

Twitter 上における「かわいい」の概念構造の可視化

大竹 恒平^{*a} 白石 諒^{**b} 中田 武太朗^{**c} 生田目 崇^{*d} 庄司 裕子^{*e}

^{*}中央大学理工学部

^{**}中央大学大学院理工学研究科

a) otake@indsys.chuo-u.ac.jp b) colorwhite32@gmail.com c) silky.takec@gmail.com

d) nama@indsys.chuo-u.ac.jp e) hiroko@indsys.chuo-u.ac.jp

概要 本稿は、ソーシャルメディア上で「かわいい」という語がどのような文脈で用いられるかを分析し、「かわいい」が有する概念構造を明らかにすることを試みたものである。具体的には、Twitter 上から「かわいい」を含む投稿を収集し、共起関係に基づく特徴語の特定及び、特徴語を用いた共起ネットワークの作成を行った。その結果、「かわいい」は、様々なトピックで用いられることがわかり、それぞれのトピック中に出現する語を評価することで、「かわいい」の概念構造の一端を捉えることが可能であることが分かった。

キーワード ソーシャルネットワーキングサービス, ネットワーク分析, Twitter

1 はじめに

近年、インターネットを介した情報発信・共有や、コミュニケーションサービス等の総称である、ソーシャルメディアが注目を集めている。ソーシャルメディア上では、一つの語でも複数の意味を持つ、多義性を有する語が多く投稿されている。その代表として、「かわいい」という語がある。「かわいい」は、「かわいらしい」という意味で使われることも多いが、「おしゃれな」といったニュアンスが含まれることも少なくない。これまで「かわいい」の持つ多義性や複雑な構造についての研究としては、人文科学的な視点からの大局的な分析研究がある[1][2]。他方で、工学的な研究としては、大倉らによる「かわいい」人工物の特徴を、実験室実験により分析した一連の研究(例えば[3]など)がある。

筆者らは、「かわいい」の多義性は、「かわいい」という語を発する個々人の個性や感性、語られる文脈に大きく依存するという考えのもと、「かわいい」が有する多義的な構造の理解を目的とした研究に取り組んでいる[4]。過去の研究では、Instagram 上に投稿された「かわいい」というハッシュタグ付きの画像を対象に、ハッシュタグ情報を用いた階層型クラスタ分析を行い、「かわいい」画像の分類を試みた。しかしながら、ユーザにより付けられたハッシュタグは属人的であり、異なる言葉(ハッシュタグ)の共通性を評価し、言葉が持つ意味を補完する必要がある事が明らかになった。また、マイクロブログの一つである Twitter 上の「かわいい」が含まれたツイートから抽出した特徴語を用いることで、ハッシュタグを補完できる可能性の一端が示された。

2 研究目的

本稿では、リアルな情報発信コンテンツとしてソーシャルメディアを取り上げ、ソーシャルメディア上で「かわいい」という語がどのような文脈で用いられるかを分析し、「かわいい」が有する概念構造を明らかにすることを試みる。具体的には、ソーシャルメディア上から収集したキーワード「かわいい」を含む投稿を対象に、共起関係に基づく特徴語を抽出する。前述の過去の研究[4]では、出現頻度のみを考慮した特徴語の選定を行っており、特徴語間の構造や、特徴語の分類については十分に検討していなかった。そこで、本稿においては、情報学、工学、心理学など、幅広い分野で活用されている[5]社会ネットワーク分析により、特徴語間の概念構造を検討する。また、これらの特徴語を共起ネットワークとして可視化し、「かわいい」の対象となるトピックの理解を試みる。なお、本稿では、10~20 代に広く利用されている Twitter を対象とする。

3 分析

3.1 「かわいい」を含んだツイート内における特徴語の特定

本稿で分析に用いたデータセットとその分析についてまとめる。はじめに、Twitter 上に 2016 年 7 月以降に投稿された、キーワード「かわいい」を含むツイートを、Twitter 社の API を用いて約 19 万件収集した。このうち、本稿ではリツイートや URL・画像のみのツイートを除いた、96,121 件のツイートを対象とした。

次に、収集した全てのツイートに対して形態素解析を行い、単語単体で意味が理解可能な名詞・固有名詞

(人名, 地名, 組織名など)を抽出した. 名詞・固有名詞は全部で, 382,822 語であった.

次に, 上記で抽出した単語から, 結びつきが強い単語間の関係を抽出するため, 共起回数を値とする信頼度 (Confidence) を算出した. 信頼度はアソシエーション分析で用いられる指標であり, 式(1)により求めた. 式(1)は, 単語Xは条件部, Yは結論部であり, $N(X)$ は単語Xの出現頻度, $N(X \wedge Y)$ は単語XとYが同時に出現した頻度を示す. 信頼度が高いルール程, 条件部から結論部への結びつきが強いことを意味する.

$$\text{Confidence}_{[X \rightarrow Y]} = \frac{N(X \wedge Y)}{N(X)} \quad \dots(1)$$

本稿では, 共起回数 5 回以上でありかつ信頼度が 0.8 以上の 587 単語 (710 ルール) を, 「かわいい」と共に利用される頻出特徴語と定める.

3.2 特徴語を用いた共起ネットワークの作成

3.1 節により抽出した特徴語及びルールを用いて, 共起ネットワークを作成し, 「かわいい」が有する概念構造の可視化を試みる. 可視化にあたり, 特徴語をノード, 特徴語間の共起関係をエッジ, 共起回数をウェイトとして用いた (図 1). なお, 可視化にはグラフマイニングツールである Gephi を用いた. 図 1 においては, 次数が 3 以上のノードのみを示す. 図 1 のノード及びエッジの色はコミュニティを表している. コミュニティはモジュラリティにより分類を行った. その結果, 図 1 においては 13 のコミュニティが確認された. これらのコミュニティのうち, 赤円及び青円で囲った 2 つのコミュニティを例として図 2 に示す. 次に, それぞれのコミュニティが有する「かわいい」が持つ概念構造を明らかにするため, 図 1 のそれぞれのコミュニティ内で利用されている特徴語が利用されているツイートに関して調査, 特徴語のクラスタリングを行った. 図 2 中左のコミュニティは, 某酒造店 (メーカー) が他企業とコラボレーションを行い発売した, 新作の日本

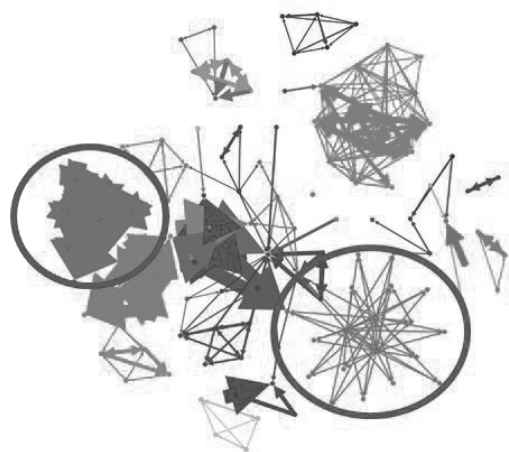


図 1. 「かわいい」が含まれるツイート内の共起関係にある特徴語を用いた共起ネットワークの全体図

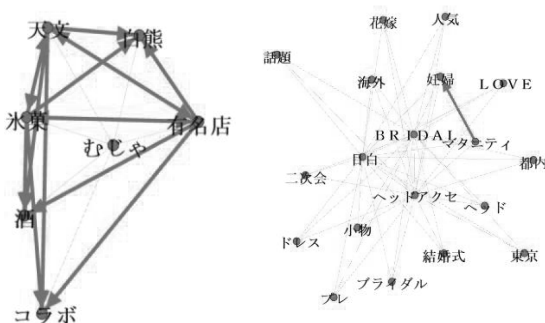


図 2. 共起ネットワーク中のコミュニティの一例

(左: 図 1 左中央の赤円で囲まれたコミュニティ, 右: 図 1 右下青円で囲まれたコミュニティ)

酒のラベルに印刷されたキャラクタに関して「かわいい」を含むツイートで出現した語により形成されたコミュニティであった. そのため, このコミュニティは酒・コラボ・有名店といった, 酒造店に関する語と, コラボレーション企業及びキャラクタに関する語により形成されているコミュニティであることが分かる. 図 2 中右のコミュニティは, 結婚式に関するツイートにより形成されたコミュニティであった. 内容に注目すると, 東京, 目白, 海外といった, 結婚式場に関する語や, ヘッドアクセ, ドレスといった結婚式で貸し出しを行っている商品に関する語により, コミュニティが形成されていることが分かる.

以上の結果から, 「かわいい」という語は, 「商品」「ファッション」「人物」「場所」など, 様々なトピックで用いられることが分かる. また, それぞれのトピック中に出現する語を用いることで, 「かわいい」の概念構造の一端を捉えることが可能であることが分かった.

4 まとめ及び今後の課題

本稿では, Twitter 上に投稿された「かわいい」が含まれるツイートを対象に, 「かわいい」が有する概念構造を明らかにすることを試みた.

今後はソーシャルメディア上に投稿された画像に対して, 本稿で得られたトピックを適用し, 分析を行いたい.

参考文献

- [1] 四方田犬彦, 「「かわいい」論」, ちくま新書, 2006.
- [2] 横幹<知の統合>シリーズ編集委員会 (編), 「カワイイ文化とテクノロジーの隠れた関係」, 東京電機大学出版局, 2016.
- [3] 大倉典子, 後藤さやか, 村井秀聡, 青砥哲朗, “バーチャルオブジェクトを利用した「かわいい」色の検討,” 日本感性工学会論文誌, Vol.8, No.3, pp.535-542, 2008.
- [4] 中田武太郎, 大竹恒平, 庄司裕子, “SNS 上の情報に見られる「かわいい」の多義性についての分析,” 第 18 回日本感性工学大会予稿集, G21, 2016.
- [5] 安田雪, 「実践ネットワーク分析—関係を解く理論と技法」, 新曜社, 2001.