翻訳のアナログ思考をデジタル技術が支える

安藤（あんどう） 進（すすむ）
翻訳家

あんどう すすむ●機械翻訳の研究開発に参加したあと、大学で英語の非常勤講師を務めた。著書に『ちょっと検索！ 翻訳に役立つGoogle表現検索』がある

1980年代半ばは翻訳ソフトの創成期だった

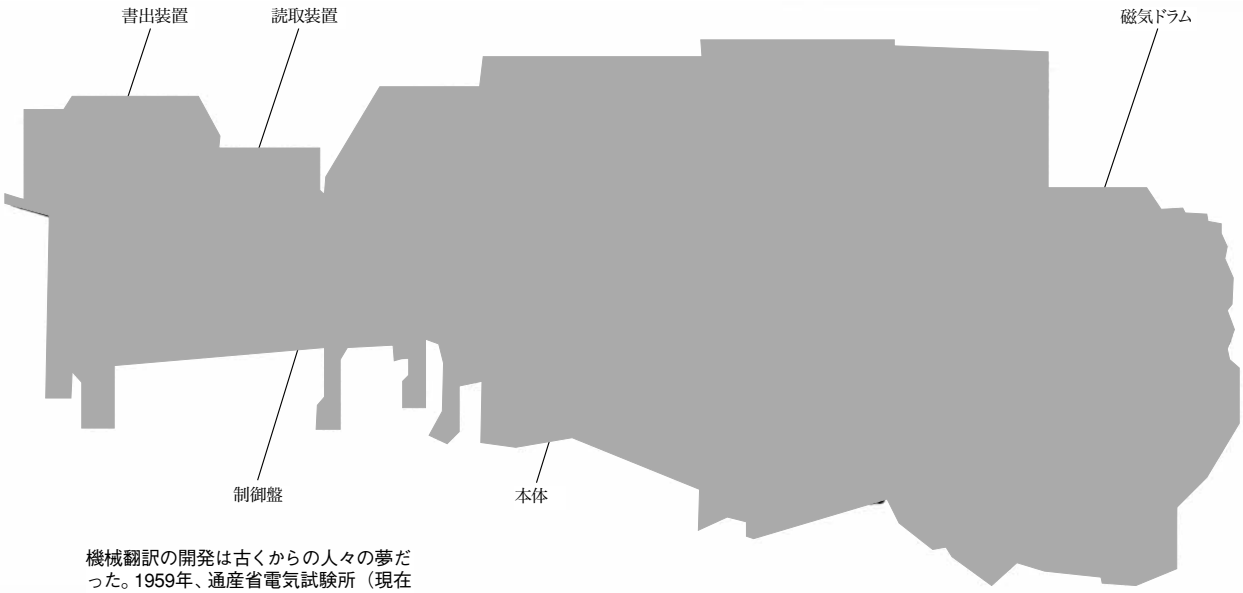
人間ではなくコンピュータが行なう翻訳が身近になってきた。日本で、コンピュータによる本格的な翻訳システムが登場したのは、1980年代の半ばである。そのころ、筆者は次のような翻訳システムの研究開発に携わっていた。

当時、「Machine translation」の訳語として「機械翻訳」が広く使われていた。「Machine」は「コンピュータ」を指す。実際には、コンピュータで動くソフトウェアのことである。

一般には「自動翻訳」という用語が使われることもあるが、業界では「翻訳支援ツール」の一つとして位置づけられている。現在では「翻訳ソフト」が広く使われている。

85年に開催された国際科学技術博覧会（つくば科学万博）で、日本の大手電算機メーカー各社が機械翻訳システムを披露した。この前の年、ブラビス社がパソコンで動く翻訳ソフトを発売したのは、業界に大きな衝撃を与えた。この影響を受けて、富士通、日立などの電算機メーカー各社がなだれを打つように、翻訳ソフトの発売を発表した。

この当時、私は富士通研究所で「ATLAS」という機械翻訳システムの開発に携わり、英語の生成を担当していた。そのときに、マスコミ各社から



機械翻訳の開発は古くからの人々の夢だった。1959年、通産省電気試験所（現在の独立行政法人産業技術総合研究所）が開発した機械翻訳専用機「やまと」。紙テープに「I like music」とパンチし入力すると、「ワレガ　オンガクヲ　コノム」とプリンタに出力できた

写真提供：産業技術総合研究所

依頼された翻訳は、なんと、小説や詩歌など、文学的なものだった。

『雪国』の書かれていない主語をどう訳すか

私に手渡されたのは、「国境の長いトンネルを抜けると雪国であった」という川端康成の小説『雪国』の冒頭の一節である。もともと、科学技術分野の翻訳を念頭においていたこともあり、これは難題だった。思案に暮れた私は、日本橋の丸善で同書の英訳*Snow Country*(Edward G. Seidensticker 訳)を買い求めた。

いま、手元の資料が散逸してしまったが、当時の記憶によって冒頭の一文の訳文を再現してみた(参考1)。

英語の主語に相当する語句が日本語には明示的に書かれていないので苦慮した。サイデンステッカー氏の訳文を見ながら、日本語の解析担当者と議論して徹夜したことを思い出す。日本語を解析し、英語を生成するまで、膨大な処理になる。結果論になるが、ATLASの訳文は、分詞構文も破格だし時制もへんだ。だが、これが精一杯だった。

ここで、インターネットで利用できる無料の翻訳サービスを利用して得られた3種類の訳文を紹介する(参考2)。参考まで、ページ下には翻訳サイトのアドレス(URL)も添えておく。

ヤフーの訳文には、主語として「it」が補われている。なんとエキサイトでは「train」が補われている。グーグルの訳文では「長い」の係り先に誤りがある。「Long Border」では「長い国境」という意味になってしまう。いずれにしても、20年以上も昔のレベルと比べても、あまり大きな進歩は見られない。

1984年、ブラビス社より発売された翻訳ソフト「マイクロ パックJ/E」。1時間に3000語を処理できた。当時の朝日新聞第一面のトップ記事に「自動翻訳機売り出す」と報道され、話題となった。現在発売されている翻訳ソフト「訳せ!!ゴマ」は、このソフト開発の流れを汲むもの　写真提供:エムティラボ(株)

アメリカのコダック社はどこにあるのか

90年代の初め、アメリカ最大の翻訳者の団体(ATA)が英日翻訳の検定試験を行なった際、私も採点者の一人として参加した。そのときの試験問題の一つに、写真で有名なコダック社の歴史的な経緯を説明している問題があった。訳された答案の日本語を読み比べ

【参考1】機械翻訳システム「ATLAS」による川端康成『雪国』冒頭の英訳

◆原文

国境の長いトンネルを抜けると雪国であった。

[ATLASの訳文]

When passing through a long border tunnel, there was a snow country.

[Seidenstickerの訳文]

The train came out of the long tunnel into the snow country.

【参考2】インターネットの翻訳サービスによる川端康成『雪国』冒頭の訳例

[ヤフー]

When it passed through the long tunnel of the border, it was a snowy district.

[エキサイト]

When the train came out of a long tunnel at the border, the snowy country there.

[グーグル]

Long border and go through a tunnel of snow country.

※インターネットで利用できる無料の翻訳サイトのアドレス(URL)

ヤフー● http://honyaku.yahoo.co.jp/transtext

エキサイト● http://www.excite.co.jp/world/

グーグル● http://translate.google.com/translate_t

て、疑問に思った箇所がある（参考3）。

【参考3】英日翻訳の検定試験に出された試験問題の例

◆原文
Rochester, New York's prosperous East Avenue

［訳文1］
ニューヨークのにぎやかなイーストアベニューのロチェスター

［訳文2］
ニューヨークのロチェスターのにぎやかなイーストアベニュー

「ニューヨーク」と「ロチェスター」と「イーストアベニュー」との相互の位置関係はどうなのだろうか。私は、地図を開いて、ニューヨーク州が五大湖のオンタリオ湖まで広がっていること、そこにロチェスターという都市があり、さらに、そこにイーストアベニューがあることを確認した。この結果、（参考3）の［訳文1］は誤訳であり、［訳文2］が正解であると判断した。

いま、ここで、グーグルの地図検索を利用すると、この事実は簡単に確認できる。さらに拡大し、［航空写真］をクリックすると、場所の様子もわかる。

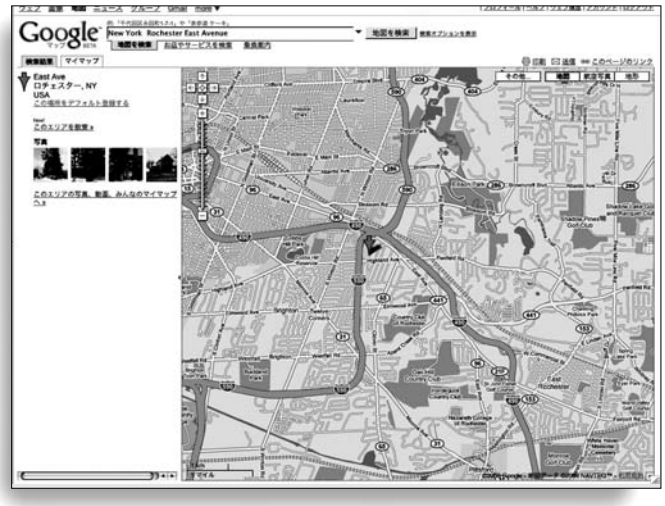
なお、「ロチェスター」という地名は、ニューヨーク州のほかに、ミネソタ州やニューハンプシャー州、さらにはイギリスにもあるので、早合点は禁物である。しかし、コダック社の話であれば、ニューヨーク州の「ロチェスター」に決まる。

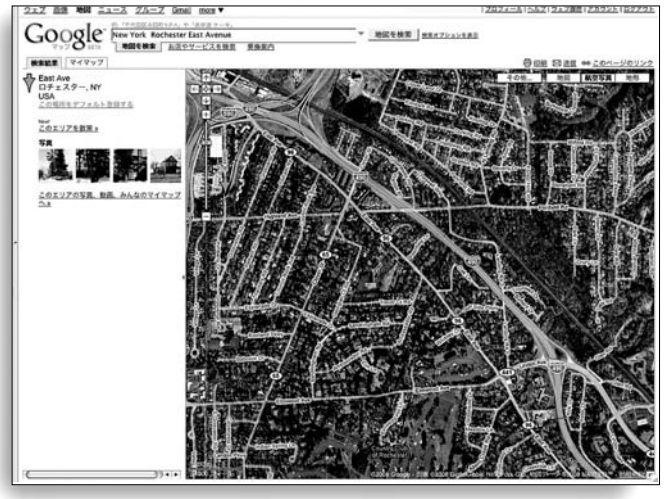
いずれにしても、日本語と英語とでは、住所の並べ方が異なることを改めて感じた。当時はインターネットで地図検索をすることはできず、そこを訪れた経験がない翻訳者には難題だったと思う。

ここで、「Rochester」の表記について考えてみたい。たとえば、インターネットで公開されている「エンカルタ（encarta）」の辞書で発音を聞いてみる（注）。

この音声をカタカナで表記すると「ラチェスタ」のように聞こえる。ところが、手元にあるいくつかの市販の辞書を引くと、どれも「ロチェスター」と記述されている。ここで、いくつかの表記をグーグルで検索してみた。

「Rochester」のさまざまな表記をグ





Googleの地図検索で表示されたロチェスターのイーストアベニュー
http://maps.google.co.jp/maps?utm_source=ja-wh-
イーストアベニューは画面の左上から右下へ走っている

注● エンカルタで「Rochester」を検索した画面
http://encarta.msn.com/dictionary_/Roch%25EF%25BD%25A5es%25EF%25BD%25A5ter.html

表1「Rochester」の訳語候補のヒット件数の比較（検索日：2008年4月）

検索表現	ヒット件数
“ロチェスター”	478,000
“ローチェスター”	4,410
“ロチェスタ”	3,100
“ラチェスター”	26
“ラチェスタ”	18

ーグルで検索してヒット件数を比較してみたのが表1だ。なお、完全一致で検索するために、半角の二重引用符で囲んである。

この結果、日本語では「ロチェスター」の使用頻度が圧倒的に高いことが推定できる。英語の発音に近い「ラチェスタ」は少ないが、今後この都市を訪れる日本人が増えれば、変わるかもしれない。

翻訳では、できるだけ広く使われている表現を使うほうが、違和感なく読者に伝わる可能性が高い。さまざまな訳語がある場合、対象読者を考慮する視点も必要になる。

「日本の東海」とはいったいどこのことか

青山学院大学理工学部で技術英語を担当していたころのエピソードを紹介する。99年9月末に、東海村で臨界事故が発生した。私は、このニュースを報じた『Nature』の記事を、学生たちに課題として翻訳させた。ある学生の訳文の一部を紹介する（参考4）。

【参考4】『Nature』の記事の原文と英訳例

◆原文
a nuclear-fuel processing plant in Tokai, Japan

[ある学生の訳文]
日本の東海の原子燃料処理プラント

[筆者の試訳]
茨城県東海村の核燃料加工施設

私は、授業でこの訳例を示して、学生全員に「日本の東海」とはどこか、「原子燃料処理プラント」でいいか、などと尋ねた。そして、「この『Nature』の記事は、日本で報道されたテレビや雑誌、新聞などの情報を参考にして、これを英語で表現したものであるはずだ」と説明した。

学生たちと一緒に、いくつかのニュースサイトにアクセスして、「Tokai」が「茨城県東海村」であることを確認した。さらに、多くの新聞では「核燃料加工施設」と書いてあることも確認した。

私は、「英語では『Tokai』だけではどこにあるのか不明なので、日本の都市であると補足しているが、日本人にとっては『日本の』は不要であり、日本人の読者には『茨城県の』と補足するほうが親切だろう」と説明した。このように、「同じ事実でも、日本語と英語では対象読者が違うので、表現も異なる」と補足した。

さまざまな表記をグーグルで検索してヒット件数を比較してみた。この結果、現在でも、「核燃料加工施設」が広く使われていると推測できる（表2）。

表2「nuclear-fuel processing plant」の訳語候補のヒット件数の比較（検索日：2008年4月）

検索表現	ヒット件数
“核燃料加工施設”	4,210
“核燃料処理施設”	1,250
“核燃料加工工場”	1,050
“核燃料処理プラント”	3
“注(1)”	1,480
“注(2)”	25

［注］次のように最大ヒット件数の表現にいくつかの語句を追加してAND検索した。
(1)“核燃料加工施設”東海村　臨界事故
(2)“核燃料処理施設”東海村　臨界事故
なお、完全一致で検索するために、半角の二重引用符で囲んである（表1も同じ）

『Nature』の記事について、先ほどと同じく、インターネットの翻訳サービスを利用して得られた3種類の訳文を紹介する（参考5）。

【参考5】インターネットの翻訳サービスによる『Nature』の記事の訳例

◆原文
a nuclear-fuel processing plant in Tokai, Japan

［ヤフー］
東海（日本）の核燃料処理植物

［エキサイト］
トーカイ、日本の核燃料加工施設

［グーグル］
核燃料加工工場では東海、日本

エキサイトでは「トーカイ」とカタカナになっている。ヤフーでは「核燃料処理植物」と思いがけない迷訳になっている。グーグルでは意味不明な訳文になっている。

辞書と文法の二刀流の限界をデジタル技術で乗り越える

たとえば、80年代の半ばに登場した富士通社のＡＴＬＡＳは1カ月の使用料金だけで１０００万円もしたといわれる。その後80年代の後半に、ワークステーションで動く機械翻訳ソフトが発売されたが、数百万円と高価であり、ハードと一緒に購入すると１０００万円近くになるため、一般の人が個人で利用するには高嶺の花であった。

90年代に入ると、各社の翻訳ソフトは数十万円ほどに下がった。さらに、90年代後半になると、インターネットの普及に伴い、無料で利用できる翻訳サービスが登場しはじめた。

そして、今日、インターネット上には、翻訳サービスを無料で利用できるサイトがたくさんあり、数千円で購入できる市販の翻訳ソフトも登場している。

書く人がいて、読む人がいて、伝える人がいる。翻訳者は、書く人の立場で原文を理解し、さまざまな読者を想定して表現を工夫しなければならない。その思考過程を工学的に追認するという意味で、機械翻訳システムの研究開発が少なからぬ貢献を果たしたことは間違いない。

しかし、従来の辞書と文法の二刀流には限界がある。この限界を乗り越えるためには、デジタル技術が役立つと筆者は考えている。たとえば、地図検索や画像検索、ヒット件数などのデジタル技術を利用すると、書く人と読む人の双方の立場を疑似的に体験することができる。

もともと翻訳はアナログの世界だが、デジタル技術を利用することで、広く読者に伝わる翻訳が可能になると思う。最近は、機械翻訳結果をそのまま出力するのではなく、対訳事例のデータベースやネット検索で収集した用例を参考にして、修正した訳文を出力するシステムも登場している。今後もデジタル技術を翻訳に活用する研究開発の動向に注目したい。

90年代後半のネット上の翻訳サービス登場で、パソコン用翻訳ソフトは廉価になり、入手しやすくなった。家電量販店のソフト売り場にて。2001年　写真提供：読売新聞社