

# 係り受け解析を用いた比喩生成支援システム

須堯大喜<sup>†1</sup> 寺井あすか<sup>†1</sup>

**概要:** 比喩は聞き手に詩的・審美的効果を喚起する効果があるが、その機能を十全に発揮する適切な喩辞の選択ができない場合がある。そこで本研究では、青空文庫コーパスから比喩を含む文を抽出し、係り受け解析を行うことで単語間の係り受け関係を明らかにした比喩表現データベースを作成した。さらに、比喩を含む入力文に対して係り受け解析を行い、入力文における被喩辞、表現したい特徴を抽出し、データベースから同様の被喩辞、あるいは表現したい特徴をもつ比喩文を検索することで、その文における喩辞を入力文に含まれる喩辞の代替案、比喩文を例文として出力するシステムを提案した。さらに、提案したシステムを用いて与えられた画像から比喩を含む文章を生成する評価実験を実施した。事後アンケートの結果により、提案したシステムの有用性が示唆された。

**キーワード:** 比喩, 直喩, 比喩生成, 係り受け解析, CaboCha, 青空文庫

## Metaphor Generation Support System using Dependency Analysis

TAIKI SUGYO<sup>†1</sup> ASUKA TERAI<sup>†1</sup>

**Abstract:** Metaphors have the function enhancing poetic and aesthetic effects to readers / listeners. However, there is a possibility that we cannot select the adequate vehicle in metaphor generation. In this research, sentences including metaphors were extracted from a literature corpus (Aozora Bunko corpus) and analyzed using dependency parsing to construct a metaphorical expression database. Then, the metaphor generation support system was proposed based on the database. In the system, given an input sentence including metaphor, a topic and expressed features are extracted using dependency analysis for the sentences. And sentences with the same topic or the features are searched from the database and the vehicles in the searched sentences are output as alternatives for the vehicle in the input sentence. Furthermore, an experiment was conducted to evaluate the system. In the experiment, participants were asked to generate a sentence including a metaphor from the shown image with or without the system. The results of the post questionnaire suggest the efficiency of the system.

**Keywords:** metaphor, simile, metaphor generation, dependency analysis, CaboCha, Aozora Bunko

### 1. はじめに

私たちは、ある物事を説明する方法として多くの修辞技法を用いる。その中でも他の事柄に見立て、喩える表現である比喩には字義通りの表現では伝達できない情報を表現し、相手に詩的・審美的効果を喚起する機能がある[1]。そのため、人の情動に影響を与える文学作品などの情緒的な文章が多く見受けられる作品の中では、比喩は文章に深みを与える重要な表現方法として多く用いられている。

しかし、比喩の生成には個人が持つ情緒的・感覚的意味が根本となっている。そのため、喩える語（喩辞）の探索においてどれだけ豊富な語彙を持っているかを含め、個人が保持する意味構造の違いが比喩生成の個人差を規定すると指摘されている[2]。

そのため、私たちがあるものごとについて説明するとき、

比喩が持つ効果を十全に生かすことができる適切な喩辞の選択ができない可能性がある。

### 2. 関連研究と目的

前章で示した可能性に対し、電子辞書を用いた比喩を含む文章作成支援システムが提案されている[3]。このシステムでは、ユーザが入力した文章から対象（喩えられる語：被喩辞）と観点（表現したい特徴）を抽出し、それに基づき喩辞の候補を提示するとともに、喩辞候補に関連する新たな特徴を提示する。しかし、このシステムでは電子辞書に基づき候補となる喩辞の推定を行なっていることから比喩として不適切な候補が出力されてしまう可能性があるとともに、通常用いられる喩辞が提案されないといった問題が残されている。

一方、Web上から収集された比喩を含む文に基づく比喩生成が試みられている[4]。この先行研究では、入力単語（被喩

<sup>†1</sup> 公立はこだて未来大学  
Future University Hakodate

辞)かつ「に例え」「のような」を含む文を抽出し、係り受け解析を行うことで文における喩辞を候補とする。その候補に対して京大格フレーム[5]を利用し、入力単語に関わる動詞群と喩辞に関わる動詞群の重複数を関連度として取得し、数値の高いものを喩辞として比喩の生成を行なった。しかしこの方法では「表現したい特徴」が考慮されていない。

また、日本語インフォ[6]や比喩表現辞典[7]等、文学作品に用いられている比喩表現が収集されているが、被喩辞、表現したい特徴を考慮した柔軟な検索が可能なシステムは提案されていない。

そこで、本研究では「のような」「ように」により比喩であることが明示される直喩を対象として、(1)文学作品から「のような」又は「ように」を含む文を収集し、係り受け解析により被喩辞、喩辞、表現したい特徴を抽出することで比喩表現データベースを構築し、(2)このデータベースに基づく比喩生成支援システムを提案する。また、本研究で提案するシステムでは、入力された比喩を含む文の被喩辞、喩辞、表現したい特徴の抽出を行うことで喩辞の代替案の提案が可能なシステムの構築を目指す。

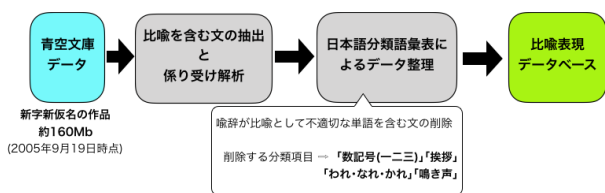


図1 データベース作成の概要

Figure 1 Overview of database creation.

### 3. 比喩生成支援システム

#### 3.1. 比喩表現データベース

本研究では比喩生成支援システムを構築するにあたり、比喩表現が実際に使用されている文を収集することで比喩表現データベースを作成した(図1)。聞き手に詩的・審美的効果を喚起する比喩の機能を重視し、本研究では情緒的な文章が多く見受けられる文学作品から比喩を収集し、比喩表現データベースを作成した。対象を2005年9月12日時点での青空文庫内の新字新仮名の作品(約160Mb)[8]とし、対象作品から「のような」「ように」をキーワードとして比喩を含む文を収集した。また、それらの文に対して日本語係り受け解析器CaboCha(以降「CaboCha」と記す)[9]を用いて係り受け解析を行い、各単語間の係り受け関係を明らかにし、被喩辞、喩辞、表現したい特徴(文中の被喩辞の特徴を表している単語)を抽出した。この時、抽出は文単位で行い、1文を改行もしくは句点を一区切りの文とした。

また、日本語分類語彙表[10]の分類項目に基づき喩辞として適さない語(「こそあど他」「数記号(一二三)」「挨拶

「われ・なれ・かれ」「鳴き声」に含まれる685語)を喩辞として持つ文を削除した。その結果、「ように」をキーワードとして持つ8546文、「のような」をキーワードとして持つ10663文、計19209文を抽出することができた。

作成したデータベースの項目は、「取得番号」「本文」「キーワード(「のような」または「ように」)」「喩辞(本システムでは名詞のみに限定)」「被喩辞(名詞)」「被喩辞の修飾語」「状態を示す語」「状態を示す語の修飾語」「副詞」とした。「被喩辞の修飾語」「状態を示す語」「状態を示す語の修飾語」「副詞」の4項目を表現したい特徴として扱うこととする。

#### 3.2. 比喩生成支援システム

本支援システムでは書き手が支援システムの入力画面から比喩を含む文を入力し、喩辞の代替案を検索する「文探索」(図2)と、被喩辞と表現したい特徴から喩辞の候補を検索する「部分探索」(図3)の2つの検索方法を用意した。システム使用者は使用用途に即したどちらかの検索方法を選択する必要がある。

文探索の場合、テキストエリアに入力された比喩を含む文(例:「彼女の手は白磁のように白く綺麗であった。」)に対してCaboChaを用いて係り受け解析を行い、被喩辞、表現したい特徴として被喩辞の修飾語、比喩で表現されている状態を示す語、状態を示す語の修飾語、副詞を要素として抽出する(例:キーワード「ように」、喩辞「白磁」、被喩辞「彼女の手」、表現したい特徴「白い」「綺麗」)。喩辞が思いつかない場合、喩辞を「\*」と入力することで同様の結果を得ることができる。抽出された語句から、被喩辞または表現したい特徴として同じ語句を1つでも含む文を比喩表現データベースから検索し、検索された文の喩辞を代替案、本文を例文として出力する。加えて、同じ喩辞を含む例文数の多いものから降順に喩辞と例文の個数を出力する(図4)。加えて、例文は出力画面内の喩辞を選択し、クリックすることで新規タブとして出力される(図5)。文探索における検索条件として、文中に(1)喩辞(\*を含む)と被喩辞が含まれているか、(2)喩辞と表現したい特徴の最低1つ以上の係り受け関係が成立しているかを文探索の検索可能な最低条件とした。

部分探索の場合、検索に必ず必要な語句としての被喩辞と任意で表現したい特徴を入力するという形を取っている。入力された要素は全て基本形に変換され、文探索同様に比喩表現データベースの検索により代替案と例文を出力する。また、表現したい特徴は最大で被喩辞の修飾語、比喩で表現されている状態を示す語、状態を示す語の修飾語、副詞をそれぞれ一つずつ入力されることを想定し、テキストエリアを4枠とした。(例として、文探索同様の結果を得るためには被喩辞「彼女の手」、表現したい特徴を「白い」「綺麗」を入力すれば良い。)

図2 入力画面（文探索）

Figure 2 Input screen (Sentence Search).

図3 入力画面（部分探索）

Figure 3 Input screen (Partial search).

### 3.3. 出力結果

支援システムを作成するにあたり、文探索において4パターンの入力文を想定した。T=被喩辞、V=喩辞、X=表現したい特徴とした場合、想定したパターンは、「TはVのようにX」、「VのようにXなT」、「VのようなXであるT」、「VのようなTはX」である。今回はそれぞれを「彼女の手は白磁のように綺麗」、「太陽のように真っ赤な林檎」、「向日葵のような笑顔が明るい」、「宝石のような美しい瞳」とした場合の結果を表1に示した。各入力文に対し、37語、15語、55語、83語の喩辞の代替案が提案された。出力数に差異が生じたが、異なる文型の入力に対して提案結果を得ることができた。また、Vを「\*」として入力した場合、文を倒置法に置き換えた場合でも同様の提案を行うことが可能であった。また、上記の入力パターンを登場人物のセリフとして入力した際、表現したい特徴を抽出できないことがあった。これはCaboChaでの係り受け解析の正答率が89.29%[11]であり、話し言葉などの文法に乱れのある文章に十分に対応出来ていない為に起った問題である。

雪	4
浮氷	2
女	2
玉	2
綿	2
銀	1
宝石	1

図4 出力画面（代替案）

Figure 4 Output screen (Alternative plans).

馬は主人の清潔好きな癖から、毎日洗い清められて、雪のように白く輝いていた。  
 ・これはこの小僧の頭が雪のように白く輝いていたからです。  
 ・かさこそと雑木の葉が、ばさりと朴の木の広葉が、……朴の木の葉は雪のように白く曝らされていた。  
 ・と、彼らの一団の中で、雪のように白く輝く物があったが、それは三頭の白馬であった。  
 ・殊に目立つのはその大きい爪で、さながら雪のように白く光っていた。

図5 出力画面（例文）

Figure 5 Output screen (Example sentences).

## 4. 実験

本章では、作成した支援システムの有用性を評価するために実施した評価実験と事後アンケートについて説明をする。実験では、システムを利用する実際の環境を想定し、提示された画像から画像を表現する比喩とその特徴を含む文章を作成する課題（比喩生成課題）を用いた。

### 4.1. 予備実験

評価実験にあたり実験で出題する画像と1問あたりにかかる時間設定のため、予備実験を行なった。

被験者は大学生4名を対象とした。予備実験で使用する画像を「ヒト」、「景色」、「植物」、「動物」の4種類のカテゴリーに分けた。各カテゴリーの画像を5枚ずつ、計20枚を準備した。予備実験で使用する画像はWeb上で公開されている画像内から画像を無作為に抽出した。それらの画像を用い比喩生成課題を被験者に行わせた。課題に対する回答時間を

表1 「文探索」での提案結果  
 Table 1 results for Sentence Search.

入力文	キーワード	表現したい特徴				提案された結果(喩辞:提案数)
		喩辞	被喩辞	修飾(被喩辞)	状態	
彼女の手は白磁のように白く綺麗	ように	白磁	彼女の手		綺麗	雪:5,綿:2,象牙:1,etc.
太陽のように真っ赤な林檎	ように	太陽	林檎	真っ赤		血:4,火:2,兎の目:1,etc.
向日葵のような笑顔が明るい	ような	向日葵	笑顔		明るい	真昼:2,春の芝生:1,etc.
宝石のような美しい瞳	ような	宝石	瞳	美しい		絵:4,花:3,海:2,etc.

表2 予備実験での解答結果  
 Table 2 results of answers in preliminary experiments.

	最短時間		最長時間		平均時間 [s]
	名称	時間[s]	名称	時間[s]	
ヒト	着物女性	17.1	顔	32.9	25.00
景色	田舎風景	34.4	日本庭園	58.9	46.65
植物	野菜	38.7	花	108.4	73.55
動物	魚	25.2	猫	62.9	44.05
平均回答時間[s]		28.9		65.8	47.31

比喩を作成する際の容易さの指標とし、各カテゴリーで解凍時間が最短・最長であった画像を本実験で使用することにした。

予備実験の結果より各カテゴリーごとに、(1)「ヒト」の場合は最短時間：「着物女性」と最長時間：「顔」、(2)「景色」の場合は最短時間：「田舎風景」と最長時間：「日本庭園」、(3)「植物」の場合は最短時間：「野菜」と最長時間：「花」、(4)「動物」の場合は最短時間：「魚」と最長時間：「猫」を評価実験で用いる画像として選出した。選出された画像と回答時間を表2に示す。また、全画像の平均解答時間が45.4秒であり、最長の解答時間が108.4秒だったことから本実験ではシステムの利用を考慮し、180秒を制限解答時間とした。

#### 4.2. 方法

被験者は大学生10名を対象とした。実験は予備実験で選出した画像を使用し、全8問から成る。内4問を支援システムを利用して比喩の特徴を含む文章の生成を行う。また支援システム利用の際、被験者にはあくまで支援システムであり、比喩の生成において提案システムが提案した結果をそのまま使う必要がないことと教示した。さらに画像の提示順序はランダムで行い、前半問題でシステムを利用する前半利用者群と後半問題でシステムを利用する後半利用者群の2群に被験者を分けた。また、周囲の環境に左右されないよう被験者と実験実施者の2名のみが在室する部屋で実験を実施した。

実験実施後に、被験者に提案システムに関する事後アンケートを質問形式と自由記述形式で実施した。

また客観的観点から支援システムを評価するために、実験

で生成された文の第三者評価を実施した。予備実験者・評価実験参加者とは異なる大学生5名(比喩に関する日本語研究を実施)により評価を行った。画像を提示し、その画像内の特徴を表現する比喩を含む文として(1)適切な表現であるか、(2)斬新であるか、(3)理解がしやすいかの3項目について5段階リッカートスケールで評価した。

#### 4.3. 結果

支援システム利用に関する事後アンケートの結果として、各評価の平均を図6、表3に示す。各項平均評定値に対して、5段階リッカートスケール(5“非常にそう思う”から1“全くそう思わない”)の「3“どちらでもない”」を基準とした1標本のt検定をおこなった。

全ての項目において評定評価値が3以上であった。特に、「比喩生成支援システムで生成された比喩の代替案は自らが思いつかなかったものが含まれている」「比喩生成支援システムを使用することで効率的に文章を作成することができる」「文章を作成する上で比喩生成支援システムは有用である。」においては5%水準で3(どちらでもない)より有意により高い。また、支援システムが良い点として、「自分のイメージを上手く言語化してもらえる」、「曖昧にしか思いつかなかったが、支援システム利用により適切に自分が見た具体的な景色を表現できる」、「自分の知らない単語があり、自身の語彙数を補填してくれている」、「きっかけをゼロから掴むより作成が容易」などが挙げられた。改善点として、「データベース中の語彙が少ない」「提案される語彙が少し硬い表現」という意見があった。改善点として、「データベース中の語彙が少ない」「提案される語彙が少し硬い表

表3 支援システムに関するアンケート調査  
 Table 3 Questionnaire survey on support system.

アンケート項目	平均	p値
比喩生成支援システムで生成された比喩の代替案は自らが思いつかなかったものが含まれている (新規性)	4.3	0.00001**
比喩生成支援システムで生成された比喩の代替案は適切なものが多い (適切性)	3.3	0.33057
比喩生成支援システムで生成された比喩の代替案は創造的なものが含まれている (創造性)	3.7	0.05100
比喩生成支援システムを使用することで効率的に比喩を含む文を作成することができる (効率的な文章)	4.1	0.0001**
比喩生成支援システムを使用することで創造的な文章が作成できる (創造的な文章)	3.8	0.02477*
文章を作成する上で比喩生成支援システムは有用である (有用性)	4.1	0.00089**
今後、文章を作成する際に比喩生成支援システムを使いたい (継続使用)	4.0	0.003533*

\*p < .05; \*\*p < .01

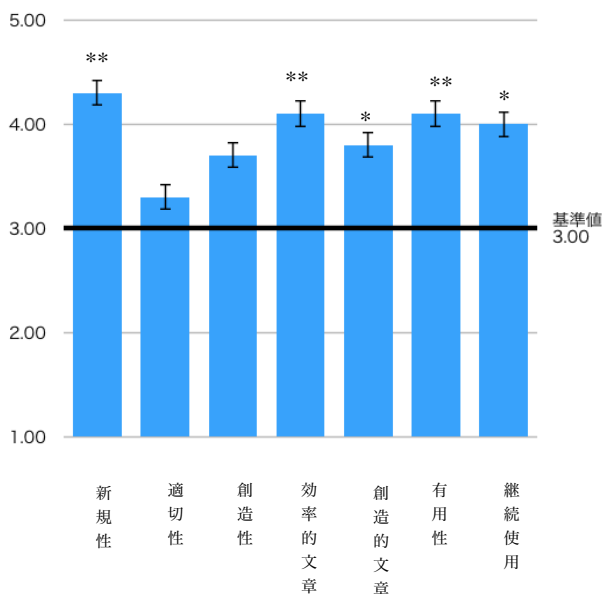


図6 システム評価におけるアンケート結果  
 Figure 6 Questionnaire results on system evaluation.

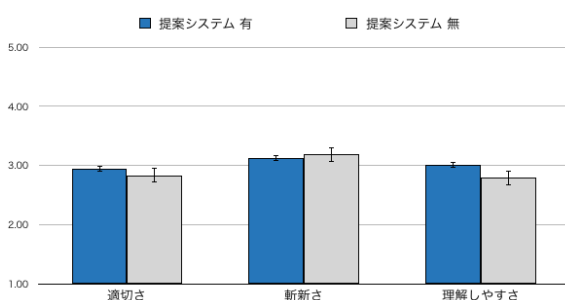


図7 生成文におけるアンケート結果  
 Figure 7 Questionnaire results on generated sentences.

現」という意見があった。

生成文に関する評価アンケートでは、各画像ごとの評定平均値をシステムの有無でt検定を行った(図7)。その結果、「画

像が持つ適切な表現である」「画像が持つ特徴を表現する理解しやすい比喩である」の2項目において支援システム利用時では非利用時より平均値がわずかに上回った。しかし、有意差は見られなかった。

## 5. 考察

本研究では、青空文庫コーパスを用いて作成した比喩表現データベースに基づく比喩生成支援システムを構築した。さらに、提案システムの有用性を検証するため、提示された画像から画像を表現する比喩とその特徴を含む文章の作成を行う比喩生成課題を用いた評価実験を実施した。その結果、システム利用者に対する事後アンケートの結果から提案システムの有用性が示された。しかし、システム利用による生成された比喩の質に対する影響は見られなかった。これは比喩表現データベースに使用されている青空文庫内の書籍データが著作権の適用外のみ掲載されているため、比喩表現データベースで使用したデータが現代の語句として適していなかったからだと考えている。

今後の展望として、事後アンケートで得られた「データベース中の語彙が少ない」「提案された語彙が少し硬い表現である」とあげられた意見に対し、比喩表現データベースのデータ量を増やし、より安定した出力の反映と、生成された比喩を含む文の質を高めるための推薦機能を実装したいと考えている。

## 参考文献

1. 内海彰. 比喩によってどのように指摘効果が喚起されるか 比喩の鑑賞過程の認知モデルについて. 第17回人工知能学会全国大会. 2003.
2. 楠見孝. 比喩の処理過程と意味構造. 風間書房. 1995.
3. 北田純弥, 萩原将文. 電子辞書を用いた比喩による文章作成支援システム. 情報処理学会論文誌, 2001, vol.42, no.5.
4. 近村亮一, ジェブラカファウ, 荒木健治. webデータを用いた単語に対する比喩的文取集法. 2011, 第27回ファジィシステムシンポジウム, 112-115.
5. “京大格フレーム検索”. <http://reed.kuee.kyoto-u.ac.jp/cf-search>.

6. “日本語表現インフォ“[.https://hyogen.info](https://hyogen.info) .
7. 中村明. 比喩表現辞典. 角川書店. 1995.
8. 野口英司. インターネット図書 青空文庫. はる書房. 2005.
9. 工藤拓, 松本裕司. チャンキングの段階適用による日本語係り受け解析. 情報処理学会論文誌, 2002, 43(6), 1834-1842.
10. 国立国語研究所編. 分類語彙表-増補改訂版. 大日本図書. 2004.
11. “Yet Another Japanese Dependency Structure Analyzer”. <http://chasen.naist.jp/chaki/t/2005-08-29/doc/CaboCha%20Yet%20Another%20Japanese%20Dependency%20Structure%20Analyzer.htm> .