

日本語の表層格構造の記述について(2)

5G-4

—— 拡張格構造の扱い ——

鈴木 克志、丸山 冬樹、太細 孝

三菱電機 情報電子研究所 知識処理開発部

1. はじめに

前報告(文献1)では、日本語の述語に対して表層格構造を記述した結果について述べた。そこでは、表層格の記述に関するいくつかの問題点(グラベル付与のゆれなど)について考察した。

本報告では、日本語の自立語と形式動詞などの共起に対する格記述や知覚動詞の格記述に対する問題を解決するために導入された拡張格構造の概念について述べる。

2. 格記述に関する問題

2-1. 形式用言

自立語と組合わされてひとつの意味を表わす以下のような表現を、日本語の形式用言と呼ぶ。

- ①一緒に+なる、一緒に+する [名詞+形式動詞]
- ②確証が+ある、確証を+持つ [名詞+形式動詞]
- ③基盤と+する [名詞+形式動詞]
- ④甘味が+ある、甘味の+ある～ [名詞+形式動詞]
- 甘味が+多い [名詞+形式形容詞]
- ⑤一杯である [名詞+ 助動詞]

これらの表現について表層格構造を定義しようとした場合、次の2つの問題がある。

1つは、構文的には自立語と形式用言とを別々の語として扱い、それら2語の組合わせに対して格構造を定義する必要があることである。これは、「確証が彼にはある」というように2語の間に他の要素が入り得るからである。

2つめの問題は、格構造を考えた場合、表層の多様性が普通の述語の格構造の場合よりも多くなることである。

例えば、図1のような表現を考えてみる。

- ①「EはOに自信がある」 : 基本型
- ②「EはOの自信がある」 : 「の」の修飾
- ③「EはOする自信がある」 : 文の修飾
- ④「EはOに大きな自信がある」 : 形容詞の修飾
- ⑤「EはCより自信がある」 : 比較対象要素の修飾
- ⑥「EはCより3倍の自信がある」 : 数量表現の修飾

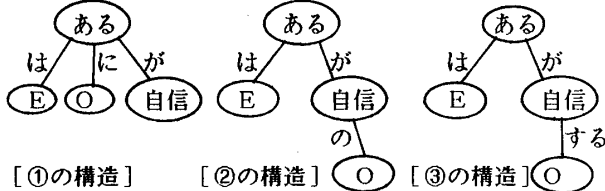


図1. 表層格構造の多様性

①が基本形である。基本形においては「自信が+ある」

に対して「AがOに」という表層格構造を与えてやれば一般の述語の表層格構造記述と同じに扱える。

しかし、②においては「自信」に対して「Oの」が修飾しており、表層格助詞が変形している。さらに③や④のように、埋込み文や形容詞句が名詞を修飾することもある。これらは「自立語と形式用言を合せて1概念として格構造を定義した場合に、表層構文において表層格が要素の埋込みの異なるレベルに散在してしまう」という問題である。

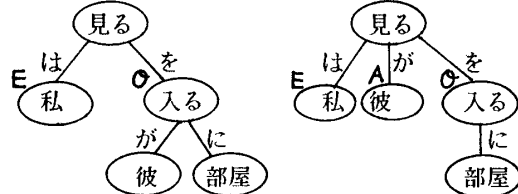
また、⑤の比較要素の修飾や、⑥の数量表現の修飾をどう扱うかという問題がある。しかし、これらの修飾要素は表層格構造の任意成分であり、一般の述語の文においても『彼は彼女より3倍本を読む』のように同じ問題があるので、ここではこれ以上触れない。

2-2. 知覚動詞

「見る」、「聞く」などの知覚動詞を代表とするある種の述語は事象を表層格としてとる。このとき、少なくとも2つの問題がある。

まず1つは、表層格構造をどのように定義すべきかということである。図2に2種類の定義法を示す。

(例文)「私は彼が部屋に入るのを見た」



[定義1: EがOを] [定義2: EはAがOスルを]

図2. 知覚動詞の表層格構造

定義1は日本語の構文感覚からすれば自然である。一方定義2は英語的である(I saw him enter the room.)。ここでは日本語固有の表層格構造を考えているので、定義1の方を採用する。しかし、このとき埋込み文の主語「彼」が主文の目的語に対応することを定義する必要がある。

2番目の問題として、埋込まれている事象の様相属性の相違により構造が変わることがある。図3に例を示す。

- ①「Eが(Oをシタイ)と思う。」
- ②「Eが(AにOをシテホシイ)と思う。」
- ③「Eが(Aがスレバヨイ)と思う。」

図3. 埋込まれている事象の様相属性

①では埋込み文の主語は主文と同一であるが、②では異なり、③では両方の場合があり得る。しかも英語への翻訳を考えると①、②はwant、③はhope、というように、対応する訳語が変わってくる。

文献3では、はじめの問題に対してはダミー格というものを導入して解決を試みている。しかし、ここでは、2番目の様相属性の制約の問題や2-1に述べた形式用言の問題など、通常の格記述で指定される制約よりも広い範囲の制約をも含めて、ある統一的な枠組み（拡張格構造モデル）の中で解決することにした。

3. 拡張格構造

3-1. 枠組み

拡張格構造の特徴は、格ラベル(AやO)と格標識(助詞の指定)から成る普通の表層格構造に加えて、上記に述べたような様々な問題を解決するための種々の制約を、統一的に次の枠組みで扱えることにある。

拡張格構造記述:=格構造の集合
格構造:=格の集合+格構造制約
格:=格ラベル+格制約

図4. 拡張格構造

3-2. 格構造制約

格構造制約は、表層格構造全体に関する制約を与える。

①概念記述 表層格構造に対応して中間言語上の概念を定義する。たとえば、図3の①~③にはそれぞれ、want1、want2、hope1の概念に対応させることができる。

②係り先との共起指定 どの形式用言と共起するかを指定し、共起に対してまとめて1つの概念を与えることができる。形式を図5に示す。たとえば、「自信がある」の表現には、「自信」の格構造制約にformal(aru, agent)を与える。また、一般の単語との共起も指定できる。

formal(Type, Case)
Type: aru, nai, yoi, warui, ooi, sukunai, suru, shinai, naru, naranai, dearu, ...
Case: 自立語と形式用言の間の表層格を指定する。
agent, object, goal, locative, ...

図5. 形式用言の指定

③態指定 普通の格構造は自立語の能動形に対して定義し、助動詞の承接による使役や受身化などには格標識の変形に対処するが、よく用いられる使役形や受身形に対して直接格構造を定義することもできる。

④品詞・活用・読み指定 構文解析および生成用の辞書としても使えるように、品詞、活用型、読みが指定できる。

⑤アスペクト指定 アスペクトの相違により概念記述が異なる場合(ex. 「覚える」と「覚えている」)や表層格構造が異なってくる場合に、この指定が有効となる。

3-3. 格制約

格制約はひとつの格に対する制約を与える。

①格ラベル指定 Agent, Objectなどの表層格ラベルを指定する。なお我々は、表層格と深層格との混同を避ける目的であえて表層格と呼んでいる(文献1)。同等の格を深層格と呼んでいる他のシステムもあるので注意する。

②格標識制約 係りの助詞を指定する。分類記号の指定により派生助詞が自動的に指定できるほか直接に助詞も指定できる。分類は14種類である。次に示すように助詞の多様性を吸収するためにはこれくらい必要である。

[分類記号]	[派生助詞]
NI	に、へ
NI-WO	に、へ、を
NI-TO	に、へ、と

③意味制約 12種類の大部分類意味マーカ、61種類の中分類意味マーカ、約700の下位分類意味マーカ、およびそれらの集合が指定できる。

④共起指定 格要素として共起する単語を直接指定可能。

⑤埋込み文の制約 2-2で述べたように、格要素に埋込み文をとる場合そのアスペクトと様相表現が指定できる。

⑥中間言語への写像 埋込み文中の格が中間言語上で主文のどの格に対応するかを指定するraise格指定により、中間言語上で表層と異なる意味構造への写像が可能である。

⑦拡張格指定 『Oの自信がある』、『Oする自信がある』のように、2-1で述べた埋込みの異なるレベルにある連体修飾の格を指定することができる。

3-4. 翻訳への応用

拡張格構造を用いて2.で述べた問題を扱うことが可能となった。以下では、翻訳処理の概要を図6に示す。

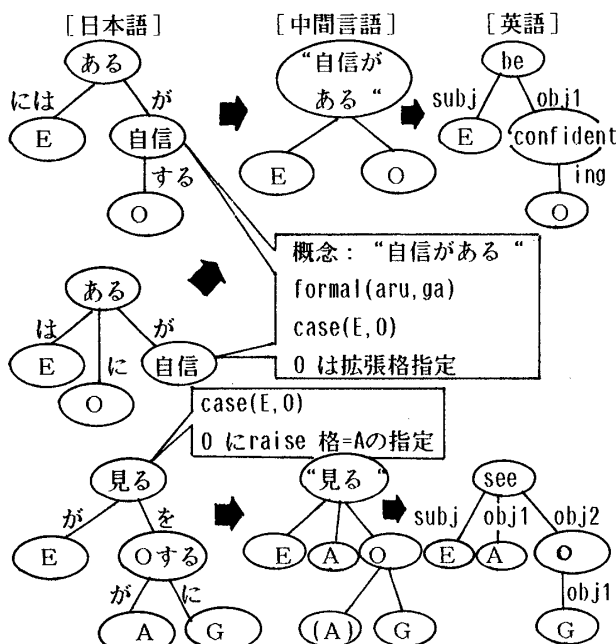


図6. 翻訳処理の概要

4. おわりに

拡張格構造の特徴は以下の諸点である。

- ①共起に対して格構造がまとめて定義できる。
- ②表層構造において異なるレベルにある格要素への制約が記述できる。
- ③中間言語表現への写像が記述できる。
- ④体言にも記述できる。
- ⑤以上の機能がひとつの記述形式に集約されている。

参考文献

- (1) 鈴木、他：日本語の表層格構造の記述について、情報処理学会第38回全国大会pp. 317-318(1989).
- (2) 計算機用日本語基本動詞辞書IPAL(Basic Verb)辞書編、情報処理振興事業協会技術センタ(1987).
- (3) 野村、他：機械翻訳システムPIVOIの日本語格フレームモデル、情報処理学会第38回全国大会(1989).