

**平成 29 年度
審決の日英機械翻訳に関する調査
調査報告書**

平成 30 年 3 月

株式会社サン・フレア

目次

1. 調査目的	3
2. 調査の概要	4
2.1 調査方法	4
2.2 実施体制	7
2.3 調査スケジュール	8
3. 審決の機械翻訳文の調査・分析	9
3.1 審決の機械翻訳文の人手評価	9
3.1.1 審決の機械翻訳文の人手評価の結果	10
3.1.2 審決の機械翻訳文の分析コメント	23
3.2 審決の機械翻訳文の自動評価結果	25
3.2.1 審決の機械翻訳文の自動評価結果	27
3.2.2 自動評価結果の分析	32
4. 昨年度事業で作成した辞書データの有効性調査（新旧機械翻訳文の比較）	35
4.1 新旧機械翻訳文の評価結果	35
4.1.1 新旧機械翻訳文の人手評価の結果	35
4.1.2 新旧機械翻訳文の自動評価の結果	38
4.2 新旧機械翻訳文の比較分析（昨年度調査対象の審決）	40
4.2.1 学習用データの反映により機械翻訳文の精度が向上した例	40
4.2.2 学習用データの反映により生じた問題点	42
4.3 新旧機械翻訳文の比較分析（今年度調査対象の審決）	44
4.3.1 学習用データの反映により機械翻訳文の精度が向上した例	44
4.3.2 学習用データが想定通りに反映されなかった例	45
4.4 辞書データによる解決が困難な例	47
5. 辞書データの作成・分析	51
5.1 辞書データの作成手順	51
5.2 作成結果	54
5.3 登録語の編集における工夫	56
5.4 辞書データ作成についての提言	57

1. 調査目的

近年、海外企業等による日本への投資を促進するための基盤整備、あるいは日本企業が海外に進出する際の進出先への情報提供等のため、我が国の法制度の運用に関する情報の英語での発信が強く求められている¹²。この要請を受け、裁判所が判決の英訳を公表する他、特許庁においても一部の審決について、人手翻訳した英文を外部に提供する等している。

人手翻訳は、費用等の面から提供できる量に限界があることから、特許庁では平成 28 年 1 月から、独立行政法人工業所有権情報・研修館の特許情報プラットフォーム（J-PlatPat）を通じて、審決の機械翻訳文を外部に提供している。現在、審決に関する情報の海外への発信を進めていくべく、機械翻訳の精度についてさらなる向上が求められているところである。

そこで本調査では、J-PlatPat を通じて審決の機械翻訳文を外部へ提供するにあたり、人手翻訳された審決や判決を利用して機械翻訳の精度向上のための分析を行い、翻訳精度を向上させる際の課題を明らかにすることを目的とする。

¹ 「国際戦略に関する提言」 自民党政務調査会 司法制度調査会 民事・国際小委員会（平成 26 年 5 月 30 日）

² 「知的財産推進計画 2015」知的財産戦略本部（平成 27 年 6 月 19 日）

2. 調査の概要

2.1 調査方法

2.1.1 評価対象文献

特許庁が指定した日本語の審決 100 件を対象に機械翻訳文の精度の評価・分析を実施した。

2.1.2 調査手順

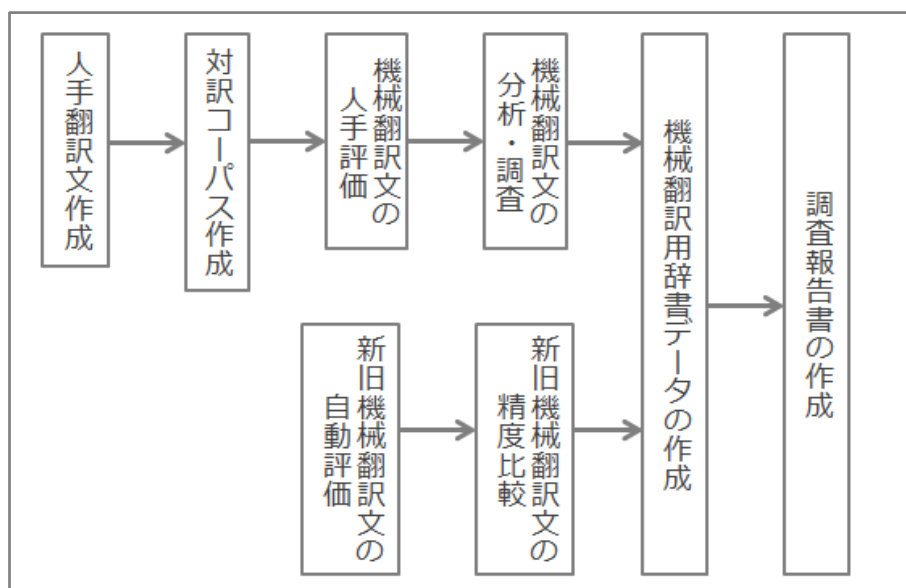


図 2.1.2-1 作業手順の全体フロー図

(1) 基準翻訳文の作成

特許庁が指定した日本語の審決 100 件を英語に人手翻訳し、その後第一校閲者及び第二校閲者による重畳的な校閲工程を経た人手翻訳文を作成した。

基準翻訳文の作成にあたっては、スケジュールの都合上、全 100 件の審決について 10 件～13 件を 1 バッチとした計 9 バッチでバッチ単位に設定したスケジュールで進捗管理を行った。

翻訳者・校閲者の採用条件：

① 翻訳者

産業財産権の出願手続等の事務手続に関する文書の日英翻訳経験を 1 年以上有する者

配置人数：約 20 名

② 第一校閲者

英語を母国語とし、産業財産権の事務手続に携わる経験（弁理士としての業務経験等）もしくはそれに足る知識を有する者

配置人数：約 5 名

③第二校閲者

産業財産権の事務手続に携わる経験もしくはそれに足る知識を有し、かつ、日英翻訳の校閲経験を 5 年以上有する者

配置人数：約 10 名

(2)機械翻訳文の評価用フォーマットの作成

基準翻訳文を特許庁に提出後、1～9 バッチの全ての審決について原文、基準翻訳文、J-PlatPat から取得した機械翻訳文のコーパスを作成した。

全 9 バッチのうち 1～4 バッチについては、昨年度の同事業にて作成した機械翻訳の学習用データを反映させる前後の機械翻訳文の精度比較を行うため、上記の 3 列コーパスに更に学習用データ反映前の機械翻訳文（旧版 MT）を加えた 4 列のコーパスを作成した。

(3)機械翻訳文の評価（人手評価及び自動評価）

(2)で作成した評価用フォーマット上で機械翻訳文の精度を一文単位で人手により評価した。1～4 バッチについては学習用データ反映後の機械翻訳文（新版 MT）と旧版 MT の両方の機械翻訳文の評価付けを行い、5～9 バッチは新版 MT のみの評価を行った。

評価手順としては、産業財産権の出願手続等の事務手続に関する文書の日英翻訳経験を 1 年以上有する者が 1 次評価を行い、その後、翻訳品質管理責任者が一元的に確認を行って個人差の是正を図るものとした。また、1 次評価の担当者には作業前に指示書と評価サンプルを配布して評価基準の摺り合わせを行った。

評価対象は、審決及び判決のうち、技術的な説明文を除いた審決特有の表現とし、評価基準は「特許文献機械翻訳の品質評価手順 Ver. 1.0」³に記載の下記「内容の伝達レベルの評価」に準じた。具体的には、人手で翻訳した訳文と機械翻訳による訳文とを比較し、原文の意味内容が正しく伝わる度合いを尺度にして機械翻訳文の精度を 1～5 の 5 段階を割り当てることで主観的に評価した。

内容の伝達レベルの評価

- | |
|--------------------------------|
| 5: すべての重要情報が正確に伝達されている。(100%) |
| 4: ほとんどの重要情報は正確に伝達されている。(80%～) |
| 3: 半分以上の重要情報は正確に伝達されている。(50%～) |

³ https://www.jpo.go.jp/shiryou/toushin/chousa/pdf/tokkyohonyaku_hyouka/01.pdf

2: いくつかの重要情報は正確に伝達されている。(20%~)

1: 文意がわからない、もしくは正確に伝達されている重要情報はほとんどない。(~20%)

特許文献機械翻訳の品質評価手順 Ver. 1.0

さらに全件について自動評価2種（BLEU 及び RIBES）による評価も行った。

(4) 機械翻訳文の分析・調査

主に(3)で評価数値が低かった文を対象に調査し、審決の機械翻訳文における問題点を分析して機械翻訳文の精度向上に寄与する機械翻訳用辞書データの作成方法を検討した。

また、1~4 バッチについては上記に加え、新旧機械翻訳文の精度を比較し、昨年度作成した機械翻訳の学習用データの適用による機械翻訳文精度への影響を調査した。

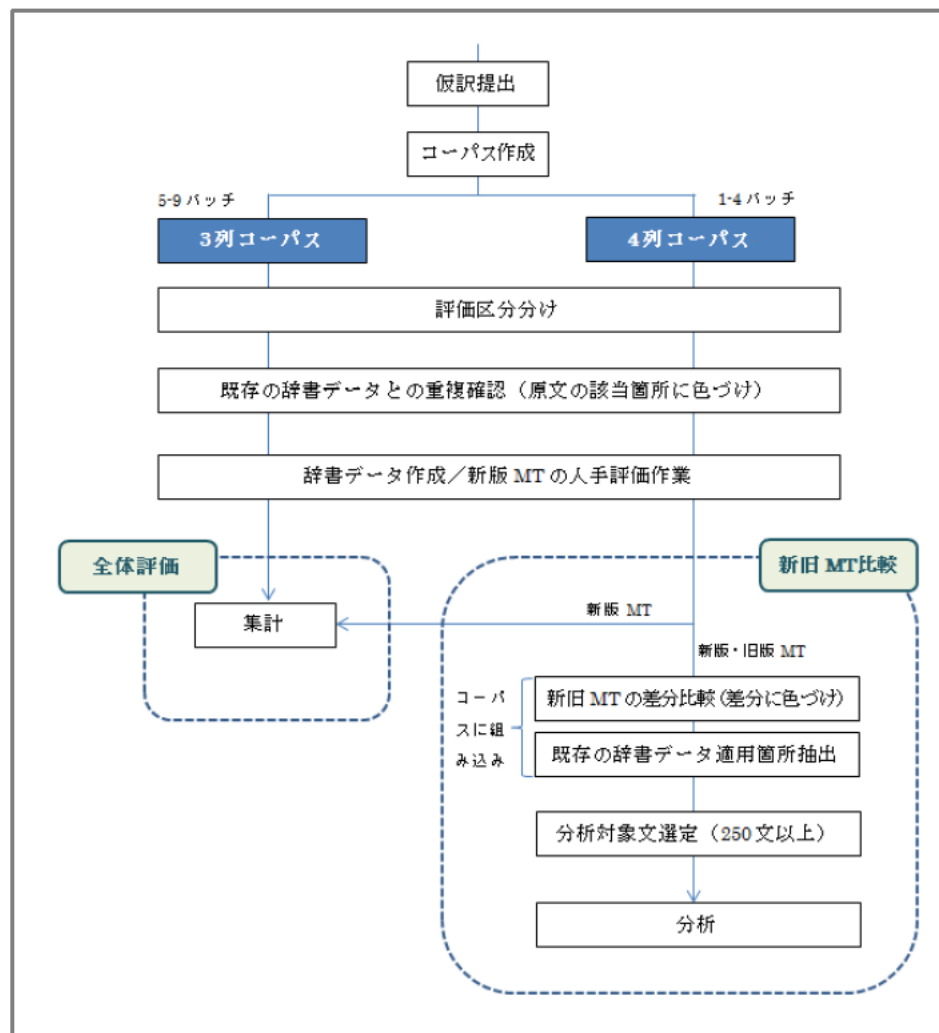


図 2.1-2-2 新旧 MT 比較の作業フロー図（今年度調査対象審決）

(5) 辞書データ作成

「(4)機械翻訳文の分析・調査」の調査結果と得られた知見をもとに、審決の機械翻訳文の精度向上に有効と思われる辞書データ作成を行った。

審決 1 件あたり 10 語程度となるよう辞書データを作成し、既存の辞書データとの重複を避けるため、作業時には原文に既存の辞書データが存在する語句に色づけを施したファイルを渡し、色づけされた語句は避けて辞書データ登録を行うよう指示した。

なお、作成した辞書の種類は以下の 3 種類である。

①ユーザ辞書

数字以外の可変部分を含まない語や句が対象となる。可変部分は数字のみ設定できるため「特許法第 17 条の 2 第 4 項の要件違反になる」のように、数字以外の訳を指定したい場合も本辞書に登録できる。

②メモリ辞書

可変部分を含まない文単位の対象となる。辞書データの原語と翻訳対象の原文が完全一致しないと適用されないため、適用される箇所が少ないというデメリットがある。一方、適用された場合は、完全な文として訳出されるので、副作用の心配が少ないというメリットがある。定型的で汎用文に近い文はメモリ辞書で登録すると効果が期待できる。また、複数のバリエーションを展開した上で登録するとより効果的である。

③パターン辞書

可変部分を含む文頭から文末までの文単位で対訳文が対象となる。原文と訳文の翻訳例に可変部分を指定した翻訳パターン（穴空き翻訳例文）に登録する。パターン辞書は、文章のひな型として利用する前提となっているため、登録する範囲は主語、述語のある完全な文とする。

2.2 実施体制

本事業に従事した担当者及び役割は下記の通りである。

責任者/担当者	作業内容	人数
統括責任者	本事業全体に対する統括を行う。	1 名
翻訳品質管理責任者	本事業の審決の人手翻訳及び辞書データ登録語句の品質に対して責任を負う。	1 名
機械翻訳用辞書責任者	対訳コーパスと機械翻訳用辞書の作成において、フォーマットデータの作成と管理、作業者への指導・監督を行う。	1 名

翻訳者	調査対象の審決 100 件を人手により英訳する。	約 20 名
第一校閲者	英訳された審決を英語表現の観点から校閲を行う。	約 6 名
第二校閲者	英訳された審決に訳抜け、誤訳等がないか原文との突き合わせによる校閲を行う。	約 10 名
コーパス作成担当者	アライメントツールで機械抽出された日英対訳コーパスの確認・修正を行う。	約 5 名
辞書データ作成担当者	機械翻訳文を文単位で評価し、主に評価数値が低かった文を対象に辞書データへの登録語句を選定する。	約 11 名

表 2.2-1 実施体制

2.3 調査スケジュール

調査対象の審決 100 件は 10～13 件を 1 バッチとして全 9 バッチに分割し、バッチ毎のスケジュールに沿って人手翻訳文作成、対訳コーパス作成、人手評価、辞書データ作成を実施した。

			2018																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
内容	作業開始日	作業終了日	10月				11月				12月				1月				2月				3月																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										
基準翻訳文作成	2017/5/1	2018/1/29	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●

表 2.2-2 調査スケジュール

3. 審決の機械翻訳文の調査・分析

本調査では、審決 100 件を対象に、審決及び判決特有の表現を含む文を選定して人手による評価を行った。その評価結果について以下に説明する。

3.1 審決の機械翻訳文の人手評価

調査対象である審決 100 件の全文のうち、技術的な説明文を除いた 15,286 文を対象に評価を行った。四法毎、審判種別毎の件数及び文数の内訳は下記の通りである。なお、今回の調査対象には実用新案は含まれていなかった。

(1) 審決 100 件中の評価対象件数

審判種別	四法			対象件数
	特許	商標	意匠	
無効	30	3	2	35
不服	40	4	2	46
判定	5	1		6
訂正	6			6
取消		3		3
異議	2	2		4
合計	83	13	4	100

表 3.1-1 四法毎／審判種別毎の評価対象件数

(2) 審決 100 件中の評価対象文数

審判種別	四法			対象文数
	特許	商標	意匠	
無効	9,396	354	776	10,526
不服	3,018	179	90	3,287
判定	504	89		593
訂正	344			344
取消		308		308
異議	129	99		228
合計	13,391	1,029	866	15,286

表 3.1-2 四法毎／審判種別毎の評価対象文数

3.1.1 審決の機械翻訳文の人手評価の結果

(1) 審決全体の人手評価結果

審決 100 件全体の評価数値の割合を下記に示す。

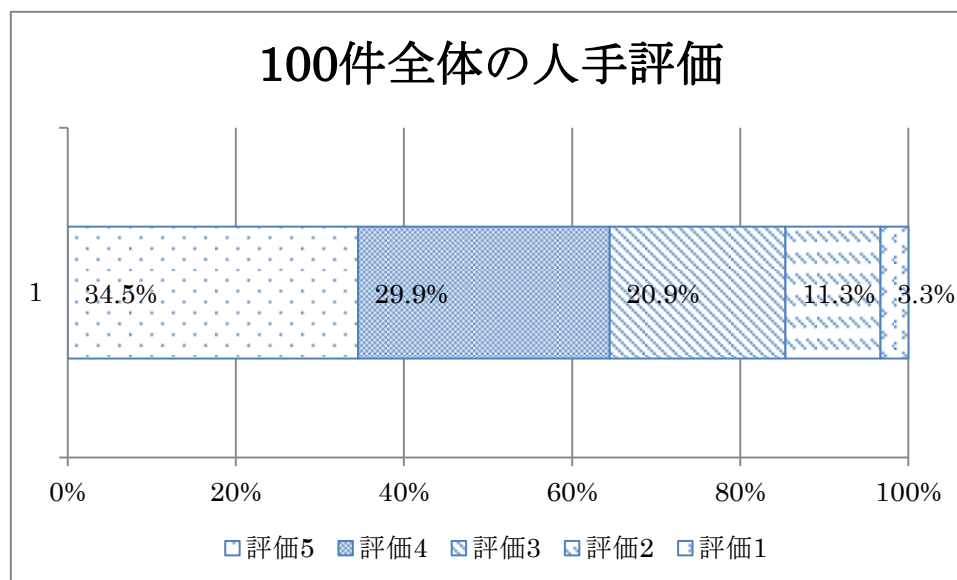


図 3.1.1-1 審決全体の人手評価結果

評価 5 が 34.5%と最も多く、続いて評価 4 (29.9%)、評価 3 (20.9%)、評価 2 (11.3%)、評価 1 (3.3%) と続いた。全体で見ると評価 4 及び評価 5 の占める割合が 64.4%となり、評価対象文の半分以上が「すべての重要情報が正確に伝達されている」あるいは「ほとんどの重要情報は正確に伝達されている」と評価される結果となった。

(2) 審決の四法毎の人手評価結果

審決 100 件の内訳として特許、商標、意匠毎に割り出した評価結果を下記に示す。

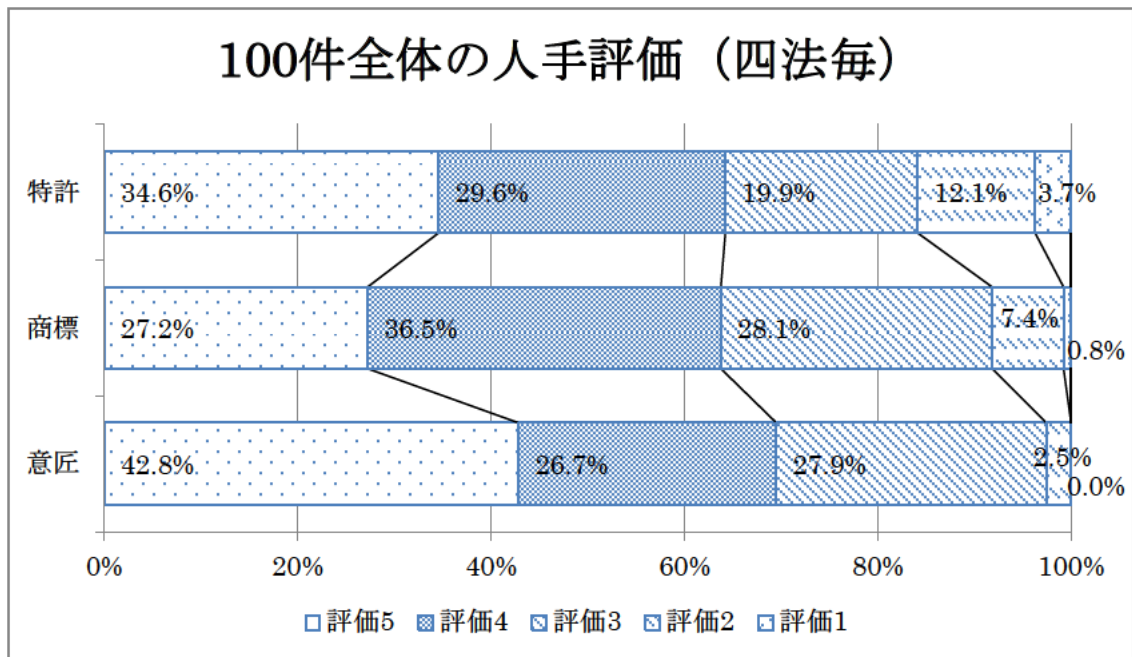


図 3. 1. 1-2 審決の四法毎の人手評価数値の割合

四法	評価 5	評価 4	評価 3	評価 2	評価 1	合計
特許	4, 630	3, 969	2, 664	1, 627	501	13, 391
商標	280	376	289	76	8	1, 029
意匠	371	231	242	22		866
合計	5, 281	4, 576	3, 195	1, 725	509	15, 286

表 3. 1. 1-1 審決の四法毎の人手評価数値別文数

「意匠」は評価 5 及び評価 4 の占める割合も 69.5%と最も高く、機械翻訳として比較的高い精度が出ているとみることができる。

評価 5 の占める割合のみに着目した場合は、「意匠」42.8%、「特許」34.6%、「商標」27.2%の順となり、「商標」の評価の低さが際立つが、評価 5 及び評価 4 の占める割合の合計値まで範囲を広げると、「意匠」69.5%、「特許」64.2%、「商標」63.7%となり、四法において大きな差は生じていないことがわかる。

一方で、評価 1 及び評価 2 の低評価の機械翻訳文に焦点を当てた場合は、評価の低い方から、「特許」15.8%、「商標」8.2%、「意匠」2.5%の順となり、「特許」の精度に問題があることが明らかとなった。

評価対象文の規模をみると全体の 15,286 文のうち大多数の 87.6%にあたる 13,391 文が「特許」であるため、「特許」の評価結果がそのまま審決文全体の人手評価の結果を表していると言える。

以上から、「特許」の審決文について評価を下げる要素及び傾向を分析して改善を図る必

要があり、「特許」を改善することで審決文全体の改善につながるものとみることができる結果となった。

(3) 審決の審判種別毎の人手評価結果

評価対象文を審判種別（無効/不服/判定/訂正/取消/異議）に分けた評価数値結果を下記に示す。

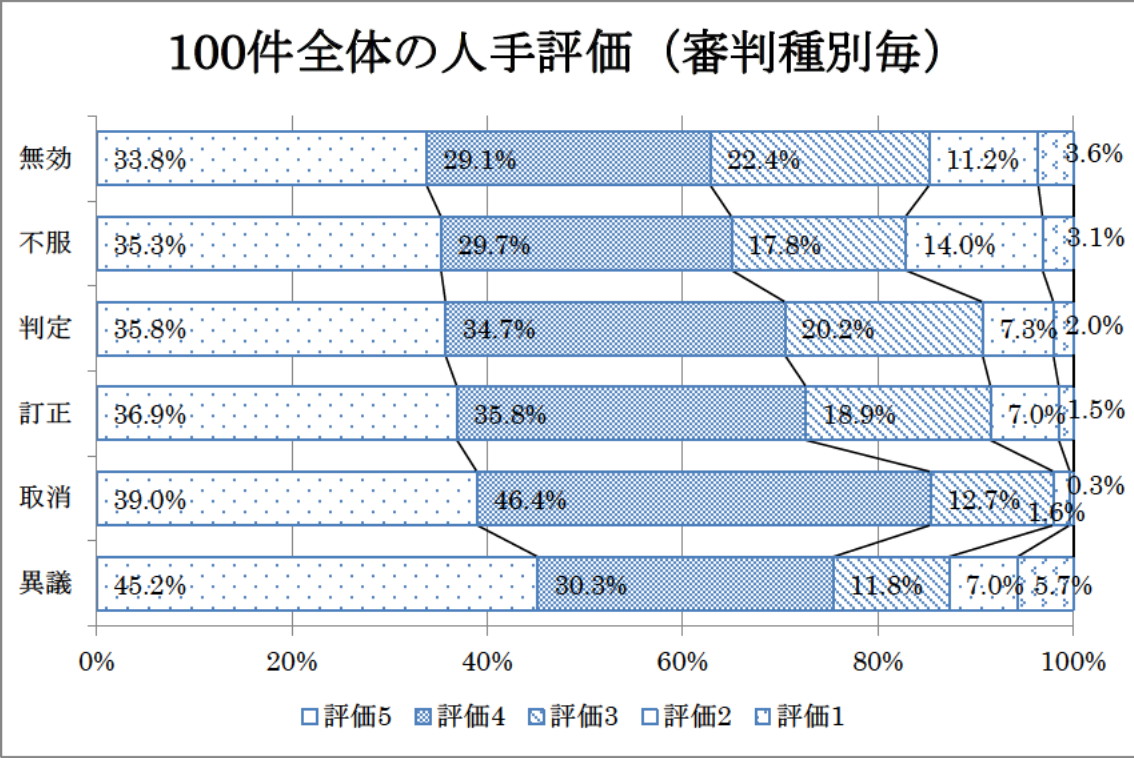


図 3.1.1-3 審決の審判種別毎の人手評価数値の割合

審判種別	評価 5	評価 4	評価 3	評価 2	評価 1	合計
無効	3, 559	3, 058	2, 358	1, 176	375	10526
不服	1, 160	977	586	461	103	3287
判定	212	206	120	43	12	593
訂正	127	123	65	24	5	344
取消	120	143	39	5	1	308
異議	103	69	27	16	13	228
合計	5, 281	4, 576	3, 195	1, 725	509	15286

表 3.1.1-2 審決の審判種別毎の人手評価数値別文数

6つの審判種別毎に評価の比較を行ったところ、評価4、5の合計の占める割合は、評価が高い順に、「取消」が85.4%、「異議」が75.5%、「訂正」が72.7%、「判定」が70.5%、「不

服」が 65.0%、「無効」が 62.9%という結果になった。

一方で評価 1 及び評価 2 が占める割合が大きく、低評価と判定された種別は、評価が低い順に、「不服」が 17.2%、「無効」が 14.8%、「異議」が 12.7%という結果になった。

6 つの審判種別のうち、高評価と低評価のいずれにおいても突出した値となった種別はなく、すべての種別において評価 5 と評価 4 の占める割合が過半数を上回る結果となっているが、評価 1～評価 3 までの合計値でみると評価の低い順に「無効」が 37.2%、「不服」が 34.9%となり、「無効」及び「不服」の精度が比較的低いものであることがわかる。

以上から、審判種別毎でみた場合に有意な差は生じていないものの、「無効」及び「不服」については評価が比較的低く、改善を図る余地があることがわかる結果となった。

(4) 調査対象範囲における審判種別の割合

今回の調査対象である 100 件の審判種別（無効/不服/判定/訂正/取消/異議）毎の件数と文数を以下の表 3.1.1-3 及び表 3.1.1-4 に示す。表 3.1.1-4 は全体に占める件数及び文数の割合を視覚的に把握できるよう、数値の大きさを色付きの帯の長さで示したものである。

	件数				文数			
	特許	商標	意匠	総計	特許	商標	意匠	総計
無効	30	3	2	35	9,396	354	776	10,526
不服	40	4	2	46	3,018	179	90	3,287
判定	5	1		6	504	89		593
訂正	6			6	344			344
取消		3		3		308		308
異議	2	2		4	129	99		228
合計	83	13	4	100	13,391	1,029	866	15,286

表 3.1.1-3 審決の審判種別毎の件数及び文数

	件数の割合			文数の割合		
	特許	商標	意匠	特許	商標	意匠
無効	30.0%	3.0%	2.0%	61.5%	2.3%	5.1%
不服	40.0%	4.0%	2.0%	19.7%	1.2%	0.6%
判定	5.0%	1.0%	0.0%	3.3%	0.6%	0.0%
訂正	6.0%	0.0%	0.0%	2.3%	0.0%	0.0%
取消	0.0%	3.0%	0.0%	0.0%	2.0%	0.0%
異議	2.0%	2.0%	0.0%	0.8%	0.6%	0.0%

表 3.1.1-4 審決の審判種別毎の件数及び文数の割合

件数では「特許」の「不服」が 40%と一番多いが、文数で見ると「特許」の「無効」が 61.5%と大多数を占めていた。これは、原文の文字数が「無効」全体で 1,125,887 字であるところ、「不服」では 518,545 字と「無効」の半分程度の分量になることによるものである。

このように、今回の調査対象の大半が「無効」と「不服」であるため、審決文全体への影響の大きさを考慮し、以下において「無効」と「不服」に特化してそれぞれの傾向を調査することとした。

(5) 審決の「無効」における四法毎の人手評価結果

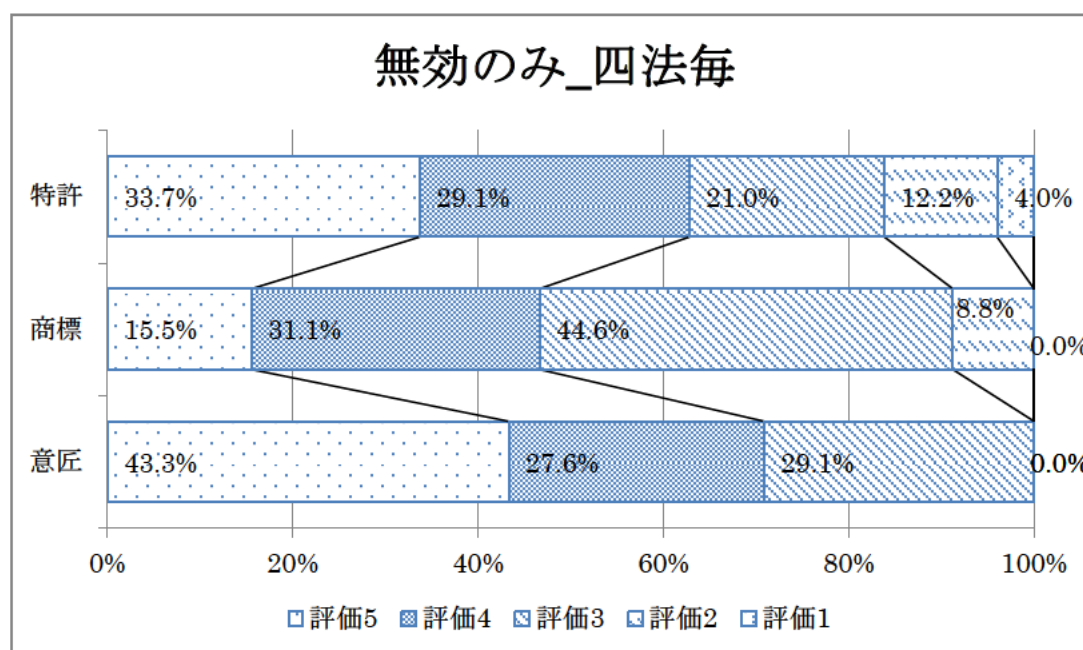


図 3.1.1-4 審決の「無効」における四法毎の人手評価数値の割合

四法	評価 5	評価 4	評価 3	評価 2	評価 1	合計
特許	3,168	2,734	1,974	1,145	375	9,396
商標	55	110	158	31		354
意匠	336	214	226			776
合計	3,559	3,058	2,358	1,176	375	10,526

表 3.1.1-5 審決の「無効」における四法毎人手評価数値別の文数

図 3.1.1-4 に示す「無効」に特化した場合の人手評価数値の割合に着目すると、「意匠」は評価 5 及び評価 4 の占める割合が 70.9%と最も高く、機械翻訳として比較的高い精度が出ているとみることができる。

反対に「商標」は、評価 5 及び評価 4 の占める割合が平均 60.1%に対して 46.6%、評価 1～3 の占める割合が平均 39.9%に対して 53.4%と、最も評価が低い結果となった。

以上から、「無効」については、特に「商標」について審決文の訳質を分析して改善を図る必要があると言える。

(6) 審決の「不服」における四法毎の人手評価結果

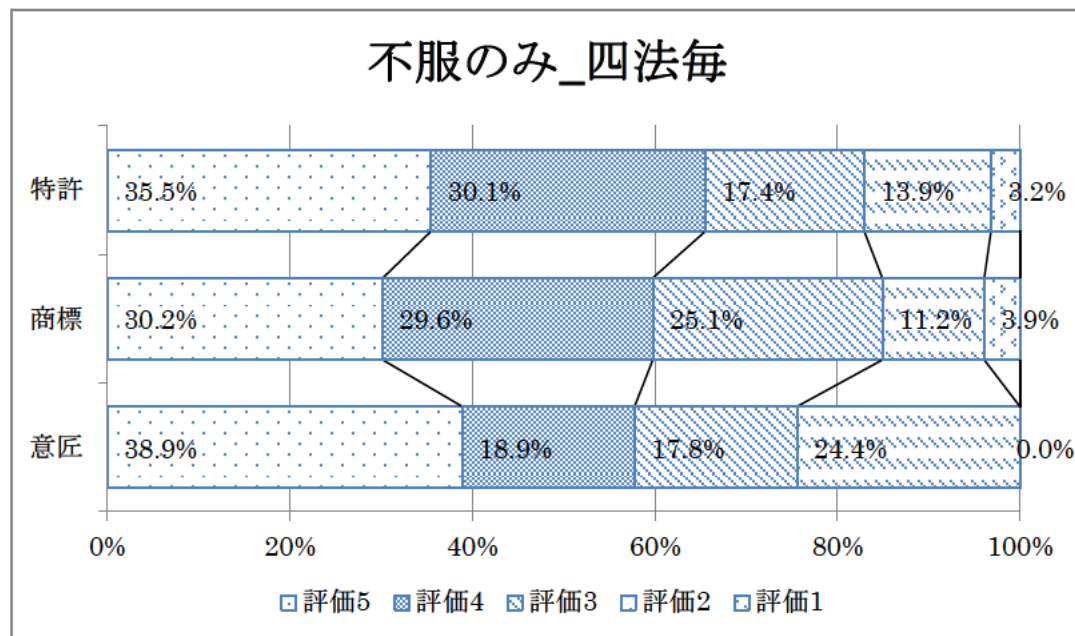


図 3. 1. 1-5 審決の「不服」における四法毎の人手評価数値の割合

四法	評価 5	評価 4	評価 3	評価 2	評価 1	合計
特許	1, 071	907	525	419	96	3, 018
商標	54	53	45	20	7	179
意匠	35	17	16	22		90
合計	1, 160	977	586	461	103	3, 287

表 3. 1. 1-6 審決の「不服」における四法毎人手評価数値別の文数

図 3. 1. 1-5 に示す「不服」に特化した場合の人手評価数値の割合に着目すると、「特許」は評価 5 及び評価 4 の占める割合が 65. 6%と最も高く、機械翻訳として比較的高い精度が出ているとみることができる。

反対に「意匠」は、評価 1 及び評価 2 の占める割合が 24. 4%と突出しており、評価 1～3 が占める割合も 42. 2%と過半数に近い値になることから最も評価が低いことがわかる。

以上から、「不服」については、特に「意匠」について審決文の訳質を分析して改善を図る必要があると言える。

(6) 原文文字数と人手評価数値の関係

一文中の原文日本語文字数と人手評価数値を比較することで、原文の長さが機械翻訳文の精度にどう影響しているかを調査・分析した。

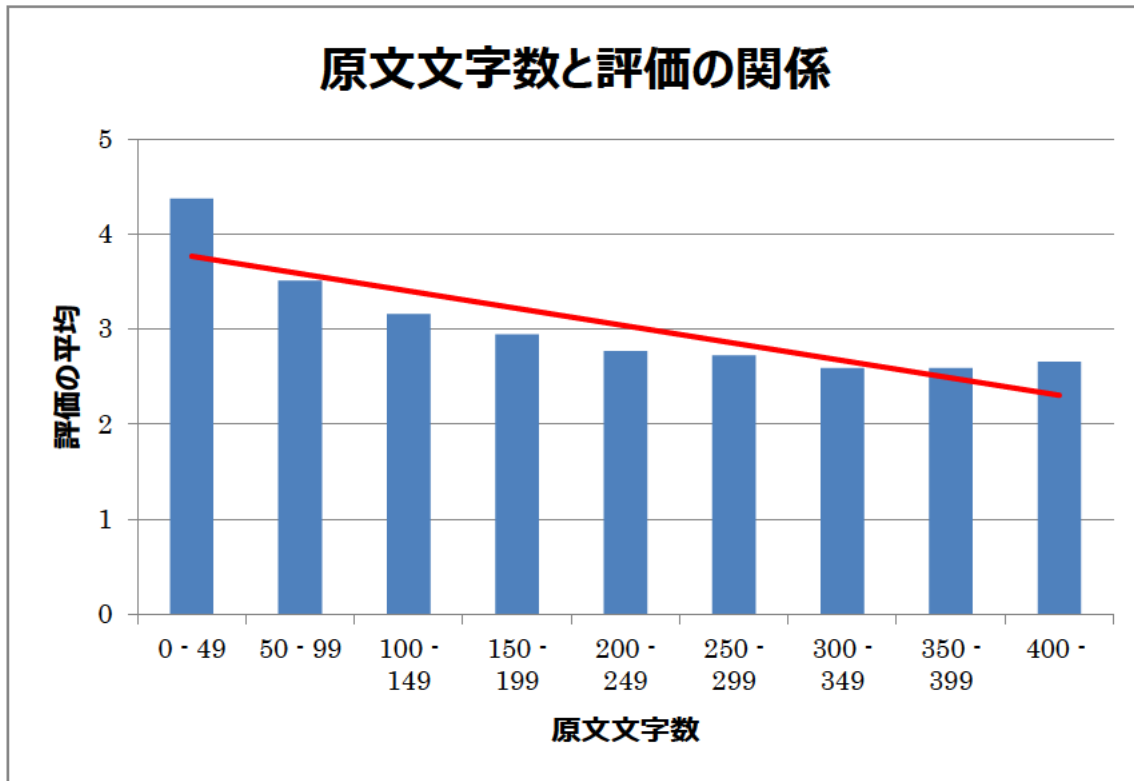


図 3.1.1-6 原文文字数と人手評価数値の関係 (評価平均)

原文文字数	評価の平均	評価 5	評価 4	評価 3	評価 2	評価 1	合計
0 - 49	4.37	4,580	2,305	820	272	77	8,054
50 - 99	3.51	448	1,182	829	431	67	2,957
100 - 149	3.16	170	565	641	416	96	1,888
150 - 199	2.94	58	265	366	252	95	1,036
200 - 249	2.77	13	112	215	150	56	546
250 - 299	2.72	6	73	122	96	41	338
300 - 349	2.59	4	25	64	42	29	164
350 - 399	2.59	1	14	32	22	14	83
400 -	2.66	1	35	106	44	34	220
合計		5,281	4,576	3,195	1,725	509	15,286

表 3.1.1-7 原文文字数と人手評価数値の関係 (文数)

表 3.1.1-7 に示したように、原文一文あたりの文字数に応じて 9 グループに分けてグループ毎の評価の平均を算出し、人手評価数値との関係性を調査した。

評価の平均値は、「0-49 字」が 4.37 と突出して高く、次いで「50-99 字」が 3.51、「100-149 字」が 3.16 という順になった。

以上から図 3.1.1-6 内の近似直線が示すように、文字量が多くなるに従って人手評価の数値も低くなる傾向があることがわかる。

グループ毎の評価の内訳は下記の通りである。

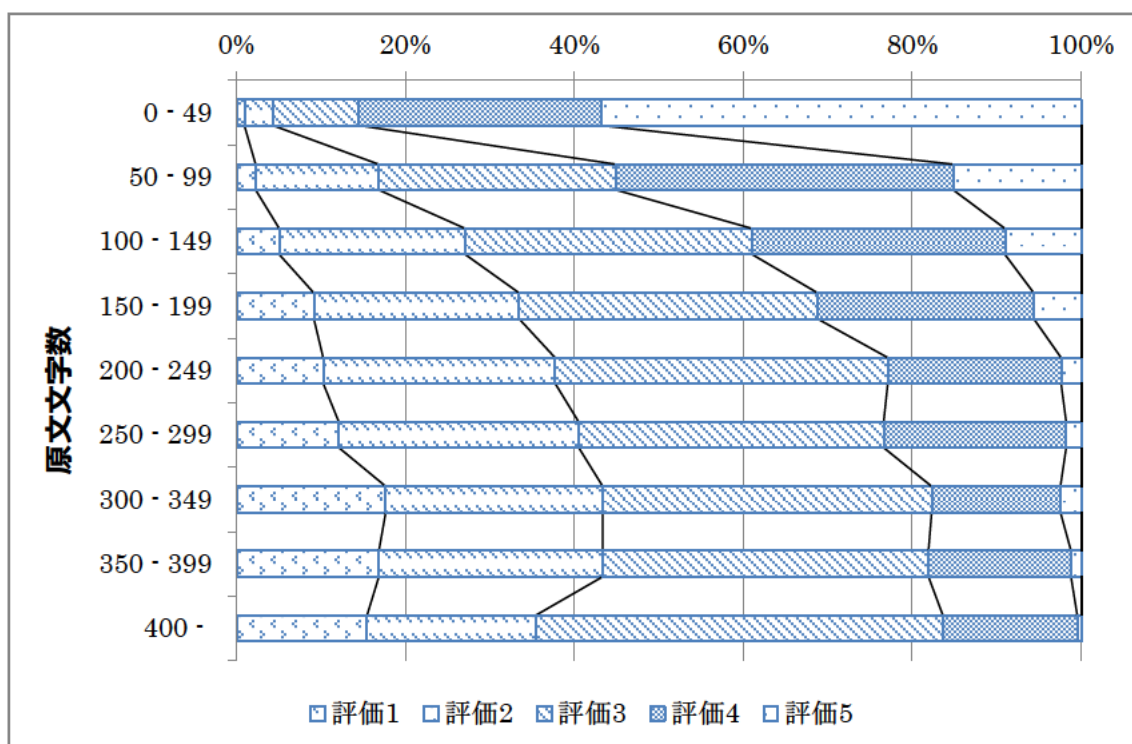


図 3.1.1-7 原文文字数と人手評価数値の関係（グループ別評価）

(7) 文中の引用文（かぎ括弧）の数と人手評価数値の関係

一文中におけるかぎ括弧の個数と人手評価数値を比較することで、一文中における引用箇所が多さが機械翻訳文の精度にどのような影響を与えるかを調査・分析した。

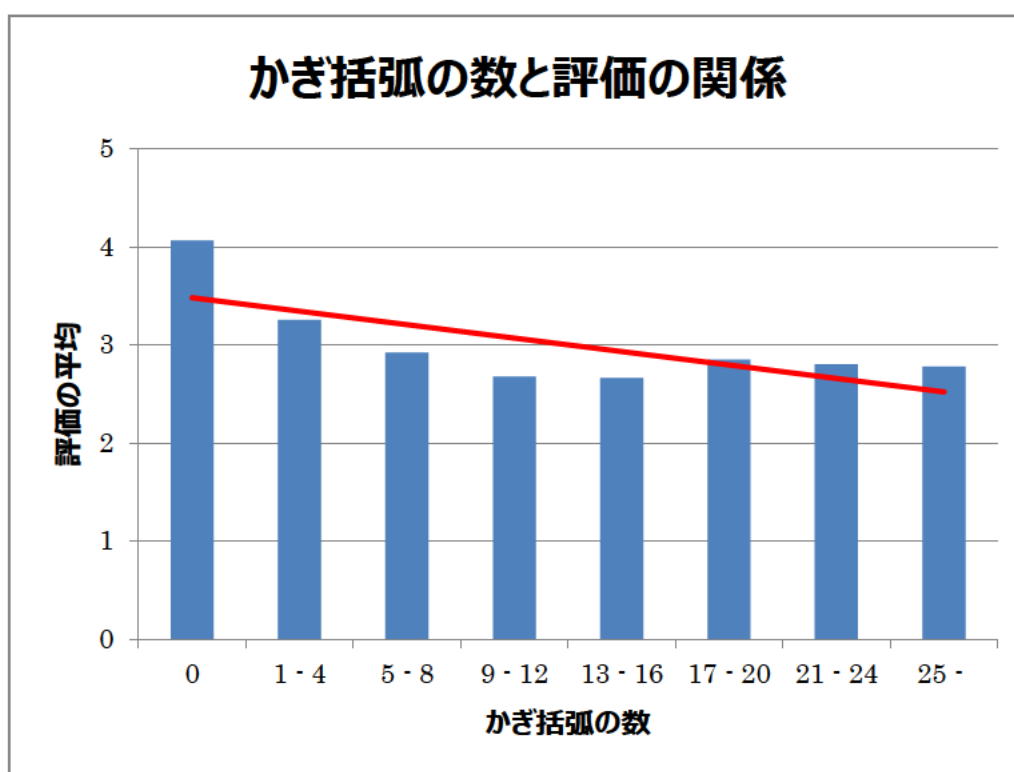


図 3.1.1-8 文中のかぎ括弧の数と人手評価数値の関係（評価平均）

かぎ括弧数	評価の平均	評価 5	評価 4	評価 3	評価 2	評価 1	合計
0	4.06	4,867	3,316	1,866	882	180	11,111
1 - 4	3.26	360	949	949	549	177	2,984
5 - 8	2.92	38	214	256	197	76	781
9 - 12	2.68	10	52	68	60	44	234
13 - 16	2.66	3	17	27	21	15	83
17 - 20	2.85	2	20	16	11	11	60
21 - 24	2.80	0	5	4	4	2	15
25 -	2.78	1	3	9	1	4	18
		5,281	4,576	3,195	1,725	509	15,286

表 3.1.1-8 文中のかぎ括弧の数と人手評価数値の関係

原文一文あたりに出現する始めかぎ括弧（「）と終わるかぎ括弧（」）をそれぞれ 1 個として計上し、表 3.1.1-7 に示したように、引用かぎ括弧の数に応じて 8 グループに分けてグループ毎の評価の平均を算出し、人手評価数値との関係性を調査した。

評価の平均値は、「0 個」が 4.06 と最も高く、次いで「1-4 個」が 3.26、「5-8 個」が 2.92 となり、図 3.1.1-8 内の近似直線が示すように、かぎ括弧の数が示す引用文の出現回数が

多くなるに従って、人手評価の数値も低くなる傾向があることがわかる。

文中のかぎ括弧の数と人手評価数値の関係におけるグループ毎の評価の内訳は下記の通りである。

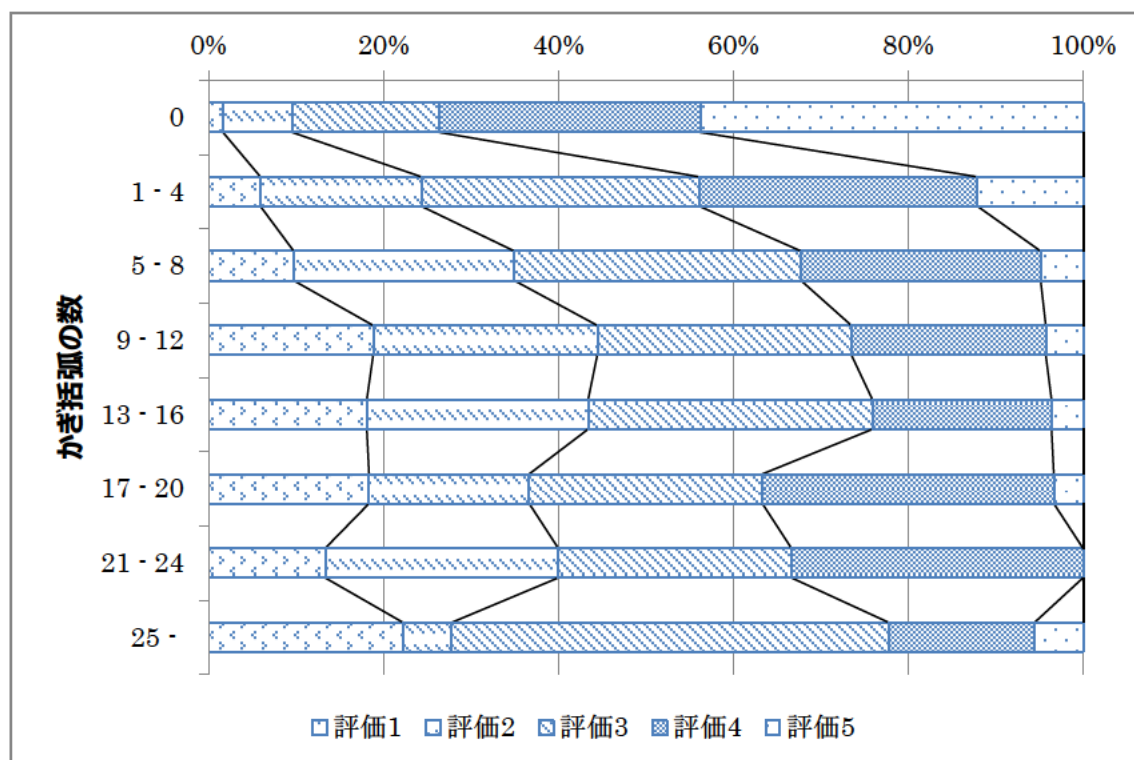


図 3. 1. 1-9 かぎ括弧数と人手評価数値の関係（グループ別評価）

(8) 文中の引用文（かぎ括弧）内の文字数と人手評価数値の関係

一文中におけるかぎ括弧の個数と人手評価数値との比較に続き、一文中におけるかぎ括弧の中での文字数と人手評価数値との関係も調査し、一文中における引用文の長さが機械翻訳文の精度にどのような影響を与えるかを分析した。

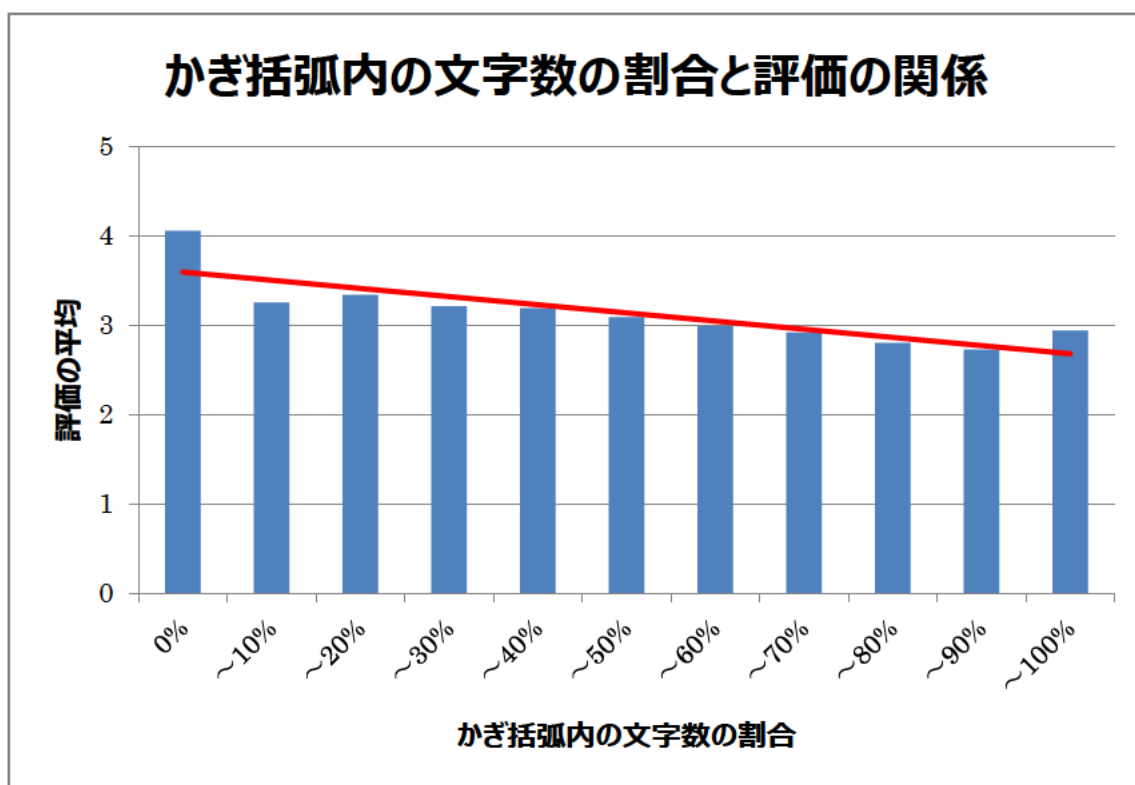


図 3. 1. 1-10 文中のかぎ括弧の文字数の割合と人手評価数値の関係（評価平均）

「」内の文字数の割合	評価の平均	評価 5	評価 4	評価 3	評価 2	評価 1	
0	4.06	4,887	3,339	1,896	905	184	11,211
~10%	3.25	79	203	185	123	39	629
~20%	3.34	115	312	226	133	60	846
~30%	3.21	63	188	181	106	42	580
~40%	3.19	43	140	143	89	26	441
~50%	3.09	33	121	119	91	30	394
~60%	2.99	26	92	97	74	36	325
~70%	2.92	11	68	102	61	25	267
~80%	2.81	15	59	106	80	33	293
~90%	2.73	5	34	86	47	25	197
~100%	2.94	4	20	54	16	9	103
合計		5,281	4,576	3,195	1,725	509	15,286

表 3. 1. 1-9 文中のかぎ括弧の文字数の割合と人手評価数値の関係

原文一文あたりに出現するかぎ括弧の範囲内の文字数が一文全体に占める割合を計上し、表 3.1.1-8 に示したように、引用かぎ括弧内の文字数の占める割合に応じて 11 グループに分けてグループ毎の評価の平均を算出して人手評価数値との関係性を調査した。

評価の平均値は、「0%」が 4.06 と突出して高く、「0%」以外の 10 グループでは割合の値が大きくなるに従って評価数値もゆるやかに下がっていく結果となった。

最後の「～100%」のグループの評価の値が最も低い値とはならず「～70%」の評価値をも上回る結果となったが、これは、一文のほぼ全体が引用かぎ括弧である文では文の構造が単純なものであることが多く、構文解析が容易であるために機械翻訳文の精度が高くなったものと考えられることができる。

文中のかぎ括弧の文字数の割合と人手評価数値の関係におけるグループ毎の評価の内訳は下記の通りである。

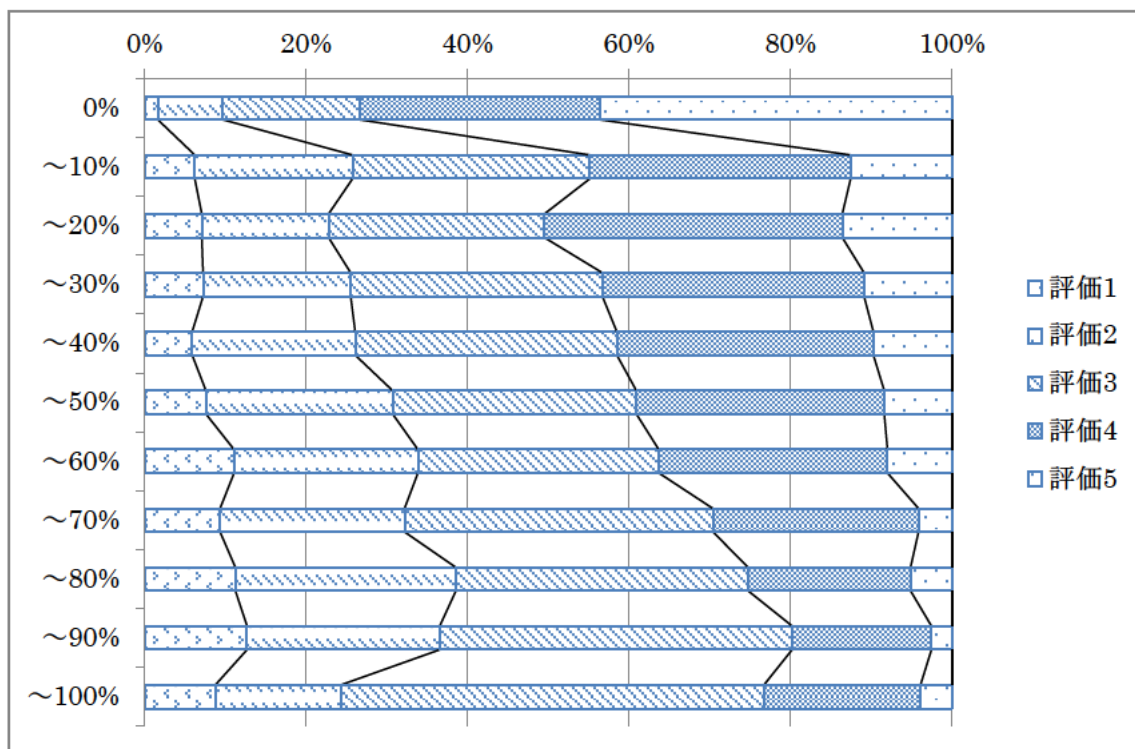


図 3.1.1-11 かぎ括弧数の文字数の割合と人手評価数値の関係 (グループ別評価)

(9) 機械翻訳文における辞書カバー率と人手評価数値の関係

特許庁より貸与された既存辞書⁴が機械翻訳文に及ぼす影響について調査した。

具体的には、機械翻訳文一文における既存辞書に一致する訳語の割合(一文内の全英文字数に対する、既存辞書に一致する文字数の割合)を「辞書カバー率」として数値化し、「辞

⁴ システム辞書を除く約 18.6 万件の辞書データであり、ユーザ辞書、メモリ辞書、パターン辞書を含む。

「辞書カバー率」の数値と人手評価数値を比較することで、既存辞書の適用状況が機械翻訳文の精度にどのような影響を与えるかを調査・分析した。

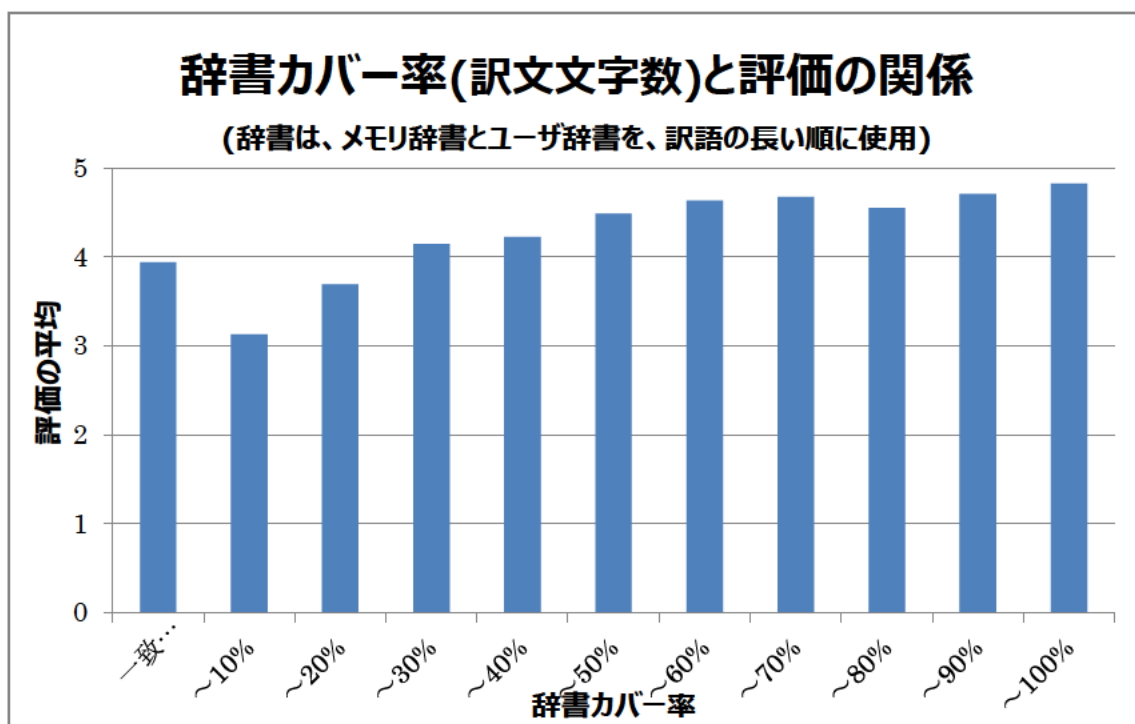


図 3.1.1-12 辞書カバー率と人手評価数値の関係

辞書 カバー率	評価の 平均	原文文字 数の平均	評価 5	評価 4	評価 3	評価 2	評価 1	合計
一致なし	3.94	57.74	3,806	2,721	1,672	950	274	9,423
~10%	3.13	169.25	245	957	1,067	631	207	3,107
~20%	3.69	73.40	260	413	298	117	23	1,111
~30%	4.14	36.24	181	161	75	17	3	437
~40%	4.22	26.37	92	110	23	7	1	233
~50%	4.49	16.95	91	61	9	0	1	162
~60%	4.63	10.62	97	21	12	1	0	131
~70%	4.68	9.89	77	37	0	0	0	114
~80%	4.55	7.82	144	51	22	1	0	218
~90%	4.71	10.26	136	23	12	1	0	172
~100%	4.83	10.74	152	21	5	0	0	178
合計			5,281	4,576	3,195	1,725	509	15,286

表 3.1.1-10 辞書カバー率と人手評価数値の関係 (文数)

既存辞書の原語とその対訳の両方が一文中の原文と機械翻訳文の両方に出現する場合に、既存辞書データと一致したものとみなし、機械翻訳文の一文において、その辞書と一致した範囲の文字数が一文全体の文字数に占める割合を算出して辞書カバー率とした。

次に、算出した辞書カバー率を表 3.1.1-9 に示したように 11 グループに分けてグループ毎の評価の平均を算出し、人手評価数値との関係性を調査した。

評価の平均値は、「～10%」が最も低く、「～10%」以降のグループでは「～20%」が 3.69、「～30%」が 4.14 と徐々に数値が高くなる結果となった。

このように「～10%」以降のグループでは、辞書カバー率が高くなるに従って人手評価も高くなることが明らかとなり、既存辞書によって機械翻訳文の精度が向上することを示す結果となった。

「一致なし」の評価の平均値が高くなる点についてさらに考察するために、各グループの一文あたりの原文文字数の平均値を調査した。その結果、上記の表 3.1.1-9 に示すように、「～10%」～「～30%」のグループに比べて「一致なし」のグループの原文文字数が比較的少ないことがわかった。この点から、「一致なし」の評価の平均値が高い理由は一文の平均文字数が少ないことが影響していると言っていることができる。

以上のように原文の文字数が比較的少ない場合には、既存辞書の適用がなくともそれなりの精度が得られる場合があるものの、原文の文字数が一定程度以上ある文章の場合には、辞書カバー率が高くなるに従って機械翻訳文の精度も上がり、機械翻訳文の品質向上において辞書の有効性が非常に高いことが明確となった。

3.1.2 審決の機械翻訳文の分析コメント

調査対象の審決より更に具体的な事例を取り上げて、J-PlatPat の機械翻訳に特徴的な訳出の結果を示す。

(1) 技術用語の訳出

人手 評価	原文	上記 (イ) (ウ) の記載、図 1 ～ 3 の記載及び技術常識を併せてみると、コイル 1 1 がケース 2 の内側に収納され、ソレノイドが、コイル 1 1 の内筒部に延出するセンタポスト部材、コイル 1 1 の内筒部に位置し往復動可能なプランジャが配置されるプランジャ室、及びプランジャに接続されバルブ部の弁体の開閉動作を可能とするロッドを備え、プレート 1 4 がケース 2 の一方の開口端部とプランジャ室との間に配置されることが理解できる。
4	MT	If a description of the above-mentioned (b) (c), Fig.1 - a description of three, and a technical common sense are combined, A center-pin component to which the coil 11 is stored inside the case 2, and a solenoid extends to an internal cylinder part of the coil 11, He has a

		plunger chamber which is placed at an internal cylinder part of the coil 11 and by which a reciprocable plunger is arranged, and a rod which is connected to a plunger and makes possible a switching action of a valve body of a valve part, and can understand that the plate 14 is arranged between one open end of the case 2, and a plunger chamber.
--	--	---

訂正 2016-390104

「ソレノイド」「a solenoid」、「コイル 1 1」「the coil 11」、「延出する」「extends」、「内筒部」「an internal cylinder part」、「往復動可能な」「reciprocable」、「プランジャ室」「plunger chamber」、「弁体」「a valve body」、「ロッド」「a rod」など、技術用語がいずれも適訳で訳出されている。さらに、上記の文以外でも本審決中では同一の技術用語が何度も繰り返されるが、目立った訳揺れもなく全体的に技術用語の訳語が統一されていた。

こうした用語の適切な訳出結果は辞書登録の成果であり、技術用語や専門用語を適切かつ訳揺れを抑えて訳出できる点は RBMT 特有の強みといえる。

(2)引用文が多い原文

人手 評価	原文	また、本願明細書中には、「臨床的に定められた再発」のない患者（「反応者」に該当）について、「ベースライン」と「2ヶ月」の時点で「IL-17 (A)、TNF- α 、IL-2 及び IFN- γ レベルの増加」がみられ、「臨床的に定められた再発」をした患者（「非反応者」に該当）について、「ベースライン」と「2ヶ月」の時点で「同一の炎症誘発性サイトカインが相当に低レベルである」、といったことが記載されているが（摘記 J）、同 J 中で引用されている図 2 の A、B を実際にみると、上の 3）（ア）～（エ）以外にも、以下の事項が読み取れる。
1	MT	[in Description of this application] [patient / (it corresponds to a "reaction person") / without "a recurrence defined clinically"] "The time of a "baseline" and "two months" IL-17 (A), TNF-alpha, [patient / (it corresponds to "a non-reacting person") / who the increase in IL-2 and IFN-gamma levels" was seen, and did "the recurrence defined clinically"] Although "the same proinflammatory cytokine is a low fairly" is described as of a "baseline" and "two months" (summary J), if A of the Fig.2 currently cited in the J and B are actually seen, the following matters can be read besides upper 3 (a)s - (d).

無効 2016-800080

かぎ括弧（「」）で括られた明細書からの引用文を原文中に多数含む文である。構文解析に失敗し、角括弧（[]）が多く残っている。

前述の通り、括弧で括られた引用文が多いほど機械翻訳の精度は低下するが、特に上記の「臨床的に定められた再発」のない患者」や、「ベースライン」と「2ヶ月」の時点で「同一の炎症誘発性サイトカインが相当に低レベルである」、といったことが記載・・・」といった、かぎ括弧で囲われた部分が独立せず地の文に続くものや、引用部分同士が複雑に入れ込まれた関係の文、引用部分が名詞ではない文などは構文解析が適切にされず精度が低下する傾向にある。

RBMT では、長文や引用文が多数含まれる文は不適切な構文解析により、機械翻訳精度が芳しくないことが多いが、構文解析の誤りにより評価数値の低い文であっても、専門用語・技術用語の訳出という観点で見ると高く評価できる文も多い。

また文献全体における用語統一や、用語単位での誤訳・訳抜けの数という点においても高く評価できる。

3.2 審決の機械翻訳文の自動評価結果

本調査では、審決 100 件を対象に、審決及び判決特有の表現を含む文を選定して BLEU と RIBES を用いた自動評価を行った。その評価結果について以下に説明する。

BLEU 及び RIBES による自動評価と人手評価との相関性

BLEU 及び RIBES の自動評価を審決 100 件に適用するにあたり、BLEU、RIBES のスコアと、人手評価の値との相関関係を事前に調査した。

具体的には、人手評価を 1～5 のグループに分け、各人手評価グループの中を BLEU 及び RIBES の値の範囲によってさらに「0.1-0.2」、「0.2-0.3」等と階層化して各区分に含まれる審決文の数を集計した。

集計結果を以下の図に示す。

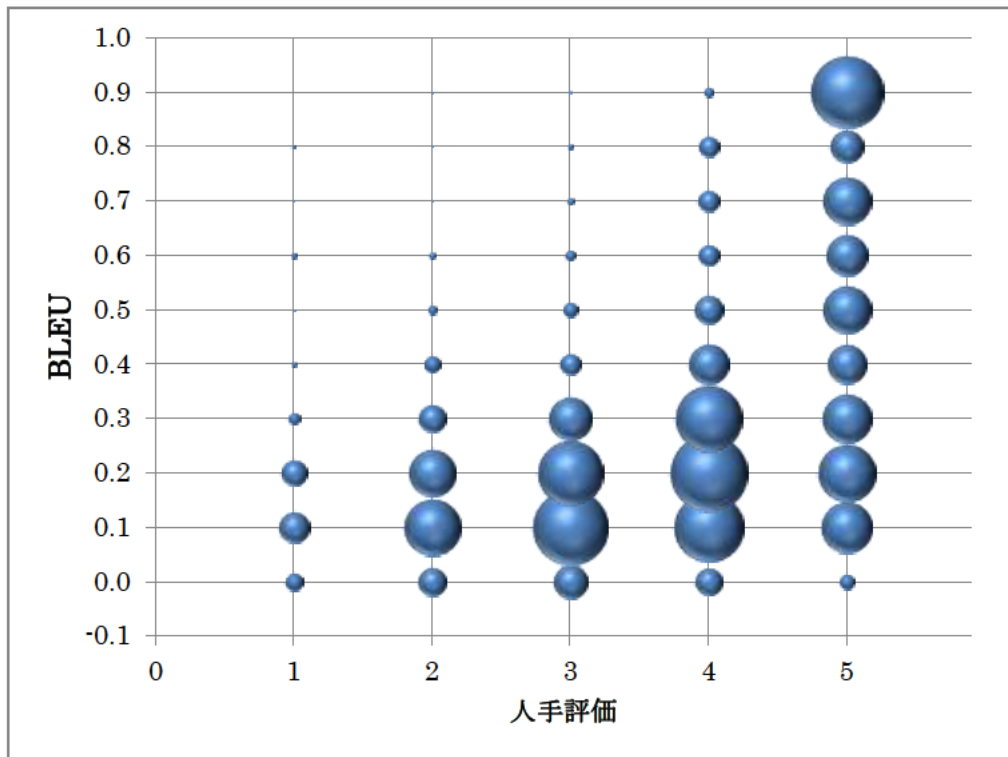


図 3.2-1 BLEU 値と人手評価との関係

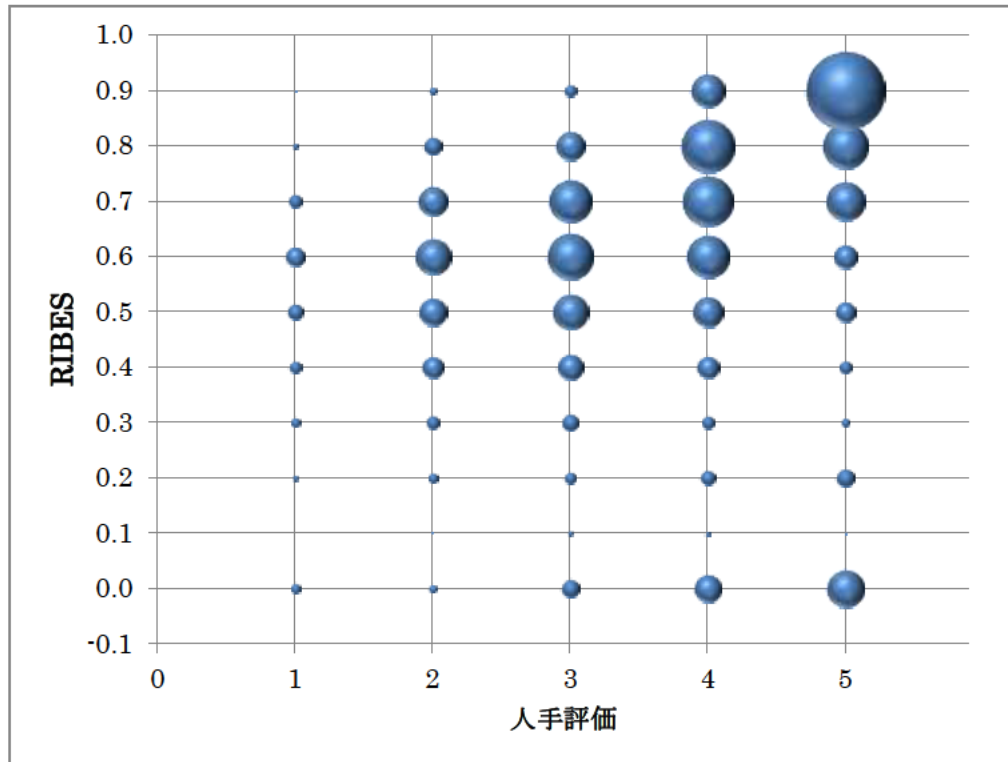


図 3.2-2 RIBES 値と人手評価との関係

自動評価と人手評価との値の相関性を調べたところ、上記の 2 つの図に示すように BLEU と RIBES の双方において人手評価との高い相関関係が示された。この点から明らかなように、自動評価と人手評価との間には高い相関関係があり、自動評価を、人手評価を補完するものとして活用することが可能であると結論することができる。

この結論を踏まえ、以下において、実際に BLEU、RIBES の自動評価手法を用いて審決文 100 件の品質評価を行うものとする。

3.2.1 審決の機械翻訳文の自動評価結果

調査対象である審決 100 件の全文のうち、技術的な説明文を除いた 15,286 文を対象に BLEU 及び RIBES を用いた自動評価を行った。四法毎、審判種別毎の件数及び文数の内訳は「3.1 審決の機械翻訳文の人手評価」に記載した通りである。

(1) 審決全体の自動評価結果

審決 100 件全体の BLEU 及び RIBES それぞれの評価数値の割合を下記に示す。

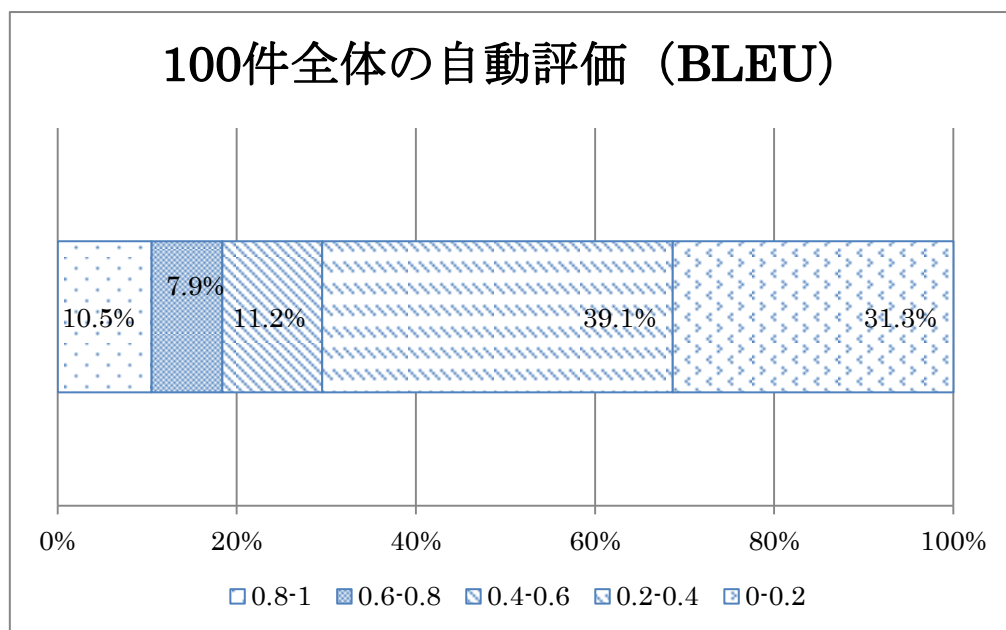


図 3.2.1-1 審決全体の自動評価結果 (BLEU)

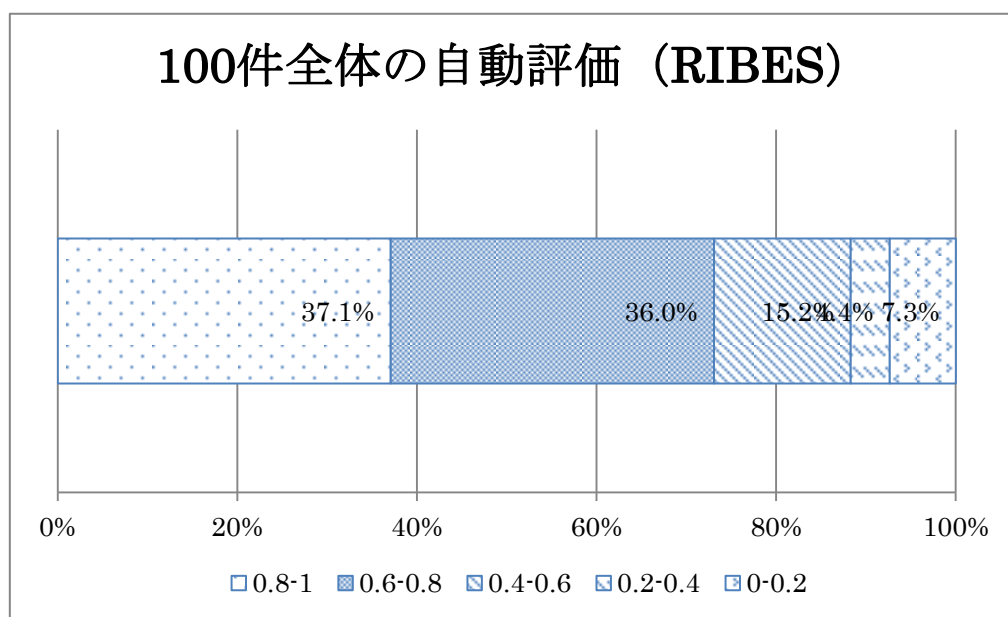


図 3. 2. 1-2 審決全体の自動評価結果 (RIBES)

評価 5 及び評価 4 の占める割合が BLEU では 18.4%、RIBES では 73.1%と大きく乖離する結果となった。両者のうち、RIBES の結果が「3.1 審決の機械翻訳文の人手評価」に示した人手評価の値に近いものとなっている。

(2) 審決の四法毎の自動評価結果

審決 100 件の内訳として特許、商標、意匠毎に割り出した評価結果を BLEU と RIBES に分けて下記に示す。

BLEU と RIBES の 0～1 までの評価数値を「0-0.2」、「0.2-0.4」、「0.4-0.6」、「0.6-0.8」、「0.8-1」の 5 つのグループに分けて各グループに占める文数の割合を算出した。

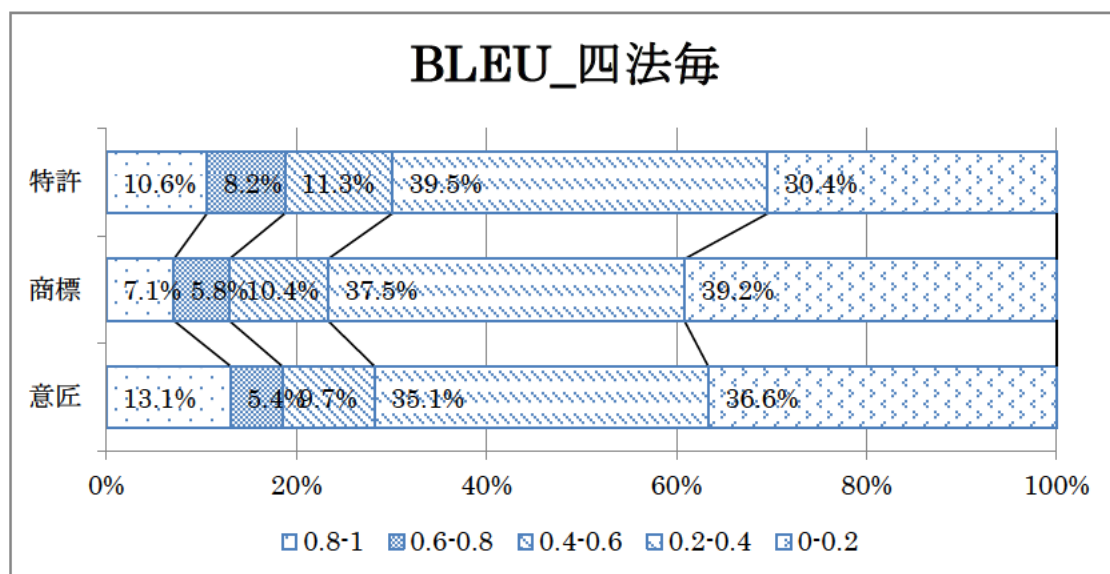


図 3. 2. 1-3 審決の四法毎の自動評価数値の割合 (BLEU)

四法	0.8-1	0.6-0.8	0.4-0.6	0.2-0.4	0-0.2	合計
特許	1,418	1,096	1,515	5,290	4,070	13,389
商標	73	60	107	386	403	1,029
意匠	113	47	84	304	317	865
合計	1,604	1,203	1,706	5,980	4,790	15,283

表 3. 2. 1-1 審決の四法毎の自動評価数値別文数 (BLEU)

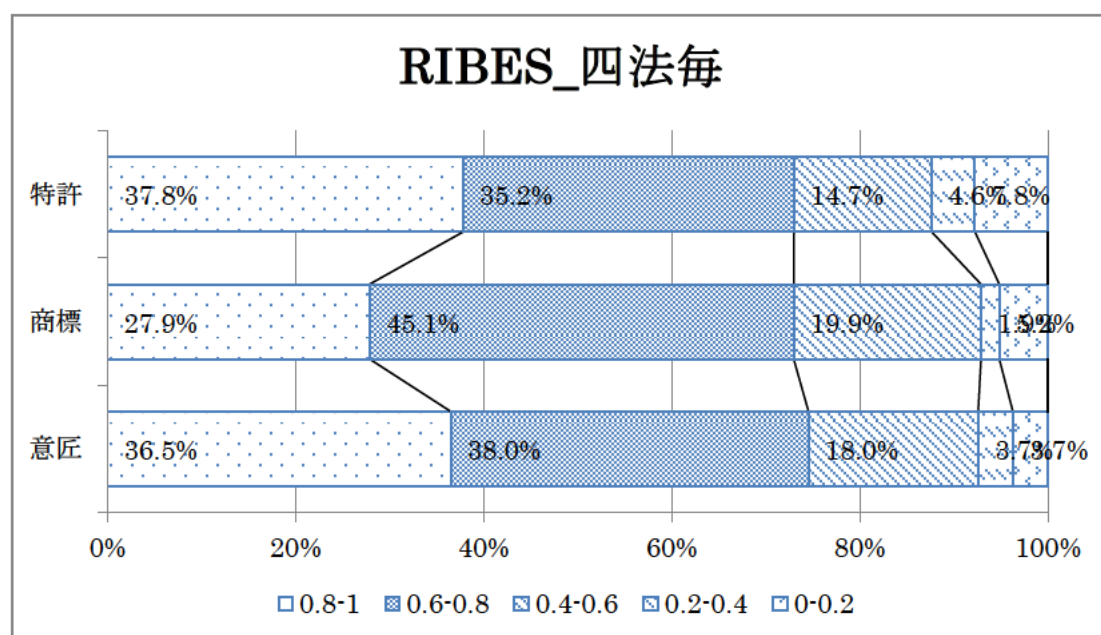


図 3. 2. 1-4 審決の四法毎の自動評価数値の割合 (RIBES)

四法	0.8-1	0.6-0.8	0.4-0.6	0.2-0.4	0-0.2	合計
特許	5,063	4,712	1,962	614	1,038	13,389
商標	287	464	205	20	53	1,029
意匠	316	329	156	32	32	865
合計	5,666	5,505	2,323	666	1,123	15,283

表 3.2.1-2 審決の四法毎の自動評価数値別文数 (RIBES)

全体評価と同様に BLEU と RIBES とで評価の値に大きな差があるものの、「特許」と「意匠」の評価が総じて高く、「商標」の評価が最も低いという結果では共通していた。

(3) 審決の審判種別毎の自動評価結果

評価対象文を審判種別（無効/不服/判定/訂正/取消/異議）に分けた評価数値結果を BLEU と RIBES に分けて下記に示す。

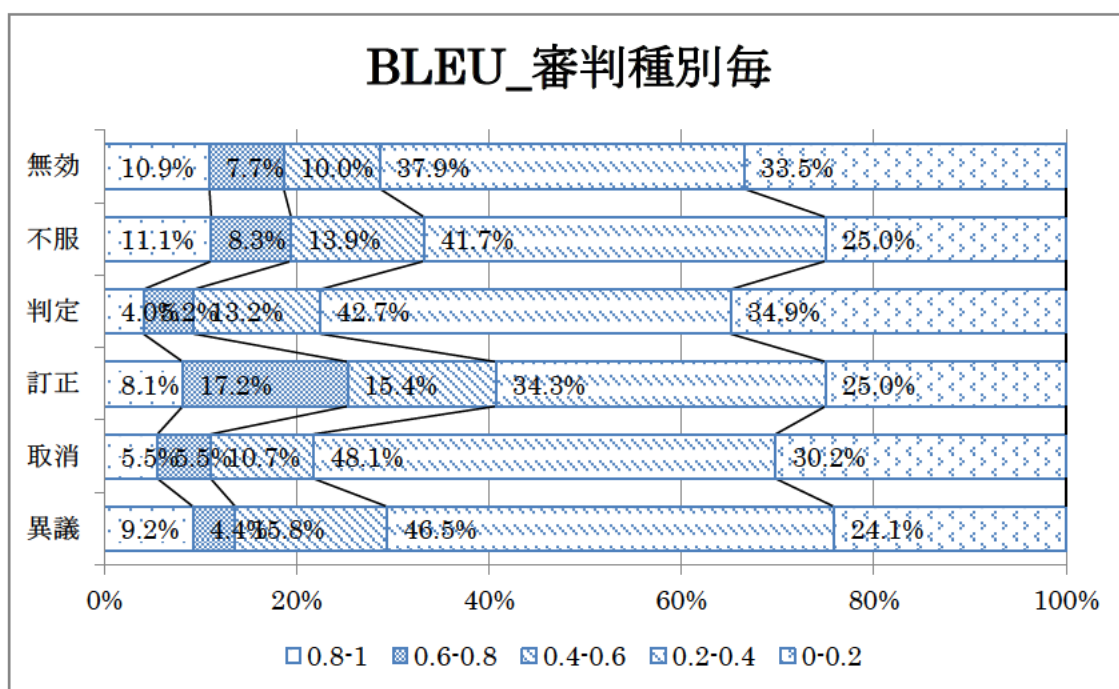


図 3.2.1-5 審決の審判種別毎の自動評価数値の割合 (BLEU)

審判種別	0.8-1	0.6-0.8	0.4-0.6	0.2-0.4	0-0.2	合計
無効	1,149	814	1,051	3,984	3,527	10,525
不服	365	272	455	1,371	822	3,285
判定	24	31	78	253	207	593

訂正	28	59	53	118	86	344
取消	17	17	33	148	93	308
異議	21	10	36	106	55	228
合計	1,604	1,203	1,706	5,980	4,790	15,283

表 3. 2. 1-3 審決の審判種別毎の自動評価数値別文数 (BLEU)

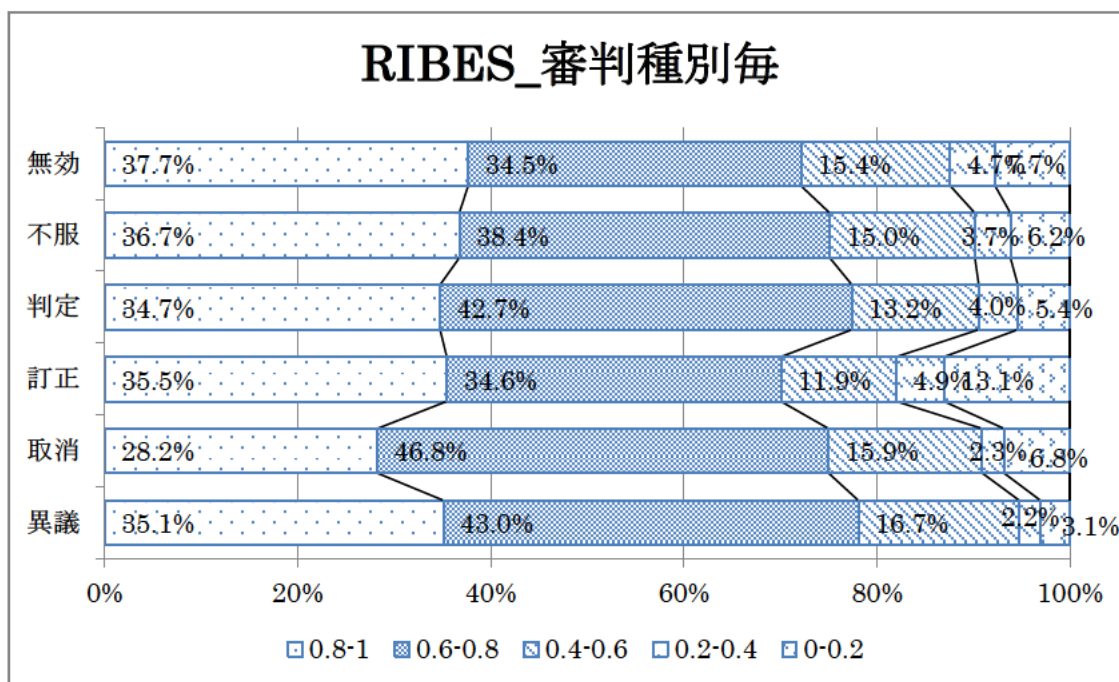


図 3. 2. 1-6 審決の審判種別毎の自動評価数値の割合 (RIBES)

審判種別	0.8-1	0.6-0.8	0.4-0.6	0.2-0.4	0-0.2	合計
無効	3,964	3,630	1,624	492	815	10,525
不服	1,207	1,261	493	121	203	3,285
判定	206	253	78	24	32	593
訂正	122	119	41	17	45	344
取消	87	144	49	7	21	308
異議	80	98	38	5	7	228
合計	5,666	5,505	2,323	666	1,123	15,283

表 3. 2. 1-4 審決の審判種別毎の自動評価数値別文数 (RIBES)

BLEU では「判定」と「取消」の評価の低さが顕著となる一方、RIBES においては審判種別毎で大きな差がない結果となり、RIBES の方が「3.1 審決の機械翻訳文の人手評価」のに近い結果となった。

3.2.2 自動評価結果の分析

自動評価と人手評価の相関に関して、自動評価数値と人手評価数値の差が大きい文を調査した。

BLUE は局所的な単語列の一致数をもとにスコアを算出する一方、BLEU は、基準翻訳文と機械翻訳文に共通して出現する単語の出現順序を順位相関に基づいてスコア化するため、一般に日本語と英語のような語順が大幅に変化する言語間の評価は、RIBES の方が人手評価との相関が高いと言われている。

本章の調査にあたって自動評価は RIBES のスコアを参照した。また、RIBES のスコアは人手評価数値と比較するため、人手評価と同じ 5 段階評価にならして行った。

RIBES は単語の出現順序に基づいて評価をしているため、基準翻訳文と構造や使われている表現が異なると、英文として文法的・意味的に妥当であっても自動評価は低くなる傾向がある。

(1) 単語の出現順序の異なる文

原文	被請求人の主張について
基準翻訳文	Demandee's allegation
MT	the allegations of the demandee

人手評価 5

自動評価 1

RIBES は共通単語の出現する順序に基づき自動評価を行うため、基準翻訳文と機械翻訳文で使用されている単語は同一で文法的・意味的に妥当であっても、単語の出現する順序が異なると自動評価数値が低い。

(2) 使用されている単語が基準翻訳文と機械翻訳文で異なるもの

原文	請求人の提出した証拠方法
基準翻訳文	Evidences Submitted by the Demandant
MT	A Means of proof which a demandant submitted

人手評価 5

自動評価 1

上記の文では「証拠方法」が基準翻訳文では「Evidences」であるが、機械翻訳文では

「Means of proof」として訳出されている。また文のスタイルも異なり、「Submitted」、「Demandant」の出現順が異なるため、自動評価数値は低い。

(3)基準翻訳文にある表現が機械翻訳文にないもの

原文	(下線部は補正箇所である。)
基準翻訳文	The underlines in the following description indicate amended portions.
MT	(An underline part is a correction part.)

人手評価 5

自動評価 1

機械翻訳文は意味的に妥当であるため人手評価は高い。一方、基準翻訳文と機械翻訳文で文のスタイルが異なっており、機械翻訳文の単語数が少なく基準翻訳文にある表現が機械翻訳文にないため、自動評価数値が低くなっている。

(4)公報の記述

原文	特開 2 0 0 3 - 2 6 1 8 2 6 号公報
基準翻訳文	Japanese Unexamined Patent Application Publication No. 2003-261826
MT	JP 2003 - 261826A

人手評価 5

自動評価 1

公報の記載方法について、基準翻訳文では全て「特開 XXXX-XXXXXXX」を「Japanese Unexamined Patent Application Publication No. XXXX-XXXXXXX」のように訳しているところ、機械翻訳では「JP XXXX-XXXXXXX A」等と訳出するルールが適用されている。訳し方のルールが違うため、使用単語や単語数も全く異なり、評価数値の差異が大きい。

(5)ハイフンでつながれた文

原文	審判事件答弁書提出
基準翻訳文	Submission of Written reply for the trial case
MT	(The-written-reply-for-the-trial-case submission

人手評価 5

自動評価 1

基準翻訳文、機械翻訳文とも「Written reply for the trial case」の順で、使用単語も単語の出現順序も同一だが、機械翻訳文は全てハイフンで繋がれているため、一単語として認識されている。

4. 昨年度事業で作成した辞書データの有効性調査（新旧機械翻訳文の比較）

昨年度作成された辞書データ等の機械翻訳の主要な学習用データは、2017 年 9 月末に J-PlatPat の機械翻訳エンジンに反映された。その際にインポートされた辞書データの有効性を検証するため、学習用データ反映前後の機械翻訳文を比較し、分析を行った。

本章では、昨年度の同事業において調査対象とされた審決と、今年度事業の調査対象である審決について、新版機械翻訳文と旧版機械翻訳文をそれぞれ取得し、比較分析することで、新たに登録された辞書データの有効性を検証する。

4.1 新旧機械翻訳文の評価結果

本調査では、審決 1～4 バッチの 40 件を対象に、機械翻訳エンジンが更新される前と後とで機械翻訳文の訳質がどのように変化したかについて、人手による評価と BLEU、RIBES による自動評価の両方を行った。各評価結果について以下に説明する。

4.1.1 新旧機械翻訳文の人手評価の結果

人手評価の対象とした件数及び文数は以下の通りである。

(1) 審決 40 件中の評価対象件数及び文数

四法	件数	文数
特許	32	4,536
商標	6	275
意匠	2	90
合計	40	4,901

表 4.1.1-1 四法毎の評価対象件数

(2) 審決全体の人手評価結果による新旧機械翻訳文の比較

審決 40 件全体の人手評価数値の割合を下記に示す。

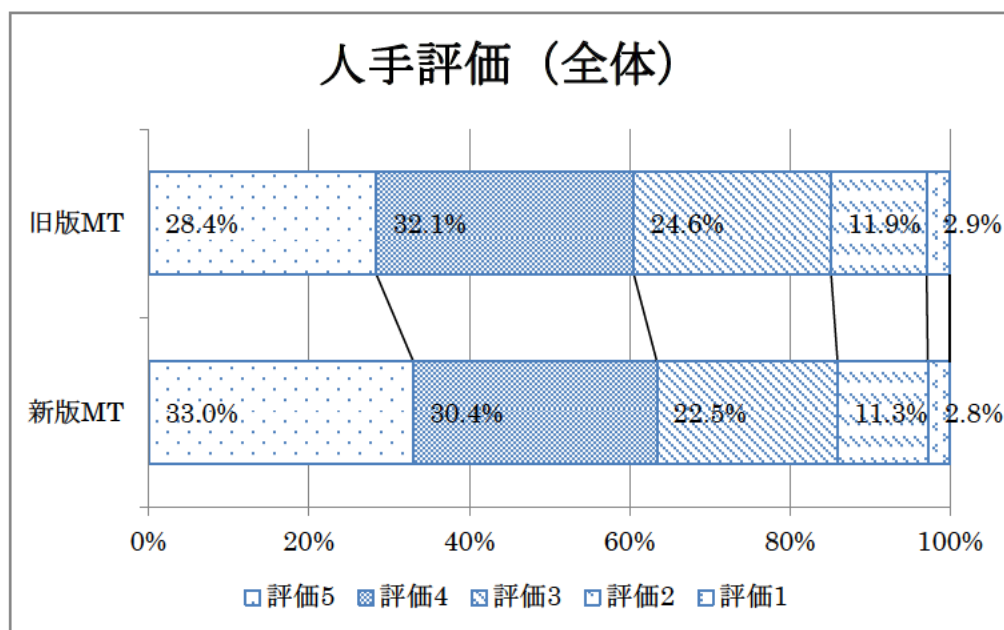


図 4. 1. 1-1 人手評価全体の新旧 MT 評価結果

「図 4. 1-1 人手評価全体の新旧 MT 評価結果」のグラフに示すように、評価 1～5 のすべてにおいて新版の機械翻訳文の評価が旧版の機械翻訳文の評価を上回る結果となった。

(3) 審決の四法毎の人手評価結果による新旧機械翻訳文の比較

審決 40 件の四法毎の人手評価数値の割合を下記に示す。

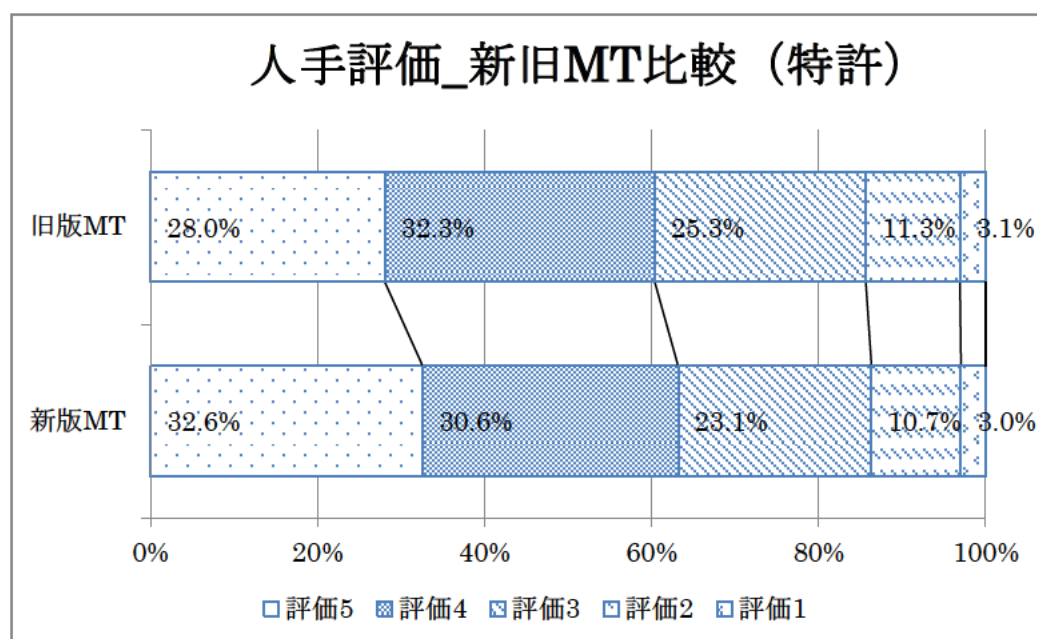


図 4. 1. 1-2 人手評価の四法毎の新旧 MT 評価結果（特許）

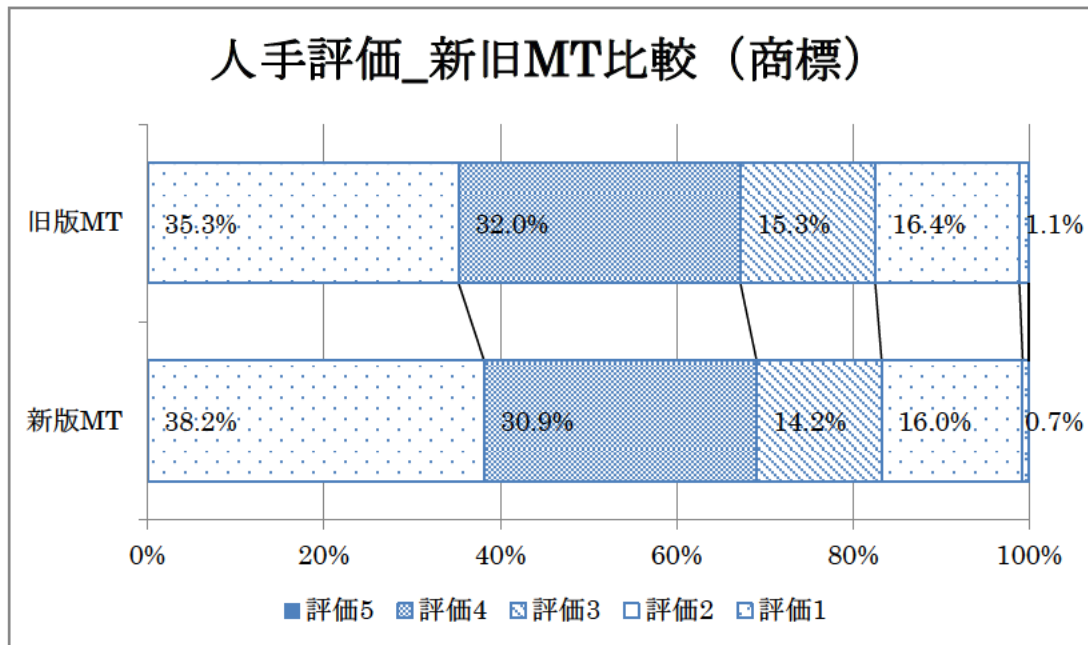


図 4. 1. 1-3 人手評価の四法毎の新旧 MT 評価結果（商標）

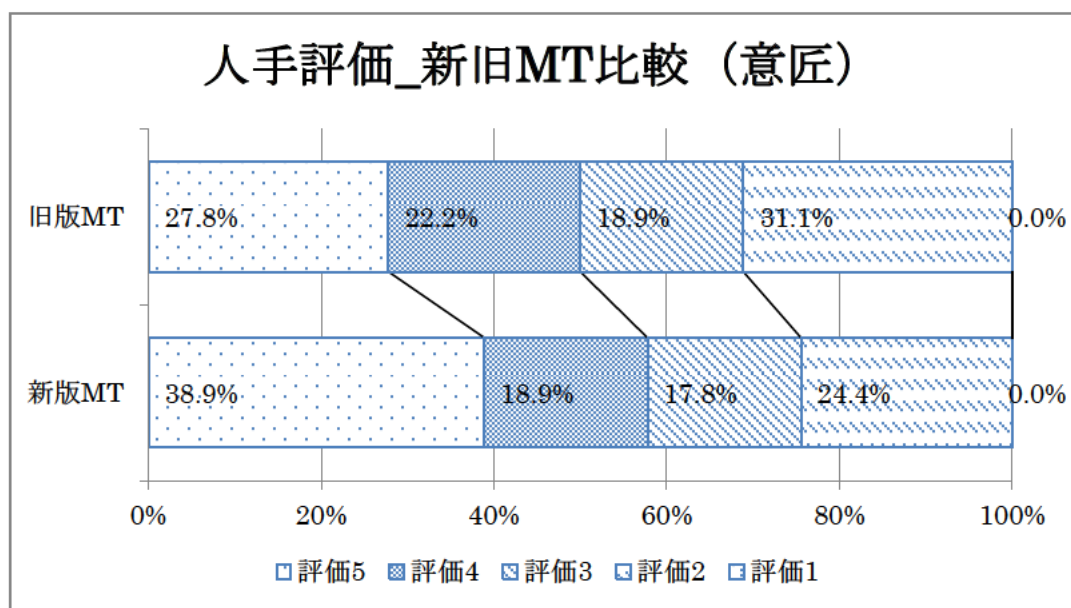


図 4. 1. 1-4 人手評価の四法毎の新旧 MT 評価結果（意匠）

「特許」、「商標」、「意匠」のいずれにおいても、全体評価と同様に、新版の機械翻訳文では旧版の機械翻訳文よりも評価が高い結果となり、改善がみられることが明らかとなった。

4.1.2 新旧機械翻訳文の自動評価の結果

次に BLEU と RIBES を用いた自動評価による新旧の機械翻訳文の比較結果を以下に示す。
人手評価の対象とした件数及び文数は「4.1.1 新旧機械翻訳文の人手評価の結果」と同様である。

(1) 審決全体の自動評価結果による新旧機械翻訳文の比較

審決 40 件全体の BLEU、RIBES それぞれの自動評価数値の平均値を下記に示す。

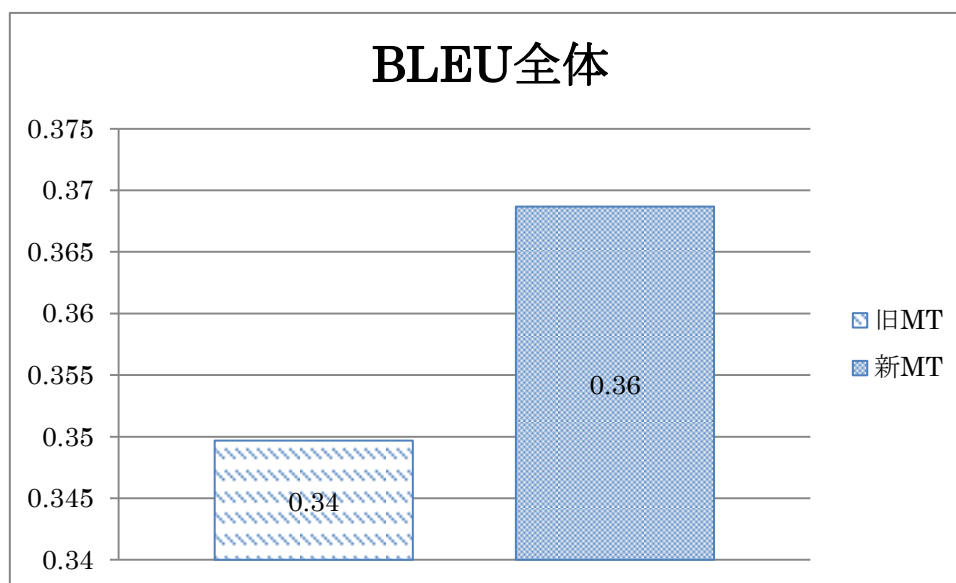


図 4.1.2-1 BLEU における対象文全体の新旧 MT の評価平均値

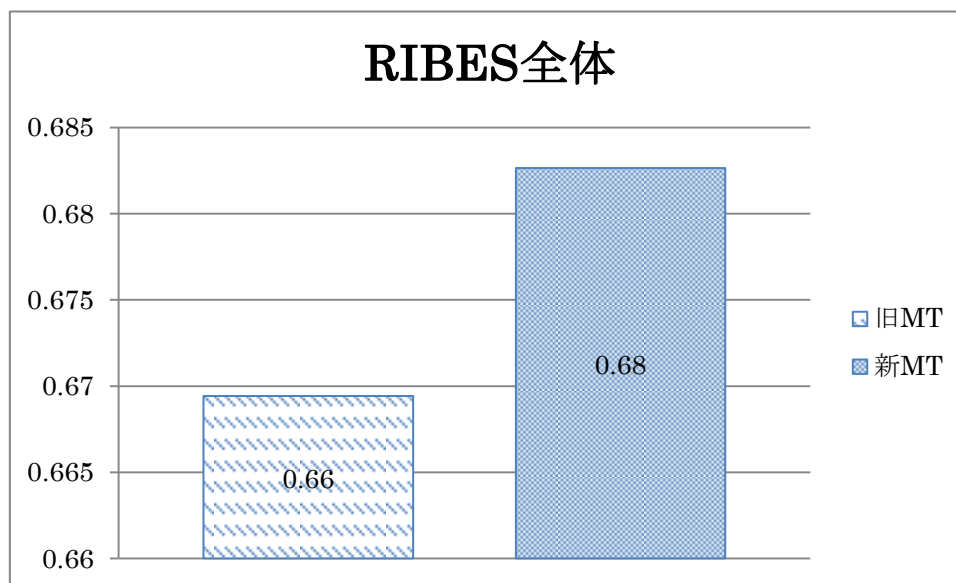


図 4.1.2-2 RIBES における対象文全体の新旧 MT の評価平均値

人手評価による新旧機械翻訳文の評価結果と同様に、新版の機械翻訳文の評価が旧版の機械翻訳文の評価を上回る結果となった。

(2) 審決の四法毎の自動評価結果による新旧機械翻訳文の比較

審決 40 件の四法毎の BLEU 及び RIBES それぞれの自動評価の平均値を以下に示す。

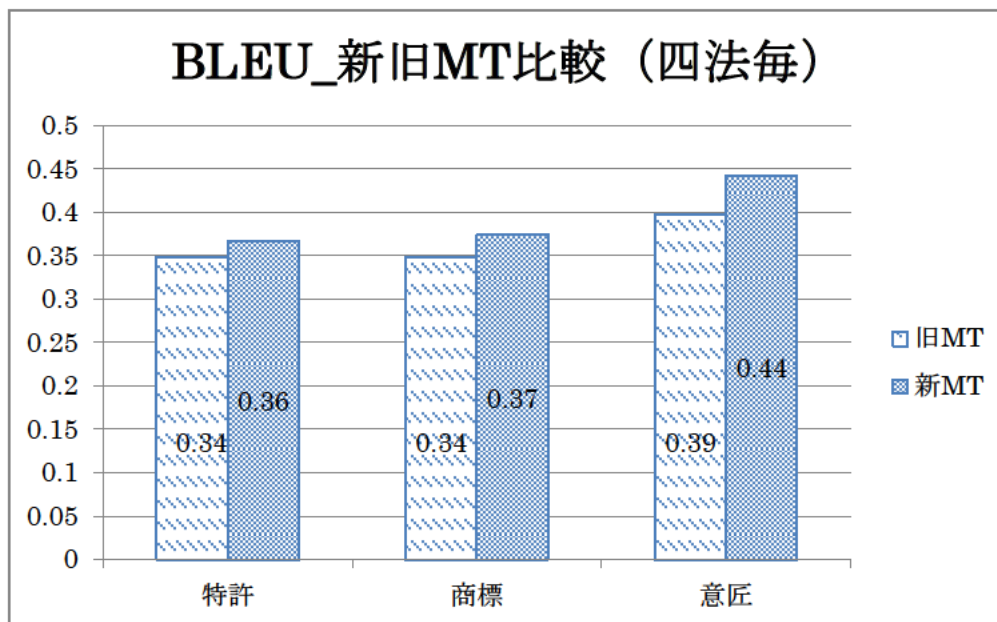


図 4. 1. 2-3 BLEU における四法毎の新旧 MT 評価の平均値

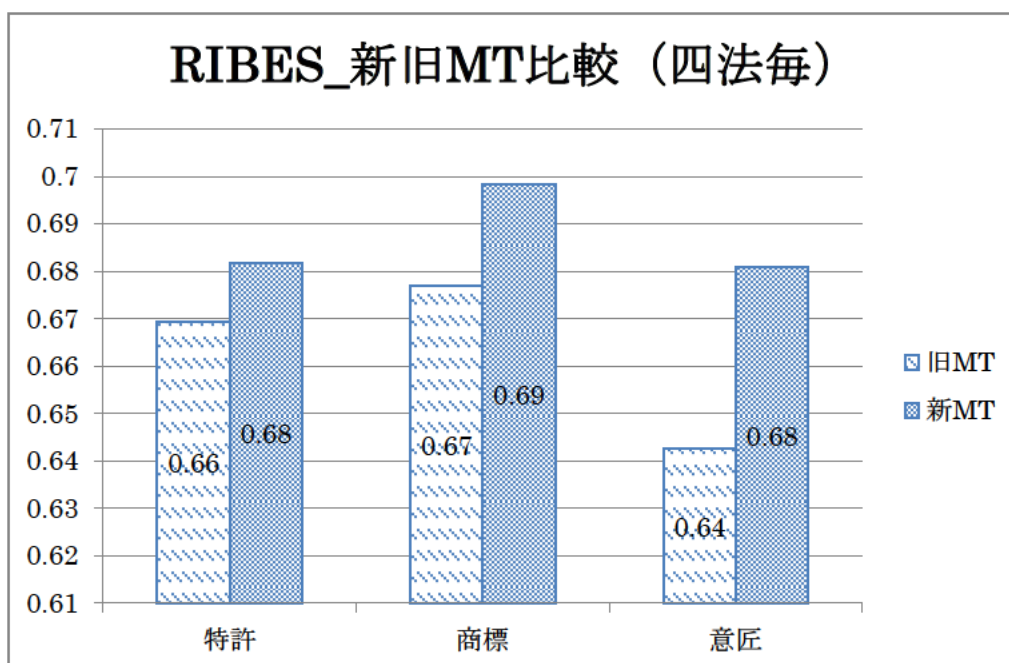


図 4. 1. 2-4 RIBES における四法毎の新旧 MT 評価の平均値

人手評価と同様に、「特許」「商標」「意匠」のいずれにおいても、新版の機械翻訳文では旧版の機械翻訳文よりも評価が高い結果となり、改善がみられることが明らかとなった。

4.2 新旧機械翻訳文の比較分析（昨年度調査対象の審決）

4.1 においては、昨年度の同事業で調査対象であった審決 100 件より、調査対象文を選定し、分析を行った結果をまとめる。

本章で取り上げる審決は、昨年度事業でそれを原文として辞書データ作成を行ったものであるから、辞書データ作成時の想定通りに精度向上が実現されたか確認することを本章の主旨とする。

具体的な対象文献としては、2016 年 9 月末から 2017 年 1 月にかけて取得した J-PlatPat の機械翻訳文（=旧版 MT）と、2017 年 9 月末から同年 12 月初めにかけて取得した機械翻訳文（=新版 MT）を比較した。なお、当該期間において J-PlatPat で入手できなかった文献は特許庁より支給された機械翻訳文を使用した。

分析対象文の選定にあたっては、昨年度の本調査事業において評価数値が低かった「特許」かつ「無効審判」の審決文を母数とし、一文単位のコーパス上で新版 MT と旧版 MT の差分を機械的に色づけし、新旧の差が大きい文を対象とした。

さらに原文のバリエーションを含むよう重複を削除し、最終的に残った 167 文を分析対象文と決定した。

4.2.1 学習用データの反映により機械翻訳文の精度が向上した例

(1) 未知語の解消

人手 評価	原文	「RainbowLoom 公式ガイドはじめてでもかんたん！かわいい！！レインボールームアクセサリー」
2	旧版 MT	It is ***** even if it "Rainbow Loom formal guide Begins! It is lovely!! Rainbow Room accessories"
4	新版 MT	"A Rainbow Loom formal guide Simple, even if it begins! It is lovely!! Rainbow Room accessories"

無効審判における証拠方法の記述箇所である。旧版 MT で未知語だった「かんたん」が新版 MT で訳出されたことで、「はじめてでもかんたん！」がまとまって訳された。

(2) 審決頻出用語の精度向上

人手 評価	原文	ウ．容易想到性について
3	旧版 MT	c. About easy idea nature
5	新版 MT	c. About an easily-conceived property

「容易想到性」を「easily-conceived property」として登録した辞書データが適切に適用された。審決特有の頻出表現が新版 MT において正確に訳される結果となった。

(3)辞書データの適用により全体の文構造が整理されたケース

人手 評価	原文	よって、特許請求の範囲の請求項 2 ないし 4 について、特許法第 1 3 4 条の 2 第 9 項において読み替えて準用する第 1 2 6 条第 7 項の規定は適用しない。
2	旧版 MT	Therefore, in the 9th item of Patent Law Article 134bis, regulation of applied mutatis mutandis Article 126(7) is not applied to Claims 2-4 of a scope of the claim.
5	新版 MT	Therefore, regulation of Article 126(7) applied mutatis mutandis by replacing certain terms pursuant to the 9th item of Patent Law Article 134bis is not applied to Claims 2-4 of a scope of the claim.

「において読み替えて準用する」を「applied mutatis mutandis by replacing certain terms pursuant to」として登録した辞書データが適用された。

これにより全体の文構造が整理され、十分文意の読み取れる文に改善された。

(4)長文の構文解析が適切になされたケース

人手 評価	原文	しかしながら、 当該判決は 、消波構造物の構成には相違点がなく、当該消波構造物の用途に相違点がある事案について、用途の進歩性を 判断 したものであって、用途及びキットの構成に相違点がある本件とは事案を異にするものであり、当審の判断を左右しない。
2	旧版 MT	However, there is no different feature in the composition of a wave absorbing structure thing, about the case which has the different feature in the use of the wave absorbing structure thing concerned, the inventive step of a use is judged , this case which has the different feature in the composition of a use and a kit differs in a case, and the court decision concerned does not influence a judgment by the body.
5	新版 MT	However, the court decision concerned judges the inventive step of a purpose about a case which does not have the different feature in composition of a wave absorbing structure thing, and has the different feature in a purpose of the wave absorbing structure thing concerned, and this case which has the different feature in composition of a purpose and a kit differs in a case. A judgment by the body is not influenced.

旧版 MT は全体の文構造を捉えられていなかったが、新版 MT では 2 文に分割され、主語と述語が整理された結果、重要事項が十分に伝達されうる内容に改善された。長文にも関わらず構文解析が適切になされ、綺麗な訳文として訳出されている。

辞書登録した登録語が文中に適用されているにも関わらず評価数値が上がらないものもあるが、登録語が正しく適用されないからではなく、文中の他の部分で悪文が残っており、依然として全体の文意を把握することは困難なためであることが多い。よって、評価数値に影響が生じるほどの精度向上はなくても、単語レベルで部分的に改善されている例は多くある。また、上記に取り上げたように辞書登録をしたために文構造が大幅に変わり、大きく改善される例も少なからず見られた。

内容伝達レベルで評価する限り、辞書データの適用により文構造が大幅に改善される文でなければ評価数値を上げることは難しいが、実際は語句レベルで部分的な改良がなされている文が多いため、数値の変化に見られる以上に改善された箇所は多く見受けられる。

4.2.2 学習用データの反映により生じた問題点

辞書データの登録により機械翻訳文の精度が向上した文が多数ある一方、少数ながら新規辞書登録により生じた新たな問題点も検出された。

(1) 不自然な名詞の修飾

人手 評価	原文	被請求人が提出した証拠方法は、以下のとおりである。
5	旧版 MT	The means of proof which a demandee submitted is as follows.
3	新版 MT	An evidence-submitted-by-the-demandee method is as follows.

「被請求人が提出した証拠方法」の部分について、旧版 MT では「The means of proof which a demandee submitted」と、後ろから前を説明する構文で正確に訳されていたが、新版 MT では「方法」にかかる修飾語がすべてハイフンで連結され、名詞の前にまとめて形容詞化された不正確な訳文となった。

この機械翻訳文は、見出し語「被請求人が提出した証拠」、訳語「evidence submitted by the demandee」として登録したユーザ辞書が適用されたものである。この辞書データのもととなった原文を調べると、「被請求人が提出した証拠によつては」「以上からすれば、被請求人が提出した証拠からは、」という文であった。そのためこの範囲の見出し語で登録がなされたが、そうすると上記の原文は「被請求人が提出した証拠」までが「方法」に係ると認識され、登録訳語全体が形容詞化されたと思われる。

新たに見出し語「被請求人が提出した証拠方法」、訳語「means of proof submitted by the demandee」として登録語を追加すれば解消される。

(2)不適切な品詞

人手 評価	原文	そして、訂正発明 1 における式 (I) で表される化合物は、外国語書面になんらの新たな技術的事項を導入するものではない。
4	旧版 MT	And the compound denoted by formula (I) in the correction invention 1 does not introduce any new technical matter into a Document in foreign language.
3	新版 MT	And the compound denoted by formula (I) in the correction invention 1 is not any introduces a new technical matter to a Document in foreign language.

旧版 MT では「導入するものではない」に対する訳が「does not introduce」と動詞の否定文として文法的に正しく訳出されていたが、新版 MT では「introduce」が名詞扱いされ、「is not any introduces」と be 動詞の否定文で不自然に訳される結果となった。

(3)不要なハイフン

人手 評価	原文	(口頭審理陳述要領書「4. 陳述の要領 1」)
3	旧版 MT	(Oral proceedings statement brief "4. Declarative point 1)")
5	新版 MT	(Oral proceedings statement brief "4. Statement-brief 1)")

「陳述の要領」を「statement brief」としてユーザ辞書に登録したデータが適用されたものである。訳語は正しいものの 2 語が不要なハイフンで繋がれているため、The 翻訳で下記の設定をしたところ、所望の訳出結果を得られた。

日英詳細辞書設定で、「活用語」の「活用変化する単語の位置」を、「2 番目の単語（規則活用）または 2 単語以上が活用」から、「2 番目の単語」に設定変更。

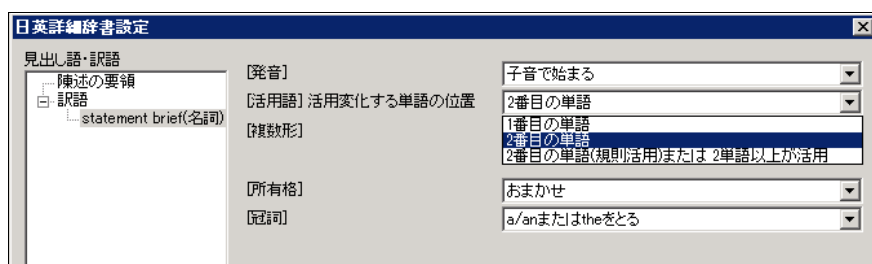


図 4.1.2-1 The 翻訳の辞書データ編集画面

出力結果：

原文	(口頭審理陳述要領書「4. 陳述の要領 1」)
修正後の MT	(Oral proceedings statement brief "4. Statement brief 1")

4.3 新旧機械翻訳文の比較分析（今年度調査対象の審決）

4.1 では、辞書データ登録のもととなった昨年度調査対象の審決を対象に分析を行った。本章では、今年度調査対象の審決について、旧版 MT と新版 MT の精度を比較、分析することで、昨年度事業での辞書データ作成時には予想しなかった効果や副作用があるかを検証する。

4.3.1 学習用データの反映により機械翻訳文の精度が向上した例

昨年度作成した辞書データの適用により、今年度調査対象の審決で新版 MT の精度が向上した例を下記に示す。

(1)

人手 評価	原文	一群の請求項ごとに訂正を請求することについて
4	旧版 MT	About requesting for correction for every a group of claims
4	新版 MT	About requesting for correction in each group of claims

訂正 2016-390104

(2)

人手 評価	原文	したがって、訂正の請求は、一群の請求項ごとにされたものである。
2	旧版 MT	Therefore, the claim of correction is carried out for every a group of claims.
2	新版 MT	Therefore, the claim of correction is carried out in each group of claims.

無効 2016-800080

(1)および(2)は、「一群の請求項ごとに」を「in each group of claims」としてユーザ辞書に登録したデータが適用された文である。(1)は見出し語(タイトル)、(2)は審決の本文だが、どちらも想定通り正しく出力される結果となった。

(3)

人手 評価	原文	また、イ号物件が掲載されている被請求人発行の商品説明書（甲3）を 参酌すると、イ号物件のフラップ取り外しミトンは、メディカルグッズ である。
4	旧版 MT	When a description-of-item document (shell 3) of demandee issue by which article A is published is taken into consideration, flap removal mittens of article A are medical goods.
4	新版 MT	Flap removal mittens of article A are medical goods in view of a description-of-item document (A-3) of demandee issue by which article A is published.

見出し語「を参酌すると」を、訳語「in view of」、機能語としてユーザ辞書に登録したデータが適用された文である。

この辞書データにより、新版 MT の構文が変化して主部と述部が文頭に移動したことで、すっきりとした英文に変わった。見出し語「を参酌すると」は文途中の表現で副作用のリスクが高そうな辞書データだったが、該当部分の訳が適切に訳出されるだけでなく、文構造全体を整理するという効果も得られた。

4.3.2 学習用データが想定通りに反映されなかった例

辞書データを機械翻訳エンジンに反映させた結果、想定通りの正しい訳で出力されなかった例を下記に示す。

(1)

人手 評価	原文	本件商標は、別掲のと通りの構成よりなるものであるところ...
3	旧版 MT	This trademark the place which is what consists of composition as indicated in the Attachment ...
3	新版 MT	This trademark the place which is what is is configured as indicated in Attached ...

異議 2016-900037

昨年度作成した辞書データにおいて、見出し語「別掲のと通りの構成よりなる」を訳語「is configured as indicated in Attachment」として登録し、さらに付属情報として見出し語品詞を五段活用動詞、訳語品詞を動詞として設定した辞書データが適用されたものである。

新版 MT を見ると、「what is is」と be 動詞が 2 つ重なっており、更に「別掲」が「Attachment」

となるべきところ、「Attacheded」と動詞の活用形のような形で訳出される結果となった。

原因として The 翻訳のユーザ辞書では動詞句を登録する際、前から何番目の語が活用するかを設定することができるが、その活用語の設定に誤りがあったと考えられる。

活用語	設定できる品詞	名詞、動詞、形容詞、副詞、目的語を必要とする動詞句、目的語を必要としない動詞句（ただし、見出し語が2語以上から構成される場合）	
	意味	見出し語を構成する語のうち、中心となって活用する語は前から何番目の語か	
	属性値	1 2 :	1 番目の単語 (例) go(es) back(動詞), part(s) of speech(名詞) 2 番目の単語 (例) blank sheet(s) of paper(名詞) :

図 4.2.2-1 活用語の設定（辞書メンテナンス読本より抜粋）

正しい訳が出力されるよう、弊社保有の The 翻訳を使用してユーザ辞書に下記の設定で再度辞書登録を行った。

見出し語：「別掲のと通りの構成よりなる」

訳語：「be configured as indicated in Attachment」

活用する語：前から 1 番目の単語（be）

活用方法：不規則変化

図 4.2.2-2 The 翻訳の辞書登録画面

結果、機械翻訳文は下記のように出力された。

原文	本件商標は、別掲のと通りの構成よりなるものであるところ...
MT	This trademark the place which is what is been configured as indicated in Attachment the character portion of "PENGUIN" under the composition,

「別掲」の「Attachment」が正しく名詞の形に修正された。

be 動詞の活用については、上記の英文では進行形となっているが、The 翻訳に受動態として「configure as indicated in Attachment」と登録すると「is configure as...」という訳出になってしまうため、この登録設定がベターと判断した。

また、修正した辞書データを登録した The 翻訳で昨年度事業対象の審決でも試したところ、原文「本件商標は、別掲のと通りの構成よりなるものであつて」が「This trademark is configured as indicated in Attachment」と思惑通りに訳される結果となったため、この修正で問題ないと判断する。

原文	本件商標は、別掲のと通りの構成よりなるものであつて...
MT	This trademark is configured as indicated in Attachment ...

辞書データの見出し語と訳語の範囲が正しくても、上記のように属性情報の設定次第で機械翻訳文が誤訳として出力される可能性があるため、辞書データ登録の際はデータ毎に属性情報を確認・設定する必要がある。

あるいは誤った属性情報によるエラーを回避する方法として、受動態での登録は極力避け、「consist of」のような能動態で登録することが考えられる。

4.4 辞書データによる解決が困難な例

前章まで新たな辞書データの適用によって改善された例、改悪となった例を挙げたが、旧版 MT、新版 MT においても誤訳箇所が変化しなかった文もある。解決されなかった誤訳は以下に示すような構文解析の誤りに起因するものが多い。

(1) 否定文が肯定文で訳出されるケース

人手 評価	原文	しかしながら、上記のとおり、かかる甲号証は、 いずれも、「ブルニア ンリンク」の作成方法について本件優先日当時の技術常識を示すもの ではないから、 当審の判断を左右しない。
2	旧版 MT	However, since all these demandant's exhibits show the technical

		common sense at the time of this priority date about the preparation method of "Bull Neang Rink" as above-mentioned, a judgment by the body is not influenced.
2	新版 MT	However, as described above, since all these demandant's exhibits show a common general technical knowledge at the time of the priority date for the Invention about the preparation method of "Bull Neang Rink", they do not influence a judgment by the body.

適用された辞書データ：

本件優先日当時の技術常識 - common general technical knowledge at the time of the priority date for the Invention

旧版 MT および新版 MT とともに「いずれも・・・技術常識を示すものではない」が、否定文で訳出されなかった。その結果、「すべての甲号証が・・・技術常識を示す」という反対の意味内容となってしまった。

試しに The 翻訳で「いずれも・・・技術常識を示すものではない。」のみの文章で機械翻訳をかけると、「Neither shows the technical common knowledge」と否定形で訳された。あるいは上記原文より「いずれも、」だけを削除した「「ブルニアンリンク」の作成方法について本件優先日当時の技術常識を示すものではないから、当審の判断を左右しない。」という文章で訳すと、「does not show」という否定文で出力された。原文の文章が長すぎたため、「いずれも・・・ない」という日本語の構文が正しく出力されなかったことが分かる。

(2)二重否定の訳出エラー

原文	しかし、このような特許請求の範囲の記載に対する実施例の多寡を理由として、本件特許が特許法第36条の規定する要件を満たしていない ということ はできない。
旧版 MT	However, the thing in which this patent is Patent Act Article 36 for reasons of the amount of an working example to the description of such a scope of the claim and to specify of not complying with requirements is made, and there is. [no]
新版 MT	However, the thing it is [a thing] such, whose Patent is Patent Act Article 36 on the ground of the number of Examples with respect to description of the scope of claims and to specify of not complying with requirements is made, and there is. [no]

辞書データ作成の結果、新規に登録した「特許請求の範囲の記載に対する実施例の多寡を理由として」「 on the ground of the number of Examples with respect to description of

the scope of claims」という辞書データは適切に適用されたものの、「満たしていないということはできない」という二重否定が正確に訳出されず、文意を汲み取ることが難しい。

(3)複数括弧がある場合の訳出エラー

人手 評価	原文	してみると、本件発明 1 の「 細いテープ状部材に粘着剤を塗着する 」とは、「 細いテープ状部材に粘着剤を塗着するもの 」、「 シート状部材に粘着剤を塗着し細片状に切断するもの 」、両者を含むと解されるから、相違点であるとの請求人の主張は採用できない。
3	旧版 MT	Then, since it is understood as "a thing which plasters a thin tape shaped member with a binder" "a thin tape shaped member is plastered with a binder", "a thing which plasters a sheet shaped member with a binder and it cuts in the shape of a split", and both being included, an appellant's allegation that it is the different feature is not employable. [of this invention 1]
3	新版 MT	Since it is understood as "a thing which plasters a thin tape shaped member with a binder" "a thin tape shaped member is plastered with a binder" in view of the above, "a thing which plasters a sheet shaped member with a binder and it cuts in the shape of a split", and both being included, an appellant's allegation that it is the different feature is not employable. [of Invention 1]

新版 MT および旧版 MT とも、括弧で括られた箇所の関係性に誤りがあり、改善されていない。そもそもの原文が各括弧内の関係性を読み取ることが難しい表現である。

(4)修飾語が多数ある場合の訳出エラー

人手 評価	原文	本件補正前の平成 27 年 4 月 3 日付けの 手続補正により補正された特許請求の範囲の請求項 1 及び請求項 3 の記載 は次のとおりである。
-	基準翻訳文	The descriptions of Claim 1 and Claim 3 of the scope of claims amended by the written amendment dated on April 3, 2015 before the Amendment are as follows.
3	旧版 MT	Claim 1 of a scope of the claim and a description of Claim 3 which were corrected by Amendment as of [of an in front of / this Amendment]Heisei 27(2015) April 3 are as follows.
3	新版 MT	Claim 1 of a scope of the claim and a description of Claim 3 which were corrected by Amendment before the Amendment and as of Heisei 27(2015) April 3 are as follows.

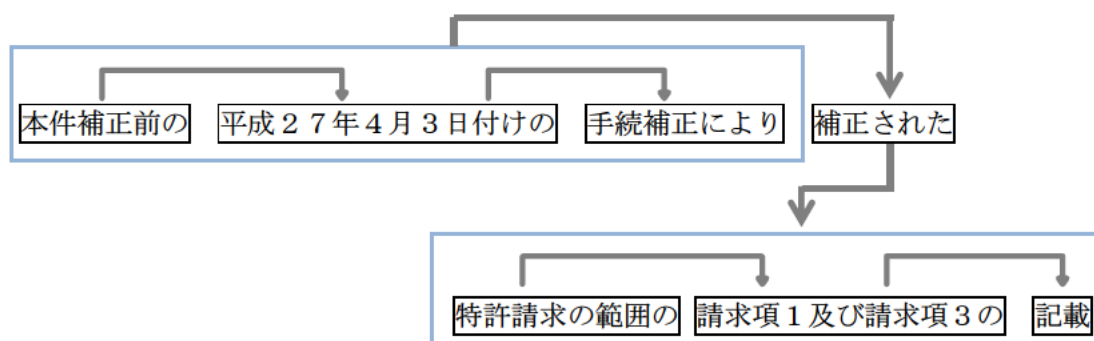
適用された辞書データ：

補正前 -before the Amendment

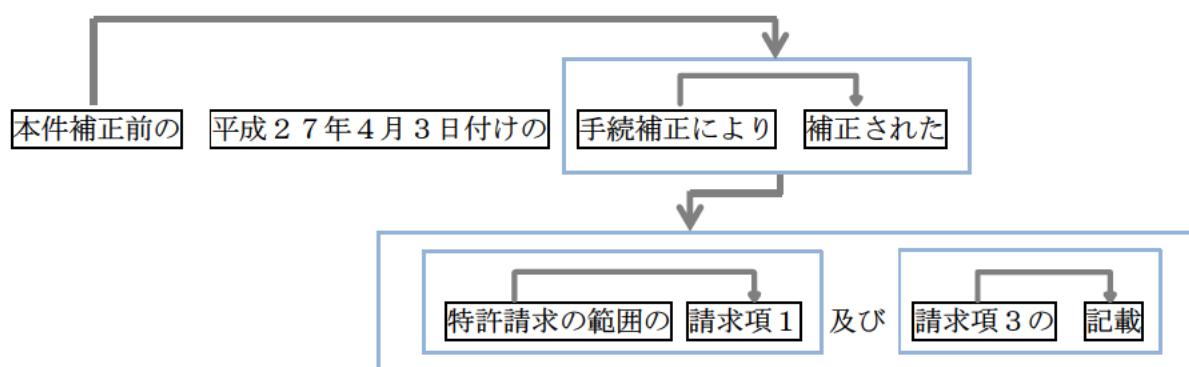
特許請求の範囲- scope of the claim

原文で名詞「記載」に係る修飾語が多く、機械翻訳で係り受けを誤っている例である。

原文の係り受け



新版 MT の係り受け



辞書データが適用されたことで、旧版 MT では「an in front of / this Amendment」と訳出された「本件補正前」が、新版 MT では「before the Amendment」と訳出されるようになった。

一方、係り受けは隣り合った語句など狭い範囲でなされる傾向があり、新版 MT でも誤ったままである。

辞書データの適用により文中の語句・表現の単位で精度向上の一定の効果は得られたが、上記に挙げたような否定文や係り受け等、構文に関するエラーは解消されず課題として残ったままの箇所が多い。辞書データによりこうした構文解析のエラーの解決を図るのは限界があるように思われる。

5. 辞書データの作成・分析

5.1 辞書データの作成手順

(1) 辞書データの登録候補語の選定

辞書データ作成担当者が、審決の機械翻訳文の中から誤訳となっている語（用語）、句（フレーズ）、又は文を審決1件につき10語程度抽出し、登録候補語リストを作成した。登録候補となる語句は基本的に審決に特有な表現に特化するものとした。

(2) 辞書カテゴリー分類の決定、登録候補語の編集

登録候補語リストの作成後、個々の登録候補語を The 翻訳で指定されている3種類の辞書データのいずれかに適合するよう登録範囲の決定や付属情報の追加等の編集を行った。

3種類の辞書（ユーザ辞書、メモリ辞書、パターン辞書）の特徴は次の通りである。

	ユーザ辞書	メモリ辞書	パターン辞書
特徴	語、句を登録 品詞の設定、活用する 動詞の設定、冠詞の種 類の設定が可能	文単位の対訳文を登録	文頭から文末までの文単 位で対訳文を登録 可変部分設定が可能
可変部分の 設定可否	数字のみ設定可能 最大4カ所	不可	複数語からなる句を可変 部分とすることが可能 最大5カ所
文字数制限	原文：全角100文字、 訳文：半角200文字	原文：全角500文字、 訳文：半角1000文字	原文：全角160文字 訳文：半角320文字

表 5.1-1 辞書データの種類

①ユーザ辞書

数字以外の可変部分を含まない語や句が対象となる。可変部分は数字のみ設定できるため「特許法第17条の2第4項の要件違反になる」のように、数字以外の訳を指定したい場合も本辞書に登録できる。

登録例：

原語	訳語
デジタル通信	Digital Communications Division
引用文献<1>、<2>参照	See cited documents <1>, <2>

②メモリ辞書

可変部分を含まない文単位の対訳文が対象となる。辞書データの原語と翻訳対象の原文が完全一致しないと適用されないため、適用される箇所が少ないというデメリットがある。一方、適用された場合は、完全な文として訳出されるので、副作用の心配が少ないというメリットがある。定型的で汎用文に近い文はメモリ辞書で登録すると効果が期待できる。また、複数のバリエーションを展開した上で登録するとより効果的である。

登録例：

原語	訳語
したがって、本願のその他の請求項に係る発明について論及するまでもなく本願は拒絶すべきである。	Therefore, the application should be rejected without examining other claims.

③パターン辞書

可変部分を含む文頭から文末までの文単位で対訳文が対象となる。原文と訳文の翻訳例に可変部分を指定した翻訳パターン（穴空き翻訳例文）を登録する。パターン辞書は、文章のひな型として利用する前提となっているため、登録する範囲は主語、述語のある完全な文とする。

登録例：

原語	訳語
<1>は、<2>の構成要件<3>に相当する構成を備えているとはいえない。	It cannot be said that <1> has a configuration corresponding to the constituent component <3> of <2>.

(3) 関連情報の記入

登録候補語リストに記載された各登録候補語について、下記項目の記入を行う。特筆すべき点について、以下でさらに説明する。

項番	連番を記入
派生番号	ユーザ辞書について、切り出し位置を変えた場合、メモリ辞書、パターン辞書について、バリエーションを展開した場合に、親となる語の「項番」を記入
辞書種別	ユーザ辞書：Y、メモリ辞書：M、パターン辞書：P
見出し語	見出し語（日本語）を記入 ユーザ辞書、パターン辞書の場合、文字数制限有
見出し語品詞	ユーザ辞書の場合記入
品詞種別	品詞が動詞の場合（五段／一段／カ変／サ変）の別、形容動詞の場合は（ダ

	型／タルト型) の別
訳語	訳語（英語）を記入 ユーザ辞書、パターン辞書の場合、文字数制限有
訳語品詞	ユーザ辞書の場合記入
活用語の位置	品詞が動詞の場合に記入（何語目か 半角数字）
冠詞設定	品詞が名詞の場合に記入（a, an／the／無冠詞）
誤訳箇所	誤訳箇所前後を摘記
誤訳文献	文献番号等
有用性	H か L
リスク	H か L
登録候補語分類	最優先登録語群（有用性高、リスク低）：S 優先登録語群（有用性低、リスク低）：Y 影響調査必要語群（有用性高、リスク高）：影響未調査 E、調査済・登録する：Eo、調査済・登録しない：Ex 登録しない（有用性低、リスク高）：X

①登録候補語句の「有用性」

登録候補語が機械翻訳の精度向上に及ぼす「有用性」については下記の基準に従って判断した。

「有用性」（＝登録することにより改善が見込める度合い）

有用性高 の代表例：

- ・未知語となっているもの
- ・現状重大な誤訳がある

有用性低 の代表例：

- ・現状の訳は不自然な訳であるが誤解を生むという程ではない
- ・誤訳原因が複雑で辞書による改善の見込みが薄い

②登録候補語句の「リスク」

登録候補語が機械翻訳の精度向上に及ぼす「リスク」については下記の基準に従って判断した。

「リスク」（＝登録することによる副作用の度合い）

リスク高 の代表例：

- ・複数の語からなる句で、修飾すべき語が中にある語
- ・訳し分けが必要（状況によって適訳が変わる可能性が高い）

リスク低 の代表例：

- ・文単位のもの（メモリ辞書候補等）
- ・単単語・2単語程度の語で、訳し分けが必要ない

③「有用性」と「リスク」に応じた分類分け

「有用性」と「リスク」に応じて下記の通り登録候補語分類の分類分けを行った。

	有用性高	有用性低
リスク低	最優先登録語群：S	優先登録語群：Y
リスク高	影響調査を実施：E	登録しない：X

(4) 影響調査

登録候補語のうち、リスク高かつ有用性高と判断された語について、当該登録候補語を辞書データとして登録した場合に悪影響が出るか否かを判断し、悪影響が出ない語は登録候補語分類を「調査済・登録する：Eo」と判断し、悪影響が出る語については登録候補語分類を「調査済・登録しない：Ex」と判断した。

(5) 辞書登録語リストの作成

翻訳品質管理責任者が、辞書登録候補語リストのうち、登録候補語分類が「登録しない：X」、「調査済・登録しないEx」と判断された語を削除して、納入物となる辞書登録語リストを完成させた。

5.2 作成結果

審決の機械翻訳用辞書の作成結果

(1) 「辞書種別」の内訳

審決における「登録候補語リスト」の「辞書種別」の内訳は、以下の通りである。

辞書種別	語数	割合
ユーザ辞書：Y	1,092	66.5%
メモリ辞書：M	284	17.3%
パターン辞書：P	265	16.1%
合計	1,641	100.0%

表 5.2-1 審決の登録候補語リストの「辞書種別」毎の語数

上記表のように、全体の 66.5%をユーザ辞書が占める結果となった。

(2) 「登録候補語分類」の内訳

「登録候補語リスト」の「登録候補語分類」の内訳は以下の通りである。

登録候補語分類	語数	割合
最優先登録語群（有用性高、リスク低）：S	964	58.7%
優先登録語群（有用性低、リスク低）：Y	609	37.1%
影響調査必要語群（有用性高、リスク高）： 調査済・登録：Eo	67	4.1%
登録しない（有用性低、リスク高）：X	1	0.1%
合計	1,641	100.0%

表 5.2-2 審決の登録候補語リストの「登録候補語分類」毎の語数

全体の 58.7%が「最優先登録語群（有用性高、リスク低）：S」に分類され、「優先登録語群（有用性低、リスク低）：Y」は 37.1%であった。

「影響調査必要語群（有用性高、リスク高）」に分類された 68 語（4.1%）に対して影響調査を行い、その結果、大半を「調査済・登録する：Eo」として登録するに至った。

(3) 「（見出し語品詞）訳語品詞」の内訳

「ユーザ辞書：Y」を原語の品詞（「見出し語品詞」）と訳語の品詞（「訳語品詞」）に振り分けたところ、以下のような分類結果となった。

（見出し語品詞）訳語品詞	語数	割合
（サ変名詞）形容詞	1	0.1%
（サ変名詞）名詞	4	0.4%
（機能詞）機能詞	66	6.0%
（形容詞・形容動詞）形容詞	45	4.1%
（形容詞・形容動詞）動詞	13	1.2%
（形容詞・形容動詞）副詞	1	0.1%
（固有名詞）名詞	37	3.4%
（動詞）形容詞	8	0.7%
（動詞）動詞	100	9.2%
（副詞）副詞	54	4.9%
（名詞）形容詞	6	0.5%
（名詞）名詞	749	68.6%
（連体詞）形容詞	7	0.6%

表 5.2-3 審決の登録候補語リストの「（見出し語品詞）訳語品詞」毎の語数

「（見出し語品詞）訳語品詞」の組合せが「（名詞）名詞」のものが全体の 68.6%と大多数を占めている。

5.3 登録語の編集における工夫

(1) 効果のある辞書登録

① 熟語単位での登録

単語単体の登録だと文脈に応じて訳し分けが必要になるなど、副作用のリスクもある。熟語単位で登録すると、ばらばらの位置で訳出されていた語がひとかたまりの語として扱われることで、文全体の構文が整理されることにも繋がる。

② 頻出する定型文言を数多く登録する

「本件審判の請求は成り立たない。(The appeal of the case was groundless.)」など、審決に頻出するフレーズを登録することで機械翻訳文の目覚ましい向上が見込まれる。

③ 表現のバリエーションを登録する

「てにをは」などを変えて辞書データが適用されるべき見出し語を網羅的に拾えるよう複数の言い換え表現を含む登録を行う。

(2) 辞書データを登録する際の注意点

① 活用語の位置を正しく設定する

動詞句や名詞句は活用する語の位置を正しく設定することが必要である。活用語の位置が誤っていると、名詞+ed など存在しない活用系で出力されたり、不必要なハイフンで語と語が連結されたりする恐れがある。

② 冠詞を正しく設定する

辞書データの訳語に冠詞を含めて名詞を登録すると、機械翻訳に掛けたとき「the the patent invention」のように冠詞が二重化してしまうので注意が必要。

③ 文脈により訳語が異なる語は複合語で登録する

「組織」のように文脈により訳し分けが必要な語は、複合名詞での登録が有効である。

例：

見出し語「生体組織」、訳語「biological tissue」

見出し語「慈善組織」、訳語「charitable organization」

(3) 正しい適用が困難な辞書登録

① 接続語

「してみると」を「in view of the above」で登録したが、辞書登録することで訳語の位置が文頭から文中に変わり、正しい位置で訳出されなくなってしまう。

②受動態

受動態など、be 動詞を含めた熟語の登録は前後の文により適切に訳出されない場合があるため、受動態にできる語句は極力能動態に修正して登録する。

例「～の構成よりなる」

× 「is configured as」

○ 「consist of」

③サポートされていない助詞の指定

The 翻訳では、見出し語の前につく助詞と、それに対応する訳語の前置詞を指定できるが、助詞のパターンが 5 種類しかないため、このパターンに該当しない助詞の設定はできない。

例えば、見出し語「検討する」で登録し、原語が「～で検討する」の場合に「consider by」と訳出させたくても、前置詞 (by) を指定することはできない。

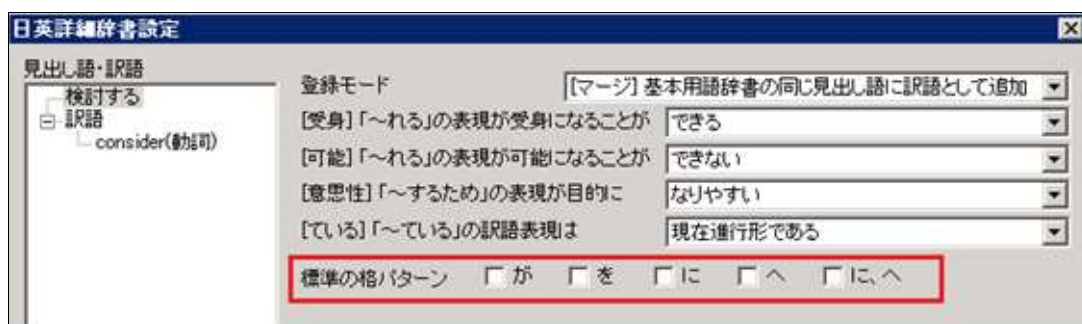


図 5.3-1 The 翻訳の辞書データ編集画面

5.4 辞書データ作成についての提言

「4.1 新旧機械翻訳文の評価結果」の比較グラフが示すように、昨年度事業で作成した辞書データを加えることによって今年度事業の機械翻訳文の精度が向上したことが明らかとなった。このように、現状の機械翻訳エンジンに登録する辞書データの量を増やすことで機械翻訳文の品質の改善を図ることは可能と考えることができる。

その一方で、「4.4 辞書データによる解決が困難な例」等で分析したように、昨年度時点での旧版機械翻訳文と今年度時点での機械翻訳文のいずれにおいても問題が解決されず誤訳となった文が一定数存在することも明らかとなった。

旧版と新版の両方の機械翻訳文において誤訳となったものの多くが原文の日本語の複雑さに起因しており、審決文の一部を占める複雑な文については現状の機械翻訳エンジンが

構文解析を適切に行っていないように見受けられた。

今後、このような複雑な構造を有する審決文においても大意のとれる機械翻訳文を作成するためには、複雑な文であっても類似した人手翻訳文を一部流用して機械翻訳文を作成することができる統計ベースの機械翻訳エンジンの適用が有効と考える。

以上により、今後、審決の機械翻訳文のさらなる向上のために、統計ベースの機械翻訳エンジンの利用または併用と、高品質かつ大容量の人手翻訳された対訳文例集の整備を検討していくことを提言したい。