

感性情報を利用したオノマトペ学習システムの開発

中部 文子[†] 浅賀 千里^{††} 渡辺知恵美^{††}

[†] お茶の水女子大学理学部情報科学科 〒112-8610 東京都文京区大塚2丁目1番1号

^{††} お茶の水女子大学大学院人間文化創成科学研究科 〒112-8610 東京都文京区大塚2丁目1番1号

E-mail: [†]{nakabe,asaga}@db.is.ocha.ac.jp, ^{††}chiemi@is.ocha.ac.jp

あらまし 日本語のオノマトペは感覚的な語であり、外国人が習得するには多くの用例に触れることが重要になる。そこで我々は、web からオノマトペを自動取得・分類し用例と共に提示するシステムの開発を行っている。本研究では、オノマトペの持つニュアンスや聞き手に与える印象に着目し、明るい印象、強い印象、といったようなオノマトペが人に与える印象・ニュアンスを自動計算し、提示する機能開発を目指す。あらかじめ用意しておいた複数の感性を評価する尺度に関して、オノマトペの感性評価値を求める。感性評価値の求め方は、オノマトペ自身の音の印象を数値化する、周辺単語を重要度に応じて数値化する、の二つの方法で行う。

キーワード web 情報システム e-ラーニング

Onomatopoeia Learning System With Aesthetic Information

Fumiko NAKABE[†], Chisato ASAGA^{††}, and Chiemi WATANABE^{††}

[†] Department of Information Sciences, Faculty of Sciences, Ochanomizu University 2-1-1 Otsuka, Bunkyo-ku, Tokyo, 112-8610 Japan

^{††} Ochanomizu University Graduate School of Humanities and Sciences 2-1-1 Otsuka, Bunkyo-ku, Tokyo, 112-8610 Japan

E-mail: [†]{nakabe,asaga}@db.is.ocha.ac.jp, ^{††}chiemi@is.ocha.ac.jp

Abstract Japanese is filled with onomatopoeic words, which describe sounds or actions like "click" or "bow-wow". In general, mastering onomatopoeia phrases is hard for foreign speakers. Then, we develop the system for foreign people to learn Japanese onomatopoeia by collecting onomatopoeia and around text from Web. Especially it is difficult to understand the feel and atmosphere which onomatopoeic word originally have. In this paper, we propose the expression method the feelings and atmosphere for onomatopoeic words as a set of adjective words such as "bright", "sad" and "intensive". We use two types of information, they are the vocal of the onomatopoeic words and words which are appeared in the example sentences which use onomatopoeic word.

Key words Web Information Systems e-learning

1. はじめに

オノマトペとは「にこにこ」「しっとり」などのいわゆる擬音語・擬態語である。微妙なニュアンスを伝えることができコミュニケーション上重要である。さらに、外国で日本の漫画やアニメが見られることが多くなり日本語のオノマトペへの関心が高まっていると考えられる。しかし、オノマトペが感覚的な語であることや、外国語に対応語が少ないこと、一つのオノマトペが複数の意味を持つことなどから、日本語学習者にとってオノマトペの習得は難しい。我々はこれまでに、オノマトペを習得するためのオンライン用例辞書「オノマトペディア」を開発してきた。オノマトペディアは Web 空間からオノマトペを含ん

だ文章を収集し、文例として適切な文章を提示している。しかし用例だけでは、日本人がそのオノマトペを聞いたときに感じる微妙なニュアンスまでは伝わりにくい。例えば、「彼はあなたを見てにっこり笑っていたよ」と「彼はあなたを見てげげら笑っていたよ」という文章では、オノマトペ部分以外はまったく同じにもかかわらず、前者は好意的、後者は否定的、とまったく印象が違ってくる。また、「雨がぼつぼつ降っている」「雨がざあざあ降っている」なども、日本人ならばまったく違う情景が想像できる。しかしオノマトペに初めて触れる人はこの感覚が得られにくいと考えられる。そこで本研究では、オノマトペを使ったり聞いたりする際重要となるオノマトペの持つ微妙なニュアンスや聞き手に与える印象に着目し、それらを提

示するシステムを提案する。本稿では、第2節にてシステム概要とユーザへの提示方法、第3節、第4節、第5節にてシステム詳細、第6節で評価を述べる。

2. オノマトペの感性情報提示システム

2.1 システム概要

本システムは、Web から抽出したオノマトペとその周辺文章を分析し、オノマトペの感性情報を提示する。これによって、オノマトペを聞きなれない人でも、そのオノマトペのニュアンスや使うべき状況が容易に想像できるようにしたい。感性を測る指標を複数用意し、オノマトペとその周辺文章から、各指標に対する感性評価値を数値として計算・集計する。その後、特徴的と判断した値のみユーザに見やすい形にして提示する。

2.2 評価尺度・ユーザへの提示方法

図1に感性情報のユーザへの提示方法を示す。まず、感性を評価するために、複数の指標を用意する。感性を評価する指標とは、形容詞とその反対語の対であり評価尺度と呼ぶことにする。例えば「明るい 暗い」「大きい 小さい」などである。感性評価値は、評価尺度の一方の形容詞に最も近い時は-2、もう一方の形容詞にもっとも近いとき2として、-2から2の値をとるようにする。オノマトペ辞典の「オノマトペ にかかる動詞」の横に、集計した感性のうち最も特徴的な評価尺度とその感性評価値が表示され、“より詳しく”をクリックすると、集計した全評価尺度についての値を見られるようにする。

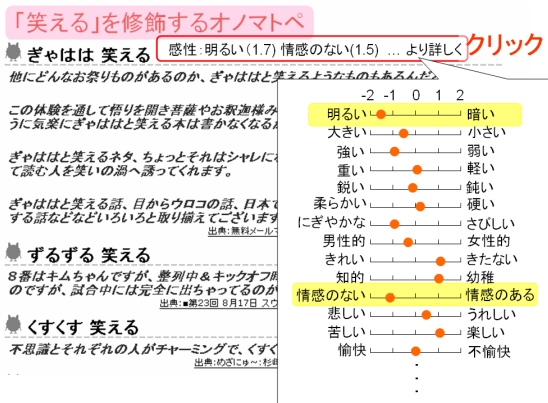


図1 ユーザへの提示方法

2.3 システム全体の流れ

図2にシステムの流れを示す。Web からオノマトペを含む文例をデータベースに格納しておき、データベースから取得した情報をもとに大きく分けて以下の二通りの方法で感性評価値を求める。

- オノマトペの語感から求める
- オノマトペ周辺の単語から求める

それぞれの詳細について3節、4節で述べる。この二通りで算出した値を集計し、最終的なオノマトペの感性評価値とする。

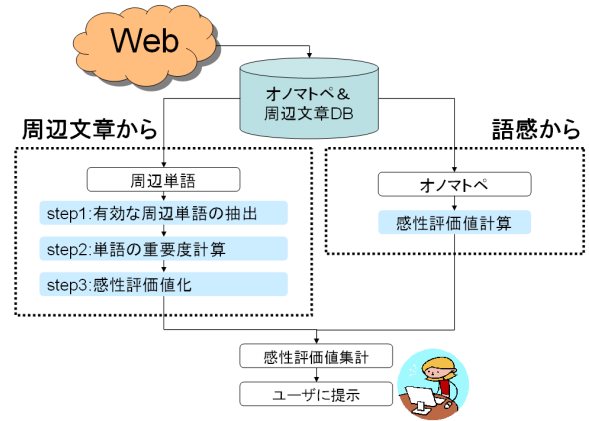


図2 システム概要

3. 語感から感性評価値を求める

3.1 語感とは

感性評価値を求める一つの方法として、「語感」を利用する。語感とは、言葉から受ける印象のことである。例えば、「あ」という音は明るくて強く大きいイメージ、「ん」は重く格調があるイメージ、のように音から人が感じるイメージが研究されている[5]。これらを組み合わせることで単語の印象を知ることができる。

3.2 語感の数値化

感性語で表現される「語感」を数値として表すために、図3に示す音のイメージ表を参考にすることにした[5]。全アルファベットの感性が10個の評価尺度に対して、○, ◯, ▲, マークなしで表現されている。をとてもそうだ、○をどちらかというとそうだ、◯をどちらでもない、▲をどちらかというとそうでない、マークなしを全くそうでないと判断し、それぞれ2, 1, 0, -1, -2と数値化し重みとし使用する。これを用いてオノマトペを構成するアルファベットの音の感性を出現割合に応じて計算し、オノマトペの語感を各評価尺度の値として表す。値は-2から2の実数になるようにする。

3.2.1 準備：評価尺度

使用する評価尺度は、音のイメージ表にある形容詞とその対義語を対にした、図4の12対とする。

3.2.2 感性評価値を求める流れ

まず、データベースからひらがなの対象オノマトペを取得しローマ字表記に変換する。これは、使用する語感情報がアルファベットについて定義されているためである。そして、オノマトペを構成する各アルファベットについて、以下の式で感性評価値を求める。

	明るさ	強さ	大きさ (元から)	重さ	軽さ	鋭さ	やわらかさ	元気	男性的	女性的
b										
e	○	△								
d				○					○	
f					○					○
g				○						
h					○					
j										
k										
l	△		△							
m										
n	△	○		△						
p					○				○	
q										
r	△		△							
s										
t		○								
v										
w										
x										
y										
z										

図 3 音のイメージ表

-2 -1 0 1 2
 フォーマル ——— カジュアル
 知的 ——— 幼稚
 情感のある ——— 粗野
 弱い ——— 強い
 陰気な ——— 陽気
 男性的 ——— 女性的
 軽い ——— 重い
 恐い ——— 優しい
 柔らかい ——— 硬い
 明るい ——— 暗い
 鈍い ——— 鋭い
 小さい ——— 大きい

図 4 語感から感性評価値を求める際の評価尺度

$$\text{感性評価値} = \text{重み} * \text{出現率} \quad (1)$$

ここでの重みとは、その文字の感性を測る各評価尺度での度合いで、3.2 項で述べたように音のイメージ表の記号を数値化したものである。オノマトペ中の出現率は、以下の式で定義する

$$\text{出現率} = \frac{\text{対象アルファベットの文字数}}{\text{対象オノマトペの文字数}} \quad (2)$$

3.2.3 計算例

ここで「ぎゃはは」というオノマトペを例に語感による感性評価値を求めてみる。まず、ローマ字にすると「gyahaha」である。その中の「g」の感性評価値は、重みが「重さ」:2「鋭さ」:1「男性的」:1 その他の項目は 0 であり、出現率は 7 文字中 1 回出てくるので $\frac{1}{7}$ より、「明るさ」:0「強さ」:0「大きさ」:0「重さ」:2 * $\frac{1}{7}$ 「軽さ」:0「鋭さ」:1 * $\frac{1}{7}$ 「やわらかさ」:0「元気」:0「男性的」:1 * $\frac{1}{7}$ 「女性的」:0「格調」:0「情感」:0「知的」:0 となる。同様にして他の文字「y」「h」「a」についても行い、評価項目ごとに合計すると「ぎゃはは」の感性評価値をもとめることができる。この例の結果の一部を図 5 に示す。

4. 周辺単語から感性評価値を求める

4.1 全体の流れ

オノマトペ周辺の単語から感性評価値を求める流れは以下の通りである。各 step の詳細については次節以降で述べる。

	明るさ	強さ	重さ	情感	格調	男性的
g			2*1/7			1*1/7
y				-2*1/7		
a	2*3/7	1*3/7	-1*3/7	-2*3/7	-2*3/7	
h				-1*2/7	-1*2/7	
gyahaha	0.857	0.428	-0.142	-1.428	-1.142	0.142

図 5 語感から「ぎゃはは」の感性評価値を求める

- step1:有効な単語の抽出
- step2:単語の重要度計算
- step3:感性評価値化

4.2 準備：評価尺度

使用する評価尺度は、語感から感性評価値を求める際に使用した 12 対の感性語を含む、図 6 の 53 対である。これは、sagara 氏らの日本語の SD 意味空間 [3] の 50 対と、語感から感性評価値を求める際に使用した 12 対をあわせたものである。

-2 -1 0 1 2	-2 -1 0 1 2	-2 -1 0 1 2
フォーマル ——— カジュアル	静か ——— うるさい	激しい ——— 穏やかな
知的 ——— 幼稚	古い ——— 新しい	好きな ——— 嫌いな
情感のある ——— 粗野	難しい ——— 易しい	楽しい ——— 苦しい
弱い ——— 強い	束縛された ——— 自由な	若い ——— 年をとった
陰気な ——— 陽気	長い ——— 短い	気持ちいい ——— 気持ち悪い
男性的 ——— 女性的	貧弱な ——— 立派な	悪い ——— 良い
軽い ——— 重い	熱い ——— 冷たい	単純な ——— 複雑な
恐い ——— 優しい	服従的 ——— 支配的	賑やかな ——— 寂しい
柔らかい ——— 硬い	間違った ——— 正しい	深い ——— 浅い
明るい ——— 暗い	活動的 ——— 不活発な	遅い ——— 早い
鈍い ——— 鋭い	ぼんやりした ——— ぼっさりした	白い ——— 黒い
小さい ——— 大きい	迷惑な ——— ありがたい	矛盾した ——— 一貫した
美しい ——— 醜い	純粋な ——— 不純な	四角い ——— 貫した
近い ——— 遠い	劣った ——— 優れた	正確な ——— 不正確な
完全な ——— 不完全な	頼もしい ——— 頼りない	すばい ——— のろい
不便な ——— 便利な	危険な ——— 安全な	愉快的な ——— 不愉快的な
狭い ——— 広い	消極的 ——— 積極的	悲しい ——— 嬉しい
汚い ——— 綺麗な	低い ——— 高い	

図 6 周辺単語から感性評価値を求める際の評価尺度

4.3 step1：有効な周辺単語の抽出

まず「オノマトペ&周辺文章データベース」から対象オノマトペを含む周辺文章を取得し、日本語係り受け解析機 CaboCha [2] を用いて文章中から単語を品詞情報と共に取り出す。有効な単語かどうかの判断は品詞によって行うことにした。品詞が「形容詞」・「形容動詞」・「名詞」である単語を感性評価値を求める際使用する単語とした。「形容詞」・「形容動詞」は直接感性を表現している語である。「名詞」はそれ自体は感性を表さないが「連想概念辞書」[5] を参照することで感性語を得ることができる。詳しくは step3 で述べる。ただし、名詞の中でも「固有名詞」「非自立」「数」「接頭詞」「接尾語」「人名」はオノマトペの感性にかかわることがないので除くことにする。

4.3.1 名詞から連想する形容詞を得る

単語 w が名詞の場合は、その名詞から連想される感性語を利用する。名詞から感性語を得る際は「連想概念辞書」を利用する。連想概念辞書は小学生の学習基本語彙中の名詞を刺激語とし、刺激語と「上位概念」「下位概念」「部分・材料概念」「属性概念」「類義概念」「動作概念」「環境概念」の7つの課題に関して連想語を抽出したものである。図7に例を示す。

刺激語	課題	連想語	距離	連想時間、連想順位、連想頻度より計算
春	1	季節	1.230	
春	1	四季	4.455	
春	2	小春	8.370	
春	2	新春	8.370	
春	3	四月	3.150	
春	3	三月	4.320	
春	4	あたたかい	1.466	
春	4	気持ちいい	4.860	
春	5	スプリング	2.295	
春	5	新春	8.370	
春	6	動く	10.260	
春	6	やって来る	4.320	
春	7	地球	1.620	
春	7	季節	4.725	

1「上位概念」2「下位概念」3「部分・材料概念」4「属性概念」5「類義概念」6「動作概念」7「環境概念」の7つの課題

属性概念に注目

図7 連想概念辞書

本研究では、感性に関する連想語の多い「属性概念」に着目する。属性概念を参照することで、名詞から連想される感性語（形容詞）を得ることができる。

4.4 step2: 単語の重要度計算

次に、取り出した単語のオノマトペの感性に対する重要度を計算する。この値は感性評価値を求める計算に使用する。オノマトペにとって重要度の高い単語の特徴として以下の2点が挙げられる。

- そのオノマトペと共に登場することが多い
- 他のオノマトペとは共に登場することが少ない

そこで今回は、TF-IDF(Term Frequency-Inverse Document Frequency) 法を利用する。TF-IDF 法は、「ある文章グループでの出現頻度が高い」単語のうち「他の文章グループにあまり出現しない」ものをその文章グループに特徴的な単語とする、というものである。

今回は、単語 w のオノマトペ o に対する重要度 $Weight(o, w)$ を以下のように定義する。

$$Weight(o, w) = TF(o, w) * IDF(o, w) \quad (3)$$

- TF

$$TF(o, w) = \frac{\text{オノマトペ } o \text{ 全文例中の単語 } w \text{ の数}}{\text{オノマトペ } o \text{ 全文例中の全単語数}} \quad (4)$$

この値が大きいほどその単語 w がそのオノマトペ o と関係があるといえる。例えば、「しとしと」を含む文章が10個あり、その中の全単語数が100個あったとする。そのうち「雨」という単語が3つ、「私」という単語が5つあったなら、 $TF(\text{しとしと}, \text{雨}) = 3/100$, $TF(\text{しとしと}, \text{私}) = 5/100$ となる。

- IDF

$$IDF(o, w) = \log\left(\frac{\text{全オノマトペ数}}{\text{単語 } w \text{ を含むオノマトペ数}}\right) \quad (5)$$

単語 w を文例に含むオノマトペが多ければ多いほど、単語 w がオノマトペ o にとって特有ではない、ということになることを表している。をするためである。例えば、先ほどのように「しとしと」を含む文章中に「雨」という単語と「私」という単語が含まれていたとする。また、全オノマトペ(800個)の文例に「雨」「私」という単語が含まれているか調べると「雨」は20個オノマトペの文例に、「私」は600個オノマトペの文例に含まれていた。そのとき、 $IDF(\text{しとしと}, \text{雨}) = \log(800/20)$, $IDF(\text{しとしと}, \text{私}) = \log(800/600)$ となり、「雨」の方が「私」より「しとしと」にとって特徴的である判断される。

最終的にオノマトペ「しとしと」に対する単語「雨」「私」の重要度は $weight(\text{しとしと}, \text{雨}) = (3/100) * \log(800/20) = 0.1106$, $weight(\text{しとしと}, \text{私}) = (5/100) * \log(800/600) = 0.0143$ となる。

また、値が小さくなりすぎることを防ぐため、このあとすべての単語の TF-IDF 中の最大値を1にするように正規化を行った。

4.5 step3: 感性評価値化

最後に、step2 で求めた「重要度」と、この節で定義する「単語と意味的な類似度の高い評価尺度の形容詞の値(2または-2)」とその「類似度」を使って単語の感性評価値を求める。オノマトペ o に対する単語 w の感性評価値 $Score(o, w)$ は次の式で表される。

$$Weight(o, w) * Relevance(w, adj) * Score(adj) \quad (6)$$

$$adj : \min(d | d = Relevance(w, adj), adj \in A)$$

A : 評価尺度の形容詞の集合

式を構成するそれぞれの値について説明する。 $Weight(o, w)$ は step2 で求めた単語 w のオノマトペ o に対する重要度である。 $Relevance(w, adj)$ は、単語 w とすべての評価尺度の形容詞を比較したとき、最も類似度の高い組 (w, adj) の類似度である。評価尺度の形容詞の値 $Score(adj)$ とは2.2項で述べた adj の端の値のことで、2または-2である。

4.5.1 周辺単語と評価尺度との類似度

単語 w と評価尺度の形容詞 adj との類似度 $Relevance(w, adj)$ は0~1の値をとる。類似度 $Relevance(w, adj)$ を求める際には、WordNet::Similarity[7]のgetRelatedness関数を利用する。英語類義語辞典 WordNet は、単語を上位概念・下位概念などの情報を元に階層化しており、WordNet::SimilarityのgetRelatedness関数はこの構造を利用して入力した2単語の類似度を測っている。getRelatednessの二つの引数に、単語 w と、評価尺度の形容詞 adj を指定し、すべての adj について

類似度を求め、最も類似度が高い組を採用する。その際、引数となる二つの単語は和英辞書 edict [4] を使って英訳したものを使用する。

5. 評価実験

5.1 アンケート

実際にそれぞれのオノマトペについて人がどのような印象を持つかを調査するため、日本人9人に対してアンケートを行った。アンケートでは、各オノマトペの印象を今回使用したすべての評価尺度53対に関して-2, -1, 0, 1, 2の5段階で評価してもらった。使用したオノマトペは、それぞれ異なる印象となると予想した「しとしと」「ざあざあ」「しーん」「がやがや」「ぎゃはは」「にっこり」「つつつつ」「がさがさ」「きびきび」「のっそり」の10語を選んだ。

5.2 プログラムの算出結果とアンケートとの比較

前項で述べたアンケート結果を評価尺度ごとに平均した値とプログラムによって算出された感性評価値を比較した。表1に差分合計と、アンケート・プログラムそれぞれにおいて特徴的であった評価尺度を示す。差分合計とは、評価尺度ごとにアンケートの平均値とプログラムによって算出された感性評価値の差分をとり、合計したものである。表2, 3には、差分合計最大・最小であった「つつつつ」「のっそり」のアンケート・プログラムそれぞれの値の絶対値の分布を示す。また、「つつつつ」

表1 アンケートとプログラムとの比較結果

オノマトペ	アンケートと プログラムとの 差分合計	アンケートで 絶対値最大 である評価尺度	プログラムで 絶対値最大 である評価尺度
つつつつ	15.22836	まるい	狭い
がさがさ	18.41166	醜い	陽気な
きびきび	22.70374	すばやい	軽い
しーん	23.16206	静かな	狭い
ざあざあ	23.33478	激しい	カジュアル
がやがや	23.38294	うるさい	激しい
にっこり	23.71869	明るい	高い
ぎゃはは	24.11194	うるさい	狭い
しとしと	26.00565	穏やかな	陽気な
のっそり	29.90302	のろい	激しい

表2 値の分布：つつつつ

	0.1 以上	0 以上 0.1 未満	0
アンケート	44	0	9
プログラム	16	13	24

表3 値の分布：のっそり

	0.1 以上	0 以上 0.1 未満	0
アンケート	40	7	6
プログラム	18	36	19

「のっそり」の全評価尺度の比較結果を図8と図9に示す。横軸が全53対の評価尺度に対応しており、プログラムの感性評価値に関して絶対値の大きい順に評価尺度を上から並べてある。図8, 図9ともに、点線がアンケートの結果を各評価尺度ごとに平均した値であり、実線がプログラムによって算出した感性評価値である。さらに、図8図9それぞれの上位20件を拡大した図を図10と図11に示す。

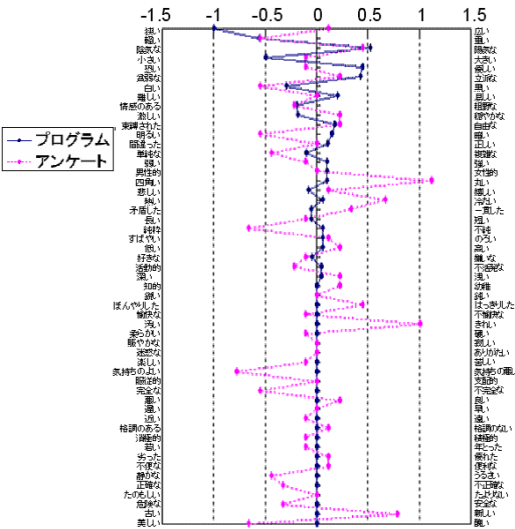


図8 アンケートとプログラムの比較 (差分最小)：つつつつ

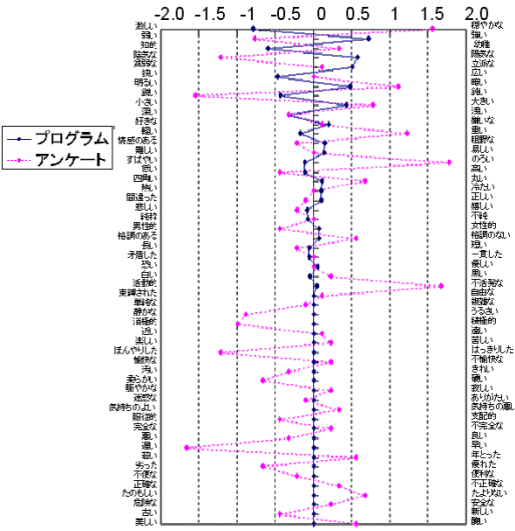


図9 アンケートとプログラムの比較 (差分最大)：のっそり

5.3 考察

表1から、アンケートとプログラムの値に差分があることがわかった。差分が生じる原因の一つとして図8, 図9からわかるように、プログラムの値の絶対値がアンケートに比べ全体的に小さいことが挙げられる。表2, 3を見ると、アンケートでは53尺度の値のうち約8割の値の絶対値が0.1以上なのに対し、プログラムの値は半数以上が0.1未満におさまっている。0.1未満の小さな値や0が多い原因として、周辺単語と評価尺度との

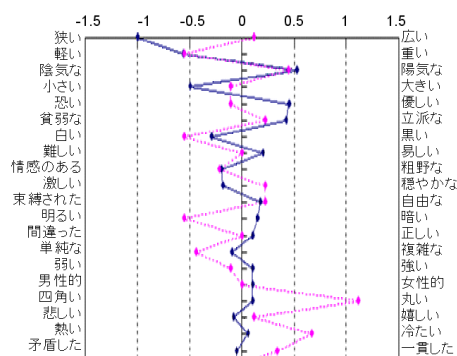


図 10 プログラムとアンケートの比較「つつる」上位 20 件

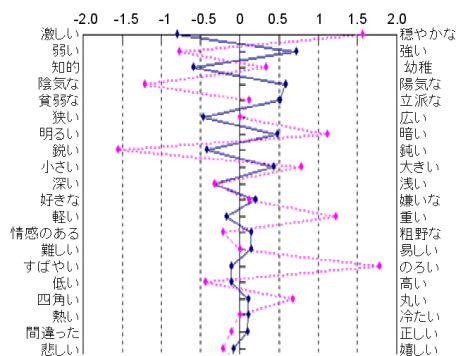


図 11 プログラムとアンケートの比較「のっそり」上位 20 件

類似度を求める際に形容詞を名詞化するためあまり一般的ではない単語になってしまい、WordNet::Similarity による類似度が小さくなるまたは 0 になることがある。一般的ではない単語になるとは例えば、つらい→つらさを *toilsome* → *toilsomeness* とするなどである。これは、今後周辺単語から得られる感性評価値の絶対値を大きくするよう重みを調節することでよりアンケートの値に近づけることができると考えられる。

また表 1 の特徴的な単語から、人がそのオノマトベから最も感じる印象とプログラムの結果による印象は異なることが多いとわかった。例えば、差分最小の例「つつる」の場合については、アンケートでの最も特徴的な尺度が「丸い」に対し、プログラムでは「狭い」であった。アンケートにおける「狭い (-2) 広い (2)」の値は 0.1111 と 0 に近い値であった。このように「つつる」の印象を表すのに有効な指標ではないにもかかわらず、その指標の値が最も特徴的な値として算出されているという結果が多く現われた。さらに、オノマトベ「のっそり」の場合、アンケートでの最も特徴的な印象が「のろい」に対しプログラムでは「激しい」であるが、アンケートでは反対語「穏やか」の印象が強く現われていた。しかし、実際の印象と合致している例もあり、「がやがや」のプログラムで特徴的とされた「激しい」はアンケートにおいても 1.1111 と比較的特徴的な値となっている。

図 10 を見ると、アンケートとプログラムで左右の判断が一致している評価尺度は 20 尺度中 11 であった。例えば、「陰気陽気」に関してアンケートが 0.4444、プログラムが 0.5202 などである。左右の判断が食い違っている評価尺度は 6 尺度あった。残りの 3 尺度は、アンケートが 0 に対しプログラムで 0 以

外の値になったものである。これより、差分最小の例「つつる」ではプログラムが特徴的と判断した上位 20 件中半数ほどが実際の印象と合致していることがわかる。

次に図 11 を見ると、アンケートとプログラムで左右の判断が一致している評価尺度は 20 尺度中 9 尺度であった。左右の判断が食い違っているのは 8 尺度あった。残りの 3 尺度は、先ほどと同様にアンケートが 0 に対しプログラムで 0 以外の値になったものである。これより、差分最大の例「のっそり」ではプログラムが特徴的と判断した上位 20 件中半数以上が実際の印象と不一致であることがわかった。

ここで、語感情の有効性について考える。全評価尺度 53 尺度のうち 12 尺度のみ語感情情報を含んだ感性評価値であり、残りは周辺文書の情報のみによる値であったが、アンケートとプログラムで左右の判断が一致している評価尺度のうち、語感情情報を含んだ感性評価値の割合は「つつる」では 11 尺度中 5 尺度であり、「のっそり」では 9 尺度中 3 尺度であった。このことから、語感情情報は有効な場合と有効でない場合があることがわかった。

最後に、プログラムから得られる結果がアンケートの結果と逆の結果になることや有効ではない指標の値が特徴的な値として算出されていることに関して考えられる原因をまとめると

- 得られる周辺単語がオノマトベの感性に関係しないことがある。
- オノマトベの感性に関する単語が得られたとしても、WordNet::Similarity を使って周辺単語と評価尺度との類似度計算をするために、単語を英訳・名詞化する過程で意味やニュアンスが変わってくる。
- 英訳・名詞化・類似度計算する過程で利用できる単語が減っていき情報が減る。

6. まとめと今後の課題

オノマトベを習得する際重要となるオノマトベの持つ感性に着目し、語感と周辺文書を分析することによりオノマトベの感性を自動算出し、ユーザに提示するシステムを提案した。今後は、アンケートの結果に反する結果を出した部分について原因を分析し精度を向上させ、インタフェース実装、カタカナへの対応などを行っていきたい。

文 献

- [1] Chisato Asaga, Yusuf Mukarramah and Chiemi Watanabe.: ONOMATOPEDIA: Onomatopoeia Online Example Dictionary System Extracted from Data on the Web, Proceedings of Asia-Pacific Web Conference (APWeb'08), pp.601-612 (2008)
- [2] Cabocha: <http://chasen.org/taku/software/cabocha/>
- [3] 岩下 豊彦: SD 法によるイメージの測定—その理解と実施の手引, 川島書店, p.204 (1983)
- [4] EDICT, <http://www.csse.monash.edu.au/~jwb/j-edict.html>
- [5] 岩永嘉弘: ネーミングの成功法則, PHP 研究所, p.241 (2002)
- [6] 岡本 潤, 石崎 俊: 連想概念辞書の距離情報を用いた重要文の抽出, 自然言語処理, Vol.10, No.5, pp139 - 151 (2003)
- [7] T. Pedersen, S. Patwardhan, and J. Michelizzi. Wordnet::similarity - measuring the relatedness of concepts. Pro-

ceedings of the Nineteenth National Conference on Artificial Intelligence (AAAI-04), pp.1024-1025(2004)

- [8] 高橋もも: イメージにあった楽曲推薦手法の提案, お茶の水女子大学人間文化研究科 修士論文 (2007)