

生成AIの時代に翻訳者が主導する新しい翻訳システム —A Translator-Led Translation System in the Era of Generative AI—

ローズ三浦 | Yukiko Miura (エッジ・トランスレーション)
津山逸 | Nogaru Tsuyama (NT Chougenbou's Medical Translation Taskforce)

【緒言】

さまざまな分野の言語サービスに人工知能（AI）の応用が進んでいる。医療関係でも、従来の機械翻訳にAIを組み合わせたシステムがすでに商業的に利用されており、一般向け医療機器の取扱説明書から高度に専門的な学術論文まで、本稿では、具体的な例文を挙げて、ポストエディット（以下、「監訳」という）を効率的に実施する方法を探る。

【方法】

津山がこれまでの実務に基づいて選択的に蓄積してきた翻訳データから、上記目的に適した日英各3個の短いテキストを選んだ。これらを原文として、一般に公開されている汎用翻訳エンジン「Google翻訳」と、三浦が開発した医療特化型AI翻訳システム「生成AIローズ三浦モデル」に掛けた。得られた出力文を津山が自身の経験に基づいて監訳した。

【結果】

6個の原文に対するMTの出力、それらに対して監訳を行った結果と要点を以下に示す。問題のある文字列は黄色で、改良後の文字列は緑色でハイライト表示する。※Google : Google翻訳、三浦モデル : 生成AIローズ三浦モデル

英日訳例1

Microscopy found irregularities in three samples.	Google翻訳 : 顕微鏡検査により、3つのサンプルに不規則性が見つかりました。 三浦モデル : 顕微鏡検査により、3つの検体に異常が認められた。 監訳 : 鏡検で3つの検体に異常を認めた。
---	---

【要点】誤訳「不規則性」を医学分野で異常所見を意味するirregularitiesに修正した。全体の言い回しも適切に直した結果、文字数を大幅に減らせた（30字／23字→16字）。

英日訳例2

She reported feeling off balance after standing, though vital signs remained within normal limits.	Google翻訳 : 彼女は、バイタルサインはまだ正常範囲内であったものの、立ち上がった後にバランスが崩れたように感じたと報告した。 三浦モデル : 患者は起立後にバランスが取れない感じがあると報告したが、バイタルサインは正常範囲内に留まっていた。 監訳 : 起立後にふらつきを訴えたが、バイタルは正常範囲内だった。
--	---

【要点】Google訳は意訳的には正しいが、医学文献における記述の標準的スタイルに反している。用語を書き換えるとともに、前半と後半の情報を提示する順序を入れ替えた。Sheに対する訳語を省いた。語数を大幅に減らせた（55字／49字→28字）。

英日訳例3

Our devices are supported for at least five years after the device model is no longer sold.	Google翻訳 : 当社のデバイスは、デバイスマodelの販売が終了してから少なくとも5年間サポートされます。 三浦モデル : 当社の機器は、機器modelの販売終了後、最低5年間はサポートされます。 監訳 :弊社機器のサポートは、モデルの販売終了後少なくとも5年間は有効です。
---	---

【要点】原文の訴求点はデバイス自体の性能や特長ではなくサポート期間の長さである。デバイスではなくサポートを主語とする文に書き換えたところ、語数を減らすことができた（43字／34字→34字）。

日英訳例1

血友病はX染色体連鎖劣性遺伝疾患であり、患者はほぼすべてが男性である。	Google翻訳 : Hemophilia is an X-linked recessive disorder that affects almost all men. 三浦モデル : Hemophilia is an X-linked recessive genetic disorder, and nearly all patients are male. 監訳 : Hemophilia is an X-linked recessive hereditary disease, and nearly all patients are male.
-------------------------------------	--

【要点】MT出力に含まれている不正確な情報を正した。まずは誤訳を正すことが大事なので、語数の微増は問題ない。（11語／13語→13語）。

日英訳例2

試料の保存条件は25±2°C、60±5%RH、暗所とし、正立で保存した。	Google翻訳 : The storage conditions were 25±2°C, 60±5% RH, and dark, and the samples were stored upright. 三浦モデル : The samples were stored upright under conditions of 25±2°C, 60±5% RH, in the dark. 監訳 : The samples were stored upright at 25±2°C and 60±5% RH in the dark.
--------------------------------------	--

【要点】情報の流れを重視して複文を単文に書き換えたことで、流暢性が上がり、語数も減った。（15語／14語→13語）。

日英訳例3

本品は海外でも中等度認知症患者を対象とする第I相臨床試験を実施した。	Google翻訳 : This product has also undergone Phase I clinical trials overseas targeting patients with moderate dementia. 三浦モデル : This product was also subjected to Phase I clinical trials targeting patients with moderate dementia overseas. 監訳 : Phase I clinical studies of this product were additionally conducted outside Japan in patients with moderate dementia.
------------------------------------	--

【要点】この原文で「本品は」は日本語にありがちな話題導入のための見た目の主語であり、そのまま英文の主語にすると不自然な訳語になってしまふ。「対象とした」は、動詞を使わずに前置詞inだけで表すことができる。語数は増えたが問題ない（15語／16語→17語）。

【考察】

- 一般的に日英翻訳の目的は、原文が伝える情報を正確、明確、簡潔（いわゆる3C）に英文化することであるが、さまざまな日本語文書で普通に見られる曖昧かつ非論理的な記述が、その達成を妨げる。主語なし、目的語なし、論理の飛躍などの要因を含んだ日本語文書が巷に溢れていることが、日英MTの一つの障壁となっている。この問題はAIを活用することで、すでにかなり解決できていると思われるが、結局は人間がMT出力を編集する「監訳」が必要である。
- 英日MTでは、構造化されていない曖昧な記述に慣れている読者を納得させる自然な日本語表現が重要である。多様な語彙や表現を瞬時に提示することができる生成AIとの対話が、訳文最適化の助けになる。
- 「ポストエディット」という概念を言語コミュニケーション行為の一つとして「翻訳」や「通訳」と等位的に位置付けるために、この冗長で関係者の受けが良くない用語に替えて、既存の言葉「監訳」を転用することを提案する。
- 機械翻訳研究者・開発者側から見ると、高レベルの翻訳者と緊密に連携し、翻訳者からのフィードバックを迅速に機械翻訳システムに実装することが、生成AI時代の新たな翻訳システムの要となる。また、翻訳者側から見ると、機械翻訳の技術開発に主体的に関与し、AIを自らの能力拡張として活用することで、翻訳スピードや効率を飛躍的に向上させることができる。本研究が提示した翻訳者主導型システムは、生成AI時代における翻訳の新たなパラダイムを示すものである。