

機械翻訳の現状と課題—新しい英文和訳法の提唱—

小島 正樹¹ 甲斐 基文²

1. 機械翻訳の現状と課題

最近の機械翻訳の精度向上は著しく、今後生成系 AI や大規模言語モデル (LLM) の適用によって、さらにその勢いは加速するものと思われる。従来の機械翻訳で苦手とされていた文章ジャンルの判断も、最近では特に明示的な指定なしでもほぼ的確になされるようになってきている。例えば、学術論文[1]からの抜粋である

We have synthesized and characterized two fluorescent rhodamine-containing PTP probes that are highly sensitive for direct in-gel visualization of PTP activity.

をデフォルトで機械翻訳にかけると (2023 年 4 月 8 日時点)、

Google 翻訳[2]: 我々は、PTP 活性の直接ゲル内可視化に高感度な 2 つの蛍光ローダミン含有 PTP プローブを合成し、特性評価しました。

DeepL[3]: 我々は、PTP 活性をゲル内で直接可視化できる高感度な 2 種類の蛍光ローダミン含有 PTP プローブを合成し、その特性を明らかにした。

T-400 (株式会社ロゼッタ): PTP 活性の直接ゲル内可視化に高感度な 2 つの蛍光ローダミン含有 PTP プローブを合成し、特性化した。

のように、いずれの場合も科学英語のジャンルに属する文として正しく翻訳している。英文和訳のテストで求められる条件が、原文の文法構造と文脈中での単語の意味を日本語に置き換えること [4, 5]とすれば、現状の機械翻訳が出力する訳文はテストの答案として合格レベルにあると言える。

一方で、上記訳文を独立した日本語の文として眺めると、関係詞節を含む英文の構造に束縛されているため、文頭の「我々は」と文尾の「合成および特性評価した」が離れすぎて文意がとりにくいなど、ややぎこちない表現になっている。こうした翻訳調の和文を、日本語の *native speaker* にとって自然な表現、英文なしで初めから日本語で書いたときに現れるような文に変換することが今後の自動翻訳の課題であると考えられる。

2. 本稿の視座

英文和訳は目的と用途に応じて、①英文の文法構造と文中での単語の意味を日本語に置き換えることによって示すレベルと、②英文を参照しなくても英文の意味がわかる自然な日本語で表すレベルに区別することができる。上述のように、テスト問題の解答や英文テキストの解説訳では①が、原書の翻訳本では②が要求されていると言ってもよい。本稿ではさらに、③英文を読むときの *native speaker* の頭の働き (情報処理) と同じ順序で日本語で表すレベル、を考えてみたい。

人間の社会的なコミュニケーション手段である言語を情報伝達という立場で捉えたとき、重要な特徴はその線条性にある。つまり、話し言葉を聞いたり書き言葉を読むときに、*native speaker* の頭の中に入力された意味情報は、文頭から文末に向かって 1 次元的に処理されていくはずであり [4]、

¹生命科学部生物情報科学研究室 ²東京家政大学第二外国語教育研究部門

一時に収容できる情報量の範囲を超えてこの処理の流れに逆行することは、言語本来の理解という観点からすると不自然であると考えられる。例えば複文中の関係詞節を後ろから返って訳すような理解は、文字が介在する構造分析ができるからこそ可能なのであって、書き言葉と話し言葉で情報処理の操作手順に乖離があるのは望ましいことではない。特に英語は文中における各語の前後配置が緊密な言語であり[6]、英文和訳ではこの緊密性のルールに従って英文の語順どおり日本語に「訳し下げて」いくことが、思考の流れに沿った意味内容の本来の受け取り方という点からも、話し言葉のリアルタイム自動翻訳という点からも重要になってくる。この段階の和訳レベルが上記の③に相当する。

本稿の目的は、英文和訳に際して機械翻訳の現状である①のレベルを、②さらには③までアップグレードするための方法論を、具体的事例に即したルールベースで提示することである。冒頭で述べた通り、生成系 AI や LLM など現在の機械翻訳の潮流は、大量データの学習に基づく知識ベースの手法の開発と適用にあるが、学習対象のデータ自体に含まれるソースに訳文として適切な日本語表現が少なければ、その効果も十分に発揮されないと考えられる。また英文和訳の汎用的ルールに関してルール化・統一化の動きがあるが[7,8]、こうした諸方面のプロダクトを参照尊重しつつ、本稿では上述した筆者独自の視点からアカデミック翻訳の事例に即して提案したい。本稿で提唱する方法論に基づく英文和訳の具体的な実践例は、文献[9]の翻訳版に本文として掲載されている。

3. 文構造の変換

3.1 限定用法の関係詞節

文全体が新しい文中の語を修飾する場合に、前文の末尾動詞の連体修飾が可能な日本語と異なり、英語で被修飾語（先行詞）の意味を限定するには後ろから関係詞を使って修飾文（従属節）をつなぐ方法が取られる。このため英文の語順通りに和訳するには、主節の構造を先行詞を連体修飾する形に組み直したうえで、従属節の中に入れてつなげばよい。例えば冒頭の英文

We have synthesized and characterized two fluorescent rhodamine-containing PTP probes that are highly sensitive for direct in-gel visualization of PTP activity. ([9]原書 p.248)

は、主節全体を先行詞 probes に送り込むことにより、

我々が合成して特性評価した 2 種類の蛍光ローダミン含有 PTP プローブは、PTP 活性のゲル内直接可視化について高感度である。

のように、英文の語順通りに訳すことができる。原文の情報提示の順序を優先したため、和訳の際に主節（被修飾文）と従属節（修飾文）の立場が入れ替わっているが、that 以下の意味によって probes の範疇が限定されている点（「高感度な 2 種類のプローブを合成した」vs「合成した 2 種類のプローブは高感度」）は訳文でも変わらない。

一方、非限定用法の関係詞節は、補足追加的に訳し下げることができる。

This paper describes the development of activity-based chemical probes for global profiling the entire PTP family on the basis of phosphatase activity, which should yield new functional insights into pathways regulated by PTPs and contribute to the discovery of PTPs as novel therapeutic targets. ([9]原書 p.246)

当論文で述べる活性に基づく化学プローブの開発は、ホスファターゼ活性に基づいて PTP ファ

ミリー全体を包括的にプロファイリングすることを目的とするが、PTPによって調節される経路への新たな機能的洞察をもたらし、新規治療標的としてのPTPの発見に寄与するはずである。

3.2 that で導かれる名詞節

S+V+that...の構造をした文において、主節のS+Vに比してthat...が長い場合は、S+Vを先にくくり出したのち、that...以下に訳し下げることができる(3.1は主節の後方への送り込み。3.2は前方へのくくり出し)。

In general it should be underlined that detailed investigations of the role of j carbides and pearlite particles on the deformation mechanisms, as well as the precipitation type and the structural nature of the j carbides, require higher resolution techniques, such as TEM or APT. ([9]原書 p.210)

一般に強調すべきは、j炭化物およびパーライト粒子の変形機序における役割の詳細な研究には、j炭化物の析出型や構造的性質と同様、TEMやAPTなどのさらに高分解能の技術が必要とされることである。

3.1と同様、語順の保存を優先するために、和訳に際して原文の構造を変換しているが、叙述の中心がthat以降にあることを踏まえると、むしろ原文の情報提示に忠実とも言える。

The simulation results showed that it was possible to reduce ATT exactly to the level calculated for the intact knee merely by increasing the magnitude of hamstrings force. ([9]原書 p160)

シミュレーション結果から示された通り、ハムストリングの力の大きさを増加させるだけで、ATTを未損傷膝の計算レベルまで正確に低下させることが可能であった。

同格節を導くthatの場合も、同様の方法で訳し下げることができる。

Here, we show a concrete example that natural stimuli that serve as proxies for memories can be bound to an underlying oscillation in a population of neurons. ([9]原書 p217)

ここで我々が示す具体例では、記憶の代理として働く自然の刺激を、ニューロンの集団で基本となる振動に結びつけることができる。

3.3 態の変換

日本語では各語の前後配置が英語ほど緊密ではなく、また態は「(ら)れる」の有無によって識別されるため、上例以外にも和訳に際して比較的柔軟に態を入れ替えることが可能である。

Our data suggest that the decrease should have occurred during the first year. ([9]原書 p.236)

我々のデータから示唆されるのは、1年目の間に減少が起こったはずだということである。

Results for individual groups are displayed in Supplementary Figure 5. ([9]原書 p.170)

各グループの結果を補足資料の図5に示す。

4. 文体の変換

和文英訳の際に、こなれた日本語の原文を英語らしい表現に翻案(文体を変換)することが推奨される[10]。②や③レベルの英文和訳ではこの逆に、元の英語の文体に引きずられない自然な日本語で表現することが求められる。

4-1. 疑問句節から名詞表現への変換

和文英訳で奨められる手順とは逆に、例えば *what the sentence is saying* は「文の内容」、*what it is doing* は「その役割」、*how to write* は「書き方」、*what's wrong with using the working title* は「仮タイトルを用いることの間違い」と訳すことができる。

4-2. 比較級から述語への変換

「より…な」は、「の方が…」と訳した方が日本語として自然な表現になる。

It's easier for the reader to track this if the writer uses the same verb, rather than switch to a so-called synonym such as find or detect. ([9]原書 p.208)

find や *detect* といったいわゆる類義語に置き換えるよりも、同じ動詞を使う方が、読者には把握しやすい。

Many more people will read the Abstract - or briefly look at it - than the whole paper. ([9]原書 p.265)
論文全体よりも、Abstract を読んだり、あるいはざっと目を通したりする人の方がずっと多いものである。

Damage accumulation evolved slowly at lower strains. ([9]原書 p.260)

損傷蓄積の進展速度は、歪みが低い方が遅い。

less などの劣勢比較級も同様の方法で訳すことができる。

In the in vivo examples, the hybrid method produces fewer obvious peak-hopping errors, ([9]原書 p.272)

in vivo 例では、ハイブリッド法の方が明白なピークホッピング誤差が少なく、

Low-complexity domain interactions that control gene transcription is less ambiguous than

Low-complexity domain interactions in gene transcription. ([9]原書 p.308)

Low-complexity domain interactions that control gene transcription の方が、*Low-complexity domain interactions in gene transcription* よりも曖昧さは少ない。

4-3. 分詞（形容詞）＋名詞の変換

分詞を名詞化することによって、例えば、*increased rate of synthesis* は「合成速度の増加」、*calculated Young's modulus* は「ヤング率の計算値」、*effects of increased pH* は「pH 上昇の影響」と訳すことができる。

4-4. 主語から主題への変換

日本語は文の基本要素として主語よりも主題を重んじる特徴がある[11]ので、原文の冒頭語を主題として提示すると自然な文体になる。

Theorem 1 should not be used unless it is apparent that the conventional assumptions are invalid. ([9]原書 p.238)

定理 1 は、従来の仮定が無効であることが明らかでない限り、使用すべきではない。

訳文は 3.3 のルールを適用した結果と変わらないように見えるが、「定理 1 を」でなく「定理 1 は」と訳すことにより、文の冒頭で主題を提示することができる。

This state can be maintained as long as needed to execute surgical and many nonsurgical procedures.

([9]原書 p.18)

この状態は、外科的処置および多くの非外科的処置を実施するのに必要な限り維持することができる。

4-5. 主語から条件節への変換

推量や可能性などの意味の助動詞を含む文では、主語が仮定条件を表す場合がある。

Therefore, the use of hESCs to investigate early human embryo development may provide valuable insights into early developmental processes, including neural induction. ([9]原書 p.23)

したがって、hESC を用いてヒト胚の初期発生を調べれば、神経誘導を含む初期発生過程への貴重な洞察が得られるかもしれない。

Extensive use of supplementary materials may have a range of negative effects; ([9]原書 p.76)

補足資料を大量に利用すると様々なマイナスの効果をもたらす可能性がある。

5. 不定詞や完了時制の語感の訳出

5-1. 現在完了形

現在完了の本質は過去の出来事の現在への反映[12]なので、「てしまった」や「してきた」よりも、完了相[13]を表す「ている」が日本語表現としてぴったりの語感の場合がある。

Little attention has been paid to the election of an alternative biocompatible material. ([9]原書 p.50)

代替となる生体適合性材料を選択することは、これまでほとんど注目されていない。

Biomass-derived polylactide (PLA) has received much attention in recent years due to its biodegradable and biocompatible properties. ([9]原書 p.4)

バイオマス由来乳酸 (PLA) は、その生物分解的および生体適合的な特性のため、近年多くの注目を集めている。

5-2. 副詞的用法の不定詞

不定詞を副詞的に使用する際の心理は不足情報の補足[14]であり、意識のうえでそれぞれの意味が必ずしも明瞭には区別されない[15]と言われている。例えば、4-5 の第 1 例文のように訳しても「hESC を用いること」の目的が「初期発生を調べること」である点は変わらないが、意味が目的であることを明示したうえで原文の語順通りに訳し下げることができる。

Each conclusion section is preceded by the Abstract of that paper and, where relevant, the synopsis or Highlights, to help you see how these differ from the conclusion section, and to emphasise that the Abstract should not be used as the primary source material for the Conclusion. ([9]原書 p.246)

Conclusion に先立って、当該論文の Abstract と、関連のある Synopsis や Highlight を掲載してあるのは、Conclusion との違いが分かるようにして、Abstract を Conclusion の主要な原資料とすべきでないことを強調するためである。

参考文献

1. Kumar, Zhou, Liang, Yang, Wang, Zhang, *J. Proteome Res.* **5**, 1898-1905 (2006)
2. <https://translate.google.co.jp/>
3. <https://www.deepl.com/translator>
4. 伊藤和夫「英語長文読解教室」研究社出版 pp.255-282 (1983)
5. 高橋善昭「英文和訳講座」研究社 pp.ix-xv (1990)
6. 大西泰斗・Paul McVay「一億人の英文法」東進ブックス pp.464-465 (2011)
7. https://www.jta-net.or.jp/useful_tools7.html
8. 駒宮俊友「翻訳スキルハンドブック」アルク(2017)
9. Hilary Glasman-Deal, *Science Research Writing Second Edition* (2021) ; 小島・甲斐訳「改訂版 理系研究者のためのアカデミックライティング」, 東京図書 (2023)
10. 富岡龍明「英語らしい英文を書くためのスタイルブック」研究社 pp.111-131 (2006)
11. 益岡隆志・田窪行則「基礎日本語文法 (改訂版)」くろしお出版 pp.2-3 (1992)
12. Hilary Glasman-Deal, *Science Research Writing* (2010) ; 甲斐・小島訳「理系研究者のためのアカデミックライティング」, 東京図書 pp.7-10 (2011)
13. 前掲書[11] pp.112-116
14. 前掲書[6] pp.464-466
15. 江川泰一郎「英文法解説 (改訂三版)」金子書房 pp.320 (1991)