

# 電子辞書を用いた比喻による文章作成支援システム

北田 純 弥<sup>†</sup>, 萩原 将 文<sup>†</sup>

本論文では、大規模な知識を有する電子辞書を利用し、認知心理学に基づく比喻による文章作成支援システムを提案する。比喻は、言語による情報伝達のみでなく、我々の思考や概念化、記憶などの認知過程にも深くかかわっていることが広く認識されている。また、自然言語処理の観点から、比喻研究の成果から反映できる技術やモデルは、応用範囲が広いと考えられている。一方、従来のコンピュータを利用した比喻研究は比喻理解に関する研究が多く、使用する知識ベースの獲得にも大きな問題があった。認知心理学においては、比喻の処理過程には情緒・感覚的類似性やカテゴリー的類似性、修飾語との共起関係の強さが比喻の生成に大きな影響を与えるとされている。そこで、本論文ではこれらの概念が工学的に導入された、電子辞書を用いた比喻による文章作成支援システムを提案する。提案システムでは、情緒・感覚的類似度、カテゴリー的類似度、共起度の3点を満たした比喻度と呼ばれる概念を導入し、柔軟な比喻の計算を行う。また、知識ベースとして大規模な知識を有するEDR電子辞書を利用している。さらに、発散的思考支援ツールとしての立場に立ち、文章作成支援システムという形で、実際にユーザが使用できるシステムを実装している。詳細な評価実験により、提案システムの有効性が確認されている。

## Figurative Composition Support System Using Electronic Dictionaries

JUNYA KITADA<sup>†</sup> and MASAFUMI HAGIWARA<sup>†</sup>

In this paper, we propose a figurative composition support system using electronic dictionaries. It is well known that simile and metaphor have deep relation not only to linguistic communication but also to human cognitive activity. However, most of the conventional computational models about simile and metaphor are for comprehension of figurative expressions and have difficulties in acquisition of knowledge. The proposed system is devised to support generation of figurative sentences using electronic dictionaries which has large-scale knowledge. It employs the following findings about understanding of simile and metaphor in cognitive psychology: (1) similarity of category influences novelty of the figurative expression; (2) emotional and sensory similarity raises understanding of the figurative expression; (3) common modifier raises understanding of the figurative expression. Based on these ideas, we define a figurative score which is calculated by categorical similarity, emotional similarity and cooccurrence score. The proposed system uses "EDR electronic dictionary" which has about 400,000 knowledge base. The proposed system can be used easily from a viewpoint of supporting human divergent thinking. A series of computer experiments shows the effectiveness of the proposed system.

### 1. はじめに

近年、人工知能の新しいあり方として知的生産性支援ツールが注目を浴びている<sup>1),4)</sup>。知的生産性支援とは人間の創造的、知的活動を助け、既知の知識の操作、加工によって新しい発想を生み出す過程を支援することである。知的生産性支援は、コンピュータによって人間の創造的な問題解決や思考活動を支援する一大研

究領域である<sup>1)</sup>。そして、コンピュータの得意なことはコンピュータに任せ、人間が得意なことは人間に任せるといった基本理念に基づき、システム設計を行うものである<sup>2)</sup>。

一方、認知心理学の分野において、比喻は言語による情報伝達のみでなく、我々の思考や概念化、記憶などの認知過程にも深くかかわっていることが広く認識されている<sup>5)</sup>。また、自然言語処理の観点から、比喻研究の成果を利用して反映できる技術やモデルは、かなり応用範囲が広いと考えられている。

このような考えから、コンピュータによる比喻の研究が増えてきている<sup>6),12)</sup>。なかでも多いのは、比喻

<sup>†</sup> 慶應義塾大学理工学部

Faculty of Science and Technology, Keio University  
現在、日本アイ・ビー・エム株式会社  
Presently with IBM Japan Ltd.

理解に関する研究である。森らは、状況理論から展開された視点を導入した比喻理解のモデル化を行っている<sup>7)</sup>。土井らは、階層型のニューラルネットワークを用いて属性層の中から比喻の意味を選択する方法を提案している<sup>8)</sup>。しかし、このモデルにおいて使用している言葉の数はたかだか 50 語である。岩山らは、顕現性を利用した視点表現の理解を取り入れ、単語の持つ情報量を利用した確率的プロトタイプモデルを提案している<sup>9)</sup>。内海は、情緒的類似性の考え方を取り入れた比喻理解のモデルを提案し、コンピュータによる比喻理解の実験を行っている<sup>10)</sup>。岩山らと内海の研究では、隠辞の属性情報として属性名とその重要度や属性値の確率を人手によって与えているため、使用している知識ベースは非常に小規模であり、知識獲得の問題が残されている。梶井らは、岩山らのプロトタイプモデルをベースとし、人手による概念の記述の代わりに自動作成した知識ベースを用いている<sup>11)</sup>。しかし、比喻性判定という立場をとっているため、比喻生成における有効性は実証されていない。一方、比喻表現の生成に関する研究としては酒匂らの研究がある<sup>12)</sup>。これは、認知心理学の心理実験から外見的特徴と心理的評価の 2 つの辞書を作成し、比喻表現を作成している。しかし、実際に使用しているデータは 145 語と小規模なものである。

そこで本論文では、電子辞書を用いた比喻による文章作成支援システムを提案する。提案システムは、梶井が行った認知心理学における実験から得られた比喻理解、生成に関する知見を工学的に導入している<sup>13),14)</sup>。これは、比喻の処理過程には、情緒・感覚的類似性やカテゴリ的類似性、修飾語との共起関係の強さが比喻の生成に大きな影響を与えるというものである。提案システムではこの考えに基づき、情緒・感覚的類似度、カテゴリ的類似度、共起度の 3 点を満たす比喻度と呼ばれる概念を工学的に導入し、柔軟な比喻の計算を行う。また、提案システムでは知識ベースとして、日本電子化辞書研究所の EDR 電子化辞書<sup>15)</sup>を使用している。この辞書は約 40 万概念の大規模な知識を有しており、従来法の欠点であった知識ベースの問題を克服している。

また提案システムは、基本的に発散的思考支援システムという立場をとっている<sup>3)</sup>。そして、ユーザの利用しやすい環境を考え、自然文を入力としている。ユーザは、比喻のもととなる文章をシステムに入力すると、システムは比喻候補および比喻特徴をユーザに提示する。そして、ユーザによって選択された言葉をもとに、最終的な比喻文章を生成する。

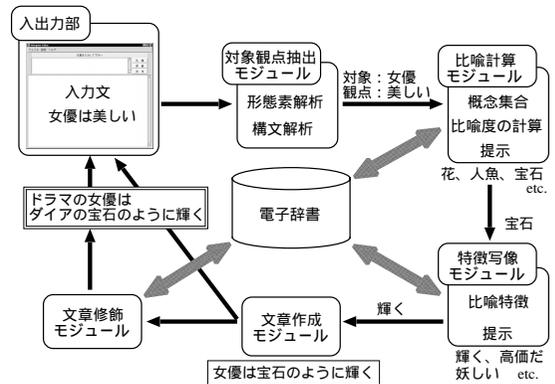


図1 提案システムの概要  
Fig.1 Proposed system.

## 2. 比喻による文章作成支援システム

### 2.1 提案システムの概略

ここでは、提案する電子辞書を用いた比喻による文章作成支援システムの概略について説明する。

提案システムの概略を図1に示す。ユーザが比喻による文章作成支援を受けたいとき、図2に示すテキストエディタ上部のテキストエリアに比喻のもととなる文章を入力する。対象観点抽出モジュールでは、入力された文章に対し形態素解析、および構文解析を行い、比喻の対象および観点を抽出する。比喻計算モジュールではこの対象と観点を利用して、観点を満たす比喻の初期概念集合を得る。そして、この集合の1つ1つの要素に対し、対象との比喻度を計算する。比喻度は、情緒・感覚的類似度、カテゴリ的類似度、および共起度を考慮に入れて計算する。システムは、比喻度の高い順にユーザに対し比喻候補を提示する。ユーザはこの中から自分が最も適すと考える比喻候補を選択する。特徴写像モジュールでは、選択された比喻候補と対象がそれぞれ関係する特徴を比較し、対象に存在しない特徴をユーザに対し提示する。ユーザはこの中から再び自分が最も適すと考える特徴を選択する。文章作成モジュールでは、入力文、対象、観点、比喻候補および比喻特徴をもとに比喻文の生成を行う。比喻の生成はここで終了するが、さらに文章の表現能力を高める目的で文章修飾モジュールが用意されている。文章修飾モジュールでは、入力された文章中の名詞的概念に対し、共起関係の高い修飾語を提示する。

提案システムでは、基本的に発散的思考支援ツールという立場で比喻文の生成支援を行う。これは、(1) 批判厳禁、(2) 自由奔放、(3) 質より量、といった原則に従い思考を支援していくことである<sup>3)</sup>。そのため、



図2 メインウインドウ(テキストエディタ)  
Fig. 2 Main window (Text editor).

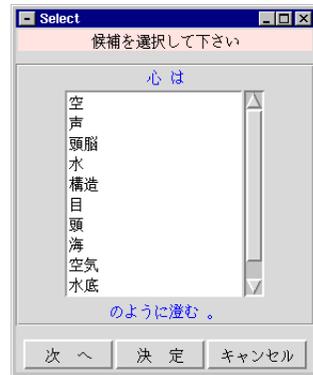


図4 比喩候補選択ウインドウ  
Fig. 4 Simile selection window.

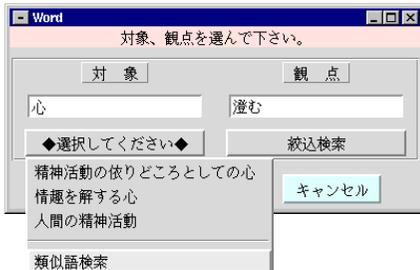


図3 対象観点抽出ウインドウ  
Fig. 3 Object and viewpoint window.

知識ベースとして膨大な知識を持った EDR 電子化辞書<sup>15)</sup>を利用している。そして、認知心理学的知見を工学的に導入して、情緒・感覚的類似度、カテゴリー的類似度、および共起度の3点を考慮に入れた比喩の計算方法を提案し、柔軟な比喩の計算を行う。

## 2.2 対象観点抽出モジュール

対象観点抽出モジュールでは、入力された比喩のもとになる文章から対象および観点の抽出を行う。ここで、対象とは何に注目して比喩文を生成するのかを表す言葉であり、観点はどのような視点で比喩文を生成するのかを表す言葉である。抽出された対象と観点は図3のようなウインドウで提示される。

文章が入力されると、初めに形態素解析を行う。提案システムでは、京都大学で開発された日本語形態素解析システムである JUMAN<sup>16)</sup>を利用している。

次に、形態素解析を行った結果をもとに構文解析、格解析を行う。自然言語処理における基本的な構文解析におけるルールを参考にし<sup>17)</sup>、表層格による構文解析、格解析を行い、基本的に主語を対象、述語を観点として抽出する。

## 2.3 比喩計算モジュール

比喩計算モジュールでは、対象、観点をもとに比喩候補の計算を行う。このとき、新たに比喩度  $R$  とい

う概念を導入する。比喩度  $R$  とは認知心理学における比喩の処理過程を工学的に取り入れたものである。これにより、図4に示すように比喩度の高い順に比喩候補を提示することができ、ユーザの選択を促すことができる。

桶見は心理実験により、(a) 比喩構成語の領域間距離を規定するカテゴリの意味と、(b) 比喩構成語の領域内距離を規定する情緒・感覚の意味を別々に測定した<sup>13)</sup>。その結果、比喩の処理過程について以下のようにまとめている。

- (1) 直喩、隠喩理解の出発点は、主題とたとえる語の字義どおりのカテゴリの意味における特徴対応によって、非類似性を検出することにある。カテゴリ的非類似性は比喩としての斬新さに影響を与える。
- (2) 主題とたとえる語の情緒・感覚の意味における特徴対応によって、類似性を検出する。この情緒・感覚的類似性が高いほど比喩が理解しやすくなる。
- (3) 比喩主題とたとえる語の共有特徴の修飾語で修飾すると、情緒・感覚的類似性が高まり、その結果として、比喩の理解容易性が高まる。

そこで、これら3つの意味を(1)情緒・感覚的類似度  $R_S$ 、(2)カテゴリ的類似度  $R_C$ 、(3)共起度  $C_o$  として求める。以降、これらを順次説明する。

### 2.3.1 情緒・感覚的類似度

まずはじめに、対象観点抽出モジュールで得た、観点および関係子から初期比喩候補を生成する。ここでは EDR 概念記述辞書から、観点と関係のある名詞的概念をすべて比喩候補として抜き出す。EDR 概念記述辞書では動詞と名詞の区別がされていないため、名詞的概念の判断は概念説明の内容をもとに行う。

次に、対象および比喩候補にそれぞれの上位概念を

表 1 EDR 概念記述辞書で用いられている関係子  
Table 1 Relation labels used in concept description dictionary.

関係子	意味	重み付け係数
agent	動作主格	$\alpha_{ag}$
object	対象格	$\alpha_{ob}$
goal	最終状態	$\alpha_{go}$
implement	道具格	$\alpha_{im}$
a-object	属性	$\alpha_{ab}$
place	場所	$\alpha_{pl}$
scene	場面	$\alpha_{sc}$
cause	原因	$\alpha_{ca}$

合わせて、関係している動詞の概念を調べる。上位概念も考慮に入れるのは、当該概念だけでなくその上位概念の性質も継承して考えた方が比較が行いやすいからである。そして、上位概念を含めた、対象と比喻候補がそれぞれ関係する動詞的概念がどれだけ一致しているかを表す情緒・感覚的類似度  $R_S$  を次式により求める。

$$R_S = \sum_i \alpha_i \times \text{関係子 } i \text{ の一致数} \quad (1)$$

ここで、 $i$  は表 1 に示す EDR 概念記述辞書で用いられている 8 つの関係子、 $\alpha_i$  は各関係子に対する重み付け係数である。また、 $R_S = 0$  となる比喻候補は順次削除する。

たとえば、図 5 に示すように「母」と比喻候補「海」を比較する場合、初めにそれぞれの概念と agent, object などの関係にある動詞的概念を 8 つの関係子について抜き出す。このとき「母」の上位概念である「父母」、「海」の上位概念である「水のある場所」についても同じように 8 つの関係子の関係を持つ動詞的概念を抜き出す。次に「母」、「父母」の動詞的概念の集まりと「海」、「水のある場所」の動詞的概念の集まりを比較し、一致している概念の数を関係子ごとに調べる。最後に、8 つの関係子の一致数を重み付けして式 (1) により情緒・感覚的類似度を求める。このようにすることでまったく異なる概念である「母」と「海」は、「育てる、穏やかだ、荒れる、…」といった同じ言葉で表されることが多いために、情緒・感覚的類似度が高くなる。

### 2.3.2 カテゴリ的類似度

次に、カテゴリ的類似度の計算を行う。認知心理学においては、比喻主題と、たとえる語のカテゴリ的非類似度が大きいほど比喻として斬新になるといわれている<sup>13)</sup>。しかし、提案システムにおいてカテゴリ的非類似度を利用すると、ユーザが求めている比喻候補と大きく異なるものが比喻候補の上位に提示されてしま

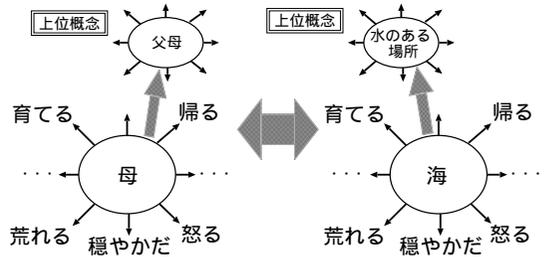


図 5 情緒・感覚的類似  
Fig. 5 Emotional and sensory similarity.

うために、違和感を与えてしまうことが多い。そこで、提案システムではカテゴリ的類似度を用いている。これにより、比喻候補の上位には一般的な比喻が、比喻候補の下位には斬新な比喻が提示されるようになる。

カテゴリ的類似度の計算には、EDR 概念体系辞書を用いる。概念体系辞書では、概念全体を木構造状に体系化して表しているため、木構造において対象と比喻候補のカテゴリ的近さを求める必要がある。中山らは、動詞的概念の類似度を求める方法を提案している<sup>18)</sup>。これを参考に、共通の上位概念の数に比例して概念間の類似度が高くなるようにしたカテゴリ的類似度  $R_C$  を次のように定義し、すべての比喻候補についてカテゴリ的類似度を求める。

$$R'_C = (dc - \gamma) \left( \frac{dc}{d_i} + \frac{dc}{d_j} \right) \quad (2)$$

$$R_C = \frac{R'_C}{R_C^{MAX}} \quad (3)$$

ここで、 $dc$  は共通の上位概念の個数、 $d_i$  は対象の上位概念の個数、 $d_j$  は比喻候補の上位概念の個数を表す。 $R_C^{MAX}$  は全比喻候補中最大の  $R'_C$  の値である。 $\gamma$  はパラメータである。 $\gamma = 0$  のときは共通の上位概念が 1 つ、 $\gamma = 1$  のときは共通の上位概念が 2 つ以上存在しないと  $R'_C = 0$  となる。また、 $R_C = 0$  となる比喻候補は順次削除する。

### 2.3.3 共起度

共起度は、比喻候補の中で日常的に頻繁に使われる組合せ、言葉が比喻候補の上位に提示されるようにするために導入する。これは、日常的に頻繁に使われる言葉を比喻候補として用いた方が理解しやすい比喻を作成できるからである。観点と比喻候補の共起度  $C_o$  は、EDR 概念共起辞書を使用し次のようにして求める。

$$C_o' = C_{ij} \times C_j \quad (4)$$

$$C_o = \frac{C_o'}{C_o^{MAX}} \quad (5)$$

ここで、 $C_{ij}$  は観点  $i$  と比喻候補  $j$  の表層共起頻度、



図6 比喩特徴選択ウインドウ  
Fig.6 Feature selection window.

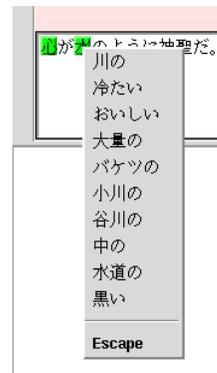


図7 修飾語の付与  
Fig.7 Addition of modifier.

$C_j$  は比喩候補  $j$  の係り側形態素頻度,  $Co^{MAX}$  は全比喩候補中最大の  $Co'$  の値である。

### 2.3.4 比喩度

比喩度は認知心理学的知見を工学的に導入し, これまでに述べてきた情緒・感覚的類似度, カテゴリ的類似度および共起度が加味されるようにしている。最終的な比喩度  $R$  は, 次式で計算する。

$$R \equiv R_S \times (1 + R_C) \times (1 + Co) \quad (6)$$

そして, 図4に示すように比喩度の高い順に比喩候補を提示する。

### 2.4 比喩特徴モジュール

比喩特徴モジュールでは, ユーザによって選択された比喩候補が関係している特徴が調べられ, その中で対象に存在しない特徴だけを図6に示すようにユーザに提示する。これは, 認知心理学における類推の機能を取り入れたものである。

類推とは, 知りたいこと, あるいはよく知らないことをよく知っていることにたとえて考えることを指している<sup>5)</sup>。比喩においても, この類推の写像は頻繁に行われる。たとえば「人が鳥のように飛ぶ」という文章よりも「人が鳥のように羽ばたく」とした方が比喩として面白く, また理解しやすいといえる。

比喩特徴モジュールでははじめに, ユーザによって選択された比喩候補に係る動詞的概念をEDR概念記述辞書を利用して調べる。同時に, 対象に係る動詞的概念についても調べる。そして, 比喩候補に係る動詞的概念の中で, 対象に係る動詞的概念に含まれないものをユーザに提示する。

### 2.5 文章作成モジュール

文章作成モジュールでは, 入力文章, 対象, 観点, ユーザにより選択された比喩候補および比喩特徴をもとに比喩文章を生成する。比喩文章は直喩とし, 観点が動詞の場合と, 形容詞・形容動詞の場合に分け, た

例えば以下のような比喩文を生成する。

観点が動詞のとき

- (1) [対象] は [比喩候補] のように [観点] 。
- (2) [対象] は [比喩候補] のように [比喩特徴] 。

観点が形容詞のとき

- (1) [対象] は [観点][比喩候補] のようだ。
- (2) [対象] は [観点][比喩候補] のように [比喩特徴] 。

入力文中の文節で, 比喩に直接係らないものはそのままの係り受け関係を保存するものとする。

生成された比喩文章は, 図2のテキストエディタ上部のテキストエリアに出力される。

### 2.6 文章修飾モジュール

文章修飾モジュールでは文章の表現能力を高めるために, 図7に示すように入力文章に含まれる名詞的概念に対し頻繁に用いられる修飾語を提示する。

修飾語はEDR概念共起辞書を利用して, 名詞的概念との共起関係の強さ  $S$  を次式により計算する。

$$S = \begin{cases} \beta_{ad} \times \text{表層共起頻度,} & \text{形容詞による共起} \\ \beta_{no} \times \text{表層共起頻度,} & \text{「の」による共起 (7)} \\ \beta_{di} \times \text{表層共起頻度,} & \text{直接共起} \end{cases}$$

ここで,  $\beta_{ad}$ ,  $\beta_{no}$ ,  $\beta_{di}$  は重み付け係数である。形容詞による共起とは, たとえば「綺麗な女優」のように形容詞によって修飾された名詞句を表す。助詞「の」による共起とは「ハリウッドの女優」のように名詞が助詞「の」によって接続された名詞句を表す。直接共起とは「ハリウッド女優」のように名詞2つが直接結合した名詞句を表す。

ユーザに対しては共起関係の強さ  $S$  の高い順に10個の修飾語を提示する。選択された修飾語は, ただちに名詞的概念の前に挿入される。

表2 評価実験の重み付け係数  
Table 2 Parameters of experiment.

$\alpha_{ag}$	$\alpha_{ob}$	$\alpha_{go}$	$\alpha_{im}$	$\alpha_{ab}$	$\alpha_{pl}$	$\alpha_{sc}$	$\alpha_{ca}$
4.0	1.0	2.0	0.0	3.0	0.0	0.0	0.0
$\beta_{ad}$	$\beta_{no}$	$\beta_{di}$					
2.0	1.5	1.0					

### 3. 評価実験

提案した比喻による文章作成支援システムに対し、様々な評価実験を行った。

#### 3.1 実験条件

評価実験に用いた条件は以下のとおりである。

- (1) 情緒・感覚的類似度を計算する式(1)の各関係子ごとの重み付け係数は表2に示すとおりである。これらの重み付け係数を決めるにあたり、あらかじめ予備実験を行った。文献19)を参考に代表的な比喻20組において、EDR概念記述辞書で用いられている8つの関係子ごとの関係概念の一致数を比較し、それぞれの関係子の比喻における重要性も考慮し重み付け係数を定めた。

- (2) カテゴリ的類似度を計算する式(3)におけるパラメータ $\gamma$ は次のとおりとした。

$$\gamma = \begin{cases} 0, & \text{比喻候補が30語以上のとき} \\ 1, & \text{比喻候補が30語以下のとき} \end{cases} \quad (8)$$

これにより、比喻候補が数多く存在するときは検索条件を厳しくし、逆に少なければ検索条件を緩くしている。

- (3) 文章修飾モジュールにおける共起関係を求める式(7)の重み付け係数は表2のとおりとした。
- (4) 比喻のもととなる入力文は、基本的に単文とした。これは、複文などの文章では文章中のどの部分を比喻にするのが決められないためである。

#### 3.2 比喻の結果

提案システムを利用して作成された比喻の結果の例を、入力文とともに示す。

- 入力 あの月は赤かった。  
出力 あの月は赤い夕日のように美しかった。
- 入力 女優は美しい。  
出力 噂の女優は宝石のように美しい。
- 入力 飛行機が飛んできた。  
出力 敵の飛行機が数多くの渡り鳥のようにやってきた。
- 入力 人生は楽しい。  
出力 僕の人生は楽しい旅のようだ。

表3 比喻候補(入力: 月光是明るい)

Table 3 Candidates of simile (Input: Moonlight is bright).

比喻候補	カテゴリ的類似度	情緒感覚的類似度	共起度	比喻度
光	1.00	2.00	1.00	8.00
陽射し	0.93	4.00	0.01	7.77
空	0.04	3.00	0.12	3.49
太陽	0.23	2.00	0.17	2.89
閃光	0.69	1.00	0.01	1.71
対象	0.38	1.00	0.00	1.38
表情	0.03	1.00	0.31	1.34
部屋	0.21	1.00	0.08	1.30
壁	0.22	1.00	0.07	1.30
天体	0.23	1.00	0.01	1.24
照明	0.19	1.00	0.02	1.21

表4 比喻候補(入力: 心が澄んでいる)

Table 4 Candidates of simile (Input: Heart is clear).

比喻候補	カテゴリ的類似度	情緒感覚的類似度	共起度	比喻度
空	0.08	5.00	1.00	10.78
声	0.05	10.00	0.00	10.45
頭脳	0.68	5.00	0.00	8.39
水	0.04	8.00	0.00	8.34
構造	1.00	4.00	0.00	8.00
目	0.04	6.00	0.00	6.21
頭	0.04	5.00	0.00	5.18
海	0.05	3.00	0.00	3.14
空気	0.05	2.00	0.00	2.09
水底	0.06	1.00	0.00	1.06
血液	0.04	1.00	0.00	1.04

- 入力 彼の心は澄んでいる。  
出力 彼の心は青い海のように澄んでいる。
- 入力 人生は果てしなく永い。  
出力 人生は果てしなく永いいばらの道のようだ。
- 入力 人生ははかない  
出力 人生ははかない生命のように尊い。
- 入力 満月は丸い。  
出力 満月は丸い船窓のようだ。
- 入力 月が光る。  
出力 月が悲しみの涙のように光る。

#### 3.3 比喻生成の評価

##### 3.3.1 比喻候補の結果

入力文を与えたときに、システムが実際に提示した比喻候補を表3、表4に示す。各比喻候補に付け加えて、カテゴリ的類似度、情緒・感覚的類似度、共起度および比喻度も示す。

##### 3.3.2 比喻候補の評価

ここでは、比喻度の計算によって求められた比喻候補についての検討を行った。比喻に対する感じ方は、個人の感性により大きく異なるため、実験ではアン

表5 比喩候補の評価の累積割合 (%)

Table 5 Accumulated evaluation of simile candidates (%).

入力文章	比喩候補数				×
人生は哀しい	3	30.0	70.0	93.3	100
人生は永い	38	18.4	46.8	77.9	100
一生ははかない	4	22.5	42.5	87.5	100
計画が広まる	23	21.7	50.0	79.1	100
天気は変わる	16	20.6	46.3	72.5	100
飛行機が飛ぶ	31	28.1	66.8	80.6	100
森は暗かった	15	16.0	48.0	77.3	100
心は澄んでいる	11	41.8	71.8	89.1	100
月光は明るい	11	30.9	61.8	82.7	100
月は明るい	19	21.6	48.9	78.4	100
月は丸い	26	9.2	34.2	67.3	100
雪が舞っている	24	25.0	49.6	78.8	100
論文が書けた	8	21.3	66.3	88.8	100
洋服が青い	22	17.3	49.5	74.5	100
平均	17.9	23.2	53.6	80.6	100

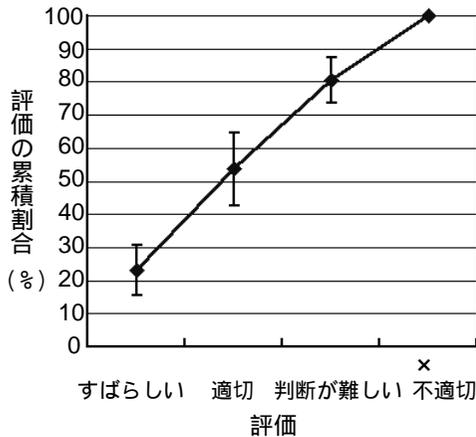


図8 利用可能な比喩候補の割合

Fig. 8 Percentage of usable simile.

ケートによる定量的な評価を行った。異なる14の文章を提案システムに入力したときに、提案システムが提示したすべての比喩候補を被験者10人に対してすばらしい比喩候補である(○), 比喩として適切である(△), 判断が難しい(◇), 比喩候補として適切でない(×), の4段階で評価してもらった。被験者は20歳から24歳の若者であり, システムの操作説明だけをあらかじめ聞いており, 実際に操作は行っていない。

表5は, 各入力文に対する4つの評価の累積割合を示している。また図8は, 平均の利用可能な比喩候補の割合を示している。各プロットには, 標準偏差もあわせて示されている。は, 評価およびを合計したものであり, あわせて比喩候補として適切である割合を示している。入力した文章によりパラツキはあ

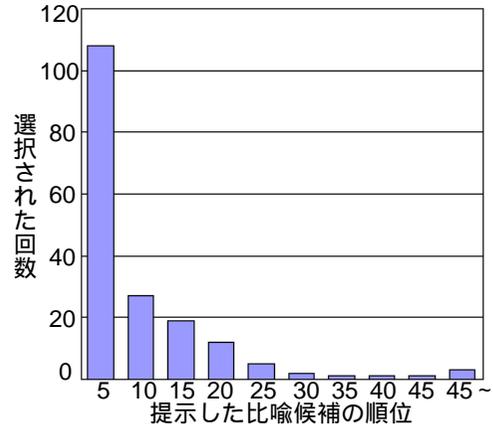


図9 提示した比喩候補の順位に対する選択された比喩候補の割合

Fig. 9 Number of selected simile with order of simile.

るものの, 平均では54%の割合で比喩候補として適切であるという結果が得られた。の判断が難しい比喩候補は多くの人に適切であると判断されるわけではないが, 何人かの被験者はこれらの候補が使えると判断したことを考え, まで累積したものは, 明らかに不適切ではない比喩候補の割合を示すと考えられる。この結果, 平均で81%の比喩候補が利用できる可能性があるといえる。

### 3.3.3 提示した比喩候補が選択された割合

ここではユーザが提案システムを利用したときに, 提案システムが提示した比喩候補の順位に対する選択された比喩候補の割合を調べた。図9に, ユーザが作成した比喩文179文に対し, システムが提示した比喩候補の順位に対する選択された比喩候補の割合を示す。また図10には, システムが提示した比喩候補の個数の割合をまとめた。

図9より, 提案システムが上位に提示した比喩候補がユーザによっても頻繁に選択されていることが分かる。これは, 比喩度が高い比喩候補はユーザによっても比喩候補として適切であると判断されていることを示している。また図10から, システムが提示した比喩候補の個数は, 入力文により変化するが15~20個前後の候補数に絞り込んでいることが分かる。

## 3.4 システムの評価

### 3.4.1 評価方法

ここでは, 提案システムによって生成された比喩文章の評価, および思考支援システムとしての評価を行った。実験は, 例題を定めて実際にユーザに比喩文を生成してもらい, その評価についてまとめた。被験者15人を3つのグループに分け, 第1グループには「人生」に関する比喩を提案システムを利用して作成

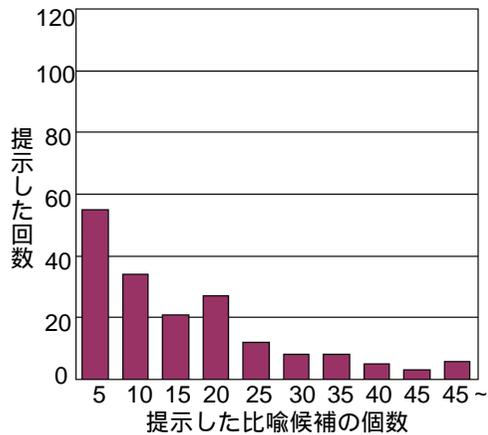


図 10 提示した比喩候補の個数  
Fig. 10 Number of shown simile.

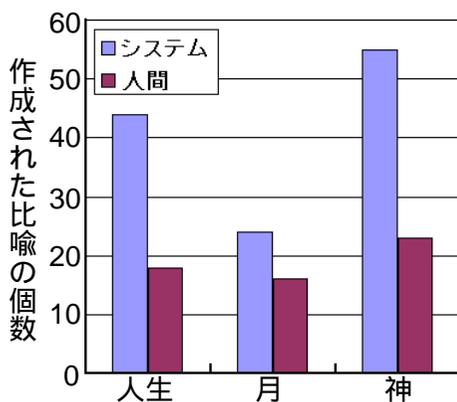


図 11 作成した比喩の数  
Fig. 11 Number of simile obtained by users.

してもらった。第2グループには「人生」に関する比喩を自分で考えてもらった。ここで、特に時間の制限は設けなかった。また人が考える比喩は、2.5節で述べたシステムが生成する比喩文と同じ形式に限定してもらっている。第3グループは、生成した比喩の評価を行ってもらった。また、グループを変え同様の実験を例題「月」「神」についても行った。被験者は20歳から24歳の若者であり、システムの操作説明だけをあらかじめ聞いており、実際に操作は行っていないものとする。

### 3.4.2 思考支援システムとしての評価

評価は、発散的思考支援システムの原則に基づき、作成することができた比喩の数の合計を比較した。その結果を図11に示す。図11より、いずれの例題でも提案システムの支援を受けることで全体としてより多くの比喩を生成できたことが分かる。一方、人が考えた比喩の数は比較的少なかった。これは、人は慣用的

表 6 比喩の平均点

Table 6 Average score of figurative sentences.

	提案システム による比喩	人が考える 比喩	比喩表現辞典 の比喩
人生	3.09	3.72	3.42
月	3.06	3.36	2.92
神	3.19	3.58	3.25
平均	3.12	3.56	3.15

な比喩表現をいくつか思い出すと、それ以上比喩を考えることが困難となる傾向があるためと考えられる。これに対し、提案システムを利用すると、簡単な文章をもとに比喩を生成することができるため、より多くの比喩を生成できている。

### 3.4.3 比喩のチューリングテスト

ここでは、比喩文章のチューリングテストを行った。チューリングテストとは、質問に対する解答がコンピュータによるものなのか人間によるものなのか判断できないとき、コンピュータは知的であると判定する方法である。初めに、提案システムを利用して作成した「人生」「月」「神」に関する比喩123文、人が考えた比喩57文、および比喩表現辞典<sup>20)</sup>より抜き出した比喩46文をランダムに提示し、第3グループに比喩としての評価を1~5の5段階でつけてもらった。比喩表現辞典は、現代日本の文学作品に記載されている比喩表現をまとめたものであり、ここでは2.5節で述べたシステムが生成する比喩文と同じ形式のものを利用した。

この結果を表6に示す。平均点に関して、提案システムにより作成された比喩文の平均点は人が考えた比喩文の平均点よりも低かったが、5段階評価で3.12と大幅に低い数字ではなかった。これは、人はある程度誰にも分かりやすい慣用的な比喩を考えるからであると考えられる。一方、提案システムを使用すると、前節で述べたように数多くの比喩を生成することができるが、反面、比喩の完成度にばらつきがあると考えられる。

そこで、評価4もしくは5という高い評価を得た比喩文について検討を行った。表7に、評価4もしくは5という高い評価を得た比喩文を合計したものの割合を示す。表7より、高い評価を得た比喩の割合と全比喩合計の割合がそれほど変わっていなかった。このことから、提案システムを利用することで、人間が考える比喩や比喩表現辞典に載っている比喩と同等な比喩表現を生成することが可能となっていることが分かる。

### 3.4.4 ユーザの感想・意見

ここでは、ユーザの主観的な感想、意見に基づく評

表7 評価4, 5の比喩の割合(%)

Table 7 Percentage of figurative sentences evaluated 4 or 5 (%).

		提案システムによる比喩	人が考える比喩	比喩表現辞典の比喩	合計
人生	評価4, 5	52.3	24.7	23.0	100
	全比喩	56.4	23.1	20.5	100
月	評価4, 5	38.7	26.6	34.7	100
	全比喩	38.7	25.8	35.5	100
神	評価4, 5	56.0	35.7	8.3	100
	全比喩	64.0	26.7	9.3	100
平均	評価4, 5	50.3	29.5	20.2	100
	全比喩	54.4	25.2	20.4	100

価をまとめてみた。

多くのユーザから、自分では思い浮かばないような比喩を作成することができたという感想を得た。一方で、ユーザが望んだ言葉が利用できない、出てこないといった意見も聞かれた。これは、提案システムの知識ベースによるところもあると考えられる。提案システムにおいて用いられるEDR電子化辞書は、新聞や学術的論文をもとに作成されているので、逆に我々が普段使用するような日常的な言葉が辞書に載っていないこともあるためであると考えられる。明らかに不適切な候補をできる限り少なくしていくことが今後求められる。

#### 4. 結 論

本論文では、大規模な知識を有する電子辞書を利用し、認知心理学に基づく比喩による文章作成支援システムを提案した。認知心理学において、比喩の処理過程には情緒・感覚的類似性やカテゴリーの類似性、修飾語との共起関係の強さが比喩の生成に大きな影響を与えるといわれている。そこで、提案システムにおいても比喩度と呼ばれる概念を導入し、認知心理学における比喩理解、生成に関する知見を工学的に導入した。また、知識ベースとして大規模な知識を有するEDR電子化辞書を利用した。そして、発散的思考支援ツールという立場に立ち、文章作成支援システムという形で、実際にユーザが使用できるシステムを実装した。詳細な実験により、提案システムの有効性が確認された。

#### 参 考 文 献

- 1) 国藤 進：オフィスにおける知的生産性向上のための知識創造方法論と知識創造支援ツール，人工知能学会誌，Vol.14, No.1, pp.50-57 (1999).
- 2) 国藤 進：発想支援システムの研究開発動向と

その課題，人工知能学会誌，Vol.8, No.5, pp.552-559 (1993).

- 3) 折原良平：発散的思考支援ツールの研究開発動向，人工知能学会誌，Vol.8, No.5, pp.560-567 (1993).
- 4) 杉山公造：収束的支援ツールの研究開発動向—KJ法の支援を中心として，人工知能学会誌，Vol.8, No.5, pp.568-574 (1993).
- 5) 鈴木宏昭：認知科学モノグラフ1類似と思考，共立出版 (1996).
- 6) 諏訪正樹，岩山 真：比喩の計算モデル，情報処理，Vol.34, No.5, pp.566-575 (1993).
- 7) 森 辰則，中川裕志：意味マッチングによる比喩理解モデル，情報処理学会論文誌，Vol.32, No.3, pp.345-353 (1991).
- 8) 土井晃一，佐川浩彦，田中英彦：隠喩理解—命題分解によるニューラルネットワークの利用，情報処理学会研究報告，90-NL-75, pp.1-7 (1990).
- 9) 岩山 真，徳永健伸，田中穂積：比喩を含む言語理解における顕現性の役割，人工知能学会誌，Vol.6, No.5, pp.674-681 (1991).
- 10) 内海 彰：比喩理解モデル，信学技報，TL97-5, pp.33-40 (1997).
- 11) 榊井文人，福本淳一，田添丈博，椎野 努：統計的手法を用いた比喩認識，信学技報，TL98-7, pp.25-32 (1998).
- 12) 酒匂孝之，中村順一，吉田 将：概念間の外見的な類似性と心理的な評価を利用した比喩表現の生成，信学技報，NLC93-20, pp.17-24 (1993).
- 13) 楠見 孝：比喩理解の構造，芳賀 純，子安増生(編)，メタファーの心理学，第3章，誠信書房 (1990).
- 14) 楠見 孝：比喩の処理過程と意味構造，風間書房 (1995).
- 15) 日本電子化辞書研究所：EDR電子化辞書仕様説明書(第2版)，EDR TR-045 (1995).
- 16) 黒橋禎夫，長尾 真：日本語形態素解析システムJUMAN version 3.61，京都大学 (1998).
- 17) 長尾 真：自然言語処理，岩波講座ソフトウェア科学，15，岩波書店 (1996).
- 18) 中山 聡，峯 恒憲，東 優，谷口倫一郎，雨宮真：EDRコーパスを利用した動詞の語義分類，信学技報，NLC95-43 (1995).
- 19) Holyoak, K.J. and Thagard, P. (著)，鈴木宏昭，河原哲雄(監訳)：アナロジーの力，新曜社 (1998).
- 20) 中村 明：比喩表現辞典，角川書店 (1995).

(平成12年3月31日受付)

(平成13年2月1日採録)



北田 純弥

昭和 51 年生．平成 10 年慶應義塾大学理工学部電気工学科卒業．平成 12 年同大学大学院理工学研究科電気工学専攻修士課程修了．現在，日本アイ・ピー・エム（株）勤務．



萩原 将文（正会員）

昭和 34 年生．昭和 62 年慶應義塾大学大学院博士課程修了．現在，同大学助教授．平成 3 年より 2 年間アメリカ Stanford 大学訪問研究員．ニューラルネットワーク，ファジーシステム，GA の研究に従事．工学博士．昭和 61 年丹羽記念賞，昭和 62 年電子情報通信学会学術奨励賞，平成 2 年 IEEE 論文賞，平成 6 年安藤記念学術奨励賞，平成 8 年ファジィ学会著述賞受賞．

---