

結合価構造に基づく日本語二重主格形容詞構文の解析

奥 雅 博†

本論文では、結合価構造に基づく二重主格形容詞構文の解析手法について述べる。多くの言語において単文は1つの主格を有する。これに対して、日本語述語は単文中において2つの表層的な主格を有することができる。このような単文のことを本論文では二重主格構文と呼び、特に述語が形容詞または形容動詞のものを二重主格形容詞構文と呼ぶ。本論文では、二重主格形容詞構文中の副助詞「ハ」の機能に着目して、日本語二重主格形容詞構文を、副助詞「ハ」が連用型の格助詞を代行するタイプ1、副助詞「ハ」が連体型の格助詞を代行するタイプ2、格助詞「ガ」が対象格を表し、副助詞「ハ」が格助詞「ガ」を代行するタイプ3、副助詞「ハ」を持つ修飾語句が時間を表し、副詞句的に働くタイプ4、の4タイプに分類する。次に、解析対象の二重主格形容詞構文がどのタイプに属するかを、述語の持つ結合価構造に基づいて判定し、判定結果に応じた適切な解析を行う手法を提案する。特に、タイプ2の解析手法については従来、工学的に述べられたことはほとんどなかった。これに対して本論文では、タイプ2の二重主格形容詞構文を解析するために、仮想的な結合価要素を導入することを提案する。そして、この仮想的な結合価要素に副助詞「ハ」を持つ修飾語句を対応付けることによって、タイプ2の二重主格形容詞構文に対する正しい結合価構造が得られることを示す。さらに、本手法の有効性を確認するためにに行った評価実験について述べる。新聞記事約1カ月分を対象に行った評価実験の結果、解析の対象とした二重主格形容詞構文の約94%を本手法によって正しく解析できることが分かった。

Analyzing Japanese Double-nominal-case Construction Having an Adjective Predicate Based on the Valency Structure

MASAHIRO OKU†

This paper describes a method for analyzing Japanese double-nominal-case construction having an adjective predicate based on the valency structure. A simple sentence usually has only one subjective case in most languages. However, many Japanese predicates can dominate two surface subjective cases within a simple sentence. Such sentence structure is called the double-nominal-case construction in this paper. This paper classifies the Japanese double-nominal-case construction into four types and describes problems arising when analyzing these types using ordinary Japanese construction approaches. This paper proposes a method for analyzing a Japanese double-nominal-case construction having an adjective predicate in order to overcome the problems described. By applying this method to Japanese sentence analysis in Japanese-to-English machine translation systems, translation accuracy can be improved because this method can analyze correctly the double-nominal-case construction having an adjective predicate.

1. はじめに

日本文解析においては、用言を述語とした関数構造として文構造をとらえる依存文法あるいは格文法的な考え方が広く用いられている。筆者らは、述語の関数構造という考えを基にした文法理論のうち、格構造よりも表層の情報を重視する結合価構造⁵⁾に着目し、これに基づく日本文解析手法について検討を進めている^{1),2),11)}。

二重主格構文(～ハ～ガ述語)は、JEIDAの「機械翻訳システム評価基準」⁶⁾において特殊構造表現の1つとしてあげられているように、機械翻訳などの実際の自然言語処理システムにおいて解析すべき特殊構文の1つである。しかし、従来の日本文解析では、ヒューリスティックルール「1文1格の原則」の制約により、格助詞「ガ」を持つ修飾語句を主格としてとらえ、副助詞「ハ」を持つ修飾語句を主格以外として解釈する手法がとられていた。このため、二重主格構文のうち、副助詞「ハ」が連体型の格助詞「ノ」を代行している場合を正しく扱うことができなかった。

一方、形容詞、形容動詞を述語とする二重主格形容

† NTTグループ企業本部新事業開発室
New Business Development Section, NTT Affiliated
Business Headquarters

詞構文(～ハ～ガ形容詞)に関しては、日本語の統語的、意味的な理論を構築するうえで重要であるため、多くの日本語学者の間で議論が行われ、それぞれの立場からその構造に関する説明が行われている^{4),7),8)}。これらの研究の中で三上⁸⁾や石上⁴⁾は、「～ハ」で表示される名詞句を副助詞「ハ」が代行している格助詞で表される格関係に対応させることにより二重主格形容詞構文の説明を行っている。この考え方は、格助詞という表層的な情報に着目している点、述語との関係を重視する、すなわち述語の関数構造として文構造をとらえるという点で筆者らが進めている結合価構造に基づく日本文解析手法と親和性が良いと考えられる。しかし、計算言語学的あるいは工学的観点から二重主格形容詞構文を扱った研究はほとんどなかった⁹⁾。

本論文では、三上⁸⁾や石上⁴⁾の研究を中心に従来の研究を工学的観点から見直し、二重主格形容詞構文の結合価構造を明らかにする手法について提案する。

本論文の構成は次のとおりである。まず2章で本論文で用いる用語について説明する。3章では筆者らが進めてきた結合価構造に基づく日本文解析について概観する。4章では、日本語二重主格形容詞構文を、構文中の副助詞「ハ」の働きに基づいて4つのタイプとその他の5つに分け、それぞれの特徴について述べる。さらに結合価構造を用いて解析可能な4タイプの日本語二重主格形容詞構文を解析する手法について提案する。5章では、新聞記事を対象に行った評価実験について述べ、結果に対して考察を加える。最後に6章において本手法を日英機械翻訳に適用した場合の効果について述べる。

2. 用語の定義

○二重主格形容詞構文

本論文では、2つの表層的な主格を有する構文「～ハ～ガ述語」を二重主格構文と呼ぶ。このうち、形容詞または形容動詞を述語とする構文「～ハ～ガ形容詞(または形容動詞)」を二重主格形容詞構文と呼ぶ。形容動詞は、形態的には形容詞とは異なるが、客観的な性質・状態あるいは主體的な感情・感覚を表すという点で形容詞と変わりはない⁴⁾。このため、以下では形容動詞を形容詞と同等に扱い、形容詞という用語には形容動詞を含むものとする。

○第一主格、第二主格

二重主格構文「～ハ～ガ述語」において、副助詞「ハ」を持つ修飾語句(「～ハ」)に相当する部分を第一主格、格助詞「ガ」を持つ修飾語句(「～ガ」)に相当する部分を第二主格と本論文では便宜的に呼ぶこと

結合価パターン:「紹介する」

N1 (主格)	SR: 【主体】 JR: が、から
N2 (対象格1)	SR: 【主体】 JR: を
N3 (対象格2)	SR: 【主体】 JR: に

凡例:

Ni: 結合価要素のラベル
SR: 主名詞に対する意味的制約
JR: 格助詞に対する制約

図1 結合価パターンの例

Fig. 1 An example of a valency pattern.

とする。たとえば、「太郎は花子が好きだ」において、「太郎は」を第一主格、「花子が」を第二主格と呼ぶ。また、「太郎は次郎も好きだ」においては、「太郎は」を第一主格、「次郎も」を第二主格と呼ぶ。

○結合価構造

文の構造は、述語とこれを修飾する語句(修飾語句)との組合せとしてとらえることができる。このような構造を「結合価構造」と呼ぶ。格構造が深層格というものを設定して述語と修飾語句との関係を表すのに対して、「結合価構造」は、どのような意味属性^{3),14)}を持つ名詞がどのような表層格マーカ(日本語の場合、助詞表現)をともなつて着目している述語と共起するかを表現する。

○意味属性

「意味属性」は名詞をその意味に基づいて階層的に分類したものである。たとえば、単語「男」も単語「女」もともに意味属性【人間】に属する。また、意味属性【人間】はその上位概念として【主体】を持つ^{3),14)}。以下では意味属性を表すのに【】を用いる。

○結合価パターン

「結合価パターン」は述語に対して可能な結合価構造を記述した構造パターンである¹⁾。結合価パターンは述語の用法ごとに定義される。図1は動詞「紹介する」に対する結合価パターンの一例である。図1は、動詞「紹介する」が、主格(N1)、対象格1(N2)、対象格2(N3)をとりうることを示している。

○結合価要素

結合価パターンにおいて、述語を修飾する要素のことを「結合価要素」という。図1に示すように結合価要素は、名詞句の主名詞に対する意味的制約(SR)と格助詞に対する制約(助詞制約:JR)とによって記述する。たとえば、主格を表す結合価要素N1は、意味的制約として【主体】を持ち、助詞制約として「ガ、カラ」を持つ。すなわち、結合価要素N1に対応する修飾語句は、助詞表現として「ガ」または「カラ」を持ち、主名詞の意味属性が【主体】を満足しなければならない。

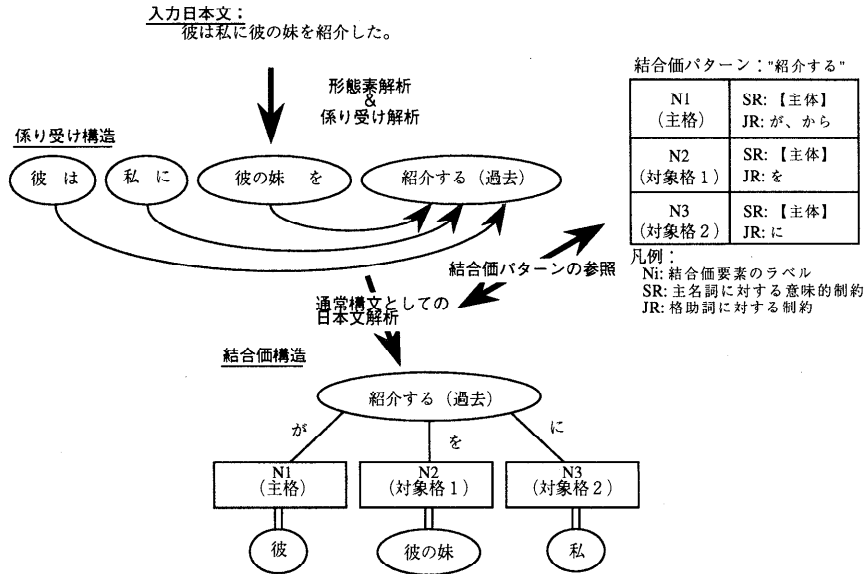


図 2 日本文解析の例

Fig. 2 An example of the Japanese sentence analysis.

3. 結合価構造に基づく日本文解析

結合価構造に基づく日本文解析では、結合価パターンを参照することによって文節間係り受け構造を結合価構造へと変換する。このとき、結合価パターンの制約を満足しない係り受け構造は誤った解釈として棄却される¹⁾。結合価構造は、入力文の述語を修飾する語句（修飾語句）を結合価パターン中の結合価要素に対応付けることによって決定される。図1に示すように、結合価パターンには各述語の持つ各結合価要素に対して、主名詞に対する意味的制約（SR）と格助詞に対する制約（助詞制約：JR）とが記述されている。入力文中の修飾語句が結合価パターン中のある結合価要素の持つ2つの制約（意味的制約と助詞制約）を満足するとき、注目している修飾語句をその結合価要素に対応付ける。日本文解析では、最初に格助詞を持つ修飾語句を結合価要素に対応付けることを試み、その後で対応付けの済んでいない結合価要素に対して副助詞を持つ修飾語句を対応付けることを試みる。これは、格助詞が表層格マーカとして表層上の曖昧さが無いのに対して、副助詞は複数の格助詞に対応しうる（たとえば、副助詞「ハ」は格助詞「ガ」、「ヲ」や「ニ」などを代行しうる）ためである。

図2に入力文「彼は私に彼の妹を紹介した」に対する日本文解析の例を示す。まず、動詞「紹介する」の結合価パターンにおける結合価要素に対して、格助詞「ヲ」、「ニ」を持つ修飾語句（「彼の妹を」と「私

に」）を対応付けることを試みる。それぞれの修飾語句は図2に示すように、それぞれ結合価要素N2、N3が持つ2つの制約（意味的制約と助詞制約）を満足する。したがって、格助詞「ヲ」を持つ修飾語句「彼の妹を」をN2に、格助詞「ニ」を持つ修飾語句「私に」をN3にそれぞれ対応付けることができる。次に、副助詞「ハ」を持つ修飾語句「彼は」を結合価パターン中で対応付けの済んでいない結合価要素に対応付けることを試みる。すでにN2とN3は対応付けが終了しているので、対応付けの済んでいない結合価要素N1との対応付けを試みる。「彼は」における「彼」の意味属性【人間】は意味属性【主体】の下位概念であるので、N1の意味的制約（SR:【主体】）を満足する。また、副助詞「ハ」は格助詞「ガ」を代行しうるので、N1の助詞制約（JR:「ガ」、「カラ」）も満足する。ゆえに、修飾語句「彼は」を結合価要素N1に対応付けることができる。以上の処理によって、入力文「彼は私に彼の妹を紹介した」に対する日本文解析結果として図2に示す結合価構造が得られる。

4. 日本語二重主格形容詞構文の解析

4.1 日本語二重主格形容詞構文の分類

多くの日本語形容詞は2つの表層的な主格を支配し、二重主格形容詞構文を構成する。二重主格形容詞構文にはいくつかのタイプが存在する。そこで本論文では、二重主格形容詞構文における副助詞「ハ」の働きに着目して、二重主格形容詞構文を以下の4つのタ

イプとその他に分類する^{4),8)}。

○タイプ 1

副助詞「ハ」が格助詞「ガ」以外の連用型の格助詞(「ニ」「デ」など)のいずれかを代行する二重主格形容詞構文をタイプ 1 とする(例 1)。このタイプの二重主格構文は動詞を述語とする文にも見られ(例:その計画は彼が実行した)、副助詞「ハ」がどの格助詞を代行しているかを解析するには、述語の持つ結合価構造が重要な手がかりとなる。

[例 1]

- その家は学校に近い ← 副助詞「ハ」が格助詞「ニ」を代行
- キャンプ場は蚊が多い ← 副助詞「ハ」が格助詞「ニ」を代行
- この地域は雨が少ない ← 副助詞「ハ」が格助詞「ニ」を代行
- この計画は見通しがいい ← 副助詞「ハ」が格助詞「ニ」を代行
- 北国は日暮れが早い ← 副助詞「ハ」が格助詞「デ」を代行

○タイプ 2

副助詞「ハ」が連体型の格助詞「ノ」を代行する二重主格形容詞構文をタイプ 2 とする(例 2)。タイプ 2 の場合、第一主格は統語的には述語を修飾しているが、意味的には第二主格を修飾しているものとして解釈しなければならない。このため、タイプ 2 の二重主格形容詞構文の解析では結合価構造の再構成が必要となる。また、タイプ 2 の二重主格形容詞構文を作る形容詞は「ガ」格以外の格を持たない場合が多い。

[例 2]

- 象は鼻が長い ← 「象」が「鼻」を修飾
- 彼は耳が遠い ← 「彼」が「耳」を修飾
- その建物は外観が良い ← 「建物」が「外観」を修飾
- 彼は教師としての経験が浅い ← 「彼」が「経験」を修飾
- 氷は摩擦係数が小さい ← 「氷」が「摩擦係数」を修飾
- 瀬戸内海は波が穏やかだ ← 「瀬戸内海」が「波」を修飾

○タイプ 3

第一主格が実際の主格を表すとともに、第二主格が対象格を表す二重主格形容詞構文をタイプ 3 とする(例 3)。タイプ 3 の二重主格形容詞構文を作る形容詞は、感情を表す情意形容詞に限られ、副助詞「ハ」が格助詞「ガ」を代行している。タイプ 1 と似ているが、

「ハ」→「ガ」だけでなく、「ガ」→「ヲ」の変換を必要とする点が異なる。

[例 3]

- 彼は彼女が好きだ ← 「彼女」が対象格
- 太郎は花子が嫌いだ ← 「花子」が対象格
- 彼はリンゴが欲しかった ← 「リンゴ」が対象格

○タイプ 4

第一主格が時間を表し、副詞句的に働く二重主格形容詞構文をタイプ 4 とする(例 4)。

[例 4]

- 6 月は雨が多い ← 「6 月は」が副詞句的に作用
- 冬はものの乾が悪い ← 「冬は」が副詞句的に作用
- 秋は紅葉がきれいだ ← 「秋は」が副詞句的に作用
- 夏休みは学校が静かだ ← 「夏休みは」が副詞句的に作用

○その他

タイプ 1~4 のいずれにも該当しない二重主格形容詞構文をその他とする。その他に分類される二重主格形容詞構文には以下のような種類がある(例文は付録参照)。

- 第一主格が第二主格以外の結合価要素を修飾
- 第二主格が第一主格の属性を表現
- 比較構文
- 副助詞「ハ」が取り立てを表現(文脈を考慮する必要がある)
- 省略によって表層上の二重主格形容詞構文を構成など。

このようにその他に分類される二重主格形容詞構文は多様であり、これらを結合価構造に基づいて解析しようとする以下のような問題が生じる。

- (1) 準体助詞や形式名詞が副助詞「ハ」や格助詞「ガ」に接続する場合がある。準体助詞や形式名詞は文脈によって指示対象が決まるために、単独で意味属性を持たない。このため、助詞表現と名詞句の持つ意味属性に基づく日本文解析を適用することができない。この場合には、文脈を考慮して指示対象を決定する必要がある。
- (2) 統語的な解析だけでなく、意味的な解析を必要とする場合がある。
- (3) 省略があって、本来その省略部分に係るべき第一主格が述語を修飾している場合がある。この場合、見かけ上は二重主格形容詞構文となっているが、省略を補うなどの語用論的な解析を行わなければならない。

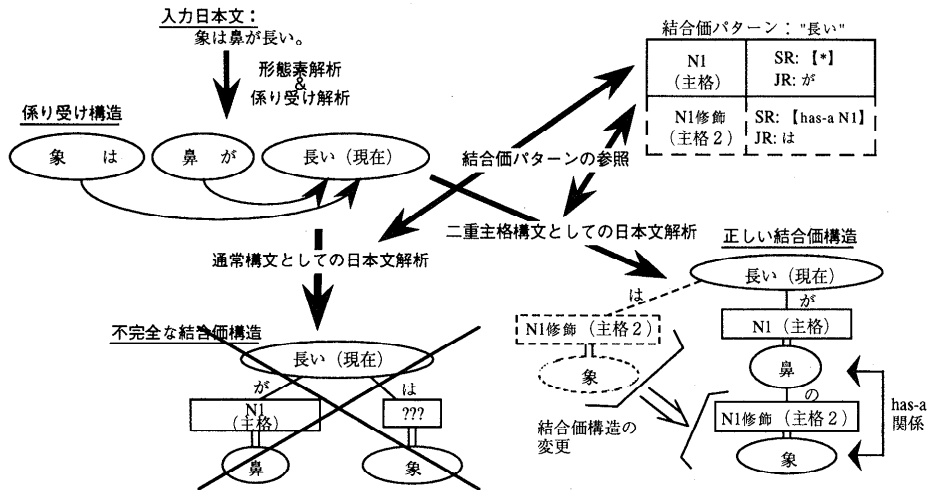


図3 二重主格形容詞構文の解析例(タイプ2)

Fig. 3 An example of analyzing double-nominal-case construction (type-2).

すなわち、その他に分類される二重主格形容詞構文は、表層格マーカである助詞表現と名詞句の持つ意味属性とを制約とした結合価構造に基づく解析では扱いきれない。そこで以下では、タイプ1~4に属する二重主格形容詞構文の解析手法について述べることにする。

4.2 日本語二重主格形容詞構文の解析における問題点

タイプ1は、副助詞「ハ」が連用型の格助詞の代用であるため、3章で述べた通常構文の解析処理により解釈可能である。また、タイプ4は、第一主格をあらかじめ副詞句的に働くものとして取り出し、残りの部分に対して3章で述べた通常構文の解析処理を行うことにより解釈可能である。これに対して、タイプ2, 3の解析には以下のような問題がある。

○タイプ2の解析における問題点

図3にタイプ2に属する二重主格形容詞構文「象は鼻が長い」の解析例を示す。述語「長い」の結合価パターンは格助詞「ガ」を助詞制約とする結合価要素N1だけを持つ(図3右上の結合価パターン中の実線枠内)。3章で述べた日本文解析に従うと、第二主格「鼻が」は結合価要素N1に対応付けられるが、第一主格「象は」がどの結合価要素にも対応付けられずに残ってしまう。すなわち、図3の左下に示すように日本文解析が不完全となる。

次節で述べるような仮想的な結合価要素を導入することで、結合価構造を正しく解析することができる(図3右下参照)。

○タイプ3の解析における問題点

図4にタイプ3に属する二重主格形容詞構文「彼は

彼女が好きだ」の解析例を示す。3章で述べた日本文解析のように格助詞を持つ修飾語句を優先的に解析すると、第二主格「彼女が」が結合価要素N1に対応付けられてしまい、図4の左下に示すような誤った解釈となってしまう。

これを避けるためには、タイプ3に属する二重主格形容詞構文を作る形容詞に対して、あらかじめ助詞の変換(ガ→ヲ, ハ→ガ)を実行した後に結合価要素との対応付けを試みる必要がある(図4右下参照)。

4.3 日本語二重主格形容詞構文の解析手法の提案

本節では、前述の問題点を克服した日本語二重主格形容詞構文の解析手法を提案する。

本論文で提案する手法は3つの処理フェーズを持つ。第1のフェーズでは入力文が二重主格形容詞構文であるか否かを表層情報を基に判定する。第2のフェーズでは入力文が二重主格形容詞構文であるときに、それが上述したどのタイプに属するかを決定する。最後に第3のフェーズにおいて各タイプに応じた解析を行う。提案する二重主格形容詞構文の解析処理フローを図5に示す。

○第1フェーズ: 日本語二重主格形容詞構文の判定

入力文の形容詞述語が、第一主格(副助詞「ハ」を持つ修飾語句)と第二主格(格助詞「ガ」を持つ修飾語句)とを有するとき、その入力文を二重主格形容詞構文と判定する(図5ステップ1)。このとき、副助詞を持つ修飾語句を2つ有する構文(例: 太郎は次郎も好きだ)も、格助詞「ガ」を持つ修飾語句を2つ有する構文(例: 兄のほうが背が高い)も二重主格形容詞

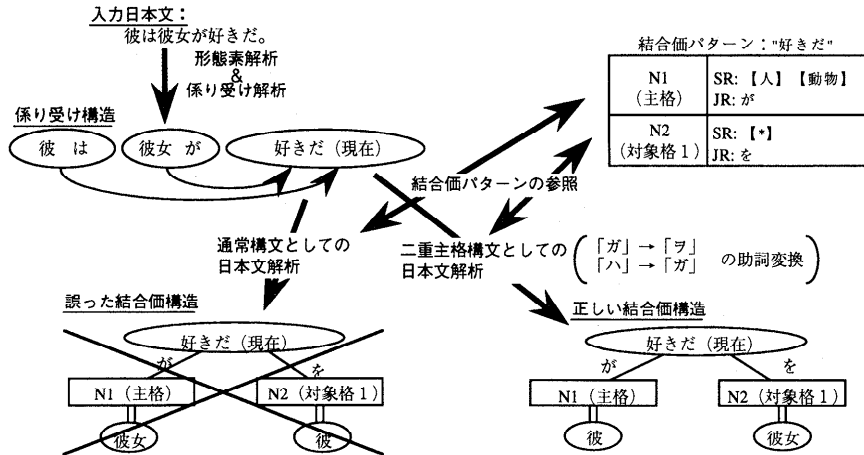


図4 二重主格形容詞構文の解析例 (タイプ3)

Fig. 4 An example of analyzing double-nominal-case construction (type-3).

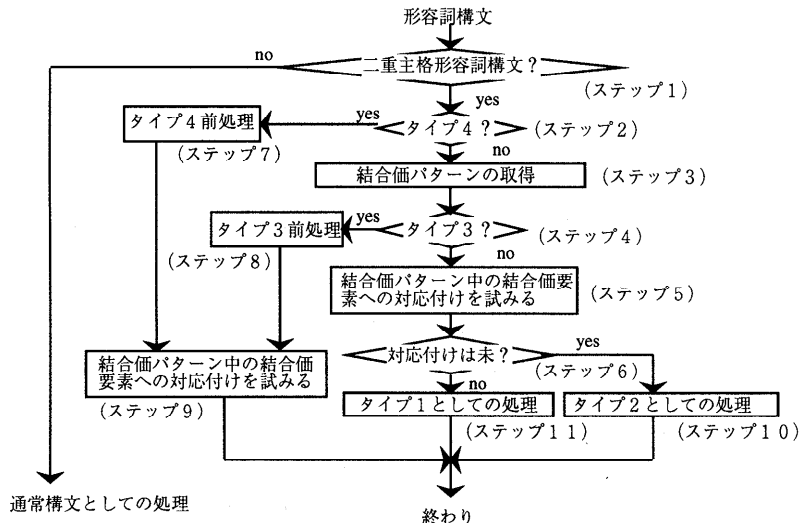


図5 二重主格形容詞構文の解析処理フロー

Fig. 5 Processing flow for analyzing Japanese double-nominal-case construction.

構文と判定する。このような場合には、語順のうえで文頭に近い修飾語句を第一主格とし、もう一方の修飾語句を第二主格とする^{*}。二重主格形容詞構文でない場合には、3章で述べた通常の日本文解析を実施する。

○第2フェーズ：タイプの決定

まず、第一主格が時間を表していればその入力文を

タイプ4に設定する (図5ステップ2)。この判定は、第一主格の意味属性^{3),14)}が【時間】を表しているかどうかによって行う。

次に入力文の述語に対する結合価パターン¹³⁾を参照することによってタイプ3を設定する (図5ステップ3, 4)。述語が情意形容詞でかつ、その結合価パターンが主格N1と対象格N2とを持っている場合に入力文をタイプ3と判定する。

最後にタイプ1とタイプ2との切り分けを行う (図5ステップ5, 6)。この切り分けは、入力文中の修飾語句と結合価パターンにおける結合価要素との対応付けの結果を基に行う。第二主格が結合価要素である主格

^{*} 副助詞 (または格助詞「ガ」) を持つ修飾語句を2つ有する場合には、どちらの修飾語句が、第一主格であるか (あるいは第二主格であるか) は文脈に依存する場合がある。ここでは、文脈を考慮せず、第1次近似として本文中のような語順に基づく判定を行っている。なお、この近似の有効性については5章の評価実験において検証する。

N1 に対応付けられ、第一主格がどの結合価要素にも対応付けられていない場合をタイプ 2 に設定する。それ以外の場合をタイプ 1 に設定する。タイプ 1 の場合には、第一主格は「ニ」格などの主格以外の格に対応付けられる。

○第 3 フェーズ：結合価構造の決定

タイプ 4 の前処理（図 5 ステップ 7）では、第一主格（時間を表現）は副詞句として解析される。次に、残りの修飾語句に対応付けの済んでいない結合価要素に対応付けることを試みる（図 5 ステップ 9）。

タイプ 3 の前処理（図 5 ステップ 8）では、対象格を表現する第二主格を結合価要素に対応付ける前に、格助詞「ガ」を格助詞「ヲ」に変換する。この変換によって、各修飾語句を正しい結合価要素に対応付けることができる。すなわち、まず格助詞「ヲ」に変換した修飾語句を結合価要素 N2 に対応付け、次に第一主格に対応付けの済んでいない結合価要素 N1 に対応付けることによって、正しい結合価構造を得ることができる（図 4 右下参照）（図 5 ステップ 9）。

タイプ 2 の処理（図 5 ステップ 10）では、副助詞「ハ」が連体型の格助詞「ノ」を代行しており、第一主格が第二主格を修飾するものとして解析しなければならない。そこで本論文では、正しい結合価構造を得るために、仮想的な結合価要素「N1 修飾」を導入する。この仮想的な結合価要素には、助詞制約として「ハ」を持たせ、意味的制約として【is-a】または【has-a】を持たせる（図 3 右上の結合価パターン中の点線枠内）。結果として、タイプ 2 の解析は結合価構造の再構成を必要とする（連用格の第一主格を、第二主格を修飾する連体格の語句として解釈するため）[☆]。図 3 右側の結合価構造において、名詞句「象の鼻」は、仮想的な結合価要素「N1 修飾」に対応付けられた副助詞「ハ」を持つ「象」と、格助詞「ガ」を持つ「鼻」とから構成される。そして、「象の鼻」が述語「長い」に対する主格に対応付けられる（図 3 右下参照）^{☆☆}。

タイプ 1 における第一主格は、連用型の格助詞「ニ」や「デ」などを代行しているの、タイプ 1 の解析は 3 章で述べた通常の日本文解析と同様に処理することができる（図 5 ステップ 11）。

上記の処理の流れによって、タイプ 1~4 に属する

二重主格形容詞構文の正しい結合価構造が得られる。特に、仮想的な結合価要素を導入することによって、副助詞「ハ」が連体型の格助詞「ノ」を代行しているタイプ 2 に対する正しい結合価構造を得ることができる。

5. 評価実験

4.3 節で提案した手法の有効性を検証するために評価実験を行った。本評価実験では以下の 2 点を明らかにする。

(1) 語順に基づく解釈に関する評価

副助詞（または格助詞「ガ」）を持つ修飾語句を 2 つ有する場合、文頭に近いほうを第一主格、もう一方を第二主格として解釈しているが、この語順に基づく解釈の妥当性を検証する。

(2) 解析精度の評価

図 5 に示す提案手法による二重主格形容詞構文の解析精度について検証する。

5.1 評価対象文の抽出

1991 年毎日新聞約 1 カ月分の新聞記事から、高頻度の形容詞を述語として含む二重主格形容詞構文を抽出した。対象とした形容詞は、用語頻度付与単語一覧¹⁰⁾に掲載されている頻度 6 以上のもの 153 語（形容詞 75 語、形容動詞 78 語）である。形容詞構文のべ 20,300 件の中から、人手によって二重主格形容詞構文 353 件を抽出した。その内訳は、副助詞「ハ」を持つ修飾語句と格助詞「ガ」を持つ修飾語句とを有する基本型が 256 件、2 つの修飾語句とも副助詞を持つものや 2 つとも格助詞「ガ」を持つもの（これらを派生型と呼ぶ）が 97 件であった。

5.2 評価手順

(1) 語順に基づく解釈に関する評価

派生型 97 件に対して、文頭に近いほうの修飾語句を第一主格とし、もう一方の修飾語句を第二主格と仮定した場合、その解釈に問題があるか否かを人手によって検証した。

(2) 解析精度の評価

上記の二重主格形容詞構文 353 件に対して、

- 人手による分類（タイプ 1~4 とその他）、
- 4.3 節で示した図 5 の解析処理フローに基づく分類、

を行った。そして、両者の結果を比較し、誤りに対する分析を行った。

5.3 評価結果

(1) 語順に基づく解釈に関する評価

派生型 97 件すべてを人手によって分析した結果、

☆ 第一主格は第二主格の主名詞を修飾するものとして構造変換を行う。

☆☆ 図 5 の解析処理フローでは、タイプ 2 として処理されるものの中には、その他に分類されるべきものも含まれてしまう。タイプ 2 からその他に分類されるべきものを精度良く選別する手法については今後の研究課題である。

表 1 二重主格形容詞構文の分類ごとの件数

Table 1 Classification results.

分類	件数 (基本型の内数)	全体に対する比率 [%]
タイプ 1	148 (117)	41.9 (45.7)
タイプ 2	75 (52)	21.3 (20.3)
タイプ 3	7 (5)	2.0 (2.0)
タイプ 4	64 (38)	18.1 (14.8)
その他	59 (44)	16.7 (17.2)
合計	353 (256)	100 (100)

表 2 提案手法の評価結果

Table 2 Experimental results.

分類	件数	提案手法による 解析成功件数	割合 [%]
タイプ 1	148	133	89.9
タイプ 2	75	74	98.7
タイプ 3	7	7	100
タイプ 4	64	62	96.9
合計	294	276	93.9

タイプ 1~4 に属する 82 件については、上記の語順に基づく解釈によって正しく解析できることが分かった。また、その他に属する 15 件の中には文脈に依存しているために上記の語順に基づく解釈では誤りを生じるものがあった。この結果から、新聞記事という分野においてタイプ 1~4 を対象とした場合、二重主格形容詞構文の解析における上記の語順に基づく解釈は第 1 次近似として妥当であると考えられる。

(2) 解析精度の評価

人手による分類の結果を表 1 に示す。表 1 より、基本型、派生型ともに、タイプ 1~4 に分類されるものが全体の約 83%、いずれにも属さないその他が約 17% 出現することが分かる。このことから、本論文で提案したタイプ 1~4 を対象にした解析手法は二重主格形容詞構文全体の 80% 以上を対象とすることができると考えられる。

提案手法による分類と人手による分類とを比較した結果を表 2 に示す。表 2 は、人手によってタイプ 1~4 に分類されたものそれぞれに対して提案手法がどのような分類を行うかという観点でまとめた。

表 2 の結果から、解析の対象としたタイプ 1~4 に分類される二重主格形容詞構文を約 94% の精度で正しく解析できることが分かる。表 1 よりその他に分類されるべきものが 59 件存在するが、これらについては提案手法ではタイプ 1 またはタイプ 2 に分類されてしまう。これを考慮すると、提案手法の解析精度は二重主格形容詞構文全体で見ると約 78% (276 件/353 件) となる。

5.4 考 察

解析に失敗した例の多くは、第一主格を自由格（たとえば場所を表す「デ」格）として解析しなければならないものを、誤った格あるいは「ガ」格を修飾するものとして解析してしまったものである。本論文で提案している手法は結合価パターンに記述されている結合価構造に基づいている。一方、結合価パターンに記述されている結合価要素はその述語を特徴付ける結合価要素のみであり、場所や時間などの自由格については記述されていない^{1),13)}。このため、自由格の解析に失敗する場合が生じる。これを避けるためには、名詞の持つ意味属性と格助詞とから自由格を推定するなど、自由格の解析に対する検討が必要である。この点については今後の検討課題である。

また、その他に分類される二重主格形容詞構文を分析すると以下のようなものが多い。

(ア) 比較構文（例：90 年のほうが上昇が大きい）。

(イ) 省略があつて本来そこに係るべき第一主格が述語を修飾してしまうために、見かけ上二重主格形容詞構文となっているもの（例：脅迫状はいたずらの可能性が高い → [[[脅迫状は [いたずらである]] 可能性が] 高い]：太字が省略箇所）。

(ア) については、見かけ上、二重主格形容詞構文と同じ形式をとるが、「～のほうが～よりも」のように 1 つの格が 2 つに分離する現象や、「～のほうが」によって比較する対象を表す現象を、二重主格形容詞構文として解釈する前に比較構文として解釈しておく必要がある。(イ) については、単純に結合価構造を用いて解析することができない。どのような場合に省略が生じるのかを分析して、結合価構造に基づく解析を行う前に省略をあらかじめ補完しておく必要がある。これらの解析については、日本文解析全体の中でどのように処理すべきかを考察する必要があり、今後の検討課題としたい。

6. 日英機械翻訳への適用

筆者らは結合価構造に基づく日本文解析を日英機械翻訳システム ALT-J/E の日本語処理部に適用している^{1),2)}。ALT-J/E において、日本文解析は入力文の述語が持つ結合価構造を基に行われており¹⁾、述語慣用句や機能動詞表現といった特殊な日本語構文を翻訳することも可能である¹¹⁾。この ALT-J/E に本論文で提案した二重主格形容詞構文の解析手法を適用することにより、次の 2 つの効果が得られる。

(1) 第一主格を正しく訳出

従来、日本文解析において第一主格が結合価構造

(あるいは格構造)に反映されなかった場合には、副助詞「ハ」の文法的機能を基に「As for ~」を一律に生成している場合が多かった。本手法によれば、第一主格を結合価構造に反映することができるので、より正確な訳出が可能となる。

(2) 訳語選択の精度向上

タイプ2において、第一主格と第二主格との意味的な関係が得られるので、よりよい訳語選択が可能となる。たとえば「象は鼻が長い、豚は鼻が短い」に対して、「Elephants have long trunks and pigs have short snouts.」(斜体部が訳語選択による訳し分け)のように英語の主語に応じた訳語を選択することが可能となる。

以上のように、本論文で述べた二重主格形容詞構文の解析手法を日英機械翻訳の日本語解析に適用することにより、翻訳精度の向上が可能となる。

7. 結 論

本論文では結合価構造に基づく日本語二重主格形容詞構文の解析手法について述べた。本論文では、日本語二重主格形容詞構文を副助詞「ハ」の機能に基づいて4タイプとその他の5つに分類し、従来の日本語解析手法を適用したときに生じる問題点について述べた。次に4タイプを対象にこれらの問題点を解決した解析手法を提案した。

本手法ではまず、表層的に第一主格と第二主格に相当する修飾語句を有する二重主格形容詞構文に対して、その述語の持つ結合価パターンに基づいてタイプ分けを行う。次に、各タイプに応じた解析を行うことにより、解析対象の二重主格形容詞構文の結合価構造を得る。特に、副助詞「ハ」が連体型の格助詞「ノ」を代行しているタイプ2を正しく解析するために、仮想的な結合価要素「N1修飾」を導入した。これにより、第一主格と第二主格との意味的な関係をとらえ、結合価構造に正しく反映することが可能となった。

さらに本論文では、本手法の有効性を示すために新聞記事を対象に行った評価実験について述べた。この結果、二重主格形容詞構文の約78%(対象とした4タイプに限れば約94%)が本手法によって正しく解析できることが分かり、本手法の有効性が確認できた。

また、本論文で述べた仮想的な結合価要素を用いるという二重主格形容詞構文解析の考え方は、埋め込み文の解析に適用することができる¹²⁾。「屋根が白い家」(「家が白い」のではなく、「家の屋根が白い」)のように形容詞の主格を修飾する名詞が外置された構造は、従来は外置された名詞(上記の例では「家」)が結合

価構造(あるいは格構造)に反映されずに残されていた。本論文で導入した仮想的な結合価要素を用いると、外置される前の元の日本語を二重主格形容詞構文「家は屋根が白い」と考えることができる。これにより、仮想的な結合価要素「N1修飾」=「家は」が外置された構造として埋め込み文「屋根が白い家」を扱うことができる。すなわち、本論文で述べた二重主格形容詞構文の解析手法によって、第一主格を結合価構造に取り込めば、結合価要素が外置された埋め込み文(例:「白い屋根」←「屋根が白い」の主格が外置)と同様に、形容詞の主格を修飾する名詞が外置された構造を扱うことが可能となる。

謝辞 本研究遂行の機会を与えていただき、つねにご指導いただいた池原悟氏、宮崎正弘氏に感謝いたします。また、研究遂行にあたり、つねに有益なご助言およびご指導をいただいた林良彦氏に感謝いたします。評価実験の遂行にあたっては東田正信氏に有益なご助言をいただきました。また、例文の収集、日英機械翻訳システムALT-J/Eへの適用に関しては松尾義博氏にお世話になりました。両氏に感謝の意を表します。

参 考 文 献

- 1) 林 良彦: 結合価構造に基づく日本語解析, 情報処理学会自然言語処理研究会, 62-6 (1987).
- 2) 池原 悟, 宮崎正弘, 白井 諭, 林 良彦: 言語における話者の認識と多段翻訳方式, 情報処理学会論文誌, Vol.28, No.12, pp.1269-1279 (1987).
- 3) 池原 悟, 宮崎正弘, 白井 諭, 横尾昭男, 中岩浩巳, 小倉健太郎, 大山芳史, 林 良彦: 日本語語彙大系, 岩波書店 (1997).
- 4) 石上照雄: 二重主格形容詞文の構造, 日本語学試験(愛知教育大学国語学研究室), Vol.3, pp.1-37 (1977).
- 5) 石綿敏雄: 文法と意味 I, 第2章 結合価から見た日本文法, 朝倉書店 (1983).
- 6) JEIDA: 機械翻訳システム評価基準, 日本電子工業振興協会 (1995).
- 7) Kuno, S.: *The structure of the Japanese Language*, MIT Press (1973).
- 8) 三上 章: 象は鼻が長い, くろしお出版 (1960).
- 9) 村田賢一: 形容詞の形式的意味記述についての考察, 第40回情報処理学会全国大会論文集, 5F-3 (1990).
- 10) 日本電子工業振興協会: 用語頻度付単語一覧 (1986).
- 11) 奥 雅博: 日本文解析における述語相当の慣用的表現の扱い, 情報処理学会論文誌, Vol.31, No.12, pp.1727-1734 (1990).
- 12) Oku, M.: A Method for Analyzing Embedded Noun Phrase Structures Derived

from Japanese Double-Nominal-Case Construction, *Proc. 12th Pacific Asia Conference on Language, Information and Computation (PACLIC12)* (1997).

- 13) 白井 諭, 横尾昭男, 井上浩子, 中岩浩巳, 池原 悟, 八木晶子: 日英機械翻訳における意味解析のための構文辞書, 言語処理学会第3回年次大会発表論文集, pp.153-156 (1997).
- 14) 横尾昭男, 宮崎正弘, 阿部さつき, 池原 悟, 白井 諭, 細井純子: 日英機械翻訳における意味解析のための単語辞書, 言語処理学会第3回年次大会発表論文集, pp.149-152 (1997).

付録 その他に属する二重主格形容詞構文の例

○第一主格が主格以外の結合価要素を修飾

- 長男は生命に別状がなかった ← 「長男は」が「生命に」を修飾

○第二主格が第一主格の属性を表現

- 辞書は新しいのが良い ← 新しい辞書が良い
- 冷蔵庫は大型が良い ← 大型の冷蔵庫が良い
- 目標は大きいほうが良い ← 大きな目標のほうが良い

○比較構文

- 90年のほうが上昇が大きい
- 兄のほうが背が高い

○副助詞「ハ」が取り立てを表現(文脈を考慮する必

要がある)

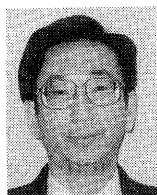
- 紅葉は秋がきれいだ ← 紅葉についていえば, 秋がきれいだ
- かき料理は広島が有名だ ← かき料理についていえば, 広島が有名だ

○省略によって表層上の二重主格形容詞構文を構成(太字が省略箇所)

- 日曜稼働させる店舗は商業地や駅周辺が多い ← [[日曜稼働させる店舗は[商業地や駅周辺にある]] ことが] 多い]
- 脅迫状はいたずらの可能性が高い ← [[[脅迫状は[いたずらである]] 可能性が] 高い]

(平成10年2月3日受付)

(平成10年9月7日採録)



奥 雅博(正会員)

昭和35年生。昭和57年大阪府立大学工学部電子工学科卒業。昭和59年同大学院博士前期課程修了。同年日本電信電話公社(現NTT)入社。入社以来平成10年3月まで研

究所において自然言語処理の研究実用化に従事。現在、NTTグループ企業本部新事業開発室担当課長。慣用表現や比喩などの非標準的な言語現象に興味を持つ。電子情報通信学会, 言語処理学会各会員。