

# 外国人のための日本語オノマトペ学習支援手法の一検討

## An Idea of Support to Learning Onomatopoeic Words in Japanese as a Foreign Language

鈴木 雅実\*<sup>1</sup>  
Masami SUZUKI

木村 寛明\*<sup>1</sup>  
Hiroaki KIMURA

\*<sup>1</sup> KDDI 研究所  
KDDI R&D Laboratories, Inc.

In this article, we discuss about an idea on support to learning Japanese onomatopoeic words that is regarded as difficult for foreign language natives. Those onomatopoeic words are frequently used in various contexts of casual conversations, while their nuances are often beyond imagination. Thus, it would be difficult for second-language learners to acquire those onomatopoeic expressions with relatively small amount of examples. For the purpose of preparing useful information for learners, we will extract specific features from the text surrounding each onomatopoeic expression. We will also discuss about related research issues to be resolved for efficient learning towards supporting cross-cultural communications.

### 1. はじめに

日本語で日常的にしばしば用いられる擬音語・擬態語などのいわゆるオノマトペ(擬声語)は, 外国人学習者にとって習得が困難とされる. 本研究では, オノマトペが実際に使用される文脈テキストを収集し, 周辺に分布する語の特徴を学習者に提示することにより, 様々な状況においてオノマトペを理解したり自己の発話で用いる上での効果的な支援方法の基礎検討を行う.

### 2. 研究の背景とねらい

#### 2.1 背景

筆者らはこれまで, 第二言語(外国語)の学習対象表現のうち, コンサイス表現という用語を用いて, 次のような論を展開している. すなわち, 各言語に共通して見られるコンサイス表現(簡潔ながらイメージ喚起力が強く人々の間で共感を得られ易い表現が様々な形式で存在)の効果的な学習方法の提案である. 母国語話者にとっては成長過程で自然に習得できる表現であっても, 第二言語(外国語)学習者が少ない例示によりそのニュアンスを体得することは難しく, どのような状況で用いるかを学ぶには多くの時間とリソースを必要とする. そこで, 大規模テキストコーパス中に現れる用例文脈から, 該当表現の近傍に出現する感情等の感性特徴の共起分布を相対的に抽出することにより, それを図示するような方法で学習補完情報として提示することが有益であると考えられる. [鈴木 2015]では, この手法提案と同仮説の検証に向けた考察を述べた.

#### 2.2 研究目標

本稿では, 外国語話者にとって習得が難しいとされる日本語の擬声語(オノマトペ=擬音語・擬態語の総称)を具体的な学習対象として取り上げ, 上記の基本構想に基づいて学習支援情報を提示することを目的として, その実現方法の検討を行う. 着眼点の第 1 点は, オノマトペが使用される状況・文脈の特徴をどのように捉えるかという問題である. 第 2 点は, コーパスを用いて特徴抽出を行う場合の留意点と結果の提示方法である.

### 3. 関連研究

#### 3.1 言語教育的な観点から

外国人にとってのオノマトペの学習に関する問題は, 日常的な言語生活における使用頻度や重要性に比して, それほど注目されて来なかったようである. その点を指摘して, 多角的な視点から, オノマトペの教育に関して体系的な調査検討を行った先行事例として, 三上の研究がある[三上 2007]. 同論文では, 使用頻度その他の根拠に基き, 外国人話者への日本語教育のための「基本オノマトペ」70 語を選定しており, 参照基準としても適切と考えられることから, その範囲での検討を試みることにする. このうち, 同音の繰り返しによるものは 32 語であり, 次のような代表的なオノマトペが含まれている(五十音順).

いらいら／うろうろ／がたがた／がやがや／きらきら／…  
ぐるぐる／げらげら／ごろごろ／ざざざ／さらさら／…  
のるのる／ばたばた／はらはら／ばらばら／…  
ふらふら／ぶらぶら／ふわふわ／…／わくわく

#### 3.2 計算論およびユーザ支援の観点から

中部らの研究では, オノマトペの持つ音韻的な特徴を含む語感に着目するとともに, オノマトペが修飾する対象となる動詞などの用例を Web から収集して, その出現頻度から予め用意した感性評価指標の評価値を計算している[中部 2009]. ただし, 出現文脈やニュアンスを正しく捉える上では, 分野依存性や多重性への対処等の課題が残されている.

また千本と竹内は, オノマトペが用いられる分野や文脈によって意味が異なる場合があることから, レビュー文書を対象に, オノマトペを含む評価語の語関連ネットワークを構築して分野間の意味的な相違を分析している[千本 2014]. さらに清水らは, 被験者による実験に基づく数量化により, オノマトペごとの微細な印象を推定するシステムを提案した[清水 2014]. 同システムでは, 未知のオノマトペに対する印象評価予測も可能としている点が興味深い. これらに対し, 本研究では外国人話者が学習する場合に参照可能な集合知の提供を指向している.

この他, 留学生を対象としたオノマトペの学習を支援するための電子絵本システムが提案されており, 比較的意味の近いオノマトペの利用状況の違いを視覚的に示すような, 対話型のユーザ・インタフェースの効果を検証した[前田 2014]. ただし, その元となる提示内容は事前に用意されたものとなっている.

連絡先: 鈴木雅実・木村寛明, KDDI 研究所メディア開発 G,  
〒102-8460 東京都千代田区飯田橋 3-10-10 ガーデン  
エアタワー, Email: msuzuki@kddilabs.jp

## 4. 学習支援情報の提供について

### 4.1 文脈関連情報の提示

筆者らの基本提案では、大規模なコーパスを用いて、該当表現を含む周辺の文脈の中から特徴的な感性語を抽出し、その分布傾向を参考情報として提示することを想定している。周辺文脈の範囲としては、内容的な一貫性(coherence)を持つパラグラフを単位とすることで、共起する感性語の出現についての信頼性が保たれるものと思われる。しかし、感性語の種類によっては近傍での共起のみが重要な意味を持つ場合も多い。例えば Positive/Negative の極性や音韻的な特徴に関しては、該当表現が現れる文と合わせて前後 3 文内の情報に限定してよいであろう。他の特徴に関しては、同様に局所的に捉えた方がよい場合と、パラグラフあるいは文書レベルの範囲内で大局的な文脈/状況を把握する方が適切な場面の両者が存在し得ると思われる。

### 4.2 オノマトペの場合～試行例に基づく考察～

オノマトペの場合はどんな様相であろうか？オノマトペは、特定の語(動詞や副詞、形容詞、名詞など)とのコロケーションの形で慣用的に用いられる場合も多い。ただし、少なからず多義性が生じることがあり、そこには比喩的理解の認知プロセスが介在していることが示唆されている[三上 2007]。ここでは、その中の考察事例から「ごろごろ」を取り上げて、その周辺文脈との関係を比較検討する。同論文に引用されている、荻坂による被験者実験(日本語を母国語とする大学生 290 名)では、「ごろごろ」から直ちに連想する一語を自由記述させた場合に、「雷」が 4 割近くとなり第 1 位であった。これに対し、「寝ている」状態を指す言葉への連想が 2 割強を占めた。両者の間には五感との対応の相違などが見られるが、メタファー的な認知プロセスが存在するとされている。これに対して、ブログコーパスを対象とした語の共起分布を観測した場合の一例を表 1 に示す。なお、対象ドメイン内の「ごろごろ」の単純出現文書数は 5060 件であった。この結果では、「ごろごろ」は音韻的な結び付きが強い「雷」よりも、抽象化された「寝る」容態と合わせて言及される頻度が相対的に高いことが解る。

表 1 オノマトペ「ごろごろ」と共起する語の分布状況

	共起頻度/共起語単独	共起度
ごろごろ+雷	1140/69300	5.125
ごろごろ+寝る	1840/46300	6.398

注)共起度は相互情報量基準に基づく計算値

一方、コーパスを限定して「ごろごろ」の出現文脈を観察するため、国立国語研究所[山口 2015]から研究利用に提供されている、全文検索エンジン「ひまわり」を用いて、同ツールのサンプル・コーパス(青空文庫)を指定した場合のコンコーダンス分析を行ったところ、「ごろごろ」が出現する文脈 9 件のうち 1 件のみ音のイメージ(荷車)と関係する他は、怠惰な状態を表すのに使用されていた。

以上のように被験者が直観的に特定のオノマトペから連想するイメージの分布には個人差があるとともに、実際に使用される文脈は必ずしも直観と対応しないことに留意する必要がある。さらに、レジスター(言語使用域)による相違点等も予想されるため、文脈を特徴づける語は対応する(話し言葉)コーパス毎に取得すべきであろう。その意味で、日常生活の様々な局面での会話を収集したコーパスの利用が期待される[小磯 2015]。

## 5. 課題と今後に向けて

筆者らの提案手法を用いて、共通した構造を持つ英語表現”Let it be”と”Let it go”各々の感性語(基本 6 感情)との共起度についてブログドメインを対象に調査した結果、それらのニュアンスの違いが把握できる可能性が示唆された[Suzuki 2014]。しかし、この例の両者に見られる宗教性の濃淡や、皮肉などの評価性に関連する高次の文脈・状況的な特徴の抽出が課題であることも明らかとなった[鈴木 2015]。

今回対象としたオノマトペの場合、直観的な五感との結びつきは、これまでに提案されているような主観評価に基づく手法でその特徴を推定することが可能と思われる。しかし、学習目的に適した提示を行うことを前提として、使用文脈のニュアンスを特徴づけるような感性語(あるいは評価語)をコーパスを用いて抽出する方法は未知の領域と言ってよい。しかも、本稿で指摘したように、オノマトペに対する主観評価と語の運用実態の間には乖離があることを考慮しつつ、外国人話者の学習支援に適した情報の取捨選択の観点から工夫が必要である。

さらに、前述したように、レジスター(言語使用域)によりオノマトペの使用状況が異なると予想される点については、本提案を実際に適用する際に、学習者の社会属性や行動範囲に対応したコーパスを用いて使用実態の分布を直観的に示すことが望ましいと考えている。特に、該当オノマトペに直接対応する表現が存在しない外国語の話者にとって、オノマトペの理解と発話を含む運用上の手引きとなるような情報を効果的に提供する方法については、視覚的な提示方法など言語に依存しないユーザ・インタフェース的な手段の活用を追求したい。

なお、国際会議 ICT4LL 2014 で発表した際に本研究への手がかりとなるコメントを頂いた、ヘルシンキ大学で日本語講座を担当する Dr. Riikka Lansialm に、ここで謝意を記すものである。

### 参考文献

- [鈴木 2015] 鈴木雅実・木村寛明: 第二言語におけるコンサイス表現の学習支援について, 信学技法, TL2014-52, 2015.
- [三上 2007] 三上京子: 日本語オノマトペとその教育, 早稲田大学博士論文, 2007.
- [中部 2009] 中部文子・浅賀千里・渡辺知恵美: 感性情報を利用したオノマトペ学習システムの開発, 第 1 回データ工学と情報マネジメントに関するフォーラム (DEIM2009), E5-1, 2009.
- [千本 2014] 千本達也・竹内和広: 分野データにより訓練した評価語関連ネットワークに対するオノマトペ表現の割り当て, 信学技報, TL2014-44, 2014.
- [清水 2014] 清水祐一郎・土斐崎龍一・坂本真樹: オノマトペごとの微細な印象を推定するシステム, 人工知能学会論文誌, Vol.29 No.1, pp.41-52, 2014.
- [前田 2014] 前田安里紗・上間大生・松下光範: 留学生を対象としたオノマトペの学習を支援するための電子絵本システムの評価, 情報処理学会第 76 回全国大会, pp.4-809, 2014.
- [山口 2015] 山口昌也: 全文検索システム「ひまわり」を用いた既存資料の活用, 第 7 回コーパス日本語学ワークショップ
- [小磯 2015] 小磯花絵, 他: 均衡会話コーパス設計のための一日の会話行動に関する調査 ―中間報告―, 第 7 回コーパス日本語学ワークショップ, 2015.
- [Suzuki 2014] M. Suzuki and H. Kimura: Support for Learning Concise Expressions with Affective Features Extracted from Context of Usage, 7th Edition of International Conference ICT for Language Learning (ICT4LL), Florence, Italy 2014.