

日本の翻訳業界における MT 使用の現状と課題

阪本 章子

英ポーツマス大学

1. 調査の概要

英ポーツマス大学と関西大学の共同で、2018年6～7月、翻訳会社でプロジェクト管理に従事する22人を対象に、グループインタビュー(フォーカスグループ)形式の聞き取り調査を行った。翻訳の現場でテクノロジーがどのように使われ、それに対し現場の人たちがどう考え、問題があればどう対応しているかを定性的に調べるのを目的とした。

グループインタビューは東京で3回、大阪で1回開催した。各グループ4～7人に6つのトピックについて2時間話し合ってもらった。トピックはCATツール、機械翻訳(MT)、テクノロジー教育・トレーニング、クラウドソーシング翻訳サービス、コミュニケーション手段、そして今後の重要課題についてだった。参加者の選考基準は、日本の翻訳会社でプロジェクト管理/コーディネイト業務を担当または管理している人とした。小規模な会社で参加者が経営者でもあるケースが1件あったが、調査対象者には、翻訳会社の代表としてではなく、業務担当者個人としての参加を依頼した。これは、同じ翻訳会社の中でも経営サイドと制作で意見が異なることが多いからである(Vieira and Alonso 2019)。あくまでも、翻訳案件に直接関与している人たちの意見を集めることを目的とした。参加者の勤務先の規模は平均85人(2～400人)、キャリア歴は平均8.6年(1～20年)だった。なお本稿では職務の名称にかかわらず、参加者をPM(プロジェクトマネージャー)と呼ぶ。

グループインタビューでの発言は、テープ起こしをしたあと、そのテキストをグラウンデッド・セオリー・アプローチ(Glaser and Strauss 1967)を用いて分析

した。Richard and Morse (2007, 177-183) に従い、まず、テキストに現れたコンセプトに応じてテキストの該当部分にコードをつけた(オープン・コーディング)。次に、頻繁に表れるコンセプトを集めてより抽象的な概念を抽出した(カテゴリー・サーチ)。そして抽出された概念ごとに結論を導いた。分析の信頼性を高めるため、Researcher-in-instrument 手法(Anderson et al. 2016)を用いた。具体的には、一人の調査員がまず分析を行い、もう一人の調査員が分析内容をチェック・確認した。調査員の一人は翻訳者、もう一人は翻訳会社経営者の経験があるため、二つの異なる視点からの分析を行い、相違点があれば合意に至るまで話し合った。

調査結果の詳細は、阪本・山田(2019)で報告した。本稿では本調査のうち、MT使用に関連する部分から、特に翻訳にかかわる「人間」の側面に関連し、かつ問題や課題を含む部分について報告する。

2. MTを組み入れた翻訳ワークフロー

MTを使った翻訳ワークフローは翻訳会社により様々だが、典型的な例として、以下の2つが考えられる。

●フリーランスの翻訳者が、人手翻訳の案件のなかで、自らの翻訳作業の一部としてMTを使用する場合。

●MTPEの案件のなかで、ポストエディターがMT出力をポストエディットする場合。

またそれ以外に、CATツールのなかで作成された翻訳メモリー(TM)のバイリンガルデータがMTの学

The current situation of and challenges faced by MT use in Japan's translation industry

Akiko Sakamoto

University of Portsmouth (UK)

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International Public License. License details: <https://creativecommons.org/licenses/by-sa/4.0/>

習に使用されることがあるので、人手翻訳案件のなかでの CAT ツールの使用も、MT ワークフローに間接的に関連してくる。

そこで本稿では、翻訳者の個人的な MT 使用（下記セクション3）、TM データの管理（セクション4）、MTPE における課題、特にポストエディターの選定（セクション5）の3分野について議論する。

3. 翻訳者の個人的な MT 使用

調査では、人手翻訳の作業のなかでの翻訳者の個人的な機械翻訳使用、とくにオンラインの無料 MT の使用について話しあってもらった。実際にどれくらいの翻訳者が個人的に MT を使っているかについては、2016 年にイギリスの調査で 22% (The European Commission et al. 2017)、2019 年の欧州の調査 (ELIA et al. 2019) では Google Translate をよく使うと答えた翻訳者が 10%、DeepL が 6% と、低い値になっている。数値が低いのは、プロの翻訳者はこの質問に対し、Yes とは答えにくい事情があるからだと思われる。

翻訳者が CAT ツール内で MT をプラグインで使用して TM 内に記録が残れば、PM は翻訳者の MT 使用を把握することができる。しかし、それ以外の場合、例えば翻訳者が MT の Web サイトで得た MT 訳をコピー&ペーストすれば、翻訳者の個人的使用を PM サイドは把握できないし、把握するシステムがないのが現状だった。しかし PM は、多くの翻訳者が翻訳会社に申告せずに無料 MT (Google Translate など) を使用していると考えていた。使用者の割合の予想値は、30% から 80~90% まで、様々だった。そう考える理由として、「トライアル試験で似た文章を複数の翻訳者が出してきた」「こっそり使っている人が多いことを個人的に聞いている」という報告があった。

また、MT 訳そっくりな低質な翻訳が人手翻訳として納品されてくるなど、その影響はすでに業務に及んでいた。気付かずにそのまま翻訳を顧客に納品をして顧客の信頼を失うことを、PM たちは恐れていた。

しかし、翻訳者の MT 使用そのものについては、PM の個人的意見は分かっていた。プロなのだから機械に頼らず自分の手で翻訳してほしい、という意見もあれば、作業効率化のために使って当然、という意見もあった。例えば韓国語への翻訳では、漢字の弱い若い翻訳者にはかえて MT を使ってもらったほうが良い、という意見があった。

調査からは次の3つの課題が見えてきた。(1) MT の普及で「人手翻訳」の定義があいまいになっていること。ISO では人手翻訳 (ISO17100) と MTPE (ISO18587) を分けて定義しているが、これだけ無料 MT が普及している今、翻訳者の仕事場での線引きが非常に難しいこと。(2) 翻訳会社側では MT 使用をチェックするシステムがまだなく、MT 使用に関する条項を契約書に盛り込んでいる翻訳会社も、一部あったが、まだ少ないこと。(3) 「信用できる翻訳者は MT を勝手に使わないだろう」という、個人のモラルや倫理観に依存した管理方法になっていること。

そこで本研究では、つぎの3点を提言する。[1] MT 使用に関するルールをきちんと契約書に入れること。[2] もし翻訳会社が翻訳者に MT 使用を許可するならば、どの MT エンジンをもどのように使っているのかをはっきり伝えること。[3] また、トラブル回避のために、それを顧客に知らせるのかどうかを検討すること。

4. 翻訳メモリー (TM) の管理

TM のアップデートや保管は誰がすべきだと思うかについて、話し合ってもらった。

まず、TM は大切な翻訳資産で、PM は継続案件を取るために TM は非常に重要なもの、いわば「人質」のような存在だと認識していた。しかし、その保管場所については、契約書上で顧客と合意があっても、実際にはあいまいになっていたり、PM 自身が契約内容をきちんと把握していないケースもあった。

また TM のアップデートや管理の難しさを嘆く声があがった。TM を使ってできあがった翻訳に修正が加

わると、その修正結果を TM に反映させるべきだが、実際には翻訳を終わらせて納品するのに精一杯で、TM アップデートの作業が手つかずに終わってしまうということだった。

この現状は、従来型人手翻訳が常に抱える問題点と言えるが、将来の MT の普及を見据えた場合も、以下の2つの理由から、非常に重要な課題となる。まず、TM データは MT の学習データになりうるので、TM の質悪化は MT の質悪化につながる。また、MT の学習に TM データを利用する際の TM データの著作権や所有権は現状では非常にグレーな部分となっている。しかし、今後 MT が増えてくると、TM の倫理的・持続的なデータ利用は避けては通れない問題となるだろう (Moorkens and Lewis 2019)。

そこで本研究では、つぎの2点を提言する。[1] TM のメンテナンスを翻訳プロジェクトの一部として認め、契約内で十分な時間と予算を当てること。[2] 今後の MT 普及を見据えて、倫理的・持続的なデータ利用につながる意識改革を人手翻訳の TM データ管理のなかで徹底させること。TM データの質、そして所有権の重要性の認識を徹底させておくことが、今後の MT 発展において適切なデータ使用の意識を高めることにつながるからだ。

5. MTPE におけるポストエディターの選定

調査ではまた、MTPE が将来、従来型翻訳に取ってかわると思うかどうかについて、話し合ってもらった。この質問は、イエスかノーかの回答を得ることが目的ではなく、話し合いの中から MTPE 業務において何が課題になっているか探るのが目的だった。

話し合いの中から、つぎの4つが課題として浮かび上がった。①顧客からの MT に対する期待度にばらつきがあり、ひとつの業務プロセスで対応しきれないこと。②公平で妥当なプライシングモデルが確立されていないこと。③ポストエディットのガイドラインがあまりないこと。④ポストエディターに負担がかかっているこ

と。④ポストエディターの適切な人材確保が難しいこと。本項では④の人材の問題に焦点を当てる。

PM は、理想としては、優秀な翻訳者にポストエディットをやってもらいたいと考えていた。しかし実際には、優秀で経験のある翻訳者はポストエディットに拒否反応を起こす人が多く、頼みにくい状況があった。その対応策として、比較的経験の浅い翻訳者や、普段から安くて速いたぐいの仕事をよく受けてくれる人にポストエディットの仕事を outsourcing している、という PM が多かった。

優秀な翻訳者に頼めないならどういう人にポストエディットの仕事をしてもらいたいのか、という質問に対しては、次のような様々な意見が出た。① 現在チェッカーやレビューアーとして働いている人 ② 翻訳者でなくてもポストエディットのトレーニングを受けた人 ③ 語学力のある PM (実際にポストエディットを仕事の一部にしている PM もいた) ④ 個人ではなく、誤訳チェッカー+専門用語チェッカー+QA エンジニアがチームで担当 ⑤ クライアントがそのテキストのことをよく知っている場合は、クライアント。また、語学の好きな高校生でもできる、と考えている PM もいた。

この調査結果を先行研究の結果と照らし合わせた場合、懸念すべき点がひとつある。ニューラル機械翻訳 (NMT) は、流暢性の高い訳出が特徴とされるが、Yamada (2019)の調査では、大学生レベルの英語では流暢性の裏に隠れたエラーに気づけないことが指摘されている。この結果から、翻訳能力の低い翻訳者にポストエディットを頼んでもエラーに気づかないかもしれない、と推測するのは妥当だろう。MTPE の目的は「MT 訳を編集して修正すること」(The British Standards Institution 2017、引用は著者訳)であるから、ポストエディターがエラーを特定できなければ当然修正もできず、MT 訳にポストエディットをかける意味が全くなってしまふ。現行のポストエディターの人選に問題があることが伺えるため、早急に再検討すべきだろう。

6. 信頼できる MTPE 需要予想の必要性

MTPE の将来に関する話し合いからは、二つの意見が見えてきた。ひとつは、MT の質はどんどん上がりポストエディットは要らなくなるだろうというもの(すでに分野によってはいらなくなっている)。もうひとつは、一部の人手翻訳はなくなることはないだろう、というもの。

この二つの予想が両方正しければ MTPE はいずれはなくなるビジネスモデルだということになる。調査では、これが原因で、MTPE の現行の問題解決にリソースをかけるよりも、様子見をしている関係者が多いと見受けられた。そしてそれが現場、つまり PM や翻訳者・ポストエディターの負担になっているようだ。

現在のところ、MTPE の案件は翻訳の現場で増えている。いずれなくなるかどうかも含め、今現場で必要になるのは、信頼できる MTPE の需要予想ではないだろうか。しかし、機械翻訳のサプライ側だけではなく、翻訳者やポストエディター、サービスのユーザーの事情や視点もすり合わせなくては、現実的な予想は立てられない。技術が役に立つかどうかは、それを使う人がどのように使うかに左右されるからだ。これからは産学、そしてすべてのステークホルダーが協力してデータや情報をすり合わせる機会を作っていくことが大切になるだろう。

尚、本稿の詳細の完全版は、阪本・山田 (2019) の報告書にまとめられている。報告書はウェブサイトより入手できる。また、本内容に基づいた発表を、2019 年 11 月の AAMT Tokyo で行った。

参考資料

Anderson, R., Guerreiro, M., & Smith, J. (2016). Are all biases bad? Collaborative grounded theory in developmental evaluation of education policy. *Journal of MultiDisciplinary Evaluation*, 12(27), 44–57.

ELIA, EMT, EUATC, FIT, GALA, & LIND. (2019). *Expectations and concerns of the European Language Industry 2019*.

https://ec.europa.eu/info/sites/info/files/2019_language_industry_survey_report.pdf.

Accessed 10 August 2019

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.

Moorkens, J., & Lewis, D. (2019). Copyright and the re-use of translation as data. In M. O'Hagan (Ed.), *The Routledge Handbook of Translation and Technology* (pp. 469–481).

Abingdon, Oxon and New York: Routledge.

<https://doi.org/10.4324/9781315311258-28>

Richards, L., & Morse, J. M. (2007). *README FIRST for a user's guide to qualitative methods, 2nd edition*. London: SAGE.

The British Standards Institution. (2017). *BS ISO 18587:2017 BSI Standards Publication Translation services — Post-editing of machine translation output — Requirements*. London.

The European Commission, The CIOL, & The ITI. (2017). *2016/17 UK translator survey: Final report*. London.

<http://www.ciol.org.uk/sites/default/files/UKTS2016-Final-Report-Web.pdf>. Accessed 23

August 2017

Vieira, L. N., & Alonso, E. (2019). Translating perceptions and managing expectations: an analysis of management and production perspectives on machine translation.

Perspectives. 28(2), 163–184.

[https://doi.org/10.1080/0907676X.2019.164677](https://doi.org/10.1080/0907676X.2019.1646776)

6

Yamada, M. (2019). The impact of Google Neural

Machine Translation on Post-editing by
student translators. *The Journal of
Specialised Translation*, (31), 87–106.

阪本章子, & 山田優. (2019). 翻訳業界におけるテクノロジー使用の現状：現場の声から フォーカスグループ報告書. Portsmouth.
<https://translation.apple-eye.com>