

上場企業開示資料の機械翻訳におけるリエディットの検討

土井 惟成 大西 恒彰 百石 弘澄 高頭 俊 山藤 敦史

株式会社 日本取引所グループ

{n-doi, n-onishi, h-hyakkoku, s-takato, a-santo}@jpx.co.jp

1 はじめに

東京証券取引所(以下、東証)上場企業の開示資料をはじめとする Investor Relations 分野(以下、IR 分野)の産業翻訳において、ニューラル機械翻訳(以下、NMT) [1]の活用が検討されている [2-5]。近年では IR 分野に適した機械翻訳モデルの精度向上と併せて、NMT を前提としたワークフローも検討が進められており、その具体例としてポストエディット (post-editing) やリエディット (pre-editing) の活用が挙げられる。

ポストエディットとは機械翻訳モデルの出力文を人手によって後編集を行うことであり、作業プロセス等が ISO 18587 [6] で規定されている。開示資料の機械翻訳では、数値及び固有名詞の誤訳、訳語の一貫性の不備、訳抜け等が問題点として指摘されている。これらの解決策として、翻訳支援ツール等を活用したポストエディットが検討されている [4,5]。一方、リエディットとは機械翻訳前に入力文を翻訳しやすいように言い換える作業である。一例として、NMT では長文の正確な訳出が難しいため、開示資料中の長文の機械翻訳においては、文意が変わらないように複数文に分けて入力することが望ましいと示唆されている [4]。しかしながら、開示資料においてどのようなリエディット手法がどの程度有効か、定量的には調査されていない。

本研究では、コーポレート・ガバナンスに関する報告書(以下、CG 報告書)を対象に、開示資料の NMT におけるリエディットの有効性について検討した。リエディットの作業ルールには、特許ライティングマニュアル第 2 版 [7] における、産業日本語への言い換えルールを利用した。産業日本語とは、「産業・技術情報を人に理解しやすく、かつ、コンピュータ(機械)にも処理しやすく表現するための日本語」と産業日本語研究会によって定義されており^{*1}、特許文書における NMT では産業日本語へのリエディットの有用性が示唆されている [8]。本研究では、図 1 のように、クラウドソーシング及び翻訳会社に言い換え作業を依頼し、汎用的な NMT サービスによる訳出文の比較を行うことで、次の観点から調査を行った。

- 開示資料におけるリエディットの有効性: 自動評価及び人手評価のスコアが向上するか。
- 各言い換えルールの有効性: どの言い換えルールがどの程度有効か。

実験より、開示資料のリエディットによる機械翻訳の品質向上と、リエディット手法として短文化が特に有効であることを示した。

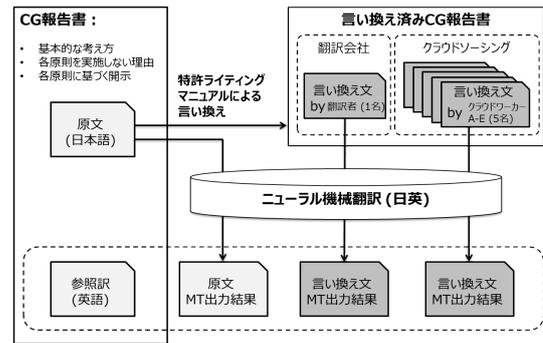


図 1 本実験の概要

2 関連研究

機械翻訳におけるリエディットの有効性評価に関する研究として、佐藤らの研究、早川らの研究、宮田らの研究が挙げられる。

佐藤らの研究 [9] では、歴史や文化に関する文書を対象に、自動生成した辞書を用いて機械翻訳しにくい表現をより一般的な表現に置換するという手法を提案し、自動評価のスコアの向上と人手評価の有意差を確認した。早川らの研究 [10] では、医学・医薬品関連の分野を対象に、NMT における 11 パターンのリエディット作業の有効性の測定を行った。その結果として、リエディット作業によって誤訳の頻度が低減し、特に、文脈情報の補足(主語や目的語の補足、指示語の具体化)や表記上の修正(誤字脱字の修正、不要な語句の削除)が有効であることを示した。また、文脈情報の補足は作業者に翻訳スキルを要するものの、表記上の修正は作業者のスキルに依らないことを示した。これらに対し、本研究との差異としては、対象文書のドメインが IR 分野である点、リエディット手法が特許ライティングマニュアルを利用した人手による言い換えである点という 2 点が挙げられる。

宮田らの研究 [11] では、病院内会話、自治体生活情報、新聞記事を対象に、統計的機械翻訳におけるリエディット事例の収集方法を提案し、収集したリエディット事例の分析結果について述べている。分析の結果、機械翻訳の訳出文の精度向上においてリエディットが有効であること、また、規定のルールが存在しない、試行錯誤的なリエディットは作業者の能力や経験に大きく依存していることを示した。本研究と宮田らの研究の異なる点としては、具体的なリエディット手法の有効性について評価している点、機械翻訳モデルとして NMT を利用している点の 2 点が挙げられる。

^{*1} <https://www.tech-jpn.jp/outline/>

表1 対象データの規模

項目	内容
パラグラフ数	218 パラグラフ
総文数	591 文 (1 パラグラフ当たり平均 2.71 文)
総文字数	63,896 字 (1 文当たり平均 108.1 字)

3 東証上場企業と英文開示

東証は 3,600 社を超える企業 (2020 年 1 月時点) が上場している世界最大の証券市場の一つであり、海外投資家に対するプレゼンスは年々高まっている。日本の上場企業の株式に対する外国法人の保有比率 (時価総額に対する金額ベース) は 2019 年現在では約 30%^{*2}であるほか、海外投資家の取引が株式の売買代金に占める割合は 2018 年時点で約 59% である^{*3}。東証上場企業は、投資家の投資判断に影響を与える情報を、適時開示資料として適時適切に開示することが義務付けられている。しかしながら、2018 年に上場企業が公表した適時開示資料では、英語の適時開示資料は全体の約 11% に留まる。

一方で、情報量、翻訳者のリソース、時間の制約により、全ての開示資料を英訳することは東証上場企業にとって負担が大きいと考えられている [12]。まず、上場企業開示資料の文量は多く、2018 年における日本語の適時開示資料及び縦覧書類の合計ページ数は 700,000 を超えており、平均すると各東証上場企業は毎年 194 ページ以上の文書を日本語で開示している。また、多くの東証上場企業は資料を開示する時期が集中しており、翻訳者のリソースを確保することは難しい。一例として、2018 年 5 月 9 日から 15 日までの 1 週間に、全体の 6 割超の東証上場企業によって計 45,000 ページ以上の決算短信が開示されている。更に、外国人投資家にとっては日本語と英語の資料の開示にタイムラグがあるのを嫌うため、東証上場企業に対して日英の開示資料を同時に開示することを要求する傾向にある。このような問題から、全ての開示資料の英訳の需要を人手翻訳だけで満たすことは容易ではなく、機械翻訳の活用が解決策の一つとして挙げられる。

4 プリエディット作業

4.1 対象データ

本研究で対象とするデータは、2019 年 7 月までに日本語及び英語で開示された CG 報告書からパラグラフ単位で抽出した、全 591 文のプレーンテキストとした。CG 報告書とは各上場企業が自社のコーポレート・ガバナンスの状況について記載する開示資料であり、投資家の投資判断における重要な情報源の一つとされている [13]。CG 報告書からの文の抽出に当たっては、パラグラフ中に 100 文字以上の文が最低 1 つ含まれるよう、「基本的な考え方」、「コーポレートガバナンス・コードを実施しない理由」、「コーポレートガバナンス・コードの各原則に基づく開示」の 3 項目から抽出した。この文量等を表 1 に示す。

*2 <https://www.jpx.co.jp/markets/statistics-equities/examination/01.html>

*3 <https://www.jpx.co.jp/markets/statistics-equities/investor-type/00-02.html>

表2 本研究で利用した言い換えルール

ルール	名称	概要
1	短文化	一文を短くして複雑な係り受けを無くす。
2	主語・目的語等の明示	主語や目的語等の省略された要素を明示して、文意を明瞭にする。
3	文構造の明確化	文の構造を工夫し、理解しやすい文にする。
4	並列要素の整理	要素を対等に並べることで、係り受けを明確にする。
5	読点の付与	適切な位置に読点を付与することで、係り受けを明確にする。
6	簡潔化	不要、冗長、難解な表現を避け、文意が変わらない簡潔な表現を用いる。
7	単語の言い換え	多義的な表現や曖昧な表現を明確・具体的な表現に改める。

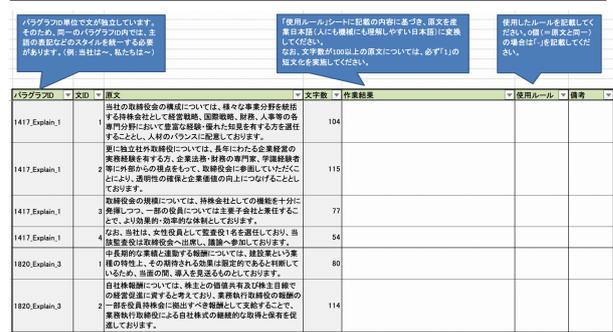


図2 言い換え作業ファイルのサンプル

4.2 言い換えルール

本研究ではプリエディット手法として、特許ライティングマニュアルに記載されている言い換えルールを利用した。この概要を表 2 に示す。

特許ライティングマニュアルでは、特許文書における産業日本語の書き方が紹介されている。この言い換えルールに従って文を修正することで、人に理解しやすく、かつ、機械にも処理しやすい日本語 (産業日本語) となることが期待される。

4.3 作業実施者

作業者間における作業結果の差異を分析するため、本研究ではクラウドワーカー (5 名) と翻訳者 (1 名) に同様の作業を依頼した。クラウドワーカーの要件には、社会人経験が 1 年以上であり、かつ、翻訳実務の経験が無いことを設けた。翻訳者の要件には、ISO 17100 [14] を参考に、次のいずれかを満たすことを設けた。

- 翻訳に関する高等教育の学歴を有すること
- 翻訳以外の分野での高等教育の学歴と 2 年間の翻訳実務経験を有すること
- 5 年間の翻訳実務経験を有すること

4.4 作業内容

各作業者には、特許ライティングマニュアルを参考にした各文の言い換え、使用した言い換えルールの記載、アンケートへの回答を依頼した。作業結果の記入に当たっては、図 2 のように、あらかじめ作業対象の文等を記載した Excel ファイルを使用した。また、言い換え作業の参考資料として、CG 報告書の言い換え文のサンプルを 18 文作成し、各作業者に連携した。

なお、特許ライティングマニュアルにて、ルール 1 の短文化を行うべき文字数の目安は 100 字と示されていることを踏まえ、100 字以上の文についてはルール 1 の短文化を必ず実施するよう作業仕様書等に明記した。

表3 プリエディットにおける適用ルール数(「ルールX」を「R.X」として表記)

対象	R.1	R.2	R.3	R.4	R.5	R.6	R.7	N/A
A	342	90	62	36	108	20	67	160
B	369	238	31	2	45	44	292	47
C	333	45	32	11	82	13	64	147
D	440	498	287	86	216	229	309	6
E	379	22	91	22	88	11	15	165
翻訳者	348	430	76	273	217	50	189	53

表4 内容の伝達レベルの評価指標

スコア	内容
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%~)
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%)

4.5 作業結果と考察

CG 報告書から抽出した 591 文を対象に、クラウドワーカーによる言い換え文を 5 名分、翻訳者による言い換え文を 1 名分作成した。本稿では各クラウドワーカーを A から E のラベルで区別する。各作業者のプリエディットにおける適用ルール数の分布を表 3 に示す。

表 3 より、ルール 1 以外の適用数は作業者による差異が大きいと言える。この要因として、ルール 1 には文字数という定量的な基準が設けられていること、それ以外のルールには適用すべきかどうかの定量的な基準が無いことの 2 点が考えられる。

5 実験

5.1 実験設定

本実験では、原文と各言い換え文を汎用的な NMT サービスに入力し、その翻訳品質を評価した。汎用的な NMT サービスには Cloud Translation API^{*4} を利用した。本稿では NMT サービスの出力文を MT 文と記載する。

MT 文の品質評価には自動評価と人手評価を行った。自動評価手法には BLEU [15] 及び NIST [16] を利用し、文単位でそれぞれのスコアを算出した。人手評価手法には、特許庁が「特許文献機械翻訳の品質評価手順^{*5}」にて公開している「内容の伝達レベル」を利用した。この評価では、文単位で 2 名の翻訳者による 5 段階の主観評価を行い、それぞれの平均値を算出する。この主観評価の指標を表 4 に示す。なお、クラウドワーカーにおける人手評価では、5 名分全ての訳文ではなく、パラグラフ単位でランダムに抽出した 1 名分の MT 文を対象とした。

その後、クラウドワーカーと翻訳者における機械翻訳文の改善の差異について検証するため、各クラウドワーカーと翻訳者の人手評価結果を比較した。

また、各言い換えルールの有効性の評価のため、クラウドワーカー及び翻訳者の人手評価結果に対して重回帰分析を行った。この時、目的変数は原文との人手評価結果の差分、説明変数は各ルールの適用状況(適用している場合は 1、それ以外は 0)とした。

表5 自動評価結果

対象	BLEU			NIST		
	スコア	差分	p 値	スコア	差分	p 値
原文	22.98	-	-	2.66	-	-
A	24.13	1.14	<0.01	2.76	0.10	<0.01
B	23.70	0.72	<0.05	2.78	0.11	<0.01
C	23.93	0.94	<0.01	2.76	0.09	<0.01
D	23.37	0.38	0.36	2.76	0.10	<0.01
E	23.55	0.56	0.05	2.75	0.08	<0.01
翻訳者	23.86	0.88	<0.05	2.77	0.10	<0.01

表6 人手評価結果

対象	スコア	差分	p 値
原文	3.78	-	-
クラウドワーカー	4.13	0.35	<0.01
翻訳者	4.23	0.45	<0.01

表7 人手評価結果における作業者間の比較

対象	文数	スコア	翻訳者のスコア	差分	p 値
A	107	4.06	4.21	-0.15	<0.01
B	125	4.24	4.23	0.01	0.89
C	119	4.05	4.21	-0.16	<0.05
D	122	4.18	4.30	-0.13	<0.05
E	118	4.14	4.22	-0.08	0.27

表8 重回帰分析の結果 ($R^2 = 0.20$ 、文数 = 1182)

ルール	名称	係数	標準誤差	t 値	p 値
1	短文化	0.73	0.05	15.8	<0.01
2	主語・目的語等の明示	0.00	0.05	-0.07	0.94
3	文構造の明確化	0.00	0.06	0.03	0.98
4	並列要素の整理	0.11	0.05	2.13	<0.05
5	読点の付与	0.02	0.05	0.46	0.64
6	簡潔化	0.07	0.07	0.96	0.34
7	単語の言い換え	0.01	0.05	0.11	0.91

5.2 実験結果と考察

自動評価の結果を表 5、人手評価の結果を表 6、人手評価結果におけるクラウドワーカーと翻訳者の比較を表 7、重回帰分析の結果を表 8、言い換え文及び MT 文のサンプルを表 9 に示す。各表における p 値は両側検定によって算出した。

表 5 より、BLEU スコアについては一部の作業者では有意な向上が見られなかったものの、NIST スコアについては全ての作業者において有意差が確認できた。NIST スコアの方がより有意に向上していることから、内容語の出力が特に改善したと考えられる。また、表 6 より、クラウドワーカーと翻訳者の両方で人手評価の有意な改善が確認できた。これらの結果は、特許ライティングマニュアルによるプリエディットが、機械翻訳の出力文の品質向上に寄与することを支持している。

表 7 より、クラウドワーカー A、C、D のプリエディット文の MT 文では翻訳者とのスコアに有意差が確認できたものの、クラウドワーカー B、E では有意差が確認できなかった。本研究では、翻訳者の作業結果は 1 名分しか対象としていないため、クラウドワーカーと翻訳者間の差異の詳細については別途分析が必要である。

次に、各言い換えルールについて分析する。表 8 より、人手評価結果の向上に特に寄与した言い換えルールは、ルール 1 の短文化だと言える。この要因として、係り受けの明確化や訳抜けの抑制が考えられる。表 9 の例 1 のように、短文化では文節の係り受けが明確に

*4 <https://cloud.google.com/translate/?hl=ja>

*5 https://www.jpo.go.jp/system/laws/sesaku/kikaihonyaku/document/tokkyohonyaku_hyouka/01.pdf

表9 言い換え文及び MT 文のサンプル (括弧中の文字は言い換え作業、数値は人手評価結果を示す)

例 1	原文	なお、当社は、女性役員として監査役1名を選任しており、当該監査役は取締役会へ出席し、議論へ参加しております。
	参照訳	The Company has appointed one female officer as a corporate auditor, and the corporate auditor attended meetings of the Board of Directors and participated in discussion.
	原文 MT 文 (4.0)	The Company has appointed one Audit & Supervisory Board Member as a female officer who attends the Board of Directors meetings and participates in discussions.
	言い換え文 (翻訳者) 言い換え MT 文 (5.0)	なお、当社は、女性役員として監査役1名を選任しております。当該監査役は取締役会へ出席し、議論へ参加しております。 The Company has appointed one corporate auditor as a female director. The corporate auditor attends the board of directors and participates in discussions.
例 2	原文	当社は、任意の指名・報酬委員会などの独立した諮問委員会を設置していませんが、取締役候補の選任や取締役の報酬については、取締役会の決議に先立ち、独立社外取締役及び親会社に対し説明を行い、適切な助言を得ております。
	参照訳	Although NTT DATA has not set up an independent advisory committee such as a voluntary nomination committee or remuneration committee, in advance of the resolution by the Board of Directors, we provide independent directors and the parent companies with explanations about the nomination of candidates for directors and remuneration for directors and receive appropriate advice from them.
	原文 MT 文 (2.0)	The Company has not established an independent advisory committee such as a voluntary nomination and remuneration committee. We give explanations and obtain appropriate advice.
	言い換え文 (D) 言い換え MT 文 (5.0)	当社は、任意の指名・報酬委員会などの独立した諮問委員会を設置していません。ですが、当社は、取締役候補の選任や取締役の報酬については、取締役会の決議に先立ち、独立社外取締役及び親会社に対し説明を行い、適切な助言を得ています。 The Company has not established an independent advisory committee such as a voluntary nomination and compensation committee. However, prior to the resolution of the Board of Directors, the Company explains the selection of director candidates and director remuneration to independent outside directors and the parent company to obtain appropriate advice.
例 3	原文	引き続き、幅広い見識に基づく助言や提言を行っていただくことを期待し、2019年の第96期定時株主総会にて取締役に再任いたしました。
	参照訳	To continue receiving advice and recommendations toward management from these wide insights, he was reappointed as External Director at the 96th annual general meeting of shareholders held in 2019.
	原文 MT 文 (5.0)	With the expectation that he would continue to provide advice and recommendations based on a wide range of insights, he was reappointed as Director at the 96th Ordinary General Meeting of Shareholders in 2019.
	言い換え文 (翻訳者) 言い換え MT 文 (3.5)	当社は、引き続き、幅広い見識に基づいて助言し、提言していただくことを期待し、同氏を2019年の第96期定時株主総会にて取締役に再任いたしました。 In the hope that we will continue to provide advice and suggestions based on a wide range of insights, we reappointed him as a director at the 96th Annual General Meeting of Shareholders in 2019.

なることで、NMTにとって文意が把握しやすくなるのが期待される。また、表9の例2のように、NMTでは長文における訳抜けの発生が知られているが、短文化によってその抑制が期待される。

短文化と同様に、文節の係り受けの明確化に繋がり得る言い換えルールとして、ルール3の文構造の明確化、ルール4の並列要素の整理、ルール5の読点の付与が挙げられる。しかしながら、表7より、文構造の明確化及び読点の付与については有意な寄与が確認できなかった。この要因として、ルールの過剰な適用や他のルールの考慮漏れによる人手評価結果の悪化が考えられる。表9の例3では、過剰な読点の付与によって文の明瞭性が損なわれていることに加え、ルール2に従って主語を補足しているものの、その係り受けの距離が離れており、ルール3の考慮が漏れている。

また、作業者のアンケート結果では、ルール3-7の適用可否の判断や言い換えが難しいといった意見が寄せられた。開示資料に即した適用の基準や言い換え例の充実により、これらの作業品質の改善が期待される。

なお、ルール2の主語・目的語等の明示において、誤った主語・目的語の補足により誤訳が生じている事例を複数確認した。この要因として、文書の作成者ではない第三者が、主語・目的語を誤解したことが挙げられる。従って、開示資料の作成者がリエディットを行う際には、この問題は生じにくいと考えられる。

6 おわりに

本研究では、開示資料のNMTにおけるリエディットの有効性について検討した。実験の結果、CG報告書を対象にしたリエディットにおいて、特許ライティングマニュアルの言い換えルールが準用可能であること、また、短文化が特に有効であることを示した。一方、複数のリエディット結果において、作業者間における品質の差異や、不適切な言い換えルールの適用が確認された。これらの改善に向け、開示資料に即した言い換え例の拡充が望まれる。

以上を踏まえ、今後の課題としては次が挙げられる。

- 開示資料に即した、リエディットマニュアルの作成 (言い換え例の拡充)
- 他の上場企業開示資料におけるリエディットの有効性の検証
- ポストエディット及びリエディットを前提とした開示資料の産業翻訳ワークフローに向けた検討

謝辞

本研究において、東京大学先端科学技術研究センターの田中久美子教授に有益なご助言を戴いた。ここに記して謝意を表する。

参考文献

- [1] Ilya Sutskever, Oriol Vinyals, and Quoc V. Le. Sequence to sequence learning with neural networks. In *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*, NIPS'14, pp. 3104-3112, Cambridge, MA, USA, 2014. MIT Press.
- [2] 土井惟成, 近藤真史, 山藤敦史. コーポレート・ガバナンス報告書における機械翻訳の検討. 言語処理学会第25回年次大会 (NLP2019), pp. 926-929, 3 2019.
- [3] 内山将夫. 適時開示専用 NMT. AAMT 2019, Tokyo ~ 機械翻訳最前線~, 11 2019.
- [4] 松本智子. 翻訳会社として、金融・IR分野の特性を踏まえたソリューションに関する見通し. AAMT 2019, Tokyo ~ 機械翻訳最前線~, 11 2019.
- [5] 三輪哲也. 開示書類翻訳における MT の可能性. AAMT 2019, Tokyo ~ 機械翻訳最前線~, 11 2019.
- [6] ISO 18587:2017. <https://www.iso.org/standard/62970.html>.
- [7] 一般財団法人日本特許情報機構特許情報研究所. 特許ライティングマニュアル (第2版)「産業日本語」, 第2版, 3 2019.
- [8] 清藤弘晃. 産業日本語の考え方と NMT. 第29回 JTF 翻訳祭, 10 2019.
- [9] 佐藤紗都, 上垣外英剛, 高村大也, 奥村学. ニューラル機械翻訳の品質向上に向けた原言語における言い換え. 言語処理学会第25回年次大会 (NLP2019), pp. 1467-1470, 3 2019.
- [10] 早川威士. 機械翻訳に対するリエディットのタスク分類に関する考察. AAMT 2019, Tokyo ~ 機械翻訳最前線~, 11 2019.
- [11] 宮田玲, 藤田篤. 機械翻訳向けリエディットの有効性と多様性の調査. 通訳翻訳研究への招待, pp. 53-72, 2017.
- [12] 山藤敦史. 開示資料と翻訳. AAMT 2019, Tokyo ~ 機械翻訳最前線~, 11 2019.
- [13] 投資家フォーラム. 投資家フォーラム-第1・2回会合-報告書. <https://investorforum.jp/>. (参照 2020-01-15).
- [14] ISO 17100:2015. <https://www.iso.org/standard/59149.html>.
- [15] Kishore Papineni, Salim Roukos, Todd Ward, and Wei jing Zhu. Bleu: a method for automatic evaluation of machine translation. pp. 311-318, 2002.
- [16] G. DODDINGTON. Automatic evaluation of machine translation quality using n-gram co-occurrence statistics. *Proceedings of the HLT Conference, San Diego, California, 2002*, pp. 138-145, 2002.