

周辺文書と語感を利用した日本語オノマトペの感性指標の提示

中部 文子 (指導教員: 渡辺知恵美)

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習得が難しい。そこで我々は、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」[1]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、「雨がぼつぼつ降っている」「雨がざあざあ降っている」は、日本人ならばまったく違う情景が想像できるが、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。

2 オノマトペの感性情報提示システム

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、感性評価値を計算・集計し、提示する。感性評価値は、53 組の評価尺度に対して計算する。53 組の評価尺度には、Sagara et al.(1961)のSD 尺度 [3] と、音のイメージ表 [5] で使われていた尺度をあわせたものを使用する。

集計結果は、図 1 のようにユーザに提示する。オノマトペディアの「オノマトペ 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。

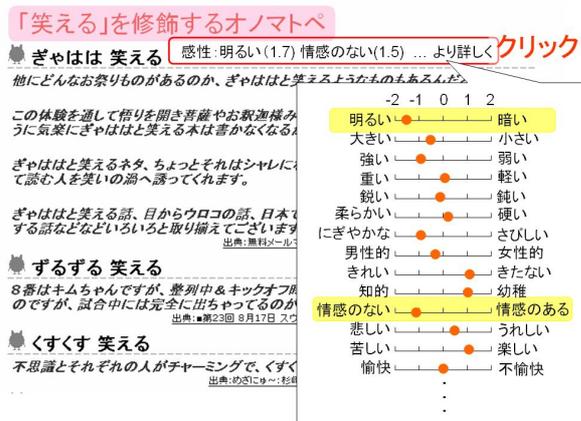


図 1: ユーザへの提示方法

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。

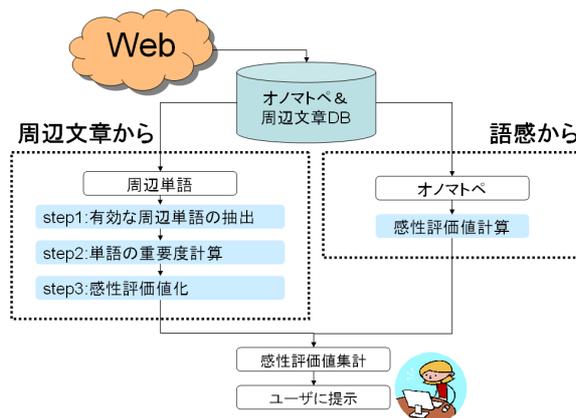


図 2: システム概要

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利用した。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。をととてもそうだ、をどちらかというそうだ、をどちらでもない、をどちらかというそうでない、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 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。

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

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、3.1 項で述べた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

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

