

## Twitter ユーザに向けたコミュニケーション支援のための言い換え表現の抽出

下吉 珠奈

近年、企業による Twitter を利用した販売促進やキャンペーンが目立つようになってきた。そのような企業アカウントの中でも、「やわらかい」表現のツイートをするアカウントは、Twitter ユーザから親しまれており、ユーザが普段なら興味にを持たない内容のツイートでも、フォローしたり、RT したりと、なにかしらのリアクションをとる傾向にある。その一方で、「かたい」文章を用いた企業アカウントによるツイートは、そのツイート内容に関心を持つユーザの目にしか留まらない。さらに、「かたい」ツイートでは、リアクションもとりにくいと考えられ、Twitter を用いて販売促進を行う意義がなくなってしまう。本研究では、このような「かたい」ツイートを「やわらかい」ツイートへ言い換えられるように、言い換え表現のペアを抽出することを目的とした。

Twitter においては、ツイートをイベントごとに収集することにより、イベントについての多様な言い換え表現を抽出することができる。本研究では「かたい」表現と「やわらかい」表現をペアにして抽出するために、イベント単位で収集したツイート集合を、ニュース系ツイートと非ニュース系ツイートに分類する。ニュース系ツイートは、名前 (user-name) とユーザ名 (user-screen\_name) に「ニュース、NEWS、News、news」というキーワードが含まれるツイート、本文に新聞社名と「https://」を含むツイート、ハッシュタグ「#news、#News、#NEWS」を含むツイート、の 3 つの基準に基づき分類する。非ニュース系ツイートは、それ以外のツイートから、アカウントの投稿件数が多いもの、および、本文に「ニュース、NEWS、News、news」が含まれるツイートを除いたものとする。次に、頻出名詞に着目し、一般名詞とサ変名詞のあとにつづく n-gram (n=3~9) のフレーズを、それぞれのツイート群から獲得し、言い換え表現のペアを抽出した。

実験では、イベントごとにツイートを収集し、頻出名詞の後の n-gram (n=3~9) に着目することで言い換え可能な多様な表現が獲得できるか調査した。2012 年から 2014 年の期間で、イベントを 18 件選択し、957,314 件のツイートを収集した。イベントをポジティブとネガティブ、中立の 3 種類に分け、イベントごとの特徴を確認した。実験の結果、ポジティブなイベントでは、お祝いツイートが多く、顔文字や感嘆符を含んだ言い換え表現が多く抽出された。ネガティブなイベントからは、「死亡 ⇄ 亡くなった」というような、表層的には似ていないが、意味は同じものや、恐怖や驚きを表す顔文字などが言い換え表現として抽出できた。中立のイベントは、「発売」や「開始」したという旨の言い換えは抽出できたが、「食事」に関するイベントなど、言い換え表現が抽出しにくいものもあった。

今後の課題として、今回の研究の洞察をもとに、絞り込みの手法を検討していくことが挙げられる。

(指導教員 関 洋平)