

概要

日英機械翻訳において結合価パターン [1] は単文レベルの翻訳精度向上に大きく寄与している。しかし、結合価パターンは基本構文で登録されているため、構文変形 (受動・可能・使役など) に伴う格助詞の変化には対応していないという問題点がある。

従来の研究では、構文変形された文を一般的な文と対応させるために格助詞の機能に応じて格助詞を変化させる方法を提案している [2, 3, 4]。しかし、抽象的な制約を用いているため機械的に処理することが困難である。

本稿では構文変形の中でも使用頻度の高い受動態から能動態への文変形について、文型パターン対を用いた格助詞の変化規則「構文変形パターン対」を作成し、その効果を検証した。

単文の受動文に対して本稿の規則を適用した結果、クローズドテスト (50 文) で正解率 100%、オープンテスト (100 文) で正解率 72%を得ることができた。

目次

1	まえがき	1
2	従来の研究	2
2.1	日本語語彙大系	2
2.1.1	結合価パターン	2
2.1.2	一般名詞意味属性	3
2.1.3	用言意味属性	4
2.2	従来の構文変形文への格助詞の変化規則	5
3	構文変形パターン対	6
3.1	構文変形文への結合価パターンの照合	6
3.2	構文変形パターン対	7
4	構文変形パターン対の作成	8
4.1	固有構文変形パターン対の作成	8
4.2	汎用構文変形パターン対の作成	9
5	構文変形パターン対の適用判定	10
5.1	用言意味属性による制約	10
5.2	結合価パターンによる制約	12
6	実験	13
6.1	入力文	13
6.2	実験条件	13
6.3	評価方法	14
6.4	実験結果	14
7	考察	15
7.1	構文変形パターン対を一意に絞りこめない	15
7.1.1	入力文：その時期に何かのパーティーが催されるようだ	16
7.1.2	入力文：軍隊は戦闘陣形に配置された	16
7.1.3	入力文：ガソリンはさまざまな化合物から調製される	16
7.1.4	入力文：車はその爆発でバラバラに吹き飛ばされた	17

7.1.5	入力文：その声明は各紙で大々的に報道された	17
7.1.6	入力文：離婚判決は公開法廷で宣言されなくてはならない	18
7.1.7	入力文：彼女の本は世界各地で読まれている	18
7.1.8	入力文：その議案は票決に付された	19
7.1.9	構文変形パターン対を一意に絞りこめない原因と考察	19
7.2	対応する構文変形パターン対が存在しない	19
7.3	用言意味属性制約付きパターン対が存在しない	20
7.4	用言意味属性による制約の失敗	21
7.4.1	入力文：ゲリラ活動が国境近くでにわかに再開された	21
7.4.2	入力文：その村はゲリラによって攻撃された	22
7.4.3	入力文：正式に両親から勘当された	22
7.5	結合価パターンによる制約の失敗	22
7.5.1	入力文：あなたの画家としての才能は絶え間ない修練によってのみ磨かれるだろう	23
7.5.2	入力文：あの場面は映画の検閲官に許可されなかった	23
7.5.3	入力文：それは第4条で規定されている	23
8	まとめ	24

目 次

1	一般名詞意味属性体系	3
2	用言意味属性大系	4
3	固有構文変形パターン対の出現頻度とパターン数	8
4	汎用構文変形パターン対の出現頻度とパターン数	9
5	用言意味属性制約付き汎用構文変形パターン対の出現頻度とパターン数	11

表 目 次

1	結合価パターン	2
2	結合価パターンと用言意味属性	4
3	構文変形パターン対	7
4	固有構文変形パターン対 (上位 5 件)	8
5	汎用構文変形パターン対 (上位 5 件)	9
6	用言意味属性制約付き汎用構文変形パターン対 (上位 5 件)	11
7	実験結果	14
8	変化規則失敗の内訳	15

1 まえがき

日英機械翻訳において結合価パターンは単文レベルの翻訳精度向上に大きく寄与している。しかし、結合価パターンは基本構文で登録されているため、構文変形(受動・可能・使役など)に伴う格助詞の変化には対応していないという問題点がある。

従来の研究では、構文変形された文を一般的な文と対応させるために格助詞の機能に応じて格助詞を変化させる方法を提案している。しかし、抽象的な制約を用いているため機械的に処理することが困難である。

本稿では構文変形の中でも使用頻度の高い受動態から能動態への文変形を対象とした。そして受動文と能動文を対応させる格助詞の変化規則として、受動文と対応する能動文から、それぞれ名詞の変数化を行ない、格助詞と用言の文構造を抽出した文型パターン対「構文変形パターン対」を作成し、その効果を検証した。

単文の受動文に対して本稿の規則を適用した結果、クローズドテスト(50文)で正解率100%、オープンテスト(100文)で正解率72%を得ることができた。

本論文の構成は以下の通りである。第2章では従来の研究について紹介する。第3章では構文変形パターン対について説明する。第4章で構文変形パターン対の作成を行なう。第5章で構文変形パターン対を適用判定する方法について考案する。第6章で実験、第7章で考察、第8章でまとめを述べる

2 従来の研究

日英機械翻訳において，日本語語彙大系に集録されている結合価パターンは翻訳精度の向上に寄与している．しかし，結合価パターンは基本構文で登録された文型パターンであるため，構文変形(受動，可能，使役など)された文に適用するには格助詞を対応させる必要がある．

以下2.1節では日本語語彙大系について，2.2節では従来の構文変形された文を一般的な文と対応させる手法について述べる．

2.1 日本語語彙大系

2.1.1 結合価パターン

池原らによってまとめられた，日本語語彙体系 [1] の構文体系には，日本語の用言(約6,000語)の結合価パターン(約14,000件)が定義されている．

結合価パターンとは，格要素と用言の意味的共起関係を体系化したものである．格要素の名詞は名詞意味属性によって規制されている．

例えば，用言「養う」の結合価パターンを検索すると，表1に示されるパターンが得られる．

表 1: 結合価パターン

「養う」	
(1)	$N1\{\text{主体}\}$ が $N2\{\text{抽象物, 人間活動, 性質}\}$ を 養う $N1$ cultivate $N2$
(2)	$N1\{\text{人}\}$ が $N2\{\text{目(器官), 耳, 感じ, 感覚}\}$ を 養う $N1$ train $N1's$ $N2$
(3)	$N1\{\text{主体}\}$ が $N2\{\text{力(主体)}\}$ を 養う $N1$ develop $N2$
(4)	$N1\{\text{主体, 施設}\}$ が $N2\{\text{人, 動物}\}$ を 養う $N1$ raise $N2$
(5)	$N1\{\text{人}\}$ が $N2\{\text{人間(親族関係)}\}$ を $N3\{\text{給与}\}$ で 養う $N1$ support $N2$ with $N3$

結合価パターンには対をなす英語表現が記されており，結合価パターンが定まれば対応する英語表現を得ることができる．また，結合価パターンの格要素は「{主体}が」

や「{ 抽象物, 人間活動, 性質 } を」のように, 一般名詞意味属性によって分類, 規制された名詞 ({ 主体 }, { 抽象物, 人間活動, 性質 }, 等) と格助詞を組み合わせ構成している. そして格要素により用言は意味的に分類, 規制され, 一つの用言に対し複数の結合価パターンが存在する.

2.1.2 一般名詞意味属性

本研究における一般名詞意味属性 [1] とは, 単語を従来からある語彙的分類ではなく, 意味的用法により階層的に分類, 体系化したものである. 約 40 万語の単語を 12 段の木構造により 2,710 の意味属性に分類しており, 日本語のシソーラスとしての規模, 語彙数, 意味分解能は世界最大級のものである.

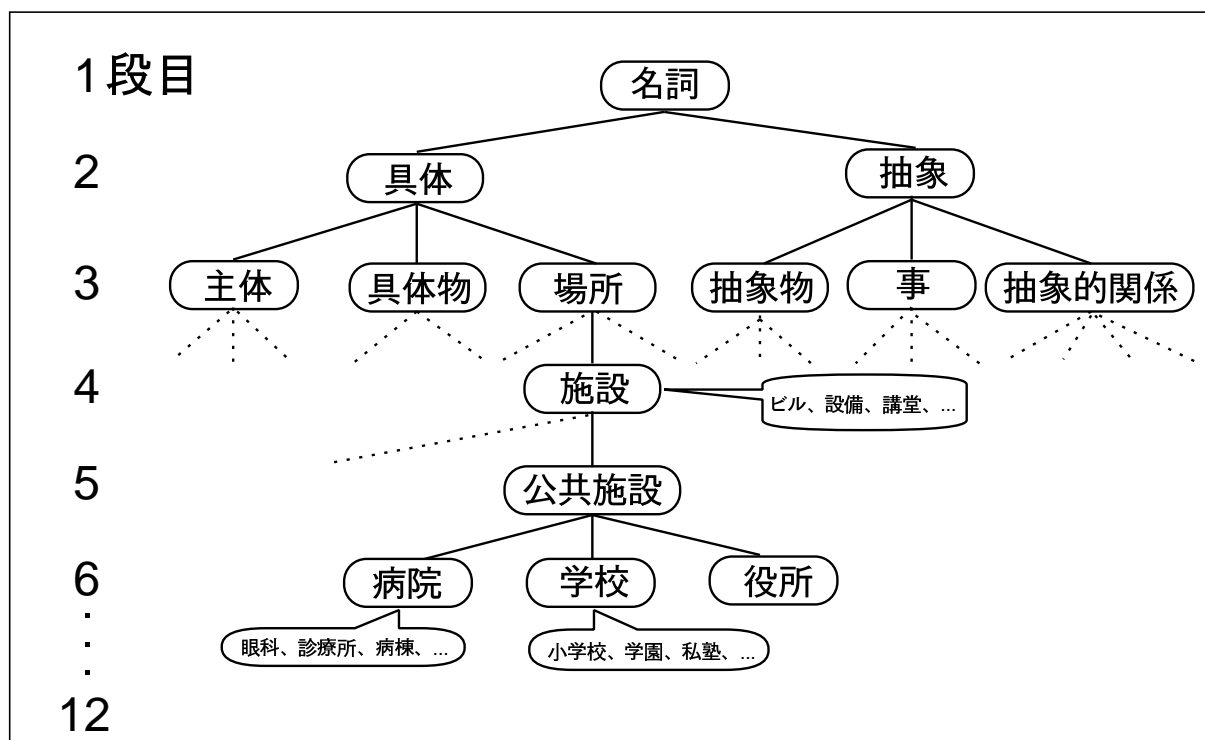


図 1: 一般名詞意味属性体系

図 1 は意味属性体系の構造の一部を記したものである. 「主体」, 「具体物」, 「施設」などの意味属性で構成され, 「施設」という意味属性には { ビル, 講堂 } といった名詞が分類されている.

また一般名詞意味属性には, 上位の意味属性の性質は下位の意味属性の性質に伝搬・継承できるという性質がある. 例えば, ビルという名詞は「施設」という上位の意味属

性を持つが、下位属性に分類される { 小学校, 学園, 私塾 } などの名詞は「施設」だけでなく下位属性の「学校」「公共施設」といった意味属性も持つ。

2.1.3 用言意味属性

日本語語彙大系 [1] に示されている用言意味属性とは、「用言が持つ動的属性の種類」と「用言の格に対する関係」の2種類の観点から分類した約 100 種類の属性からなる。

本稿では、「用言が持つ動的属性の種類」の観点から分類した 36 種類の属性からなる属性大系を使用する。

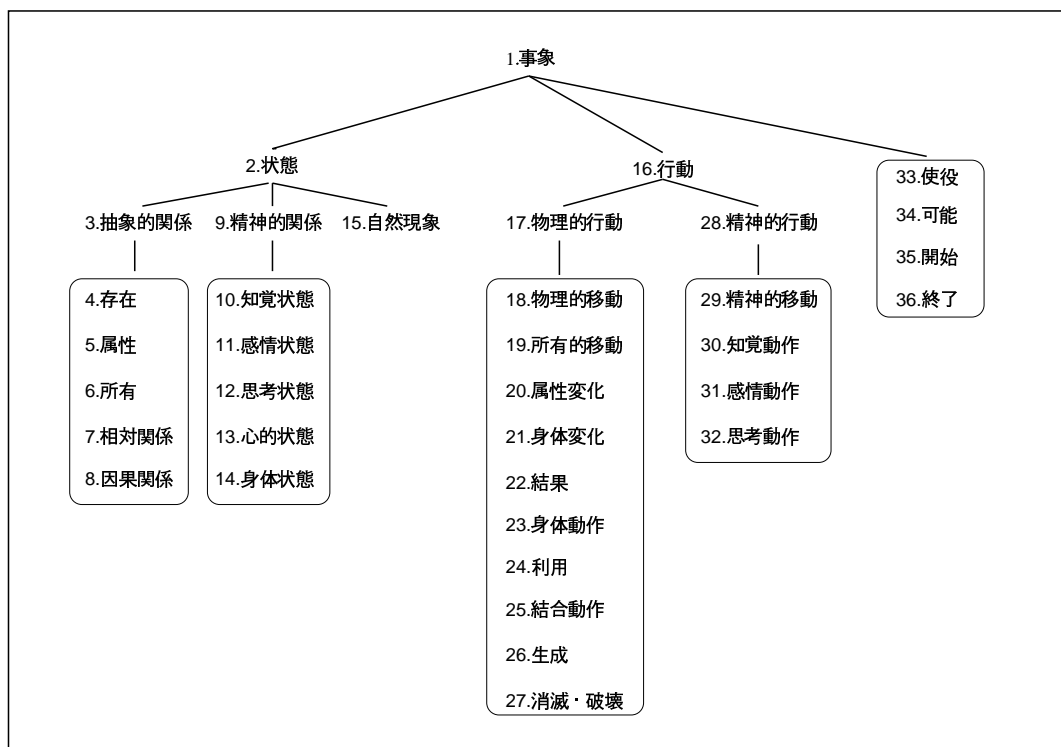


図 2: 用言意味属性大系

結合価パターンには、用言としての意味的用法が用言意味属性により示されている。例を表 2 に示す。

表 2: 結合価パターンと用言意味属性

結合価パターン	用言意味属性
{ 主体 } が { 動物 } をかわいがる	感情動作
{ 主体, 具体物 } が { 抽象 } に成功する	結果

2.2 従来の構文変形文への格助詞の変化規則

言語学的に，仁田ら [2, 3] や寺村 [4] は格助詞や動詞の機能に応じて，構文変形文と基本構文を対応させる方法を提案している．

例えば「 X が Y に V 」という基本構文に対する受動文は， V が2者の関係を表し，物理的・心理的働きかけを行なう動詞の場合「 Y が X に $V.rareru$ 」となり， V が2者の関係を表し，創造を行なう動詞の場合「 Y が X によって $V.rareru$ 」となるという規則を示している．しかし，これらの規則は抽象的であり，機械処理に実装するための具体的な手法が示されていない．また文意を考えなければならないため，機械処理で扱うのは困難である．

工学的な方法として，吉田 [6] は構文変形文に対して，代表的な格助詞「が」，「を」，「に」，「と」，「から」のみを扱い，受動文の「が」格は能動文の「を」格に変化するというように，格助詞が一意に変化すると仮定して処理を行なっている．しかし，実際の格助詞は一意に変化しないため，この方法には問題がある．

3 構文変形パターン対

構文変形文から基本構文を得るために、格助詞と用言の文構造に着目した格助詞の変化規則を提案する。

以下、2.1 節では構文変形文へ結合価パターンの照合する方法について述べ、2.2 節では格助詞の変化規則「構文変形パターン対」を提案する。

3.1 構文変形文への結合価パターンの照合

結合価パターンは基本構文で登録されている。そのため、構文変形(受動・可能・使役など)された文に対して直接には照合することができない。

構文変形文に結合価パターンを照合する方法を以下の受動文を例に説明する。

<受動文> 子供が両親に養われる

上記の受動文の用言 { 養われる } の原型 { 養う } の結合価パターンは以下の5つが検索される。

パターン1	{主体}が{抽象物, 人間活動, 性質}を養う
パターン2	{人}が{目(器官), 耳, 感じ, 感覚}を養う
パターン3	{主体}が{力(主体)}を養う
パターン4	{主体, 施設}が{人, 動物}を養う
パターン5	{人}が{人間(親族関係)}を{給与}で養う

5つのパターンのうち、受動文を満たす結合価パターンとして次のパターンが選ばれる。

<結合価パターン> {主体, 施設}が{人, 動物}を養う

しかし、上記の受動文は構文変形されることで格助詞が変化しているため、結合価パターンと字面では照合できない。そこで、受動文を以下の能動文に変形する。

<能動文> 両親が子供を養う

この能動文は基本構文となっており、結合価パターン「{主体, 施設}が{人, 動物}を養う」と照合することができる。

3.2 構文変形パターン対

本稿では，構文変形文に対して結合価パターンを照合するため，格助詞と用言の構造に着目した構文変形パターン対による格助詞の変化規則を提案する．

構文変形パターン対は，構文変形文と対応する基本構文からそれぞれ名詞を変数化して格助詞と用言の構造を抽出した変形構文パターンと基本構文パターンにより構成される．また，構文変形パターン対には用言を字面で制約した固有構文変形パターン対と，用言を変数化した汎用構文変形パターン対がある（表3）．

表 3: 構文変形パターン対

構文変形パターン対	変形構文パターン	基本構文パターン
固有構文変形パターン対	$N1$ は $N2$ に養われる	$N2$ が $N1$ を養う
汎用構文変形パターン対	$N1$ は $N2$ に $V.rareru$	$N2$ が $N1$ を V

固有構文変形パターン対に比べ汎用構文変形パターン対は高いカバー率を有する．しかし，格助詞の変化規則が一意に定まらないため，より強い制約と組み合わせる必要がある．

4 構文変形パターン対の作成

構文変形パターン対を作成するために、まず日英対訳データベース [5] からランダムにレル・ラレルを含む 1,000 文を抽出し、人手により単文の受動文 (633 文) を収集する。次に結合価パターンに一致するよう基本構文に変化させた構文変形文対を作成する。

4.1 固有構文変形パターン対の作成

構文変形文対から格要素と用言を抽出し、名詞を変数化した固有構文変形パターン対を作成する。その結果、554 個の規則が得られた。得られた規則の中で最も出現頻度の高かった上位 5 件を表 4 に示す。

表 4: 固有構文変形パターン対 (上位 5 件)

頻度	変形構文パターン	基本構文パターン
8	$N1$ は $N2$ で行われる	$N1$ を $N2$ で行う
7	$N1$ は $N2$ に包まれる	$N2$ が $N1$ を包む
5	$N1$ は $N2$ で覆われる	$N2$ が $N1$ を覆う
5	$N1$ には $N2$ が含まれる	$N1$ が $N2$ を含む
3	$N1$ は読まれる	$N1$ を読む

固有構文変形パターン対の出現頻度とパターン対の種類数の関係を図 3 に示す。

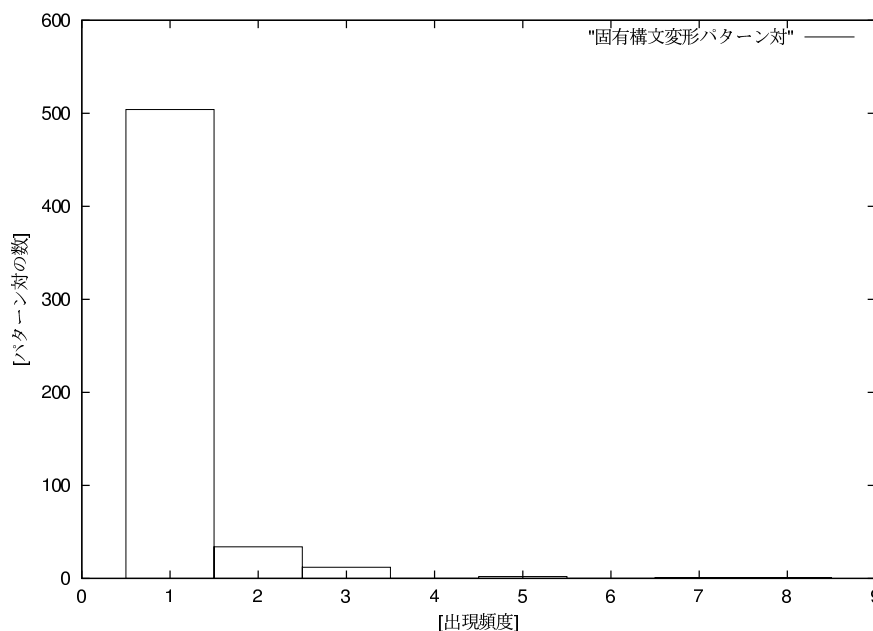


図 3: 固有構文変形パターン対の出現頻度とパターン数

4.2 汎用構文変形パターン対の作成

固有構文変形パターン対は用言が強く制約されているため，登録されていないパターン対が多く，全ての用言をカバーすることができない．そこで用言を汎化し，格助詞の変化構造を抽出した汎用構文変形パターン対を作成する．構文変形文対から 153 個の規則が得られた．表 5 に一部を示す．

表 5: 汎用構文変形パターン対 (上位 5 件)

頻度	変形構文パターン	基本構文パターン
85	$N1$ は $V.rareru$	$N1$ を V
75	$N1$ は $N2$ に $V.rareru$	$N1$ を $N2$ に V
46	$N1$ は $N2$ に $V.rareru$	$N2$ が $N1$ を V
44	$N1$ は $N2$ で $V.rareru$	$N1$ を $N2$ で V
34	$N1$ が $V.rareru$	$N1$ を V

汎用構文変形パターン対の出現頻度とパターン対の種類数の関係を図 4 に示す．

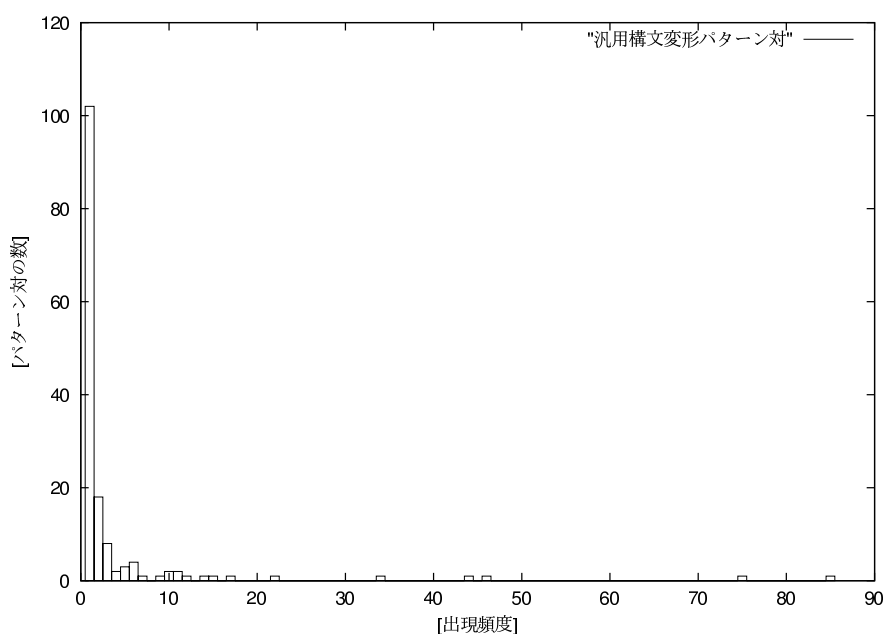


図 4: 汎用構文変形パターン対の出現頻度とパターン数

5 構文変形パターン対の適用判定

このままの汎用構文変形パターン対は用言の制約が無いため，様々な用言を用いた構文変形パターン対が候補としてあがる．そこで，5.1 節の制約で大きくパターン対を制限し，5.2 節の方法で意味的に正しい候補を選択する．

5.1 用言意味属性による制約

[2, 3, 4] では用言の意味的用法により分類した格助詞の変化規則が示されており、用言の意味が格助詞の変化規則に制約を与えていると考えられる．そこで本稿では，用言に制約を持つ用言意味属性を用いることで汎用構文変形パターン対を制約する手法を提案する．

用言意味属性は基本構文の例文から判断し，汎用構文変形パターン対に制約条件として追加する．具体的な例として，以下の構文変形文対を考える．

[構文変形文対]

< 構文変形文 > 猫が妹にかわいがられる

< 基本構文 > 妹が猫をかわいがる

上記の構文変形文対から，以下の汎用構文変形パターン対が得られる．

[汎用構文変形パターン対]

変形構文パターン	基本構文パターン
$N1$ が $N2$ に $V.rarereru$	$N2$ が $N1$ を V

また，ここで基本構文「妹が猫をかわいがる」を満たす結合価パターンと用言意味属性は，以下が検索される．

[結合価パターンと用言意味属性]

結合価パターン	用言意味属性
{ 主体 } が { 動物 } をかわいがる	感情動作

得られた用言意味属性と汎用構文変形パターン対を組み合わせることで，以下の用言意味属性制約付き汎用構文変形パターン対を作成する．

[用言意味属性制約付き汎用構文変形パターン対]

変形構文パターン	基本構文パターン	用言意味属性
$N1$ が $N2$ に $V.rareru$	$N2$ が $N1$ を V	感情動作

上記の規則により用言意味属性制約付き汎用構文変形パターン対の作成を行なった結果、373 個の規則が得られた (表 6) . なお、用言意味属性の約 7 割をカバーしている .

表 6: 用言意味属性制約付き汎用構文変形パターン対 (上位 5 件)

頻度	変形構文パターン	基本構文パターン	用言意味属性
14	$N1$ は $N2$ に $V.rareru$	$N1$ を $N2$ に V	精神的移動
13	$N1$ は $V.rareru$	$N1$ を V	属性変化
12	$N1$ は $V.rareru$	$N1$ を V	思考動作
12	$N1$ は $N2$ に $V.rareru$	$N1$ を $N2$ に V	思考動作
11	$N1$ は $N2$ に $V.rareru$	$N1$ を $N2$ に V	感情動作

用言意味属性制約付き汎用構文変形パターン対の出現頻度とパターン対の種類数の関係を図 5 に示す .

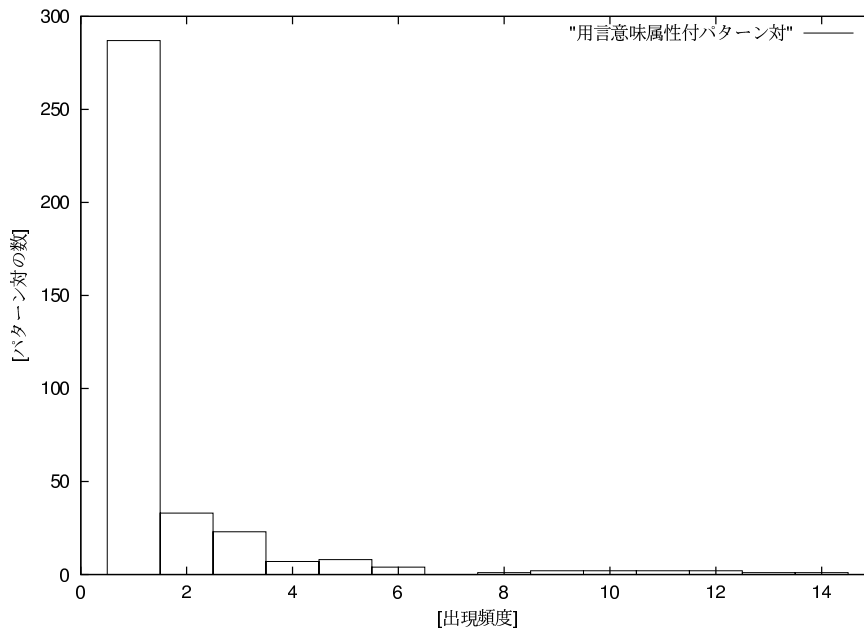


図 5: 用言意味属性制約付き汎用構文変形パターン対の出現頻度とパターン数

5.2 結合価パターンによる制約

構文変形された文が入力され，対応する構文変形パターン対が決まると，基本構文の文が候補として複数得られる．最終的に出力される文は意味的に正しくなければならない．そこで候補文に結合価パターンを照合し，意味的結合性から候補を選択する．以下の構文変形文を例に具体的方法を説明する．

<例文> 草花は彼女に育てられた

例文に対して，まず以下の汎用構文変形パターン対を適用することで2つの候補文を得ることができる．

[汎用構文変形パターン対]

	変形構文パターン	基本構文パターン
パターン対1	$N1$ は $N2$ に $V.rareru$	$N2$ が $N1$ を V
パターン対2	$N1$ は $N2$ に $V.rareru$	$N1$ を $N2$ に V

候補文1: 彼女が草花を育てる

候補文2: 草花を彼女に育てる

次に，例文2で用いられている用言 [育てる] の結合価パターンを調べると以下の5つパターンが得られる．

- {主体} が {抽象} を育てる
- {主体} が {植物, 力・能力等} を育てる
- {主体} が 伝統芸能を育てる
- {主体} が {人(専門的技術的職業), 人(支配的管理的職業)} を育てる
- {人, 動物} が {人, 植物} を育てる

最後に，結合価パターンに格要素の名詞意味属性が最も一致することを条件に，パターン”{主体} が {植物, 力・能力等} を育てる” に一致する候補文1を選択する．なお，この一致度は [6] のアルゴリズムで求めることができる．

6 実験

構文変形パターン対を用いた格助詞の変化規則の有効性を調べるために以下の実験を行った。

6.1 入力文

日英対訳コーパスから抽出した単文の受動文に対して実験を行う。また、時制や副詞に関しては考慮しないものとする。受動文の例を以下に示す。

- ガソリンはさまざまな化合物から調製される
- 上腕骨は肩甲骨と結ばれている

6.2 実験条件

構文変形パターン対を選択するために本稿で提案した規則を以下の4つの手法で適用する。

手法 (a) : 固有構文変形パターン対を適用する

手法 (b) : (a) を適用し、できなかった文に対しては汎用構文変形パターン対を適用する

手法 (c) : (b)+用言意味属性による制約を行う

手法 (d) : (c)+結合価パターンによる制約を行う

上記の4つの手法を以下の実験に適用する。

1. クローズドテスト :

構文変形パターン対の作成に用いた構文変形文からランダムに抽出した単文の受動文 (50 文) を対象とする。

2. オープンテスト :

日英対訳コーパスからランダムに抽出した単文の受動文 (100 文) を対象とする。ただし、構文変形パターン対の作成に使用した文は除く。

6.3 評価方法

構文変形パターン対により受動文から能動文への変形を行い，以下の3つ条件を満たす文を正解とする．

1. 一意に絞り込まれている
2. 人手で意味的に適切であると判断される
3. 結合価パターンに最も格要素が多く一致するよう格助詞が変化している

<例>

[結合価パターン]{主体，人工物}が{具体物}を{具体，場}に寄せる

○ブルドーザーが残骸を端へ寄せる

×ブルドーザーで残骸を端へ寄せる

結合価パターンに対して，{ブルドーザーが残骸を端へ寄せる}はガ格・ヲ格の2つの格要素が一致している．一方，{ブルドーザーで残骸を端へ寄せる}はヲ格のみしか格要素が一致していない．よって，最も格要素が多く一致する{ブルドーザーが残骸を端へ寄せる}を選択する．

6.4 実験結果

実験の結果を表7に示す．その結果，最も正解率が高くなったのは全ての規則を適用した手法(d)で，クローズドテストで100%，オープンテストで72%となった．

表 7: 実験結果

	手法 (a)	手法 (b)	手法 (c)	手法 (d)
クローズドテスト	98%	98%	98%	100%
オープンテスト	19%	27%	50%	72%

7 考察

オープンテストにおいて最も正解率の高かった手法 (d) を用いた場合, 100 文中 28 文が格助詞の変化に失敗した. 内訳を表 8 に示す.

表 8: 変化規則失敗の内訳

分類	原因	頻度
1	構文変形パターン対を一意に絞りこめない	8
2	対応する構文変形パターン対が存在しない	8
3	用言意味属性制約付きパターン対が存在しない	6
4	用言意味属性による制約の失敗	3
5	結合価パターンによる制約の失敗	3

分類 1, 2, 3, 4, 5 について, 以下の節でそれぞれ例をあげて説明する.

7.1 構文変形パターン対を一意に絞りこめない

分類 1 の構文変形パターン対を一意に絞りこめなかった文は 8 文存在した. 以下に絞りこめなかった 8 文を示す.

- その時期に何かのパーティーが催されるようだ
- 軍隊は戦闘隊形に配置された
- ガソリンはさまざまな化合物から調製される
- 車はその爆発でバラバラに吹き飛ばされた
- その声明は各紙で大々的に報道された
- 離婚判決は公開法廷で宣言されなくてはならない
- 彼女の本は世界各地で読まれている
- その議案は票決に付された

絞りこめなかった原因について以下に考察する.

7.1.1 入力文：その時期に何かのパーティーが催されるようだ

入力文：その時期に何かのパーティーが催されるようだ

上記の入力文に対して、構文変形パターン対を用いることで得られる候補文を以下に示す。

候補文1：その時期が何かのパーティーを催す

候補文2：その時期に何かのパーティーを催す

この2つの候補文に対し、吉田のアルゴリズム [6] を適用すると両文とも以下のパターンが最も一致する。

{主体} が {式・行事等, 言動} を催す

吉田のアルゴリズムでは、候補文1, 2はともに同じ一致度を出力する。これは両候補文がヲ格の格要素のみが結合価パターンと一致し、候補文1ではガ格の格要素が異なり、候補文2では二格の格要素に関する制約が無いと判別するためである。そのため、これ以上の絞り込みを行うことができない。

7.1.2 入力文：軍隊は戦闘陣形に配置された

入力文：軍隊は戦闘陣形に配置された

上記の入力文に対して、構文変形パターン対を用いることで得られる候補文を以下に示す。

候補文1：戦闘陣形が軍隊を配置する

候補文2：軍隊を戦闘陣形に配置する

この2つの候補文に対し、吉田のアルゴリズム [6] を適用すると両文とも以下のパターンが最も一致する。

{主体} が {主体} を {場所, 場, 人<職業, 地位, 役割>} に配置する

吉田のアルゴリズムでは、候補文1, 2はともに同じ一致度を出力する。これは両候補文がヲ格の格要素のみが結合価パターンと一致し、候補文1ではガ格の格要素が異なり、候補文2では二格の格要素が異なると判別するためである。そのため、これ以上の絞り込みを行うことができない。

7.1.3 入力文：ガソリンはさまざまな化合物から調製される

入力文：ガソリンはさまざまな化合物から調製される

上記の入力文に対して、構文変形パターン対を用いることで得られる候補文を以下に示す。

候補文1：ガソリンをさまざまな化合物から調製する

候補文2：ガソリンをさまざまな化合物で調製する

この2つの候補文に対し、吉田のアルゴリズム [6] を適用すると両文とも以下のパターンが最も一致する。

{ 主体 } が { 人工物 } を調製する

吉田のアルゴリズムでは、候補文1, 2はともに同じ一致度を出力する。これは両候補文がヲ格の格要素のみが結合価パターンと一致し、その他の格要素に関する制約が無いと判別するためである。そのため、これ以上の絞り込みを行うことができない。

7.1.4 入力文：車はその爆発でバラバラに吹き飛ばされた

入力文：車はその爆発でバラバラに吹き飛ばされた

上記の入力文に対して、構文変形パターン対を用いることで得られる候補文を以下に示す。

候補文1：車をその爆発がバラバラに吹き飛ばす

候補文2：車をその爆発でバラバラに吹き飛ばす

この2つの候補文に対し、吉田のアルゴリズム [6] を適用すると両文とも以下のパターンが最も一致する。

{ 風 } が { 具体 } を吹き飛ばす

吉田のアルゴリズムでは、候補文1, 2はともに同じ一致度を出力する。これは両候補文がヲ格の格要素のみが結合価パターンと一致し、候補文1ではガ格の格要素が異なり、候補文2ではデ格の格要素に関する制約が無いと判別するためである。そのため、これ以上の絞り込みを行うことができない。

7.1.5 入力文：その声明は各紙で大々的に報道された

入力文：その声明は各紙で大々的に報道された

上記の入力文に対して、構文変形パターン対を用いることで得られる候補文を以下に示す。

候補文1：その声明を各紙に大々的に報道する

候補文2：その声明を各紙で大々的に報道する

この2つの候補文に対し、吉田のアルゴリズム [6] を適用すると両文とも以下のパターンが最も一致する。

{ 主体 } が { 抽象 } を { 主体 } に報道する

吉田のアルゴリズムでは、候補文1, 2はともに同じ一致度を出力する。これは両候補文がヲ格の格要素のみが結合価パターンと一致し、候補文1では二格の格要素が異なり、候補文2ではデ格の格要素に関する制約が無いと判別するためである。そのため、これ以

上の絞り込みを行うことができない。

7.1.6 入力文：離婚判決は公開法廷で宣言されなくてはならない

入力文：離婚判決は公開法廷で宣言されなくてはならない

上記の入力文に対して、構文変形パターン対を用いることで得られる候補文を以下に示す。

候補文 1：離婚判決を公開法廷で宣言する

候補文 2：離婚判決を公開法廷に宣言する

この 2 つの候補文に対し、吉田のアルゴリズム [6] を適用すると両文とも以下のパターンが最も一致する。

{ 主体, 命令, 論理, 情報 } が { 名詞 } を宣言する

吉田のアルゴリズムでは、候補文 1, 2 はともに同じ一致度を出力する。これは両候補文がヲ格の格要素のみが結合価パターンと一致し、その他の格要素に関する制約が無いと判別するためである。そのため、これ以上の絞り込みを行うことができない。

7.1.7 入力文：彼女の本は世界各地で読まれている

入力文：彼女の本は世界各地で読まれている

上記の入力文に対して、構文変形パターン対を用いることで得られる候補文を以下に示す。

候補文 1：彼女の本を世界各地が読む

候補文 2：彼女の本を世界各地で読む

候補文 3：彼女の本を世界各地に読む

この 3 つの候補文に対し、吉田のアルゴリズム [6] を適用すると 3 文とも以下のパターンが最も一致する。

{ 主体 } が { 抽象物 (精神), 精神 } を読む

吉田のアルゴリズムでは、候補文 1, 2, 3 はすべて同じ一致度を出力する。これは全ての候補文がヲ格の格要素のみが結合価パターンと一致し、候補文 1 ではガ格の格要素が異なり、その他の候補文では格要素に関する制約が無いと判別するためである。そのため、これ以上の絞り込みを行うことができない。

7.1.8 入力文：その議案は票決に付された

入力文：その議案は票決に付された

上記の入力文に対して、構文変形パターン対を用いることで得られる候補文を以下に示す。

候補文1：その議案を票決が付す

候補文2：その議案を票決で付す

候補文3：その議案を票決に付す

この3つの候補文に対し、吉田のアルゴリズム [6] を適用すると3文とも以下のパターンが最も一致する。

{ 主体 } が { 抽象 } を不問に付す

吉田のアルゴリズムでは、候補文1, 2, 3はすべて同じ一致度を出力する。これは全ての候補文がヲ格の格要素のみが結合価パターンと一致し、候補文1ではガ格の格要素が異なり、その他の候補文では格要素に関する制約が無いと判別するためである。そのため、これ以上の絞り込みを行うことができない。

7.1.9 構文変形パターン対を一意に絞りこめない原因と考察

これらの絞り込みの失敗は、文の意味的結合性を結合価パターンのみで判別し、最も結合価パターンに一致している文を正解としていることが大きな原因と考えられる。この問題に対しては、結合価パターンに依存しない格助詞の変化規則を組み合わせることで改善できると考えられる。例えば、入力文の格要素 { 戦闘陣形に } の格助詞二は副詞的用法で用いられている。そこで、「副詞的用法の格助詞二は変化しない」などの規則を加えることで対処できると考えられる。

7.2 対応する構文変形パターン対が存在しない

分類2の対応する構文変形パターン対が存在しなかった文は8文存在した。以下に存在しなかった8文を示す。

- プリンセスの流産は一部のタブロイド紙ではおいしいニュースとして扱われた
- 先祖伝来の土地が世代から世代へと引きつがれた
- 円高ドル安によりアメリカの強気の債券市場の利益は一掃された
- ギリシャ人の間では俳優は立派な職業と見なされていた

- さし絵に対しても同様多くの注意が払われている
- 九竜半島は1860年にイギリスに割譲された
- 年にこの会社に雇われた
- 夜6時から朝6時まで夜間外出禁止令がしかれている

原因について，以下の入力文を例に考察する．

入力文：先祖伝来の土地が世代から世代へと引きつがれた

上記の入力文から格助詞と用言の構造を抽出すると次の文型パターンが得られる．

文型パターン： $N1$ が $N2$ から $N3$ へ引きつがれる

しかし，作成した構文変形パターン対の中に上記の文型パターンをもつパターン対が存在しないため，出力文を得ることができなかった．この問題は，構文変形パターン対を拡充することで改善できると考えられる．

7.3 用言意味属性制約付きパターン対が存在しない

分類3の用言意味属性制約付きパターン対が存在しなかった文は6文存在した．以下に存在しなかった6文を示す．

- 銃は正規の方法で保管されていた。
- それは贈り物として大いに喜ばれるであろう
- 時計つきラジオには目覚ましが入蔵されている
- われわれ人間の利益のために何千ものこれらの動物が殺害されている
- 人の応募者の中から選ばれた
- ゴリアテは一騎打ちでダビデに殺された

原因について，以下の入力文を例に考察する．

入力文：銃は正規の方法で保管されていた

上記の入力文に対して，汎用構文変形パターン対を用いることで得られる3つの候補文を以下に示す．

候補文1：正規の方法が銃を保管する

候補文 2 : 銃を正規の方法で保管する

候補文 3 : 銃を正規の方法に保管する

ここで，用言 { 保管する } は用言意味属性 { 所有 } を持つことから，{ 所有 } の制約を持つ用言意味属性制約付きパターン対の検索を行う．

変形構文パターン	基本構文パターン	用言意味属性
$N1$ は $N2$ で $V.rareru$	$N2$ が $N1$ を V	自然現象
$N1$ は $N2$ で $V.rareru$	$N2$ が $N1$ を V	行動
$N1$ は $N2$ で $V.rareru$	$N1$ を $N2$ で V	属性
$N1$ は $N2$ で $V.rareru$	$N1$ を $N2$ で V	身体動作
$N1$ は $N2$ で $V.rareru$	$N1$ を $N2$ に V	生成
⋮	⋮	⋮

しかし，入力文に対応する文構造の { 所有 } の制約を持つパターン対が存在しないため，出力文を得ることができなかった．この問題も 7.2 節の問題と同様に，構文変形パターン対を拡充することで改善できると考えられる．

7.4 用言意味属性による制約の失敗

分類 4 の用言意味属性による制約が失敗した文は 3 文存在した．以下に失敗した 3 文を示す．

- ゲリラ活動が国境近くでにわかには再開された
- その村はゲリラによって攻撃された
- 正式に両親から勘当された

失敗した原因についてそれぞれ以下に考察する．

7.4.1 入力文：ゲリラ活動が国境近くでにわかには再開された

入力文：ゲリラ活動が国境近くでにわかには再開された

上記の入力文に対し，用言意味属性付き汎用構文変形パターン対を用いることで得られる出力文と，正解文を以下に示す．

出力文：国境近くがゲリラ活動をにわかには再開する

正解文：ゲリラ活動を国境近くでにわかには再開する

入力文の用言意味属性を満たす誤った構文変形パターン対が存在したため、正解とは異なる文を出力してしまった。

7.4.2 入力文：その村はゲリラによって攻撃された

入力文：その村はゲリラによって攻撃された

上記の入力文に対し、用言意味属性付き汎用構文変形パターン対を用いることで得られる出力文と、正解文を以下に示す。

出力文：その村にゲリラが攻撃する

正解文：その村をゲリラが攻撃する

入力文の用言意味属性を満たす誤った構文変形パターン対が存在したため、正解とは異なる文を出力してしまった。

7.4.3 入力文：正式に両親から勘当された

入力文：正式に両親から勘当された

上記の入力文に対し、用言意味属性付き汎用構文変形パターン対を用いることで得られる出力文と、正解文を以下に示す。

出力文：正式に両親から勘当する

正解文：正式に両親が勘当する

入力文の用言意味属性を満たす誤った構文変形パターン対が存在したため、正解とは異なる文を出力してしまった。

7.5 結合価パターンによる制約の失敗

分類5の結合価パターンによる制約が失敗した文は3文存在した。以下に失敗した3文を示す。

- あなたの画家としての才能は絶え間ない修練によってのみ磨かれるだろう
- あの場面は映画の検閲官に許可されなかった
- それは第4条で規定されている

失敗した原因についてそれぞれ以下に考察する。

7.5.1 入力文：あなたの画家としての才能は絶え間ない修練によってのみ磨かれるだろう

入力文：あなたの画家としての才能は絶え間ない修練によってのみ磨かれるだろう
上記の入力文に対し構文変形パターン対を用いた格助詞の変化を行ない，結合価パターンによる制約を用いることで得られる出力文と，正解文を以下に示す．

出力文：あなたの画家としての才能を絶え間ない修練が磨く

正解文：あなたの画家としての才能を絶え間ない修練によって磨く

出力文は意味的にはおかしくないが，主体が擬人化されており，入力文の意味と異なっている．

7.5.2 入力文：あの場面は映画の検閲官に許可されなかった

入力文：あの場面は映画の検閲官に許可されなかった

上記の入力文に対し構文変形パターン対を用いた格助詞の変化を行ない，結合価パターンによる制約を用いることで得られる出力文と，正解文を以下に示す．

出力文：あの場面を映画の検閲官に許可する

正解文：映画の検閲官があの場面を許可する

出力文は意味的にはおかしくないが，入力文の意味と異なっている．

7.5.3 入力文：それは第4条で規定されている

入力文：それは第4条で規定されている

上記の入力文に対し構文変形パターン対を用いた格助詞の変化を行ない，結合価パターンによる制約を用いることで得られる出力文と，正解文を以下に示す．

出力文：それを第4条が規定する

正解文：それを第4条で規定する

出力文は意味的にはおかしくないが，主体が擬人化されており，入力文の意味と異なっている．

8 まとめ

本稿では，文型パターンを用いた構文変形に伴う格助詞の変化規則を提案し，効果を検証した．具体的には，日英対訳データベースから抽出した単文の受動文 633 文から，受動文と対応する能動文の格助詞と用言の文構造を抽出した構文変形パターン対を作成し，その適用判定の手法を提案した．その結果，固有構文変形パターン対を 554 個，汎用構文変形パターン対を 153 個，用言意味属性制約付き汎用構文変形パターン対を 373 個作成することができた．また効果を検証するため，日英対訳データベースから抽出した単文の受動文 100 文に対して構文変形パターン対を用いた格助詞の変化規則を適用した．この結果，受動文の約 72% を結合価パターンに最も一致する能動文にすることができた．今後は，構文変形パターン対を拡充すると共に結合価パターンに依存しない格助詞の変化規則を作成する予定である．

謝辞

本論文作成に際して、多大なる検討と助言をしてくださった池原悟教授ならびに村上仁一助教授、徳久雅人助手そして池原研究室の方々に対し深く感謝します。altjaws は NTT との共同研究の下に使わせて頂きました。

また、参考にさせて頂いた文献の著者の方々に対しても感謝します。

参考文献

- [1] 池原, 宮崎, 白井, 横尾, 中岩, 小倉, 大山, 林: 日本語語彙体系, 岩波書店 (1997)
- [2] 仁田, 山梨, 城田, 定延, まつもと, 金田, 金水, 塚本: 日本語の格をめぐって, くろしお出版 (1993)
- [3] 仁田, 村木, 竹沢, 外池, 益岡, 定延, 井島, 天野, 野本, 杉本: 日本語のヴォイスと他動性, くろしお出版 (1991)
- [4] 寺村: 日本語のシンタクスと意味, くろしお出版 (1982)
- [5] 村上, 池原, 徳久: 日本語英語の文対応の対訳データベースの作成, 「言語、認識、表現」第7回年次研究会 (2002)
- [6] 吉田, 池原, 村上: 入力文に対する結合価パターン対の選択方法について, 言語処理学会第8回年次大会, pp299-302(2002)