

5N-4

# 広告などにおける感性つきことばの概念ベースによる生成の可能性

阿部 明典

NTT コミュニケーション科学基礎研究所

## 1 はじめに

インターネットの発展に伴い、マルチメディアが注目されているが、実は画像以上に「言葉」の重要性が、その膨大なデータ量と共に再認識されつつある。このような大衆発信的な情報 / 知識の伝達においては、言葉は、古典的コミュニケーションの定義で使われているような単に情報を伝えるメディアとしてだけではなく、旧来より、広告等で実践されてきた多分に感性的 / 感性的な側面 [4] が意味を持って利用されている。言葉のこうした側面を扱うには、従来の工学的な発想だけでは限界があり、非工学分野との連携、それも心理学や社会学等のように今まで AI が連携を深めてきた分野だけでなく、文学や芸術等の比較的疎遠だった分野をも巻き込んだ検討が実りある発展をもたらすと考える。我々は、このような知見に基づき「ことば工学 [1]」という研究分野を提唱した。

ことば工学では、感性的 / 感性的なことばの扱いを目指す書いたが、当然、ことば自体の性質（世界の中のことばなど）、特性、更に、ことばを使う時の特徴なども研究対象として考えており、これらの性質などの工学的実現を目指している。更に、ことばと身体の関係など、ことばが発生する意味、理由なども研究対象として考えているが、本論文ではこれら、ことばの本質に関しては言及しない。

一方、これらの性質を踏まえたことば工学の応用としては様々なことが考えられるが、例えば、ことばの意味を見せるチャット、「くりはくりでも口からでるくりはなあにと」いったタイプのなぞなぞ（一種の Rebus Riddles）を解く、クロスワードを解く [2]、駄洒落や humour の自動生成を行なうといったことば遊び、流行することばの予測、ことばのデザイン [5] などといったことがこれまでに行なわれている。

本論文では、それ自体の見た目、更に、読んだ時の韻、想起されることなどを効果的に表している広告や詩などで見られる効果を持つことばを半自動的に生成する手法を示す。具体的には、ことばの感性的な面を扱えると思える概念ベース [3] を利用することにより、感性を持つと思われることば（キャッチコピー）の生成について述べる。

*Generation of Words with Sensitivity by using the Concept base*  
Akinori Abe  
NTT Communication Science Laboratories  
2-4, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0237  
JAPAN

## 2 感性を持つことばの扱いについて

### 2.1 広告のことば

広告では、見る人に二重、三重の意味を与えられるようなキャッチコピーをよく使う。限られた時間、スペースで多くのことを伝えるという意味もあるのであるが、それ以上に、いかに宣伝しているものを印象に残らせるかという意味からも、そのようなキャッチを使う。例えば、最近の広告の例では、「男女驚愕」（大手前大学）がある。もともと女子大だったのが共学になったので、びっくり という意味であろう。又、最近の東京近郊の JR 東日本のスキー関係の広告に「愛に雪、恋を白」がある。「会いに行き、恋をしろ」となるはずの文章がスキーとはカップルで行くものか (?) という前提のもとに恋愛の「愛」という文字が与えられ、雪山に行くという context の下に「雪」、「白」という文字が与えられていて、ごく普通、いや、陳腐といってもいいかも知れないフレーズが印象的なフレーズになっている。このように、広告等では、別の漢字を使うことで、二つの意味を持たせたり、別の意味を強調させたり、更には、そこには使われるはずのない、同じ音の漢字を使うことで、印象の広がりをもたせたりといった、ことばに感性を与える手法が良く用いられる。

JR 東日本の使ったキャッチを使って感性を持つことばの一つを説明をした。扱、この感性は、人間の経験から想出されるものであるが、アルゴリズムでもある。従って、（余り良くないが、）工学的にいつてしまえば、“ある文脈のもとに、文脈から想起され、且つある「音」を持つことばをさがせ” というタスクに相当しよう。更に、キャッチは駄洒落、ニアミスのことばとも親和性が良い。この場合は、“ある文脈のもとに、文脈から想起され、且つある音に似た「音」を持つことばをさがせ” というタスクに相当しよう。

### 2.2 概念ベース

概念ベース [3] は語をベクトルで表現し、ベクトル間の距離（内積）で語同士の類似性（関連度とも考えられる）を判別するための辞書である。概念ベースでは、 $N$  個の語の各々の意味（概念  $g_i$ ）を、同じく  $N$  個の語から選んだ  $g_i$  の特徴を表す語（属性  $p$ ）、および  $p$  の重要度  $q$  の対の集合で表す。すなわち、

$$\text{概念 } g_i = \{(p1, q1), (p2, q2) \dots\} \quad (1)$$

となる。ここで各  $p, q$  は、国語辞典等における見出し  $g$  に対する語義文から、自立語とその出現頻度として機械的に抽出する。例えば語「馬」について「首と尾に毛が生えた動物で、走るのが早い動物・・・」云々との語義が書かれていた場合、

$$\text{「馬」} = \{(\text{首}, 1), (\text{動物}, 2) \dots\} \quad (2)$$

等となる<sup>11</sup>。

この結果、概念ベースは  $N=4$  万の語を各々  $N$  次元の属性ベクトルとして表現した  $N \times N$  の行列として構築される。これを用いて、ある語  $a$  に対し  $a$  に関連した語  $K(a)$  を、語  $a$  のベクトルが含む属性  $p$ 、語  $a$  を属性  $p$  として含む語  $b$ 、語  $a$  と似た属性ベクトルを持つ語  $c$ 、等として検索することができる。更に、このベクトル間の関係は、観点というものの導入により、観点に従って変わる。

本節では概念ベースの特性に就いて述べたが、概念ベースによって計算される語同士の関連度はことばの感性の一面を表しているものである。以下の節では、この概念ベースにより計算される語同士の関連度を活用してことばの感性的な面を扱うことを示す。

### 3 感性を持つことばの生成例

前節では、感性を持つことばを広告のことばを例に就いて述べた。本節では、感性を持つことばの自動生成の可能性に就いて述べる。前節に示したように、概念ベースでは、語同士の関連度を計算することが出来る。従って、これを利用して、あることばに、関連のある言葉を生成することで、感性つきのことばを生成するのであるが、簡単な例としては、以下の手法が考えられる。以下の例では、さしあたって、入力文として、ことわざのようなものを使う。

- ある「音」を持つ語で、あることに関連のあることばを使う

ある文(例えば、ことわざなど)を品詞ごとに分け、名詞(句)、動詞(句)などを対象にして、文中の語の変換を行なう。この場合、観点により、語同士の関連性が変わることを利用する。

例えば、「かわいい子には旅をさせよ」という文を使った場合、この文は

かわいい / 子 / に / は / 旅 / を / させ / よ

と分解出来るが、文中に「旅」があるので<sup>12</sup>、水辺へ<sup>13</sup>の「旅(行楽)」という観点でキャッチを作成して見る。「行楽」という観点で、「水辺」に関連があつて、夫々の語(名詞と、動詞に限る)と同じ音を(語の一部にでも)持つことばを概念ベースを使って検索すると、

川(かわ(い)), 湖(こ), 瀬((さ)せ)が見つかり、それらの文字に該当の語を変換することにより、

「川い湖には旅をさ瀬よ」

というキャッチが出来上がる。

- ある音と似た「音」を持つ語で、あることに関連のあることばを使う

<sup>11</sup>実際には、まずこのようにラフな概念ベースを作成した後で、概念ベース自身を参照しながら属性データの質の向上をはかっている

<sup>12</sup>観点は、文の中にある必然性はない。他からとってきてよい

<sup>13</sup>これは、自由に選んでいる

上記の過程ではことばが見つからない場合に行なう。但し、こつちを使った方が響きなどが良いこともある。実際の広告などでは、駄洒落などが多用されている。従って、駄洒落を生成するとか、Rebus Riddles 型のなぞなぞを解くことと同じような技術を使うことが可能である。

上記の手法では、観点はまだ手入力であり、出来上がったキャッチも余り美しくないかも知れないが、概念ベースを利用することで、(広告という文脈で)感性を持つことばが半自動的に生成出来るという可能性は示していると思う。

### 4 まとめ

本論文では「ことば工学」の一環として、感性を持つことばの生成を示した。概念ベースにより得られる関連度により、ある意味を持つ言葉を生成し、それを一種の感性を持つことばとして考えた。当然、全てがうまくいくわけではないが、概念ベースを関連語検索に使うだけでもある程度簡単な感性を持つことばを生成出来ると思われる。

しかし、本論文に示したものは、ほんのプロトタイプなので、おもちゃ程度の動きしかしない。更に様々な戦略などをつけることにより、更に豊かな感性を持つことばを生成でき、キャッチなどを作るときに支援出来るものになると思うとともに、キャッチの生成だけではなく、更に創発的なことも出来るようになると思われる。

広告等を作る場合、観点は文脈や絵の情報から得られることが多い。ここに示したものは、観点は与えるものとして考えているが、観点を絵などの情報から得る研究もこれからの課題である。

尚、本論文で言及している「ことば工学研究会」(<http://www.kecl.ntt.co.jp/banana/Workshop/>)では、これだけではなく、様々なことばに関する話題を議論している。

### References

- [1] 阿部 明典: ことば工学の地平線 — “あとがき” にかえて... —, 人工知能学会研究会資料, SIG-LSE-9901-12 (1999)
- [2] 金杉 友子, 松澤 和光 and 笠原 要: アバウト推論の「ことば遊び」への適用, 信学技報, NLC96-31 (1996)
- [3] 笠原 要, 松澤 和光, 石川 勉: 国語辞書を利用した日常語の類似性判別, 情報学会論文誌, Vol. 38, No. 7, pp. 1272-1283 (1997)
- [4] McLuhan M. and Fiore Q. (南博 訳): メディアはマッサージである, 河出書房新社 (1995; 原書: 1967)
- [5] 須永 剛司: 情報の形は何処にあるの? デザインの新しい学びの視点から, 情報デザイン学論考, pp. 149-164, 多摩美術大学 (1998)