

機械翻訳（A I 翻訳）に関する調査研究業務
調査報告書

令和4年12月23日

凸版印刷株式会社

目次

1.	はじめに.....	1
2.	機械翻訳活用に関する法務省のこれまでの取組	2
3.	本調査研究業務の内容	5
3.1	概要	5
3.2	機械翻訳結果の誤訳パターン抽出	6
3.3	翻訳システムの構築	8
3.3.1	翻訳システム	8
3.3.2	翻訳エンジンの概要	9
3.3.3	翻訳モデルの概要	10
3.3.4	対訳コーパス及びその整備方法	12
3.3.5	本調査研究業務で構築した翻訳システムの概要	14
3.4	翻訳品質改善プログラム	16
3.4.1	概要	16
3.4.2	追加学習による改善（翻訳品質改善プログラム【1】）	18
3.4.3	前後処理プログラムによる改善（翻訳品質改善プログラム【2】～【7】）	22
4.	翻訳結果に対する評価	34
4.1	評価対象	34
4.2	客観評価	35
4.3	主観評価1（スコア評価）	36
4.4	主観評価2（エラーカウント）	37
4.5	主観評価3（詳細評価）	38
4.6	各課題への対応結果	41
4.7	翻訳サンプル再生成と評価	48
5.	翻訳品質の改善	55
5.1	評価結果の概略（初回翻訳結果）	55

5.2	評価結果の概略（2回目翻訳結果）	61
5.3	効率的かつ効果的なA I学習の強化方法及び翻訳品質改善プログラムの提案	64
	参考資料	66

1. はじめに

日本法令外国語訳は、平成16年11月の司法制度改革推進本部決定により政府の施策として積極的に取り組むことが決定され、従前は内閣官房において業務を行っていたが、平成18年12月の関係省庁連絡会議の決定により、平成21年度から法務省が業務を承継している。

法令外国語訳では、①法令所管府省庁による英訳法令の原案作成、②法務省における品質チェック、③法令所管府省庁による最終確認という工程を経て、翻訳法令公開の専用ホームページである日本法令外国語訳データベースシステム（以下「JLT」という。）に英訳法令の最終公開を行っている。

社会経済のグローバル化が進む中、日本法令の外国語訳を整備して国際発信することは、日本企業の海外進出や対日投資等の促進を図るためのインフラ整備として大変重要であり、法令外国語訳整備を一層加速することの必要性が各所で指摘されているところ、現状、法令の改正（公布）からその最終公開まで、平均約2年半を要しており、特に法令所管府省庁が法令の英訳原案を作成して法務省に提出するまでに平均して約2年もの長い期間を要している。

このような状況から、翻訳工程におけるAI（Artificial Intelligence）の技術を利用した高度な機械翻訳（以下、単に「機械翻訳」という。）を活用することの必要性が各所で指摘されており、これにより将来的に英訳法令等の公開までの期間について抜本的な改善を図ることができる可能性も考えられるため、法務省において、機械翻訳を活用して英訳法令の原案を作成する際の、翻訳品質改善に係る調査研究を実施するものである。

2. 機械翻訳活用に関する法務省のこれまでの取組

法務省では、日本法令外国語訳整備プロジェクトの一環として、英訳法令の公開の迅速化を目的とした、機械翻訳の活用・導入に向けた検証を実施してきた。

平成31年度における検証では、機械翻訳によって得られた英訳は、文法上の誤りは基本的にはないが、翻訳の原文となる日本語の構造が複雑な場合に不正確な英訳となることがあることがわかった。しかし、令和3年度における検証では、機械翻訳エンジン（翻訳処理装置）の性能向上により、平成31年度に比して文法の誤りや不正確な訳が減少したことで、翻訳の質が飛躍的に向上していることを確認した。一方で、日本語と英語の構造の違いから生じる不適切な主語の補い、訳語の不統一、並びに、法令翻訳の統一的指針である「法令用語日英標準対訳辞書」及び「法令翻訳の手引き」への非準拠といった課題があることが判明した。以下に課題の実例を記載する。

課題① 不適切な主語の補い

※日本語では、主節・条件節で主語が同じ場合、一方の主語を省略することがあるところ、英語では、英文として成立させるため、主語を補足する必要がある、その場合に、前後の文脈の関係で不適切な主語が補われることがある。

原文に存在しない単語が訳文に追加されることから、機械翻訳においては「湧き出し」と呼ばれる。

(例) 裁判外紛争解決手続の利用の促進に関する法律

【日本語原文】

第二十三条四項 法務大臣は、第一項又は第二項の規定により認証を取り消したときは、その旨を官報で公示しなければならない。

【機械翻訳による翻訳結果】

(4) If the Minister of Justice rescinds certification pursuant to the provisions of paragraph (1) or (2), the Prime Minister must issue public notice of this in the official gazette.

主語のない日本語から英語への翻訳過程で、「首相」を意味する「Prime Minister」が主語として補完されているが、正しくは、「法務大臣」を指す語句（単数の” they” や” That Minister” 等）にすべきであり、主語が誤って補われている。

課題② 訳語の不統一、「法令用語日英標準対訳辞書」・「法令翻訳の手引き」への非準拠

(1) 訳語の不統一

同じ法令内で使用されている同じ語には、同じ訳語を用いるべきであるところ、別の訳語が用いられることがある。

(例) 裁判外紛争解決手続の利用の促進に関する法律

【日本語原文】

第六条七号 民間紛争解決手続の開始から終了に至るまでの標準的な手続の進行について定めていること。

【機械翻訳による翻訳結果】

(vii) Standard procedures for the process from the commencement to the termination of private dispute resolution procedures have been established.

【日本語原文】

第十四条三号 第六条第七号に規定する認証紛争解決手続の開始から終了に至るまでの標準的な手続の進行

【機械翻訳による翻訳結果】

(iii) the standard operation process from the commencement to the termination of certified dispute resolution procedures as prescribed in Article 6, item (vii) ;

「標準的な手続の進行」の翻訳が、いずれも誤りではないものの、訳語が異なっている。法令翻訳においては、異なる訳語を用いることにより、意味が異なるものであるという誤解を生じさせるおそれがある。

(2) 「法令用語日英標準対訳辞書」、「法令翻訳の手引き」への非準拠

(例) 債権管理回収業に関する特別措置法施行規則

【日本語原文】

第四条二項 債権回収会社は、法第七条第一項に規定する届出をしようとするときは、別紙様式第四号により作成した変更等届出書に、前条各号に掲げる書類のうち変更に係るものを添付して、法務大臣に提出しなければならない。

【機械翻訳による翻訳結果】

(2) When a claim management and collection company intends to make a notification as prescribed in Article 7, paragraph (1) of the Act, it shall submit to the Minister of Justice a written notification of change, etc. prepared using appended form 4, and attach thereto the documents listed in the items of the preceding Article which pertain to the change.

【日本語原文】

第十条 法第十七条第二項に規定する法務省令で定める事項は、次に掲げる事項とする。

【機械翻訳による翻訳結果】

Article 10 (1) The matters specified by Ordinance of the Ministry of Justice as prescribed in Article 17, paragraph (2) of the Act shall be the following :

いずれも英訳として誤りではないが、「法令翻訳の手引き」において、義務の表現（～しなければならない）や定義規定（～とする）では用いないこととしている“shall”が含まれている。

3. 本調査研究業務の内容

3.1 概要

本業務においては、前記2つの課題を重点解決事項として、その解決に当たっては、翻訳エンジンに翻訳品質改善を実現するプログラム（以下「翻訳品質改善プログラム」という。）を実装することとした。

どのような翻訳品質改善プログラムを実装するのが有用かを検討するに当たり、法令の翻訳としての品質改善を図るため、効率的かつ効果的なAI学習の方法や、日本法令の英訳としての翻訳の品質を高める方法を調査すべく、下記の内容を実施した。なお、日本法令の翻訳については、統一的で信頼できる法令の翻訳が継続的に行われるようにするため、「法令用語日英標準対訳辞書」や「法令翻訳の手引き」等の翻訳ルールに準拠した英訳原案を作成することとしており、本業務における「翻訳品質改善」とは、単なる英文としての品質だけでなく、これらの翻訳ルールへの準拠性を高めることも含んでいる。

（1）機械翻訳の誤訳パターン抽出及び改善事項の選定

どのような法令翻訳における改善点（以下、「誤訳パターン」という。）について翻訳品質改善を図り、どのような翻訳品質改善プログラムを実装するかを検討するに当たり、現状の翻訳エンジンにおいて法令を翻訳した結果を基に、誤訳パターンを抽出した。誤訳パターンとは、英作文としての誤訳だけではなく、法令翻訳のルールに準拠していない状態も含む。

（2）翻訳システムの構築

誤訳パターンに対し、翻訳品質改善プログラムを実装した翻訳システムを構築した。

（3）翻訳サンプルの生成

法務省指定の法令データを（2）で構築した翻訳システムで機械翻訳し、翻訳サンプルを生成した。

（4）調査分析

（3）で生成した翻訳サンプルデータを用いて、翻訳結果の考察及びより効率的かつ効果的なAI学習の強化の方法や、より精度の高い翻訳品質改善のためのプログラムについて調査・検討を実施した。

（5）調査報告書の作成

上記（1）ないし（4）の成果を本調査報告書に取りまとめた。

3.2 機械翻訳結果の誤訳パターン抽出

令和3年度に法務省において実施した機械翻訳の活用・導入に向けた検証では、3本の法令を対象として機械翻訳により英訳を作成し、その結果について誤訳パターンを抽出した。

本業務において取り扱う具体的な改善事項を検討するに当たり、抽出した誤訳パターンについての分類を実施した（表3.1参照）。

本業務で翻訳品質改善を行う事項は、複数回出現した誤訳パターンのうち、プログラム実装の実現性があるものを凸版印刷にて抽出し、法務省と協議の上、決定した。（誤訳パターンのうち、「法令用語日英標準対訳辞書」及び「法令翻訳の手引き」の指針に関連すると思料されるものについては、凸版印刷において抽出し、「関連指針」欄に記載した。）

表 3.1 誤訳パターンの分類

項目	内容	令和3年度の翻訳結果で見られた誤訳パターン	関連指針
A.正確さ	原文に対する訳文の正確さ		
a.誤訳	原文と訳文の意味が異なる。	・数値の誤り(三桁の数字の誤訳)	手引き 31 「〇〇大臣」と「〇〇省」の誤訳に注意
b.訳語の追加/削除	原文に含まれる単語に対応する訳語が訳文に含まれない。	・主語省略の影響による湧き出し(不適切な主語の補い) ・訳漏れ	
B.用語	用語の適切さ		
a.不適切な用語訳	規定の用語訳が使われていない。	・辞書への非準拠 ・「当該」の英訳として said は使わない(手引き 25)というルールへの非準拠	法令用語日英標準対訳辞書 手引き 1 法令名中にある「法律」の英訳は"Act" 手引き 2 法令名中にある「法律」の英訳に"Law"が使える例外 手引き 3 法令表題の Act に定冠詞"The"は不要 手引き 4 「…に関する法律」の英訳は"Act on…" 手引き 21 引用「法令名」の正確性・統一性保持 手引き 22 府省庁の「部署名」「役職名」の一貫性保持 手引き 23 「条約名」の英語の正確性・統一性保持 手引き 25 「当該」の英訳 手引き 28 「…の既定の/による/により」の使い分け及び用法 手引き 29 「…定める」「…規定する」の英訳 手引き 32 「会議」「総会」の英訳についての注意
b.用語訳の不統一	同一法令中で用語訳が統一されていない	・訳語の不統一	手引き 19「目次」「定義語」「訳語」の一貫性保持
C.スタイル	スタイルへの準拠		
a.条・項・号等の英訳表記		手引きへの非準拠 ・1項しかないものに(1)と入っている ・(i)が大文字になっている。	手引き 5 条・項・号の英訳表記 手引き 6 条・号の枝番号 手引き 7 条・項・号の表記順序

		・(1)の抜け ・「〇項から〇項まで」の表記では through を用いる(手引き 7)	
b.号の文頭・文末表記		・最終号の1つ前の号の文末の表記(and 又は or) ・文末にはピリオドをつける ・各号の先頭文字は小文字(手引き 8)	手引き 8 各号の冒頭及び末尾の表現
c.「等」の英訳		・条見出し中の「等」の英訳 “etc.”は原則不要(手引き 16)	手引き 15 法令名中の「等」の英訳” etc.”は原則不要 手引き 16 条見出し中の「等」の英訳” etc.”について 手引き 17 条文中の「等」は英訳する
d.大文字・小文字表記		・定義語は小文字にする(手引き 20)	手引き 20 定義された語や略称の表記
e.定義規定の shall		・定義規定に shall は使わない(手引き 34)	手引き 34 “shall”の用法②
f.性別を示す表現		・him/her など性別を示す表現は使用しない(手引き 37)	手引き 37 人称代名詞の使い分け
g.条件を示す表現		・in cases where は原則使わない(手引き 38)	手引き 38 条件を示す場合(「場合」・「とき」)
h.スタイルの不統一	同一法令中でスタイルが統一されていない	・大文字小文字の不統一	手引き 24 英訳スタイルの一貫性保持

以上を基に、本業務においては、表 3.2 に記載の誤訳パターンについて対応することとした。

表 3.2 本業務において改善対象とする誤訳パターン

番号	誤訳パターン	備考
1	主語省略の影響による湧き出し	前記2課題①に該当
2	定義規定の”shall”の訳出	前記2課題②に該当
3	性別を示す表現の訳出	前記2課題②に該当
4	条・項・号等の英訳表記	前記2課題②に該当
5	号の文頭・文末表記	前記2課題②に該当
6	定義語は小文字にする	前記2課題②に該当
7	「〇項から〇項まで」の表記	前記2課題②に該当
8	条見出し中の”etc.”の訳出	前記2課題②に該当
9	数値の誤り	翻訳の正確さ・信頼性の観点から実施
10	法令用語日英標準対訳辞書への非準拠	前記課題②に該当

3.3 翻訳システムの構築

3.3.1 翻訳システム

翻訳システムとは、システムに入力された翻訳原文（例：日本語）に対し、翻訳エンジンが翻訳処理を実行し、翻訳文（例：英語）が出力される仕組みのことである（図 3.1(a)参照）。実用される翻訳システムにおいては、翻訳品質の改善を目的として、この翻訳エンジンの前段階と後段階に何らかの処理を組み込むことがある（図 3.1(b)参照）。翻訳前後処理を組み込む目的は様々であるが、例えば、翻訳対象文書の分野に適した翻訳出力を得るための整形処理や、翻訳の誤りを低減するための処理が実装されることがある。

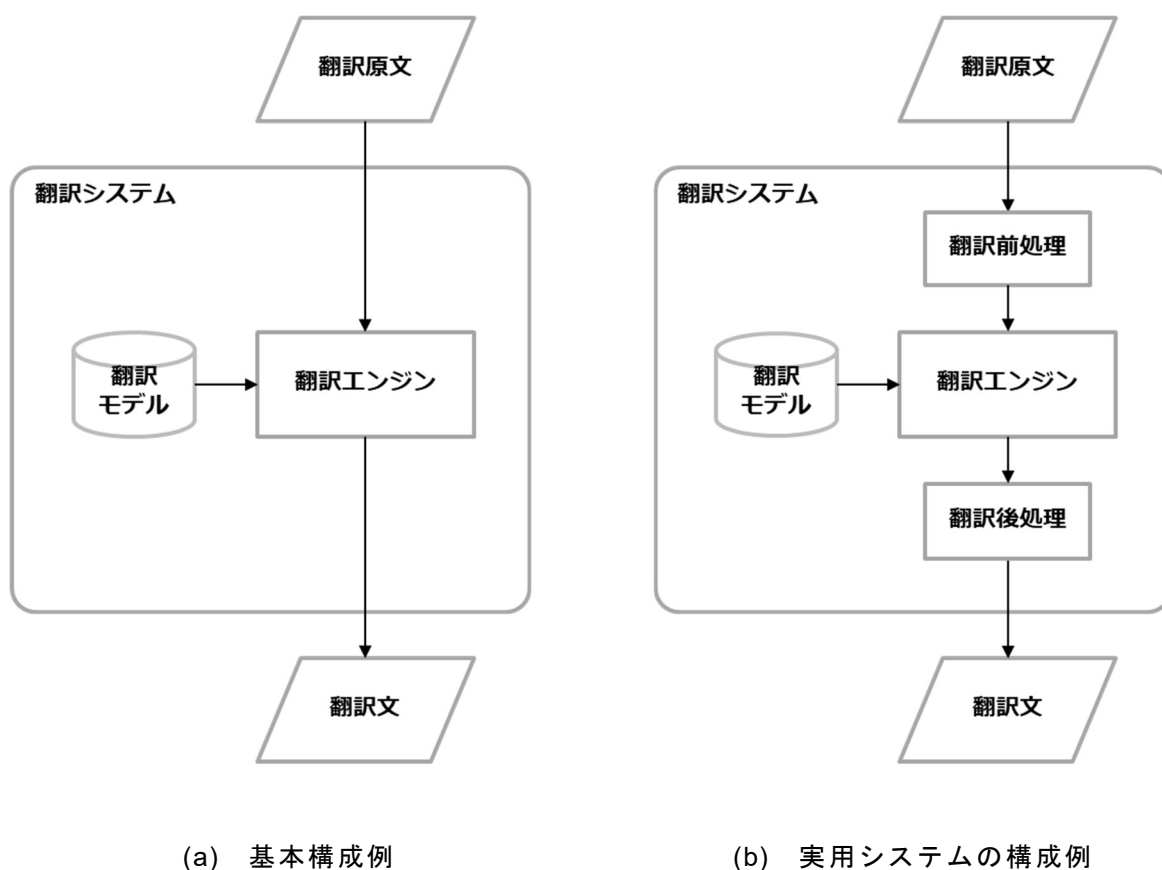


図 3.1 翻訳システムの構成例

3.3.2 翻訳エンジンの概要

翻訳エンジンは、原言語（例：日本語）で記述された入力テキストを目的言語（例：英語）のテキストに翻訳して出力するものである。翻訳エンジンを実現する機械翻訳技術は、長年の研究開発の積み重ねの中で様々な方式が提案されてきた（表 3.3 参照）。これらの方式のうち、現在は、深層学習（ディープラーニング）を利用したニューラル機械翻訳（A I 技術を利用していることから A I 翻訳とも呼ばれる）が主流となっている。深層学習とは、脳の神経回路の一部を模したニューラルネットワークを用いた機械学習¹の手法の 1 つであり、多層構造のニューラルネットワークを用いることが特徴となっている。

表 3.3 機械翻訳技術の方式

方式	内容
ルールに基づく機械翻訳 Rule-Based Machine Translation: RBMT	言語の専門家が作成したルールや辞書に基づいて訳文を出力
用例に基づく機械翻訳 Example-Based Machine Translation: EBMT	翻訳用例（対訳コーパス ² ）を組み合わせることで訳文を出力
統計機械翻訳 Statistical Machine Translation: SMT	対訳コーパスを統計的に解析し、入力された原文に対して対訳となる確率が最も高い訳文を生成
フレーズベースの SMT Phrase-Based SMT: PBSMT	SMT にフレーズ概念を導入
ニューラル機械翻訳 Neural Machine Translation: NMT	ニューラルネットワークを多層に結合した深層学習（ディープラーニング）を導入

表 3.3 に示す方式は、いずれも、ルールや対訳コーパスといったデータと、翻訳エンジンプログラムで実行される統計処理やニューラルネットワークをモデルとしたアルゴリズム（データからパターンや特徴を発見し表現するための仕組み）との組み合わせで実現される。データの改善とアルゴリズムの改善のそれぞれを実施することにより、翻訳システム全体の性能（翻訳結果の品質）を向上させることができる。

本業務においても、ニューラル機械翻訳を用いて、翻訳システムを構築したものである。

¹ 機械（コンピューター）が自動で学習し、データの背景にあるルールやパターンを発見する方法。

² 機械翻訳の学習データとして利用するためなどに構築された、異なる言語の文と文が対訳の形でまとめられたもの。

3.3.3 翻訳モデルの概要

(1) 翻訳モデルの構築

ニューラル機械翻訳では、図 3.2 に示すように、対訳コーパスを使用して、深層学習を行うことによって、翻訳モデルと呼ばれる一種のデータベースを構築する。

構築された翻訳モデルを利用して翻訳エンジンが翻訳処理を実行するが、データの改善の実施の観点では、翻訳処理の性能は、一般的には、翻訳モデルの構築に使用する対訳コーパスの規模（対訳文の数）に依存することが知られている。実用化されている翻訳サービスの翻訳モデルの構築で使用されている対訳コーパスの規模は多くの場合は非公開であるが、数百万文対から数億文対以上の対訳コーパスを利用して翻訳モデルが構築されていると推測される。高性能な翻訳モデルを構築するには大規模な対訳コーパスが必要となるが、本業務でベースとしたニューラル翻訳エンジンを開発している国立研究開発法人情報通信研究機構（以下、「NICT」という。）では、「翻訳バンク」の取り組みを通じて大量の対訳コーパスの集積を行い、翻訳モデルを構築している。

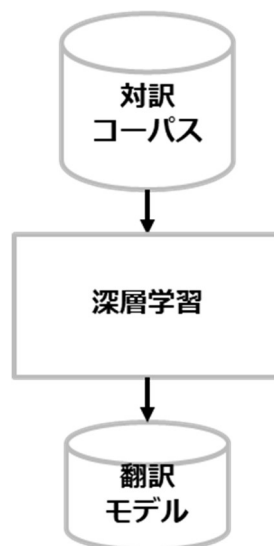


図 3.2 対訳コーパスを用いた翻訳モデルの構築

(2) 追加学習による翻訳モデルの構築

上記 (1) の構築方法のほか、あらかじめ構築済みの翻訳モデルをベースに、特定の対訳コーパスを追加学習³（アダプテーション）することで翻訳モデルを構築する手法も知られている（図 3.3 参照）。追加学習の手法では、例えば法令分野に特化した対訳コーパスを使用することで、ベースとなる翻訳モデルによる翻訳出力が、使用した対訳コーパスの分野に適した出力結果となるように適応させることができる。大規模な対訳コーパスによって構築された翻訳モデルをベースとすることができると、使用する対訳

³ 対訳コーパスを使用し、あらかじめ構築済みの翻訳モデルを機械学習によって調整する方法のこと。

コーパスの対訳文数が比較的少ない場合でも、分野に適応した翻訳モデルを構築することができる（図 3.4 参照）。

すなわち、追加学習をさせる対訳コーパスの違いにより、異なる翻訳結果が得られることとなる。

本業務では、構築済みの翻訳モデルに J L T から収集した対訳コーパスを追加学習させた。

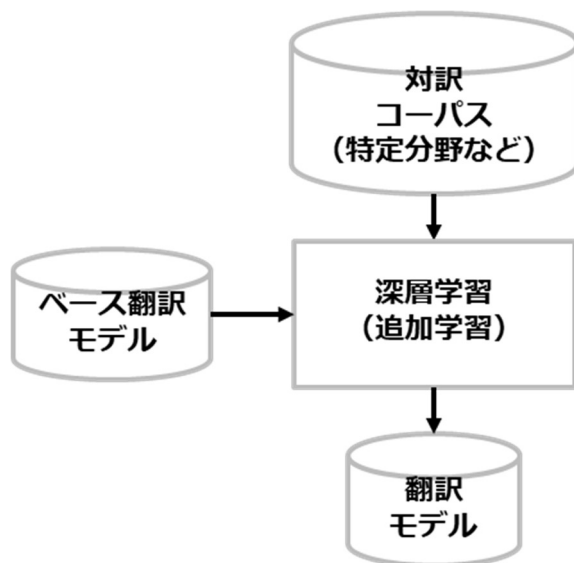


図 3.3 追加学習による翻訳モデルの構築

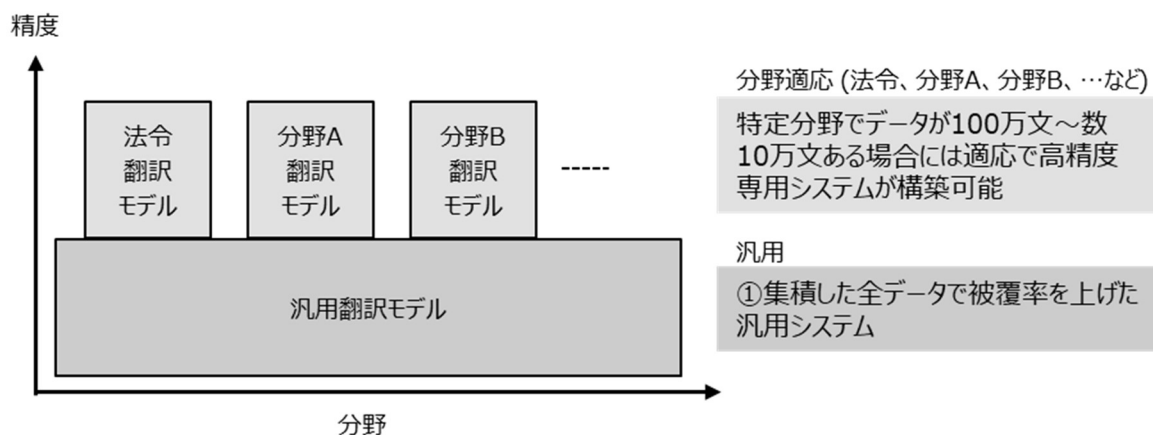


図 3.4 翻訳モデルの分野適応

3.3.4 対訳コーパス及びその整備方法

高性能な翻訳モデルを構築するには、品質の高い対訳コーパスを大量に準備することが必要不可欠である。しかし、対訳コーパスに誤りが含まれていると、その誤りも正しいものとして翻訳モデルが学習されてしまう。ここでいう、「品質の高い対訳コーパス」とは、余分な空白や記号など不要な文字が含まれていないだけでなく、

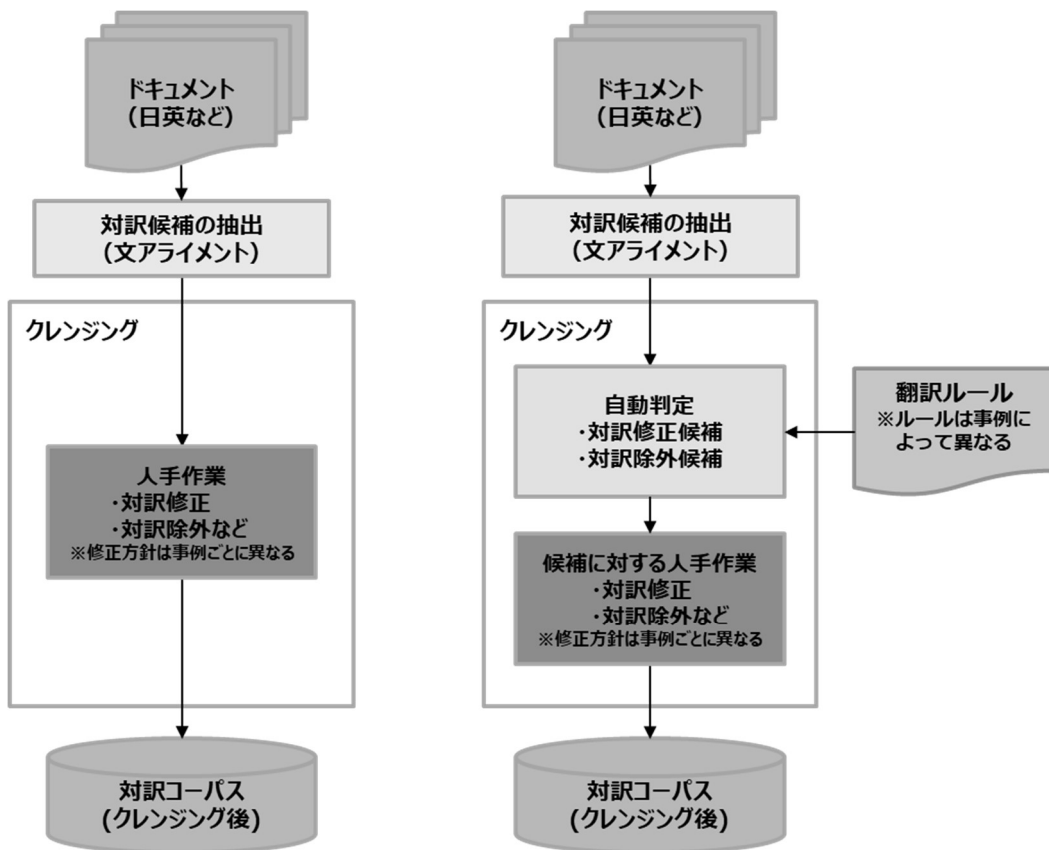
- ・対訳文が過不足のない翻訳となっている
- ・当該分野では避けるべき表現などが対訳文に含まれていない

といった条件が求められる。

対訳コーパスは文単位で翻訳原文と翻訳文が対になっている必要があるが、例えば日本語版と英語版とで一つずつの文書ファイルとなっているような場合、それぞれの文書ファイルからテキストデータを読み込み、対訳対の候補を抽出する（「アライメントを取る」と呼ばれることがある。）必要がある。抽出された対訳候補は必ずしも対訳となっていない場合があることから、対訳となっていない訳文について修正や除外をするなどの処理が必要となる。なお、不適切な対訳文を対訳コーパスから取り除く、あるいは適切に修正する作業は「クレンジング」と呼ばれており、対訳コーパスの構築とクレンジング作業のワークフローを図 3.5(a)に示す。

対訳コーパスが小規模の場合、クレンジング作業は人手で実施することも可能であるが、対訳コーパスが大規模になるとクレンジング作業の自動化技術の導入が不可欠となる。図 3.5(b)に示すように自動化技術を導入することにより、クレンジング作業の効率化を図ることができる。本業務で用いた対訳コーパスについても、自動化技術を導入し整備した。

このように、対訳候補の抽出とクレンジングを行った対訳コーパスを追加学習させ、翻訳モデルを構築するものであるが、通常、翻訳モデルの構築作業を実施する場合は、構築時に指定する各種パラメータを変化させて構築作業を繰り返し実施し、最適な翻訳品質が得られるパラメータを決定する作業を伴う。ここで、パラメータにはクレンジングの条件（どのような条件で除外するかなど）も含まれる。



(a) 一般的なフロー

(b) 自動化を導入したフロー

図 3.5 対訳コーパスの構築とクレンジングワークフロー

3.3.5 本調査研究業務で構築した翻訳システムの概要

(1) 翻訳エンジン

本システムでは、NICTの法令契約モデル（NICTの汎用翻訳エンジンに対し、法令や契約分野の対訳コーパスを利用して追加学習し翻訳精度を高めた翻訳エンジン）をベースとして利用した。

(2) 翻訳品質改善プログラム

前述した誤訳パターンに対応する（翻訳品質改善）ため、翻訳エンジンに追加学習をするとともに、翻訳エンジンの翻訳前処理及び翻訳後処理として翻訳品質改善プログラムを実装した。

(3) 設置場所

クラウドサーバーとしてAWS（Amazon Web Service）を利用して構築した。アジアパシフィック（東京）のリージョンを利用しており、データセンターは日本国内に設置されている。

(4) 入出力形式

本システムは、日本語で記述されたMicrosoft Word形式及びMicrosoft Excel形式ファイルを入力対象とし、英語に機械翻訳された結果を同じファイル形式で出力することとした。

(5) 翻訳システムの処理フロー

本システムの処理フローを図3.6に示す。翻訳品質改善プログラムは、必要に応じて翻訳前処理及び翻訳後処理として実装する。

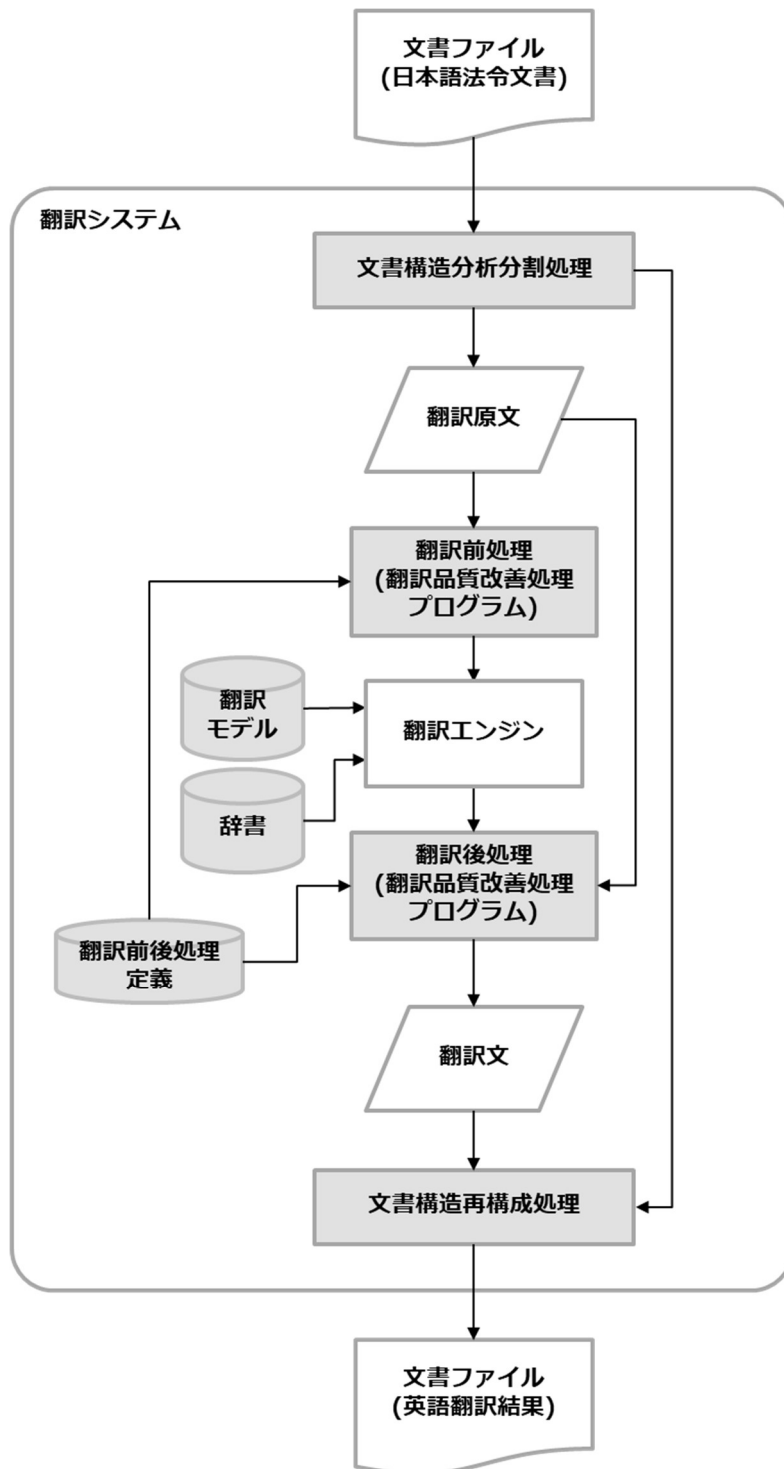


図 3.6 翻訳システムの処理フロー

3.4 翻訳品質改善プログラム

3.4.1 概要

表 3.2 で示した改善事項として選定した誤訳パターンに対し、下記の翻訳品質改善プログラムを実装することとした（表 3.4 参照）。

なお、改善事項として選定した誤訳パターンのうち、項番 10 「法令用語日英標準対訳辞書への非準拠」については、辞書登録（単語リストの事前登録）により改善を図ることとし、翻訳品質改善プログラムには含めていない。

表 3.4 誤訳パターンに対する翻訳品質改善プログラム

番号	誤訳パターン	改善内容	翻訳品質改善プログラムの名称	翻訳品質改善の手法
1	主語省略の影響による湧き出し	“Prime Minister”の湧き出しを抑制する	翻訳品質改善プログラム【1】 翻訳品質改善プログラム【5】	追加学習による改善 前後処理プログラムによる改善
2	定義規定の“shall”の訳出	“shall”の訳出抑制	翻訳品質改善プログラム【1】	追加学習による改善
3	性別を示す表現の訳出	性別を示す表現（“his/her”など）の訳出抑制	翻訳品質改善プログラム【1】	追加学習による改善
4	条・項・号の英訳表記	条・項・号の記載方法について、法令翻訳のルールに準拠させる	翻訳品質改善プログラム【2】	前後処理プログラムによる改善
5	号の文頭・文末表記	号の文頭・文末の記載方法について、法令翻訳のルールに準拠させる	翻訳品質改善プログラム【2】	前後処理プログラムによる改善
6	定義語は小文字にする	“hereinafter”で定義されている用語を小文字に統制する	翻訳品質改善プログラム【3】	前後処理プログラムによる改善
7	「○項から○項まで」の表記	「○項から○項まで」の表記では“through”を用いる	翻訳品質改善プログラム【4】	前後処理プログラムによる改善
8	条見出し中の“etc.”の訳出	条見出し中の“etc.”の訳出抑制	翻訳品質改善プログラム【6】	前後処理プログラムによる改善
9	数値の誤り	数値の誤り（三桁の数字の誤訳）の改善	翻訳品質改善プログラム【7】	前後処理プログラムによる改善

<追加学習による改善>

翻訳品質改善プログラム【1】 追加学習

<前後処理プログラムによる改善>

- 翻訳品質改善プログラム【2】 条・項・号の英訳表記、号の文頭・文末表記
- 翻訳品質改善プログラム【3】 定義語は小文字にする（大文字・小文字の表記）
- 翻訳品質改善プログラム【4】 「○項から○項まで」の表記（定型表現への対応）
- 翻訳品質改善プログラム【5】 主語省略の影響による湧き出し（湧き出しへの対応）
- 翻訳品質改善プログラム【6】 条見出し中の“etc.”の訳出（訳出への対応）
- 翻訳品質改善プログラム【7】 数値の誤りへの対応

3.4.2 追加学習による改善（翻訳品質改善プログラム【1】）

（1）改善対象とする誤訳パターン

- ① 主語省略の影響による湧き出し
- ② 定義規定の“shall”の訳出
- ③ 性別を示す表現の訳出

1) 実施方法

JLT対訳コーパスをクレンジングし、クレンジングした対訳コーパスを使用して翻訳モデルの追加学習（アダプテーション）を実施する。

なお、①について、湧き出しの問題については“Prime Minister”に限った話ではないが、有効性が認められれば、将来的に横展開し、同様の事例に対して実施することが期待できるため、本業務では“Prime Minister”に絞って実施した。

2) 期待される効果

- ① 主語を補完する際に省略された主語（例：「法務大臣」）とは異なる主語（例：“Prime Minister”）が訳文に湧き出す現象を抑制する。
- ② 英訳内に shall が出現しづらくなる。
- ③ 単数の人称代名詞を訳出する際に、性別を示す表現が出現しづらくなる。

3) 期待される効果の説明

①について

JLT対訳コーパスの各対訳において、日本語側における「大臣」と英語側における“Minister”の出現数が一致しない場合がある（表 3.5 参照）。特に日本語側の出現数が少ない対訳文を翻訳モデルの訓練に多く使用すると、「湧き出し」の問題が起きてくる。

この問題に対応するため、JLT対訳コーパスの各対訳において、日本語側における「大臣」と英語側における“Minister”の出現数が一致しない場合に、当該対訳を対訳コーパスから除外して翻訳モデルのアダプテーションを実施することにより、翻訳品質の改善を図る。

表 3.5 日本語側における「大臣」と英語側における“Minister”の出現数が一致しない対訳例

日本語	英語
4 内閣総理大臣は、政令で定めるところにより、定期的に、及び任命権者の求めがある場合その他必要があると認める場合には随時、適格性審査を行い、幹部候補者名簿を更新するものとする。	(4) The Prime Minister is to implement the eligibility screening and renew the executive candidates list periodically, whenever an appointer requests to implement it and whenever the Prime Minister finds it to be necessary, pursuant to the provisions of Cabinet Order.

このような対訳文を2022年6月に収集したJLT対訳コーパスでカウントしたところ全304,486文中4,592文あり、1.5%ほどとなる。なお、この数字は実作業の際の精査にともない、変更されるものである。

②③について

例として、人称代名詞の対応のうち、“he/she”について説明する。

JLTに掲載されている英訳法令のうち、比較的古い時期に登録された法令には、“he/she”などの性別を示す表現が含まれている場合がある。これらの表現が含まれる対訳コーパスを使用して翻訳モデルを構築すると、翻訳結果に性別を示す表現が訳出されてしまう可能性がある。

なお、2022年6月に収集したJLT対訳コーパスで“he/she”の表現が含まれる行数をカウントしたところ、表3.6のような結果となった。

表 3.6 JLT対訳コーパス

収集時期	ファイル数	行数	“he/she”が含まれる行数
2022年6月	684	304,486	510 (0.17%)

この問題の改善に向けて、NICTの法令契約モデルに、2022年6月に収集したJLT対訳コーパスからさらに“he/she”を含む人称代名詞を削除したコーパスを用いてアダプテーションを行い、翻訳モデルを試作した。2022年5月31日版のNICTの法令契約翻訳モデルの翻訳結果と、この試作モデルの翻訳結果を表3.7に示す。

二つの翻訳結果を比較すると、同じ文を翻訳しても、2022年6月のコーパスを用いて試作した翻訳モデルでは、性別を示す表現について翻訳品質が改善されている。

このように、翻訳ルールに即していないと思われる対訳をJLT対訳コーパスから除外して翻訳モデルのアダプテーションを実施することにより、翻訳品質の改善が図られる。

表 3.7 法令契約モデルと J L T 対訳コーパスによる試作翻訳モデルの比較

原文	翻訳結果 (NICT法令契約翻訳モデル)	翻訳結果 (JLT対訳コーパス試作翻訳モデル)
<p>四 法第五条第七号に規定する役員等(以下「役員等」という。)が、自ら事業を営み、若しくは事業を営む者の使用人となり、又は他の法人の役員若しくは使用人となっているときは、当該役員等の氏名並びにこれを使用する者の氏名又は当該法人の商号若しくは名称、その住所及び当該事業の種類又は当該法人の業務の種類</p>	<p>(iv) when an officer, etc. prescribed in Article 5, item (vii) of the Act (hereinafter referred to as an "officer, etc.") operates a business by himself / herself, becomes an employee of a person who operates a business, or becomes an officer or employee of another juridical person, the name of said officer, etc., as well as the name of the person who employs him / her, or the trade name or name, address of said juridical person, and the type of said business or the type of business of said juridical person ;</p>	<p>(iv) if an officer, etc. prescribed in Article 5, item (vii) of the Act (hereinafter referred to as an "officer, etc.") operates a business by themselves or serves as an employee of a person who operates a business, or serves as an officer or employee of another corporation, the name of the officer, etc., as well as the name of the person who employs the officer, etc., or the trade name or name of the corporation, the address of the person, and the type of the business or the type of the business of the corporation ;</p>
<p>第十一条 債権管理回収業の実施業務に従事する者は、その業務を行うに当たり、債権回収会社の商号、許可番号及び自己の氏名を記載した身分を示す証明書を携帯し、相手方の請求があったときは、これを提示しなければならない。</p>	<p>Article 11 (1) A person engaged in the implementation of a claim management and collection business shall, in carrying out his / her business, carry an identification card containing the trade name and license number of the claim management and collection company and his / her own name, and present it at the request of the other party.</p>	<p>Article 11 (1) A person engaged in the implementation business of claim management and collection must, in carrying out the business, carry an identification card stating the trade name and license number of the claim management and collection company and the person's own name, and present it at the request of the counterparty.</p>
<p>第六条 法務大臣は、前条の認証の申請をした者(以下「申請者」という。)が行う当該申請に係る民間紛争解決手続の業務が次に掲げる基準に適合し、かつ、申請者が当該業務を行うのに必要な知識及び能力並びに経理的基礎を有するものであると認めるときは、当該業務について認証をするものとする。</p>	<p>Article 6 (1) The Minister of Justice shall certify private dispute resolution services provided by a person who has applied for certification under the preceding Article (hereinafter referred to as to as "applicants") when he / she finds that such services conform to the following standards and that the applicants have the necessary knowledge, skills and financial basis to perform such services.</p>	<p>Article 6 (1) The Minister of Justice is to certify the private dispute resolution services provided by a person who has applied for the certification referred to in the preceding Article (hereinafter referred to as the "applicant"), if the Minister finds that the services pertaining to the application conform to the following standards and that the applicant has the necessary knowledge, skills and financial base to perform the services :</p>

(2) クレンジング実施内容

J L T から取得した対訳コーパスを対象として、表 3.8 に示す事項について自動処理によるクレンジングを実施した。クレンジングによって、有効な対訳 281,498 文のうち 7,623 文 (2.7%) の対訳文を除外した。

表 3.8 クレンジングの実施内容

ステップ	説明	除外ルール等	除外した文数	合計文数
-	JLTサイトから取得した対訳コーパス (取得: 2022年6月)			303,812
0	予備クレンジング	・ 対訳として収集されていない対訳文（主に別表）を除外	22,284	281,528
1	条項号の分離	・ 条項号の情報が日英で不整合な対訳文を除外	30	281,498
2	対象事項のクレンジング	<ul style="list-style-type: none"> ・ him/herなど性別を示す表現は使用しない（手引き37） - he/him/she/herなど、性別を示す表現が英文に含まれている対訳を除外（836文） ・ 定義規定にshallは使わない（手引き34） - 英文にshallが含まれる対訳を除外（6,557文） ・ 湧き出しへの対応 - 英文の“prime minister”の出現数が日文の「総理大臣」の出現数より多い場合に対訳を除外、あるいは、英文の“minister of XXX”の出現数が日文の「XXX大臣」の出現数より多い場合に対訳を除外（752文） 	7,623	273,875

※補足 2021年9月に収集したJLTコーパスでは、he/sheなどが含まれる対訳の文数は2,663文であった。

3.4.3 前後処理プログラムによる改善（翻訳品質改善プログラム【2】～【7】）

(1) 翻訳品質改善プログラム【2】（条・項・号の英訳表記、号の文頭・文末表記）

1) 改善対象とする誤訳パターン

条・項・号の英訳表記について、法令翻訳のルールに準拠させる。

具体的には、以下の事象について改善する。

- ・ 項の訳抜け（(1)表記抜け）
- ・ 号の表記については小文字で記載するべきであるところ、大文字になる事象
- ・ 1項しかない場合は（1）の表記は不要であるところ、(1)と入る事象

2) 実施方法

翻訳原文を翻訳する前に、条・項・号を取り除き、本文のみ翻訳する。その後、条・項・号記述ルールに従って取り除いた条・項・号を翻訳し、本文翻訳結果と条・項・号翻訳結果を統合する方法で修正する。各翻訳結果を適切に統合して修正するために、当該文書における条・項・号の構造を事前に解析しておく。

3) 期待される効果

条・項・号の英訳表記について、法令翻訳ルールに則り、適切に訳出される。

4) 期待される効果の説明

表 3.9 に条・項・号の記述ルールを示す。

表 3.9 条・項・号の記述ルール

	日本語表現	英語表現	英語詳細
条	第二条	Article 2	項がある場合、文末は: 項がない場合、文末は、 枝番号はハイフンのあとに算用数字
項	算用数字（1, 2）	(1), (2)	2項以上ある場合は条の最初に(1)を訳出する 1項しかない場合は、(1)は訳出しない
号	漢数字（一、二）	(i), (ii)	最後の号以外、文末は; 最後の号の一つ前の文末は"; and" または"; or" 最後の号の文末は、 枝番号はハイフンのあとに算用数字 先頭文字は小文字
号以下	10/a	(a), (b), .. (aa)	最後の号以下以外、文末は; 最後の号以下の一つ前の文末は"; and" または"; or" 最後の号以下の文末は、
	(1), (2)	1., 2.	
	(i), (ii)	i., ii.	

本プログラムを適用する前には、以下の対訳コーパスを作成し、事前にアダプテーションを行う。

- ・ 日本語側から条・項・号を取り除き、本文のみにする。

・英語側から条・項・号並びに文末の.:;等を取り除き、本文のみにする。

以下に本プログラムの動作を簡単に述べる。

1. 翻訳原文をすべて読み込み、条・項・号の構造を解析する。
2. 翻訳原文から条・項・号を取り除き、法令本文のみを翻訳する。
3. 法令本文の翻訳結果に対し、条・項・号の記述ルールに基づいた英語表現を付与する。

ここで、条・項・号の構造解析では、入力された法令から分離した条・項・号のテキストについて、3回の処理パスに分けて解析処理を実施し、判定した結果を記憶（マーク）した。記憶した情報を使用して、本文の翻訳結果に条・項・号や文末の punctuation 等を付与した。処理パスごとの解析内容は以下の通りである。

パス0：条・項・号と本文を分離

パス1：条・項・号について、号部分を抜き出して、階層ごとに位置（最後の一つ前、最後、それ以外）をマーク

パス2：さらに2項以降があるものについて条にマーク。最上層の号を参照する条または項にマーク（:処理のため）

パス3：最後の一つ前の号について、and または or を推定（「いずれか」が本文中にあれば or それ以外は and）

図 3.7 に示す「民事訴訟費用等に関する法律」の第9条第1項から第5項までについて、条・項・号の構造解析処理を行った例を表 3.10 に示す。

...
第九条 手数料が過大に納められた場合においては、裁判所は、申立てにより、決定で、過大に納められた手数料の額に相当する金額の金銭を還付しなければならない。
2 前項の規定にかかわらず、支払督促若しくは差押処分の申立てで手数料又は別表第二の上欄に掲げる事項の手数料が過大に納められた場合の還付は、申立てにより、裁判所書記官が行う。
3 次の各号に掲げる申立てについてそれぞれ当該各号に定める事由が生じた場合においては、裁判所は、申立てにより、決定で、納められた手数料の額（第五条の規定により納めたものとみなされた額を除く。）から納めるべき手数料の額（同条の規定により納めたものとみなされた額を除くものとし、民事訴訟法第九条第一項に規定する合算が行われた場合における数個の請求の一に係る手数料にあつては、各請求の価額に応じて案分して得た額）の二分の一の額（その額が四千元に満たないときは、四千元）を控除した金額の金銭を還付しなければならない。
一 訴え若しくは控訴の提起又は民事訴訟法第四十七条第一項若しくは第五十二条第一項の規定若しくはこれらの規定の例による参加の申出の口頭弁論を経ない却下の裁判の確定又は最初にすべき口頭弁論の期日の終了前における取下げ
二 民事調停法による調停の申立て却下の裁判の確定又は最初にすべき調停の期日の終了前における取下げ
三 労働審判法による労働審判手続の申立て却下の裁判の確定又は最初にすべき労働審判手続の期日の終了前における取下げ
四 借地借家法第四十一条（大規模な災害の被災地における借地借家に関する特別措置法第五条第二項（同条第四項において準用する場合を含む。）において準用する場合を含む。以下この号において同じ。）の事件の申立て、借地借家法第四十一条の事件における参加の申出（申立人として参加する場合に限る。）又はその申立て若しくは申出についての裁判に対する抗告（次号に掲げるものを除く。）の提起却下の裁判の確定又は最初にすべき審問の期日の終了前における取下げ
五 上告の提起若しくは上告受理の申立て又は前号の申立て若しくは申出についての裁判に対する非訟事件手続法（平成二十三年法律第五十一号）第七十四条第一項の規定による再抗告若しくは同法第七十五条第一項の規定による特別抗告の提起若しくは同法第七十七条第二項の規定による抗告の許可の申立て原裁判所（抗告の許可の申立てにあつては、その申立てを受けた裁判所。以下この号において同じ。）における却下の裁判の確定又は原裁判所が上告裁判所若しくは抗告裁判所に事件を送付する前における取下げ
4 前項の規定は、数個の請求の一部について同項各号に定める事由が生じた場合において、既に納めた手数料の全部又は一部がなお係属する請求についても納められたものであるときは、その限度においては、適用しない。同項第五号に掲げる申立てについて同号に定める事由が生じた場合において、既に納めた手数料の全部又は一部がなお係属する他の同号に掲げる申立てについても納められたものであるときも、その限度において、同様とする。
5 支払督促の申立てについて、却下の処分確定又は支払督促の送達前における取下げがあつた場合においては、裁判所書記官は、申立てにより、第三項の規定に準じて算出した金額の金銭を還付しなければならない。ただし、前項前段に規定する場合には、その限度においては、この限りでない。
...

図 3.7 条・項・号の構造解析処理の入力例

表 3.10 条・項・号の構造解析処理例

(a) 構造解析結果の一例

条項号	判定したマーク
第九条	F
2	
3	S1
一	b
二	b
三	b
四	B1,A
五	E
4	
5	

(b) マークの意味

マーク	説明
F	2項以降の記述がある条
Sn	階層構造の開始
b	階層構造の中間
Bn	階層構造の中間（終了のひとつ前）
A	付与する接続詞がand
O	付与する接続詞がor
E	階層構造の終了
	※"n"は、ひとつの条文内の階層番号を示す

(2) 翻訳品質改善プログラム【3】(定義語は小文字にする(大文字・小文字の表記))

1) 改善対象とする誤訳パターン

定義語は小文字にする

事象：定義された語や略称の翻訳において、(hereinafter … “xxxx”)と訳される際、xxxx は小文字とすべきだが、xxxx が小文字になっていない場合がある。

2) 実施方法

翻訳後処理で定義箇所(hereinafter … “xxxx”)が存在する部分を検索し、xxxx を小文字に変換する。また、定義箇所以外の文書内の xxxx も小文字にする。

3) 期待される効果

hereinafter で定義された語が小文字で適切に訳出される。

4) 期待される効果の説明

翻訳後処理で、翻訳文中の (hereinafter … “xxxx”)を正規表現を使用して検出し、xxxx を小文字に変換する。その後、翻訳文全体における xxxx を小文字に変換する。

(3) 翻訳品質改善プログラム【4】(「○項から○項まで」の表記(定型表現への対応))

1) 改善対象とする誤訳パターン

「○項から○項まで」の表記では“through”を用いる
事象:「○項から○項まで」の英訳に、“through”ではなく“to”が使用されている場合がある。

2) 実施方法

対象とする語句と置換条件等を定義ファイルに記述し、翻訳後処理で当該語句を置換、あるいは、削除する。

- ・ 翻訳結果に“paragraph (m) to (n)”が出現したら“paragraphs (m) through (n)”に置換する。

3) 期待される効果

「○項から○項まで」の英訳について、“paragraphs ○ through ○”の形で適切に訳出される。

4) 期待される効果の説明

対象とする語句は正規表現を使用して記述する。

- (4) 翻訳品質改善プログラム【5】(主語省略の影響による湧き出し(湧き出しへの対応))
- 1) 改善対象とする誤訳パターン
主語省略の影響による湧き出し
 - 2) 実施方法 ※前述のとおり、Prime Ministerの湧き出しに絞って実施する。
対象とする語句と置換条件等を定義ファイルに記述し、翻訳後処理で当該語句を置換、あるいは、削除する。
 - ・翻訳結果に“Prime Minister”が出現しているが翻訳原文に「総理大臣」が含まれていない場合、“Prime Minister”を“Minister”に置換する。
 - 3) 期待される効果
不適切な湧き出し語句(“Prime Minister”)が翻訳結果から除外される。
 - 4) 期待される効果の説明
上記事象は、まず、翻訳品質改善プログラム【1】で改善を試みるが、当該事象を完全に排除することは難しいと思われる。そこで、二段階の対策として、本対応を実装するものである。

(5) 翻訳品質改善プログラム【6】(条見出し中の“etc.”の訳出 (訳出への対応))

1) 改善対象とする誤訳パターン

条見出し中の“etc.”の訳出

事象：日本語原文において条見出し中に「等」が含まれていても、条見出しにおいては訳出しないこととなっているが、訳出される場合がある。

2) 実施方法

条見出し英訳中の“etc.”を削除する。

3) 期待される効果

条見出しの英訳中に“etc.”が訳出されなくなる。

4) 期待される効果の説明

対象とする語句は正規表現を使用して記述する。

(6) 翻訳品質改善プログラム【7】(数値の誤りへの対応)

1) 改善対象とする誤訳パターン

数値の誤り

事象：三桁以上の数字を漢数字から算用数字に翻訳するに当たり、誤訳される場合がある。

2) 実施方法

翻訳前処理で条の数字部分を算用数字に置き換えてから翻訳を実行する。

※アダプテーションは算用数字に置き換える前のコーパスで実施した。

3) 期待される効果

本文内で参照される条の数字が適切に訳出される。

4) 期待される効果の説明

ベースとした翻訳システムについて予備評価を行ったところ、条の数字部分が「十条」、「十六条」、「十七条」、「十八条」などで終わる場合、誤訳する場合があることが分かった。

そこで、本プログラムは、翻訳前処理で、文中の「第<漢数字列>条」について正規表現を使用して検出し、<漢数字列>を算用数字に置き換えることにより、上記の問題を回避する。

従来の機械翻訳結果(数値が誤って翻訳される例)と、本提案によって数値が正しく翻訳される例とを表 3.11 と 3.12 に示す。

表 3.11 数値が誤って翻訳される例(従来)

日本語	英語
一 民事訴訟法第二百七十五条第二項又は第三百九十五条若しくは第三百九十八条第一項(同法第四百二条第二項において準用する場合を含む。)の規定により和解又は支払督促の申立ての時に訴えの提起があつたものとみなされたとき。	(I) where it is deemed, pursuant to the provisions of Article 275, paragraph (2) of the Code of Civil Procedure, or Article 395 or Article 309, paragraph (1) of said Code (including cases where applied mutatis mutandis pursuant to Article 402, paragraph (2) of said Code), that an action was filed at the time when a petition for a settlement or a demand for payment was filed ;

表 3.12 数値が正しく翻訳される例(提案手法)

日本語	英語
一 民事訴訟法第275条第二項又は第395条若しくは第398条第一項(同法第402条第二項において準用する場合を含む。)の規定により和解又は支払督促の申立ての時に訴えの提起があつたものとみなされたとき。	(I) where it is deemed, pursuant to the provisions of Article 275, paragraph (2) of the Code of Civil Procedure or Article 395 or Article 398, paragraph (1) of said Code (including cases where applied mutatis mutandis pursuant to Article 402, paragraph (2) of said Code), that an action was filed at the time when a petition for a settlement or a demand for payment was filed ;

(7) 定義ファイル等の作成

前述した翻訳品質改善プログラムの実装方針等を踏まえて、以下の辞書ファイルや定義ファイルを整備する。辞書ファイルや定義ファイルを更新することにより、翻訳品質改善プログラム等を修正することなく、用語や定義を追加・更新することができる。

1) 作成する定義ファイル等

① 法令用語辞書ファイル

令和3年度の検証において、「法令用語日英標準対訳辞書」への非準拠が見られた事例のうち、「法令用語日英標準対訳辞書」を参照し、翻訳エンジンの辞書として登録可能な語句を抽出し記述した（表 3.13 参照）。

表 3.13 辞書ファイル

#	日本語	英語
1	施行規則	Enforcement Regulation
2	法務省令	Order of the Ministry of Justice
3	代表権	right to represent
4	当該	relevant
5	重要な使用人	significant employees
6	登記事項証明書	certificate of registered information
7	住民票	resident records
8	賠償	compensation
9	保証契約	guarantee contract
10	委託者	entrustor
11	担保権の実行	enforcement of a security right
12	特定金銭債権	specified monetary claim
13	職員	employee
14	報酬	reward
15	暴力団員	member of an organized crime group
16	警察庁長官	Commissioner-General of the National Police Agency
17	訴訟手続	litigation proceedings
18	過料	civil fine
19	一部改正	partial amendment
20	不変期間	unextendable period
21	収入印紙	revenue stamp
22	民事訴訟	civil suit
23	督促手続	demand procedures

② 翻訳前後処理定義ファイル

以下の翻訳品質改善プログラムで使用する定義を記述する。

【3】定義語は小文字にする（大文字・小文字の表記）

【4】「○項から○項まで」の表記（定型表現への対応）

【5】主語省略の影響による湧き出し（湧き出しへの対応）

【6】条見出し中の etc. の訳出（訳出への対応）

2) 実施方法

【3】定義語は小文字にする（大文字・小文字の表記）

正規表現と処理する位置番号(n)
<置き換え正規表現> ||| n
...

例として定義を以下とする。

¥(hereinafter (.*)“(.*)”¥) ||| 2

¥((I)¥) ||| 1

上の定義は、例えば翻訳文中の

(hereinafter referred to as a “Written Application for Permission”)

にマッチし、Written Application for Permission が小文字化の対象となる。

【4】「○項から○項まで」の表記（定型表現への対応）

置き換え元単語列の正規表現と置き換え先単語列の正規表現の組
<置き換え元単語列正規表現> ||| <置き換え先単語列正規表現>
...

例として定義を以下とする。

Articles? (¥d+) to (¥d+) ||| Articles ¥1 through ¥2

上の定義は、例えば翻訳文中の

Article 2 to 3

にマッチし、Article が Articles に、to が through に置き換わる。

【5】主語省略の影響による湧き出し（湧き出しへの対応）

置き換え元単語の正規表現と、処理する位置番号(m, n)、置き換え先単語及び入力文においてこの置き換への偽条件となる正規表現の組
<置き換え元単語正規表現> ||| <置き換え先単語> | <入力文単語偽条件正規表現>
...

例として定義を以下とする。

Prime Minister ||| Minister | 総理大臣

上の定義は、例えば翻訳文

... the Prime Minister must consult ...

中の Prime Minister にマッチし、入力文中に「総理大臣」の単語がないときに限り、Prime Minister が Minister に置き換わる。

【6】条見出し中の etc. の訳出（訳出への対応）

例として定義を以下とする。

`^¥(. *?(, etc¥.) . *?¥)$ | | | 1`

上の定義は、例えば翻訳文中の

(Application for Authorization of Business Transfer, etc.)

にマッチし、最後の” , etc.” が削除される。

(8) 前後処理プログラムの実行順

翻訳品質改善プログラムの前後処理プログラムは表 3.14 及び表 3.15 に示す順番で実行した。

表 3.14 前後処理プログラムの実行順(前処理)

順番	改善処理	説明
1	【7】数字の表記	三桁以上の「第 xxx 条」を算用数字に置き換える

表 3.15 前後処理プログラムの実行順(後処理)

順番	前後処理	説明
1	【6】条見出し中の etc.への対応	ルールに基づき削除(, etc.)
2	【2】条・項・号の対応	ルールに基づき小文字化 ((l) -> (i))
3	【3】大文字・小文字の表記	hereinafter 処理
4	【4】定型表現への対応記	ルールに基づき置き換え(Articles m to n -> Articles m through n)
5	【5】湧き出しへの対応	ルールに基づき条件付き置き換え(minister)

4. 翻訳結果に対する評価

4.1 評価対象

評価作業で対象とした法令等を表 4.1 に示す。評価対象のうち、S01、S02-1、S02-2 及び S03 については、法務省における品質検査（ネイティブチェック及び専門家による検査等）を受けた英訳が J L T に掲載されている。S04、S05、S06、及び、S07 については、品質検査後の英訳は J L T に未掲載である。3.4 節に示した翻訳品質改善プログラムを実装した翻訳システム（以下、本システムと言う。）に評価対象法令を入力し、出力された翻訳結果について評価を行った。

表 4.1 評価対象とした法令等

番号	法令名	法律番号	行数
S01	債権管理回収業に関する特別措置法施行規則	平成十一年一月二十九日法務省令第四号	139
S02-1	民事訴訟費用等に関する法律	昭和四十六年四月六日法律第四十号	160
S02-2	民事訴訟費用等に関する法律（別表）	昭和四十六年四月六日法律第四十号	158
S03	裁判外紛争解決手続の利用の促進に関する法律	平成十六年十二月一日法律第百五十一号	202
S04	相続等により取得した土地所有権の国庫への帰属に関する法律	令和三年法律第二十五号	93
S05	情報通信技術を利用する方法による国の歳入等の納付に関する法律	令和四年法律第三十九号	75
S06	確定拠出年金法	平成十三年法律第八十八号	986
S07	こども家庭庁設置法	令和四年法律第七十五号	103

各法令の選定理由については下記の通りである

- ・ S01～03 ……令和 3 年度の検証で使用した法令。
課題①及び課題②の改善状況について確認するため。
- ・ S04～07 …… J L T 未掲載につき、対訳コーパスを有していない法令。
新しい法令を翻訳した際の結果を評価するため。
- ・ S04 ……予備評価を行った際に「Prime Minister」の湧き出しが認められたもの。
その事象が改善されたかを評価するため。
- ・ S06 ……日本語原文に「内閣総理大臣」が含まれているもの。
湧き出しへの対応を行った後も、「内閣総理大臣」が正しく訳出されるかを評価するため。

4.2 客観評価

本システムの翻訳品質を、客観評価の指標であるBLEU (bilingual evaluation understudy)によって評価した。BLEUは、機械翻訳の精度を評価する際に広く利用されている訳質を示すスコアで、値が大きいほど高精度である。BLEUスコアは、原文に対する人間による模範訳（参照訳）と自動翻訳の訳文の類似度を計算することによって得られる。訳質に関する人手評価の結果との相関を高くする工夫として、BLEUスコアは1文単位で計算するのではなく多数の文（例えば数百文）から計算する。BLEUスコアを算出するには参照訳が必要であることから、評価対象とした法令のうち英訳がJLTに掲載されている3つの法令（S01～S03）を評価の対象とした。

BLEUスコアの結果については、表4.2のとおりである。ここで、表中の「ベース」は本業務の翻訳品質改善プログラムを適用する前のNICT法令契約翻訳モデルによる翻訳結果から算出したスコアで、「改善後」は本業務の翻訳品質改善プログラムを適用した本システムによる翻訳結果から算出したスコアである。

表4.2から、翻訳品質改善プログラムによって、評価対象のいずれの法令についても、BLEUスコアの観点から翻訳品質が向上したといえる。

表 4.2 BLEU スコア

番号	法令名	NICT法令契約		
		ベース	改善後	差分
S01	債権管理回収業に関する特別措置法施行規則	65.97	70.25	4.28
S02-1	民事訴訟費用等に関する法律	49.76	59.32	9.56
S02-2	民事訴訟費用等に関する法律（別表）	59.03	61.93	2.90
S03	裁判外紛争解決手続の利用の促進に関する法律	43.96	48.13	4.17

4.3 主観評価1（スコア評価）

評価対象とした法令について、主観評価（スコア評価）を行った。具体的には、「正確さ」（原文の意味するところを、英文として過不足なく正しく伝えているか）と「流暢さ」（原文に忠実でありながら、正しい英文となっているか）の2つの観点について5段階で評価を行った。評価は、日本法令の英訳業務の実績を有する1名の翻訳者が行った。評価対象の法令それぞれからランダムに25文ずつ抽出し、合計200文について評価を行った。なお、評価対象文の抽出にあたっては31文字以上の文を対象とした。

評価結果は、表4.3のとおりである。

表4.3から、翻訳品質改善プログラムによって、正確さが大幅に向上したといえる。また、流暢さについてもベースシステムの品質が概ね維持できているといえる。

表 4.3 主観評価結果（スコア評価）

	ベース		改善後	
	正確さ	流暢さ	正確さ	流暢さ
S01	3.16	4.44	4.20	4.88
S02-1	2.84	4.36	3.28	4.24
S02-2	3.20	4.40	3.28	3.96
S03	2.88	4.40	3.84	4.36
S04	2.96	3.96	3.12	3.88
S05	2.76	4.12	2.96	3.64
S06	2.96	4.08	3.64	4.00
S07	3.08	4.50	3.96	4.60
全体	2.97	4.26	3.54	4.20

4.4 主観評価2（エラーカウント）

評価対象とした法令について、主観評価（エラーカウント）を行った。具体的には、本業務において改善対象とした9つの誤訳パターンについて、それぞれ翻訳品質改善プログラム適用前の機械翻訳文でのエラー発生件数及び翻訳品質改善プログラム適用後の機械翻訳文での翻訳品質改善件数をカウントした。評価結果は、表4.4のとおりである。評価は、日本法令の英訳業務の実績を有する1名の翻訳者が行った。

表 4.4 主観評価結果（エラーカウント）

改善事項	誤訳パターン	発生件数	翻訳品質改善件数	翻訳品質改善率 (%)
1	主語省略の影響による湧き出し （「法務大臣」の英訳として“Prime Minister”の湧き出しに限る）	4	4	100.0
2	定義規定の“shall”の訳出	561	561	100.0
3	性別を示す表現の出現の訳出	162	162	100.0
4	条・項・号の英訳表記 （項の訳抜け／号が大文字になる事象／1項しかないものも(1)と入る事象）	231	231	100.0
5	号の文頭・文末表記 （各号の先頭文字小文字、最終号の1つ前号の文末表記、最終号の文末の表記）	360	350	97.2
6	定義語は小文字にする	197	123	62.4
7	「○項から○項まで」の表記	8	8	100.0
8	条見出し中の“etc.”の訳出	36	36	100.0
9	数値の誤り	22	22	100.0

改善事項5では一部未改善の訳文が見られた。これは、原文の条・項・号の構成が特殊であり、プログラム開発時に想定したルールを外れていたことによると考えられる。

また改善事項6でも未改善の訳文が見られた。これは“hereinafter”で定義された語が同一法令内に再度出現した際、該当の語を小文字に変換するプログラム処理に問題があったためであり、今後のプログラムの改修を要する。

4.5 主観評価3（詳細評価）

法務省の法令翻訳アドバイザー（英訳法令の品質検査を担当している者）等による主観評価（詳細評価）を行った。具体的には、前記2記載の課題（「不適切な主語の補い」及び「訳語の不統一、「法令用語日英標準対訳辞書」・「法令翻訳の手引き」への非準拠）が改善されているかについて確認し、その結果を誤訳パターンごとの品質改善結果としてまとめた（表4.5参照）。また、構築前、構築中の工程（アダプテーション、クレンジング、前後処理）、及び構築後の機械翻訳文を比較し、いずれの工程が有用であったかについても確認を行った。

誤訳パターンごとにカウントした翻訳品質改善結果を表4.5及び表4.6に示す。

表 4.5 翻訳品質改善対象とした誤訳パターン別 カウント数

番号	誤訳パターン	既存エラー 一件数	改善 件数	改善率 (%)	新出エラー 一件数	備考
1	主語省略の影響による湧き出し （「法務大臣」の英訳として“Prime Minister”の湧き出しに限る）	1	1	100.0	0	翻訳品質改善プログラム【1】+【5】
2	定義規定の“shall”の訳出	84	84	100.0	0	翻訳品質改善プログラム【1】
3	性別を示す表現の出現の訳出	16	15	93.8	0	翻訳品質改善プログラム【1】
4	条・項・号の英訳表記					
	・1項しかないものに(1)が入る	23	23	100.0	0	翻訳品質改善プログラム【2】
	・号が大文字になる（(i)が(1)になる）	30	30	100.0	0	翻訳品質改善プログラム【2】
	・(1)の挿入	5	5	100.0	0	翻訳品質改善プログラム【2】
5	号の文頭・文末表記					
	各号の頭文字を小文字	43	43	100.0	0	翻訳品質改善プログラム【2】
	号の文末表記	44	44	100.0	0	翻訳品質改善プログラム【2】
6	定義語の大文字小文字	15	11	73.3	2	翻訳品質改善プログラム【3】
7	「○項から○項まで」の表記	5	5	100.0	0	翻訳品質改善プログラム【4】
8	条見出し中の“etc.”の訳出	15	15	100.0	0	翻訳品質改善プログラム【1】

						翻訳品質改善プログラム【6】
9	数値の誤り（三桁の漢数字）	8	8	100.0	0	翻訳品質改善プログラム【7】
10	法令用語日英標準対訳辞書への非準拠					
	・当該の訳として“said”を使用	8	6	100.0	0	未対策
	・用語の統制（法務省令）	22	0	0.0	2	用語登録
	・用語の統制（登記事項証明書）	3	3	100.0	0	用語登録
	・用語の統制（法令名）	3	3	100.0	0	用語登録
	・用語の統制（住民票）	2	2	100.0	0	用語登録
	・用語の統制（職員）	0	0	-	1	用語登録
	・用語の統制（その他）	9	5	55.6	3	未対策

表 4.6 本業務の改善対象外の誤訳パターン別 カウント数

誤訳パターン	既存エラー 件数	改善 件数	改善率(%)	新出エラー 件数
訳抜け	15	12	80.0	11
番号の訳抜け（別表）	0	0	-	9
番号の誤訳（三桁の漢数字以外）	0	0	-	14
枝番号の誤訳	3	0	0.0	0
主語省略の影響による湧き出し （Prime Minister 以外）	9	6	66.7	0
主語以外の湧き出し	0	0	-	3
訳文の重複	0	0	-	3
定義語以外の大文字・小文字	14	0	0.0	9
法令翻訳の手引きへの非準拠				
・ in cases where の使用	4	3	75.0	0
・ 見出し語は大文字に	3	1	33.3	7
定義語の訳の不一致	5	1	20.0	4
項番表記の不統一（別表）	5	4	80.0	0
訳語の不統一	3	0	0.0	4
文法上の問題				
・ When 節では現在形	3	1	33.3	1
・ 単語の単数形・複数形	4	2	50.0	2
・ 冠詞の誤り	1	0	0.0	6
辞書登録による悪影響 （relevant（当該）を含む語の訳出）	0	0	-	23

ルールが競合したときのエラー (号の文頭に法律名が訳出される場合 の表記)	0	0	-	9
---	---	---	---	---

4.6 各課題への対応結果

(1) 課題① 「不適切な主語の補い」への対応結果

- ・内容

主語省略の影響による湧き出しのうち、“Prime Minister”の湧き出しを抑制する。

- ・対策方法及びその方法を選択した理由

翻訳品質改善プログラム【1】+【5】

- ・実施した結果

“Prime Minister”が不適切に湧き出されることについては全体を通じて改善されていた。

(改善事例) 相続等により取得した土地所有権の国庫への帰属に関する法律第13条第4項

日本語原文	翻訳品質改善プログラム 実施前英訳サンプル	翻訳品質改善プログラム 実施後英訳サンプル
4 法務大臣は、第一項の規定により第五条第一項の承認を取り消したときは、法務省令で定めるところにより、その旨を同項の承認を受けた者に通知するものとする。	(4) If the Minister of Justice rescinds the approval referred to in Article 5, paragraph (1) pursuant to the provisions of paragraph (1), <u>the Prime Minister</u> is to notify the person that has obtained the approval referred to in that paragraph of this pursuant to the provisions of Ministry of Justice Order.	(4) Having rescinded the approval referred to in Article 5, paragraph (1) pursuant to the provisions of paragraph (1), <u>the Minister of Justice</u> is to notify the person that has obtained the approval referred to in that paragraph of this, pursuant to the provisions of the Order of the Ministry of Justice.

(2) 課題② 訳語の不統一、「法令用語日英標準対訳辞書」・「法令翻訳の手引き」への非準拠への対応結果

②-1. 定義規定の“shall”の訳出訳制

- ・内容

翻訳文中に“shall”が訳出されないようにする。

- ・対策方法

翻訳品質改善プログラム【1】

- ・結果

全体を通じて改善されていた。クレンジング前の対訳コーパスを使用してアダプテーションした翻訳モデルの時点で大幅な改善が見られていたため、本件における

クレンジング効果は限定的であったが、クレンジング後の対訳コーパスを使用してアダプテーションした翻訳モデルでは“shall”を含む表現は完全に訳出されなくなっていた。

(改善事例) 相続等により取得した土地所有権の国庫への帰属に関する法律 第5条第2項

日本語原文	翻訳品質改善プログラム 実施前英訳サンプル	翻訳品質改善プログラム 実施後英訳サンプル
2 前項の承認は、土地の一筆ごとに行うものとする。	(2) The approval set forth in the preceding paragraph <u>shall</u> be made for each parcel of land.	(2) The approval set forth in the preceding paragraph <u>is to be</u> granted for each parcel of land.

②-2. 性別を示す表現の訳出抑制

・内容

翻訳文中に“him/her”などの性別を示す表現が訳出されないようにする。

・対策方法

翻訳品質改善プログラム【1】

・結果

全体を通じて改善されていた。クレンジング前の対訳コーパスを使用してアダプテーションした翻訳モデルの時点で大幅な改善が見られていたため、本件におけるクレンジング効果は限定的であったが、クレンジング後の対訳コーパスを使用してアダプテーションした翻訳モデルでは“him/her”を含む表現は完全に訳出されなくなっていた。

(改善事例) 債権管理回収業に関する特別措置法施行規則 第11条

日本語原文	翻訳品質改善プログラム 実施前英訳サンプル	翻訳品質改善プログラム 実施後英訳サンプル
第十一条 債権管理回収業の実施業務に従事する者は、その業務を行うに当たり、債権回収会社の商号、許可番号及び自己の氏名を記載した身分を示す証明書を携帯し、相手方の請求があったときは、これを提示しなければならない。	Article 11 (1) A person engaged in the implementation of a claim management and collection business shall, in carrying out <u>his / her</u> business, carry an identification card containing the trade name and license number of the	Article 11 A person engaged in carrying out the operations for conducting the claim management and collection business must carry an identification card stating the trade name and license number of the claim management and collection

	claim management and collection company and <u>his / her</u> own name, and present it at the request of the other party.	company, and <u>the person's</u> own name, and present it when requested by the other party, in the course of carrying out the operations.
--	--	--

②-3. 条・項・号の英訳表記

・内容

条・項・号の記載方法について、法令翻訳のルールに準拠させる。

具体的には、

- ・ 項の訳抜け（(1)表記抜け）
- ・ 号の表記については小文字で記載すべきところ、大文字になる事象
- ・ 1項しかない場合は（1）の表記は不要であるところ、(1)と入る事象

を改善する。

・対策方法

翻訳品質改善プログラム【2】

・結果

ほぼ改善されたことを確認しているが、法令 S02-2 において条項番号の訳抜けや誤訳が見られた。

（改善事例）債権管理回収業に関する特別措置法施行規則 第6条第1項第1号

日本語原文	翻訳品質改善プログラム 実施前英訳サンプル	翻訳品質改善プログラム 実施後英訳サンプル
二 合併の経緯等を説明した書面	(I) a document explaining the particulars, etc. of the merger ;	(i) a document explaining the particulars, etc. of the merger ;

（要修正事例）民事訴訟費用等に関する法律 別表第一 一七イ(ロ)

日本語原文	翻訳品質改善プログラム 実施前英訳サンプル	翻訳品質改善プログラム 実施後英訳サンプル
(ロ) 非訟事件手続法又は国際的な子の奪取の民事上の側面に関する条約の実施に関する法律の規定による忌避の申立て、特別代理人の選任の申立て、弁護士でない者を手続代理人に選任	(b) The filing of a petition for a challenge under the provisions of the Non-Contentious Cases Procedures Act or the Act for Implementation of the Convention on the Civil	A petition for challenge under the provisions of Non-Contentious Case Procedures Act or Act for Implementation of the Convention on the Civil Aspects of International

<p>することの許可を求める申立て、裁判所書記官の処分に対する異議の申立て、これらの法律の規定による強制執行の停止、開始若しくは続行を命じ、若しくは執行処分の取消しを命ずる裁判を求める申立て又は受命裁判官若しくは受託裁判官の裁判に対する異議の申立て</p>	<p>Aspects of International Child Abduction, a petition for the appointment of a special agent, a petition for permission to appoint a person who is not an attorney at law as a counsel, an objection to a disposition made by a court clerk, a petition for a judicial decision to order a stay, commencement, or continuation of compulsory execution or to order the revocation of a disposition of execution under the provisions of these Acts, or an objection to a judicial decision made by an authorized judge or commissioned judge under the provisions of these Acts</p>	<p>Child Abduction, a petition for the appointment of a special agent, a petition for permission to appoint a person who is not an attorney at law as counsel, an objection to a disposition made by a court clerk, a petition for a judicial decision to order a stay, the commencement, or a continuation of compulsory execution or order the revocation of a disposition of execution, which is made under the provisions of these Acts, or an objection to a judicial decision made by an authorized judge or commissioned judge</p>
--	---	---

②-4. 号の文頭・文末表記

・内容

号の文頭・文末の記載方法について、法令翻訳のルールに準拠させる。

具体的には、

- ・各号の先頭文字は小文字に統制する
- ・最終号の1つ前の号の文末の表記（;and 又は ;or を付す）改善
- ・文末にはピリオドをつける

点について実施する。

・対策方法

翻訳品質改善プログラム【2】

・結果

ほぼ改善されていることを確認しているが、号の文頭が法令名から始まる場合には大文字であるべきところ、法令 S-02 において小文字になっている箇所があることを確認した。

(改善事例) 裁判外紛争解決手続の利用の促進に関する法律 第6条第1項第10号

日本語原文	翻訳品質改善プログラム 実施前英訳サンプル	翻訳品質改善プログラム 実施後英訳サンプル
十 民間紛争解決手続において提出された資料の保管、返還その他の取扱いの方法を定めていること。	(x) <u>A</u> method has been established for retaining, returning and other handling of materials submitted in the course of private dispute resolution procedures ;	(x) <u>establishes</u> methods for retaining, returning, and other handling of materials submitted through private dispute resolution procedures ;

(要修正事例) 民事訴訟費用等に関する法律 第3条第2項第3号

日本語原文	翻訳品質改善プログラム 実施前英訳サンプル	翻訳品質改善プログラム 実施後英訳サンプル
三 消費者の財産的被害の集団的な回復のための民事の裁判手続の特例に関する法律第五十二条第一項の規定により債権届出の時に訴えの提起があつたものとみなされたとき。	(iii) procedure to determine the amount of execution costs and money to be refunded as prescribed in Article 42, paragraph (4) of the Civil Execution Act ;	(iii) <u>civil</u> Execution Act proceedings for determining the execution costs and the amount of money to be returned prescribed in Article 42, paragraph (4); and

②-5. 号の文頭・文末表記

- ・内容

“hereinafter” で定義されている用語を小文字に統制する。

- ・対策方法

翻訳品質改善プログラム【3】

- ・結果

改善されたことを確認したが、一部で未改善箇所があることも確認した。

(改善事例) 債権管理回収業に関する法律施行規則 第6条第1項第4号

日本語原文	翻訳品質改善プログラム 実施前英訳サンプル	翻訳品質改善プログラム 実施後英訳サンプル
-------	--------------------------	--------------------------

四 合併後存続する会社又は合併により設立される会社（以下「合併会社」という。）に係る次に掲げる書類	(iv) the following documents pertaining to the company surviving the merger or company incorporated upon the merger (hereinafter referred to as the “ <u>Merging Company</u> ”) :	(iv) the following documents concerning the company surviving the merger or the company incorporated in the merger (hereinafter referred to as the “ <u>merging company</u> ”) :
---	---	--

②-6. 「○項から○項まで」の表記

・内容

「○項から○項まで」の表記では“to”ではなく、“through”を用いる。

・対策方法

翻訳品質改善プログラム【4】

・実施した結果

全体を通じて改善されていた。

(改善事例) 相続等により取得した土地所有権の国庫への帰属に関する法律 目次

日本語原文	翻訳品質改善プログラム 実施前英訳サンプル	翻訳品質改善プログラム 実施後英訳サンプル
第四章 雑則（第十三条— 第十六条）	Chapter IV Miscellaneous Provisions (Articles 13 <u>to</u> 16)	Chapter IV Miscellaneous Provisions (Articles 13 <u>through</u> 16)

②-7. 条見出しの“etc.”の訳出抑制

・内容

条見出し中に「等」が含まれていても“etc.”を訳出しないようにする。

・対策方法

翻訳品質改善プログラム【6】

・結果

全体を通じて改善されていた。

(改善事例) 裁判外紛争解決手続の利用の促進に関する法律 第11条

日本語原文	翻訳品質改善プログラム 実施前英訳サンプル	翻訳品質改善プログラム 実施後英訳サンプル
(<u>認証の公示等</u>)	(Public Notice of Certification, <u>etc.</u>)	(Public Notice of Certification)

その他

③-1. 数値の誤訳

・内容

数字（条番号）を正しく訳出する。

・対策方法

翻訳品質改善プログラム【7】

・実施した結果

ほぼ改善されていることを確認している。

一方で枝番号（例：〇〇条の2）が訳せていない箇所が見られた。

（例）民事訴訟費用等に関する法律 第13条の2第1項第4号

日本語原文	翻訳品質改善プログラム 実施前英訳サンプル	翻訳品質改善プログラム 実施後英訳サンプル
四 少額訴訟債権執行（民事執行法 <u>第一百六十七条の二</u> 第二項に規定する少額訴訟債権執行をいう。以下同じ。）の手續	(iv) proceedings for execution against a claim relating to an action on small claim (meaning the execution against a claim relating to an action on small claim prescribed in <u>Article 106, 17-2</u> , paragraph (2) of the Civil Execution Act ; the same shall apply hereinafter) ;	(iv) a procedure for execution against a claim relating to a small claims action (meaning execution against a claim relating to a small claims action prescribed in <u>Article 167, paragraph (2)</u> of the Civil Execution Act; the same applies hereinafter).

4.7 翻訳サンプル再生成と評価

4.6に示した翻訳結果を踏まえ、さらに、改善が必要な事項に対応した上、再度翻訳サンプルを生成し、その翻訳結果について再評価（主観評価）を行った。

翻訳サンプル再生成に当たっては、改善が必要な事項のうち、下記（2）記載の事項について翻訳品質改善プログラムを修正した。

また、追加学習については、

- ・学習内容の違い（クレンジング項目の違い）により翻訳結果が異なること
- ・誤訳パターンのうち“shall”及び性別を示す表現の訳出については、クレンジング前の対訳コーパスをアダプテーションした状態で改善が見られたこと
- ・“shall”を含む対訳コーパス（5600文以上）を除外することによる影響が大きいと考えられること

から、①クレンジングを一切行わない場合、②クレンジング項目から“shall”を除外した場合の2種類の追加学習を実施し、下記（1）のとおり、それぞれ前記の前後処理を施した環境での翻訳結果を確認した。

（1） 評価対象

A：「クレンジングを一切行わず、アダプテーション+前後処理を施した環境で、定義ファイル（辞書登録）の修正を行った翻訳サンプル」

B：「shall以外のクレンジングを行った上でアダプテーション+前後処理を施した環境に、定義ファイル（辞書登録）の修正を行った翻訳サンプル」

（2） 改善すべき課題及びその対応方法

- ① 定義語の小文字化
定義語の小文字化の文書全体への適用
- ② 見出し語は頭文字を大文字にするというルールに準拠していない点
見出し内キャピタライズ化
- ③ 文頭でルールが抵触した際の処理
各号の先頭は小文字、ただし法令名ときは大文字とする処理を行う
- ④ 表内の二桁の漢数字への対応
表内の二桁の漢数字（「一一」など）への対応
- ⑤ 「〇〇大臣」の〇〇の訳出が抜け落ちてしまう課題
定義ファイル修正（ministerの過置換）
- ⑥ 「当該」を辞書登録したことにより、「当該〇〇」の英訳に問題が見られる点
辞書登録用語の更新

（3） 改善内容と結果

- ① 定義語の小文字化
定義された語や略称の翻訳において、(hereinafter … “xxxx”)と

訳され、その際に xxxx は小文字とすべきだが、xxxx が小文字になっていない場合がある。

1 回目で改善されたことを確認したが一部で未改善箇所があることも確認した。

・実施方法

翻訳後処理で定義箇所 (hereinafter … “xxxx”) が存在する部分を検索し、xxxx を小文字に変換する。また、“xxxx” について、大文字であるか小文字であるかを区別せずに定義箇所以降の文書全体を検索し、文書内の xxxx も小文字にする。

・実施した結果

おおむね改善していた。

(例) 債権管理回収業に関する特別措置法施行規則 第 4 条第 2 項

日本語原文	翻訳品質改善プログラム 実施英訳サンプル	翻訳品質改善プログラム 修正後英訳サンプル (A)	翻訳品質改善プログラム 修正後英訳サンプル (B)
<p>2 債権回収 会社は、法第 七条第一項に 規定する届出 をしようとする ときは、別 紙様式第四号 により作成し た変更等届出 書に、前条各 号に掲げる書 類のうち変更 に係るものを 添付して、法 務大臣に提出 しなければならない。</p>	<p>(2) When a <u>Claim management and collection company</u> intends to make a notification prescribed in Article 7, paragraph (1) of the Act, it must submit to the Minister of Justice a written notification of change, etc. prepared using Appended Form No. 4, and attach thereto the documents listed in the items of the preceding Article which pertain to the change.</p>	<p>(2) If a <u>claim management and collection company</u> intends to make the notification prescribed in Article 7, paragraph (1) of the Act, it must submit to the Minister of Justice a written notice of change, etc. prepared using Appended Form No. 4, and attach thereto the documents listed in the items of the preceding Article which pertain to the change.</p>	<p>(2) If a <u>claim management and collection company</u> intends to make a notification as prescribed in Article 7, paragraph (1) of the Act, it must submit to the Minister of Justice a written notification of change, etc. prepared using Appended Form No. 4, and attach thereto the documents set forth in the items of the preceding Article which pertain to the change.</p>

② 見出し語は頭文字を大文字にするというルールに準拠していない点

・実施方法

見出しの文字列の各単語について、接続詞等は除いて頭文字を大文字にす

る。

・実施した結果

ほぼ修正されていたが、法令 S-03 の中で一部修正されていないものもあった。

(改善事例) 民事訴訟費用等に関する法律 第 13 条

日本語原文	翻訳品質改善プログラム 実施英訳サンプル	翻訳品質改善プログラム 修正後英訳サンプル (※)
(郵便切手等による予納)	(Prepayment by <u>Postal</u> postage stamp)	(Prepayment by <u>Postal</u> Postage Stamp)

※ A の翻訳結果と B の翻訳結果は同一であった。

(要修正事例) 裁判外紛争解決手続の利用の促進に関する法律 第 15 条

日本語原文	翻訳品質改善プログラム 実施英訳サンプル	翻訳品質改善プログラム 修正後英訳サンプル (※)
(暴力団員等の使用の禁 止)	(Prohibition of Use of <u>member of an organized crime</u> group)	(Prohibition of Use of member of an organized crime group)

※ A の翻訳結果と B の翻訳結果は同一であった。

③ 「各号の頭文字は小文字にする」というルールと、「法令名の頭文字は大文字にする」というルールが競合した時の対応

・実施方法

各号の先頭の単語は小文字とする。ただし、法令名のときは大文字とする。

・結果

a Civil～のように、法令名に冠詞がつく形で回避されていることから不適切であり、更なる改善が必要。

また、「〇〇法第〇条〇項の」の表記について、in Article ○, paragraph (○) of the 〇〇Act と記載すべきところ (例: 「民事執行法第四十二条第四項」 → Article 42, paragraph (4) of the Civil Execution Act)、法令名と条・項の数字の分離が見られた。

さらに、新たな問題点として、法令名の説明をしている箇所ではないにもかかわらず「法令名: 〇〇～」という書き出しが複数見られ、日本語の内容と一致せず誤訳が見られた。

(要修正事例) 民事訴訟費用等に関する法律 第 3 条第 2 項第 3 号

日本語原文	翻訳品質改善プログラム 実施英訳サンプル	翻訳品質改善プログラム 修正後英訳サンプル (A)	翻訳品質改善プログラム 修正後英訳サンプル (B)
三 消費者の財	(iii) <u>act</u> on Special	(iii) <u>a Act</u> on Special	(iii) <u>a Act</u> on Special

産的被害の集団的な回復のための民事の裁判手続の特例に関する法律第五十二条第一項の規定により債権届出の時に訴えの提起があつたものとみなされたとき。	Measures Concerning Civil Court Proceedings for the Collective Redress for Property Damage Incurred by Consumers where an action is deemed to have been filed at the time of the Filing of Proofs of Claims, pursuant to the provisions of Article 52, paragraph (1).	Measures Concerning Civil Court Proceedings for the Collective Redress for Property Damage Incurred by Consumers if it is deemed, pursuant to the provisions of Article 52, paragraph (1), that an action has been filed at the time of the Filing of Proofs of Claims.	Measures Concerning Civil Court Proceedings for the Collective Redress for Property Damage Incurred by Consumers where an action is deemed to have been filed at the time of the Filing of Proofs of Claims pursuant to the provisions of Article 52, paragraph (1).
--	---	---	--

(要修正事例) 民事訴訟費用等に関する法律 第13条の2第1項第3号

日本語原文	翻訳品質改善プログラム 実施英訳サンプル	翻訳品質改善プログラム 修正後英訳サンプル (A)	翻訳品質改善プログラム 修正後英訳サンプル (B)
三 民事執行法 第四十二条第四 項に規定する執 行費用及び返還 すべき金銭の額 を定める手続	(iii) <u>civil</u> Execution Act proceedings for determining the execution costs and the amount of money to be returned prescribed <u>in Article 42, paragraph (4)</u> ; and	(iii) <u>Civil Execution Act</u> : procedures for determining the amount of the execution costs and the money to be refunded as prescribed <u>in Article 42, paragraph (4)</u> ; and	(iii) <u>a Civil Execution Act</u> : a procedure for determining the amount of the execution costs and the money to be refunded as prescribed <u>in Article 42, paragraph (4)</u> ; and

④ 表内の二桁の漢数字への対応

・実施方法

「一一（じゅういち）」などのように一桁の漢数字が連続して記述されている漢数字表記があれば、二桁の数であるとみなして翻訳を実行する。

・結果

改善されていた。

ただし、「条・項・号の英訳表記」で見られた、表内の条項番号の訳抜けについては、改善されていなかった。

(改善事例) 民事訴訟費用等に関する法律 別表第一

日本語原文	翻訳品質改善プログラム 実施英訳サンプル	翻訳品質改善プログラム 修正後英訳サンプル (※)
--	(i)	(xi)

※ A の翻訳結果と B の翻訳結果は同一であった。

⑤ 「〇〇大臣」の〇〇の訳出が抜け落ちてしまう課題

・実施方法

翻訳品質改善プログラム【5】湧き出しへの対応の定義ファイルを修正する。

・結果

改善されていた。

一方で「財務大臣」の表現や大文字としなくて良い単語を大文字とするなど、処理が進むにつれて品質が落ちていると思われるものもあった。

(改善事例) 相続等により取得した土地所有権の国庫への帰属に関する法律 第8条

日本語原文	翻訳品質改善プログラム 実施英訳サンプル	翻訳品質改善プログラム 修正後英訳サンプル (※)
<p>第八条 法務大臣は、第五条第一項の承認をするときは、あらかじめ、当該承認に係る土地の管理について、財務大臣及び農林水産大臣の意見を聴くものとする。ただし、承認申請に係る土地が主に農用地（農地法（昭和二十七年法律第二百二十九号）第二条第一項に規定する農地又は採草放牧地をいう。以下同じ。）又は森林（森林法（昭和二十六年法律第二百四十九号）第二条第一項に規定する森林をいう。以下同じ）</p>	<p>Article 8 The Minister of Justice is to, when granting approval referred to in Article 5, paragraph (1), hear the opinions of <u>the Minister</u> and the Minister of Agriculture, Forestry and Fisheries in advance with regard to the management of land pertaining to relevant approval; provided, however, that the land pertaining to the application for approval is mainly agricultural land (meaning cropland or meadow / pastureland prescribed in Article 2, paragraph (1) of</p>	<p>Article 8 When granting the approval referred to in Article 5, paragraph (1), the Minister of Justice is to hear the opinions of <u>the Minister of Finance</u> and the Minister of Agriculture, Forestry and Fisheries in advance with regard to the management of land pertaining to the approval; provided, however, that the land pertaining to an application for approval is mainly agricultural land (meaning cropland or meadow / pastureland prescribed in</p>

	the Cropland Act (Act No. 229 of 1952); the same applies hereinafter) or forest (meaning forest prescribed in Article 2, paragraph (1) of the Forest Act (Act No. 249 of 1951); the same applies hereinafter).	Article 2, paragraph (1) of the Cropland Act (Act No. 229 of 1952); the same applies hereinafter) or forest (meaning forest prescribed in Article 2, paragraph (1) of the Forest Act (Act No. 249 of 1951); the same applies hereinafter).
--	--	--

※ A の翻訳結果と B の翻訳結果は同一であった。

- ⑥ 「当該」を辞書登録したことにより、「当該〇〇」の英訳に問題が見られる点
- ・実施方法
辞書登録から「当該」を削除する。
 - ・結果
「当該」の辞書登録解除後、修正ができていることを確認。

(改善事例) 民事訴訟費用等に関する法律 第 12 条第 2 項

日本語原文	翻訳品質改善プログラム 実施英訳サンプル	翻訳品質改善プログラム 修正後英訳サンプル (A)	翻訳品質改善プログラム 修正後英訳サンプル (B)
2 裁判所は、前項の規定により予納を命じた場合においてその予納がないときは、 <u>当該費用</u> を要する行為を行なわないことができる。	(2) When the court has ordered prepayment pursuant to the provisions of the preceding paragraph but the prepayment is not made, the court may choose not to conduct acts that require the <u>Cost of relevant.</u>	(2) If the court has ordered prepayment pursuant to the provisions of the preceding paragraph but the prepayment has not been made, the court may choose not to perform the act that requires <u>the relevant costs.</u>	(2) If the court has ordered a prepayment pursuant to the provisions of the preceding paragraph but the prepayment has not been made, the court may choose not to perform the act that requires <u>the costs.</u>

(その他)

辞書登録用語について、「当該」を含む文については前記のとおり訳質の改善が見られた。

一方で、新たに「弁護士 (attorney)」を辞書に登録したところ、「弁護士」を含む言葉 (例：弁護士会) を含む文が誤訳となったものや、「所有権 (ownership)」

を辞書登録したことによって、訳が悪化していると思われる条文が新たに発生するなど、一部悪化したものも見受けられた。

(要修正事例) 債権管理回収業に関する特別措置法施行規則 第3条第4項

日本語原文	翻訳品質改善プログラム 実施英訳サンプル	翻訳品質改善プログラム 修正後英訳サンプル(※)
四 <u>取締役である弁護士</u> が 法第六条第二項ただし書に 定める所属弁護士会の推薦 を受けた者であるときはそ の旨を証明する書面の写し	(iv) if <u>an attorney who</u> <u>serves as a director</u> is recommended by the bar association of which the attorney holds a membership as prescribed in the proviso to Article 6, paragraph (2) of the Act, a copy of the document certifying to that effect;	(iv) if <u>a attorney who</u> <u>serves as a attorney</u> is recommended by the board of company auditors to which the director belongs as prescribed in the proviso to Article 6, paragraph (2) of the Act, a copy of the document certifying to that effect;

※ Aの翻訳結果とBの翻訳結果は同一であった。

※冠詞の問題も新たに発生している。

- ・ 1回目の翻訳サンプル結果及び再評価用の翻訳サンプル(2種)をそれぞれ比較した結果、「法令翻訳の手引き」へ準拠度合いや英語としての読みやすさの優劣は法令によって異なっていた。

5. 翻訳品質の改善

5.1 評価結果の概略（初回翻訳結果）

<課題1 不適切な主語の補い>

日本語では、主節・条件節で主語が同じ場合、一方の主語を省略することがあるところ、英語では、英文として成立させるため、主語を補足する必要があり、その場合に、前後の文脈の関係で不適切な主語が補われることがある。

今回の検証では、日本語の「大臣」が英語で「総理大臣」を示す“Prime Minister”に置き換えられる事象を取り上げた。

【対策方法及びその方法を選択した理由】

A. 追加学習

JLT対訳コーパスの各対訳において、日本語側における「大臣」と英語側における“Minister”の出現数が一致しない場合には、不適切な補いを発生させる要素となる可能性があるため、当該対訳を除外して翻訳モデルの再学習を実施することにより、翻訳品質の改善を図った。

B. 翻訳後処理で当該語句を置換、あるいは、削除する

上記の対策で全ての原因コーパスを除外できない可能性があるため、翻訳結果に“Prime Minister”が出現しているが翻訳原文に「総理大臣」が含まれていない場合、“Prime Minister”を“Minister”に置換する。

【実施した結果】

“Prime Minister”の湧き出しは見られず、改善されていた。

【課題】

今回の検証では、想定していた効果を得た。しかし、日本語側と英語側とで主語の数が異なることが最適な対訳となるような条文が引き続き存在する場合、ニューラル翻訳の原理上、不適切な主語の補いが出現する可能性がある。

また、主語省略の影響による湧き出しは、“Prime Minister”に限られる事象ではなく、本システムにより作成した翻訳結果の中にも、他の湧き出しは見られた。これら全ての事象に対応をすることは困難であり、課題解決に向けて、引き続き検討を要する。

【対応策】

AI学習手法の更なる改良や、誤訳の可能性を示すチェックツール等による補完が考えられる。

【考察】

“Prime Minister”の湧き出しについては、追加学習に加え、翻訳結果を後処理で補

完したことが効果的であったため、翻訳システムに採用すべきであると言える。

ただし、日本語側と英語とで主語の数が異なることを避けた結果、構文に変化が見られるため、法務省における品質検査の過程では、自然な英訳であるかどうかの確認は引き続き実施されるべきである。

<課題2 訳語の不統一、「法令用語日英標準対訳辞書」・「法令翻訳の手引き」への非準拠>

① 定義規定の“shall”の訳出

【対策方法】

JLT掲載法令のうち、比較的古い時期に登録された法令には、“shall”が含まれている場合がある。これらの表現が含まれる対訳コーパスを使用して翻訳モデルを構築すると、翻訳結果に“shall”を含む文が訳出されてしまう可能性があることから、英文に“shall”が含まれる対訳を除外し、除外した対訳コーパスを使用して翻訳モデルの追加学習を実施した。

【実施した結果】

クレンジング後の対訳コーパスを使用してアダプテーションした翻訳モデルの翻訳結果から“shall”は完全に訳出されなくなっていた。クレンジング前の対訳コーパスを使用してアダプテーションした翻訳モデルの時点で大幅な改善が見られていたため、本件におけるクレンジング効果は限定的であった。

【課題】

対訳コーパスを除外した結果、一部の翻訳結果の悪化も見られた。

【対応策】

JLTの継続的な点検・改善に加え、対訳コーパス除外の影響を考慮しながら再学習を行うことが必要である。

【考察】

“shall”を含む対訳を対訳コーパスから除外することで、定義規定の“shall”の訳出に対しては根本的な対策を行うことが可能だが、追加学習させる対訳コーパスの内容によっては翻訳結果に悪影響を及ぼす場合もある。

そのため、定義規定の“shall”の訳出を防止したい場合は、全体的な翻訳結果に悪影響を及ぼす場合があることを念頭に、翻訳システムに採用すべきであると言える。

また、翻訳システムに採用する際は、運用開始前に、翻訳品質について繰り返し確認する等の調整を行うべきである。

② 性別を示す表現の訳出

【対策方法】

上記①と同様に対訳コーパスから英文に he/him/she/her など、性別を示す表現が含まれている対訳を除外し、除外した対訳コーパスを使用して翻訳モデルの追加学習を実施した。

【実施した結果】

改善されていた。

【考察】

性別を示す表現を含む対訳を対訳コーパスから除外し再学習を実施したことが効果的であった。

改善効果が得られており、翻訳システムに採用すべきであると言える。

③ 条・項・号の英訳表記

【対策方法】

翻訳エンジンのみでの改善が難しい事項であることから、翻訳対象文書を機械翻訳する前に、条・項・号を機械的に取り除き、本文のみ機械翻訳を実施した。その後、ルールに従って条・項・号を機械翻訳し、本文の翻訳結果に条・項・号の翻訳結果を統合した。

なお、条・項・号を除去した対訳コーパスを作成し、事前に追加学習を行った。

【実施した結果】

おおむね改善されたことを確認しているが、別表においては一部、条項番号の訳抜けが見られた。

【課題】

別表内の番号等の表記にも対応できるようにすること

【考察】

表内の条項番号の訳抜けについては、表内の箇条書きの番号等の表記ルールが定めれば、当該ルールを翻訳前後処理に実装することによって本事象が改善することが期待できる。

基本的には改善効果が得られており、翻訳システムに採用すべきであると言える。

④ 号の文頭・文末表記

【対策方法】

上記③と同様、翻訳原文を翻訳する前に、条・項・号を取り除き、本文のみ翻訳する。

その後、条・項・号の記述ルールに従って翻訳し、本文翻訳結果を条・項・号と統合する。翻訳結果を適切に修正するために、当該文書における条・項・号の構造を事前に解析しておく。

【実施した効果】

ほぼ改善されていることを確認したが、法令 S-02 の中で、本来は大文字で法律名は記載されるべきところ、法律名が号の文頭に表記される際にも小文字で表示される事象が見られた。

【課題】

翻訳ルールが抵触した場合の対応

【対応策】

法令名の大文字を優先し、各号の先頭は小文字、ただし法令名の際は小文字とする前後処理を実施。

【考察】

各処理の特性を踏まえ、前後処理と再学習を組み合わせたことが効果的であった。

課題はほぼ改善されていることから翻訳システムに採用すべきと言える。

ただし、翻訳ルールが抵触した場合については、下記 5.2 のとおり、現時点では採用すべきであるとは言えない。

⑤ 定義語は小文字にする

【対策方法及びその方法を選択した理由】

正規表現を活用し、翻訳後処理で、小文字に変換した。

【実施した結果】

おおむね改善されていた。

【課題】

一部で未改善箇所があることも確認した。

【対応策】

定義語の小文字化の文書全体への適用

【考察】

一部未改善箇所があるものの、基本的には改善効果があったため、翻訳システムに採用すべきであると言える。

⑥ 「○項から○項まで」の表記

【対策方法及びその方法を選択した理由】

翻訳エンジンでの対応が難しい事項であることから、対象とする語句と置換条件等を定義ファイルに記述し、翻訳後処理で当該語句を置換、あるいは、削除した。

【実施した結果】

改善されていた。

【考察】

翻訳後処理での対処が適切であった。

改善効果が得られており、翻訳システムに採用すべきであると言える。

⑦ 条見出し中の“etc.”の訳出

【対策方法及びその方法を選択した理由】

翻訳エンジンで対応が難しい事項であることから、翻訳前処理で条の文字部分を算用数字に置き換えてから翻訳を実行する。

【実施した結果】

改善されていた。

【考察】

翻訳後処理での対処が適切であった。

改善効果が得られており、翻訳システムに採用すべきであると言える。

⑧ 数値の誤り

【対策方法及びその方法を選択した理由】

翻訳エンジンで対応が難しい事項であることから、翻訳前処理で条の数字部分を算用数字に置き換えてから翻訳を実行する。

【実施した結果】

三桁の数字の翻訳については改善されていた。

一方、枝番号（例：〇〇条の２）が適切に訳出されていない箇所が見られた。

【課題】

枝番号（例：〇〇条の２）が適切に訳出されていない箇所が見られた。

【考察】

三桁の漢数字に対する処理については、翻訳後処理での対処が適切であり、翻訳システムに採用すべきであると言える。

一方、枝番号の表記については、法令の信頼性担保の観点から改善が必要な事項であるため、課題解決に向けて、引き続き検討を要する。

5.2 評価結果の概略（2回目翻訳結果）

① 定義語の小文字化

【対策方法】

改善処理を文書全体へ適用した。

【実施した結果】

改善されていた。

【考察】

改善効果が得られており、翻訳システムに採用すべきと言える。

② 見出し語は頭文字を大文字にするというルールに準拠していない点

【対策方法】

見出し内の文字列の各単語について、接続詞等は除いて頭文字を大文字にする。

【実施した結果】

ほぼ修正されていたが、一部修正されていないものもあった

【考察】

上記 4.7 で示した一部修正されていなかった事象については、小文字になっていた“member of an organized crime group”（暴力団員）は辞書登録をしていた用語であるため、辞書登録上の表記である小文字が優先されたものと考えられる。

おおむね改善効果が得られており、翻訳システムに採用すべきと言える。

③ 文頭でルールが抵触した時の対応

【対策方法】

法令名の大文字を優先し、各号の先頭は小文字、ただし法令名の際は小文字とする前後処理を実施。

【実施した結果】

法律名の大文字は改善されたが、冠詞の不適切な使用や誤訳が見られた。

【考察】

改善が見られた部分もあるが、前記結果のとおり、翻訳結果に悪影響を及ぼすことも確認できた。

対策方針の検討には更なる調査を要するが、法令名が文頭に記述された場合は冠詞を除外するなど、法令名の取り扱い方法のルールを定めて翻訳品質改善プログラムに採用することによって本事象が改善することを期待できる。

英文として成立することを優先すべき指針とするならば、現時点では採用すべきとは言えない。

④ 表内の二桁の漢数字への対応

【対策方法】

表内の二桁の漢数字表記（「一一」など）について、二桁の数であると認識させる処理を追加した。

【実施した結果】

改善されていた。

【考察】

改善効果が得られており、翻訳システムに採用すべきと言える。

⑤ 「〇〇大臣」の〇〇の訳出が抜け落ちてしまう課題

【対策方法】

定義ファイル修正

【実施した結果】

改善していた。

【考察】

改善効果が得られており、翻訳システムに採用すべきであると言える。

⑥ 「当該」を辞書登録したことにより、「当該〇〇」の英訳に問題が見られる点

【対策方法】

辞書登録用語からの削除

【実施した結果】

改善していた。

【考察】

「当該」のように、単体で使用できない用語を辞書登録すると、「当該」が名詞であるかのような翻訳結果になっていたことから、辞書登録が有効であるのは、単体で使用できる名詞であると考えられる。

したがって、「当該〇〇」を含む文の英訳に対しては、「当該」を辞書登録用語から削除することで改善効果が得られた。一方、他の用語（単体で使用できると考えた名詞絵）を辞書登録した結果、訳が悪化していると思われる条文が新たに発生するなど、一

部悪化したものも見受けられたため、辞書登録の活用に関しては、訳出結果への影響を逐一確認の上、翻訳システムへ採用することが必要だと思われる。

5.3 効率的かつ効果的なAI学習の強化方法及び翻訳品質改善プログラムの提案

上記の結果から、課題となっていた不適切な主語の補いや shall や him/her といった法令翻訳において使用すべきではない表現については、適切に修正・除去されたJLT対訳コーパスを翻訳エンジンに追加学習させることで改善が期待できるとわかった。また、同じく課題となっていた訳語の不統一、「法令用語日英標準対訳辞書」・「法令翻訳の手引き」への非準拠に関しても、適切な翻訳品質改善プログラムを構築し、翻訳エンジンと適切に組み合わせて翻訳システムとして実装することで、改善を期待できることを確認した。

法令に特化した翻訳システムを作成する際には、翻訳品質改善プログラム【1】の追加学習については、学習させる対訳コーパスによって翻訳結果が異なることから、訳出結果への影響を逐一確認の上、慎重に追加学習に用いる対訳コーパスの内容を決定することが必要と考えられる。

また、前後処理プログラムについては、本業務において有用であった、

- ・ 翻訳品質改善プログラム【2】
- ・ 翻訳品質改善プログラム【3】 + 文章全体への適用
- ・ 翻訳品質改善プログラム【4】
- ・ 翻訳品質改善プログラム【6】
- ・ 翻訳品質改善プログラム【7】 + 二桁の漢数字への対応
- ・ 見出し内キャピタライズ化

を採用することが有用であると考えられる。

本業務において、複数の法令について翻訳サンプルを生成・翻訳結果確認を実施し、さらに、改善が必要な事項に対応した上、それぞれの法令について再度2種類の翻訳サンプルを生成・翻訳結果確認を実施したものであるところ、どの翻訳システムを採用するかについては、翻訳ルールへの準拠度合いや、英語としての読みやすさが法令によって異なっていたことから、翻訳システムの導入に当たっては、翻訳品質についてのテストを複数の法令を用いて複数回実施した上で導入することが望ましいといえる。

なお、クレンジングについては、実施したことで訳質が改善しているものもあれば、訳質が悪化していると思われるものもあり、クレンジングの実施可否及び実施する際の除去する対訳については慎重な検討を要すると思われ、翻訳システム導入に当たってクレンジングを実施する際は、リリース前に入念な調整を要するものと考えられる。さらに、リリース後も年に1回程度コーパスの整備及び追加学習を予定しているため、最適な翻訳システムを維持するため、調整が必要になることが見込まれる。

さらに、新たな課題も発見されたが、翻訳の品質は、翻訳品質改善プログラムを搭載することにより向上しており、法令外国語訳業務において、英訳法令の原案作成に有用であることが確認できた。

ただし、日本語としても構造が分かりにくい文については、英文として成立してい

たとしても、係り方（解釈）が正しいか等について人の目で確認する必要があるほか、大文字/小文字、単数型/複数型、文脈に応じた訳語の選定等については、機械的に判断することは難しいと考えられるため、人による確認・修正を前提とし、チェックの在り方を充実させることを検討すべきである。

以上を踏まえると、本業務で翻訳品質改善に繋がった処理に加え、法令外国語訳における翻訳品質のチェックの在り方を充実させた上で英訳法令の原案の作成に臨むこととすれば、英訳法令の原案作成までの期間を抜本的に改善でき、英訳法令の原案作成期間短縮の結果及びチェック体制の充実等により、公開までの期間も短縮できると評価できる。

参考資料

- [1] 法務省大臣官房司法法制部, “法令翻訳の手引き,” 法務省大臣官房司法法制部 (2018)
- [2] 日本法令外国語訳推進会議, “法令用語日英標準対訳辞書 (令和4年3月改訂版),” 法務省大臣官房司法法制部 (2022)
- [3] 法務省, “日本法令外国語訳データベースシステム,”
<https://www.japaneselawtranslation.go.jp/ja>
- [4] 法務省, “日本法令の国際発信の推進に向けた官民戦略会議 (第3回),”
https://www.moj.go.jp/housei/hourei-shiryou-hanrei/housei03_00033.html
- [5] 国立研究開発法人情報通信研究機構, “みんなの自動翻訳@TexTra,”
<https://mt-auto-minhon-mlt.ucri.jgn-x.jp/>
- [6] 国立研究開発法人情報通信研究機構, 総務省, “翻訳バンク,”
<https://h-bank.nict.go.jp/>
- [7] 金融庁, “「自動翻訳エンジンの金融専用モデル構築に関する委託研究」についての報告書,”
<https://www.fsa.go.jp/common/about/research/20220311/20220311.pdf>