

日英翻訳における連体修飾句の訳し分け

北村 博 荻野 紫穂

日本アイ・ビー・エム 東京基礎研究所

日本アイ・ビー・エムで研究開発中の日英機械翻訳システム(JETS)における名詞句の翻訳処理、特に連体修飾された名詞句に対する処理を発表する。形式名詞、副詞的名詞、用言との間に特別の関係を持ちうる名詞、等、これ等を総称して特殊名詞と呼ぶことにすると、被修飾名詞がこれら特殊名詞である時の統一的な翻訳方式はこれまで困難と見られていたが、今回それを可能にする方法が実現できたので、その方法、及び、中心となるテーブルのデザイン、及びそれ以外の一般的な連体修飾句に対する翻訳方法につき発表する。

Translation of Japanese Adnominals

Hiroshi Kitamura Shiho Ogino

Tokyo Research Laboratory, IBM Research
IBM Japan, Ltd. No. 36 Kowa building
5-19, Sanbancho, Chiyoda-ku
Tokyo, 102, JAPAN

We have been developing a Japanese to English machine translation system for the last several years. Recently we developed a method to transfer Japanese adnominals. When a Japanese sentence is adnominalized by some special important nouns, these nouns are translated into English adverbials, prepositional phrases, conjunctions, auxiliary verbs, or other expressions. We listed and classified special nouns, clarified the conditions that determine the corresponding English phrase types, and developed a table to process these nouns and conditions uniformly. We now give details of the table and a process algorithm for transferring adnominals.

1 はじめに

日本語から英語への機械翻訳において名詞の持つ構文上の働きの日英両言語での差異をどう処理するかは大きな課題である。日本語の名詞には動詞句に代表される用言句を、連体修飾により、名詞句に変える働きがあり連体修飾句を作る。この語法は、日本語では連用中止法と並んで単文をつないで普通の長い文を作る重要な方法であり一般文で多用されている。このため連体修飾句の処理は日本語から英語への機械翻訳において極めて重要である。

連体修飾句を作る名詞の機能は、英語では典型的には関係代名詞(relative-pronoun)がこれに対応する機能である。しかし日本語の形式名詞を含む重要名詞群が連体修飾されている時、英語では関係代名詞句、関係副詞句、前置詞句、同格のthat句、助動詞(auxiliary verb)、接続詞(conjunction)、前置詞句(prepositional phrase)等に訳されなければならない場合が数多くあり、この訳し分けは困難なことが多かった。

(例)

水が滴る音 the sound of water dripping
彼が死んだ(という)ニュース the news that he died
雨が降った場合私は行かない If it rains, I shall not go.
日本に行く場合私は飛行機で行く When I go to Japan, I will go by airplane.
彼の話し方は少しおかしい His way of speaking is a little funny.
大切なのは彼に話すことだ Important thing is that (you) should talk to him.
いま日本人は平均して1年に1万円以上も援助として払っている勘定だ。

(JETS-output)

Japanese is considered to pay (dir-obj) on the average now in a year more than 10000 yen as assistance.

2 特殊名詞

機械翻訳の立場からみれば、関係代名詞を用いて翻訳できない用法を持つ連体修飾句の被修飾名詞を特殊とみなすことが実用的である。この立場から、言語学の本を探すと、[1]に詳しい記述がされている。Martinはそこで、連体修飾句V Nにおいて名詞Nが内包している(Vからextrudeされない)epithemeに基づき、特殊名詞を3分類した。

summational

状況 事実 報告 経験 例 実行 表情 類似 必要 態度 価値 計画 希望 考え 夢 感覚
常識 象徴 etc

「という」が入っても省略されてもよい

resultative

状態 産物 表出 認識(匂い 景色 音 etc) 印象 反応(喜び etc) etc

transitional(relational)

場所 近く 辺り 下 左側 角 etc 「ときの」「ところの」が省略された
時間 今日 昨年 前 後 翌日 etc

理由 原因 目的 etc 「ための」が省略された

範囲 程度 etc 「ほどの」が省略された

形式名詞はpostadnominalとして別個に各語別に詳しく扱われている

挙げ句 間 場合 他 段 位 限り 方 一方向き 際 様 毎 以外 以上 etc

もの ところ とき よう まま はず つもり ため

一方、この問題で、これまでに発表された最も網羅的な機械翻訳からの研究は、Muプロジェクトのもので、大略以下のようなものであった。[2]（原文献では理めこみ文の処理となっている。）

連体修飾のタイプを4分類する。

(a) type 1

(イ) 修飾用言句が必須である格を空にしている。

(ロ) 被修飾名詞が原因、目的、動機、条件、方法、等の特殊名詞であっても、自由格を埋めると考えられるとき。

(b) type 2

(イ) 用言句の主格の名詞の意味属性が部分、属性、動作であり、用言が状態性の意味属性を持っている。

(c) type 3

(イ) 「という」、「などの」等の表現が付随している。

(ロ) 被修飾名詞が、こと、場合、点、必要性、状況、事実、方法、等の特殊名詞である。

(d) type 4

変形ルールを個別に書くことで訳す。用言と被修飾名詞が意味的な関係を持つ。

(イ) ～を用いた方法、行為、生産物、

～を考慮した方法、行為、測度

のようなパターンに当てはまる。

(ロ) 被修飾名詞が、音、匂い、結果、等の特殊名詞である。

これらに対し具体的にどう変換処理、生成処理をするかの記述はされていないが、一般的に考えれば、type1はrelative clauseとして、type2はgenitive relative clauseとして、type3は同格that clauseとして、変換処理では訳し、生成ないしは、変換後処理で[2]にあるようないい訳を作ることが1つの方法であると思われる。我々はtype1,2に関しては、このように処理し、type3,4に関しては辞書（テーブル）を中心とした処理をおこなっている。

3 J E T Sにおける変換処理

J E T Sでは深層格を一切扱っておらず、すべてを表層の語と構造の解析結果の分析のみで変換処理している。この理由の一端は[3]で発表した。本題の連体修飾句の処理においても、特殊名詞をプログラムコーディングでなく、辞書（テーブル）情報のみで処理し、一般名詞についても用言の必須格情報と名詞の荒い意味属性だけで連体修飾句の翻訳が可能となる処理法の開発を行ってきた。細かい条件は繁雑すぎてすべては書ききれないが、

大要は以下のようである。

連体修飾句処理ロジック : 入力データ : 用言句B → 名詞A

「名詞Aは特殊名詞テーブルに登録されていない」とき

(A) 関係代名詞処理

(1) 用言句Bは形容詞、形容動詞文でBに空格があり、しかもga-格がない

(2) 動詞句Bに空格があるか、またはBの英訳語に省略不可のcase-frameがあって埋まっていない。

このときさらに次のいずれかの条件が満たされるとき

(2-1) : 用言句Bが受身形でなく主格以外の空格があるか、又は受身形で主格の空格がある。

(2-2) : 用言句Bで主格だけが空格であり名詞Aの意味属性が人/組織である

(B) 所有格関係代名詞処理

形容詞句、形容動詞句Bに空格がないか、(推定された格も含めて) ga-格がある

(C) 同格処理 : (Bに+subjectのときN that S, -subjectのときN of Ving..)

(1) 動詞句Bに空格がない

(2) 「という」相当表現がBの末尾にある

(3) (A-2) が満たされて、かつ関係代名詞処理にならなかったもの

「名詞Aは特殊名詞テーブルに登録されていた」とき

(D) 特殊名詞テーブルの内容に従い処理される

但し、名詞Aの使われ方が名詞用法であり、かつ用言句Bに主格以外の空格がある時は、関係代名詞処理とする

(注) (C-3) の条件の時は特殊名詞として登録するかどうかの判断を人間が後にする必要がある。

(注) これらはheuristic rule である。機械翻訳では完全な語のリストを得るのは何時の時点でも不可能であり、テーブルに未登録であっても、不足した情報の下での最善の結果を出すために考えだされた。また日本語では主格が高頻度で省略され、その推定を正確に機械でするのは困難であり、誤って推定した時は、大きく誤った英語を出力してしまう。これを避けるために考えだされた規則も含まれている。

4 特殊名詞テーブル

現在我々は、数百の日本語名詞につき、登録をしており、拡張中であるが、これらの作成にあたっては、インターグループ社の側島康博氏をはじめとする多くの同社の方の提案、協力があつたことを特記しておく。

(a) 特殊名詞の定義

- (イ) 連体修飾された名詞が関係代名詞を用いて英訳できない用法があるもの。
- (ロ) 同格表現をとることができる名詞

(b) 特殊名詞の用法の分類

- (イ) 名詞用法：当連体修飾句が名詞句として文中に現れている

判別基準

格助詞（が、を、に、と、より、から、まで、で、へ）

係助詞（は、も）

副助詞（さえ、ばかり、のみ、すら、しか、こそ、だけ、ぐらい、など）

又はそれらの相当表現と共に使用されており、その名詞を支配する親の用言句においてこの名詞節が格要素になっているもの

(例)

彼が会議に出席する場合を考慮した。

社長が欠席した例はない。

所長が述べる意見ばかり採用される。

- (ロ) 連用用法：当連体修飾句が連用修飾句として文中に現れている

判別基準

この名詞節が支配する親の用言句において格要素でなく連用修飾句と考えられるもの。

格要素（に、で）を伴う時も伴わない時も有るが、主に時間、空間に関係した語である。これらの名詞は日本語の辞書においてマークされる。

(例)

彼は会議に出る場合十分に準備をする。

父が死んだ後彼が生まれた。

彼が来た翌日事件が起きた。

- (ハ) 述語用法：当名詞が述語になっている

判別基準

名詞節が言い切りの助動詞、またはそれに相当する次の表現を含む用法である。

だ、で(す)、であ(る)、でな(い)、かもしれ(な(い))、ぬ、ず)、らし(い)、
だろう、であろう、でしょう、だった、なのだ、なのです、なのに、なので、
どころか、がてら、かとばかり、のついでに、のことだ、といえる、というが
とはいえ、、、、等

(例)

彼が述べた考えだ。

彼がやった感触だ。

首相と親しい間柄かもしれない。

(c) テーブルの項目

各特殊名詞毎に以下の情報を持たせる

用法区分 : 名詞節の用法=名詞用法、連用用法、述語用法

用言接続条件 : 用言句Bの文節末の形

=連用形接続、infinitive形接続(終止連体形)、
finite形接続(完了の助動詞「た」に接続)

用言主格 : あり/なし

英訳語 : 名詞の英訳語

品詞 : 英訳語の品詞

名詞句前置詞: 名詞節につく前置詞

complementizer : verb-phraseを親のclauseと結びつけるための英語

subject-form : verb-phraseのsubjectがとる人称形

=nominative, accusative, genitive

verb-form : verb-phraseのverbがとるべき形

=finite, infinitive, ing, base それにtense

modal : verb-phraseに付け加えられるmodal

=may, must, can, would, should, , , , ,

(注) これらは同一語につき下線部付き欄をkeyとして複数登録される。また処理時には正確に条件にあてはまらなくても、適宜条件を弱めて、検索される。

次に典型的な用例につき、テーブルの内容がどうなっているか、あげてみる。なおこれ等は、transfer処理のみを書いており、この後の英文生成については[4]を参照されたい。

(例)

(イ) 私は食べる以外は必要でない。

以外 : 名詞用法、infinitive形接続、主格なし

英訳語=other, 品詞=noun, complementizer=than, verb-form=ing

必要でない need, not

私は 以外wo I(subject) other(direct-object), comp=than
食べる eat, verb-form=ing

-----> I do not need other than eating.

(ロ) 彼が来た翌日事件が起きた。

翌日 : 連用用法、finite形接続、主格あり

英訳語=the day, 品詞=noun-phrase, 前置詞=on, complementizer=after,
subject-form=nominative, verb-form=finitive, +past

起きた happen, +past

翌日 事件が the day, prep=on, nom=after event(subject)

来た arrive, verb-form=finitive, +past

彼が he(subject), nominative

-----> An event happened on the day after he arrived.

(ハ) 彼が言った感触だ。

感触：述語用法、finite形接続

subject-form=nominative, verb-form=+present-perfect, modal=may

感触だ (nil)

言った say, +pres-perfect, modal=may

彼が he(subject), nominative

-----> He may have said.

5 結論

機械翻訳でも「図形分析のキーは何か」というと、図形の本質が分るような非常に特殊な言語をつくって、それでプログラミングすることです。・・・10年かかると思われていたことが、1年でできるようになることがあるんです。」という西氏の意見[5]はあてはまると思われ、その立場からみれば、今回の方法は範囲が連体修飾文の処理と限定されているが、多少なりとも意味があると思う。例えば、「音」に対する観察から、次のように処理したくなったとする。

(1) 水が滴る音 --- sound of water dripping

(2) 風船が破裂した音 --- sound that a balloon burst

(3) 私が聞いた音 --- sound that I heard

これにたいし「音」のテーブルに次の内容を登録すればよい

(1) 名詞用法、infinitive形接続、主格あり、"sound", verb-form=ing,
subject-form=acc

(2) 名詞用法、finite形接続、主格あり、"sound", verb-form=finite,
complimentizer="that", subject-form=nom

(3) 登録不要

もう1つ例をあげると、「規則」に対する処理要求(観察)が次のようであったとする。

(1) チェックアウトする規則を／が／、 --- reguration to check out

(2) 客がチェックアウトする規則を／、 -- reguration that a visitor checks out

(3) 客がチェックアウトする規則だ-- a visitor is obligated to check out

「規則」のテーブル

(1) 名詞用法、infinitive形接続、主格なし、"reguration", verb-form=infinitive

(2) 名詞用法、infinitive形接続、主格あり、"reguration", verb-form=finite,
complimentizer="that"

(3) 述語用法、infinitive形接続、"obligated"(adj), verb-form=infinitive

参考文献

[1] Samuel Martin, 「A Reference Grammar of Japanese」 Yale Univ. Press, 1975

[2] 長尾 真、辻井 潤一「機械翻訳における訳語選択と構造変換過程」情報処理vol-26, 1985

- 辻井、長尾“日英翻訳過程での処理とその翻訳結果への反映”自然言語処理47-10, 1985
- [3]北村 博“日本語語彙の代数構造と機械翻訳への応用”自然言語処理71-4, 1989
- [4]Johnson,D,E,“A Transportable English Generator for Machine Translation”,
IBM TRL Research Report TR87-1023,1988
- [5]横井俊夫、神田泰典、坂村健、西和彦“日本型プロジェクト論、プロデューサ論”
A I ジャーナル No-12, 1987