

日本語の論述様態の定型化について

A draft proposal on Converting Japanese to English in Information-unit Order

岩垣 守彦

Morihio Iwagak

前玉川大学教授

Former Professor at Tamagawa University

Abstract

I earnestly hope that a world may come into existence where everybody communicates with each other through their own language: When a young man wants to befriend another young man from a different culture and in a different language, they can both communicate in their own language as smoothly as if they were people from the same country.

As a tool to assist in the realization of this alchemical dream, I would like to use computers whose software is based on a series of devices common to all languages in the world.

But my hope has not yet been realized, even by Machine Translation software. In fact, we cannot translate Japanese to English effectively.

Therefore, I'd like to make a proposal concerning the idea of converting Japanese into English, in information-unit order.

1. 「単語」は「集合イメージの符牒」

日本の幼児にいろいろな種類の犬を見せて「ワンワン」と命名する。足の短い胴長のダックスフントを見せて「これは何だ」と尋ねると子供たちは「ワンワン」と真似して命名する。このとき子供たちの脳に出来ているのは「犬の集合イメージ」であり、「ワンワン」はその「集合イメージに付けられた符牒（グループ共通の呼称）」である。

「同一集合イメージ」に別のグループが別の「符牒」を付けるのは自由である。「ワンワン」を別のグループの符牒（フランス語）の *tou-tou* と付け替えることができるし、また別のグループ（英語）の *woof-woof* と付け替えることもできる。「言葉」はグループの構成員個々人の脳に蓄積された共通の集合イメージに付けられた「グループ共有の符牒の総体」であり、「意味」とは「集合イメージ」のことである。

したがって、「同一集合イメージ」に付けられた「符牒」なら、どのグループの「符牒」を使っても、発信者は「同一集合イメージ」を受信者に喚起することができる。したがって、発信は「理解」される。

また、「集合イメージ」が同じであるなら「符牒」は付け替えることができるので、「ワンワン」を *tou-tou* と、あるいは *woof-woof* と付け替えることもできる。したがって、未知の符牒（フランス語・英語）を、対応する既知の符牒（日本語）と付け替えることによって、未知の符牒の持つ意味を「既知の符牒」のもたらず「集合イメージ」で理解することができるかと仮定することができる。

2. 未知の「符牒列」も既知の「符牒列」に変えて理解することができる

「犬の集合イメージ」を「空間」の中に描く。それに「時間」を加えると、「犬の集合イメージ」は動き、吠える。「吠える」も集合イメージである。「ワンワン（集合イメージ）」と「吠える（集合イメージ）」の二つの集合イメージに符牒を付け、{ワンワン（符牒）；吠える（符牒）}と組み合わせると、「日本語ルール」で「ワンワンは吠える」と出力すると日本語の「符牒列」になる。この集合イメージの組み合わせは { *woof-woof* ; *bark* } と符牒を付け替えることができる。それに「英語ルール」を加えて“*Woof-woofs bark.*”と出力すると英語の「符牒列」になる。したがって、「同一集合イメージ」を背後に持つ「符牒列」の間では、相互に対応する「符牒付け替えルール」

を決めると、一つの言語が発信した「符牒列」(ワンワンは吠える。)は他の言語の「符牒列」(Woof-woofs bark.)に変換することができる。

ということは、背後に共通の「集合イメージ列」があるなら、「意味」がわからない「未知の符牒列」(I love you.)でも、符牒を張り替えることで「既知の符牒列」(私はあなたが好きです。)に変換できて「意味」がわかると仮定することができるということである。

3. 「集合イメージ」の処理

言語はそれぞれが個別の「符牒配列ルール」(I love you./ Je t'aime./ 私はあなたが好きです。)で処理されているが、どのような「符牒列」でも、そのルールを相互に対応させると、実際に使用できるかどうかとは関係なく機械的に「あるグループの符牒列」を「別のグループの符牒列」と付け替えることが出来ると仮定することができる。I love you forever. 私はあなたを永遠に愛します。あそこにお巡りさんが立っている。 A policeman is standing over there. のように。

4. 「単位情報」が一つの場合

「言語情報」は「動詞を含む複数の集合イメージ(共通) + 符牒(個別) + 符牒配列規則(個別)」で成り立っているため、人間しかできないような翻訳を望まなければ、符牒の付け替えで日本語に変換することが可能である。実際、[池原悟(2005)]によると「翻訳実験では、IPAL (IPA 1987) の例文 5000 件に対して、90%の正解率が得られており、方式限界は 97%であることが報告されている。(金出地ほか 2001)これにより、日英機械翻訳における単文の意味的な訳し分けの問題はほぼ解決できたとみられるので... (pp.1-2)」とのことであるから「単位情報」が一つの場合は符牒の張り替えに問題ないように思われる。

5. 「単位情報」が複数ある場合

「単位情報」が複数ある場合はどうであろうか。「言語情報」の基本は次のように考えることができる。

1. 「言語情報」は「符牒(語彙)」と「符牒配列規則(文法)」に基づいて作られる「単位情報」(動詞を含む複数の集合イメージ)を、前から順に並べて提示する。

2. 「単位情報」が複数ある場合には「つなぎ」を使うこともできる。

3. 「つなぎ」は「機能語」で情報と情報の関係性を明瞭にする機能を持っている。(なお「つなぎ(conjunctives)」とは「単位情報をつなぐもの」のことで、接続詞、関係詞のほか、準動詞も含まれる。)

この観点から日本語における「単位情報」と「つなぎ」の関係を、例文をあげて考えると、

私は昨日街に行きました(単位情報) + {通りを歩いていた(単位情報) + 時に(つなぎ)} + 花子に会いました(単位情報)

のようになっている。それに対して、上の日本語を英語にして示すと、

A: I went downtown yesterday. I was walking along the street when I happened to meet Hanako.

B: I went downtown yesterday. When I was walking along the street, I met Hanako.

の二つの符牒列が可能である。

情報伝達に際してどちらの型を選ぶかは「伝達する情報の比重によって決まる。「個々の単位情報を同じ比重で伝える」場合は(A)の型(したがって、上の例で言えば I went downtown yesterday. は、次に続く単位情報に影響を及ぼさない)であり、「単位情報の一つを了解情報として比重を軽くして伝える(時の設定、条件の設定も含む)」場合は(B)の型(前に I went downtown yesterday. があるから、When I was walking... と続けることができる)である。

したがって、日本語と英文の「単位情報」と「つなぎ」の関係は、

日: 「単位情報 + つなぎ」 + 「単位情報」

という一つの型に対して、

英 A: 「単位情報」 + 「つなぎ + 単位情報」

B: 「つなぎ + 単位情報」 + 「単位情報」

と二つの型で対応することになる。

ただ、問題は、英語では「了解情報として単位情報の比重が軽い場合(時の設定、条件の設定も含む)」は(B)「つなぎ + 単位情報」 + 「単位情報」(When I got to the office, he was already drinking whisky.)として表層に現れるが、日本語では「情報の比重」は文脈の中に示されるだけという点である。

6. 論理展開・情報・イメージ順に変換する

翻訳家が日本語を英文に翻訳する場合は、日本語の文脈を読み取って「了解情報」として単位情報の比重を低くする型を選ぶことができる。しかし、コンピュータに「文脈を読み取って情報に比重を付ける」という作業を求めているのは、絶対に不可能とは言えないにしても、当分は無理ではないかと思われる。

したがって、日本語を英文に変換する場合には
日「単位情報+つなぎ」+「単位情報」
英 A「単位情報」+「つなぎ+単位情報」
と対応させるのが原文（日本語）の持っている論理展開・情報・イメージの順序を崩さないで英語に変換するプロの翻訳家の英文に近づく方法である。

7. 日本語の論述様態をパターン化

日本語の情報やイメージの順序を損なうことなく、出来る限り英語らしい英文にするために

日「単位情報+つなぎ」+「単位情報」
英 A「単位情報」+「つなぎ+単位情報」
と対応させるには、「日本語の論述様態」をパターン化して「英語の論述様態」のパターンと対応させなければならない。たとえば、日本語の

「単位情報+ので」+「単位情報」

に対して、英語では、

彼は大正の初めに貧しい農家に生まれたので、学校教育をほとんど受けられなかった。

He was born into a poor farming family at the beginning of the Tisho era, and had hardly any formal education. (ここのような因果関係を暗に示す内容の「ので」には and とか so などを使う。)

「今日は午後から大阪へ行くので、帰りは最終になるかもしれない」夫はそう言って玄関を出た。

“Today I have to go to Osaka in the afternoon,” said Masako's husband as he went out through the front-door. “I may not be back till the last train.” (話し言葉では、短い情報の積み重ねが普通なので 情報, (so) + 情報 となる。)

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。

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. (日本語と同じよ

うに文を引き締めるには、関係代名詞を使うのが最も適切である。)

犬が迷いこんできた。餌が欲しくて人に媚びてくるので、捨てられたばかりだとわかった。

The dog came wandering in from nowhere. It made up to people in the hope of getting food, which suggested that it had only just lost its owner. (この関係代名詞も and/so it...としても問題ない。)

のように

「単位情報+ので」+「単位情報」

に対して、英語では多様な「つなぎ」が使われている。しかし、どの場合にも、基本的には

英 A「単位情報」+「つなぎ+単位情報」

の構造で「つなぎ」は so で処理できる。つまり、「ので」に関しては

日「単位情報+ので」+「単位情報」

英 A「単位情報」+「so + 単位情報」

と集約すればよいことになる。

このように、日本語の論述様態と英語の論述様態を、日本語の前の単位情報の文末の「つなぎ」を英文では次の単位情報の先頭で処理するという対応関係でとらえると、たとえば、

「うれしいわ」と言って、彼女は膝頭で彼の膝を押しした。

“Nice man!” she said, nudging his knee with hers.

は

「・・・い(て・し・して・ながら),・・・た」

は、英語では

past tense,...ing,....

である。これを基本にすると

エミ子は曖昧に言い、言葉を濁して、その話題をさけた。

Emiko spoke noncommittally, deliberately using vague language, so as to steer off the subject.

は、それに「・・・しよう」とが加味されただけで、

「・・・い,・・・して,(・・・しようとし)た」

past tense,...ing,...(so as) to....

と集約的にパターン化できるなら、かなり通用する英文に日本語を変換することができるのではないかと思われる。

なお、この考えは日本語の文の論理展開・情報・イメージの順序を壊すことなく英語の文に変換して「等価翻訳」を作るという作業を、20年ほど、さるイギリス人と共同でしているうちに生まれたものであるが、[池原, 2005]の試みは、「パターン化」という点で、基本的には同じアイ

ディアではないかと思われる。

8. 「言語情報の並置化」

「論述パターン」に合わない日本語は、すべての情報を同じ比重で伝えるとして、日英両文をすべて

日 { 単位情報 + つなぎ } + 単位情報

英 単位情報 + { つなぎ + 単位情報 }

と対応させて、and, or, but, so, for などで結ぶか、あるいは、「つなぎ」を思いつかない場合には、すべての単位情報をそのまま並列すれば、ほとんどすべての言語情報を前から順に変換できるということである。たとえば、

昼食を終わって、光子叔母と二人で奥の部屋で荷物を整理していると、庭の方から口笛を吹く音が聞こえて来た。

(90EA6-1)128 テープ済み

は

「(~時,) ~していると, ~した」

(When....) ...was ...ing, when past tense.

であるから

When lunch was over, I was sorting out my things in a back room with Aunt Mitsuko, when we heard someone give a whistle in the garden.

であるが

I had the midday meal, and I was putting my belongings in order in a backroom with Aunt Mitsuko, then someone whistled in the garden.

としても通じる英語になるし、たとえば、

三年ばかりニューヨークで暮らした私には、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。

も、現在の翻訳ソフトのように

*I who lived in New York only in about three years understand becoming an internationally-minded person well in a however serious thing.

と誤訳しないで

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

と処理することができる。もっとも、プロの翻訳家は

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'.

と訳すのであるが。

参考文献

[池原悟(2005)]：「非線形性に着目した言語表現モデルと重文と複文に対するパターン辞書の開発」(第10回LACE研究会(2005.12.24-25) pp.1-2

[岩垣守彦(1993-a)]『英語の言語感覚----ルイちゃんの英文法』(玉川大学出版部, 1993)

[岩垣守彦(1993-b)]『日本人に共通する和文英訳のミス』(ジャパントイムズ, 1993)

[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, 1994)

[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, 1996)

[岩垣守彦(2003a)]「機械翻訳の精度を上げるための構文解析の提案」(自然言語処理研究会, 東京工科大学 2003/03/25)

[岩垣守彦(2003b)]「日本語の「つなぎ」と英文の「つなぎ」の対応に関して」(自然言語処理研究会, 東京工業大学 2003/05/26)

[岩垣守彦(2003c)]：An Idea of Translation of Japanese contact-clauses into English (第一回アジア辞書学会, 明海大学 2003/08/26)

[岩垣守彦(2003d)]「連体修飾節の英訳に関して」(自然言語処理研究会, 山形大学 2003/07/25)

[岩垣守彦(2006a)]「コンピュータによる異言語人とのコミュニケーション」(人工知能学会全国大会ワークショップ 2006/06/13)

[岩垣守彦(2006b)]「言語情報を前から単位情報順に変換する」(言語・認識・表現」第11回年次研究会, 明治学院 2006/12/19「)

[斉藤俊雄(編)1998]『英語コーパス言語学』(研究社 1998)

[佐藤理史(1997)]『アナロジーによる機械翻訳』(日本認知科学会編, 共立出版, 1997)