

言語の壁を超えるAI翻訳

AI Translation Breaks Language Barrier



大阪大学 名誉教授

成田 一

英日対照構造論・機械翻訳・言語教育/外国語習得論専攻。大阪大学功績賞受賞。著書『パソコン翻訳の世界』（講談社）、『日本人に相応しい英語教育』（松柏社）ほか、編著『こうすれば使える機械翻訳』（バベルプレス）、『英語リフレッシュ講座』（大阪大学出版会）ほか、共著『英語教育徹底リフレッシュ』（開拓社）ほか。『翻訳技術の言語的な基盤』（51ヶ月連載：バベルプレス）、『総合的な翻訳による英語教育』（50ヶ月連載：バベルプレス）、『Translators』（日本翻訳協会）、『Cat』（アルク）の連載、『SPA』（扶桑社）、『英語教育』（大修館書店）、『新英語教育』（三友社出版）などの雑誌記事、英文テキスト、新聞コラム、論文など多数。英語教育総合学会会長。情報通信技術研究交流会監事。

1 翻訳方式の変遷

日本における機械翻訳研究の切っ掛けとなったのが、（京大総長となる）京都大学の長尾真教授と辻井潤一助教授を中心に進められた科学研究費特定研究「言語情報処理の高度化」（1986-1988）¹である。これには主要な大学や国立国語研究所の研究者などが参加したが、研究室の大学院生は電気通信系企業の研究員となり、企業における翻訳システム開発の任にあたった。筆者も「機械翻訳前処理自動化のための基礎研究」研究班長として一翼を担い、その後、大手電気通信系メーカー各社が研究部長を送り込む電子工業振興協会の機械翻訳専門委員会の学術顧問となった²。

日本では80年代中旬から①ルールベース（1000以上の文法規則を作成）の翻訳システムの開発が始まった。90年代初頭までは業務用コンピュータはワークステーション中心で100万円前後し、漢字プリンターも

数十万円していた³。大手電気通信系企業の翻訳ソフトも日英・英日それぞれ50万円した。バブルの崩壊した90年代半ばまでに⁴、パソコンが20～30万円、漢字プリンターも数万円に、91年頃から発売されたベンチャー系の翻訳ソフトは92年に20万円弱に、93年には10万円弱に値下がりしたが、個人ユーザーを掴むことはできなかった。しかし、94年末に翻訳品質の比較的高かったロゴヴィスタの『コリヤ英和』⁵が9800円で発売されたのを機に個人の間にも機械翻訳ブームが到来し、全ての翻訳ソフト併せても年間400～500本だったのが、『コリヤ英和』だけで月間4000～5000本売り上げた。このため、ほかの翻訳ソフトも数万円までに値下げした⁶。低価格化に拍車をかけたのがWindowsの登場によるインターネットブームで95年末からインターネット対応版の翻訳ソフトが各社から2万円前後で売り出され、店頭販売で50万本、プリインストール分がそれ以上出荷されたと推定される。

1 機械処理のための言語構造の言語間対照研究、言語解析のシステムを作成し、機械翻訳における訳語選択の方法を提案した。また日本語と英語その他で、訳語対応をさせたときの意味のずれや、日本語の指示詞の使い方、談話理解への適用を行ない、多くの成果を得た。隣接する文相互間の文脈の関係性、さらに文章理解の計算機モデルの作成、人間と機械の対話システムの作成などを行なった。

2 NEC、富士通、東芝、シャープ、リコー、三洋電機、沖電気、ATR、日本IBMなどが開発中の翻訳・通訳システムの評価、改善点の指摘なども行い、20回余り研究助成金を供与された。

3 機械翻訳に不可欠な日本語処理技術として漢字への変換があった。これ以前はコンピュータで扱う文字はアルファベットか仮名のみ。1978年に東芝が発表し翌年出荷した世界初の日本語ワードプロセッサ『JW-10』である。単漢字変換機ではあったが、価格は630万円、重さは220kg片袖机ほどの大きさの筐体に、キーボード、ブラウン管、10MBのハードディスク、8インチフロッピーディスクドライブ、プリンターが収められていた。

4 バブル崩壊以降、電気通信系企業各社は機械翻訳開発部門の縮小廃止に向かった。

5 間もなく『コリヤ英和 一発翻訳』に改称。

6 野村証券も翻訳ソフト市場の今後の展開に関心を持ち、筆者に将来的可能性について見解を求めた。

90年代中旬から②統計翻訳（大量の対訳データから抽出した統計パターンを基に行う翻訳）が情報通信研究機構（NICT）や国際電気通信基礎技術研究所（ATR）などにおいて研究の緒についた。しかし、近年、③ニューラル翻訳⁷（膨大な学習データを用いた人工知能の深層学習⁸の技術による翻訳＝AI翻訳）が成果を出し、『Google翻訳』⁹、『DeepL翻訳』¹⁰などが従来のルールベースの翻訳精度を遙かに超える異次元の翻訳精度で、機械翻訳は新時代を迎えた。AI翻訳が言語の壁を打ち破り、主要な言語間の瞬時の翻訳を実現し、音声認識合成と連結すればほぼリアルタイムの同時通訳が行える状況だ。

2 翻訳ソフトの文法的評価

翻訳ソフトがどの程度正しく翻訳できるかを表すのに「翻訳率」という用語を使っていたが、1980年代中層までは明確な規定のないまま漠然とした感覚で使われてきた。そこで1980年代後半に筆者は学会で「文法項目ごとに対応できているか」を見る「文法的評価法」を提案した。文法的評価法では、英日翻訳の場合、たとえば、準動詞構文（不定詞句&分詞句）、関係節など主要な文法構造や用法が適切に分析できているか「対応する

- 7 ニューラル翻訳（Neural Machine Translation）は深層学習技術を用いて翻訳を行う最新の機械翻訳技法。最近ではAI翻訳と呼ぶことが多い。文の意味や文法関係を網羅的なベクトルで捉える。学習データから自動的に文法や表現を獲得でき、翻訳の精度が高く自然で洗練された訳文になる。ただし、モデルの内部がブラックボックス化されているため、エラーの原因の特定や修正が難しい場合もある。
- 8 深層学習（Deep Learning）とは、多層のニューラルネットワークから構成され、対象の全体像から細部まで各々の粒度の概念を階層構造として関連させて学習する手法のこと。十分なデータ量があれば、機械が自動的にデータから特長を抽出してくれる、脳神経回路をモデルとしたアルゴリズム。
- 9 Google社が2006年にリリースした機械翻訳サービスを2016年に深層学習を用いたシステムに刷新したGoogle翻訳は、中国語と日本語の精度が大きく向上。130を超える言語に対応。
- 10 ドイツのDeepL社が2017年にリリースした機械翻訳サービス。当初は欧州諸語にのみ対応したが、2020年に日本語に対応。英語から他言語への翻訳は人間の翻訳者に匹敵する自然な翻訳が特徴だ。

日本語の表現形式¹¹」に適切に翻訳できているかなどを、文法項目や表現に分けて用例を作り、翻訳結果を検証するというものだ。この検証により、開発中の翻訳システムの文法規則の修正や追加が進み、翻訳システムの改善に使われた。

翻訳能力の評価は、例文の翻訳ごとに文法解析面と語彙面について見ることになるが、翻訳システムの文法解析面を評価するというのは、文の構造を内部の節や句構造に分けて解析し、修飾関係の成否を見ていく作業になるが、これには言語分析の専門知識が必要だ。

それぞれの専門分野に特有の語彙や定型表現（特定の言い回し）などもあるので、多角的に見る必要がある。分野ごとに専門用語や定型表現を組み込んだ翻訳システムであれば、語彙面でも文法解析面でも高い翻訳率になる。したがって、分野特有の文書ではなく、一般的な文書で翻訳率を比較するのだが、医薬歯、理化学、法学など、分野ごとの文書に関して、どの程度の翻訳率になるかを比較するのがより正確な評価になる。

また、「翻訳方向」によっても翻訳率に差が出る。日英翻訳では、原文の日本語に主語や目的語などが欠けることが多いが、英訳するには欠落成分を補わないといけな。ところが、それには文脈情報も必要になるが、文脈情報を活用する人工知能機能はルールベースの翻訳システムには備わっていない。欠落成分を補う際に誤りを犯すことが多いので、英日翻訳よりも10～20%ほど翻訳率は下がる。

翻訳対象となる文書の種類、難易度によっても違うが、従来のルールベースの翻訳システムの精度はロゴヴィスタの『コリヤ英和 一発翻訳』など、比較的高品質のソフトでも英日翻訳で70～85%程度だ。ただし、日本語は主語や目的語など文の主要成分が省略されることが普通だが、欧米語では省略しないで「代名詞で

11 たとえば、「I read three books.」が「3つの本」ではなく「三冊の本」を読んだ」のように、英語の数詞を[数詞+類別詞]に翻訳しているか、更に「の」を介しない「本三冊」を読んだ」としているか、「本」を「三冊」読んだ」ないし「三冊」「本」を読んだ」など、「類別詞の名詞句外への分離操作」を実行し、自然な日本語表現への翻訳ができているかを評価する。

残す」ので¹²、日本語で省略された成分を代名詞で復元しなければならないが、その適切な復元が困難¹³で誤訳になることが多く、日英翻訳では55～70%に留まる。

欧州諸語は同じ言語グループだと方言差程度の違いしか無いのだが、元来ゲルマン語族だった英語は、歴史的にフランス語を話すノルマン王朝に長年支配されて、語彙だけでなく構文的にもフランス語の影響を受けた混合語になったと言っても過言ではない¹⁴。このため、英仏／仏英翻訳で90～95%、英独／独英翻訳¹⁵で85～90%と実用レベルにあった。

3 日韓翻訳

日本語と韓国語は文法の95%が共通の双子関係にある。このため、深い構造分析は必要なく、どちらも主語、目的語などを欠いても文が成立するため、欠落成分を補わないで翻訳が成立する。また漢字を背景とする高級語も共通だ。(ただし、読み方は、漢字の借用された時代の音韻を担っているの、日韓語に違いがある。)このため、日韓翻訳は90～95%の翻訳率になる。しかし、韓国語はハングルで表記されるため、韓日翻訳において同音異義語が誤訳される可能性が少なくない。これに対して、日本語では同音異義語が異なる漢字で表記されるため、日韓翻訳では誤訳は起こらない。概して、日韓翻訳の方が韓日翻訳よりも翻訳精度が5～10%ほど高い。

12 スペイン語では動詞の活用で主語が1人称、2人称、3人称か単数か複数かが分かる。

13 文脈から欠如成分を求める機能が備わっていない。

14 1066年のノルマン征服により、英国では公用語がフランス語になった時代が300年ほど続いた。この間に日常語以外の語彙はフランス語が取り入れられ、分詞構文など文法面でもフランス語化が進んだ。

15 英語は本来ドイツ語と同じゲルマン諸語に属し、名詞句内の冠詞、形容詞、名詞が活用していたが、歴史的に活用が失われ、かつ、ノルマン征服後のフランス語化によりドイツ語との遊離が加速した。

4 AI翻訳の実力

AI翻訳としては『Google翻訳』、『DeepL翻訳』のほか、ロゴヴィスタの『翻訳 powered by chatGPT¹⁶』(以降、本文では『翻訳 chatGPT』と略称表記)を取り上げ、その訳文を載せ評価する。なお、参考までに従来のルールベースの翻訳(ロゴヴィスタの『コリア英和一発翻訳』(以降、本文では『一発翻訳』と略称表記))の訳文も記載した。

AIの深層学習によるMTは進化する

読売新聞のコラム¹⁷によると、『Google翻訳』に下記の日本語文を入力すると、①の誤った翻訳が翌日には筋の通った②に変わっていたという。

「キリマンジャロは雪に覆われた1万9710フィートの山で、アフリカで最も高い山と言われている」(ヘミングウェイ短編より)

- ① Kilimanjaro is 19,710 feet of the mountain covered with snow, and it is said that the highest mountain in Africa. (『Google翻訳』1回目)
- ② Kilimanjaro is a mountain of 19,710 feet covered with snow, and it is said to be the highest mountain in Africa. (『Google翻訳』2回目)
- ③ Kilimanjaro is a snow-capped 19,710-foot mountain and is said to be the highest mountain in Africa. (『Google翻訳』3回目)
- ④ Kilimanjaro is a snow-capped 19,710-foot mountain and is said to be the highest mountain in Africa. (『DeepL翻訳』)
- ⑤ Mount Kilimanjaro is a 19,710 feet high mountain covered in snow, and is said to be the tallest mountain in Africa. (『翻訳 chatGPT』)

16 オープンAI社が2022年11月に開発した『chatGPT』の根幹が「翻訳を含む言語処理」だ。この翻訳エンジンを『chatGPT翻訳』と称して良い。ロゴヴィスタの『翻訳 powered by chatGPT』は実質的に翻訳エンジン『chatGPT翻訳』と同じものだ。

17 4月頃のコラムだったかと思うが日付は不詳。

英訳①は元の日本語文の翻訳としては成立しない。誤訳が①には2箇所ある。日本語配列のまま英語化したかのような「非文法的で意味をなさない 19,710 feet of the mountain」それに（節が続くことを表す文法標識である）[that]の後に名詞句しか続かず文法的に許容されない that the highest mountain in Africa だ。一方、②は一応正しい英訳だが、it is said to be...のitは先行文の Kilimanjaro だと考えられる。しかし、このitは省略可能で、is said to be...の方が先行文への繋がりが緊密に感じられ、文体的には優れている。

念のため、本日（2023/05/31）『Google 翻訳』で翻訳したところ、③の英訳が出力された。英訳③では英訳②の covered with snow と of 19,710 feet が前置され、snow-capped と 19,710-foot に改められているが、これは洗練された翻訳である。AIの深層学習による翻訳プログラムは新規の言語データを日々消化して更に高度化するのが実感できた。（因みに、④は『DeepL 翻訳』の訳だが③と全く同じだ。⑤は『翻訳 chatGPT』の訳だが、述部に is...mountain とある以上 Mount は不要ではあるが、Kilimanjaro が山であることを認識した訳となっている点で評価できる。ただし、19,710 feet high は 19,710-foot high だろう。また、山が tallest というのは頂けない。前置詞も間違っており covered in snow という訳文になっている。

大学入試問題を AI 翻訳で解く

AI 深層学習による機械翻訳の英語力を日本の大学受験生の英語力と比較する意味でも、大阪公立大学英語試験（2023）の機械訳を予備校の解答例と比べて解説したい。

英日翻訳

The more educated you are, and the higher your grades, the more you participate in after-school classes and tutorials, the more likely you are to be myopic. (大阪公立大学 2023 英語試験より)

『Google 翻訳』

「学歴が高く、成績が高いほど、放課後のクラスやチュートリアルに参加するほど、近視眼的になりやすくなります。」

『DeepL 翻訳』

「学歴が高く、成績が良く、放課後の授業や個別指導に参加している人ほど、近視になる可能性が高いのです。」

『翻訳 chatGPT』

「学歴が高く、成績が良いほど、放課後の授業や補習に参加するほど、近視になる可能性が高くなります。」¹⁸

予備校解答例

「教育を受ければ受けるほど、成績が上がれば上がるほど、課外授業と個別指導に出席すればするほど、近視になる可能性は高くなります。」

AI 翻訳については、myopic は「近視眼的」ではなく「近視」である。そこは『Google 翻訳』が誤訳になるが、ほかの箇所はいずれも基本的に正解だ。「the+ 比較級…」という相関接続詞の率いる節が3つ連続するが、「ほど」を繰り返さず最後の節だけで訳出する流暢さにおいて『DeepL 翻訳』が一枚上だ。夕陽丘予備校解答例も超える。

『一発翻訳』

「あなたがより見込みが高く近視であるはずであると、それだけ、あなたがいっそう教養を身につけている、そしてあなたの学年がより高いと、それだけあなたももっと学童保育クラスとチュートリアルに参加します。」

下線箇所は原文では主文の the more likely you are to be myopic を訳したもののだが、the more likely you are が「あなたがより見込みが高く」と訳され、to be myopic が「近視であるはず」と訳されるなど、be likely to be という構文の訳も間違っている。[the+ 比較級… (従文), the+ 比較級… (主文)] という相関接続詞構文も主文が従文に訳されるなど、文の構造が誤って解析されている。ただし、この構文で従文が3つもあるのは珍しく、意地の悪い試験問題だ。

18 この訳には「学歴や成績が高いほど、放課後の授業や補習に参加すると近視になる可能性が高くなる。」という要約が添えてある。人工知能として載せられた「要約する機能」があるのだが、原文と意味が違っている。訳文は正しいが要約は間違っている。この程度の劣悪なものだと必要の無い機能だ。



日英翻訳

① 「確かな感触を得た。これは自分さえ頑張れば叶わないことじゃない。何かが始まりそうな予感に胸がはずんだ。」(大阪公立大学 2023 英語試験より)

『Google 翻訳』

I got a certain feeling. This isn't something that can't be achieved if you do your best. My heart fluttered with a premonition that something was about to begin

『DeepL 翻訳』

I got a solid feeling. This is not something that can't be achieved if only I try hard enough. I was excited by the feeling that something was about to begin.

『翻訳 chatGPT』

I got a definite feeling. This is not something that can't be achieved as long as I work hard. My heart leaped with a premonition that something is about to begin.

予備校解答例

I felt I could somehow do it. I thought it would not be impossible if I did my best. I got excited to find something wonderful would begin.

AI 翻訳については、いずれも基本的に正解だ。「胸がはずんだ」が『Google 翻訳』では My heart fluttered、『翻訳 chatGPT』では My heart leaped のように直訳調なのに対し、『DeepL 翻訳』は I was excited と I を主語としているが、これは文体上の問題で優劣はない。ただ、『翻訳 chatGPT』の is about to は was about to のように過去形にするのが正しい。予備校解答例 I felt I could somehow do it. は「確かな感触を得た」の模範訳として受験生に示すのは如何なものか。AI 翻訳の方が素直な翻訳で受験生も自分の解答と較べやすい。

『一発翻訳』

"I acquired the clear feel. This is not suiting, if he does his best. The breast bounced to the premonition that something is likely to start." (feel は「感触」にならない。This is not suiting,

は「適する」なので「叶う」に対応しないし、「叶わないことじゃない」の訳にはならない。また、「胸がはずんだ」が文字通り「胸＝乳房が上下に揺れる」を表す The breast bounced と訳されているが、やはり AI 翻訳の My heart fluttered や I was excited、My heart leaped などの訳が望ましい。

② 「国際英語論の考え方にもとづけば、[英語はその普及された土地の言語文化の影響を受けて発展する] ことを肯定的にとらえます。」(大阪公立大学 2023 英語試験より)

『Google 翻訳』

Based on the idea of international English theory, it is positive [that English develops under the influence of the linguistic culture of the land in which it is disseminated].

『DeepL 翻訳』

Based on the conception of International English Theory, it is positive [that English develops under the influence of the linguistic culture of the land in which it is disseminated].

『翻訳 chatGPT』

Based on the perspective of international English discourse, [English is seen positively as developing under the influence of the language and culture of the land where it has spread].

予備校解答例

Thinking from the viewpoint of globalized English, we have to accept the idea [that any version of English [developed uniquely under the influence of linguistic culture of each region] should be respected as such].

「ことを肯定的にとらえます」は『Google 翻訳』や『DeepL 翻訳』の it is positive [that English より『翻訳 chatGPT』の English is seen positively as の方が滑らかな訳であるし、it has spread も it is disseminated ほど堅苦しくない。問題文の「…ことを肯定的にとらえます」自体あまり自然な表現ではないが、これを夕陽丘予備校では上記のように the idea

[that any version of English [developed...each region] should be respected as such]. と翻訳している。この英訳では「[それぞれの地域の言語文化の影響を受けて発展した英語はどれも（地域）英語として尊重されなければならない] という考え方を受け入れなければならない。」という意味になり、問題文とは多少違った意味合いになる。これはかなり解釈を入れた翻訳だ。『Google 翻訳』と『DeepL 翻訳』が「明確だ」という意味合いの it is positive に、『翻訳 chatGPT』が is seen positively as に素直に訳しているのとはかなり異なったものだ。予備校の教員が相談して仕上げたものだろうが、これが原文に対応する訳として適切かどうかは疑念がある。

『一発翻訳』

"If based on the view of an International English theory, I will catch in the affirmative what [English develops in response to the influence of the language culture of the diffused land].

（「こと」を develops の目的語として what と訳しているが、それだと what English develops 「英語発達させたもの」と言った意味になる。that 節で訳すのが正しい。しかし、ほかの部分に間違いはない。）

更に別の例文

大学入試問題の日本語文は高校までの中堅以上の学生の英語力と AI 翻訳を比べる意味で取り上げたのだが、不自然な表現も含まれていたため、更に別の日本語文も取り上げてみたい。例文①は 3 つの AI 翻訳と『一発翻訳』で見ると、例文②は『Google 翻訳』をやや超えると思われる『DeepL 翻訳』と『翻訳 chatGPT』で比べたい。なお、例文③では、AI 翻訳の「意味理解」に関して、興味深い翻訳箇所があったので、『DeepL 翻訳』の訳を取り上げる。

例文①

「理工系の研究者の多くは海外の論文を MT で翻訳して概要を把握し、重要箇所は研究グループで英文を読み込む。また英語で論文を書く場合も、MT で翻訳した上で英訳を仕上げていく。専門分野の用語や表現は

人間と違い MT だと一貫性があるのだ。」（筆者の論考より）

『Google 翻訳』

Many researchers in the field of science and engineering translate overseas papers with MT [to grasp the outline], and the research group reads the English text for important parts. Also, when writing a paper in English, we will translate it with MT and then complete the English translation. The terminology and expressions of specialized fields are consistent with MT, unlike humans.

『DeepL 翻訳』

Many researchers in the science and engineering fields translate foreign papers by MT [to get an overview], and read the important parts in English with their research group. When writing papers in English, they also use MT to translate and then finish the English translation. Unlike humans, MT allows for consistency in terminology and expressions in specialized fields.

『翻訳 chatGPT』

Many researchers in the science and engineering field use machine translation (MT) to translate overseas papers and [grasp the overview], and important parts are read in English by research groups. When writing papers in English, they also use MT to translate and then refine the English translation. Unlike humans, machine translation provides consistency in specialized terms and expressions.

AI 翻訳はいずれも中堅大学の平均的な学生以上の翻訳を出力している。『Google 翻訳』が前置詞の使い方に難があるが、『DeepL 翻訳』と『翻訳 chatGPT』はそれも問題ない。更に、『翻訳 chatGPT』は「英訳を仕上げていく」に refine を選ぶなど他より優れた翻訳となっている。



『一発翻訳』

Many of researchers of a science-and-engineering system **translate an overseas paper by MT**, they grasp an outline, and ① **an important part reads English by a research consortium**. Moreover, also when writing a paper in English, after translating by MT, I **finish** English translation. Unlike man, ② **if the term of a special field of study and expression are MT(s)**, they are consistent.

逆翻訳（『一発翻訳』に備わった機能¹⁹）

理工システムの研究者の多くがMTによる海外の論文を翻訳します、（彼・それ）らはアウトラインを把握します、そして①**重要な部分が研究グループによって英語を読みます**。さらに、同じく、英語でペーパーを書くとき、MTによって翻訳した後で、私は英語の翻訳を終えます。人間と異なり、②**もし専門分野の用語と表現がMTであるなら**、（彼・それ）らは一貫しています。

『一発翻訳』は下線の箇所の訳が全く原文の意味を伝えていない。特に、① **an important part reads English by a research consortium** は逆翻訳の下線①であり、原文にも対応しないし英文としても成立しない。② **if the term of a special field of study and expression are MT(s)** は逆翻訳の下線②であり、原文にも対応しない。

例文②

文部科学省は17日、[4月からの学校における**新型コロナウイルス対策では、マスクの着用を求めないことを基本とする**] **方針**を各教育委員会などに**通知した**。[**運動会など行事の縮小は不要とし**]、新学期からはコロナ前の学校生活に徐々に戻る**見通し**だ。通知では、[[マスクの着用を求めないことを**基本としつつ**]]、[基礎疾患など様々な理由で、①着用を希望したり、②着用できなかったりする] **児童生徒**もいることから、**学校や教職員が児童生徒に着脱を強いることがないよう**] **明記**。[③

19 『一発翻訳』の翻訳画面は、日英翻訳なら、[日本語][英語][日本語]となっており、それぞれの欄に[原文の日本語文][英訳][英訳の和訳]が表示。この[英訳の和訳]が逆翻訳である。

登下校時に混雑した電車やバスに乗る場合や、④**校外学習で医療機関などを訪ねる際は** **着用を推奨する**。

『DeepL 翻訳』

On **April 17**, the Ministry of Education, Culture, Sports, Science and Technology (MEXT) **notified** boards of education and other organizations of **its policy** of **[not requiring the wearing of masks]** as a basic measure against the new coronavirus at schools starting in April. It is **expected** [that school life will gradually return to the pre-coronavirus lifestyle from the new school term, with **no need** [to reduce events such as athletic meets]]. **The notice states** [that [while **the basic policy** is **not to require** students to wear masks], **schools and teachers should not force** students to put on and take off their masks, [as **some students** may [① wish to wear masks or ② be unable to do so for various reasons, such as basic medical conditions]. It is **recommended** [that students **wear the mask** ③ when riding crowded trains and buses to and from school, and ④ when visiting medical institutions **for field trips**.]

『翻訳 chatGPT』

On the 17th, the Ministry of Education, Culture, Sports, Science and Technology **notified** each board of education [that **the policy** for new coronavirus measures in schools from April onwards is **[to not require the wearing of masks]** as a basic principle]. It is **expected** [that **events such as sports festivals** **will not be** scaled down], and [that school life will gradually return to pre-coronavirus conditions from the new school term]. **The notification specifies** [that while **the basic principle** is **to not require** the wearing of masks, there may be **students** [who ① wish to wear masks or ② are unable to do so due to various reasons such as underlying medical conditions], and therefore **schools and staff**

should not force students to wear or remove masks. Wearing masks is recommended ③ when commuting on crowded trains or buses, or ④ when visiting medical institutions during extracurricular activities.

表現形式や文体に違いはあるものの、どちらも全体的には適切な翻訳になっているが、『翻訳 chatGPT』の to not require は非文法的で not to require でなければならない。また、『DeepL 翻訳』では On April 17 となっているが、「4月からの…」という内容からすると日程的に合わないので、前月の On March 17 が正しいだろう。『翻訳 chatGPT』のように「17日」という日付だけを訳して On the 17 としただけで良いのに、人工知能システムらしく「気を回して」間違ってしまったのだろうか。賢い！

意味を理解した翻訳ではない。

例文③

職場におけるダイバーシティ、特に [女性が働きやすい] 環境整備 については、私も様々な仕事を通じてお話を伺っていますが、御社はかなり早くから取り組まれていたようですね。 そうですね。 もともと風通しの良い会社で、性別関係なく活躍の機会がありましたが、以前から、積極的に女性の基幹職採用を行っており、同時にキャリアと育児・介護との両立支援制度や、事業所内保育所の設置なども整備してきました。

『DeepL 翻訳』

I have heard through my various jobs about diversity in the workplace, especially in creating an environment [where women can work comfortably], but it seems that your company was involved in this from a very early stage. That's right. We have always been an open company, with opportunities for women to play an active role regardless of gender, but we have been actively hiring women for key positions for some time, and at the same time, we have developed a support system for balancing career, childcare, and nursing care, as well as an on-site daycare center.

「そうですね。」に続く文は、話を聞かれている会社側の発言に切り替わって We を主語に立て We have always... と翻訳されているが、これは「そうですね」(That's right.) という「つなぎ」表現をきっかけに「聞き手と話し手が入れ替わる」ことを膨大な言語データから AI が学習しているためではないかと考えられる。こうした「話者を切り替える」能力は『DeepL 翻訳』だけではなく『Google 翻訳』や『翻訳 chatGPT』にも見られるが²⁰、現象的には「文脈を読み込んだ」翻訳になっていると言えないこともない。

ただし、これはあくまで対応する文脈を含む膨大な言語データを解析して翻訳できたということで、「原文の意味を理解して」翻訳したということではない。また、opportunities for women to play an active role regardless of gender では regardless of gender (「男や女といった性に関係なく」) となっているのだから、活躍は「社員」の活躍なので女性社員に限定しているわけではない。したがって、for women to play an active role とは矛盾しており for workers to play an active role が適切だろう。このケースでも AI に「意味理解」はできていないのだ。現段階での人間と AI の決定的な違いは「意味理解」にあると言えるだろう。

逆翻訳で誤誤を検出

「意味理解」できないことが人工知能翻訳の限界であったとしても、それが究極的な欠点かということ、そのようにネガティブに捉える必要はない。AI 翻訳が意味を理解できないことから誤誤になったとしても、それをチェックする方策はある。日英翻訳の場合、①英語力の極めて高い日本人（ないし日本語力の極めて高い英語ネイティブスピーカー）が原文と訳文を比べて誤誤を探す

20 『Google 翻訳』 … and it seems that your company has been working on it from quite an early stage. Let's see. Originally, we were an open company with opportunities to play an active role regardless of gender, 「翻訳 chatGPT」 … and it seems that your company has been working on this for quite some time. Yes, that's right. Our company has always had a good working atmosphere where both men and women have opportunities to excel.



ことができる。

私が提案したいのは、②訳文を元の言語に逆翻訳²¹して原文と比べることだ。日本人が英語を良く分からなくても、逆翻訳文を原文の日本語と対照して、意味の違った箇所があればそこが誤訳部と認識できる。この逆翻訳という方策により、いかなる言語への翻訳であっても誤訳部を検出できるのである。誤訳部は原文を編集して改めて翻訳すれば良い。しかし、こうした誤訳検出の作業をやるかどうかは別にしても、AI翻訳の高い精度と圧倒的な処理速度のメリットは、意味理解の欠落による希な誤訳を補って余りある。

文化の位相翻訳

読売新聞の『編集手帳』には下記のくだりがある。

日本文学研究者のドナルド・キーンさんが著した英訳版、太宰治「斜陽」には、よく知られた“超訳”がある。原文の「白足袋」を「白手袋」に置き換えた◆村医者が、敗戦後の没落貴族である婦人を往診する際、礼を尽くそうと白足袋を履いた。しかし、英語圏の読者には「白い靴下」と直訳すると礼節が伝わらない。西洋式の正装である白手袋に変えたのは、苦肉の策だったという◆「空気を読む」の英訳も難しいとされる。「read the room (場の空気を読む)」といった例文がいくつかはある。しかし、どこかしら同調圧力をまとう「KY」のニュアンスに、ぴったりの表現ではなさそうだ。言葉は国民性を色濃く反映するものであるらしい。

AI翻訳であっても、文化の違いを睨んで、原文の「白足袋」を「白手袋」に置き換えるといった機転は利かない²²。それが、データに依存した言語処理によって異次元の高度な翻訳精度を誇るが、「真の意味理解に基づく翻訳をしているわけではない」AI翻訳の限界だろう。翻訳者通訳者は機転を利かせられる点で出番はあるのだ。

しかし、AI翻訳にこのような限界があったとしても、これはその効能を貶めるものではない。このような文化

21「逆翻訳」は原文を翻訳した訳文を翻訳し直す作業だが、日英翻訳なら【日本語】【英語】【日本語】と三つの画面が現れる『一発翻訳』のように、翻訳ソフトに最初からこの機能が付いているものもある。

22ただし、「斜陽」の原文と“超訳”をAI翻訳データとして与えていれば、同じ訳を出力するだろう。

の違いを睨んだ機転の利いた高度な翻訳が必要な場面は多くはない。また、上記の場合は、注釈の形で「白足袋」を履いたことが礼節を尽くすことになることを解説すれば済むことでもある。AI翻訳が「機転の利いた翻訳」ができなくても、ここまで正確な言語処理ができていたのであれば、「人間が足りないところを補う」という形で活用することが適切なのではないだろうか。全て丸投げするのではなく、共存するのがAI翻訳活用のあるべき姿だろう。

同時通訳の実現

話者が話し始めて二、三秒遅れで訳す同時通訳になると、とにかく（言語対によって）「基本語順の違い」、そして「言語差」が、決定的に作業の難易度を左右する。それも日英通訳が英日通訳よりも遙かに難しいのである。

文の構造は構文情報を担う述語（動詞／形容詞）で決まる。英語は「主語—述語—目的語（SVO）」という配列なので、主語に続く述語が聞こえた時点ですぐに文の骨組みの情報が分かるので通訳が始められるが、日本語は「主語—目的語—述語（SOV）」といった配列なので、述語が最後になる。日本語は文末にある述語を聴くまでは文の骨組みの情報が分からないので訳せない。だが、現実には見切り発車で述語を聴く前に主語に続いて先行文脈から予測した述語に訳すこともある。

また、最後の述語に否定辞が付く場合もあるが、否定辞 not や n't は主語に続く助動詞に付くので、否定文でなかった場合、急遽訳文を調整しなければならない。それだけではない。日本文は修飾句や主語、目的語などが節により修飾されていることも多々ある（「僕が [[昨日新宿で買った] 本] を読み終えたら、君に貸してあげるよ」）。そうなると、後ろに続く要素の予測はほとんど不可能になる。正確に訳そうとすれば、一文遅れで翻訳するしかない。

だが、①一文全体を最初から最後まで聴いて、構文・意味解析し知識システムと照合して理解し、②（それを記憶に保持しつつ、）リアルタイムで訳文を生成する間に、③後続く文をリアルタイムで聴き取って、構文・意味解析し、知識システムと照合して理解し、それを記憶する。聴き取り中に通訳を始め、同時に聴き取りを継

続するのだ。こうした聴き取りと通訳を同時並行的に遂行する二重処理作業を繰り返し、継続しなければならないことになる。これは超人的な負荷のかかる²³至難の業²⁴だ。

AI 翻訳では、発話を音声認識し文字化された文ごとに訳すのだが、その処理が次の文の始まり数語で終わる。従って、話者が普通で話している間に、1 文ごとに通訳されることになる。つまり、発話を文ごとに瞬時に訳しそれを音声化する作業を連続的に行うのだが、これは時間的には同時通訳に匹敵する。これには音声認識⇒語連鎖（文）への変換⇒構造分析⇒翻訳文の生成⇒音声化を瞬時に連続的に行うシステムを組むことが必要だが、技術的には実現可能だ。

AI 翻訳は数多くの言語間の翻訳が可能で音声システムと組めば多言語通訳システムになる。グローバル時代においては、AI 通訳システムさえあれば、お互い自分の母語（地域語）で話してコミュニケーションできる。いわば、グローバルな時代なのである。「グローバル時代には日本人も英語の習得が必要だ」というのは過去になりつつある。

参考文献

- 『こうすれば使える機械翻訳』（成田一編著）バベルプレス 1994.4
- 『パソコン翻訳の世界』（成田一）講談社 1997.10
- 「機械翻訳はどこまで人間に迫れるか」（成田一編著）（『AI JAPAN』）白夜書房 2000.1 「特別講座：機械翻訳とはじめ」（成田一）（『翻訳辞典 2002』）アルク 2001.11
- 「特許文の現代化と機械翻訳」（成田一）（『Japio 創立 20 周年記念誌』（財）日本特許情報機構 2005.10
- 「特許文の多言語機械翻訳」（成田一）（『Japio YEAR BOOK 2006』（財）日本特許情報機構 2006.11
- 「機械翻訳の歴史と今後の展望」（成田一）（『Japio YEAR BOOK 2007』（財）日本特許情報機構 2007.11

23 英仏通訳、日韓通訳は、文成分の配列、文法が極めて近いので、なぞるように通訳でき、負担は少ない。

24 国際会議の場だと日英間翻訳の通訳は複数人が 15 分ほどで交代するが、通訳しない時間はメモなど用意して随時通訳者がそれを見て通訳に役立てる。

「翻訳ソフトあれこれ」（成田一）（『私のおすすめパソコンソフト』）岩波書店 2002.8 『英語リフレッシュ講座』（成田一編著）大阪大学出版会 2008.4

「日本語編集の視座」（成田一）（『Japio YEAR BOOK 2008』（財）日本特許情報機構 2008.11

『日本人に相応しい英語教育』（成田一）松柏社 2013.8

「翻訳技術の言語的な基盤（長期連載）」（成田一）（『The Professional Translator』）バベルプレス 2013.6～2017.6

「総合的な翻訳による英語教育（長期連載）」（成田一）（『The Professional Translator』）バベルプレス 2018.3～2023.5

「自動翻訳の高度化と英語教育」（成田一）（『Japio YEAR BOOK 2019』（財）日本特許情報機構 2019.11

「人間翻訳と自動翻訳」（成田一）（『Japio YEAR BOOK 2020』（財）日本特許情報機構 2020.11

「機械翻訳の目標と設計デザインー言語的基盤を踏まえてー」（成田一）（『Japio YEAR BOOK 2021』（財）日本特許情報機構 2021.11

「翻訳と通訳における機械翻訳の活用」（成田一）（『Japio YEAR BOOK 2022』（財）日本特許情報機構 2022.11

