

# 概要

## 第 I 部概要

近年、機械翻訳の方式として等価的類推思考の原理に基づく機械翻訳方式が提案されている。この方式の実現に向けて、日本語の重文・複文を対象とした文型パターンを大量に蓄積した文型パターン辞書の構築が進められている。文型パターンは、言語表現を、字面・変数・関数・記号で記述したものであり、パターンマッチングにより入力文を解析する。現在、文型パターン辞書には単語レベル・句レベル・節レベルが存在する。この単語レベル文型パターンの問題点の一つに、入力文に対し約 48%しか文型パターンが出力されておらず、現状では適合率が低いことがあげられる。また、現在の単語レベル文型パターン辞書には、入力文に対する適合率を向上させる手段として、表記のゆらぎを吸収するために、選択記号が記述されている。

そこで、本研究では単語レベル文型パターンにおける選択記号の効果を「文型パターン拡大率  $\eta$ 」、および「適合率  $R1$ 」を用いて、定量的に評価し、改良の可能性を検討した。また、現在の単語レベル文型パターン辞書は、選択記号になるべき箇所が記号になっていなかったり、表現要素の表記が不足している。そこで、それらの箇所に対し、既存の選択記号で最も表現要素数が多い選択記号による均一化、および既存の選択記号から新たに作成した選択記号による均一化を行うことで選択記号を増加したときの文型パターン拡大率と適合率も同様に評価した。

調査の結果、現在の単語レベル文型パターン辞書に付与されている選択記号は、適合率を約 2%向上させていることが分かった。さらなる適合率の向上を狙い、既存の選択記号のさらなる付与を行ったが、適合率はそれほど向上しなかった。

これにより、現在の付与されている選択記号に関しては表現のゆらぎを吸収するに十分な効果があることが示された。しかし、単語レベル文型パターン中に選択記号となるべき箇所が残っていることも示された。

## 第 II 部概要

最近、分かりやすい文への書き換えや要約や機械翻訳への適用を目指した言い換え技術に関する研究が盛んである。言い換えのための知識(事例や規則)としては、語彙資源やコーパスからの収集が代表的である。このうち、コーパスからの収集の方法において、対訳コーパスから同一の英文に対する複数の日本文を用意し、日本文どうしのアライメン

トを取る手法が提案されている。しかしこの手法は、得られる日本文の組の数が少なく、かつ組内の日本文の文数も少ないと考えられる。

そこで本研究は対訳コーパスから作成された文型パターンから、同一の英語パターンに対する複数の日本語パターンを用意し、日本語パターンのアライメントを取る手法を提案した。

今回は、単文文型パターン辞書に提案手法を用い、英語パターンが同一で日本語パターンが異なる日本語パターンの組を収集したときの、日本語パターンの組の量を調査した。また、得られた日本語パターンの組に対し、言い換えの知識が収集できた割合を調査した。

調査方法として、単文パターン辞書において英語パターンが同一で日本語パターンが異なる日本語パターンの組をランダムで100組抽出し、その日本語パターンに対し言い換えの知識が存在するかを調査した。

調査の結果、単文文型パターン215,342件から提案手法によって日本語パターンの組を4,077組収集することができた。また、収集した日本語パターンの組に言い換えの知識が含まれている割合を調査したところ、ランダムで100組中、71組の日本語パターンは言い換えの知識が含まれていることが分かった。

# 目次

<b>第I部 選択記号による文型パターンの汎化の効果</b>	<b>1</b>
1 はじめに	1
2 単語レベル文型パターンにおける選択記号	2
2.1 単語レベル文型パターン辞書の概要	2
2.2 選択記号	3
3 選択記号の効果の調査	4
3.1 調査方法	4
3.2 調査対象	6
3.2.1 選択記号の変数化	10
3.3 調査結果	14
4 考察	15
4.1 選択記号の表現要素数に関する調査	15
4.2 人手で言い換えた入力文を用いた調査	16
4.3 選択記号の表現要素を新たに発見する方法	18
5 おわりに	21
<b>第II部 単文文型パターンの言い換えの抽出</b>	<b>23</b>
1 はじめに	23
2 対訳パターンを用いた言い換え知識の収集方法	24
2.1 原理	24
2.2 特徴	24
3 調査方法	25
3.1 調査対象	25
3.1.1 単文の条件	25

3.1.2	単文文型パターンの例 . . . . .	25
3.2	言い換えの知識の収集方法 . . . . .	25
3.3	言い換えの知識の有無の調査 . . . . .	26
4	調査結果	<b>27</b>
5	英語原文が同一の場合の言い換えの収集	<b>29</b>
5.1	調査目的 . . . . .	29
5.2	調査方法 . . . . .	29
5.3	調査結果 . . . . .	30
6	考察	<b>32</b>
6.1	英語原文の同一性の調査 . . . . .	32
6.2	言い換えの知識の種類調査 . . . . .	32
6.3	文型パターン作成方法との関係 . . . . .	33
7	おわりに	<b>35</b>

## 表目次

1	記号の一覧	2
2	各辞書に対する文型パターン拡大率	14
3	各辞書に対する適合率	14
4	選択記号で使用される表現要素の位置の平均	15
5	SENAQX の表現要素の使用頻度	16
6	言い換えにより受理された文数	17
7	人手で言い換えた日本文に対する適合率	17
8	(て   で) のかわりに受理された表現要素の (一部)	20
9	言い換えの知識の有無	27
10	英語原文が同一のときの言い換えの可能性	30
11	英語原文の同一性	32
12	言い換えの知識の種類	32

## 第I部

# 選択記号による文型パターンの汎化の 効果

## 1 はじめに

近年、機械翻訳の方式として等価的類推思考の原理に基づく機械翻訳方式が提案されている [1]. この方式の実現に向けて、日本語の重文・複文を対象とした文型パターンを大量に蓄積した文型パターン辞書の構築が進められている [2]. 文型パターンは、言語表現を、字面・変数・関数・記号で記述したものであり、パターンマッチングにより入力文を解析する。現在、文型パターン辞書には単語レベル・句レベル・節レベルが存在する。この単語レベル文型パターンの問題点の一つに、入力文に対し約 48%しか文型パターンが出力されておらず、現状では適合率が低いことがあげられる。また、現在の単語レベル文型パターン辞書には、入力文に対する適合率を向上させる手段として、表記のゆらぎを吸収するために、選択記号が記述されている。

そこで、本研究では単語レベル文型パターンにおける選択記号の効果を実験的に評価し、「文型パターン拡大率  $\eta$ 」、および「適合率  $R1$ 」を用いて、定量的に評価し、改良の可能性を検討する。また、現在の単語レベル文型パターン辞書は、選択記号になるべき箇所が記号になっていなかったり、表現要素の表記が不足している。そこで、それらの箇所に対し、既存の選択記号で最も表現要素数が多い選択記号による均一化、および既存の選択記号から新たに作成した選択記号による均一化を行うことで選択記号を増加したときの文型パターン拡大率と適合率も同様に評価する。

結果として、現在付与されている選択記号は単語レベル文型パターン辞書の日本語パターン数を 2 倍相当にしている、かつ適合率をおよそ 2%向上させていることが分かった。しかし、人手による入力文を用いた調査において選択記号となるべき箇所がまだ残っていることも示された。

本稿第一部の構成は以下の通りである。第 2 章で本研究で扱う単語レベル文型パターン辞書と選択記号について述べる。第 3 章で選択記号の効果をはかるための方法について述べる。第 4 章で選択記号の効果について実験の結果をもとに考察を行う。第 5 章でまとめを述べる。

## 2 単語レベル文型パターンにおける選択記号

### 2.1 単語レベル文型パターン辞書の概要

文型パターンは、日英対訳標本文を、変数化および関数化、任意化している。その中で単語レベル文型パターンは、表現に使用される名詞、動詞などの自立語の線形な表現要素を変数化している。また、変数化すると対訳の訳出が困難になる部分は非線型な表現要素として字面、あるいは関数の形式で残されている。単語レベル文型パターンの例を以下に示す。

- ・日本語原文 自分ひとりで何でもやるのが彼の主義だ。
- ・日本語パターン  $N1(\text{ひとりで}|\text{一人で})(\text{何でも}|\text{なんでも})\text{やるのが}/\#2[N3]\text{の}/N4.da$
- ・英語原文 It is his principle to do anything whatever for himself.
- ・英語パターン It is  $\#N2[N3.poss]N4$  to do anything whatever for  $N1.reflex$ .

変数には名詞や動詞の変数を表す  $N_n$  や  $V_n$  など 8 種類がある。関数には  $.da$  や  $.kako$  などがあり、字面の指定や表現を指定している。記号は、パターン記述要素の受理の仕方について、任意化、選択および順序変更などの制御を行う。

表 1: 記号の一覧

記号名	表記	意味
離散記号	/	文型に無関係な要素
選択記号	(...   ...)	いずれかの要素列と受理
任意記号	[...]	文型選択上、任意の表現要素
補完要素記号	<...>	ゼロ代名詞など
順序任意要素 指定記号	{...   ...}	順序入れ換え可能な範囲 (例 各要素の順序)
位置変更可能 要素指定記号	$\$n^{\{...\text{定義}...\}}\$n$	指定位置に入れ換え可能 (例 副詞の位置)
文節境界記号	!	文節境界と受理
記憶記号	$\#n$	受理内容を記憶

本研究では、[2] の文型パターン辞書のうち文法・単語レベルの文型パターン辞書 (122,619 パターン) を調査対象とする。この辞書を用いて選択記号の効果を求めるために、新たな単語レベル文型パターン辞書を作成する。

## 2.2 選択記号

選択記号とは、表現要素のグループ化を行うため、助詞、助詞相当語、または、副詞などの字面のうち、同一の意味で異なる表記を持つものを対象に、置き換え可能な表現として指定したものであり、(... | ...) のように表記する。

2.1 節の例において日本語パターンに (ひとりで | 一人で) や (何でも | なんでも) という選択記号を付与することにより、日本語原文の「自分ひとりで何でもやるのが彼の主義だ。」だけでなく、「自分一人で何でもやるのが彼の主義だ。」や「自分ひとりでなんでもやるのが彼の主義だ。」のように表現がゆらいでも文型パターンパーサで受理可能になる。

以下に、選択記号の具体例を示す。

### < 選択記号の具体例 >

(て | で)

(まで | 迄)

(*V6.kako* | *ND6* をした)

(ようだ | 様だ)

(*V5.tekureru^meirei* | *V5.tekureru.meireigo*)

(とても | 逆も)

(ような | 様な)

(*V4^meirei* | *V4.meireigo* | *ND4* をしなさい)

(行く | いく)

(てき | て来)

(けれど | けど)



### 3 選択記号の効果の調査

#### 3.1 調査方法

選択記号自体の効果を、文型パターン拡大率  $\eta$  および適合率 R1 を用いて定量的に評価する。以下に、拡大率と適合率を評価するためのパラメータの計算方法を示す。

< 文型パターン拡大率  $\eta$  >

$\eta$  は「評価対象の文型パターン辞書の文型パターンが基準となる文型パターン辞書の文型パターン数に換算して、何倍に相当するか」を表したものである。定義を次式に示す。

$$\eta = X/B$$

$X$  : 対象文型パターン辞書の選択記号を全て展開したときの文型パターン数

$B$  : 基準文型パターン辞書の文型パターン数

< 例 >

例1の日本語パターンは、例2のように3つの日本語パターンに展開できるため、このときの  $\eta$  は 3.00 となる。

< 例1 >

/ytkTIME1 も/cf あい(変わり | かわり | 変り) ませず!お付き合いの/k ほど < /tkN2 は > /tcfk お願い申し上げます。

< 例2 >

/ytkTIME1 も/cf あい変わリませず!お付き合いの/k ほど < /tkN2 は > /tcfk お願い申し上げます。

/ytkTIME1 も/cf あいかわりませず!お付き合いの/k ほど < /tkN2 は > /tcfk お願い申し上げます。

/ytkTIME1 も/cf あい変りませず!お付き合いの/k ほど < /tkN2 は > /tcfk お願い申し上げます。

< 適合率  $R1$  >

適合率  $R1$  は, 入力文に対して受理された文型パターンが存在する割合を文単位で集計したものである. 定義を次式に示す.

$$R1 = M/I$$

$M$ : 「自己パターン」以外に受理された文型パターンが存在する入力文の数

$I$ : テスト用入力文の数

本研究では入力文として, 単語レベル文型パターン辞書作成に使用した日本語原文 123,451 文を使用する. 入力文と単語レベル文型パターン辞書を文型パターンパーサ jpp[5] を用いて照合を行ない, 照合結果から適合率を求める. 文型パターンパーサは入力文が受理できる文型パターンを全て出力するプログラムである.

< 例 >

例として, 文型パターンパーサに 3 文入力したとき, 「自己パターン」以外に受理された文型パターンが存在する文が 2 文有れば, 適合率は 67%となる.

$$R1 = 2/3 = 0.67$$

## 3.2 調査対象

選択記号の効果を求めるため、以下の単語レベル文型パターン辞書を作成し、文型パターン拡大率と適合率を評価する。

### (1) 選択記号を無くした単語レベル文型パターン辞書 (選択記号無し)

選択記号自体の効果を求めるため、選択記号を日本語原文と同じ表現要素のみにし、選択記号を無くした単語レベル文型パターン辞書を作成する。作成手順を以下に示す。

手順 1 現在の選択記号で、日本語原文と同じ表現要素だけを残し、残りの要素を削除する。

現在の選択記号は、原文と同じ要素を必ず一番左の要素としている。そこで選択記号内の一番左の要素だけを残し、残りの要素を削除する。

< 例 >

*/ytkTIME1* も */cf* あい(変わり | かわり | 変り)ませず!お付き合いの */k* ほど < */tkN2* は > */tcfk* お願い申し上げます。

*/ytkTIME1* も */cf* あい変わリませず!お付き合いの */k* ほど < */tkN2* は > */tcfk* お願い申し上げます。

### (2) 現在の単語レベル文型パターン辞書 (オリジナル)

本研究では、[2] で作成された単語レベル文型パターン辞書 (ver.5.3.1) を使用する。なお、この単語レベル文型パターン辞書のパターン総数は 122,619 パターンである。

この単語レベル文型パターン辞書において、選択記号の述べ数は 72,208 個、種類数は 3,652 種類であった。種類の詳細は付録 1 に示す。

(3) 既存の選択記号で最も表現要素数が長い選択記号に均一化した単語レベル文型パターン辞書 (最長均一化)

現在の単語レベル文型パターン辞書には、例3の選択記号を付与された日本語パターンがあるにも関わらず、例4のような同じ表現要素を持ちながら表現要素数が少ない選択記号が付与されている日本語パターンがある。また、選択記号となるべき表現要素が選択記号になっておらず字面で残っている日本語パターンもある。そこで該当する要素に既存の選択記号で最も表現要素数の長い選択記号を日本語パターン付与した単語レベル文型パターン辞書を作成する。

< 例 3 >

(会う | あう | 逢う)

< 例 4 >

(会う | あう | 逢う | 遇う | 遭う | 會う | 遘う)

作成手順を以下に示す。

手順1 日本語パターンを形態素解析し品詞番号を付与する。なお、選択記号内の他の単語は原文内の単語と同じ品詞番号を付与する。

< 例 >

/ytkTIME1 も /cf あい(変わり | かわり | 変り)ませず!お付き合いの /k ほど < /tkN2 は > /tcfk お願い申し上げます。

RSNYTK3;TIME1; も@7530@;SNCF4; あい@6100@;(変わり@2183@ | かわり@2183@ | 変り@2183@); ませ@7232@; ず@7193@;!; お@6100@; 付き合い@1400@; の@7410@;RSNK5; ほど@1100@;< RSNTK6;N2; は@7530@>; RSNTCFK7; お願い@1220@; 申し上げ@2413@; ます@7236@;。 @0110@

ここで”;" は形態素の境界を示し, “@...@" は品詞番号を意味する。また, “/” は、文型パターンパーサの仕様に合わせ”RSN” とし、変数と同様に番号を付与した。

手順2 手順1で作成した日本語パターンから選択記号を抽出する。

抽出した選択記号の例 (一部)

(*AJ1*; *.sugiru*; *.kako^rentai* | *AJ1^sugiru*; *.kako^rentai*)

(*ND1*; を @7430@; *RSNCF10*; し @2433@ | *V1*) |

(あげ@2413@ | 挙げ@2413@ | 擧げ@2413@ | 擧げ@2413@)

(憚る@2387@ | はばかり@2387@)

… など

手順 3 *N1* や *V2* などの変数はどの変数にでも受理できるように変数番号を *N\**, *V\** のように汎化する.

< 例 >

(*V3.reru.teiru^rentai* | *ND3* をされている)

(*V\* .reru.teiru^rentai* | *ND\** をされている)

手順 4 抽出した選択記号から, 単語を選択記号に, あるいは選択記号を同じ表現要素を持ちながらさらに表現要素数が多い選択記号に置き換える辞書を作成する.

ただし, *V\** と *ND\** に関してはさまざまに受理され置き換わってしまう可能性があるため辞書から削除した.

< 例 >

(変わり@2183@ | かわり@2183@ | 変り@2183@) (変わり@2183@ | かわり@2183@ | 変り@2183@)

(変わり@2183@ | かわり@2183@) (変わり@2183@ | かわり@2183@ | 変り@2183@)

; 変わり@2183@; (変わり@2183@ | かわり@2183@ | 変り@2183@)

; かわり@2183@; (変わり@2183@ | かわり@2183@ | 変り@2183@)

; 変り@2183@; (変わり@2183@ | かわり@2183@ | 変り@2183@)

手順 5 辞書に従い, 選択記号を置き換える.

この単語レベル文型パターン辞書において, 選択記号の述べ数は 190,239 個であり, 種類数は変数を汎化したため 2,669 種類に減少した. 種類の詳細は付録 2 に示す.

(4) 既存の選択記号から、新たに作成した選択記号に均一化した単語レベル文型パターン辞書(新作成)

例えば、例5の選択記号と例6の選択記号は、一部同じ表現要素を持っている。この2つの選択記号は例7のように一つにまとめることができる。

このように一部同じ表現要素を持っている選択記号どうしを組み合わせることで新たに選択記号を作成し、日本語文型パターンに付与することで新たな単語レベル文型パターン辞書を作成する。

< 例5 >

(上がっ | 上っ | あがっ | のぼっ | 上ぼっ | 躋っ | 躋ぼっ | 隣っ | 隣ぼっ)

< 例6 >

(拳がっ | あがっ | 拳っ | 上がっ | 上っ | 擧がっ | 擧っ | 擧がっ | 擧っ | 驤がっ | 驤っ)

< 例7 >

(上がっ | 上っ | あがっ | のぼっ | 上ぼっ | 躋っ | 躋ぼっ | 隣っ | 隣ぼっ | 拳がっ | 拳っ | 擧がっ | 擧っ | 擧がっ | 擧っ | 驤がっ | 驤っ)

作成手順を以下に示す。

手順1~3 単語レベル文型パターン辞書(3)の作成手順1~3に同じ。

手順4 既存の選択記号のうち、同じ表現要素を一つでも持っている選択記号があれば選択記号どうしを合わせ新たな選択記号を作成する。

ここで、変数を含む選択記号は、同じ表現要素を持つものが多く非常に長い選択記号となることが予想されたため、2,669種類の選択記号から変数を持たない選択記号2,131種類を使用し新たに選択記号を作成した。

< 例 >

(空い@2114@ | すい@2114@)

(空い@2114@ | あい@2114@ | 開い@2114@ | 明い@2114@)

(空い@2114@ | あい@2114@ | すい@2114@ | 明い@2114@ | 開い@2114@)

手順5 手順4で作成した選択記号と既存の選択記号を合わせ、単語レベル文型パターン辞書(3)の作成手順と同様に選択記号に置き換える辞書を作成する。

<例>

```
(空い@2114@ | あい@2114@ | すい@2114@ | 明い@2114@ | 開い@2114@)
(空い@2114@ | あい@2114@ | すい@2114@ | 明い@2114@ | 開い@2114@)
(空い@2114@ | あい@2114@ | 開い@2114@ | 明い@2114@) (空い@2114@ |
あい@2114@ | すい@2114@ | 明い@2114@ | 開い@2114@)
(空い@2114@ | すい@2114@) (空い@2114@ | あい@2114@ | すい@2114@ |
明い@2114@ | 開い@2114@)
;空い@2114@; (空い@2114@ | あい@2114@ | すい@2114@ | 明い@2114@ |
開い | @2114@)
;あい@2114@; (空い@2114@ | あい@2114@ | すい@2114@ | 明い@2114@ |
開い@2114@)
;すい@2114@; (空い@2114@ | あい@2114@ | すい@2114@ | 明い@2114@ |
開い@2114@)
;明い@2114@; (空い@2114@ | あい@2114@ | すい@2114@ | 明い@2114@ |
開い@2114@)
;開い@2114@; (空い@2114@ | あい@2114@ | すい@2114@ | 明い@2114@ |
開い@2114@)
```

手順6 辞書に従い、選択記号を置き換える。

この単語レベル文型パターン辞書において、選択記号の述べ数は190,239個であり、種類数は新たに779種類作成し合計3,448種類になった。種類の詳細は付録3に示す。

### 3.2.1 選択記号の変数化

文型パターンパーサで選択記号のある日本語パターンを照合する際、選択記号の表現要素を全て展開し、全ての組合せの日本語パターンを作成してから圧縮をする作業を行う。しかし、その圧縮のルーチンにバグが有るため、本研究のように表現要素数が多い選択記号を扱う際は圧縮をしない方が良いと考えられる。

そこで、各辞書に付与された選択記号を、[5]で示された「S型変数」および「A型変数」に置き換えた。ただし、関数をA型変数で定義することが困難なことから、変数を含んだ選択記号は変数化せず全て字面の表現要素のみで構成されている選択記号のみを変数化した。手順を以下に示す。

手順1 まず、最も選択記号の種類数が多い単語レベル文型パターン辞書(4)において使用している選択記号から、変数を含まない全ての選択記号を抜き出す。

なお、抜き出した選択記号の種類数は2,910個であった。

手順2 A型およびS型変数の定義ファイルを作成する。

<例：選択記号(NO | No | no | ノー)の場合>

A型変数の定義ファイル

SENAAA

arc(1,2,SENAAA)

arc(2,0,pop)

end

S型変数の定義ファイル

SSENAAA

NO

No

no

ノー

end

なお、A型変数は「SEN + アルファベット3文字」で名前を付けており、アルファベット3文字の箇所は「AAA」から順にAAB,AAC…と26進数で番号を示している。また、S型変数にはA型変数の名前の頭にSを付けてある。

手順3 選択記号をA型変数に置き換える辞書を作成する。

<例>

(NO | No | no | ノー) SENAAA



手順4 手順3で作成した辞書を用い, 各単語レベル文型パターン辞書の選択記号をA型変数に置き換える.

以上の手順を日本語パターンに使用した例を以下に示す.

単語レベル文型パターン辞書 (2) の日本語パターン

*/y \$ 1/tk#1[N2の]/kN3を/cf(食つ|くつ)て\$ 1^{!この}/kあたりは/cf どの/kN4も/cf(V5.teiru | ND5をしている)。*

単語レベル文型パターン辞書 (4) の日本語パターン

*RSNY6 \$ 1RSNTK7#1[N2の]RSNK8N3をRSNCF9(食つ|喰つ|くつ|くらつ|喰らつ|食らつ)(て|で) \$ 1^{!この}RSNK10あたりはRSNCF11 どのRSNK12N4もRSNCF13(V5.teiru | ND5をしている)。*

単語レベル文型パターン辞書 (4) の日本語パターンの選択記号を変数化

*RSNY6 \$ 1RSNTK7#1[N2の]RSNK8N3をRSNCF9SENCWR15SENANJ14 \$ 1^{!この}RSNK10あたりはRSNCF11 どのRSNK12N4もRSNCF13(V5.teiru | ND5をしている)。*

また, 今回作成したA型変数およびS型変数の定義ファイルの例を以下に示す.

A型変数の定義ファイルの例

```
SENAAB
arc(1,2,SSENAAB)
arc(2,0,pop)
end
SENAAC
arc(1,2,SSENAAC)
arc(2,0,pop)
end
```

## S型変数の定義ファイルの例

SSENAAB

あい

すい

空い

明い

end

SSENAAC

あい

すい

空い

明い

開い

ひらい

啓い

拆い

擺い

攤い

闡い

end

### 3.3 調査結果

(1) から (4) までの単語レベル文型パターン辞書に対する、文型パターン拡大率を表 2 に、適合率を表 3 に示す。この結果、現在の単語レベル文型パターンに付与されている選択記号は単語レベル文型パターン辞書の日本語パターン数を 2 倍相当にしている、かつ適合率をおよそ 2% 向上させている。しかし、さらなる適合率の向上を狙い作成した辞書は、付与を最も多く行った辞書 (単語レベル文型パターン辞書 (4)) で日本語パターン数が 7 倍近くに相当するにもかかわらず、適合率がほとんど向上していない。

この結果、現在の単語レベル文型パターン辞書に付与されている選択記号は、表現のゆらぎを吸収するのに十分な効果があると言える。

表 2: 各辞書に対する文型パターン拡大率

辞書	展開パターン数	文型パターン拡大率
(1) 選択記号無し	122,619	1.00(122,619/122,619)
(2) オリジナル	245,850	2.00(245,850/122,619)
(3) 最長均一化	711,055	5.80(711,055/122,619)
(4) 新作成	826,758	6.74(826,758/122,619)

表 3: 各辞書に対する適合率

辞書	自己以外に受理	適合率
(1) 選択記号無し	56994	46.330(56994/123451)
(2) オリジナル	60180	48.748(60180/123451)
(3) 最長均一化	60243	48.779(60243/123451)
(4) 新作成	60248	48.803(60248/123451)

## 4 考察

### 4.1 選択記号の表現要素数に関する調査

単語レベル文型パターン辞書(4)は、既存の選択記号から新たに選択記号を作成し、その中で最も長い選択符号に均一化している。この単語レベル文型パターン辞書(4)の選択記号の表現要素が適合率を向上させる効果があるかを調査するため、以下の実験を行った。

入力文12万文に対し文型パターンパーサで照合を行い、各選択記号に対して照合の際に使用された頻度をとった。そして各表現要素位置における使用された頻度の平均値を求めた。調査結果を表4に示す。

表 4: 選択記号で使用される表現要素の位置の平均

表現要素位置	割合
第1要素	86.800
第2要素	11.655
第3要素	1.267
第4要素	0.202
第5要素	0.059
第6要素	0.013
第7要素	0.002
第8要素以降	0

表4より、全体の98%は第2要素までに使用している。これにより、最も表現要素数が長い選択記号に均一化し、選択記号の表現要素数を増加させても適合率の向上が低いことが分かった。

以下に、例としてSENAQX(取っ | とっ | 奪っ | 採っ | 撮っ | 執っ | 獲っ | 捕っ | 盗っ | 奪っ | 攬っ)の表現要素の使用頻度を示す。

表 5: SENAQX の表現要素の使用頻度

区別	表現要素	使用回数	使用割合
+	取っ	919	53.711280%
+	とっ	651	38.047925%
+	奪っ	99	5.786090%
+	採っ	22	1.285798%
+	撮っ	15	0.876680%
+	執っ	5	0.292227%
!	獲っ	0	0.0%
!	捕っ	0	0.0%
!	盗っ	0	0.0%
!	奪っ	0	0.0%
!	攪っ	0	0.0%
	計	1711	

ここで, “+” はその表現要素が 1 度でも使用されていることを示し, “!” はその表現要素が 1 度も使用されていないことを示す.

## 4.2 人手で言い換えた入力文を用いた調査

本研究で使用した入力文, すなわち単語レベル文型パターンを作成するために使用した標本文は, 辞書や語学教育用の教科書, 機械翻訳機能評価用の試験文などで構成されている. これらは, 日本語の基本的な表現で収録されているため表現のゆらぎが少ないと考えられる. そこで, 入力文に対し人手で言い換えを行い, その文における受理パターン率を調査した.

調査対象として, 入力文からランダムで 114 文を抽出し, 人手で 641 文に言い換えた. 言い換えた 641 文のうち, 形態素解析で誤った 40 文を除いた 601 文を調査対象とした. 各辞書において受理された文数を表 5 に示す.

この結果を見ると, 選択記号が無い辞書 (1) に比べ現在の辞書 (2) ではわずかに受理パターン数が増加するが, 辞書 (2) と既存の選択記号を増加させた単語レベル文型パターン辞書 (3), (4) とを比較しても受理された文数は変わらない. なお, 単語レベル文型パターン辞書 (2) に比べ単語レベル文型パターン辞書 (3), (4) の受理された文数が減っている理由は, 選択記号を増加させたことによる文型パターンパーサ jpp のバグだと考えられる.

表 6: 言い換えにより受理された文数

辞書	受理された文数	受理されなかった文数
(1) 選択記号無し	401	200
(2) オリジナル	432	169
(3) 最長均一化	425	176
(4) 新たに作成	425	176

次に、元々の日本語原文と人手で言い換えた日本文における適合率を比較した。使用した辞書は選択記号を最も多く付与した単語レベル文型パターン辞書(4)である。結果を表6に示す。

表 7: 人手で言い換えた日本文に対する適合率

入力文	入力文数	自己以外に受理	適合率
日本語原文	114	57	50.00
言い換えた日本文	601	305	50.75

この結果、入力文の種類が変わっても適合率にそれほどの差は無いことが分かった。

次に、単語レベル文型パターン辞書(4)において受理されなかった176文と見ると、例8や例9のように、わずかな表現のゆらぎしかないが受理不可能になる文が63文存在した。この63文の受理不可能になった箇所は選択記号に置き換えることで受理できると考えられる。また、残りの113文に関しては、例10や例11のように、「サ変名詞+する」と動詞の変化や、名詞と名詞の変化などがほとんどであった。この結果、選択記号の箇所が不足していると考えられる。また、これらの箇所が全て改善されれば、適合率はおよそ80% $((305 + 176)/601 = 0.8003)$ まで向上が期待される。

<例8>

- ・そうするのはどうしてもいやだと言う。(受理)

日本語パターン：RSNY3 < RSNTK4N1 は > RSNTCFK5 そう RSNF6  
(V2<sup>rentai</sup> | ND2 をする)!のは RSNCF7 どうしても RSNF8 いやだ SENAQJ10  
RSNCF9 言う。

- ・そうするのはどうしても 嫌 だと言う。(受理不可)

< 例 9 >

- ・彼はあまりなれなれしいから人に嫌われる。(受理)

日本語パターン:RSNY4 \$ 1^RSNTK5N1 は#2[RSNCF6 あまり] \$ 1RSNF7 なれなれしいから \$ 1RSNTK8 人に RSNCF9(V3.reru | ND3 をされる)。

- ・彼はあまりなれなれしいので 人に嫌われる。(受理不可)

< 例 10 >

- ・目的地まで遠いから、時々休みながら行く。(受理)

日本語パターン:RSNYTK3N1SENAXA9RSNCF4遠いから、< RSNTK5N2は > #3[RSNTCFK6 時々]RSNTK7休みながらRSNF8SENBYQ10。

- ・目的地まで遠いから、時々休憩しながら行く。(受理不可)

< 例 11 >

- ・頭痛は明るる日になっても直らなかった。(受理)

日本語パターン:RSNY4 < RSNTK5N1は > RSNTCFK6N2はRSNCF7明るるRSNTK8TIME3に RSNCF9 なっSENAPX11RSNF10 直らなかった。

- ・頭痛は翌日になっても直らなかった。(受理不可)

### 4.3 選択記号の表現要素を新たに発見する方法

選択記号内の表現要素を発見する方法として、現在の単語レベル文型パターン中の選択記号が付与されている箇所を2形態素までなら文型パターンパーサで受理できるようにし、置き換えた表現要素が選択記号の表現要素として使用可能かどうかを調査した。

< 例 12 >

日本語原文：そこで笑ってはだめだ。

日本語パターン：/y#1[そこで]/fV2(て|で)は/cfだめだ。

日本語パターン：/y#1[そこで]/fV2\*は/cfだめだ。

具体的な例を示す。例 12 において、日本語パターン中の選択記号(て|で)のかわりにどのような表現要素でも受理可能である\*の記号をつけ、入力文 12 万文と照合を行った。

\*で受理された表現要素を受理された回数でソートし、上位 100 件ほどを調べたところ、「ていて」という表現要素を発見した。この表現要素は、例 13 の文を入力しても受理するよう(て|で|ていて)のように新たに選択記号の表現要素として追加しても構わないと考えられる。

< 例 13 >

そこで笑っていてはだめだ。

しかし、選択記号(て|で)の箇所かわりに受理された表現要素は本来 103,964 件あり、その中から人手で追加可能な表現要素を捜すのは困難であった。以下に、(て|で)のかわりに受理された表現要素の一部を示す。



表 8: (て | で) のかわりに受理された表現要素の (一部)

区別	表現要素	使用回数	使用割合
+	て	6590959	48.446188%
+	で	456923	3.358567%
-	た	328845	2.417142%
-	と	316633	2.327379%
-	ようだ	199257	1.464619%
-	、	179919	1.322477%
-	ために	177066	1.301506%
-	に	143150	1.052210%
-	を	129763	0.953810%
-	て、	127985	0.940741%
-	て其の	111007	0.815946%
-	ながら	94424	0.694054%
-	れるて	93699	0.688725%

ここで、元々存在していた表現要素と新たに受理された表現要素を区別するため、元々の表現要素には”+”を、新たに受理された表現要素には”-”を付けている。

## 5 おわりに

本研究では単語レベル文型パターンにおける選択記号の効果を「文型パターン拡大率  $\eta$ 」、および「適合率  $R1$ 」を用いて、定量的に評価した。その結果、単語レベル文型パターンにおける選択記号は単語レベル文型パターン辞書の日本語パターン数を2倍相当にしている、かつ適合率をおよそ2%向上させていることが分かった。

また、現在の単語レベル文型パターン辞書は、選択記号になるべき箇所が記号になっていなかったり、表現要素の表記が不足している。そこで、それらの箇所に対し、既存の選択記号で最も表現要素数が多い選択記号による均一化、および既存の選択記号から新たに作成した選択記号による均一化を行うことで選択記号を増加したときの文型パターン拡大率と適合率も同様に評価した。その結果、文型パターン拡大率がおよそ7倍程度になるにも関わらず、適合率はほとんど向上しなかった。

これにより、現在の付与されている選択記号に関しては表現のゆらぎを吸収するに十分な効果があるといえる。また、人手で言い換えた日本文に対しても、日本語原文と同じだけの適合率が示され、母集団が変わっても日本語原文を入力文とした際と同じだけの効果があることが分かった。しかし、単語レベル文型パターン中に選択記号となるべき箇所が残っていることも示された。

## 参考文献

- [1] 池原ほか:等価的類推思考の原理による機械翻訳方式, 電子情報通信学会技術研究報告, TL2002-34, pp.7-12, 2002.
- [2] 池原ほか:非線型な表現構造に着目した重文と複文の日英文型パターン化, 言語処理学会論文誌, Vol.11, No.3, pp.69-95, 2004.
- [3] 遠藤ほか:文型パターンにおける任意要素の記述方法とその効果, 言語処理学会第 11 回年次大会発表論文集, pp.368-371, 2005.
- [4] 池原ほか:日本語重文・複文を対象とした文法レベル文型パターンの被覆率特性, 言語処理学会論文誌, Vol.11, No.4, pp.147-178, 2004.
- [5] 徳久ほか:文型パターンパーサの試作, 言語処理学会第 10 回年次大会発表論文集, pp.608-611, 2004.

## 第II部

# 単文文型パターンの言い換えの抽出

## 1 はじめに

最近、分かりやすい文への書き換えや要約や機械翻訳への適用を目指した言い換え技術に関する研究が盛んである。言い換えのための知識(事例や規則)としては、語彙資源やコーパスからの収集が代表的である[5]。このうち、コーパスからの収集の方法において、[1]は対訳コーパスから同一の英文に対する複数の日本文を用意し、日本文どうしのアライメントを取る手法を提案している。この手法は、得られる日本文の組の数が少なく、かつ組内の日本文の文数も少ないと考えられる。

本研究は対訳コーパスから作成された文型パターンから、同一の英語パターンに対する複数の日本語パターンを用意し、日本語パターンのアライメントを取る手法を提案する。

今回は、単文文型パターン[2]に提案手法を用い、英語パターンが同一で日本語パターンが異なる日本語パターンの組を収集したときの、日本語パターンの組の量を調査する。また、得られた日本語パターンの組に対し、言い換えの知識が収集できた割合を調査する。

結果として、単文文型パターン 215,342 件から提案手法によって日本語パターンの組を 4,077 組収集することができた。また、収集した日本語パターンの組に言い換えの知識が含まれている割合を調査したところ、ランダムで 100 組中、71 組の日本語パターンは言い換えの知識が含まれていることが分かった。

本稿第二部の構成は以下の通りである。第 2 章で本研究で提案した手法の原理と特徴を述べる。第 3 章で調査対象について述べる。第 4 章で調査結果を述べる。第 5 章で第 4 章で得られた結果よりさらなる調査を行う。第 6 章でそれぞれの調査結果から考察を行い、第 7 章でまとめを述べる。

## 2 対訳パターンを用いた言い換え知識の収集方法

### 2.1 原理

句型パターンは多くの日本文をカバーできるように変数によって表現されている。よって提案手法で日本語パターンの組を収集すれば、従来手法よりも多く日本語パターンの組の数が収集できると思われる。したがって、言い換えの知識を [1] の手法よりも多く収集できると考えられる。

### 2.2 特徴

提案手法は、原文を変数に汎化することによって、変数に当てはまる単語の言い換えが、収集できない欠点があると思われる。しかし、単語の言い換えはすでに同義語辞書が存在するため、わざわざ収集する必要性は低いと考えられる。

## 3 調査方法

### 3.1 調査対象

本研究では, CREST 対訳例文 100 万件 [3] から単文を収集して作成された単文文型パターン [2] を用い, 本手法の有効性を調査する. 以下に, 本研究で用いる単文文型パターンの概要を示す.

#### 3.1.1 単文の条件

[2] で用いられた単文は, 基本的に文中に動詞が一つだけある文であり, 疑問文・命令文・会話文は対象外としている. 以下に例を示す.

(例 1) 彼は毎日自転車に乗る。

(例 2) この林檎はややすっぱい。

#### 3.1.2 単文文型パターンの例

以下に, [2] で作成された単文文型パターンの例を示す.

日本語原文：妹は私と同じくらい一所懸命勉強する。

英語原文：My sister studies as hard as I.

日本語パターン：N1 は PRO2 と同じくらい一所懸命勉強する。

英語パターン：My N1 studies as hard as PRO2.

単文文型パターンは, 7 つの日英対訳辞書を用いて原文中の対応関係が決定できた単語を *N*(名詞を表す) や *PRO*(代名詞を表す) のような変数に自動的に置き換えている. また, 日本語と英語の対応が取れるように変数に番号が付けられている.

### 3.2 言い換えの知識の収集方法

本研究では, 3.1.2 節で述べた単文文型パターン 215,342 件から, 提案手法を用い日本語パターンの組を 4,077 組収集した.

### 3.3 言い換えの知識の有無の調査

本研究では、3.2 節で収集した 4,077 組のうちランダムで選んだ 100 組の日本語パターンに対し言い換えの知識が存在するかを調査する。調査は ・ ・ x の 3 段階で評価を行う。以下に具体的に説明する。

：日本語パターンの組内の全ての日本語パターンにおいて言い換えの知識が存在

(例 3-1) *PRO1* はよく *N2* を 休む。

(例 3-2) *PRO1* はよく *N2* を 欠席する。

例 3 では、例 3-1 と例 3-2 の両方から言い換えの知識が収集できるため、評価 となる。なお、この例では「休む」と「欠席する」という言い換えの事例が収集できる。また、例 3-1 と例 3-2 そのものが言い換えの規則としても収集できる。

：日本語パターンの組内の一部の日本語パターンにおいて言い替えの知識が存在

(例 4-1) *N1* が *N2* を押し流した。

(例 4-2) *N1* が *N2* を吹きとばした。

(例 4-3) *N1* でその *N2* が流された。

例 4 では、例 4-1 と例 4-3 からは言い換えの知識が収集できるが、例 4-2 では収集できないため評価 となる。なお、この例では言い換えの規則が収集できる。

x：日本語パターンの組内の全ての日本語パターンにおいて言い換え知識が存在しない

## 4 調査結果

調査対象に提案手法を用いて収集した日本語パターンの組 4,077 組中ランダムで選んだ 100 組の日本語パターンに対し、言い換えの知識が存在するかを調査した。調査結果を表 9 に示す。

表 9: 言い換えの知識の有無

言い換えの知識の有無	組数
	63
	8
×	29

この結果、およそ 7 割の組の日本語パターンには言い換えの知識が存在することが分かった。以下に各評価の例を示す。

評価 の例

(例 6-1)  $N1$  と  $N2$  は相入れない概念だ。

(例 6-2)  $N1$  は  $N2$  と両立しない。

この例では日本語パターン全体が言い換え表現となっており、文レベルでの言い換える規則が収集できた。

評価 の例

(例 7-1)  $PRO1$  は  $N2$  がよい。

(例 7-2)  $PRO1$  は 幸せな $N2$  の 下 に生まれた。

(例 7-3)  $PRO1$  は 幸運な $N2$  の もと に生まれた。

この例では、例 7-2 と例 7-3 から「幸せな」と「幸運な」および「下」と「もと」という言い換える事例が収集できた。



評価×の例

(例 8-1) N1 は不在です。

(例 8-2) N1 関節がはずれた。

この例では、例 8-1 と例 8-2 の意味が異なるため評価×とした。

## 5 英語原文が同一の場合の言い換えの収集

### 5.1 調査目的

今回調査対象とした単文文型パターンには、英語原文が一致する場合でも例9のように対応する日本語原文の表現がわずかに異なるために変数化の度合が異なり、英語パターンが一致しない単文パターンの組合せがある。

前章までで調査を行った提案手法では、英語パターンが同一で日本語パターンが異なるという条件で日本語パターンの組を収集しているため、例9のような組合せは収集することができない。

そこでこの章では、単文文型パターンから英語原文が同一で日本語パターンが異なるという条件で日本語パターンの組を収集する。また、得られた日本語パターンの組内の日本語パターンに対し3章と同様に評価を行う。

#### (例 9-1)

日本語原文：道路がカチカチに凍っている。

英語原文：The road is frozen hard.

日本語パターン：N1がカチカチに凍っている。

英語パターン：The N1 is frozen hard.

#### (例 9-2)

日本語パターン：道がかちかちに凍っている。

英語原文：The road is frozen hard.

日本語パターン：N1がADV2に凍っている。

英語パターン：The N1 is frozen ADV2.

### 5.2 調査方法

単文文型パターン 215,342 件から、5.1 節の手法を用い日本語パターンの組を 2,638 組収集した。

### 5.3 調査結果

英語原文が同一で、日本語パターンが異なる日本語パターンの組 2,638 組中ランダムで選んだ 100 組に対し、言い換えの知識が存在する日本語パターンの数を調査した。調査結果を表 10 に示す。

表 10: 英語原文が同一のときの言い換えの可能性

言い換えの可能性	組数
	74
	5
×	21

この結果、5.1 節の手法で得た日本語パターンの組のうち、およそ 8 割の組の日本語パターンには言い換えの知識が存在することが分かった。

以下に各評価の例を示す。

評価 の例

(例 10-1) *PRO1* は 難聴だ。

(例 10-2) *PRO1* は 耳が不自由だ。

この例では、「難聴だ」と「耳が不自由だ」という言い換えの事例が収集できた

評価 の例

(例 11-1) *N1* は きゅう覚が *ADJ2*。

(例 11-2) *N1* は 鼻がよく利く。

(例 11-3) *N1* は 嗅覚が *ADJ2*。

この例では、例 11-1 と例 11-3 からは言い換えの知識が収集できるが、例 11-2 では収集できないため評価 として。なお、この例では言い換えの規則が収集できた。

評価×の例

(例 12-1) その 会社は *N1* を *VERB2* た。

(例 12-2) 会社は *N1* を *VERB2* た。

この例では、例 12-1 の日本語パターンしか比較要素が存在しないため評価×とした。

## 6 考察

### 6.1 英語原文の同一性の調査

今回対象とした日本語パターンの組には、英語原文が同一の組も含まれている。そこで、言い換えの知識が存在すると評価した日本語パターンの組のうち、英語原文が同一である組の割合を調査した。結果を表 11 に示す。

この結果、提案手法は [1] の手法の倍程度の量の日本語パターンの組が収集できていることが分かる。しかし [1] の手法よりも言い換えの知識が存在する割合が低い。よって提案手法にさらなる制約を加えることが必要であると考えられる。

表 11: 英語原文の同一性

評価	英語原文：同	英語原文：非同
	41	22
	4	4
×	5	24

### 6.2 言い換えの知識の種類調査

言い換えの知識には、言い換えの事例と言い換えの規則が存在する。そこで、本研究で得た言い換えの知識が存在する日本語パターン 71 組に対し言い換えの知識の種類を調査した。調査結果を表 12 に示す。

表 12: 言い換えの知識の種類

言い換えの知識の種類	組数
規則	37
事例	34
計	71

この結果、提案手法で得た言い換えの知識のおよそ半分は、規則としての言い換えの知識であった。なお、事例の言い換えの知識の内、文そのものが言い換えの知識となっている組が 10 組存在した。以下にそれぞれの言い換えの種類例を示す。

### 言い換えの規則の例

(例 13-1)  $N1$  が  $N2$  で  $VERB3$  た。

(例 13-2)  $N1$  は  $N2$  を  $VERB3$  た。

### 言い換えの事例の例

(例 14-1)  $PRO1$  は  $PNOUN2$   $N3$  に たった。

(例 14-2)  $PRO1$  は  $PNOUN2$   $N3$  に 出発した。

### 文そのものが言い換えの知識の例

(例 15-1) われながら腑甲斐ないと思う。

(例 15-2) 赤面の至りです。

## 6.3 文型パターン作成方法との関係

提案手法は、文型パターンの作成方法によって収集結果が変わると考えられる。例 8 で言い換えの知識が存在せず × と評価した例の単文パターンの詳細を例 16 に示す。例 16-2 を見ると、本来「肩関節」までが変数化され  $N1$  となるべきはずが「肩」しか変数化されていない。

この理由として、今回用いた単文文型パターンは、変数化を自動的に行ったため変数化が正しく行われていないためだと考えられる。人手により正しく変数化が行われれば、より言い換えの知識を収集できると考えられる。

(例 16-1)

日本語パターン： $N1$  は不在です。

日本語原文：タバコが無くなった。

英語原文：My tobacco is out.

英語パターン：My  $N1$  is out.

(例 16-2)

日本語パターン：N1 関節がはずれた。

日本語原文：肩関節がはずれた。

英語原文：My shoulder is out.

英語パターン：My N1 is out.

また、文型パターンを作成する際に名詞や用言に対し、名詞や用言の意味属性 [4] を付与することによって、より信頼性の高い言い換えの知識が獲得できると考えられる。

## 7 おわりに

本研究では、言い換えの知識を収集するため、対訳コーパスから作成された文型パターンから、同一の英語パターンに対する複数の日本語パターンを用意し、日本語パターンのアライメントを取る手法を提案した。

提案手法の効果をはかるため、単文文型パターン 215,342 件から提案手法によって日本語パターンの組を 4,077 組収集した。また、収集した日本語パターンの組に言い換えの知識が含まれている割合を調査した。4,077 組からランダムで 100 組を選び、その日本語パターンに言い換えの知識が含まれているかを調べた結果、71 組の日本語パターンは言い換えの知識が含まれていることが分かった。

今後の予定として、本研究で収集した全ての日本語パターンに対し調査を行っていきたいと考えている。



## 参考文献

- [1] 乾:言語表現を言い換える技術, 言語処理学会第8回年次大会チュートリアル, pp.1-21, 2002.
- [2] 西山ほか:単文文型パターン辞書の構築, 言語処理学会第11回年次大会発表論文集, pp.372-375, 2005.
- [3] 村上ほか:日本語英語の文対応の対訳データベース, 「言語・認識・表現」研究会, 2002.
- [4] 池原ほか:「日本語語彙大系」, ISBN4-00-130101-6, 岩波書店, 1997.
- [5] 今村ほか:階層的句アライメントを用いた日本語翻訳文の換言, 言語処理学会第7回年次大会ワークショップ, pp.15-20, 2001.

## 謝辞

この研究は、科学技術振興事業団「JST」の戦略的基礎研究推進事業「CREST」における研究領域「高度メディア社会の生活情報技術」の研究課題「セマンティックタイポロジーによる言語の等価変換と生成技術」の支援によるものです。

本研究を進めるにあたり、種々の御助言をいただきました鳥取大学工学部知能情報工学科計算機工学講座池原研究室の池原悟教授、村上仁一助教授、徳久雅人助手に心からお礼申し上げます。

また、研究に協力していただいた研究室メンバーの片山慶一郎君に感謝します。

その他、本研究に使用させて頂いた本の著者の方々、および様々な場面で御助力いただいた計算機工学講座池原研究室の皆様に深く感謝の意を表します。

## 付録

付録 1 現在の単語レベル文型パターン辞書における選択記号の種類

付録 2 作成した単語レベル文型パターン辞書 (3) における選択記号の種類

付録 3 作成した単語レベル文型パターン辞書 (4) における選択記号の種類

付録 4 英語パターンが同一で日本語パターンが異なるランダム 100 組

付録 5 英語原文が同一で日本語パターンが異なるランダム 100 組