

科学研究費助成事業（基盤研究（S））公表用資料
〔令和3（2021）年度 中間評価用〕

令和元年度採択分
令和3年3月31日現在

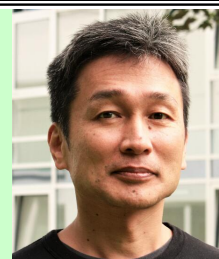
翻訳規範とコンピテンスの可操作化を通じた翻訳プロセス・モデルと統合環境の構築

Developing a translation process model and constructing an integrated translation environment through detailed descriptions of translation norms and competences

課題番号：19H05660

影浦 峡 (KAGEURA Kyo)

東京大学・大学院教育学研究科・教授



研究の概要（4行以内）

翻訳プロセスをアクタの行為とアイテムの操作として高解像度で記述し、行為と操作を翻訳規範とコンピテンスに対応づけた翻訳プロセス・モデルを構築する。モデルの記述と実務・教育プロセスでの利用のためにメタ言語を構築し自動化可能な操作の開発・実装を行う。これらを統合環境に組み込み、それを介して、モデルの有効性を教育・実務・技術の観点から評価する。

研究分野：応用情報学

キーワード：翻訳テクノロジー、メタ言語、翻訳プロセス研究、機械翻訳

1. 研究開始当初の背景

翻訳論において翻訳規範とコンピテンスの理論が発展してきたが、実務や教育には十分活用されていない。翻訳教育において学習されるべきコンピテンスは外的プロセスを明示的に含むようになってきたが対応するカリキュラムの構築は十分ではなく翻訳者要請の質的・量的需要に応えられていない。翻訳産業は急速に拡大しているが、翻訳実務において高品質の翻訳は熟練した翻訳者の暗黙知に依存している状況がある。技術領域ではニューラル機械翻訳（NMT）は特に訳文の流暢さを大幅に向上させ、また、翻訳プロセスで活用可能な言語処理・文書処理技術も進展してきているが、それらの利用と翻訳実務と技術の関係について共通理解は十分でなく実践のモデルも確立していない。

理論・教育・実務・技術の各面で知見と技術は展開しているが、翻訳実践全体としてはそれに応じた展開が十分に見られていない。拙速なNMTの利用による問題など報告されている。関与するアクタが「翻訳とは何か」を十分に共有していないことからこうした状況の背後には、翻訳のプロセスと規範・コンピテンスが共有可能な外在的操作として具体的に記述されてこなかったことがある。

2. 研究の目的

本研究では、(1) 翻訳プロセスに翻訳規範とコンピテンスの要素を対応づけた翻訳プロセス・モデルを構築すること、(2) 翻訳プロ

セス・モデルを組み込んだ統合的環境をシステムとして構築し、翻訳教育と翻訳実務における有効性を実証的に検証すること、を目的とする。翻訳プロセス・モデルの記述はアクタの行為とアイテムに対する操作、アクタ間のコミュニケーションという二つのレベルで行い、その際に、モデルの記述を担うとともに実務・教育プロセスで運用可能なメタ言語の体系を構築する。翻訳プロセスにおいて自動化可能な部分を定義し自動化手法を開発する。研究終了時までモデルとメタ言語、環境、データをリリースする。

3. 研究の方法

翻訳プロセス・モデルとメタ言語は、関連領域の文献レビューと翻訳産業界への聞き取り・アンケートの結果を統合し繰り返し精緻化することで構築する。規範とコンピテンスは最終品質基準を起点に各プロセスの要素と関連づける。関連する技術開発は必要技術に応じて主にデータに基づく機械学習を利用する。統合的環境は本研究グループの主要メンバーでこれまで開発してきた環境を起点とし、翻訳プロセス・モデルと実利用環境の双方を考慮したシステムを実現するためにユースストーリーマッピングの手法を用いる。

4. これまでの成果

(a) 翻訳実務プロセスの明確化

プロジェクト・フローレベルでは、ISO17100を起点にステップとアクタを詳細化すると

もに MT の位置付けを明確にした。起点言語文書に対して文書属性と文書要素を同定し翻訳における記号操作の要件を明らかにした。一部は自動化タスクとして定式化し手法の開発と学習データ構築を進めた。アクタ間コミュニケーションについては翻訳業界における主観語彙の調査と定式化、プロジェクト・シナリオ、翻訳教育におけるクラスルーム・シナリオの定義を行なった。

(b) 規範とコンピテンスの可操作化

起点言語文書属性と要素を考慮し翻訳プロセスに対応した規範とコンピテンスを明らかにし、一部を自動化タスクとして定式化した。

(c) メタ言語の構築

(a)(b)の定式化を進めつつ、プロセス語彙約 180 語、文書属性と文書要素語彙約 300 語、方略語彙約 30 語、効果語彙約 30 語、評価語彙約 50 語からなるメタ言語の体系を構築した。

以上(a)(b)(c)の成果は論文として発表している他に、本研究の参加者と協力者による編著書 *The Technologicalisation of Translation: Developing and Using Metalanguage in the Translation Process* としてまとめ、現在の、英国の学術出版社で査読中である。また、日本通訳翻訳学会・言語処理学会でのテーマ講演、複数の大学における翻訳授業での試験利用を通して社会展開を進めている。

(d) 自動化技術の開発・実装

起点言語文書属性・要素分析に関しては、文書属性コーパスと文書要素タグ付きコーパスの構築、書体や句読点等のパラ言語要素の機能・意味の整理とコーパス設計を進めた。

目標言語文書要素への変換手法に関して、効果メタ言語の構築および修正トリガの定式化と学習用に MT+X 文 Y 訳事後編集 (PE) + 翻訳 PE の対照コーパスを構築した。また、独立言語要素としての専門用語は文書における同定だけでなく別のデータ管理が翻訳において必要となるため、独立タスクとして専門語彙における概念体系の情報を組み込んだ多言語専門語彙構築の手法を開発した。また翻訳者の行為をエミュレートする生成・検証型多言語語彙確認システムの試行運用を開始した。

NMT に関しては、(a)(b)(c)の知見を踏まえ、文書単位 MT の実装、非翻訳要素の組み込み、内容語注目型 NMT の開発、低資源学習と高資源学習の差異を通じた NMT の分析的診断、事前編集を介した NMT の品質評価、タグ付き逆翻訳によるデータ拡張の効果の評価を進めている。また、文書単位 MT 用コーパス ParaNatCom は、課題共有型ワークショップ WAT のタスクに活用されている。

(e) 統合的翻訳環境・翻訳学習環境の構築

過去に代表者・分担者で開発してきた「みんなの翻訳実習」を起点にユーザストーリーマッピングを通して翻訳プロセス・モデルと活用形態を整理し、学習用にクラスルーム・レイヤを導入した。またメタ言語接続モジュール

を実装した。これに基づく新システム第一版の試験稼働を開始した。

5. 今後の計画

今後も引き続き、当初の計画に従って研究を進める。翻訳プロセス・モデルとメタ言語については当初予定通り第 2 年度までで全体の構築を終えたので、統合的翻訳環境に組み込み、実証的な検証と精緻化の段階に入る。自動化手法については起点言語文書分析、個別起点言語文書要素処理の MT との接合を継続するとともに、自動評価を展開する。統合的翻訳環境についてはユーザストーリーマッピングを反映した環境を実装し、それを用いて実証実験と評価・フィードバックに基づくプロセス・モデルとメタ言語の精緻化を進める。コロナウィルスの感染状況により第 3 年度以降も国際的な実証ワークショップを対面で開催することが困難である可能性があるため、オンラインによる国際実証実験の計画を立て実施する予定である。

6. これまでの発表論文等 (受賞等も含む)

Miyauchi, T., Miyata, R. and Kageura, K. "Constructing a metalanguage for analyzing source documents in translation practice," *III International Conference EnTRetextos* (2021).

Kageura, K., Yamada, M. and Abekawa, T. "Developing a translator training platform by clarifying translation process through user story map," *Translation 4.0: Training, Research and Practice*, pp. 187-201 (2020).

Yamada, M., Yamamoto, M., Onishi, N., Fujita, A., Miyata, R. and Kageura, K. "Metalanguage for the translation process," *Translation in Transition (TT5)*, pp. 46-51 (2020).

Chen, K., Wang, R., Utiyama, M. and Sumita, E. "Content word aware neural machine translation," *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 358-364 (2020).

Marie, B., Rubino, R. and Fujita, A. "Tagged back-translation revisited: Why does it really work?" *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 5990-5997 (2020).

Nanami, O. and Yamada, M. "Why translator competence in information searching matters," *Invitation to Interpreting and Translation Studies*, 22, pp. 1-23 (2020).

Piao, H., Han, S. and Kageura, K. "The use of meta-language in translation revisions," *2019 International Conference on Translation Education* (2019).

7. ホームページ等

<https://tntc.p.u-tokyo.ac.jp/>

<http://mnh-tt.org/>