

カスタマーレビューにおけるオノマトペとその活用についての調査

An Investigation of Onomatopoeic Expression in Online Costumer Reviews

岡田 真^{*1}
Makoto Okada

橋本 喜代太^{*2}
Kiyota Hashimoto

^{*1} 大阪府立大学理学系研究科 Graduate School of Science, Osaka Prefecture University
^{*2} 大阪府立大学現代システム科学域 College of Sustainable System Sciences, Osaka Prefecture University

Today, many colloquial text data such as microblogs and costumer reviews are available on the Internet, and there have been a number of studies to extract useful information from them. Particularly, costumer reviews are important research targets because they contain authors' evaluations and/or impressions on items, though their extractions are not trivial. Onomatopoeic expressions are widely used in colloquial Japanese, and, despite the difficulty of describing their meaning, most Japanese people share their meanings at a surprising level, and thus they are to be more utilized in customer review analysis. In this study, we chose customer reviews on traveling as example data and extracted onomatopoeic expressions found in them. We analyzed co-occurrence patterns and discussed the possibility of more detailed analysis on customers' evaluation.

1. はじめに

オノマトペ(擬音語・擬態語などを総称して本稿ではそう呼ぶ)は現実世界の音を模倣したり, 行為や状態の様態を音表象したりする表現である[橋本 2011]. 近年の電子化文書の増加に伴い, コンピュータで文書データを扱う機会が増加していることは周知の事実であるが, その対象として, ウェブログや Twitter などのマイクロブログや Facebook などの SNS サイトなどに記載される文書群が増加している. これらの文書群は従来の新聞記事などと異なり, ユーザが自由に記述する場合が多く, 口語的な表現が多用されている. その中でオノマトペも多く見受けられる.

一方, カスタマーレビューを用いた研究も近年盛んになされている. カスタマーレビューは商品や宿泊施設などに対する消費者自身の評価や感想を記入した文章であり, Amazon .co.jp や価格.com や旅行情報ポータルサイトであるトリップアドバイザーといったウェブ上の商業サイトや商品評価サイトや旅行情報サイトで参照することができる. カスタマーレビューは消費者側と経営者側の双方にとって有益な情報源である. 新規の消費者にとっては, 商品などを購入する際に利用することで有益な情報を取得することができ, 経営者側にとっては消費者の評価や嗜好を調査し, 経営に有効に反映できるという意味で利用価値が高いと考えられている.

これらのカスタマーレビューも, レビュー筆者の感情や感覚を表す表現として口語的な表現を多く含む傾向がある. そのため, カスタマーレビュー中の筆者の感情や感覚を自動的に推定し, 他の利用者の要求に応じてレビューを有効に提示するためにそれらの表現を利用することは有用だと考えられており, それらの効率的な抽出や利用を狙い, さまざまな研究がなされてきている[飯田 2005], [杉木 2009]. オノマトペについても利用に対する要求は強くなっているが, カスタマーレビュー中のオノマトペの利用状況やレビューの内容との関連などについての詳しい調査はまだ十分になされていないのが現状である.

そこで本稿では, カスタマーレビュー中でオノマトペがどのように用いられているかを調査し, その特徴を考察した. その上でレビューの自動要約手法において利用する設計を提示する.

本稿ではまず 2 章で旅行関係ポータルサイトであるトリップアドバイザーとそこで参照できるカスタマーレビューの内容について説明する. 3 章ではカスタマーレビューにおけるオノマトペの特徴について述べ, 4 章では応用として, オノマトペを利用したカスタマーレビューの自動要約手法について提案し, 今後の課題についても述べる.

2. トリップアドバイザー

トリップアドバイザーは世界中の旅行関連情報を集めたポータルサイトである. 日本を始め, 世界中のホテルとレストランの情報が収集されており, それらの施設の利用者が自らの感想などをレビュー文書として付け加えることができる.

このサイトには 5 千万以上のレビューが 20 カ国語以上のさまざまな言語で登録されている. 日本に観光に来た外国の旅行者が利用したホテルやレストランに自分の母語でレビューを記入することができる. そのために, 登録されている施設には複数



図 1. トリップアドバイザーのレビューの例

の言語でレビューが記載されている場合が多々見受けられる。

レビュー文書には「タグ」と呼ばれるキーワードを付加することができ、それにより利用者がどのような状況や目的で施設を使用したかを表すことができる。トリップアドバイザーから提供されているタグとしては、「仕事」、「カップル」、「家族」、「友人」、「一人」という5種がある。

さらに、「価格」、「立地条件」、「清潔さ」、「サービス」などといったさまざまなテーマについて5段階で評価を加えることが可能である。

図1にトリップアドバイザーのレビューの例を示す。

この図にはレビューの文書、タグ、テーマが示されている。

3. カスタマーレビューにおけるオノマトペの特徴

カスタマーレビューにおけるオノマトペの特徴について調査を行った。調査のために用いたデータとしては、トリップアドバイザーから抽出した宿泊施設に関するカスタマーレビューを用い、それらを人手により調べた。

日本語によるレビュー中でオノマトペを含むと判断された表現の例を図2に示す。図2中の下線部がオノマトペと判断された表現である。図3には今回調査したレビュー中で出現頻度の上位10個のオノマトペ表現を示す。

今回調査したレビューにおいて、オノマトペ表現は宿泊施設全体に対する印象に対して記述されている場合と、各部分について表現する場合に用いられている場合が見受けられた。筆者の意図の推定に用いる場合には、好意的か否定的か、その強弱の度合いはどうかといった点に加え、対象の範囲を確定しなければならない。

また、同じオノマトペ表現がいろいろな対象や状況で使用されている場合もよく観察された。図2の例では「しっかり」という表現が適合すると考えられる。この場合は、共起表現などを利用して対象が何かや、どのような状況下で用いられているかを推定しなければ、正確な評価ができない。そのための推定ルールやルール構築のための手法について考える必要がある。

図3の頻度上位のオノマトペ表現は、大別して「しっかり」のように事物や行為の描写に用いられる場合と、「ゆっくり」のように記述者の気持ちの記述に用いられる場合がある。前者の場合は前述したような共起表現を利用した対象やその範囲の推定を行う必要が見受けられるため、考慮しなければならない点が多い。後者の気持ちの表現に用いられる場合はそのみで筆者の気持ちを推し量れる場合が多く、レビューの評価に直接反映させることが可能であると考えられる。例えば「ゆったり」であれば「ベッド」や「部屋」などの設備が「ゆったりしていた」と用いられることがよくあるが、その際にはそれらの設備に空間的な余裕があったため、筆者が「ゆったり」とくつろげたので「良かった」と評価している場合が多く、オノマトペ表現が直接関連している単語などを無視して、レビュー全体の評価とみなしても問題がない。

4. オノマトペを利用したレビューからの情報の自動抽出

本章ではオノマトペの特徴と、カスタマーレビューに付与されたタグ情報などを利用したレビューからの情報の自動抽出手法について提案する。

まずカスタマーレビューからあらかじめオノマトペ表現を抽出しておく。その際にレビューに付与された評価やタグ情報などをともに用いて、オノマトペ表現が好評・不評といった評価のどれをどの程度の強さで表しているのか、またどのような分野に関連した内容かを推測可能なように情報を準備し、これらのオノマト

- パリッとした真っ白なテーブルクロス
- テーブルがベスト!
- 恋人と泊まりましたが、一度は泊まって見たかったのでわくわくしてチェックインしました。
- 対応しておらずちよつとがっかり
- 旅先で洗濯が出来るのは、気持ち的にもさっぱりするし、
- 部屋も大変広く清掃等もしっかりしていたので清潔感がありました。
- セキュリティもしっかり
- 値段の割には朝食バイキングではしっかり食べられますし、
- 上の階の歩く足音が相当大きく聞こえました。どんどん、とどうしてあんなに響くのかと思う程でした。
- ほっと心安らぐ場所

図2. カスタマーレビュー中のオノマトペ表現の例

ゆったり、ゆっくり、のんびり、しっかり、ほっ(ホッ)、
がっかり、さっぱり、わくわく、びっくり、ぐっすり

図3. 出現頻度上位のオノマトペ表現

ペ表現を中心としたデータ群をオノマトペ表現辞書として保存しておく。

次に、レビュー文書全体からユーザの求める情報を抽出する。この際にどのような観点で抽出するかが重要となる。レビューの場合、対象となる商品や施設の全体と各部の評価とその度合い、およびその理由などが抽出されるべき情報となる。この時に名詞など以外に、形容詞などのレビュー筆者の感情が出やすい表現が手掛かりとしてよく用いられる。これらにオノマトペ表現を加えることで、レビューの筆者が評価したいと考えた対象を絞り込み、より詳細な評価を抽出することが可能となると予想される。

最後に抽出した情報を整理し、ユーザに直接提示するか、他の手法へと応用する。具体的な応用手法としては情報検索や自動要約などが挙げられる。

課題としては、オノマトペのさらなる調査とその応用としてのシステムの構築があげられる。旅行情報におけるオノマトペの利用とそれとユーザが付与した情報との関連の調査と、その結果を生かした筆者の感情解析手法などへの応用である。さらに、日本語以外の言語についてもさらなる調査をおこない、日本語のオノマトペおよびカスタマーレビューで用いた手法の多言語における有効性についてもさらなる調査を行いたいと考えている。

参考文献

- [橋本 2011] 橋本喜代太, 竹内和弘: ジャンル、文脈に応じたオノマトペ表現の感性極性値の推定手法の提案, 人工知能学会第25回年次大会, 1C2-OS4b-6, 2011.
- [飯田 2005] 飯田龍, 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一: 意見抽出を目的とした機械学習による属性-評価値対同定, 情報処理学会研究報告 NL165-4, pp.21-28, 2005.
- [杉木 2009] 杉木 健二, 松原 茂樹: カスタマーレビューに基づく商品検索のための感性表現ソーラスの構築, 言語処理学会第15回年次大会, P3-19, pp.781-784, 2009.