

人間の翻訳におけるコーパスの位置づけ

影浦 峽（東京大学大学院教育学研究科）

テキストは引用の織物である¹。

かつて頻繁に引用され人口に膾炙したこの言葉は、しかしながら、言語の集積体たるコーパスを前にし、「言語学」と「自然言語処理」を両脇に置いたとき、有限の語彙と規則にもとづく無限の構築物としての言語表現に対置されるべき、歴史的に語られ記録され相互に関連しつつ集積してきたモノとしての言語表現全体という領域をそのモノ性を強調しつつ指し示すものとしてとらえ返すことができる。

翻訳者は——そして私たちはオンライン上で活動しているボランティアのまさに翻訳者を支援するシステムを構築しようとしているのだが——、翻訳者としてある限り、いささかも言語や意味を扱いはしない。そうではなく、それが言語的にどのような単位として規定されるとしても、テキストを、そしてテキストのみを、あるいはテキストとしてあらゆるものを、扱うのだ。

そんなはずはない、翻訳者ほど辞書を——そして辞書はいわば歴史上唯一人間の言語活動に役立つ言語ツールだったはずだ——頻繁に参照する職業はほとんどほかにはないではないか。そんな反論が——その反論は空疎なイメージをもてあそぶのでもなければ確かに有効性をもつ場を与えられるのであるが——すぐさま聞こえてもこよう。それにもかかわらず、翻訳者としてある限り、という担保のもとでは、翻訳者がテキストのみに関係しているという宣言は、ほとんど分析的に言ってよいまでに、つまりほとんどトートロジカルに、真実なのである。

そのように宣言される翻訳者の翻訳者としての活動はどのようなものなのか、そしてそうした翻訳者の活動においてコーパスはどのように位置づけられるのか、言語学や自然言語処理のそとを見ればすでに拡散したかたちではあれかなりの議論がなされているこれらの問題をあらためて整理し、翻訳支援という極めて世俗的な枠組みに翻訳活動におけるコーパスを位置づけることが、本稿の目的である。

* * *

マルチン・ルターが聖書をゲルマンの俗語で翻訳し、それが標準ドイツ語をつくりあげたことはよく知られ

¹ ロラン・バルト, 1977. 『テキストの快楽』東京: みすず書房.

ている²。実際、そう古くない過去に成立しロマン主義的潮流の中で口語による生き生きとした精神の表出であるかのように見なされ現在あたかも自明の存在とされている「ドイツ語」「日本語」「フランス語」といった「個別言語」は、その起源を書き言葉および翻訳に負っている。

逆に、欧米（日）列強の植民地主義の脅威にさらされながら「近代の成立」を急いだ地域において、これが極めて意識的なプロセスとしてなされたことは、魯迅の「翻訳にかんする通信」の中に見られる次のような言葉からもうかがえる³。

翻訳には——原書の内容を中国の読者に紹介する以外に——さらに重要な役割があります。それはすなわち、わたしたちが中国の新しい現代語を創り出すのに役立つということです。

日本語においても、事情はかわらない。たとえば福沢諭吉は次のように述べていた⁴。

世の中に原書が読めて翻訳のできぬ人は、唯むづかしい漢文のやうな訳文ができぬと云ふまでのことで、原文の意味はよく分つて居ることだから、其意味を口で云ふ通りに書くことは誰にもできませう。して見ればこの後は世の中の原書よみは其まゝ翻訳者になられる、そこで世間に翻訳書はふえて、其書は読み易く、何ほどの便利かしれませぬ。翻訳書のかしいと云ふのは、漢文のやうな文章の中にはなしのことばがまじるからこそをかしかれ、これをまるではなしの文にすればすこしもをかしいわけはありますまい。

日本近代の言文一致運動が、それ自体新たな言語の創設だったことを考えるならば⁵、福沢のこの言葉からも、言語の創生を読みとることができよう。

² 森安達也, 2002. 『近代国家とキリスト教』東京: 平凡社ライブラリー.

³ 魯迅, 1985. 「翻訳にかんする通信」『魯迅全集 6』東京: 学研. オリジナルは1932年「翻訳を論ず」として『文学月報』に発表されたもの.

⁴ 福沢諭吉, 1874. 「明治七年六月七日集会の演説」小森陽一, 2000. 『日本語の近代』東京: 岩波. より再引用.

⁵ 柄谷行人, 1999. 「エクリチュールとナショナリズム」『ヒューモアとしての唯物論』東京: 講談社学術文庫.

これは、ベンヤミンがルードルフ・パンヴィッツの次のような一節を引きつつ論じた「翻訳者の使命」に対応する正真の翻訳行為である⁶。

[それらの翻訳は]ドイツ語をインド語化、ギリシア語化、英語化するかわりに、インド語、ギリシア語、英語をドイツ語化しようとするのである。外国語の作品の精神に対してよりも、自国語の用語法に対してはるかに多大な畏敬の念を抱いているのだ。……翻訳者の原則的な誤謬は、自国語を外国語によって激しく揺さぶるかわりに、自国語の偶然的状態をあくまで保持しようとするところにある。……翻訳者は自国語を外国語によって拡大し深めなければならない。

このように論ぜられる翻訳は、その本質において既存の言語に関わるものであるよりも、私たちが今となっては所与の言語とみなしているような言語そのものの創出に関わる（その意味で、既存の言語の自明性が成立した現在私たちが普通に「言語」と呼ぶものに関わるものではない）。これに関わって、フランス古典主義文学を「ポーランド語」に翻訳するにあたりポーランド語古典文体をさえ創り出したタデウシ・ボーイ・ジェレニスキーの存在を思い起こしておくのも無用ではなからう⁷。

とはいえ、パンヴィッツが批判した翻訳、「外国語の作品の精神に対してよりも自国語の用語法に対してはるかに多大な畏敬の念を抱いている」翻訳者たちこそが、通常の翻訳実務において大多数を占め、また現実的にその有用性が疑いようもないかたちで成立している現状を踏まえるならば、私たちは、言語の創生という近代において本質的でありながら現在ではいささか例外的にも見なされがちな「翻訳者の使命」を離れて、個別言語の自明性を前提とした上での翻訳行為の検討に入らねばなるまい。

* * *

翻訳者にとって、「翻訳」は、原文を理解し、その意味を対象言語で表す行為では、ない⁸。

原文を理解すること自体は、翻訳活動を行わなくても、およそ原文そのものを読む者ならば誰でも行っている行為であり、翻訳行為そのものの前提条件ではあっても、いささかも翻訳行為の一部を構成するものではない。その意味で、原文の意味を理解することは、読

⁶ヴァルター・ベンヤミン、1996。「翻訳者の使命」『ベンヤミン・コレクション2』東京：ちくま。オリジナルは1923年に発表、より再引用。

⁷辻由美、1993。『翻訳史のプロムナード』東京：みすず書房。

⁸門翻訳者・ボランティア翻訳者8名への著者による聞き取り。

み書きができることとほぼ同様に、翻訳行為を行うために必要な条件でありながら、それだけでは十分な条件でもなく、また、ほかの多くの行為にも必要であるという意味で翻訳に固有の条件でさえない。

同様に、意味を対象言語で表す行為も、それ自体としては、文字通りに近いかたちならば英文和訳（英日を想定した場合）であり、より自由度が高ければかみ砕いた解説と呼ばれるものになり、翻訳そのものを構成する本質的なプロセスではない。

要するに、翻訳は、翻訳そのものとして見るならば、およそ認知的なプロセスではまったくないのである。翻訳に必要とされ翻訳者が備えていなくてはならない認知的なメカニズムはむしろ存在するだろう。けれどもそれはほかの多くの行為にも同様に必要であると同時に、翻訳の翻訳たる側面とは本質的に無関係である。翻訳を認知的プロセスと考える限り、永遠に翻訳はできない。

積極的な特徴付けに移ろう。一般の翻訳とは、その本質において、歴史上実際に記録されたテキストの集積体の中に、新たなテキストを付加するという、あくまで社会的な出来事である。その意味では、徹頭徹尾、同様に社会的な出来事である小説の創作や論文の執筆に類する行為であり、ただその初期条件/入力と出力の配置が、いささか異なっているだけなのだ。

敷衍しよう。

翻訳行為の出発点には原言語のテキストが存在する。そのテキスト自体は、その言語において歴史的に記録され集積されてきたテキストの総体の中に新たな場所を占めるかたちで追加されたものであり、ロラン・バルトの言を待つまでもなく、使われている言語がそもそも存在していることからテキストとして意味を持つことに至るすべてのレベルで、テキストの連関の中で、そしてそこにおいてのみ存在するものである。

一方で対象言語におけるテキストの総体がこれまでその言語において歴史的に記録され集積されてきたものとしてある。それは、その言語がそもそも存在していることからその言語で書かれたテキストが意味を持つことに至るすべてのレベルにおいて準拠枠となっているものである。

今、原言語のテキストを対象言語に翻訳するという行為は、原言語においてそのテキストが原言語における歴史的に記録され集積されてきたテキストの総体に対して持つ位置と対応する位置を、原言語におけるテキストの総体と対象言語におけるテキストの総体とのあいだにある一致とズレを踏まえた上で、対象言語のテキストを創り出しつつ与えてやるという、徹頭徹尾社会的な行為なのである。

異なる歴史を有し異なるかたちを有する異なる言語

のテキストの集積の中に対応するかたちを認め対応する位置づけを対象言語の対応する文書を創り出しつつ与えてやること、これが翻訳としての翻訳の行為であり、理論的・微視的言語学で言われるところの言語や意味の対応を手がかりにすればそれなりに翻訳が進むことは、すでに翻訳によって諸言語が成立した近代という歴史的な偶然に支えられてのことではしかない。逆に言語レベルの移し替えによって翻訳が成立するような言語間には、本来の意味で翻訳は必要がないとも言えよう。

いささか抽象的に述べてきた以上のことは、通訳者や翻訳者のエッセイを見れば、多くの例を通じて論ぜられているところでもある。たとえば米原万里は、「他人のふんどしで相撲をとる」をロシア語に訳すときに「他人のパンツでレスリングをする」とするといかにズレてしまうかについて語っているが⁹、これは、社会構成の相違に応じた「ふんどし」という語を含むテキストの連関が、ロシア語における「パンツ」のそれと大きく違うことこそが翻訳上の問題であることを示唆する一例である。

さらにまた、これをもって翻訳に見られるいくつかの点を説明することができる。第一は、翻訳者は、対象言語の表現を求めために多大な時間を使うこと。これは、対象言語において集積されたテキストの総体に対して今まさに翻訳者がその位置を与えつつ創り出している翻訳テキストの、位置づけを示すアンカーを求める行為であるからこそ、多大な時間と労力を投下するプロセスになるのである。第二は、翻訳者の極めて多くは、対象言語のネイティブであること。これもまた、翻訳が、いずれの言語におけるテキストの集積上の位置づけにより配慮しなくてはならないかに対応している。

とはいえ、翻訳をこのように捉えるならば、しばしば文化の相違を盾に翻訳不可能性を主張する言説を導きかねないことにもなるので、それについて付言しておかねばなるまい。

翻訳不可能性を検証不能な神学的主張として提出するのでない限り、その主張が成立しないことを論証するのは容易である。神学的主張でないならば、一つのテキストに対して提出された潜在的に無限なあらゆる「正しい翻訳」の候補に対して、人間が、検証可能なかたちで、それらは「正しい翻訳」ではないということ判定できなくてはならない。無限の候補に対して検証可能なかたちで判定することの条件は、端的に、少なくともこの世界の誰か一人は「正しい翻訳」を知っているか知りうるということであるから、翻訳が不可能であるとは言えないことになる。かくして翻訳の不可

能性の条件はまた翻訳の可能性の条件でもある。

実際のところ、翻訳とはパフォーマンス的な行為であり、翻訳を通じて翻訳可能性の限界がその都度拡張されるような行為である。それをその最基層において見るならば、第2節で見たような「言語の創出」が観察されることになる。

翻訳を認知的作業と見なす立場の完全なる否定。少なくともこの点に、言語の創生を担った翻訳と個別言語の自明性を前提としながらもなお翻訳としてなされる普通の翻訳とのあいだに存する共通性を見て取ることができる。これを暫定的な了解事項として翻訳者が求めるコーパスの検討に移ろう。

* * *

翻訳は、言語学が言う意味での言語にではなく、テキストに関わる。言語という言葉の意味を拡張するならば、このことはまた、次のようにも言えるだろう。すなわち、翻訳者は、一般に、有限の語彙と有限の文法規則に基づく無限の文の生成ではなく、歴史的に語られ表現され記録され人目に触れてきたテキストとそこにおける言語表現の蓄積とをこそ、言語ととらえる。

翻訳者が想定するコーパスの位置づけもこれに対応して理解することができる。すなわち、自然言語処理におけるコーパスが、統計的学習のサンプルと見なされることを典型として、さらに用例ベース翻訳や用例ベース翻訳支援を含め、一般的な言語の性質を表現する「例」であるのに対し、翻訳者が翻訳時に想定し、求めているコーパスは、現在翻訳中の文書との関係で、歴史の中の出来事として実際に語られ記録された、参照すべき具体的な表現の集積なのである。いわゆる意味での翻訳者が機械翻訳だけでなく翻訳支援システムをさえあまり使わないことには¹⁰様々な要因があろうが、翻訳メモリー的なシステムを含めて、一定以上の広がりを見せないのは、提供しうるコーパスが翻訳者の要求仕様と質的にずれていることも一因であろう。

フーコーの言葉を借りるならば¹¹、翻訳者にとって、コーパスはドキュマンであると同時にモニュマンでもある。あるいは、モニュマンであることの度合いに応じて、コーパスへの参照要求の高さも異なってゆくことになる。たとえば「To be or not to be, that is the question」ならば、その意味をとるために参考になる何らかのコー

¹⁰ 専門翻訳者・ボランティア翻訳者 13 名へのアンケート。および Fulford, H. and Granell Zafra, J. 2004. "The uptake of online tools and web-based language resources by freelance translators: implications for translator training, professional development, and research," Second International Workshop on Language Resources for Translation Work, Research and Training, p. 37-44.
Fulford, H. 2001. "Translation tools: An exploratory study for their adoption by UK freelance translators," *Machine Translation*, 16(3), p. 219-232.

¹¹ ミシェル・フーコー, 1981. 『知の考古学』東京: 河出書房新社。

⁹ 米原万里, 1998. 『不実な美女か貞淑な醜女か』東京: 新潮文庫。

パスではなく、坪内逍遙や福田恆存、小田島雄志を参照できることが、翻訳者の要求なのである。

とはいえ、翻訳がテキストを生成しつつテキスト集積の中に位置づける行為であるとして、テキストは様々な言語単位の引用からなるのだから、コーパスへの参照は、一定範囲の言語単位の参照としてなされる。その参照がどこまで歴史的固有性を持ったテキストに対してなされなくてはならないかは、すべての翻訳が文学の翻訳に類するものなのではないのだから、様々であるし、また、言語表現の単位に応じて異なる。

大まかに類型化を試みるならば、次のようになろう。

まず、文学作品や人文系の哲学・思想関係のもの。これらが参照するコーパスは、固有のテキストと固有の人とに強く結びつく。たとえば、朱牟田夏雄のスターン、米川正夫のドフトエフスキー、原佑のカント、など。参照コーパスの単位は、基本的に固有のテキストである。

次に、より広義の社会関係のもの。そのトピックに関わる様々な表現が引用等も含めて要請される。参照コーパスの単位は、関連する具体的なテキストである。

そして、科学技術文献に代表されるものがある。この場合は、分野に関わる専門用語やコロケーション等、かなりの程度、小さな言語単位に分割されて参照がなされる。したがって、参照コーパスの単位はテキストではなく、用語や連語等の分断された言語単位になっても大きな不都合がない。

第一の類型をカバーするコーパスは、古今東西のあらゆるテキストと翻訳されたテキストとを、その固有性とともに探せるような、普遍的図書館となる。第二の類型においては、それよりも断片化されてもかまわないが、トピックごとに十分な量の標準的なテキストをカバーするコーパスが求められるだろう。ここまでは、コーパスはあくまでもテキストとして位置づけられるものであり、後に言語の要素に断片化されるべき言語データのサンプルではない。そして、計算機処理で一般に想定されてきたコーパスがカバーできるのは、最後のケースのみであり、この最後の類型は、翻訳の概念自体から見れば、いわば例外なのである。

付言するならば、一見テキストのレベルを回避して言語単位をもとに構成された辞書が、それにもかかわらず今のところ唯一翻訳者が活用するツールであるのは、それ以外に使えるツールがないというだけでなく、「活字になる用例の何倍かの、日の目を見ない用例があってこそすぐれた辞書ができる」¹²という言葉に端的に示されているように、良くできた辞書が、テキストに結びついた用例のレベルで様々な消化と取捨選択を行っているからでもある。機械処理用に「概念」にまで分解した辞書に言語実践者から何一つ使い道がないのも、

¹²松井栄一, 2005. 『国語辞典はこうして作る』東京: 新宿書房.

そして人間向けの辞書が権威主義と結びつく本質的な性格を有していることも¹³、また同じ理由による。

* * *

Martin Kay は、英仏翻訳において“The man looked at the girl with the telescope”とその翻訳例を、“He looked at the girl with penetrating eyes”に用いることはできないと正しく論じている¹⁴。本稿で展開してきた議論に照らして見るならば、一見、伝統的な言語学や言語処理の範囲に収まるかに見える Kay の議論は、最終的には、適格な文を構成する単語と文法規則、それに対応する意味という言語のモデルから、歴史的に発話され記録されたテキストという言説の歴史的集積に移行する過渡段階の一例と捉えることができる。

そして、本稿で示してきたように、「言語」の範囲で操作可能かつ有用なコーパスの応用範囲は、翻訳の場においては、科学技術翻訳のように限られたもの——その範囲では、テキストのレベルを二次的な位置づけにして、専門用語等の単位を一次的な収集対象とするツールは極めて重要であろう——である。それ以外の膨大な翻訳においては、翻訳者が参照すべきコーパスは、たまたま翻訳対象テキストに出てくる表現と類似の表現が使われている任意のコーパスではなく、翻訳者が翻訳しつつあるテキストがまさにその中に密に位置づけられるような、これまでのテキストの、ドキュマンであると同時にモノユマンの、集積でなくてはならない。

このことは、翻訳者にとってのコーパスは、一部の確定的な言語単位の情報源として以上のものを求めるならば、翻訳対象テキストに応じて動的に決まるものであることを意味する。それを事前に準備できるようにすることを考えるならば、コーパスは、そこで何か見つければよいという基準ではなく、そこで見つからなければあきらめてよいという、理念としてこれまで図書館が担ってきた基準を満たすべく構成されなくてはならないことになろう。歴史的に集積されてきたテキストの総体に新たなテキストを位置づける行為が究極的に参照すべきコーパスとして歴史的に集積されてきた翻訳テキストの総体が求められること。当然と言えばこれ以上当然のことはない。

* * *

本研究の一部は、日本学術振興会科学研究費補助金基盤(A)「翻訳者を支援するオンライン多言語レファレンス・ツールの構築」(課題番号 17200018)の支援を得て行われた。

¹³Wells, R. A. 1973. *Dictionaries and the Authoritarian Tradition*. The Hague: Mouton.

¹⁴Kay, M. 1997. “The proper place of men and machines in language translation,” *Machine Translation*, 12(1), p. 3–23.