6Y-02 形容詞の修飾関係を使用した日本語名詞の訳語選択

竹信伸介　藤井龍　池原悟　村上仁一
鳥取大学工学部知能情報工学科

1 はじめに

機械翻訳を行う際の問題点として、多義を持つ単語の訳語選択がある。ある単語に対して複数の訳語が存在する場合，正しい翻訳結果を得るためには，その中から適切な訳語を選択する必要がある。これまでの訳語選択の研究としては，格の情報と意味属性を用いた方法 [1] などが提案されているが，名詞は動詞に比べて種類が多く訳し分けが困難である。そこで本研究では，対象を形容詞に修飾された名詞に限定し，形容詞と名詞の意味的関係から名詞の訳語を決定する方法を考案し，その精度を評価する。

2 「形容詞+名詞」を含む文の分類

訳語選択を行う場合，名詞と名詞を修飾する言葉との意味的関係から、名詞の訳語候補を絞り込める可能性がある。また ,「形容詞+名詞」という組の日本語に英語が対応するパターンも複数あると考えられるので、本研究では「形容詞+名詞」を含む文を，日本語と英語の対応関係から 5 つに分類する。分類の対象としてアンカー和英辞典 46108 文中より ,「形容詞+名詞」の形になっている 1054 文を使用する。

(1) 形容詞と名詞それぞれに英語訳語が対応 (680 文)
例:選手たちは全員 青いユニフォーム を着ていた
The players all wore blue uniforms.
(2) 形容詞と名詞の組に対して英語訳語が対応 (141 文)
例:あいつはまだ 青臭い男 だよ
That guy's still green.
(3) 当該表現の前後を含む表現に英語訳語が対応 (165 文)
例:少女は 赤い着物を着た 人形を抱いていた
The girl carried a doll in red in her arms.
(4) 慣用表現 (28 文)
例:痛い目に合う
have a bitter experience
(5) (1) ~ (4) に分類されず文全体で意訳 (40 文)
例:セールスマンはうまい言葉で学生を釣って英会話のカセットテープを買わせた
The salesman talked the student into buying cassette tapes for English practice.

3 翻訳規則の作成

3.1 研究の範囲

2 章の分類のうち，使用頻度の高い分類 (1) の「形容詞 + 名詞」を含む表現を扱い、アンカー和英辞典の例文で頻度の高い形容詞 (216 種類中 19 種類) の場合の翻訳規則を作成する。

一般名詞意味属性体系

翻訳規則は、一般名詞意味属性を使用して作成される。一般名詞意味属性体系とは、対象と単語の関係を、対象の捉え方に着目して分類したもので、図 1 のように 12 段の木構造 (属性数 約 2700) になっている。また、上位属性は下位属性を内包する。

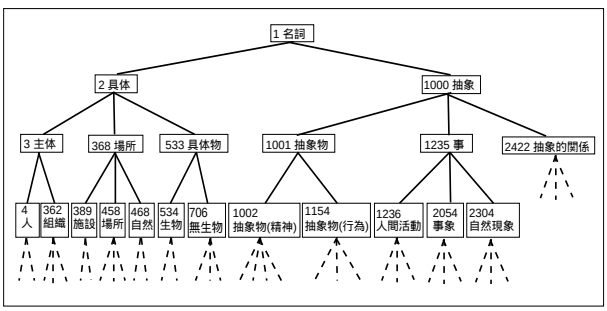

図 1. 一般名詞意味属性体系 (上位部位)

3.2 翻訳規則

2 章で分類 (1) に定義した翻訳規則を述べる。
アンカー和英辞典の例文に出てくる、形容詞の後にくる名詞の意味属性と、名詞の意味属性ごとの訳語を調べて，それぞれ対応表を作成する。表 1 及び表 2 はその一部である。

Translation of Japanese Nouns modified by Adjective into English.
Shinsuke TAKENOBU,Ryu FUJII,Satoru IKEHARA,Jin'ichi MURAKAMI
Tottori University

表 1. 形容詞の後にくる名詞の意味属性

形容詞	名詞の意味属性
古い	8都、１２５友人、１５４国民、４２０橋、４３９工場 ４４７居住施設(その他)、８６５家屋(本体)、８９４家財道具、灯火等 ９０４容器、９０８袋、９１９帳、９７０通信機器 ９８９乗り物(本体(移動(水圏)))、１００２抽象物(精神) １０５６歌謡、１０６１映画、映像、１０７６諺、１１１９本(内容) １１６３法律、１２１７風俗、１２２１慣例、２３６５雪、２５８２列

表 2. 名詞の意味属性ごとの訳語

名詞	意味属性	名詞の訳語
家	７３家族 ３８７家庭 ４４７居住施設 （その他） ８６５家屋（本体）	home home house,dwelling,residence house
噂	１０４３説話 １１４９噂 １５５０伝達	report rumor,gossip hearsay,communication

表 1 は、形容詞「古い」の後にくる名詞を例文から抽出し、形態素解析結果から求められた名詞の意味属性を列挙したものである。また、表 2 は名詞を意味属性ごとに分け、それぞれ対応する訳語を付けたものである。例えば、形容詞「古い」の後に名詞「家」がきて、その意味属性が形態素解析時に「865 家屋 (本体)」となった場合、「house」と訳されることを示す。

4 評価

3 章で作成した分類 (1) の翻訳規則の精度を、使用頻度の高い 19 種類の形容詞を含む、①分類に使用した 343 文，②機能試験文集 [3] と基本語用例辞典 [4] よりランダムに抽出した 79 文で評価する。評価結果の分類は、4 つに分ける。

◎ 選択された候補が正解のみの場合
○ 選択された候補の中に正解を含む場合
△ 候補の中に正解がない場合
＊名詞を登録すればルールが適用できる場合

○ 、△ については、複数の結果が出力される。また、正解とは、対訳例と作成した規則での結果が一致したものとする。

(◎ の例) 聴衆は 美しい旋律 に酔いしれた
The audience was enraptured by the beautiful melody.

(○ の例) 父は 厚い本 を枕の代用にした
実験結果：
My father used a thick book as a pillow.
My father used a thick magazine as a pillow.
対訳例：
My father used a thick book as a pillow.

(△ の例) わたしは 高い石段 を 1 段ずつ登った
実験結果：
I went up the high stone steps step by step.
対訳例：
I went up the high stone stairway step by step.

(＊の例)1 通の 厚い封筒 を受け取った
I received a fat envelope.

5 結果

4 章での評価結果を表 3 に示す。また，デフォルト訳語 (訳語として辞書の先頭にきている単語) で評価した結果も示す。括弧内はそれぞれの度数である。選択された候補に正解を含む文 (◎ 又は○) は，アンカー和英辞典より抽出した文で約 85%，機能試験文集と基本語用例辞典より抽出した文においては約 30%であった。

表 3. 翻訳結果

評価	①和英辞典(標本内)		②機能試験文集 基本語用例辞典(標本外)	
	翻訳規則	デフォルト訳語	翻訳規則	デフォルト訳語
◎	21.3%(73)	15.5%(53)	5.1%(4)	6.3%(2/32)
○	65.3%(224)	41.4%(142)	27.8%(22)	43.8%(14/32)
△	13.4%(46)	43.1%(148)	7.6% (6)	50.0%(16/32)
＊	0 (0)	0 (0)	59.5%(47)	――

6 考察

候補の中に正解を含む場合 (◎ 又は○) について，形容詞と名詞の意味的関係を使用した結果，名詞の平均多義 (複数の意味や用法) 数は，①では 6.85 から 4.19 に減少し，②では 9.42 から 5.25 に減少した。また，②では，“名詞を登録すればルールが適用できる場合” の文について表 1 に名詞の意味属性を登録すれば，80%程度の名詞で正しい訳語が選択されると推定される。

7 おわりに

本研究において，名詞と名詞を修飾する形容詞との意味的関係から訳語候補を平均で 4 から 5 個程度に絞って選択できることが分かった。今後は，他の用言 (動詞，形容動詞) と名詞の意味的関係についても調べ，訳語選択を行う際に考慮することが必要だと考えられる。

参考文献

[1] 北野，萩野：日英翻訳における連体修飾句の訳し分け，情報処理学会研究報告書，Vol.90,No5,75-10(1990)
[2] 池原，宮崎，白井，横尾，中岩，小倉，大山，林：日本語語彙体系 1．意味体系，岩波書店 (1997)
[3] 池原：日英機械翻訳機能試験文集,NTT-CS 研 (1994)
[4] 文化庁編：外国人のための基本語用例辞典, 大蔵省印刷局 (1994)