

受動態に変換した結合価パターン辞書

福田 泰介[†] 徳久 雅人^{††} 村上 仁一^{††} 池原 悟^{††}

[†] 鳥取大学工学部知能情報工学科

^{††} 鳥取大学大学院工学研究科情報エレクトロニクス専攻

〒 680-8552 鳥取県鳥取市湖山町南 4-101

E-mail: [†]s062048@ike.tottori-u.ac.jp, ^{††}tokuhisa@ike.tottori-u.ac.jp

あらまし 本稿では、日本語語彙大系における結合価パターンを能動態から受動態へと変換する方法を示す。日本語では能動態から受動態へと変換する際、動詞ごとに助詞の変換内容が異なるが、ある程度の法則性があると言われている。そこで、結合価パターンを変換する規則を作成し、日本語語彙大系に適用することを試みる。9件の変換規則を適用した結果 11,310 件のパターンを変換することができた。受動文から能動文に変換する実験において、変換の精度が約 5 割であった。誤り分析によると、動作主、対象物、あるいは手段を表すための多様な格助詞を網羅するという問題、および、能動態では必須ではなくとも受動態では必須となる格要素を追加するという問題があることがわかった。キーワード 受動態、受動パターン、態の変換、パターンの汎化

Valency Pattern Dictionary Converted to Passive Voice

Taisuke FUKUDA[†], Masato TOKUHISA^{††}, Jin'ichi MURAKAMI^{††}, and Satoru IKEHARA^{††}

[†] Department of Information and Knowledge Engineering, Faculty of Engineering, Tottori University

^{††} Department of Information and Electronics, Graduate School of Engineering, Tottori University

4-101, Koyama-Minami, Tottori, 680-8552, Japan

E-mail: [†]s062048@ike.tottori-u.ac.jp, ^{††}tokuhisa@ike.tottori-u.ac.jp

Abstract This paper proposes a method to convert valency patterns in A-Japanese-Lexcon to the passive voice. It is said that on converting the voice in Japanese the tendency of the change are observed, while the postposition to be changed depends on the verb. Therefore, we try to convert the valency patterns from the active voice to the passive voice based on rules. As the result of applying the 9 rules, 11,310 patterns are acquired. In the experiments of the conversion, the accuracy is 50%. According to the error analysis, we find two major problems. One is to cover the diversity when postposition suggests agent/target/method-case. The other is to add case elements to original valency patterns, if the case elements are necessary on the passive voice while they are not necessary on the active voice.

Key words passive voice, passive pattern, voice change, pattern generalization

1. はじめに

文型パターンを用いた言語表現の解析において、パターンの網羅性を高めるために、パターンで使用される格助詞や様相表現に自由度を持たせる必要がある [1]。とりわけ、受動態と使役への対策には格助詞の変形が伴うため、単なる過去形や推量の表現に対応するよりも対策が複雑である。

ここで、日本語の単文について、代表的な文型パターン辞書として日本語語彙大系が挙げられる [2]。これは多段翻訳方式において日本語の命題表現をカバーするためのものであり、原則としてパターンは能動態で記述されている (図 1)。

結合価パターン: $N1$ が $N2$ を 愛する
意味属性制約: $N1(4 \text{ 人}) N2(*)$

図 1 結合価パターンの例

そのため、受動態の文に直接的には対応しておらず、パターン外の処理において対処する必要がある。ところが、受動態に対処する方法は [2] において具体的に示されておらず、この辞書の利用者が各自で対処しなければならないという問題がある。

受動文を結合価パターンと適合させる方法として、受動文の格助詞に対応した結合価パターンを用意する方法と、パターン照合の際に入力文を書き換えて結合価パターンの格助詞と対応

させる方法の2つがある。

ここで、寺村の考察 [3] を参照すると、『日本語において、格の移動により能動文と受動文が対応する以上、一定の規則で表すことのできる変化であるが、動詞の性質や名詞の性質によって規則の適用される範囲は限定的である』と述べている。この考察は、動詞ごとに格助詞変化があるのだが、ある程度の規則性があると解釈できる。つまり、動詞ごとに受動態のパターンを用意すべきであるが、その多くは規則的に能動態のパターンから作成できるかもしれないといえる。そこで、本稿では、受動文の格助詞に対応した結合価パターンを能動態のパターンごとに用意するという前者の方法を試みる。

2. 能動パターンから受動パターンへの変形

能動態で書かれた文を能動文と呼ぶことから、本稿では、能動態で書かれたパターンを能動パターン、同様に、受動態で書かれたパターンを受動パターンと呼ぶことにする。

本稿では、能動態から受動態への変換規則を使用し、日本語語彙大系における動詞の能動パターンのそれぞれに対して、受動パターンを作成する。

2.1 変換規則

変換規則は、[3] [4] をもとに作成した表 1 を使用する。表中の「*A*, *B*, *C*」は格要素の名詞変数、「*V*」は述語となる動詞変数、「*.reru*」は受動表現の接辞を意味している。

表 1 変換規則

典拠	条件パターン	帰結パターン
庭氏, 寺村	<i>A</i> が <i>V</i>	<i>A</i> に <i>V.reru</i>
	<i>A</i> が <i>B</i> と <i>V</i>	<i>A</i> に <i>B</i> が <i>V.reru</i>
	<i>A</i> が <i>B</i> に <i>C</i> を <i>V</i>	<i>A</i> に/から <i>B</i> が/に <i>C</i> を/が <i>V.reru</i>
	<i>A</i> が <i>B</i> を <i>C</i> に <i>V</i>	<i>A</i> に/から <i>B</i> を/が <i>C</i> が/に <i>V.reru</i>
	<i>A</i> が <i>B</i> に <i>V</i>	<i>A</i> に <i>B</i> が <i>V.reru</i>
	<i>A</i> が <i>B</i> の <i>C</i> を <i>V</i>	<i>A</i> に <i>B</i> が <i>C</i> を <i>V.reru</i>
	<i>A</i> が <i>B</i> を <i>C</i> と <i>V</i>	<i>A</i> に <i>B</i> が <i>C</i> と <i>V.reru</i>
	<i>A</i> が <i>B</i> を <i>V</i>	<i>A</i> に <i>B</i> が/を <i>V.reru</i>
寺村	<i>A</i> が <i>B</i> を <i>C</i> から <i>V</i>	<i>A</i> に <i>B</i> が/を <i>C</i> から <i>V.reru</i>

2.2 変換規則の運用

能動パターンを条件パターンと照合し、適合する場合に、帰結パターンを用いて受動パターンを生成する。生成した受動パターンの例を以下に示す。

能動パターン: *N1* が *N2* に *N3* を伝える
 受動パターン: *N1* から/に *N2* に/が *N3* を/が ‘伝える’.*reru*
 意味属性制約: *N1*(3 主体) *N2*(3 主体) *N3*(“ 悲報/訃報 ”)

上記の能動パターンは、一見すると「*A* が *V*」「*A* が *B* に *V*」「*A* が *B* に *C* を *V*」の要素を持つ3つの条件パターンに適合しうる。ここでは、3つの条件パターンの中で要素の並びが最も多く適合する「*A* が *B* に *C* を *V*」という条件パターンを適合し、受動パターンを生成する。また、意味属性制約は受動パターンにも対応している。

一方で、能動パターンの格要素が条件パターンの格要素よりも多い場合、適合部分のみから帰結パターンに従い格助詞を変換し、受動パターンを作成する。例えば、以下に示す能動パ

ターンでは、格助詞が「が」「を」「から/より」「に/へ/まで」の並びとなっている。この場合は、「*A* が *B* を *C* から *V*」に対する条件パターンが該当し、受動パターンが作成される。

能動パターン: *N1* が *N2* を *N3* から/より *N4* に/へ/まで移転する
 受動パターン: *N1* に *N2* が/を *N3* から/より *N4* に/へ/まで
 移転する'.*reru*

日本語語彙大系の動詞の能動パターン 12,297 件に対して、表 1 を使用したところ、11,310 件の受動パターンが得られた。

3. 受動文の解析

受動パターンを使用し、受動文の意味解析を行う手順を以下に示す。

- (1) 入力文(受動文)を形態素解析
- (2) 文法レベルで入力文と受動パターンを照合
- (3) 照合結果に意味属性制約をかけて候補選択

この結果パターンと入力文が対応する。パターンに基づく意味解析では、「パターンが意味に対応している」とみなすので、(3)の完了により意味解析が完了となる。また、次の手順を踏むことで、機械的に受動文を能動文へと変換することができる。

- (4) 受動パターンと対になる能動パターンを用いて文生成

4. 予備実験

能動文と受動文が格の移動で対応することから、受動文と能動文の対応を評価することで、変換規則により作成された受動パターンの正しさを判断することができると考えられる。そこで [3] より抽出した受動文を第 3 章の手順で解析し、能動文へ意味的に正しく復元できるかを評価することで、変換規則による受動パターンの作成が有効であるかを検証する。ただし、変換規則の正しさを調べるのが目的であるので、第 3 章の手順 (3) における複数候補の選択の性能は問わないこととする。

4.1 評価方法

第 3 章で述べた処理方法にて、入力文を受動文から能動文に復元した結果を、以下の方法で評価する。

- 受動パターンが入力文の格要素を満たしているか評価
- 能動文に変換し、意味的に正確であるかを評価

また、1つの入力文に対して、複数の候補が得られるが、そのうち1つでも、上記の2つの評価を満たした受動パターンがあれば、変換規則は正しく適用されたと判断する。

4.2 入力文

予備実験では [3] より、受動態の例文 57 文を入力文として使用する [3] では例文がカタカナで表記されている場合があるが、形態素解析を正しく行うため、ひらがな表記に直して入力文として使用する。

彼八刑事二怪シマレタ。 彼は刑事に怪しまれた。

また、基本的な受動文を得るために、受動文の文末を「れる」「られる」に加工して実験に使用する。

直孝は祖母に育てられた。 直孝は祖母に育てられる。
 多くの人に愛されている。 多くの人に愛される。

4.3 実験結果

入力文 57 文に対し、37 文が受動文から能動文に復元するこ

とができ、カバー率は65%(37文/57文)となった。表2に復元に失敗した原因と頻度を示す。

表2 変換失敗の内訳

分類	原因	頻度
1	形態素解析の誤りがある	1
2	受動パターンの助詞が不足している	15
3	格要素の意味属性制約が満たされていない	3
4	格要素の個数に変化を要する	1

4.4 誤り分析

誤り事例を分類ごとに分析する。

4.4.1 分類1「形態素解析の誤り」

以下に誤り事例を示す。

入力文：日本の領空は外国機に飛ばれる。

復元文：外国機に [が日本の領空を から/より へ/に/まで 飛ば]。

正解文：外国機が日本の領空を飛ば。

復元文の括弧内は、受動パターンが適合した範囲を表す「 ϕ 」は結合価パターンの格要素に名詞が照合されていないことを表す。

入力文の「外国機」は、能動文でのガ格の名詞として解析されるべきである。しかし、実際の出力ではそのようになっていない。そこで、使用したパターンを参照したところ、次の通りであった。

能動パターン：N1がN2をN3から/よりN4へ/に/まで飛ば
 受動パターン：N1にN2がN3から/よりN4へ/に/まで「飛ば」.reru
 意味属性制約：N1(3主体 536動物(個体)990乗り物(本体(移動(空圏)))
 N2(388場所 2610場)N3(388場所 2610場)
 N4(388場所 2610場)

上記に示される助詞の対応では、入力文の「外国機」は受動パターンの変数N1またはN4が選択候補となり、意味属性制約により変数N1の名詞として適合されるべきである。しかし、名詞の形態素解析の結果を参照したところ、次の通りとなっていた。

/外国(1100,NI:385,NI:463)
 +機(6280,NI:2595,NI:2588)

「機」は接尾辞として解析されており、一般意味属性番号の2595, 2588が付与された。これらは「単位」、「個数・回数等」を意味する。したがって、「飛行機」という意味で解析されておらず、形態素解析における名詞意味属性の付与誤りであることがわかった。これは変換規則では対応できない。

4.4.2 分類2「受動パターン」の助詞が不足

以下に誤り事例を示す。

(2a)

入力文：彼は刑事から怪しまれる。

復元文：刑事から [が彼を怪しむ]。

正解文：刑事が彼を怪しむ。

以下に使用されたパターンを示す。

能動パターン：N1がN2を怪しむ
 受動パターン：N1にN2が/を「怪しむ」.reru
 意味属性制約：N1(4人)N2(*)

使用された受動パターンの変数N1とN2の助詞には、助詞「から」がどちらにも存在しない。そのため、名詞「刑事」が受動パターンの名詞変数と照合されなかった。

入力文が受動パターンと正しく対応するためには、「N1に」ではなく「N1から」が受動パターンに必要である。同様に誤り事例を示す。

(2b)

入力文：金閣は義満によって建てられる。

出力文：義満によって [が金閣を に/へ 建てる]。

正解文：義満が金閣を建てる。

以下に使用したパターンを示す。

能動パターン：N1がN2をN3に/へ建てる
 受動パターン：N1から/にN2が/をN3が/に/へ「建てる」.reru
 意味属性制約：N1(3主体)N2(389施設 938像・書画 863建造物 928碑)
 N3(388場所 2610場)

(2a)と同様に、受動パターンの変数の助詞に、複合助詞「によって」が存在しない。そのため、名詞「義満」が受動パターンの名詞変数と照合されなかった。入力文が受動パターンと正しく対応するためには、「N1に」ではなく「N1によって」が受動パターンに必要である。

以上(2a)と(2b)により、表1の変換規則に改良するべき点が示された。

4.4.3 分類3「格要素の意味属性制約が満たされていない」

以下に誤り事例を示す。

入力文：家の流儀が当主から次に当主となる子に受け継がれる。

出力文：家の流儀が [次に当主となる子が を当主から/より 受け継ぐ]。

正解文：次に当主となる子が家の流儀を当主から受け継ぐ。

以下に使用したパターンを示す。

能動パターン：N1がN2をN3から/より受け継ぐ
 受動パターン：N1にN2が/をN3から/より「受け継ぐ」.reru
 意味属性制約：N1(3主体)N2(362組織 1936仕事)N3(3主体)

入力文を意味的に正しく能動文へ変換するためには、名詞句「家の流儀」が受動パターンのN2に適合しなければならない。ここで、「家の流儀」の形態素解析結果を参照したところ、次の通りであった。

/家(1100,NI:447,NI:865,NI:73,NI:387)
 +の(7410)
 /流儀(1100,NI:1222,NI:1035)

名詞「流儀」の意味属性番号1222, 1035は「流儀」、「方法」を意味する名詞として正しく解析されている。しかし、受動パターンの変数N2は意味属性制約により「流儀」を選択することができなかった。これは表1の規則の問題ではなく、日本語彙大系における問題である。なお、日本語彙大系の一般名詞意味属性に対する制約条件の変更は、本稿では対象としないため、対策を設けない。

4.4.4 分類4「格要素の個数に変化を要する」

以下に誤り事例を示す。

入力文：アーサー王子が両親をラビック王に殺される。

出力文：アーサー王子が [ラビック王が 両親を 殺す]。

正解文：ラビック王がアーサー王子の両親を殺す。

ここで使用したパターンは次の通りである。

能動パターン：N1 が N2 を殺す
受動パターン：N1 に N2 が/を '殺す'.reru
意味属性制約：N1(4人 535 動物)N2(4人 535 動物)

入力文の格要素として考えられる名詞は「アーサー王子」, 「(アーサー王子の) 両親」, 「ラビック王」の3つである。しかし, 動詞「殺す」の全ての能動パターンは, 2つの格要素しか選択できない。これは, そのため, 入力文の格要素を, 能動文として意味的に正しく変換できなかつた。

仮に, 入力文を正しく能動文に変換するならば, 日本語語彙大系に以下の結合価パターンが必要になると考えられる。なお, 名詞変数 N3 を囲む「#n[]」は, 任意のパターン記述要素を表す記号である。

能動パターン：N1 が #1[N3 の] N2 を殺す
受動パターン 1：N1 に N3 が N2 を '殺す'.reru
受動パターン 2：N1 に N3 の N2 が '殺す'.reru
意味属性制約：N1(4人 535 動物)N2(4人 535 動物)N3(*)

しかし, 能動パターンの変形までは本稿の対象としないため, 本稿ではこの誤り事例に対策を設けない^(注1)。

5. 変換規則の改良

本章では, 予備実験において受動パターンがカバーできていなかった「から」と「によって」に対応させるために, 変換規則の改良を行う。

5.1 改良の方針

- 助詞「から」

表1の変換規則で作成した受動パターンは「から」が含まれる受動文と適合させる場合, 以下のパターンでしか対応できていなかった。

能動パターン：A が B に C を V
受動パターン：A に/から B が/に C を/が V.reru
能動パターン：A が B を C から V
受動パターン：A に B が/を C から V.reru

そこで, 他の受動パターンに対して「から」を付与するため, 次の基本的な構文を考える。

X が Y を V

この構文にて [3] では『V が感情・感覚・思考の動きを表す動詞をとる場合, Y を主役に立てた受動表現では, 感情や感覚が X から発して Y に到ると, 通常見立てることができる。思考の動きを表す動詞も同様に考えることができる』と述べられている [3] を踏まえ,

(注1): この事例より類推すると次のような受動文が考えられる: 「この象は鼻が覆められた。」 「この象は鼻を覆められた。」

感情・感覚・思考の動きを表す動詞が含まれる受動パターンで, 動作主となる格要素に格助詞「から」を追加する。

- 複合助詞「によって」

受動パターンは入力文に複合助詞「によって」が使用されている場合, 全く対応できていない。入力文に「によって」が付属する名詞は能動文において主格となっている。そのため, 受動パターンにおいて動作主となる格要素に複合助詞「によって」を追加する。

5.2 改良を行った変換規則

改良を行った変換規則を表3に示す。

表3 改良した変換規則

条件パターン	帰結パターン
A が V	A に/によって/から* V.reru
A が B と V	A に/によって/から* B が V.reru
A が B に C を V	A に/から/によって B が/に C を/が V.reru
A が B を C に V	A に/から/によって B を/が C が/に V.reru
A が B に V	A に/によって/から* B が V.reru
A が B の C を V	A に/によって/から* B が C を V.reru
A が B を C と V	A に/によって/から* B が C と V.reru
A が B を V	A に/によって/から* B が/を V.reru
A が B を C から V	A に/によって B が C から V.reru

助詞「から*」は V が「感情動作・知覚動作・思考動作」となるパターンのみ有効である。この判別には, 日本語語彙大系の結合価パターンにおける用言意味属性を参照する。

5.3 改良の効果

表3の変換規則を使用し実験を行った結果, 入力文57文に対し, 49文を正しく能動文に復元することができ, 分類2の問題が改善されたことがわかる。カバー率は86%(49文/57文)に向上した。

表4 変換失敗の内訳

分類	原因	頻度
1	形態素解析の誤りがある	1
2	受動態パターンの助詞が不足している	3
3	格要素の意味属性制約が満たされていない	3
4	格要素の個数に変化を要する	1

6. 評価実験

受動パターンの網羅性と正確性を検証するために評価実験を行う。

6.1 評価方法

評価手順を以下に示す。

- (1) 受動パターンが入力文の格要素を満たしているか評価
- (2) 能動文が意味的に正確であるか評価

以上の評価を満たした出力があれば, 受動パターンが正しく作成されていると見なし, 正解とする。変換規則により作成した受動パターンのカバー率と精度を以下の方法で評価する。

6.1.1 カバ ー 率

受動パターンが入力文に対応しているかを判断する際に、入力文に対して複数の能動文を出力する場合がある。出力のうち1つでも正解となる出力が得られた場合、受動パターンが入力文を網羅していると考え、以下の式でカバー率を求める。

$$\text{カバー率} = \frac{\text{1つでも正解となる出力が得られた入力文数}}{\text{全入力文数}} \times 100[\%]$$

6.1.2 精 度

以下の式で作成した受動パターンの精度を求める。

$$\text{精度} = \frac{\text{正解と判断した出力件数}}{\text{全出力件数}} \times 100[\%]$$

6.2 入 力 文

本実験では、一般的な学習用辞書から収集した単文 [5] の受動文 100 文を使用する。以下に例を示す。

- 4月に新しいお札が発行される。
- 泥棒は近所の人に捕まえられた。

6.3 実 験 結 果

入力文 100 文のうち、56 文が意味的に正しく受動パターンで判断できた (カバー率 56%)。また、入力文 100 文に対し、157 件の出力候補が得られ、62 件が意味的に正しかった (精度 40%)。

6.4 誤 り 分 析

本実験において能動文への出力が失敗したもの、および、意味的に正しいと判断できなかった出力の内訳を表 5 に示す。

表 5 変換失敗の内訳

分類	原因	頻度
1	形態素解析の誤りがある	6
2	適切な受動パターンが不足している	1
3	受動パターンの助詞が不足している	11
4	格要素の意味属性制約が満たされていない	9
5	複数の候補からの選択に誤りがある	10
6	格要素の個数に変化を要する	5
7	曖昧性による問題がある	1
8	その他	1

予備実験では見られなかった新たなタイプの誤り事例について分析を行う。

6.4.1 分類 2 「適切な受動パターンが不足している」

以下に誤り事例を示す。

入力文：その囚人は終身刑を言い渡された。

出力文：出力なし

正解文：その囚人に終身刑を言い渡す。

入力文に対して出力文を得ることができなかった。「言い渡す」の受動パターンを参照したところ、次の通りであった。

能動パターン：N1 は N2 を言い渡す
意味属性制約：N1(4 人)N2(“ 判決 ”)

名詞変数 N1 の助詞が「は」となっており、表 3 の条件パターンに該当しないため、受動パターンが作成されていなかった。

また「言い渡す」という能動パターンには、入力文の「囚人」に対応する格要素が存在しない。正解文の出力を得るためには、「囚人」に対応する格要素を追加した以下に示すパターン対が

必要となり、「分類 6、格要素の個数に変化を要する」問題へと転化する。

能動パターン：N1 は N2 を₁[N3 に/へ] 言い渡す
受動パターン：N1 から/によって N2 が/を N3 が/に/へ
‘言い渡す’.reru
意味属性制約：N1(4 人)N2(“ 判決 ”)N3(*)

6.4.2 分類 3 「受動パターンの助詞が不足している」

第 5 章の予備実験の結果より、変換規則に「から」、「によって」を追加したが、それでも助詞が不足した。以下に誤り事例を示す。

(3a)

入力文：あの大リーガーの場外ホームーには度胆を抜かれた。

出力文：あの大リーガーの場外ホームーには [が の度胆を 抜く]。

正解文：あの大リーガーの場外ホームーが度胆を抜く。

以下に使用したパターンを示す。

能動パターン：N1 が N2 の度胆を抜く
受動パターン：N1 に/によって/から N2 が度胆を ‘抜く’.reru
意味属性制約：N1(*)N2(4 人)

使用した受動パターンの変数 N1 には、助詞「には」が存在しない。そのため、名詞句「あの大リーガーの場外ホームー」が受動パターンの名詞変数 N1 に適合しなかった。入力文を正しく解析するには、受動パターンを「N1 には」とする必要がある。

次に (3a) と性質の異なる助詞の問題を次の事例で考える。

(3b)

入力文：あの政治家はテロにより命を奪われた。

この入力文は次に示す 2 つの能動文に復元されることが考えられる。

復元文 1：(組織が) テロで政治家の命を奪う。

復元文 2：テロが政治家の命を奪う。

復元文 1 では、入力文に明記されていない「組織」という動作主が「テロ」という手段を用いて政治家の命を奪うと解釈できる。一方、復元文 2 では、「テロ」という擬人化された主格が政治家の命を奪うと解釈できる。したがって、能動パターンと受動パターンの対として、少なくとも、以下のパターンが必要であり、さらにいくつかの受動パターンを追加しなければならない。

既存パターン

能動パターン：N1 が N2 を奪う

受動パターン：N1 に/によって N2 が/を ‘奪う’.reru

意味属性制約：N1(4 人*)N2(“ タイトル ” 2385 生命現象)

追加すべきパターン

能動パターン：N1 が N3 で N4 の N2 を奪う

受動パターン 1：N4 は/が N3 により N2 を ‘奪う’.reru

受動パターン 2：N4 は/が N1 に/から/により N3 で N2 を ‘奪う’.reru

受動パターン 3：N4 は N1 の N3 に/から/により N2 が/を ‘奪う’.reru

受動パターン 4：N1 に/から/により N3 で N4 の N2 が/を ‘奪う’.reru

意味属性制約：N1(4 人*)N2(“ タイトル ” 2385 生命現象)N3(*)N4(*)

表5の分類3に該当する入力文を、助詞の追加で意味を正しく解析できるものと、できないものに分類する。

助詞の追加で解析できるもの
・この文書には最終条件が含まれる。
・その件については銘々の判断に任せられた。
・その雑誌のクイズにはたくさんの回答が寄せられた。
・その事件については裁判所の判断に委ねられた。
助詞の追加だけでは解析できないもの
・あの政治家はテロにより生命を奪われた。
・昨日その工場で役人による抜き打ち監査が行われた。
・今日の交通ストで約9万人の通勤客が足を奪われた。
・その言葉は別の意味にも用いられる。
・これは速度の意味で用いられる。
・このシステムにより交換機からデータが収集される。

太字で強調した助詞は、その助詞が受動パターンに存在しなかったものを表す。本実験では、「により・による・で・には・については・にも」の助詞が不足していることが分かった。

6.4.3 分類5「複数の候補からの選択に誤りがある」

以下に誤り事例を示す。

入力文：このホテルはよく商談の場に使われる。
 出力文：よく [このホテルが を 商談の場に 使う]。
 正解文：このホテルをよく商談の場に使う。
 以下に候補として使用した2つのパターンを示す。

能動パターン1: N1 が N2 を N3 に使う
受動パターン1: N1 から/に/によって N2 が/を N3 が/に '使う'. <i>reru</i>
意味属性制約: N1(3 主体)N2(-2702 順序 1190 金銭 2670 時間)N3(*)
用言意味属性: 利用
能動パターン2: N1 が N2 を N3 に使う
受動パターン2: N1 から/に/によって N2 が/を N3 が/に '使う'. <i>reru</i>
意味属性制約: N1(3 主体 760 人工物*)N2(*)N3(*)
用言意味属性: 利用

上記に示した2つのパターンは、意味属性制約のみが異なる。受動パターン2は、受動パターン1と比べ、自由度の高い意味属性制約が掛けられていることがわかる。正解文を出力するには、受動パターン2が使用されるべきである。しかし、今回候補選択に問題があり、意味属性を一部でも厳密に満たす受動パターンが優先されたため、正解となるパターンが出力されなかった。

6.4.4 曖昧性による問題がある

以下に誤り事例を示す。

入力文：あの会社は誇大広告で非難された。
 出力文：誇大広告で [が あの会社を 非難する]。
 正解文1:(ある団体が) 誇大広告のあの会社を非難する。
 正解文2:(ある団体が) あの会社の誇大広告を非難する。
 正解文3:(ある団体が) あの会社を誇大広告で非難する。
 入力文に対し、正解文が二通り考えられる。正解文を出力するには、入力文に対し、以下の解釈がそれぞれ必要である。

解釈1: あの会社は誇大広告が原因で(ある団体に)非難された。
解釈2:(ある団体が) あの会社の誇大広告を(対象として)非難した。
解釈3:(ある団体が) 誇大広告を利用してあの会社を非難した。

常識的には、入力文の文意は解釈1である。常識というのは、入力文の外の知識を含んでいる。解釈1に限定する処理は常識や文脈をとらえた上での処理であり、意味理解という範囲である。パターンに基づく意味解析は、文のもたらず解釈を出力することを処理範囲としているため、解釈1から解釈3まですべて出力されるべきである。しかし、解釈の仕方が複数ある場合、受動パターンに対して助詞の付与や、格要素の個数を補うだけでは、正しい意味解析が困難である。したがって、この誤り事例への対策は処理の段階ごとに設ける必要がある。

7. おわりに

本稿では、能動パターンから変換規則に受動パターンの作成を行い、受動パターンの網羅性を検証した。具体的には、能動態から受動態への9つの変換規則を示し、能動パターンから受動パターンを11,310件作成した。その結果、受動文100文に対する受動パターンのカバー率は50%であった。また、誤り事例を分析することで、作成したパターンに不足していた助詞を発見することができた。同時に、助詞による文の曖昧性が存在する場合、能動文では任意格となる格要素の追加や、格助詞の示す意味を正確に捉えたパターンの追加が必要となることがわかった。今後は、受動パターンへ不足していた助詞の付与を行うと共に、動詞ごとに助詞の意味を正確にとらえた受動パターンの作成を行う。

文 献

- [1] Masato Tokuhisa, Kumiko Endo, Yuya Kanazawa, Jin'ichi Murakami, Satoru Ikehara: Evaluation of Pattern Generalization Effect under Development of Pattern Dictionary for Machine Translation, PAFLING2005, pp.311-318, 2005.
- [2] 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.
- [3] 寺村秀夫: 日本のシンタクスと意味, くろしお出版, 1982.
- [4] 庭三郎: 現代日本語文法概説, <http://www.geocities.jp/niwasaburoo/index.html>, 2004.
- [5] 西山七絵, 村上仁一, 徳久雅人, 池原悟: 単文句型パターン辞書の構築, 言語処理学会第11回年次大会, pp.372-375, 2005.