

# 機械翻訳の精度を上げるための構文解析の提案

岩垣 守彦

前玉川大学教授

## 1. 機械翻訳ソフトで英語を日本語に訳させてみる

現在、日本には、日英・英日の翻訳ソフトがいくつもあるが、最新版のソフトでも「小説・ドラマなどの芸術の分野、会話文は対象としていない」と断っている。では、普通の英文なら、正しく日本語に訳すことができるのだろうか。（使ったのは、市販のソフトとインターネット経由のプロバイダーの翻訳サービスであるが、会社名やプロバイダー名は伏せておく。）  
コンピュータに訳してもらおう。

### 1. He went to London to study English.

彼は、英語を勉強するために、ロンドンに行きました。  
彼は、英語を研究しにロンドンに行った。

これをみると、「翻訳ソフト」は私たちの英文解釈と合っているように感じられるが、はたしてこれでよいのか。

### 2. The train will make a brief stop at Machida before reaching Sinjuku terminal.

列車は、Sinjuku ターミナルに到着する前に、Machida に  
おいて短い停止をします。  
新宿端末に達する前に、列車は町田で簡潔な停止をするだろう。

この文は小田急のロマンスカーの車両の前の扉の上に出る電光テロップであるが、文の全部が一度に提示されるわけではない。この訳し方で訳すと、before 以下が出るまで全文を理解することができないことになる。このような約仕上げの欠陥は関係詞を含む文になるとさらにはっきりする。

### 3. There was once a Goldfish who lived in the sea in the days when all fishes lived there.

すべての魚がそこに生きていた日に海に住んでいたキンギョが昔いたことがあります。  
すべての魚がそこに生息した数日には海の中に住んでいた Goldfish がかつてあった。

構文は正しくとらえているように思える。しかし、訳文としては駄目であり、これも、最後まで文を読まないと全文を訳すことができない。関係詞節の取り扱いがこれでよいのかという問題が含まれている。

次は、もう少し複雑な新聞の記事である。これを機械翻訳にかけてみる。

### 4. It was almost midnight on a Saturday evening in December 1994 when General Colin Powell received a telephone call from President Clinton asking the former chairman of the Joint Chiefs of Staff to drop into the White House for a chat.

The Need for New Software to convert Information in Information-unit Order between Japanese and English  
IWAGAKI Morihiko Ex-professor of Tamagawa University  
iwagaki@lit.tamagawa.ac.jp

一般 Colin Powell がクリントン大統領から電話呼び出しを受け取った 1994 年 12 月の土曜日の夕方、雑談のためのホワイトハウスに立ち寄るために統合参謀本部に前者の議長を要求することが、ほとんど真夜中でした。

私たちの構文理解の不備が「翻訳ソフト」の上にはっきりでているような気がする。これは、人間の脳では可能であるがコンピュータではできない「仕組み」をもとにしてソフトが作られたのではないかと疑わざるをえない。と同時に、そもそも、私たちの構文理解は正しいのだろうか、私たちは英語と日本語の対応関係を正しくとらえているのだろうかという問題が浮かんでくるのである。

## 2. 機械翻訳ソフトで日本語を英語に訳させてみる

### 1 「あそこにお巡りさんが立っている」

There a/the policeman is standing.  
The policeman is standing there.

### 2 三年ばかりニューヨークに暮らした私には、国際人になるということがどんなに大変なことかよくわかる。

主語なし A serious case how or I understand that becomes, to me which lived only 3 in New York a cosmopolitan well.  
I who lives only by three years in New York am serious becoming to an international person or understand well.

どちらもうまくいかない。特に後者の文では関係代名詞の使い方がおかしい。これらの文を人間が訳すと次のようになる。

### 1 There's a policeman standing over there. Look! A policeman is standing over there.

### 2 Having once lived for three years or so in New York, I know very well how difficult it is to become an "internationally minded person."

これもコンピュータに正しい情報をインプットしていないのではないかと考えざるをえない。コンピュータが正しい翻訳文を提示できないというのは、日本語の情報が正しくインプットされていないし、それに対応する英語の情報も正しくインプットされていないからである。

## 3. 等価翻訳とはどういうものか

正しい等価関係を機械翻訳で示すためには、「言語」（ここでは「日本語」と「英語」）を「情報伝達の素材」として見て、その「仕組み」を改めて考え直さなければならないのではないと思う。しかし、翻訳ソフトの訳文を見て推察されることは、古くから異国語に対して私たちが使ってきた「返り点式解釈法」をベースにして「ソフト」が作られているのではないかとということである。

私たちは、漢文の文構造を分析して「返り点式解釈法」を使えば、日本語の文法に従って「意味」がわかるということを知っている。このような漢文の処理法は、朝鮮半島や日本に限ったこと

ではないらしい。最近の研究によると、漢字文化圏の西の端ウイグルでも、漢字の仏典が古代ウイグル語で訓読されていたことがわかってきている。(雑誌『言語』1996 年 8 月号 p.37 庄垣内正弘『漢字文化圏西端にも存在した「漢文訓読」 ウイグル僧は漢文をどう読んだか』)

となると、外国語の文構造を見極めて母国語の論理・文構造で理解するというのは、「外国語理解の一般的な方法」なのかもしれない。ただし、外国語を構文分析して、「返り点」を使って文の意味を自国語で理解するという方法は、必ずしも、当該外国語をその意味する通りに理解することができるということを保証するものではない。それは、「誤解」も含めて、外国語を「納得する」ための方法でしかないのかもしれないのである。

では、そもそも「翻訳する」とはどういうことであろうか。極めて簡単な例で示そう。

ある事象が起きる。——お客が店に入ってくる。  
イギリス人の店員は英語で——“Can I help you?”  
日本人の店員は日本語で——「いらっしゃいませ」

ある事象が起きる。——パーティーで誰かを探しているような人がいる。

イギリス人の給仕は英語で——“Can I help you?”  
日本人の給仕は日本語で——「どなたかお探ですか」

つまり、等価翻訳とは、一つの事象に対して、同時に、ほとんど無意識に二つの言語で表現することである。このような翻訳を機械的に行うことは、無限の定形表現に対応して無限の相当表現を集積をしない限りできない。理論的には可能であろうが、実際には不可能と思う。それでも、コンピュータを中において、異国語同士の若者が母国語を捨てることなく、即時に直接的に話しかけ答え合う場面を見たい。

では、どうすればこのような錬金術的な夢が可能になるだろうか。

#### 4. 文を「単位情報」順に変換する

「言葉」による伝達では動詞が重要である。たとえば、  
“I love you.”  
という英語では、love にストレスがある。この love は動詞である。フランス語では、語順が変わって  
“Je t'aime.”  
であるが、aime にストレスがある。aime は動詞である。日本語の場合はもっとも極端で、  
「好きです」  
と動詞だけになってしまう。これは、言語では、「動詞を中心にして一つの単位情報が形成される」ということであり、その場合、文に必要な基本的要素は「動詞 + 補助(名詞・代名詞)」ということである。これは世界の「言語」に共通することである。(1830年に英国聖書協会が「ヨハネによる福音書」第3章16節「神は、その独り子をお与えになったほどに、世を愛された。独り子を信じる者が一人も滅びないで、永遠の命を得るためである」(新共同訳)を630の言語と方言に訳したが「名詞」と「動詞」のない言語はなかった。)

これはどういうことかと言うと、たとえば、最初に示した

He went to London to study English.

は、学校英文法では「単文(simple sentence)」であるが、動詞が二つ(went と to study) 含まれているので、この文は二つの「単位情報」からできているということになる。したがって、この文の発言意図は、単位情報順に二つの事象を伝えることにある。つまり、

He went to London (彼はロンドンに行くという行為を過去にした)

(He) to study English there. (それ以後そこで英語を勉強するためだった)

の二つの事象が to 不定詞という「つなぎ」で結ばれている

「彼がロンドンに行ったのは(それ以後そこで)英語を勉強するためでした」

という情報を提示していることになる。  
したがって

The train will make a brief stop at Machida before reaching  
Sinjuku terminal.

という文は

The train will make a brief stop at Machida  
reaching Sinjuku terminal

という二つの単位情報が before という「つなぎ」で結ばれていて、未来形 will および、動名詞 reaching の情報を組み入れると

「電車は(スケジュールに従って)町田にちょっと止まってから、(いつもの通り)終点新宿に着きます」

という情報を提示していることになる。  
関係詞を含む次の文

There was once a Goldfish who lived in the sea in the days  
when all fishes lived there.

も

There was once a Goldfish  
who(=he) lived in the sea  
All fishes lived there

という三つの単位情報が who; in the days when という二つの「つなぎ」で結ばれているのであるから、

「昔、金魚がいました」(There was once a Goldfish.)  
「その金魚は海に住んでいました」(who(=he) lived in the sea.)  
「その当時、お魚はみんな海に住んでいたのです」(All fishes lived there.)

となり、これに「つなぎ」を加えて、前から順に単位情報を日本語に変換すると

昔、金魚がいて、それは海に住んでいました。そのころお魚はみんな海に住んでいたのです。

となる。先に出したもう少し複雑な新聞記事も同じように処理できる。

It was almost midnight on a Saturday evening in December 1994/ when General Colin Powell received a telephone call from President Clinton/ asking the former chairman of the Joint Chiefs of Staff/ to drop into the White House for a chat.

これを動詞を中心にした単位情報の順に日本語に変換すると

1994 年の 12 月ある土曜日の真夜中ごろのこと、コリン・パウエル將軍はクリントン大統領からの電話を受けた、クリントン大統領は前統合参謀本部議長に、ホワイトハウスにちょっと立ちよっておしゃべりしていかないかと言うのであった。

となる。暖炉のそばで練りに練った19世紀の英文と違って、現代の英文はほとんど前から順に単位情報をひろって読めば、情報がとれるように書かれていると言ってよい。

したがって、もし情報を英文からできるだけ正しく得たいなら、英語によって与えられた情報を単位情報順に日本語に変換することによって「筋・イメージ」を正しく追うことができるのではないかと思うのである。このことは「イメージ」を大切に詩に於いても有効である。

詩はイメージの順序が大切と思うのであるが、現在、旧来の「返り点式解釈法」によって訳されている。次の二行は、ある有名は詩の原文の一部と市販されている訳詩の訳文である。

We passed the school where children played;  
Their lessons scarcely done.

子供たちが遊んでいる学校のそばを通りました  
授業はほとんど終わっていなかった。

これではイメージの順序が原文の順序と違ってしまふ。したがって、訳すなら

私たちが学校のそばを通ったら子どもたちが遊んでいました  
ああ、やっと授業が終わったところなんだ。

としなければならないと思う。

このように、動詞は中心になって「単位情報」を作るのであるが、それだけでなく、動詞は、さらに、「述語動詞」として、「言語情報」にとって大切な「現実の時間との関係」をあらわす。

「いつものこと」は現在時制  
「今現在のこと」は現在進行形  
「すんだこと」は過去時制  
「今以後のこと」は未来形

さらに、重要なことは、この「現実の時間との関係」は、述語動詞を軸として、準動詞や助動詞にも適用されるということである。たとえば、準動詞は

「いつものこと」は動名詞 Seeing is believing.  
「今現在のこと」は現在分詞 Take a look at the baby sleeping in the cradle.  
「すんだこと」は動名詞（能動的結果）と過去分詞（他動詞は受動的結果・自動詞は能動的結果） I remember seeing him somewhere before. / fallen leaves  
「（今）以後のこと」は(to)不定詞（一時性もあらわす） To see is to believe.  
と対応している。

この「動詞を中心にした単位情報の順に処理する」という方法は「英語から日本語」だけでなく「日本語から英語へ」の変換にも利用できる。

晩夏の夜空はいちめんの星明かりで、空気に強い草の香がまじった。広場の奥は雑草が群がり、しきりに地虫の声が沸いている。——伊藤桂一『水と微風の世界』

この日本文の単位情報は、動詞を中心に「イメージ」の順に分解すると、四つの単位情報に分けることが出来る。それに相当する英語をつけると、次のようになる。

晩夏の夜空はいちめんの星明かりであった。  
The late-summer night sky was suffused with starlight.  
空気に強い草の香がまじった。  
A strong fragrance of grass mingled in the air.  
広場の奥は雑草が群がっていた。  
The far side of the open space was crowded with weeds.  
（そこから）しきりに地虫の声が沸いていた。  
From there arose a steady chirruping of insects

この単位情報に「つなぎ」を加えてまとめると

The late-summer night sky was suffused with starlight and a strong fragrance of grass mingled in the air. The far side of the open space was crowded with weeds, from which arose a steady chirruping of insects.

となる。

次の文は五木寛之の小説の一部である。

波留子はあらかじめ電話で連絡を受けていたので、夕食の支度のための買物をすませ、部屋を掃除して布由子の帰りを待っていた。

この文は動詞を中心にした単位情報は四つである。それに相当する英語を与える。

波留子は前もって電話で連絡を受けていた  
Haruko had been warned in advance by telephone  
（彼女は）夕食の買い物をした  
she had done the shopping for dinner  
（彼女は）部屋を掃除した  
(she had) cleaned the room  
（彼女は）布由子が帰宅するのを待っていた  
(she was) in readiness for Fuyuko's homecoming

これに、日本文の「つなぎ」（ので・・・；・・・、・・・して）に相当する英語の「つなぎ」を加えてまとめると、

Haruko, who had been warned in advance by telephone, had done the shopping for dinner and cleaned the room in readiness for Fuyuko's homecoming.

となる。これは原日本文と同じイメージの順序であり、内容もほぼ等価の変換になっている。

## 5. おわりに

考えてみると、このような「動詞中心の単位情報の順に変換する」という方法は、新しいものではない。「英語とフランス語の変換」では、当たり前である。たとえば、

Please/ give me a glass of water.  
Donnez-moi un verre d'eau / s'il vous plait.

Please/ call someone/ who speaks English.  
Appelez-moi quelqu'un/ qui parle anglais./ s'il vous plait.

のように。したがって、これらの「返り点史記英文解釈」に関する問題は「日本語と英語の間」でのことかもしれないが、文中の動詞の処理は考えを改めなければならない。というのは、動詞情報と関連するのであるが、特に「翻訳ソフト」で一番気になるのが「関係詞」の処理だからである。すでに示した例文もそうである。

3. There was once a Goldfish who lived in the sea in the days when all fishes lived there.

すべての魚がそこに生きていた日に海に住んでいたキンギョが昔いたことがあります。

すべての魚がそこに生息した数日には海の中に住んでいたGoldfish がかつてあった。

2 三年ばかりニューヨークに暮らした私には、国際人になるということがどんなに大変なことがよくわかる。

主語なし A serious case how or I understand that becomes, to me which lived only 3 in New York a cosmopolitan well.

I who lives only by three years in New York am serious becoming to an international person or understand well.

のように、どちらもだめである。一般に「・・・した～は・・・」に類する文をコンピュータで訳すと、必ずミスが生じる。これは、翻訳のプロは分詞構文で訳す。たとえば

1949 年に二度目の渡仏をした藤田嗣治は、その年のうちに、レストランにひとりで座っている女性を描いた。

は、翻訳ソフトでは

Fujita Tuguharu who did the Watari Buddha of the eye twice in 1949 years painted the woman who is sitting in a/the restaurant alone in the year.

と訳してしまうが、そうではなくて、正しくは

Visiting to France for the second time in 1949, Fujita Tsuguharu, within the same year, painted a picture of a woman sitting alone in a restaurant.

である。

日本で、翻訳ソフトに関する作業が始まって30年くらい経っただろうか。コンピュータの素人として、専門家をお願いしたいことが二つある。一つは、動詞を中心にした単位情報を順に処理して「文書の筋道」を正しくつかむことが出来るような受信用の「翻訳要約ソフト」である。たとえば、

Many people mistakenly suppose that teaching is a professional activity requiring long, complicated training and an official teacher's license.

というような文をコンピュータにかけると、まず、文中の主観的修飾語句が一瞬に消えて

people suppose  
teaching is an activity  
the activity requires a training and a license

という骨格を捕らえ

人々は、教えることは一つの活動で、それには訓練と免許がいると思っている。

という主旨を伝えることはできないだろうかということである。つまり、この際、「言葉の等価性」を犠牲にしても「事象と筋道の等価性」を重視した「異国語間変換ソフト」を作ることはできないだろうかということである。もう一つは、発信の手段として、先に示したような「日本語から英語へ」の変換ソフトが出来ないかということである。

外国語の学習に情熱を捧げる若者がいてもいいが、文化間の交流に言葉の障壁があるなら、また、若者たちが母国語を捨てることなく異国の若者とコミュニケーションできるようになるなら、その願いをコンピュータが叶えてくれるとうれしいと思う。

#### 【参考文献】

- 岩垣守彦(1993) 『英語の言語感覚----ルイちゃんの英文法』(玉川大学出版部、東京)(本論で述べた文法の基礎はこの本に示されている。また、その実際の運用は、以下の本の中で例示している。)
- 岩垣守彦(1993) 『日本人に共通する和文英訳のミス』(ジャパントイムズ、東京)
- 岩垣守彦(1994) 『よい英文を書くための和文英訳のテクニック』(ジャパントイムズ、東京)
- 岩垣守彦(1996) 『辞書ではわからない英語の

使い方』(ジャパントイムズ、東京)

岩垣守彦(1999) 文法理論の文学への適応について、日本認知科学会「文学と認知・コンピュータ」研究分科会第一回「文学の設計と実装」プロジェクト研究会。

岩垣守彦(2000) 「イメージの形成と言語発生モデル」から「文学のモデル」へ。日本認知科学会テクニカルレポート「文学と認知・コンピュータ6--ことばと文学--」,(32),69-82.

佐藤理史(1997) 『アナロジーによる機械翻訳』(共立出版、東京)