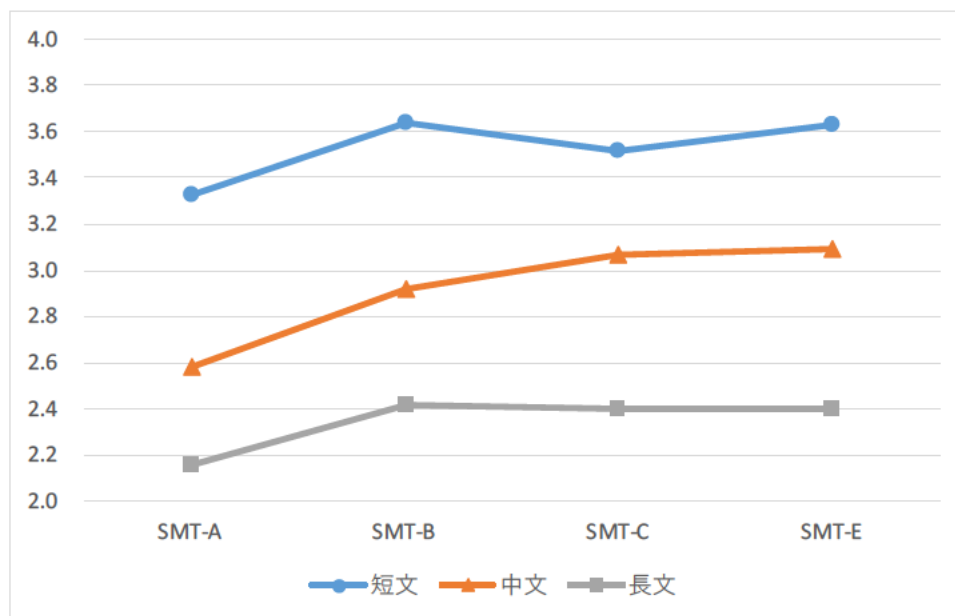


図 3.1-3 人手評価（文長別）

#### (4) 文の種類別

文の種類別に結果を見ると、「判断」の文は「認定」、「対比」に比べ評価値が高い傾向が見られた。これは、「判断」の文長が平均 78.1 文字と、「認定」の文長 79.7 文字、「対比」の文長 107.8 文字に比べ短く、また文の表現が比較的定型的なことがその要因と考えられる。

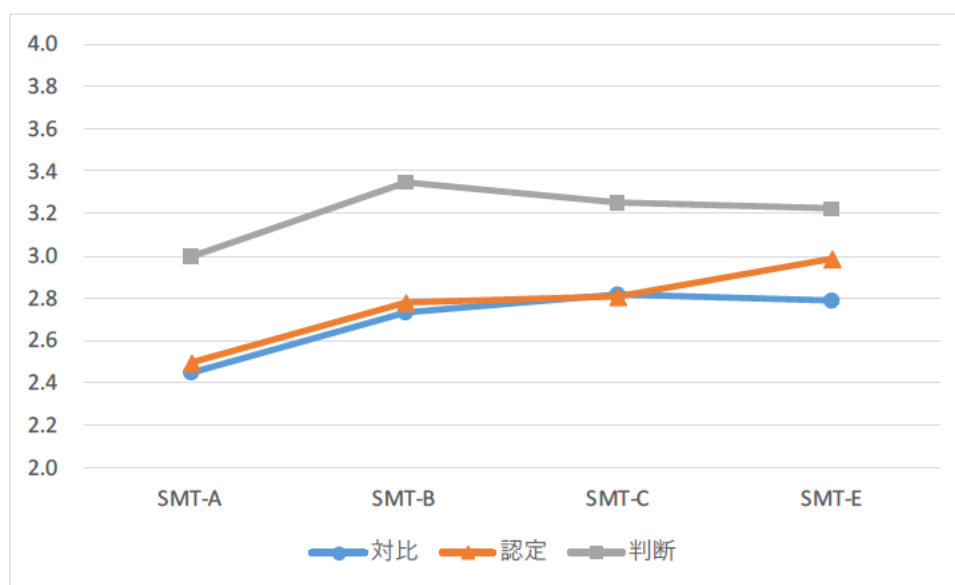


図 3.1-4 人手評価（文の種類別）

#### 3.1.2 自動評価

##### (1) 学習コーパス量と自動評価値との関係（SMT・NMT）

自動評価結果を見ると、BLEU, RIBES とともに学習コーパス量の増加に従い、評価値の上昇が見られた。SMT/NMT-A と SMT/NMT-C とでは、コーパスの学習量の差が 100 万コーパス以上あるのに対し、SMT/NMT-C と SMT/NMT-E とでは、コーパスの学習量の差は 22,500 コーパス程度であるのにもかかわらず、BLEU, RIBES の上昇量は後者の方が大きい。このことは、本事業で作成した日英対訳コーパスが、見解書コーパスと比較して、拒絶理由通知の翻訳精度の効率的な向上に寄与することを示唆するものと考えられる。

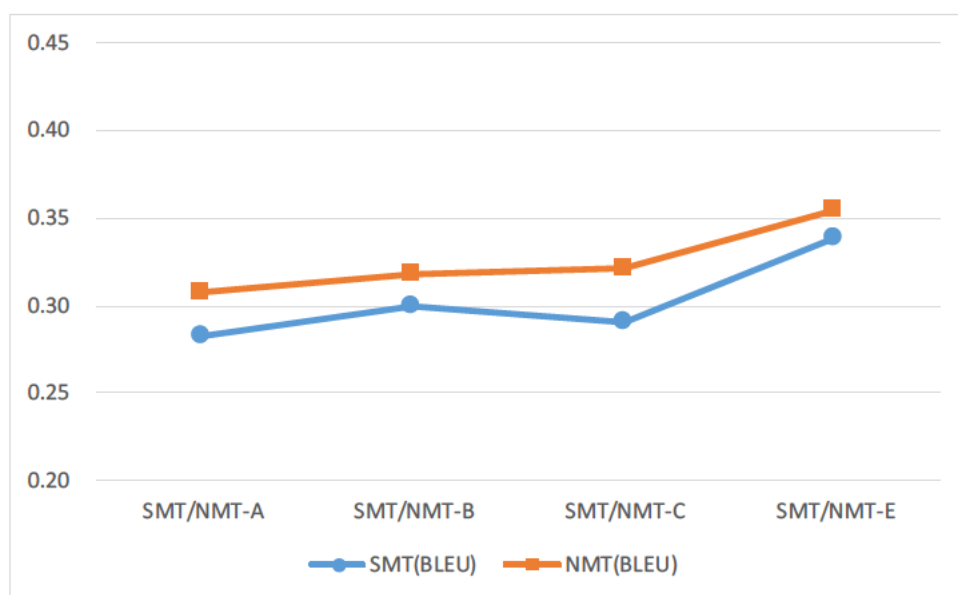


図 3.1-5 自動評価結果 (BLEU)

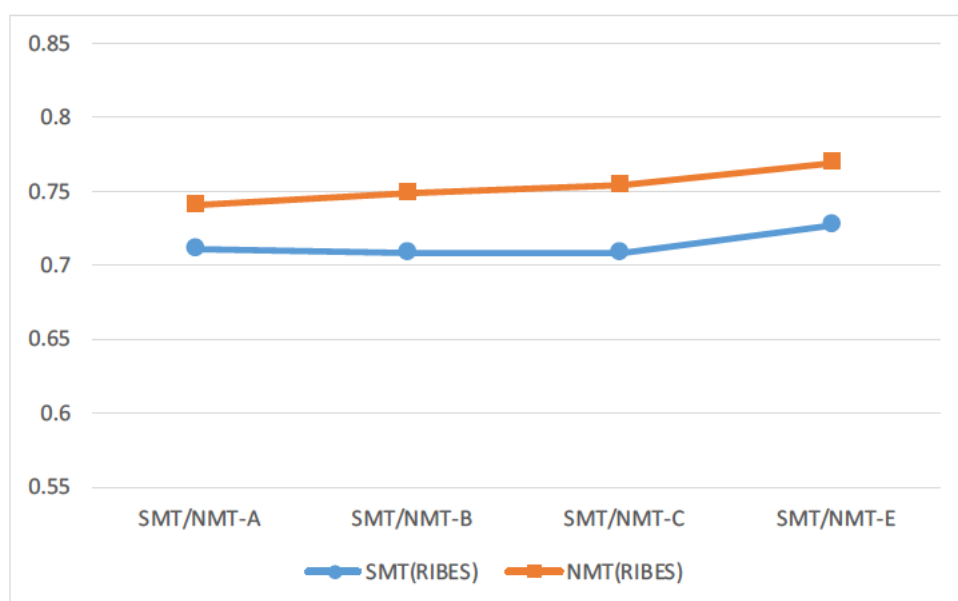


図 3.1-6 自動評価結果 (RIBES)

## (2) 人手評価と自動評価の関係 (SMT)

統計機械翻訳の人手評価と自動評価とを比較してみると、全体的には概ね学習コーパスの増加に従って上昇傾向となる点は同じであるが、人手評価は学習コーパス B 追加時の伸びが大きかったのに対し、BLEU では、学習コーパス E 追加時の伸びが大きく、また、RIBES ではいずれも大きな変化は見られない等、高い相関は見られなかった。

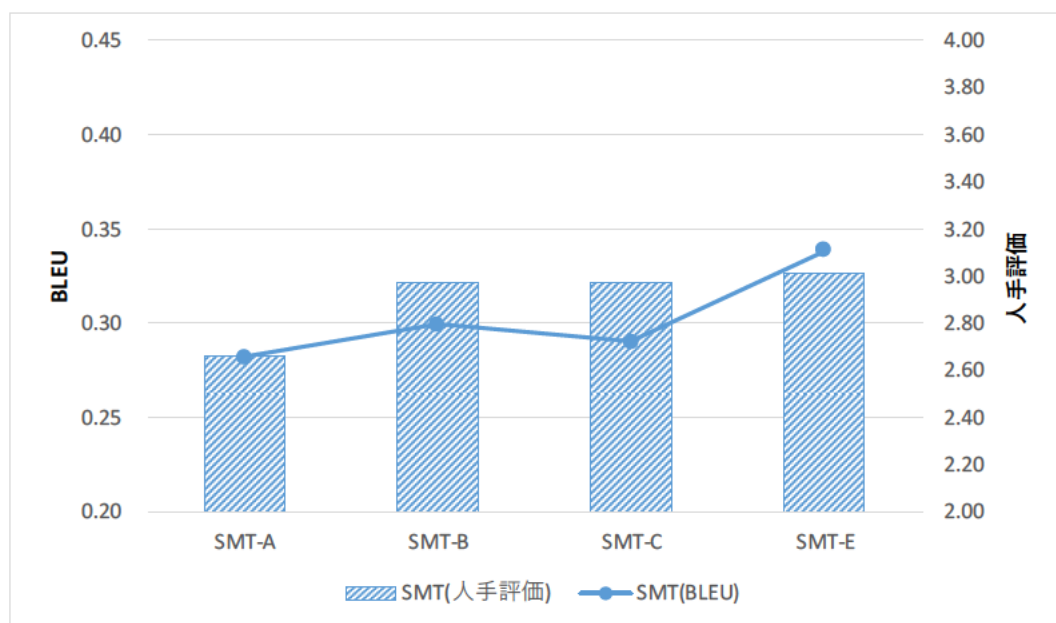


図 3.1-7 人手評価および自動評価結果 (BLEU)

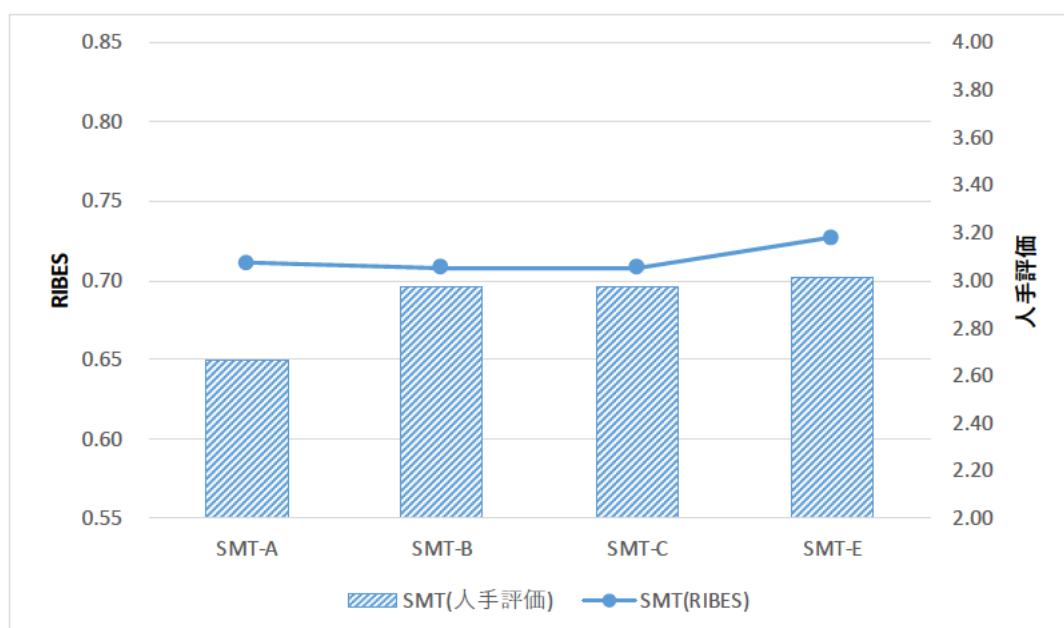


図 3.1-8 人手評価および自動評価結果 (RIBES)



## 3.2 機械翻訳文の分析（統計機械翻訳）

### 3.2.1 機械翻訳の誤りの傾向

人手評価文について評価者からのコメントを元に統計機械翻訳 A から E に出現した誤りを分析した。その結果、頻出の誤りとして、文中の各用語や句の間の関係（係り受け）の誤り、原文にある内容が訳文から欠落しているもの（訳抜け）、逆に原文に存在しない内容が訳文に存在しているもの（湧き出し）、そして訳語の誤りが多く見られた。機械翻訳結果を分析すると 1 文中にこれらの誤りが複合的に存在した。誤りの割合は係り受け誤りが人手評価文の半分以上に含まれており最も多く、湧き出しと訳抜けはそれぞれ評価文の 2 割程度の量を占めた<sup>13</sup>。以下に各誤りの例を誤りの原因分析<sup>14</sup>とともに示す。

#### (1) 係り受け誤り

係り受け誤りは評価文の半数以上に存在し、最も多く見られた。また、係り受け誤りを含むものは文意が通りにくくなり、人手評価も低い傾向が見られた。

拒絶理由通知に限らず特許文書では、構成要素や引用箇所を列挙して表現することが多く、こうした並列表現の解釈範囲の誤りが多く見られた。

下記例は、「参照」が「段落【0047】」と「図面等」に係らなければならないところ、機械翻訳結果では「refer to the drawings and the like」となっており、「参照」が「図面」だけに係った訳となっている。

また、「ランプヒータ～備えた液体加熱ユニット」の部分についても本来「ランプヒータ、流体通流部、及び反射板」が並列となり動詞「備えた」の目的語となるべきだが、機械翻訳文では動詞「備えた / provided」の係り先が「反射板 / reflection plate」だけになる誤りが見られた。

例 1.

また、引用文献 2（ <u>段落【0047】、図面等参照</u> ）には、 <u>ランプヒータ、流体通流部、及び反射板を備えた液体加熱ユニット</u> 、の発明が記載されている。	In addition, document 2 discloses a (, <u>paragraph [0047], refer to the drawings and the like</u> ), <u>a lamp heater fluid flow section, and a liquid heating unit provided with a reflection plate</u> of the present invention.
---	---

上記例は、比較的狭い範囲内の並列関係の誤りであるが、長文になるに従い並列関係と

<sup>13</sup> 誤りは 1 文中に複数かつ複合的に存在するため、厳密に誤り個数を集計することはできなかった。ここでは 1 文に同種の誤りが複数存在しても 1 つとして集計したため、分量はおおよその目安程度の値となる。

<sup>14</sup> 係り受け誤りの原因は機械翻訳システム内の構文解析によるものと推測されるが、本調査で調査対象とした「みんなの自動翻訳」では、構文解析結果の詳細な情報を得ることができないため誤りの例示のみとした。

なる句の組み合わせが複数存在し、並列関係の解釈範囲が広がるケースが多くなり、これらが係り受け誤りとなることによって翻訳文の内容が全く理解できなくなる。

下記例は、本願と引用文献との間の対比文である。文の構造は本願と引用それぞれに並列句があり、その両者を対比したものである。

この文は2つの係り受け誤りが存在した。1つ目が引用文献1の「扉体 15L」と「ヒンジ軸 18」が並列に扱われていない点、2つ目が本願請求項の「第1の回転プレート」、「第1の軸」が並列に扱われていないために、「第1の回転プレート / 1 rotation plate<sup>15</sup>」と「第1の軸 / 1 shaft」の間に、副詞「それぞれ / respectively」が割り込んでしまった点であり、これらによって意味を理解できない翻訳結果になってしまった。

例 2.

<p>~ 引用文献 1 に記載された発明の「扉体 15L」、「ヒンジ軸 18」は、それぞれ、請求項 1 に係る発明の「第 1 の回転プレート」、「第 1 の軸」に相当する。</p>	<p>~ "door body 15 l" disclosed in document 1 in the invention "hinge shaft 18", "1 rotation plate" of respectively in the invention as in claim 1, corresponds to the "1 shaft".</p>
--	---

## (2) 訳抜け

訳抜けのうち、「引用文献 / cited document」の「cited」が訳出されないものが多く見られた。

例 3.

<p>引用文献 3 には、樹脂成形部材の材料として、~</p>	<p>Document 3 discloses, as a material for a resin molded member, ~</p>
---------------------------------	---

これは、学習コーパスにおける「引用文献」に対する訳が「document」となっているものが多いためと考えられる。学習コーパス A のうち「引用文献」を含む文がおよそ 3,000 文存在し、そのうち 2,700 弱の文が下記に示すような「cited」を含まない文であった。他の見解書コーパスからの学習コーパスにおいても同様の傾向であった。

<sup>15</sup> 「第 1」が「first」でなく「1」と訳される点については、(4)「訳語誤り」の節を参照の事。

学習コーパス	
上記請求の範囲に係る発明は引用文献 1 により進歩性がない。	The invention as in the abovementioned claims does not involve an inventive step in view of <u>document</u> 1.
上記請求の範囲に係る発明は引用文献 1 により新規性がない。	The invention as in the abovementioned claims lacks novelty in view of <u>document</u> 1.
引用文献 1 には、上記( 1 )で指摘した点が記載されている。	<u>Document</u> 1 describes the features pointed out under (1) above.

また、拒絶理由通知では頻出する「参照」、「図」についても訳抜けが多く見られた。

例 4.

引用文献 1：特に、段落【 0 0 4 7 】、第 5 図参照。	Document 1: in particular, paragraph [0047], figures 5 and.
----------------------------------	---

これらについても、学習コーパスでは必ずしも「参照」が訳出されていないことが原因と予想される。下記の 1 番目と 2 番目は英語のスタイルとしてあえて「参照」を訳出していないケース、最後の例の日本語は「～。(【 0 4 6 4 】参照)」となっていたものがコーパス作成過程において句点「。」で文分割されたために、英文との対応関係が取れない誤った文分割のケースと考えられる。

学習コーパス	
そして、文献 1 と同様なイオン対は公知( 例えば、 <a href="#">文献 2 - 5</a> , 等参照 ) であることから、該公知のイオン対を採用してみることも、当業者であれば適宜選択し得る事項である。	Therefore, because using an ion pair similar to that of document 1 is well known (for example, <a href="#">documents 2-5, etc.</a> ), using the well-known ion pair is also a matter a person skilled in the art
光ピックアップの技術において、オフセットの許容値が波長と開口数に関係することは、文献 4 に記載されているように、当業者に広く知られている技術である ( <a href="#">段落 0 0 5 7 を参照</a> ) 。	Among the features of optical pickup devices, the feature wherein the tolerance value for an offset is related to the wavelength and the numerical aperture is widely known to a person skilled in the art, as described in document 4 ( <a href="#">paragraph [0057]</a> ).
( <a href="#">【 0 4 6 4 】参照</a> ) 文献 1 , 4 に記載されたそれぞれの発明の低屈折率層を、文献 3 に記載された発明の塗布組成物を用いて形成することは、当業者にとって容易である。	It would be easy for a person skilled in the art to use the coating composition in the invention described in document 3 to form the low refractive index layer in each of the inventions described in documents 1 and 4.

その他、頻出する「発明」の訳抜けも見られた。

例 5.

引用文献 1 には、＜理由 1＞で認定した通りの <a href="#">発明</a> が記載されている。	Document 1 discloses, a set in <Reason 1>.
---	--

学習コーパスから類似表現の文を調査すると、「発明」が英語では省略されている文が存在した。このような文を学習したことで訳抜けが発生したと考えられる。

学習コーパス	
文献 3 には、立体表示装置の <a href="#">発明</a> が記載されている。	Document 3 discloses a stereoscopic display device.
文献 1 には、積層型半導体装置の <a href="#">発明</a> が記載されている。	Document 1 discloses a laminated-type semiconductor device.
さらに、文献 3 - 5 にも、タイヤの形状データのデータ処理装置の <a href="#">発明</a> が記載されている。	Moreover, documents 3-5 also describe data processing devices for tire shape data.

### (3) 湧き出し

拒絶理由通知では、「当業者」の訳が重複して出力される湧き出しが多く見られた。

例 6.

引用発明 1 と引用発明 2 は、共に複数のキャラクターが対戦するゲームの表示制御に関するものであり、引用発明 1 によるゲームの進行状況が所定の条件を充たした場合に、引用発明 2 の「行動指令（コマンド）」の制御方法を採用することは、 <u>当業者</u> に容易である。	Invention 1 and invention 2 relate to display control of a game in which a plurality of characters play a match, and when the progress of the game by the cited invention 1 充たし prescribed conditions, <u>a person skilled in the art</u> to apply the control method of "action command (command)" of cited invention 2 would be easy for <u>a person skilled in the art</u> .
引用文献 1 に記載された発明のレンズアクセサリーに、上記引用文献 2 に記載された技術を適用することは、 <u>当業者</u> が容易になり得ることである。	<u>A person skilled in</u> the lens accessory of the invention disclosed in document 1, applying the feature disclosed in document 2, <u>a person skilled in the art</u> could easily achieve.
また、引用例 1 記載の発明は、インターバル時間 $T_i$ 経過後コンデンサの充電時間によって第 2 時間を作るものであるが、引用例 1 の【0025】 - 【0029】を参照すれば、コンデンサの充電時間をタイマーで計測することは、 <u>当業者</u> であれば容易に想到できる。	In addition, the invention disclosed in document 1, <u>a person skilled in the art</u> to form the 2 time by charging time interval time $T_i$ after the elapse of a capacitor, in view of the disclosures in paragraphs [0025] - [0029] of document 1, the charging time of a capacitor and measuring the timer can be easily conceived of by <u>a person skilled in the art</u> .
大電力に用いられるスイッチング素子に関する技術分野に属する引用文献 1 ないし 4 に記載された発明を、三相交流回路において用いることは、 <u>当業者</u> が容易になり得た範囲内のことである。	The inventions disclosed in documents 1 to 4 belong to the technical field of switching devices used in high power switching, <u>a person skilled in the art</u> could use a 3 phase alternating circuit, is within the range of what <u>a person skilled in the art</u> could easily.

この誤りの原因を調べるため、まず原文のどの用語がどの英語に翻訳されているかを、「みんなの自動翻訳」で確認した。その結果、下記図の網掛けで示すとおり、原文「こと」の翻訳結果が「a person skilled in the」に対応していることが分かった。

引用発明 1 と引用発明 2 は、共に 複数の キャラクタ が対戦 する ゲームの 表示制御 に関するものであり、引用発明 1 による ゲームの進行状況 が所定の条件 を満たし た 場合に、引用発明 2 の 「 行動 指令 ( コマンド ) 」 の 制御方法を採用する ことは、当業者に 容易である。
Invention 1 and invention 2 relate to display control of a game in which a plurality of characters play a match, and when the progress of the game by the cited invention 1 満たし prescribed conditions, a person skilled in the art to apply the control method of "action command (command)" of cited invention 2 would be easy for a person skilled in the art.

この結果を参考に、湧き出しの原因となった学習コーパスを調べた。その結果、下記に示すように、「当業者」が存在しないにも関わらず、「a person skilled in the art」を含む文が複数見つかった。これらの文は英語に翻訳する際に主語「a person skilled in the art」を補ったものと考えられ、このような学習コーパスを学習した結果が湧き出しの原因と考えられる。

学習コーパス	
分断層の膜厚の範囲を規定する <u>こと</u> は、適宜なし得る事項である。	Specifying the range of the film thickness of the segmentation layer is a matter that <u>a person skilled in the art</u> could achieve, as appropriate.
なお、所定の時間を所定の距離に置き換える <u>こと</u> は適宜なし得る事項である。	In addition, <u>a person skilled in the art</u> could substitute the predetermined period of time with a predetermined distance, as appropriate.
請求項 2 1 に対して、検出部を設ける <u>こと</u> は、適宜なし得ることに過ぎない。	With regard to claim 21, <u>a person skilled in the art</u> could provide a detection section as appropriate.

また、「同文献」の訳が「document 1」となり、原文にない「1」の湧き出しが見られた。

例 7.

また、 <u>同文献</u> の図 5 及び関連説明箇所には、加速度センサで持ち方を判定することが記載されている。	Further, figure 5 and related explanatory sections of <u>document 1</u> that it has the determined using an acceleration sensor.
---	--

「同文献」を含む学習コーパスを調べると、下記に示すように「同文献」の「同」に対応する英語は文脈によりバリエーションがあることが分かった。その結果、最も出現頻度の多かった英語「document 1」が訳語となったと考えられる。



学習コーパス	
また、 <u>同文献</u> の図示されたガスタービンにおける寸法関係との差異もないといえる。	In addition, there are no differences between the dimensional relationships in the gas turbines illustrated in <u>document 1</u> .
また、 <u>同文献</u> の【0022】には、メッキに用いる金属としてニッケルが挙げられている。	In addition, <u>document 2</u> (paragraph [0022]) presents nickel as a metal to be used for plating.
文献 1 には、 <u>同文献</u> の一般式 ( 1 ) から誘導される基を構造単位として有する高分子化合物については記載されていない。	Document 1 does not disclose a polymer compound that has a group derived from general formula (1) of <u>the same document</u> as a structural unit.

その他、冠詞「a」や記号「)」等の湧き出しが多く見られたが、湧き出しの原因となった学習コーパスを特定することはできなかった。

例 8.

引用文献 1 にはアンカー部材が突設された立体ブロック ( 1 ) の内部にコンクリートを打設してコンクリート架台 ( 2 ) に固定する技術手段は記載されていない。	The technical means for securing the concrete frame (2) <u>a</u> , and pouring concrete into the interior of the (1) 3 d block in which an anchor member protruding is not disclosed in document 1.
本願の請求項 1 - 3 に係る発明と引用文献 1 ( 特に、第 7 図を参照。) に記載された発明とは、次の点で相違しその他の点は一致する。	The invention as in claims 1-3 of the present application and the invention disclosed in cited document 1 (refer in particular to figure 7 <u>).</u> ) and differ in the following other features being coincident.

#### (4) 訳語誤り

その他に見られた誤りを以下に示す。

「第 1 」「第 2 」の訳が序数「first」「second」と訳されずに数字の「1」「2」に訳された。

例 9.

～ <u>第 1</u> の切換部材 3 1 と <u>第 2</u> の切換部材 3 2 とを～
～ the <u>1</u> switching material 31, a <u>2</u> switching member 32 ～

例 10.

～ コンデンサの充電時間によって第2時間を作るもの～
～ the 2 time by charging time ～

「第2」を含む学習コーパスを見ると下記例に示すような文の「第2の実施形態/embodiment 2」、「第2頁 / page 2」の部分から「第2」、「2」の対応を学習したことが原因と考えられる。

学習コーパス	
文献1における、第2の実施形態と第3の実施形態とを単に組み合わせることは、当業者に自明である。	Simply combining embodiment 2 and embodiment 3 in document 1 would be obvious to a person skilled in the art.
文献5の第2頁右上欄第12行 - 左下欄第3行に記載された流量計が検出した流量による制御手法を適用することは当業者にとって容易である。	It would be easy for a person skilled in the art to use the control technique using the flow rate detected by the flowmeter disclosed in document 5 (page 2, upper right column, line 12 to lower left column, line 3).

### 3.2.2 学習コーパスの登録による改善点や改悪点

3.1 では人手評価、自動評価ともに学習コーパスの増加により評価値の向上が確認できた。しかし、評価値は1文全体の評価であるため、3.2.1 で分析した各誤りの傾向が、学習コーパス登録の増加に従ってどのように変化しているかについて分析した。

#### (1) 係り受け誤り

下記例の統計機械翻訳 A では、「段落【0047】、図面等参照」の翻訳が「, paragraph [0047], refer to the drawings and the like」となり「参照 / refer」の係り受け誤りがあったが、学習コーパスを増やした統計機械翻訳 B 以降は「, paragraph [0047], and drawings」と動詞「refer」が訳出されなくなることで係り受け誤りが解消した<sup>16</sup>。

また「ランプヒータ・・・」部分の係り受け誤りについては各構成要素の冠詞の有無、定冠詞/不定冠詞の違いや訳語の変化「流体通流部 / fluid flow section / fluid conduction part」はあるものの、訳語の並びに変化はないため、係り受け誤りの解消は見られなかった。

<sup>16</sup> 日本語と英語との逐語訳の観点からは「refer」が訳出されないのは訳抜けとなるが、ここでは人間が読んだ場合に特に問題はないので係り受け誤りが解消したとした。



例 11.

原文	また、引用文献 2 ( <u>段落【0047】、図面等参照</u> ) には、 <u>ランプヒータ、流体通流部、及び反射板を備えた液体加熱ユニット、の発明</u> が記載されている。
統計機械翻訳 A	In addition, document 2 discloses a ( <u>, paragraph [0047], refer to the drawings and the like</u> ), <u>a lamp heater fluid flow section, and a liquid heating unit provided with a reflection plate of the present invention.</u>
統計機械翻訳 B	In addition, cited document 2 discloses ( <u>, paragraph [0047], and drawings</u> ), <u>a lamp heater, the fluid flow section, and liquid heating unit provided with a reflecting plate of the present invention.</u>
統計機械翻訳 C	In addition, cited document 2 discloses ( <u>, paragraph [0047], and drawings</u> ), <u>lamp heater, the fluid flow section, and a heating unit provided with a reflecting board of the invention.</u>
統計機械翻訳 E	In addition, cited document 2 describes ( <u>, paragraph [0047], and drawings</u> ), <u>a lamp heater, a fluid conduction part, and a liquid heating unit comprising a reflection plate of the invention.</u>

## (2) 訳抜け

3.2.1 で分析した「引用文献/cited document」の「cited」の訳抜けについて、学習コーパスの増加による機械翻訳結果の変化を調査した。下記例の統計機械翻訳 A から C では、訳抜けの改善は見られなかったが、統計機械翻訳 E で「cited document 3」となり訳抜けが改善された。統計機械翻訳 A から C の学習コーパスは特許庁から貸与された「見解書コーパス」から学習した機械翻訳であるが、統計機械翻訳 E の学習コーパスは本事業で作成した日英対訳コーパスを追加し学習したものである（2.2.1 を参照のこと。）。このコーパスでは「引用文献 / cited document」と翻訳されているため訳抜けが改善したと考えられる。

例 12.

原文	<a href="#">引用文献 3</a> には、樹脂成形部材の材料として、熱硬化性樹脂や熱可塑性樹脂や炭素繊維強化合成樹脂を用いた複合成型材料を使用する技術が記載されている。
統計機械翻訳 A	<a href="#">Document 3</a> discloses, as a material for a resin molded member, and a feature wherein a composite molding material which uses a thermosetting resin, a thermoplastic resin or a carbon fiber reinforced synthetic resin.
統計機械翻訳 B	<a href="#">Document 3</a> describes, as a material for a resin molded member, the feature of using a composite molding material which uses a thermosetting resin or a thermoplastic resin or a carbon fiber reinforced synthetic resin is used.
統計機械翻訳 C	<a href="#">Document 3</a> discloses, as a material for a resin molded member, and a feature wherein a composite molding material which uses a thermosetting resin or a thermoplastic resin or a carbon fiber reinforced synthetic resin.
統計機械翻訳 E	The <a href="#">cited document 3</a> describes, as a material for a resin molded member, a technique of using composite molding material which uses a thermosetting resin or a thermoplastic resin or a carbon fiber reinforced synthetic resin.

### (3) 湧き出し

3.2.1 で分析した「a person skilled in the art」の湧き出しについて学習コーパスの増加による変化を調査したところ、下記例では、統計機械翻訳 A に見られた湧き出しが学習コーパスを追加した統計機械翻訳 B では改善された。さらに、学習コーパスを追加した統計機械翻訳 C においても湧き出しは見られなかったが、学習コーパス全てを学習した統計機械翻訳 E では湧き出しが 2 か所見られたことにより最も評価が低くなった。

例 13.

原文	引用文献 1 に記載された発明のレンズアクセサリーに、上記引用文献 2 に記載された技術を適用することは、 <u>当業者</u> が容易になし得ることである。
統計機械翻訳 A	<u>A person skilled in</u> the lens accessory of the invention disclosed in document 1, applying the feature disclosed in document 2, <u>a person skilled in</u> the art could easily achieve.
統計機械翻訳 B	To the lens accessory of the invention described in document 1, applying the technique disclosed in the abovementioned document 2, <u>a person skilled in</u> the art could easily achieve.
統計機械翻訳 C	The lens accessory of the invention described in document 1, applying the feature disclosed in document 2, <u>a person skilled in the art</u> could easily achieve.
統計機械翻訳 E	<u>A person skilled in the art</u> to the lens accessory of the invention described in document 1, <u>a person skilled in the art</u> could easily conceive of applying the technique described in the cited document 2, <u>a person skilled in the art</u> could easily achieve.

統計機械翻訳 E で再び湧き出しが発生した原因を調査したところ、湧き出しと原文との対応は下記図に示すように原文の「に」や句点「、」と対応していた。この点を参考に統計機械翻訳 E で追加した学習コーパスを調べると、「当業者」を含まず「a person skilled in the art」を含む文が存在した。これらの文が湧き出しの原因と考えられるが、該当する文数は 26 文と少なく、明確な湧き出しのプロセスは明確とならなかった。

翻訳
引用文献 1 に記載された発明のレンズアクセサリーに、上記引用文献 2 に記載された技術を適用することは、当業者が容易になし得ることである。
<u>A person skilled in the art</u> to the lens accessory of the invention described in document 1, <u>a person skilled in the art</u> could easily conceive of applying the technique described in the cited document 2, <u>a person skilled in the art</u> could easily achieve.

翻訳
引用文献 1 に記載された 発明 のレンズアクセサリー に、上記引用文献 2 に記載された技術を適用すること は、当業者が容易になし得る ことである。
A person skilled in the art to the lens accessory of the invention described in document 1, a person skilled in the art could easily conceive of applying the technique described in the cited document 2, a person skilled in the art could easily achieve.

学習コーパス	
ここで、どのタイミングで印刷キュー監視を始めるか／終わらせるかは、必要に応じて適宜選択し得る事項である。	When to start / stop monitoring of the print queue is the matter that <u>a person skilled in the art</u> could have appropriately selected in accordance with need.
さらに、撮像された画像を二値化する際の閾値として、画像の輝度レベルの強度ヒストグラムに基づいて設定されたものが引用文献 4 に記載されており、文献 1、2 における閾値としてこのように設定されたものをを用いるよう想到することは、これら文献から容易である。	Cited Document 4 discloses that when the obtained image is binarized the threshold is set based on an intensity histogram of the brightness of the image, and it is easily possible for <u>a person skilled in the art</u> to use the threshold in Cited Documents 1 and 2 based on the Cited Document 4.
しかしながら、引用文献 2 の特許請求の範囲、引用文献 3 の特許請求の範囲、引用文献 4 の特許請求の範囲、第 3 頁第 4 行～第 5 行に記載されているように、ピッチ除去剤をフローテーションにより脱墨を行う工程及び／又はその前工程に添加することは、当該技術分野において周知の技術に過ぎない。	However, it is a well-known technique for <u>a person skilled in the art</u> to perform a deinking step by means of flotation using the pitch remover and/or an addition of the pitch remover prior to the deinking step, as described in claims in Cited Document 2, claims in Cited Document 3, and claims, lines 4-5 on page 3 in Cited Document 4.

3.2.1 で分析した「同文献 / document 1」の湧き出しについても同様に調査したところ、湧き出しの改善は見られなかった。

例 14.

原文	また、 <u>同文献</u> の図 5 及び関連説明箇所には、加速度センサで持ち方を判定することが記載されている。
統計機械翻訳 A	Further, figure 5 and related explanatory sections of <u>document 1</u> that it has the determined using an acceleration sensor.
統計機械翻訳 B	Further, figure 5 and the explanations related thereto in <u>document 1</u> , and has the for determination of an acceleration sensor.
統計機械翻訳 C	Further, figure 5 and the explanations related thereto in <u>document 1</u> that it has the determination of an acceleration sensor.
統計機械翻訳 E	In addition, figure 5 and the explanations related thereto in <u>document 1</u> has a determination with an acceleration sensor.

#### (4) 訳語誤り

その他の誤りとして分析した、序数「第 1」が first とならない誤りについても同様に各機械翻訳を調査したが、改善は見られなかった。

例 15.

原文	～ <u>第 1</u> の切換部材 3 1 と <u>第 2</u> の切換部材 3 2 とを ～
統計機械翻訳 A	～ the <u>1</u> switching material 31, a <u>2</u> switching member 32 ～
統計機械翻訳 B	～ a <u>1</u> switching member (31) and a <u>2</u> switching member 32 ～
統計機械翻訳 C	～ a <u>1</u> switching member (31) and a <u>2</u> switching member 32 ～
統計機械翻訳 E	～ a <u>1</u> switching member 31 and the <u>2</u> switching member 32 ～

#### (5) 未知語

その他学習コーパスの追加で効果が期待される未知語について調査を行った。今回機械翻訳文中に原文の表記がそのまま出力された未知語が 7 語存在した。これら未知語の変化を見ると、「虚偽」は学習コーパス B の追加により「nuisance」と訳されるようになり、学習コーパス D, E を追加したことにより「虚偽申告/ false declaration」と適訳となり訳語の改善が見られた。「敷く」は学習コーパス B の追加により「placing」と訳されるようになったが、学習コーパス D, E の追加により「the placing」と名詞句として訳されるようになってしまった。

表 3.2-1 未知語一覧

統計機械翻訳 A	統計機械翻訳 B	統計機械翻訳 C	統計機械翻訳 E	備考
索敵	索敵	索敵	索敵	
プラマリハイパー	プラマリハイパー	プラマリハイパー	プラマリハイパー	原文の用語「プラマリハイパーバイザ」の一部が未知語 なお「プラマリ」は「 <u>ラ</u> マリ」の誤記
充たし	充たし	充たし	充たし	
虚偽 reporting	nuisance reporting	nuisance reporting	false declaration	nuisance は迷惑、妨害の意味。
踏角	踏角	踏角	踏角	
敷く	placing	placing	the placing	
足跡	足跡	足跡	足跡	

また、未知語ではないが、日本語のローマ字表記が訳語となったものが存在した。これは学習コーパスにある引用文献名の表記が学習されたことが原因と考えられる。学習コーパス自体に問題はないが、このような対訳の扱いは注意が必要となる。

例 16.

ハイスコア大会	high score <u>Taikai</u>
---------	--------------------------

学習コーパス	
文献 5 : 日本農芸化学会 <u>大会</u> 講演要旨集, 2011.03.05	Document 5: Japan Society for Bioscience, Biotechnology, and Agrochemistry <u>Taikai</u> Koen Yoshishu, 05 March 2011
文献 2 : 日本菌学会第 54 回 <u>大会</u> 講演要旨集, 2010, p.31, 2SA2	Document 2: The Mycological Society of Japan Dai 54 Kai <u>Taikai</u> Koen Yoshishu, 2010, pp. 31, 2SA2
文献 4 : 日本臨床検査自動化学会第 39 会 <u>大会</u> 抄録集, 2007.09.01, p.434	Document 4: The Japan Society for Clinical Laboratory Automation Dai 39 Kai <u>Taikai</u> Shorokushu, 01 September 2007, page 434

### 3.2.3 同一の語の翻訳によるぶれの程度

人手評価文のうち任意の 50 文について同一語の翻訳によるぶれを調査した。調査は各評価文に対して統計機械翻訳 A-E の翻訳結果を比較し、訳が異なる部分を抽出した。但し、

冠詞の有無や異なり（a, an, the）、単数形/複数形の違いは不問とした。また、未知語については 3.2.2 (5)にて分析を行っているため対象外とした。

調査の結果、50 文中 32 文にぶれが存在した。用語のぶれを分析すると、表現は異なるものの、意味が近いものが多数を占めた。

例 17.

外周面	outer <u>peripheral</u> surface outer <u>circumferential</u> surface
揮発性記憶手段	volatile <u>memory</u> means volatile <u>storage</u> means
規則変更手段	<u>rule modification</u> means <u>regular change</u> means

「そして」、「また」など多くの文で頻出する接続表現にも訳語のバリエーションが見られた。また、文中で頻出する参照表現「前記」、「上記」についても単に冠詞「the」と訳されるものや「the abovementioned」と訳されるものが見られた。

例 18.

そして	In addition Furthermore
また	In addition Also
前記 上記	the the <u>abovementioned</u>

また、拒絶理由通知で頻出する用語については下記例の用語に訳のバリエーションが見られた。

例 19.

刊行物	Publication document
特定事項	<u>defining</u> matter <u>specifying</u> matter
従来周知	well known <u>conventionally</u> well known



同一語のぶれの多くは、比較的類似した概念の用語であったため内容理解に大きな影響を与えないと考えられるが、調査結果の中には文によって訳し分けが必要となるような語も存在した。

例 20.

パイロット式	pilot- <u>type</u> pilot <u>formula</u>
課題	problem goal
方向	orientation direction

その他、1 文中に出現する同一語の訳にぶれが生じているものも見られた。下記の例は「切替部材」が2回出現するが、異なる訳「switching material」、「switching member」となっている。

例 21.

<p>下記文献 1 には、マニホールド電磁弁（切換弁組立体）において、給気ブロック（給排気ブロック 4 a）の側方に第 1 の<u>切換部材</u> 3 1 と第 2 の<u>切換部材</u> 3 2 とを選択的に取付けることにより、内部パイロット式と外部パイロット式とに切換える構成が記載されている。</p>
<p>The following documents 1 manifold electromagnetic valve (switching valve assembly); wherein, the 1 <u>switching material</u> 31, a 2 <u>switching member</u> 32 selectively attached to the side of the air supply block (supply block (4 a)), switching from an internal-pilot-type external pilot-type).</p>

### 3.2.4 日英対訳コーパスの学習量に応じた機械翻訳精度の変化

3.1 では人手評価、自動評価ともに学習コーパスの増加により機械翻訳精度の向上が確認できた。ここでは人手評価対象文について精度が向上した文、逆に精度が下がった文の分析を行う。人手評価である内容伝達レベルの評価値は 1 から 5 までの 5 段階評価であるが、人手評価値が同値であっても精度差が生じていることもあることから、各評価文について 4 つの機械翻訳（統計機械翻訳 A, B, C および E）の翻訳精度を比較し順位を付け、その内容を元に学習コーパス量の増加により精度が上がったもの、逆に精度が下がったものを抽出した。



### (1) 精度が上がったもの

下記例は統計機械翻訳 E で精度が上がった例である。各機械翻訳の相対的な優劣は (A=B=C<<E)となる。人手評価では英語構造(語順)や文法的に必須な要素(主語、動詞...)を読み取れないと内容理解が難しくなる。そのため、下記例の統計機械翻訳 E のように訳文に動詞が出力されるだけで人手評価は大きく改善する。一方自動評価では、機械翻訳 C と E との差は僅かであり、このような違いを検出することは難しいと考えられる。

例 22.

原文	(相違点 2) 搬送手段が、上記請求項記載の発明は『回転体』であるのに対して、上記引用文献記載の発明は『搬送コンベア』である点。
統計機械翻訳 A	(Difference 2) conveying means"according to the invention as recited in the above claims"rotating body"of the inventions disclosed in the abovementioned documents"conveyor".
	英文は動詞が存在せず文として成立していない。引用符の付き方も不安定。
統計機械翻訳 B	(Difference 2) conveyance means"according to the invention as recited in the above claims"rotating body"of the inventions disclosed in the abovementioned documents"conveyor".
	A と実質的に同レベル。
統計機械翻訳 C	(Difference 2) conveyance means"according to the invention as recited in the above claims"rotating body"of the inventions disclosed in the abovementioned documents"conveyor".
	A,B と実質的に同レベル。
統計機械翻訳 E	(Difference 2) conveyance means"according to the invention as recited in the above claims <b>is</b> "rotary body", the inventions disclosed in the abovementioned documents"conveyor".
	動詞「is」が訳出された分、前半に関しては原文のような意味にもとれるようになり、後半も推定は可能となる。A-C より改善した。

下記例は相対評価が（ $B < A < C = E$ ）となる例である。この例では統計機械翻訳 B で動詞「referring to」が訳出される一方で、訳語誤りが発生しているため、文全体としては翻訳精度が一旦下がり、統計機械翻訳 C で訳語誤りが改善され精度が再び上がった例となる。

例 23.

原文	特表平 1 1 - 5 0 9 8 8 3 号公報の表 1 に対する注記等を参照するに、 C a b - o - s i l T S - 7 2 0 はコロイダルシリカと認められる。
統計機械翻訳 A	Notes, etc. to table 1 of <a href="#">JP flat [11] - [509883]</a> , [Cab (silTS 720 is considered to colloidal silica.
	文献番号や Cab-o-silTS-720 が不正確だが、前半、後半とも構造的にはそれなり。但し、前半と後半との関連性は全く分からない。
統計機械翻訳 B	<a href="#">Referring to</a> notes, etc. to table 1 of <a href="#">JP flat 11 A 509883</a> Publication No., Cab (o- <a href="#">aminophenol</a> silTS 720 is considered to colloidal silica.
	冒頭の <a href="#">referring to</a> で前後半の関係性がはっきり分かるようになった。ただその反面、Cab-o-silTS-720 はさらに崩壊し、 <a href="#">aminophenol</a> という無関係の訳語に変換されてしまった。誤解を招くという意味で、トータルでは A よりも劣る。
統計機械翻訳 C	Referring to the note to table 1 of <a href="#">Japanese 11 509883</a> paragraph, <a href="#">TS-720</a> sil <a href="#">TS 720</a> is recognized to be colloidal silica.
	B と比べると、aminophenol は消え印象は良くなったものの、Cab-o がなくなり <a href="#">TS-720</a> が繰り返されるなど、不正確な点では変わりなし。
統計機械翻訳 E	Referring to notes, etc. to table 1 of <a href="#">International Patent 11 - 509883 A</a> , TS-720 silTS 720 is considered to be colloidal silica.
	C と同等。番号部分も、一見改善しているように見えて、JP が無かったり H が無かったりするので結局不完全。

## (2) 精度が下がったもの

下記例は学習コーパスを増やしたにも関わらず翻訳精度が下がったものである。各統計機械翻訳間の相対評価は（ $E < A < C < B$ ）の順となり、学習コーパス量が最も大きな統計機械翻訳 E が最下位となった。この文は、本願と引用文献との対比文であるため、文中の各用語の訳の正確性は勿論であるが、比較対象の関係性が明確に読み取れる文に訳されるか否かがポイントとなる。そのため、統計機械翻訳 B のように文の構造を示す「whereas」などの機能語の有無が評価に大きな差をもたらすと考えられる。

例 24.

原文	(相違点 1) 外観検査対象が、上記請求項記載の発明は『対向する側面 A、B 及び対向する側面 C、D を備えた平面視して矩形あるいは正方形の電子部品』であるのに対して、上記引用文献記載の発明は『物品』である点。
統計機械翻訳 A	<p>(Difference 1) appearance inspection object according to the invention set forth in the claims states "an electronic component of a rectangular or square in a planar view provided facing side surfaces (A) and (B) and facing side surfaces (C), (D)", and therefore it is found that the inventions disclosed in the abovementioned documents "article".</p> <p>「～であるのに対して…」という対比の文であるが、「～で、従って…」と訳されている。「対向する側面」は「電子部品」に係るが、「平面視」(名詞として認識)に係っている。「発明は『物品』」の箇所もうまく訳されていない。原文後半の技術内容部分も非常に複雑で MT にはハードルが高いが、何を言っているのか全く分からない。</p>
統計機械翻訳 B	<p>(Difference 1) appearance inspection object, <u>whereas</u> the invention set forth in the abovementioned claims is "an electronic component of a rectangular or square in a planar view <u>and</u> provided <u>with</u> a facing side surfaces A and B and facing side surfaces (C), (D)" of the inventions disclosed in the abovementioned documents "article".</p> <p>位置はともかく <u>whereas</u> が訳出された分、文の主旨は A よりは推測しやすい。後半も、<u>and</u> や <u>with</u> が入ることで、文法的に多少改善した。</p>
統計機械翻訳 C	<p>(Difference 1) appearance inspection object"according to the invention as recited in the above claims"the electronic component of a rectangular or square in a planar view and provided with a facing side surfaces (A, B) and the facing side surfaces, C and D"of the inventions disclosed in the abovementioned documents"article".</p> <p>前半は A に戻ってしまった。後半は B のままなので、A&lt;C&lt;B。</p>
統計機械翻訳 E	<p>(Difference 1) appearance inspection object, the invention as set forth in the claims states "an electronic component of a rectangular or square shape in a plan view and that is provided with a side surface A, B and opposing sides C, D" of the inventions disclosed in the abovementioned documents "article".</p> <p>前半は A と同等。後半は、shape はあったほうが正しいものの、「側面」を最初は side surface、2 回目は opposing sides としてしまったため、A と C、B と E が対向するという誤解を招きやすくなった。内容理解の観点からは大きな改悪となった。</p>

### 3.3 統計機械翻訳とニューラル機械翻訳の比較

#### 3.3.1 比較対象とした機械翻訳

本調査では統計機械翻訳に加え最近翻訳精度の向上が著しいニューラル機械翻訳の自動評価を行った。ここではニューラル機械翻訳の自動評価値が高いもの、低いものについてその翻訳結果を分析するとともに統計機械翻訳との比較を行った。なお、比較は自動評価値の平均値が一番高かった、統計機械翻訳 E とニューラル機械翻訳 E との間で行った。

#### 3.3.2 評価値の比較

人手評価対象の 100 文について各文の統計機械翻訳 E とニューラル機械翻訳 E の BLEU, RIBES とを比較した。その結果、BLEU では統計機械翻訳とニューラル機械翻訳がほぼ同等、RIBES では 6 割以上の文でニューラル機械翻訳の方が高い結果となった。

表 3.3-1 BLEU 値の比較結果

#	BLEU 値の比較結果	件数
1	統計機械翻訳 E > ニューラル機械翻訳 E	48
2	統計機械翻訳 E = ニューラル機械翻訳 E	1
3	統計機械翻訳 E < ニューラル機械翻訳 E	<b>51</b>
	合計	100

表 3.3-2 RIBES 値の比較結果

#	RIBES 値の比較結果	件数
1	統計機械翻訳 E > ニューラル機械翻訳 E	37
2	統計機械翻訳 E = ニューラル機械翻訳 E	0
3	統計機械翻訳 E < ニューラル機械翻訳 E	<b>63</b>
	合計	100

#### 3.3.3 個別例の分析

さらに統計機械翻訳とニューラル機械翻訳の特徴を把握するため、ニューラル機械翻訳の自動評価値の高かったもの、低かったものをそれぞれ数件ずつ選び比較・分析した。分析の結果、ニューラル機械翻訳の自動評価値が高いものを見ると比較的短文でかつ、文の内容も単純なものが多く見られ、カッコ記号など細かな点を除き、ほぼ統計機械翻訳と同様な翻訳結果となった。

一方、自動評価値の低いものを見ると訳抜けが多く見られた。特にニューラル機械翻訳の訳抜けは、統計機械翻訳のそれとは異なり、フレーズ単位での訳抜けが見られた。以下に比較結果を示す。

(1) ニューラル機械翻訳の BLEU が高いもの

例 25.

原文	(特に、引用文献 1 の段落【 0 0 7 5 】、引用文献 2 の段落【 0 0 3 6 】、【 0 0 3 7 】、引用文献 3 の段落【 0 0 5 1 】、引用文献 4 の段落【 0 0 2 6 】参照)
基準翻訳	(Refer to paragraph [0075] in the cited document 1, paragraphs [0036] and [0037] in the cited document 2, paragraph [0051] in the cited document 3, and paragraph [0026] in the cited document 4, in particular.)
統計機械翻訳 E (B: 0.6804) (R: 0.718236)	(, in particular, paragraph [0075] of Cited Document 1, and paragraph [0036] of document 2, paragraph [0037]), cited document 3 (paragraph [0051] and paragraph [0026] in Cited Document 4, see)
ニューラル機械 翻訳 E (B: 0.7948) (R: 0.899632)	(In particular, see paragraph [0075] in cited document 1, paragraphs [0036] and [0037] in cited document 2, paragraph [0051] in cited document 3, and paragraph [0026] in cited document 4.)
備考	統計機械翻訳はカッコの湧き出しがあるものの、ニューラル機械翻訳と大差なし。

例 26.

原文	引用文献 1 : 特に、段落【 0 0 4 7 】、第 5 図参照。
基準翻訳	Cited Document 1: Especially, see paragraph [0047] and Fig. 5.
統計機械翻訳 E (B:0.5761) ( R:0.896883)	Cited document 1: In particular, paragraph [0047], Fig. 5).
ニューラル機械 翻訳 E (B:0.7474) (R:0.575467)	In particular, see paragraph [0047] and fig. 5 of cited document 1.
備考	短文かつ文の内容も比較的単純であるため、統計機械翻訳、ニューラル機械翻訳とも高い評価値となった。

例 27.

原文	引用文献 1 には、＜理由 1＞で認定した通りの発明が記載されている。
基準翻訳	The cited document 1 describes the invention as identified in <Reason 1>.
統計機械翻訳 E (B: 0.268) (R: 0.889051)	In the cited document 1, as in the <Reason 1>.
ニューラル機械 翻訳 E (B: 0.7298) (R: 0.984733)	Cited document 1 describes the invention identified in < reason 1 >.
備考	統計機械翻訳は「発明」、「記載」が訳抜け。ニューラル機械翻訳はほぼ基準翻訳どおりの翻訳結果となった。

## (2) ニューラル機械翻訳の BLEU が低いもの

例 28.

原文	<u>特表平 1 1 - 5 0 9 8 8 3 号公報の表 1</u> に対する注記等を参照するに、C a b - o - s i l <u>T S</u> - 7 2 0 はコロイダルシリカと認められる。
基準翻訳	Cab-o-silTS-720 is recognized as colloidal silica by referring to the note to Table 1 of National Publication of International Patent Application H No.11-509883.
統計機械翻訳 E (B: 0.1853) (R: 0.454573)	Referring to notes, etc. to table 1 of International Patent 11 - 509883 A, TS-720 silTS 720 is considered to be colloidal silica.
ニューラル機械 翻訳 E (B: 0.053) (R: 0.27493)	In view of the note of table 1 in <u>table 1</u> , the Cab-o-silTS <u>TS</u> -720 is considered to be colloidal silica.
備考	ニューラル機械翻訳は「特表平 1 1 - 5 0 9 8 8 3 号公報」が訳抜け。「表 1」、「TS」table 1、TS が重複。

例 29.

原文	そして、 <u>工具とワークの接触による過負荷を抑制することは周知の課題であるから、引用文献 1 に記載のワークとドリルとの接触時に、引用文献 3 に記載されたとおり、両者が離れる方向に所定距離だけ退避することは、当業者が容易に想到し得ることである。</u>
基準翻訳	In addition, since inhibiting overload due to contact of a tool and a work piece is a well-known problem, those skilled in the art could have easily conceived that, when the work piece and the drill disclosed in Cited Document 1 are contacted, they retract in a direction away from each other by a predetermined distance as disclosed in Cited Document 3.
統計機械翻訳 E (B:0.3022) (R:0.702255)	In addition, controlling the excessive load due to the contact of a tool and a workpiece is a well-known goal, when the workpiece contacts described in the cited document 1 and the drill and, as described in the cited document 3 to retreat by a predetermined distance in a direction away, a person skilled in the art could easily conceive of.
ニューラル機械 翻訳 E (B:0.0785) (R:0.208339)	Therefore, it would be easy for <u>a person skilled in the art</u> to apply the feature disclosed in document 3 to the invention disclosed in document 1. <u>A person skilled in the art</u> could easily conceive of this feature.
備考	ニューラル機械翻訳は評価文の主要な部分「工具とワークの接触による過負荷を抑制すること」、「両者が離れる方向に所定距離だけ退避すること」が訳抜けとなった。また「当業者 / a person skilled in the art」の湧き出しも見られた。

例 30.

原文	下記文献 1 には、マニホールド電磁弁（切換弁組立体）において、給気ブロック（給排気ブロック 4 a）の側方に第 1 の切換部材 3 1 と第 2 の切換部材 3 2 とを選択的に取付けることにより、内部パイロット式と外部パイロット式とに切換える構成が記載されている。
基準翻訳	The following document 1 describes constitution for switching between an internal pilot system and an external pilot system by selectively mounting the first switching member 31 and the second switching member 32 on the side of the air supply block (air supply/discharge block 4a) in the manifold electromagnetic valve (switching valve assembly).
統計機械翻訳 E (B: 0.392) (R: 0.579981)	Document 1 discloses a manifold electromagnetic valve (switching valve assembly) by a 1 switching member 31 and the 2 switching member 32 are selectively mounted on the side of the air supply block (supply / discharge block 4 a), switching to an internal-pilot-type external pilot-type).
ニューラル機械 翻訳 E (B: 0.0853) (R: 0.565009)	<u>The following documents are cited in the ISR.</u> Document 1 discloses a configuration of switching between an internal pilot type and an external pilot type.
備考	ニューラル機械翻訳の 1 文目は原文と関係のない文の湧き出しが見られた。2 文目も必要な情報のかなりの部分が訳抜けとなった。

### (3) ニューラル機械翻訳の RIBES が高いもの

ニューラル機械翻訳の RIBES が高いものを見ると、BLEU が高いものと同様に比較的短文かつ単純な文の評価が高くなった。



例 31.

原文	そして、高濃度に形成される領域は、本発明の「第 1 の半導体層」に相当する。
基準翻訳	Here, the region made to have high concentration of impurities <u>corresponds to the "first semiconductor layer" of the present invention.</u>
統計機械翻訳 E (B: 0.1427) (R: 0.831083)	In addition, the region that is formed in a high concentration corresponds to "a 1 semiconductor layer" in the invention of the present application.
ニューラル機械 翻訳 E (B: 0.5983) (R: 0.936479)	In addition, the region formed at a high concentration <u>corresponds to the "first semiconductor layer " of the present invention.</u>
備考	<p>人手評価は統計機械翻訳、ニューラル機械翻訳とも「高濃度に形成される領域」の訳以外は正しく訳出された。</p> <p>自動評価は RIBES 値が 0.8 (統計機械翻訳) 0.9 (ニューラル) とどちらも高い値となり人手評価と同様の結果となった。</p> <p>一方、BLEU 値は 0.1 (統計機械翻訳) 0.6 (ニューラル) と大きな差が見られた。これはニューラル機械翻訳の後半部分「corresponds to the " first semiconductor layer " of the present invention.」が基準翻訳と完全に一致しているために BLEU 値が高くなったと考えられる。</p>

例 32.

原文	本願請求項 1 - 1 5 に係る発明と、引用文献 1 に記載の発明とを対比すると、両者は以下の点において相違する。
基準翻訳	Comparison of the invention in Claims <u>1-15 of the present application and the invention described in</u> cited document 1, both inventions differ in the following aspects.
統計機械翻訳 E (B: 0.3852) (R: 0.833552)	The invention as in claim 1-15 of the present application is compared with the invention disclosed in document 1, the two inventions differ in the following points.
ニューラル機械 翻訳 E (B: 0.5951) (R: 0.92429)	Comparing the invention as in claims <u>1-15 of the present application and the invention described in</u> document 1, the two differ in the following points.
備考	人手評価は、統計機械翻訳、ニューラル機械翻訳ともにほぼ問題ない翻訳結果。自動評価は、RIBES 値が 0.8 (統計機械翻訳) \ 0.9 (ニューラル) とどちらも高い値となり人手評価と同様の結果となった。BLEU はニューラル機械翻訳の方が基準翻訳と一致する単語列が多かったことにより、BLEU 値に差が生じた。

#### (4) ニューラル機械翻訳の RIBES が低いもの

例 33.

原文	また、本願請求項 4、6 に係る発明も、同様にして当業者が容易に想到し得たものである。
基準翻訳	The inventions according to Claims 4 and 6 of the present application could have been also likewise easily conceived by <u>a person skilled in the art</u> .
統計機械翻訳 E (B:0.3015) (R:0.640273)	In addition, the invention as in claims 4 and 6, <u>a person skilled in the art</u> could easily conceive of the same.
ニューラル機械 翻訳 E (B:0.4562) (R:0.372071)	In addition, <u>a person skilled in the art</u> could also easily conceive of the invention as in claims 4 and 6 of the present application.
備考	統計機械翻訳の RIBES 値が 0.6 であるのに対しニューラル機械翻訳の RIBES 値は 0.3 と比較的大きな差が見られた。これは、ニューラル機械翻訳の「当業者が ~ / a person skilled in the art ~」の位置と基準翻訳の「a person skilled in the art」の位置が大きく異なるため、ニューラル機械翻訳の RIBES 値が低くなったと考えられる。なお、機械翻訳を人手で一对評価したところ、どちらもほぼ内容伝達がされており、RIBES の値ほど差はなかった。

## 4. 精度向上のための検討

### 4.1 係り受け誤りへの対応

#### 4.1.1 日本語構文解析の精度向上

3.2.1 では、機械翻訳の誤り傾向を分析し、係り受けに関する誤りが多くあることが分かった。また 3.2.2 では、学習コーパスを増加しても係り受けに関する誤りの改善が見られないケースがあることが分かった。特に拒絶理由通知は特許文の中でも構造が複雑であることから、原文（日本語）の構文解析精度の向上は重要な課題と考えられる。

#### 4.1.2 日本語の書き方の工夫

3.2.1 の分析では、係り受け誤りの例として並列関係にある要素が適切に扱われない例を示した。これは日本語解析が並列関係にある要素を適切に扱うことができなかったことが原因であるが、日本語原文の書き方を工夫することで改善する場合もある。

下記例上段の書き換え前の機械翻訳は「ランプヒータ、流体通流部、及び反射板」が並列に翻訳されない訳となっている。これを下段のように並列要素の係り受けが明確となるよう原文を書き換えることで正しい係り受けの訳となる。

例 34.

原文 (上段が書き換え前、下段が書き換え後)	機械翻訳文
また、引用文献 2（段落【0047】、図面等参照）には、 <u>ランプヒータ、流体通流部、及び反射板</u> を備えた液体加熱ユニット、の発明が記載されている。	In addition, document 2 discloses a (, paragraph [0047], refer to the drawings and the like), <u>a lamp heater fluid flow section, and a liquid heating unit provided with a reflection plate</u> of the present invention.
また、引用文献 2（段落【0047】、図面等参照）には、液体加熱ユニットは、ランプヒータ、流体通流部及び反射板を備える発明が記載されている。	In addition, cited document 2 discloses (, paragraph [0047], and drawings), liquid heating unit is provided with <u>a lamp heater, the fluid flow section and a reflective plate</u> .

### 4.2 学習コーパスの整備

本調査で調査対象とした統計機械翻訳、ニューラル機械翻訳はどちらも学習コーパスを学習することで翻訳を実現しているため、学習コーパスに翻訳に適さないデータが含まれていると翻訳精度に悪影響を及ぼすと考えられる。そのため、学習コーパスを整備することが翻訳精度向上のための課題となる。ここでは、3.2.1 の分析結果を参考にコーパスの整備方法案を検討した。

#### 4.2.1 文分割に誤りがある学習コーパスの除外

2.2.3 で学習コーパスを準備するにあたり、下記に示すような英語「fig.」で分割される等、英文の文分割が不適切なデータは除外した。

文献 1 を参照のこと（特に F i g . 4 ）。	Refer to document 1 (particularly fig.
特に請求項 5、6 について。 /// 周知技術である。	4). /// In particular, claims 5 and 6 are well-known features.

他にも、下記のような日本語文の文分割が不適切なデータも存在した。このようなデータは学習コーパスに含めると湧き出しの原因にもなるため、可能な限り学習コーパスから除外したほうがよい。

（請求項 4、段落【0020】、段落【0022】参照。）	(See claim 4 and paragraphs [0020] and [0022].)
「圧縮機 3 内の油量が一定量以上減少していると判断された場合」とは、請求の範囲 1 に係る発明の「油量算出手段(51)の算出値が所定値以上となったとき」に相当する。	"When the amount of oil within the compressor (3) is determined to have decreased by at least a certain amount" corresponds to "when the calculated value of the oil amount calculation means (51) reaches or exceeds a predetermined value" in the invention as in claim 1.

#### 4.2.2 数字が対応しない学習コーパスの除外

3.2.1 の湧き出しで分析した「同文献 / document 1」のような湧き出しを防ぐため、学習コーパスの日本語と英語との間で数字の対応を確認し、対応していないデータを学習対象外とする対応が有効と考えられる。

例えば、下記の学習コーパスの日本語文の数字「0002」は英語文にも存在するが、英語文の「2」は日本語文に存在しない。このようなデータは数字が対応していないデータとして学習コーパスから除外したほうがよい。

例 35.

また、同文献の【0022】には、メッキに用いる金属としてニッケルが挙げられている。	In addition, document 2 (paragraph [0022]) presents nickel as a metal to be used for plating.
---	---

その他、引用文献名など固有名詞がローマ字表記されたコーパスを学習すると誤訳の原因になるため注意が必要となる。

例 36.

原文	機械翻訳文
ハイスコア大会	high score <u>Taikai</u>

学習コーパス	
文献 5 :日本農芸化学会 <u>大会</u> 講演要旨集, 2011.03.05	Document 5:Japan Society for Bioscience, Biotechnology, and Agrochemistry <u>Taikai</u> Koen Yoshishu, 05 March 2011
文献 2 :日本菌学会第 54 回 <u>大会</u> 講演要旨集,2010,p.31,2SA2	Document 2:The Mycological Society of Japan Dai 54 Kai <u>Taikai</u> Koen Yoshishu, 2010, pp. 31, 2SA2
文献 4 :日本臨床検査自動化学会第 39 会 <u>大会</u> 抄録集, 2007.09.01, p.434	Document 4:The Japan Society for Clinical Laboratory Automation Dai 39 Kai <u>Taikai</u> Shorokushu, 01 September 2007, page 434

#### 4.3 未知語の取り扱いについて

本調査は比較的小規模の学習コーパスを使用したこともあり、翻訳結果にいくつかの未知語が見られた。未知語は学習コーパスを追加することで減少していくと思われる。

その一方で、追加した学習コーパスにより、新たに訳語のぶれなどが発生する可能性も生じる。また、学習コーパスを追加し、再度学習させる手法は多くの処理時間を要するため、未知語対応が迅速に行えないという課題が存在する。この対策の一つとして、未知語となった用語の対訳辞書を作成し、機械翻訳の前処理または後処理において対訳辞書の内容で置換する手法が挙げられる。

## 5. 付録

### 5.1 自動評価手法の概要<sup>17</sup>

機械翻訳の評価手法として古くから人手による評価が行われ、現在も行われている。しかしながら、人手による評価はコストがかかるため、コンピュータを用いた自動評価ができれば望ましい。そのような動機から自動評価が考えられ、これまでに多くの自動評価基準が提案されている。自動評価では、まず試験文（原文）に対して基準翻訳文と呼ばれる、いわば理想的な訳文を用意しておく。その後、原文を機械翻訳し、機械翻訳文と基準翻訳文とがどれほど近いかで翻訳精度を評価するものである。基準翻訳文に近い機械翻訳文ほど翻訳精度が高いと判断する。このときの「近さ」をどのようにして測るかで様々な自動評価基準が考えられる<sup>18</sup>。ここでは、今回使用した 3 種類の自動評価基準である BLEU、NIST、RIBES について、具体例を交えて説明する。

#### 5.1.1 BLEU

BLEU (Bilingual Evaluation Understudy) は機械翻訳の自動評価基準として最初に提案された基準であり、現在でも多く使われている。BLEU は以下のようにして求められる。基準翻訳文と機械翻訳文の双方を単語分割し<sup>19</sup>、得られる単語  $n$  グラム（連続する  $n$  個の単語） $w_1, w_2, \dots, w_n$  に対して、基準翻訳文と機械翻訳文中の頻度を求める。

$C^{tst}(w_1, w_2, \dots, w_n)$  = 機械翻訳文中の  $w_1, w_2, \dots, w_n$  の頻度

$C^{ref}(w_1, w_2, \dots, w_n)$  = 基準翻訳文中の  $w_1, w_2, \dots, w_n$  の頻度

そして、 $C_{clip}^{tst}(w_1, w_2, \dots, w_n)$  を次のようにして求める。

$$C_{clip}^{tst}(w_1, w_2, \dots, w_n) = \min(C^{tst}(w_1, w_2, \dots, w_n), C^{ref}(w_1, w_2, \dots, w_n))$$

つまり、機械翻訳文中の頻度と基準翻訳文中の頻度のうちの小さい方の値である。具体例で説明する。

[ 例文 1 ]

原文 コネクタ本体 14 が基板 12 上に実装される

基準翻訳文 The connector main body 14 is mounted on the substrate 12

機械翻訳文 The connector body 14 is mounted on the substrate body 12

<sup>17</sup> 5.1.1 から 5.1.4 までは「平成 27 年度 特許審査関連情報の日英機械翻訳文の品質評価に関する調査報告書

( [https://www.jpo.go.jp/shiryou/toushin/chousa/pdf/kikai\\_honyaku/h27\\_03.pdf](https://www.jpo.go.jp/shiryou/toushin/chousa/pdf/kikai_honyaku/h27_03.pdf) )」からの引用となる。

<sup>18</sup> 原文を翻訳する場合、理想的な訳文は一つとは限らない。そこで、基準翻訳文を複数用意して近さを測る方法もあるが、特許翻訳の場合、単一の基準翻訳文を用いることが多く、今回の評価でもそのようにしている。そこで、本文では単一の基準翻訳文の場合についてのみ説明する。また、1 文単位で評価するのではなく複数の文からなるドキュメント単位で評価する場合があるが、ここでは文単位の評価について考える。

<sup>19</sup> ここでは、大文字と小文字を区別せず、“The”と“the”は同一の単語とみなす。また、ピリオドやコンマなどはその前の単語と分離して単独の単語として扱う。

とするとき、 $n=1$  として  $w_1 = \text{"body"}$  の場合を考える。このとき、機械翻訳文中の  $\text{"body"}$  の頻度は 2 であり、基準翻訳文中の  $\text{"body"}$  の頻度は 1 である。そこで、 $C^{tst}(\text{body}) = 2$ 、 $C^{ref}(\text{body}) = 1$ 、 $C_{clip}^{tst}(\text{body}) = 1$  となる。

次に、機械翻訳文に含まれるすべての  $n$  グラム  $G_n$  を用いて  $n$  グラムの適合率  $p_n$  を以下の式で求める。

$$p_n = \frac{\sum_{w_1, w_2, \dots, w_n \in G_n} C_{clip}^{tst}(w_1, w_2, \dots, w_n)}{\sum_{w_1, w_2, \dots, w_n \in G_n} C^{tst}(w_1, w_2, \dots, w_n)}$$

例文 1 の場合、 $G_1 = \{\text{the, connector, body, 14, is, mounted, on, substrate, 12}\}$  であり、

$$p_1 = \frac{2 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1}{2 + 1 + 2 + 1 + 1 + 1 + 1 + 1 + 1} = \frac{10}{11} = 0.909$$

となる。 $p_n$  の相乗平均をもとに BLEU は以下のように計算される。

$$BLEU = BP \cdot \sqrt[N]{p_1 p_2 \dots p_N}$$

ここで、 $N$  は考慮する最大のグラム数であり、通常は 4 が用いられる。また、BP は brevity penalty と呼ばれる補正項であり、機械翻訳文が短い場合に BLEU の値が大きくなってしまふ不具合を修正するもので

$$BP = \begin{cases} 1, & c > r \\ e^{(1-r/c)}, & c \leq r \end{cases}$$

で定義される<sup>20</sup>。ここで  $c$  は機械翻訳文の単語数であり、 $r$  は基準翻訳文の単語数である。つまり、 $c$  が  $r$  より少ないときにペナルティを課す項である。たとえば、例文 1 の場合、機械翻訳文が  $\text{"body 14 is mounted"}$  という断片的な訳文でも  $N=4$  としたときには  $p_1 = p_2 = p_3 = p_4 = 1$  となり  $\sqrt[4]{p_1 p_2 \dots p_N} = 1$  となってしまうが、 $r/c = 11/4 = 2.75$  となり、 $BP = 0.174$  であるから  $BLEU = 0.174$  と小さな値となる。

$0 \leq p_n \leq 1$  が成り立ち、 $0 < BP \leq 1$  であるから、BLEU は 0 と 1 の間の数値をとり、大きいほど機械翻訳文が基準翻訳文に近いことになる。

例文 1 の場合、 $p_1 = 0.9091$ 、 $p_2 = 0.7$ 、 $p_3 = 0.5556$ 、 $p_4 = 0.5$  となり、 $\sqrt[4]{p_1 p_2 p_3 p_4} = 0.6484$  である。さらに  $BP = 1$  であるため  $BLEU = 0.6484$  となる。

### 5.1.2 Smoothed BLEU

BLEU の問題点のうちの一つは、 $p_n$  の相乗平均を用いているため、 $p_n = 0$  となる  $n$  が一つでもあると  $BLEU = 0$  となってしまうことである。このため、オリジナルの BLEU では文単位の評価に用いるのではなく、ドキュメント単位の評価に用いるべきものとしている。ドキュメントにおいては 4 グラムのような比較的大きな  $n$  に対しても  $p_n = 0$  となることは考えにくく、その場合  $BLEU = 0$  となることもないからである。しかし、BLEU を文単位の評価に用いたいという要望は多く、上記の問題に対処する方法がいくつか提案されている。ここでは NIST からフリーの評価ツール mteval-v13a.pl として提供されている方法について説明

<sup>20</sup>  $c/r$  に対する BP の値は後述する図 5.1-1 に示されている。



する。n グラムの適合率を以下のように修正する。

$$\tilde{p}_n = f(x) = \begin{cases} \frac{\sum_{w_1, w_2, \dots, w_n \in G_n} C_{clip}^{tst}(w_1, w_2, \dots, w_n)}{\sum_{w_1, w_2, \dots, w_n \in G_n} C^{tst}(w_1, w_2, \dots, w_n)}, & p_n \neq 0 \\ \frac{1/2^{(n-k+1)}}{\sum_{w_1, w_2, \dots, w_n \in G_n} C^{tst}(w_1, w_2, \dots, w_n)}, & n \geq k \text{ where } p_{k-1} \neq 0 \text{ and } p_k = 0 \end{cases}$$

例えば、 $p_4 = 0$  で  $p_3 \neq 0$  の場合、 $\tilde{p}_4$  の分子は、 $1/2^{(4-4+1)} = 1/2$  となる。 $\tilde{p}_n$  を用いて

$$BLEU^{\sim} = BP * \sqrt[N]{\tilde{p}_1 \tilde{p}_2 \dots \tilde{p}_N}$$

として smoothed BLEU が求められる。

具体例を示す。以下の例文 2 を考える。なお、以下で用いる例文の原文と基準翻訳文はすべて例文 1 と同じであるので記述を省略する。

[例文 2]

機械翻訳文 Connector main part 14 is mounted in substrate 12

機械翻訳文に含まれる 4 グラムは {Connector main part 14, main part 14 is, part 14 is mounted, 14 is mounted in, is mounted in substrate, mounted in substrate 12} であるが、これらはいずれも基準翻訳文には含まれていない。したがって  $p_4 = 0$  となり、オリジナルの BLEU は 0 となってしまう。一方、 $\tilde{p}_4 = 1/(2 \times 6) = 0.0833$  となり、 $BLEU^{\sim} = 0.2089$  となる。

### 5.1.3 NIST

NIST (National Institute of Standards and Technology) <sup>21</sup> は基本的に BLEU の考え方を継承し、改良を加えた評価基準であり、以下の式で定義される。

$$NIST = BP_{NIST} \sum_{n=1}^N q_n$$

ここで  $q_n$  は

$$q_n = \frac{\sum_{w_1, w_2, \dots, w_n \in G_n} \text{Info}(w_1, w_2, \dots, w_n)}{\sum_{w_1, w_2, \dots, w_n \in G_n} 1}$$

で与えられる。N は通常 5 が用いられる。NIST は 0 以上であるが、BLEU と異なり 1 以上の値をとることもある。定義式において BLEU と異なる点は 3 点ある。

- (1) BLEU は相乗平均を用いていたが NIST では和を用いている。
  - (2) brevity penalty の計算方法が異なる。
  - (3)  $q_n$  の式の分子の関数形が BLEU では  $C_{clip}^{tst}(w_1, w_2, \dots, w_n)$  であったが NIST では  $\text{Info}(w_1, w_2, \dots, w_n)$  となっている。分母も BLEU では機械翻訳文に含まれる n グラムの延べ度数であったが、NIST では異なり度数となっている。
- (1) によって  $q_n = 0$  となる n があっても NIST=0 となるとは限らない。つまり smoothing を必

<sup>21</sup> NIST は組織の名前であるが、この自動評価基準は、NIST が主催している評価型 workshop で用いられており、NIST の名で呼ばれている。

ずしも必要としない。(2)に関しては

$$BP_{NIST} = \exp\{\beta \log^2 \left[ \min\left(\frac{c}{r}, 1\right)\right]\}$$

である。 $\beta$ は  $c/r=2/3$  のときに  $BP_{NIST} = 0.5$  となるように定められる。 $c/r$  に対する  $BP$  と  $BP_{NIST}$  の値の変化を図 5.1-1 に示す。 $c/r > 0.761$  の時に  $BP_{NIST} \geq BP$  となり、 $c/r < 0.761$  の時には逆になる。

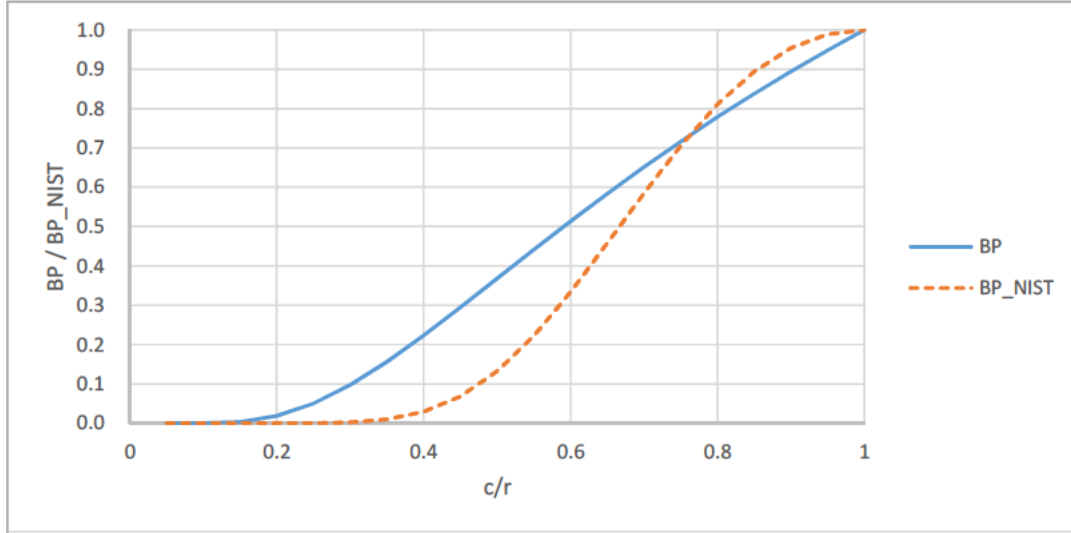


図 5.1-1 基準翻訳文の単語数( $r$ )と機械翻訳文の単語数( $c$ )の比に対する brevity penalty

(4)に示す Info は以下のように定義され、基準翻訳文において、 $w_1, w_2, \dots, w_{n-1}$  が持つ情報量から  $w_1, w_2, \dots, w_n$  が持つ情報量への増加分を表している。ここで  $C^{ref}(w_1, w_2, \dots, w_n)$  は前述した基準翻訳文に含まれる  $w_1, w_2, \dots, w_n$  の頻度である。

$$\text{Info}(w_1, w_2, \dots, w_n) = \begin{cases} 0, & C^{ref}(w_1, w_2, \dots, w_n) = 0 \\ \log_2\left(\frac{C^{ref}(w_1, w_2, \dots, w_{n-1})}{C^{ref}(w_1, w_2, \dots, w_n)}\right), & \text{else} \end{cases}$$

但し、 $n=1$  の場合は、 $C^{ref}(w_1, w_2, \dots, w_{n-1})$  を基準翻訳文に含まれる全単語数とする。

例文 1 において Info は以下ようになる。

$$\text{Info}(\text{the}, \text{substrate}) = \log_2\left(\frac{C^{ref}(\text{the})}{C^{ref}(\text{the}, \text{substrate})}\right) = \log_2\left(\frac{2}{1}\right) = 1$$

$$\text{Info}(\text{mounted}, \text{on}) = \log_2\left(\frac{C^{ref}(\text{mounted})}{C^{ref}(\text{mounted}, \text{on})}\right) = \log_2\left(\frac{1}{1}\right) = 0$$

$$\text{Info}(\text{the}) = \log_2\left(\frac{11}{C^{ref}(\text{the})}\right) = \log_2\left(\frac{11}{2}\right) = 2.46$$

$$\text{Info}(\text{mounted}) = \log_2 \left( \frac{11}{C^{\text{ref}}(\text{mounted})} \right) = \log_2 \left( \frac{11}{1} \right) = 3.46$$

となる。例文 1 に対して、 $q_1 = 2.9631$ ,  $q_2 = 0.2$ ,  $q_3 = 0$ ,  $q_4 = 0$  となり、さらに  $BP_{NIST} = 1$  となるため  $NIST=3.1631$  である。

#### 5.1.4 RIBES

BLEU と NIST のもう一つの問題点は、n グラムという局所的な情報のみを用いているため、語順の大幅な入れ替えがある日英翻訳などの評価に用いた場合、評価の信頼性が劣る点である。実際 NTCIR-9 の特許翻訳タスク、日英サブタスクでは、BLEU や NIST の値と人手評価結果 (Adequacy) との相関係数が負になっている。そこで大域的な語順の類似性も考慮した自動評価基準として RIBES が提案されている。RIBES では基準翻訳文と機械翻訳文との語順の類似性を Kendall の順位相関係数を用いて求めている。Kendall の順位相関係数は次のようにして計算される。まず機械翻訳文と基準翻訳文との間に単語対応を求め、2 個の単語対応の組み合わせについて単語の出現順序が機械翻訳文と基準翻訳文との間で同順か逆順かの場合数を計数する。

具体例で説明する。

[例文 3]

機械翻訳文 On the substrate 12 , the connector main body 14 is mounted  
 の場合を考えると、単語対応が図のように得られる。

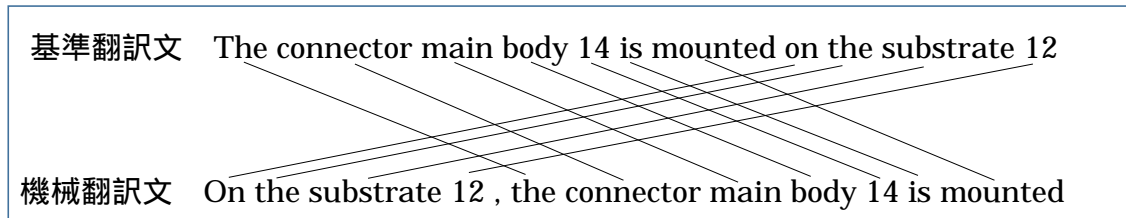


図 5.1-2 単語対応例

このとき 2 個の単語対応の組み合わせに対して同順(S)と逆順(I)が図 5.1.3 のように求まる。例えば、on と substrate の組み合わせの場合、基準翻訳文での出現位置は、おののお 8 と 10 であり、機械翻訳文での出現位置は 1 と 3 である。つまり基準翻訳文と機械翻訳文のいずれでも on より substrate のほうが後に出現しており、同順となる。一方、on と connector では、基準翻訳文での出現位置が、おののお 8 と 2 であるのに対して、機械翻訳文での出現位置は 1 と 6 である。そこで逆順となる。

	the	substrate	12	the	connector	main	body	14	is	mounted
on	S	S	S	I	I	I	I	I	I	I
the		S	S	I	I	I	I	I	I	I
substrate			S	I	I	I	I	I	I	I
12				I	I	I	I	I	I	I
the					S	S	S	S	S	S
connector						S	S	S	S	S
main							S	S	S	S
body								S	S	S
14									S	S
is										S

図 5.1-3 例文 3 の単語対応の同順(S)と逆順(I)

同順の個数を $C_S$ とし、逆順の個数を $C_I$ とする。また、単語対応の個数を $n$ とすると、2 個の単語対応の組み合わせの総数は $n(n-1)/2$ である。このとき Kendall の順位相関係数( )は

$$= \frac{C_S - C_I}{\frac{n(n-1)}{2}}$$

で与えられる。2 個の単語対応の組み合わせについて、すべてが同順であれば  $= 1$  となり、全てが逆順になれば  $= -1$  となる。さらに の範囲を $[0,1]$ に正規化した NKT (Normalized Kendall's ) を

$$\text{NKT} = \frac{\tau + 1}{2}$$

と定義する。

例文 3 の場合は

$$= \frac{27 - 28}{\frac{11(11-1)}{2}} = -0.018$$

$$\text{NKT} = \frac{-0.018 + 1}{2} = 0.491$$

となる。

や NKT には次の性質があり、このままでは自動評価基準として不適切である。

- (1) 単語対応の個数が少なく、2 個の単語対応の組み合わせに同順が多い場合に高い値となる。
- (2) 機械翻訳文が短く、2 個の単語対応の組み合わせに同順が多い場合に高い値となる。

例えば、

[例文 4]

機械翻訳文 A connected major part 15 was inserted in the substrate 12

とすると、単語対応は ( the, substrate, 12 ) のみであり、2 個の単語対応の組み合わせは全て同順であるので NKT=1 となる。また

[例文 5]

機械翻訳文 is mounted

に対しても、NKT=1 となる。例文 4 は(1)の結果であり、例文 5 は(2)の結果である。

RIBES では、上記問題点の(1)に対応するために以下に述べるペナルティ P を用いている。また、(2)に対応するために BLEU の brevity penalty と同じ BP を用いている。ここで、P は以下の式で定義される。

$$P = \frac{n}{c}$$

n は単語対応の個数、c は機械翻訳文の単語数である。最終的に RIBES は

$$\text{RIBES} = \text{NKT} \cdot P^\alpha \cdot BP^\beta$$

として定義される。ここで  $\alpha \geq 0$  と  $\beta \geq 0$  はパラメータであり、 $\alpha = 0.25$ 、 $\beta = 0.1$  が初期設定されている。RIBES は 0 以上 1 以下の値をとり、大きいほど基準翻訳文と機械翻訳文が近いと評価される。

[例文 6]

機械翻訳文 The substrate 12 is mounted on the connector main body 14

とすると、例文 6 は基準翻訳文とまったく異なる意味を持つが、局所的な n グラムの範囲では基準翻訳文と類似している。その結果、BLEU = 0.375 となる。一方、大域的語順を考慮した RIBES では、RIBES = 0.273 と BLEU に比較して小さい値となる。

実際、前述した NTCIR-9 のデータでは、BLEU や NIST の値と人手評価結果 ( Adequacy ) との相関係数が負になっていたが、RIBES の値と Adequacy との相関係数は正となっている。

### 5.1.5 BLEU と RIBES の特徴

本調査では機械翻訳の自動評価として BLEU と RIBES を使用したが、ここでは各自動評価の特徴を把握するため、各自動評価の値の乖離が大きい文について原因を調査した。調査の結果、BLEU は重大な訳抜けが評価値に影響を及ぼし、RIBES は語順が評価値に影響を及ぼす傾向にあることが分かった。

下記例の ニューラル機械翻訳は BLEU 値が 0.13, RIBES 値が 0.7 と値が大きく異なった。 ニューラル機械翻訳を調べると原文「(特に、段落【0074】 - 【0098】、【0

103】、【図5A】 - 【図5B】、【図8】参照。)」の部分が訳抜けしていることがわかった。次にこの訳抜けが、BLEU が低くなった原因であるのかを確認するため、訳抜け部分を補った文を作成し、その内容を評価した。なお、補った内容と位置は 統計機械翻訳の内容を利用した。

その結果、BLEU 値が 0.13 から 0.59 となり、BLEU が RIBES に対して極端に低くならず、且つ の統計機械翻訳の BLEU 値が 0.56 と近い値となったことから、前記推測のとおり、訳抜けにより BLEU が低くなったと考えられる。

一方、RIBES 値を見ると 0.74 だったものが 0.58 と逆に値が低くなった。この原因を調べるため ニューラル機械翻訳で補った訳抜けを基準翻訳の位置に合わせて移動した文をニューラル機械翻訳として作成し、その内容を評価したところ、RIBES 値が 0.58 から 0.80 になった。このことから、訳語の出現位置が RIBES の値に影響する度合いが強いことが分かった。

例 37.

原文	刊行物 A に記載された発明（特に、段落【0074】 - 【0098】、【0103】【図5A】 - 【図5B】、【図8】参照。）における「励磁巻線」、「検出巻線」、「フィードバック巻線」は、それぞれ、請求項 1、4、7 - 8 に係る発明における「励磁コイル」、「検出コイル」、「帰還用コイル」に相当する。
基準翻訳	An "excitation winding," a "detection winding," and a "feedback winding" in the invention described in cited publication A ( <a href="#">refer to paragraphs [0074]-[0098], [0103], [Fig. 5A]-[Fig. 5B], and [Fig. 8] in particular</a> ) correspond to an "exciting coil," a "detection coil," and a "feedback coil" in the inventions claimed in claims 1, 4, 7-8, respectively.
統計機械翻訳 (B: 0.567) (R: 0.567226)	"excitation winding" in), "detection winding", "feedback winding" respectively "excitation coil" in the invention as in claims 1, 4, 7-8, and "detection coil" corresponds to "feedback coil" described in Publication A ( <a href="#">in particular, paragraphs [0074] - [0098], [0103], [5], Fig. 5 b], [Fig. 8].</a> ).
ニューラル機械翻訳 (B: <b>0.1316</b> ) (R: 0.74019)	The " exciting winding wire ", " detection winding wire " and " feedback winding " in the invention described in publication A respectively correspond to the " excitation coil, " " detection coil, " and " feedback coil " in the invention as in claims 1, 4, and 7-8.
ニューラル機械翻訳（訳抜け補正） (B: 0.5996) (R: <b>0.583401</b> )	The " exciting winding wire ", " detection winding wire " and " feedback winding " in the invention described in publication A respectively correspond to the " excitation coil, " " detection coil, " and " feedback coil " in the invention as in claims 1, 4, and 7-8 ( <a href="#">in particular, paragraphs [0074] - [0098], [0103], [5], Fig. 5 b], [Fig. 8].</a> ).
ニューラル機械翻訳（訳抜け補正 2） (B: 0.6034) (R: <b>0.801803</b> )	An "excitation winding," a "detection winding," and a "feedback winding" in the invention described in cited publication A ( <a href="#">refer to paragraphs [0074]-[0098], [0103], [Fig. 5A]-[Fig. 5B], and [Fig. 8] in particular</a> ) correspond to an "exciting coil," a "detection coil," and a "feedback coil" in the inventions claimed in claims 1, 4, 7-8, respectively.

## 5.2 日英対訳コーパスの内訳（フォーマット A）

表 5.2-1 日英対訳コーパスの内訳（フォーマット A）

審査第一部	認定	対比	判断	合計
文数	311	58	386	755
1 文あたりの文字数（日本語）平均	110.4	115.4	103.2	107.1
1 文あたりの文字数（日本語）最小	8	29	14	8
1 文あたりの文字数（日本語）最大	634	425	455	634
審査第二部	認定	対比	判断	合計
文数	371	21	326	718
1 文あたりの文字数（日本語）平均	113.0	126.2	96.6	105.9
1 文あたりの文字数（日本語）最小	7	35	15	7
1 文あたりの文字数（日本語）最大	564	544	515	564
審査第三部	認定	対比	判断	合計
文数	215	31	252	498
1 文あたりの文字数（日本語）平均	122.0	120.2	111.5	116.6
1 文あたりの文字数（日本語）最小	8	36	16	8
1 文あたりの文字数（日本語）最大	785	326	427	785
審査第四部	認定	対比	判断	合計
文数	266	59	375	700
1 文あたりの文字数（日本語）平均	136.5	116.6	113.3	122.4
1 文あたりの文字数（日本語）最小	8	28	14	8
1 文あたりの文字数（日本語）最大	567	759	813	813
合計文数	1,163	169	1,339	2,671