

カスタマーレビューにおけるレビュー評価点とオノマトペの関係の分析

佐藤 宏亮^{*1} 内田 ゆず^{*2} 長谷川 大^{*2} 米山 淳^{*2}
 Kosuke Sato^{*1} Yuzu Uchida^{*2} Dai Hasegawa^{*2} Jun Yoneyama^{*2}

^{*1} 青山学院大学大学院理工学研究科

^{*1} Graduate School of Science and Technology, Aoyama Gakuin University

^{*2} 青山学院大学

^{*2} Aoyama Gakuin University

Nowadays, the internet has been developed. Therefore, people use it more and more. For example, the persons using Facebook or Twitter have been increasing. Moreover, we shop many things on the internet so that there are a lot of customer reviews there. The feature of Facebook, Twitter and customer reviews is that people write a review in a colloquial style. Consequently, people tend to use onomatopoeias when they write reviews. Customer reviews are very useful information. However, there are too many customer reviews on the internet. That is why sometimes it is very difficult to find beneficial information. Our purpose is to get the customer reviews customers or companies want to see from a lot of them by using onomatopoeias.

1. はじめに

日本語の語彙には、動作や状態、事物の姿形を感覚的に表す擬態語、事物の音や声を表す擬音語が豊富に存在している。この擬態語・擬音語をまとめてオノマトペと呼ぶ。

近年、インターネットの発展によりWebを介した情報流通が盛んに行われている。その対象として、FacebookやTwitterなどのSNSサイトが挙げられる。これらのサイトでは口語的な表現が多く用いられるのが、一つの特徴である。

同様に、インターネット上の商取引も増えている。インターネットを介した購買行動では、外出する手間が省けるだけでなく、自分の気になった商品に対する消費者のレビューが見られるのが特徴である。このカスタマーレビューもFacebookやTwitterと同様に、口語的な表現が多く見受けられる。

カスタマーレビューは消費者と企業両者にとって有益な情報である。消費者は自分の気になった商品情報を、その商品を購入した消費者から得ることが可能である。企業にとっては、消費者からの率直な意見を得ることが出来る。

ウェブ上には、膨大な量のカスタマーレビューが存在する。そのため、自分の得たい情報を探す事は困難である。そこで、我々は、レビュー内に出現するオノマトペと、その商品に対する評価の関係が評価文分析等の分野で有用であると考えている。そこで本稿では、カスタマーレビューからオノマトペを抽出し、レビュー評価点とオノマトペの関係を考察する。

2. オノマトペの収集

2.1 商品レビューデータ

Yahoo! ショッピングから商品レビューを収集した。商品番号等の不必要な情報は削除し、<ReviewTitle>要素、<Description>要素、<Rate>要素を用いる。これらの要素はそれぞれ、レビューのタイトル文、レビュー本文、商品に対する満足度を意味する。Rateは、1から5で評価され、数が大きい程、満足度も高くなる。

この調査において我々は、70万件以上の商品レビューを用いる。商品は9,170のカテゴリに分類されている。例えば、「本・雑誌・コミック」、「レジャー・アウトドア」等である。

また、<ReviewTitle>要素と<Description>要素に含まれる単語数は35,720,231語である。

2.2 オノマトペの抽出

合計140万件以上の要素から、オノマトペを抽出した。抽出対象としたオノマトペは175個である。抽出対象としたオノマトペは日本語オノマトペ辞典[2]を用い、決定した。表1に抽出対象とした一部のオノマトペと、その動作カテゴリーを示す。

次にオノマトペを自動抽出する際に、使用したアルゴリズムについて説明する。まず始めに、

Yahoo! ショッピングの商品レビューからオノマトペの関係性を調べるために必要とした <ReviewTitle>, <Description>, <Rate> を抽出する。次に, <ReviewTitle>, <Description> に形態素解析を施し, 上述した 175 個のオノマトペを含むレビューを抽出する。形態素解析器には, 日本語形態素解析システム JUMAN7.0[3] を用いる。

2.3 抽出したオノマトペを含むレビュー

抽出したオノマトペを含むレビューの件数は, 796 件であった。また, 今回抽出対象とした 175 個のオノマトペのうち, 47 個が抽出された。このうち最も出現頻度が多かったのは, 「しっかり」であった。表 2 に抽出された 47 個のオノマトペのうち使用頻度が高かった上位 5 語を示す。ここで総数はそのオノマトペを含むレビューの件数を意味する。

3. オノマトペとカスタマーレビュー

3.1 Rate とオノマトペ

出現頻度が高い 5 語のオノマトペと Rate の関係を図 1 に示す。この図は各 Rate に分布するオノマトペの割合を示したものである。図 1 において, Rate の値が 1 の「がっかり」を例にとると, 抽出した Rate の値が 1 のオノマトペのうち 40% 以上が「がっかり」であることがわかる。

表 1 オノマトペ収集対象

オノマトペの分類	オノマトペの例
食べる・噛む・舐める	ばくばく, ごくごく
笑う	にやにや, にこり
怒る・不機嫌	かちん, むかむか
元気がない	がっかり, がっくり
驚く	びっくり, どつきり

表 2 オノマトペ収集結果

オノマトペ	総数
しっかり	313
ぴったり	117
びっくり	93
すっきり	39
がっかり	36

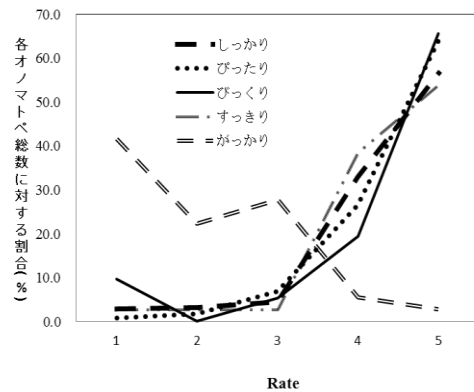


図 1 各 Rate の値におけるオノマトペの割合

3.2 考察

図 1 において, ポジティブな意味を持つものは, Rate の値が高い方が, 出現頻度が高い。一方, ネガティブな意味を持つものは, Rate の値が低い方が, 出現頻度が高い事がわかる。これは, Rate が高ければ消費者の満足度が高く, 逆に低ければ満足度が低い事を意味するので, 妥当な結果である。しかし, この 5 語のオノマトペは Rate の値が 5 から 1 まで全ての Rate に出現している。これは“オノマトペ+否定語”の表現が原因の一つであると考えられる。例えば, 「この洋服はしっかりしていないため, すぐに穴があいた」等である。

4 まとめ

今回 Yahoo! ショッピングの商品レビューを対象として, オノマトペとの関係性を調査した。この調査の結果, Rate によってオノマトペの出現頻度が異なる事がわかった。今後の課題としては, “オノマトペ+否定語”の分析が必要である。

参考文献

- [1] 岡田真, 橋本喜代太: 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, 第 26 回人工知能学会全国大会, 1M1-OS-8a-1, 2012.
- [2] 小野正弘: 擬音語・擬態語 4500 日本語オノマトペ辞典, 小学館, 2007.
- [3] 日本語形態素解析システム JUMAN 7.0, <http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN>