

# 語構成パターンに応じた変換規則による形容詞の言い換え

増野 成章<sup>†</sup> 浦田 隆史<sup>‡</sup>  
佐藤 理史<sup>†</sup> 宇津呂 武仁<sup>†</sup>

<sup>†</sup> 京都大学情報学研究科知能情報学専攻

<sup>‡</sup> 京都大学工学部電気電子工学科

{masuno, urata}@pine.kuee.kyoto-u.ac.jp, {sato, utsuro}@i.kyoto-u.ac.jp

## 1. はじめに

各種の言い換えのうち、内容語の置換に重点をおいたものを、語彙的言い換えと呼ぶ<sup>1)</sup>。この語彙的言い換えを実現するためには、置換可能な内容語のペアからなる言い換え辞書が不可欠であるが、それ以外にも、限られた数の置換ペアで、より多くの言い換えを実現することができるような、ある種の「生産的な」枠組が必要であることが、麻岡ら<sup>2)</sup>によって指摘されている。これは、たとえば、以下のような言い換えを、「迅速→すばやい」という1つの置換ペアによって実現することを目指すものである。

- (1) a. 迅速な 対応 ⇒ すばやい 対応  
b. 迅速に 対応する ⇒ すばやく 対応する  
c. 対応が 迅速だ ⇒ 対応が すばやい  
d. 対応の 迅速さ ⇒ 対応の すばやさ  
e. 対応が 迅速化した ⇒ 対応が すばやくなった

本研究は、麻岡らの後継研究として、彼らが提案した「見立て」という抽象的な概念に基づく操作を使わずに、表層的な変換写像のみで同等の機能を実現する方法を提案する。この表層的な変換写像は、語構成パターンに基づく変換規則として定式化する。

本論文では、まず、本研究で対象とする言い換えるの範囲を示した後、変換規則による言い換えるの定式化について述べる。次に、形容詞・形容動詞を中核語とした内容語の変換規則について述べる。最後に、作成した言い換えるシステムについて述べる。

## 2. 変換規則に基づく言い換え

本研究が対象とする語彙的言い換えるの範囲は、おおよそ「いわゆる長単位の語のうち、短単位の内容語（中核語）を1つ含むもの」の言い換えである。すなわち、言い換えるの対象範囲は、文節を越えない範囲である。たとえば、中核語が形容動詞「明確」であれば、次のような表現（用法）が、その範囲に含まれることになる。（厳密な範囲を定義することは、それ自身、研究項目である。）

- (2) a. 明確な 答え — 連体修飾用法  
b. 明確に 答える — 連用修飾用法

- c. 答えは 明確だ — 述語用法  
d. 答えを 明確にする  
— 「二する」を伴って用言を構成  
e. 答えが 明確になる  
— 「二なる」を伴って用言を構成  
f. 答えの 明確化 が必要である  
— 「化」を伴って体言を構成  
g. 疑問点を 明確化しよう  
— 「化する」を伴って用言を構成

なお、本研究では、言い換え対象となる長単位の語に対して、可能な言い換えを生成することに焦点を絞る。文節を越えた前後の文脈を見て適切な言い換え表現を選択する問題は扱わない。

さて、上記のような言い換えるを機械的に実現するためには、次のことが必要である。

- (i) 言い換え対象とする範囲の指定  
対象範囲を、形態・品詞パターンとして記述するのであれば、その範囲指定を入力パターンと呼ぶことができよう。言い換え対象が長単位の語であるとするならば、これは、**語構成パターン**となる。  
(ii) その言い換えるに使用する（できる）言い換え辞書のエントリの指定  
これを**基本置換パターン**と呼ぼう。  
(iii) 基本置換パターンから言い換え対象全体に対する言い換え表現を生成する方法の指定  
どのような表現を生成するかを、形態・品詞パターンで記述するのであれば、この指定は、**生成パターン**（出力パターン）と呼ぶことができよう。

この3つの要素を記述したものを、変換規則と呼ぶ。変換規則は、以下のような記法で記述する。

- (3) [ 語構成パターン ] < ( 基本置換パターン )  
⇒ [ 生成パターン ]

この記法に従えば、たとえば、「迅速 → すばやい」を用いて、「迅速さ ⇒ すばやさ」、「迅速化する ⇒ すばやくなる」を実現する規則は、それぞれ、次のように書ける。（「形動」は形容動詞を、「形容」は形容詞を形容詞を表す。）

- (4) [ X:形動/語幹 さ ] < ( X:形動 → Y:形容 )

⇒[ Y:形容/語幹 さ ]

- (5) [ X:形動/語幹 化 する ] < ( X:形動 → Y:形容 )  
⇒[ Y:形容/連用形 なる ]

次に、このような変換規則で、文中の一部を書き換えることを考えよう。言い換え元（の末尾）が活用する場合、活用形の調整（写像）が必要となる。最も簡単な場合は、次のような同一品詞語（正確には、同一活用型）の言い換えの場合である。

- (6) a. 容易な 作業 ⇒ 簡単な 作業  
b. 容易に 解決できる ⇒ 簡単に 解決できる  
c. それは 容易だ ⇒ それは 簡単だ

このような同一品詞語の言い換えでは、単に活用形を言い換え元から言い換え先にコピーすればよい<sup>3)</sup>。そのため、次のような一つの変換規則で、それらを代表させることにしよう。

- (7) [ X:形動 ] < ( X:形動 → Y:形動 ) ⇒ [ Y:形動 ]

一方、言い換え元と言い換え先で品詞（正確には活用型）が異なる場合は、両者の活用形が必ずしも完全に一致しないため、単純なコピーではうまくいかない。

- (8) a. 迅速な 対応 ⇒ すばやく 対応  
(ダ列基本連体形 → 基本形)  
b. 迅速に 対応する ⇒ すばやく 対応する  
(ダ列基本連用形 → 基本連用形)  
c. 対応が 迅速だ ⇒ 対応が すばやく  
(基本形 → 基本形)

しかしながら、このような現象も、活用型の変更に對する活用形の対応表を用意することで、ほとんどの場合は解決できると考えられる（これを活用形標準変換と呼ぶ）。そこで、このような活用型の変更を伴う場合も、次のような1つの変換規則で、すべての活用形を代表させることとする。

- (9) [ X:形動 ] < ( X:形動 → Y:形容 ) ⇒ [ Y:形容 ]

つまり、語構成パターンの末尾の形態素の活用形を陽に指定しなかった場合は、すべての活用形が指定されたものとみなす。また、生成パターンの末尾の形態素が陽に指定されなかった場合は、活用形の標準変換を自動的に行なうものとする。

上記は、1つの短単位の語の場合について述べたが、長単位の語に対しても、同様の解釈を採用する。たとえば、先に示した変換規則(5)は、「迅速化した ⇒ すばやくなった」という言い換えも実現するものとする。

### 3. 形容詞・形容動詞のための変換規則

語彙的言い換えを、実際に、前節で述べた変換規則として定式化するためには、次の3つのステップを行なう必要がある。

- (i) 対象とする基本置換パターンの種類を確定する。
- (ii) 対象とする語構成パターンの種類を確定する。
- (iii) それぞれの語構成パターンに、それぞれの基本置換パターンが適用できるかどうか検討し、適用で

きる場合は、その生成パターンを作成する。

本研究では、内容語が形容詞と形容動詞（活用型がナ形容詞またはナノ形容詞である形容詞）の場合について、上記の作業を行なった。また、下記の記述では、形容動詞という品詞を除き、形態素文法体系としてJUMAN体系を採用する。

#### 3.1 基本置換パターンの整理

基本置換パターンの整理は、麻岡ら<sup>2)</sup>が作成した言い換え辞書を対象に行ない、生産性および頻度等を勘案して、採用すべき基本置換パターンを決定した。得られた基本置換パターンの概要を表1に示す。

#### 3.2 語構成パターンの整理

語構成パターンは、いわゆる用法（単独、あるいは、助詞を伴って文節を構成する場合）と、いわゆる語構成（接辞等を伴ってより大きな語を構成する場合）に大きく分けられる。語構成パターンの整理は、京大コーパスを用いて、以下のように行なった。

- (1) 京大コーパスから、形容詞（あるいは形容動詞）を含む文節を抜き出す。
- (2) 抜き出した文節の末尾の句読点等は削除する。その後、記号等を含む文節を除外する。
- (3) こうして得られた文節を、形態素数別にグループ化し、それらの類形化を手作業で行なう。最後に、頻度等を勘案して、採用すべき語構成パターンを決定する。

#### 3.3 作成した変換規則集合

こうして得られた語構成パターンと基本置換パターンを組み合わせ、変換規則を作成した。作成した変換規則の概要を表2に示す。この表では、変換規則を次のタイプに分類して示した。

**タイプ A** 単純な語の置換、あるいは、活用形のコピーを伴った語の置換を行なう規則

**タイプ B** 活用形標準変換を伴った中核語の置換を行なう規則

**タイプ C** それ以上の処理を行なう規則

この表から分かるように、全く同じ性質を持つ語に置換する基本置換パターンを使用する場合は、言い換え処理は単純となるが、性質の異なる語に置換する基本置換パターンを使用する場合には、より複雑な処理が必要となる。

タイプ C の規則は、さらに、次のように分類される。

**タイプ C1** 活用形標準変換を逸脱した活用形変換が必要な変換規則。

**タイプ C2** 判定詞や助動詞が、活用形として吸収される変換規則。（「形容詞+助動詞『だろう』⇒形容動詞/ダ列基本推量形」のような言い換えを実現する）

**タイプ C3** 新たな形態素の導入を必要とする変換規則。（「形容詞+接尾辞『すぎる』⇒「形容詞+『ない』+接尾辞『さ』+『すぎる』」のような言い換えを実現する）

表 1 形容詞・形容動詞の基本置換パターンの概要

大分類	中分類	置換先パターン	形容詞		形容動詞	
			個数	例	個数	例
形容詞	S	→ 形容詞	190	後ろめたい → やましい	239	新鮮 → 新しい
		→ 名詞+形容詞	2	奥深い → 意味深い	1	現金 → 計算高い
	C	→ ~+らしい			1	まことしやか → ほんとうらしい
		→ 形容動詞+っぽい	1	生臭い → 俗っぽい		
→ 動詞+にくい		6	煙たい → 近づきにくい	2	難解 → わかりにくい	
→ 動詞+やすい		5	人なつこい → 親しみやすい	3	通俗的 → わかりやすい	
		→ 動詞+づらい	1	言いにくい → 言いづらい		
		→ 動詞+たい	1	恨めしい → 恨みたい		
	N	→ ~+ない	9	生ぬるい → 厳しくない	58	不相応 → ふさわしくない
形容動詞	S	→ 形容動詞	64	煩わしい → めんどろ	445	あいにく → 残念
		→ 名詞+的			12	オープン → 開放的
	C	→ ~+そうだ	3	弱々しい → 弱そうだ	8	高圧的 → 偉そうだ
→ 形容詞+めだ				1	控え目だ → 少なめだ	
動詞	基本形	→ 動詞	2	しんどい → つかれる	8	懲り懲り → 懲りる
		→ 名詞+する			1	貧弱 → 見劣りする
	→ 名詞+できる			1	有望 → 期待できる	
	→ 形容動詞+なる			1	徳用 → 得になる	
		→ ~+すぎる	2	くどい → 濃すぎる	7	過少 → 少なすぎる
	テイル形	→ 動詞/テ形+いる	2	ふさわしい → 合っている	46	月並み → ありふれている
	タ形	→ 動詞/タ形	3	差し出がましい → でしゃばった	14	大胆だ → 思いきった
→ 動詞+られる/タ形				1	正式 → 決められた	
		→ 動詞+動詞/タ形	2	白々しい → 見え透いた		
連体詞		→ 連体詞	1	何気ない → ちょっとした		
副詞	S	→ 副詞			20	非常 → かなり
		→ 名詞+副詞			2	汗だく → 汗びっしょり
	S スル	→ 副詞			21	気長 → のんびり
名詞	様態	→ 名詞	2	危なげない → 余裕	22	長命 → 長生き
		→ 名詞+名詞			2	世間知らず → 経験不足
		→ 接頭辞+名詞			1	まとも → 真正面
		→ その他			2	順当 → 予想通り
	普通	→ 名詞+家			1	プロフェッショナル → 専門家
		→ 動詞+名詞			1	乗り気 → やる気
	サ	→ 形容詞+さ			2	潔白 → 正しさ

#### 4. 実 装

3章で述べた変換規則に基づいて言い換えを行なうシステムを作成した。システムは、言い換え辞書、変換規則集合、変換規則実行系から構成され、入力として与えられた文節（形態素列）に対して、その言い換えの候補となる文節を出力する。言い換え辞書のエントリ数は、形容詞 296 エントリ、形容動詞 923 エントリの計 1219 エントリである。変換規則の数は、形容詞のための変換規則 108 個、形容動詞のための変換規則 130 個の計 238 規則である。

形態素列に対する規則の適用は、次のような手順で行なう。

- (1) 形態素列全体とマッチする語構成パターンを持った規則を探す。
- (2) そのような規則が見つかった場合、それぞれの規則に対して以下の処理を行ない、ステップ (4) へ進む。
  - (a) 規則で指定された基本置換パターンとマッチするエントリを言い換え辞書から探す。
  - (b) 見つかったエントリのそれぞれに対して、生成パターンを用いて、言い換え出力を生成

する。

- (3) 形態素列から末尾の形態素を削除し、形態素列がまだ存在する場合は、ステップ (1) へ戻る。
- (4) 終了。

この手順のステップ (2) で、複数の規則、あるいは、複数のエントリが見つかった場合、複数の言い換え候補が出力される。この手順は、より長い語構成パターンを持つ規則を優先するため、たとえば、「迅速化する」という文節に対しては、[ X:形容/語幹 化 する ] という語構成パターンを持った規則が適用され、[ X:形容/語幹 化 ] と語構成パターンを持った規則は、適用されないことになる。本システムの実行例を表 3 に示す。

#### 5. 検 討

本論文で提案した方式は、言い換えの対象範囲（語構成パターン）と使用する言い換え辞書エントリ（基本置換パターン）を変換規則に明示的に記述するため、システムの挙動を制御しやすい。これにより、インクリメンタルなシステム開発が可能となる。

現在までに実装した部分は、形容詞・形容動詞の中核的・典型的用法（語構成パターン）の範囲に留まっており、今後、拡充すべき点は多い。表 2 に示したとおり、現

表 2 変換規則の概要

		基本置換パターン分類																																																														
		形容詞			形容動詞		動詞			連体詞	副詞		名詞																																																			
		S	C	N	S	C	基本形	テイル形	タ形		S	S スル	様態	普通	サ																																																	
形容詞の語構成パターン	連体修飾用法	A			B		B		C3	A	-	-	B	-	-																																																	
	連用修飾用法	A			B		B		C3	×			-			-	B	-	-																																													
	述語用法	A			B		B		C3								×			-	-	B	-	-																																								
	-だろう	A			C2		B		C3													×			-	-	A	-	-																																			
	-です	A			C2		△		△																		×			-	-	A	-	-																														
	-んだ	A			C1		B		C3																							×			-	-	C1	-	-																									
	-する/なる	A			B		△		△																												×			-	-	B	-	-																				
	-ない	A			C1		C1		C3																																	×			-	-	C1	-	-															
	-そうだ	A	C3	B	△	B		C3	×																																						-			-	A	-	-											
	-すぎる	A	C3	B	△	B		C3																																											×			-	-	A	-	-						
	-さ	A	×	B	△	×		×																																																×			-	-	A	-	-	
	-め/げ	A	×	×	×		×																																																						×			×
形容動詞の語構成パターン	名詞用法	×			A	×	×				-	-		A																																																		
連体修飾用法	B			A		B		C3		×			-	-	B	×		×																																														
連用修飾用法	B			A		B		C3							×	-	-	B	×	×																																												
述語用法	B			A		B		C3										×	-	-	C3	B	×																																									
-する/なる	B			A		△		△													×	-	-	△	B	×																																						
-ない	C1			A		C1		C3																×	-	-	C3	B	×																																			
-そうだ	B	C3	A	△	B		C3	×																			-	-	C3	△	×																																	
-すぎる	B	C3	A	△	B		C3																						×	-	-	C3	△	×																														
-さ	B	×	A	△	×		×		×																							-	-	×	△	C2																												
-化	△			A		△																												×	×	-	-	△	△	×																								
-性	△			A		△																												×				×	-	-	△	△	×																					
-視	△			A		△																												×							×	-	-	△	△	×																		

A：タイプ A、B：タイプ B、C1~5：タイプ C、△：規則作成保留中、×：言い換えできない、-：基本置換パターンに該当なし

表 3 システムの実行例

入力 (入力文節=ゴシック体)	使われた規則・エントリ		出力 (出力文節=ゴシック体)
	規則タイプ	エントリ	
とても後ろめたかった。	A	〈後ろめたい → やましい〉	とてもやましかった。
そのアイデアは新鮮だった。	B	〈新鮮 → 新しい〉	そのアイデアは新しかった。
	A	〈新鮮 → きれい〉	そのアイデアはきれいだった。
その仕事はしんどくなかった。	C1	〈しんどい → つかれる〉	その仕事はつかれなかった。
彼がいれば、心強だろう。	C2	〈心強い → 安心〉	彼がいれば、安心だろう。
練習が生ぬるすぎた。	C3	〈生ぬるい → 厳しくない〉	練習が厳しくなすぎた。
形がいびつだった。	C3	〈いびつ → ゆがんだ〉	形がゆがんでいた。
彼の行動は大胆でない。	C3	〈大胆 → 思いきった〉	彼の行動は思いきっていない。

在、作成を保留している変換規則があると同時に、語構成パターンとしては、形容動詞（語幹）が複合名詞の造語成分となる場合、および、接頭辞を伴って語を構成する場合などが未実装である。これらのものを実装していく必要がある。

一方、基本置換パターンの類型化においても、置換先が動詞、および、名詞の場合の類型化が不十分である。たとえば、動詞においては、状態動詞・動作動詞の区別、自動詞・他動詞の区別などが正しい言い換えの生成には不可欠である。このような語彙的情報を整備するとともに、これらをどのように利用するかを検討していく必要がある。

本研究が対象としているのは、文節の範囲を越えない範囲であるが、一つの文節を複数の文節で言い換えることも頻繁に発生しうる。このような言い換えも生成できるように、今後、枠組を拡張していく予定である。

本研究の一部は、次の研究費による；基盤研究 (A) 「円

滑な情報伝達を支援する言語規格と言語変換技術」(課題番号 16200009)、21 世紀 COE プログラム「知識社会基盤構築のための情報学拠点形成」。

### 参考文献

- 乾健太郎, 藤田篤: 言い換え技術に関する研究動向, 自然言語処理, Vol. 11, No. 5, pp. 151-198 (2004).
- 麻岡正洋, 佐藤理史, 宇津呂武仁: 語構成を利用した言い換え表現の自動生成, 言語処理学会第 10 回年次大会発表論文集, pp. 488-491 (2004).
- 近藤恵子, 佐藤理史, 奥村学. 「サ変名詞+する」から動詞相当句への言い換え. 情報処理学会論文誌, Vol. 40, No. 11, pp. 4064-4074 (1999).