

SNSにおける大学生の投稿の分析

An Analysis of College Students Posts in SNS

テーマ：次世代教育・学習環境のデザイン
指導教員：松本 章代

教養学部 情報科学科
1357141 豊島 琢磨

1. 研究背景および目的

2016年時点で、日本人のスマートフォン所持率は72.2%と国民の7割が所有しており、20代の所持率は84.5%と普及している[1]。大学生のSNSアカウントの所持率はLINEで97.9%、Twitterは80.9%と8割以上の学生がSNSを利用しているとの結果が出ている[2]。また、大学生の情報を得るメディアとしてTwitterは男子で52.1%、女子で64.6%と半数以上の学生が日常的にSNSを利用していることが分かる。

SNS上で発信されている情報は、ニュースや公式アカウントからの情報が挙げられるが、話題の投稿や炎上している投稿、様々な意見が飛び交う投稿などユーザーから発信される情報も挙げられる。

例えば、2020年東京五輪で使用される新しいエンブレムのデザインに対し「ダサイ」や「情緒がある」など意見が投稿された。他には、キングコング西野氏の公式ブログ内での「お金の奴隷解放宣言」発言に対して賛否の意見が多く投稿された。

そういった投稿や記事を対象に、本研究では、Twitter上で情報を得ている「大学生のアカウント」と「大学生以外のアカウント」を対象にして投稿を収集し、大学生の投稿にどのような特徴や違いがみられるのか分析を行う。

2. 関連研究

Twitter上での大学生についての研究はすでに行われており、水沼ら[3]は「ソーシャルメディアにおける大学生の“すべき”、“すべきでない”といった意識」に着目し、アンケート形式で調査を行った。これにより、大学生がどういった意識を持ちながらTwitterを利用しているか実態の把握ができています。本研究でも、大学生の投稿から文末表現に注目することで、書き手の気持ちや文章の意味を分析を行う。しかし、大学生のアカウントから投稿を実際に収集することに大きな違いがある。

3. 文末表現

本研究では、松本先生の携わった研究[4]を参考にし、最も文末にある自立語に加え、助動詞、終助詞の含まれている文節を「文末表現」と定義する。

日本語の文では、最も文末にある自立語が文の中心要素になるため、その語が含まれている文節は、その文において書き手が伝えたいことと深く結びついている。そこに注目することでユーザーがどういった内容を投稿しているのかを理解できるのではないかと考えた。また、自立語に加え助動詞、終助詞に注目すると、助動詞は、過去、進行、使役、命令などの意味を文に

付け加える働きを持ち、終助詞は、希望、疑問、禁止などの書き手の気持ちを表す意味を付け加える働きを持っている。文末の持つ意味に注目することで、ユーザーがどういった表現を使っているのか、どういった時に投稿を行っているのか詳しく知ることができるのではないかと考えた。

4. 手法

本研究では、Twitter上の「大学生のアカウント」と「大学生以外のアカウント」の投稿を対象に、文末表現に着目する。

Twitter上での様々な意見が投稿されているツイートとネットニュースに加え、ユーザーがツイートをまとめることができるサイト「Togetter(トゥギャッター)」の記事を対象に、RTや対象について投稿しているユーザーを自動で収集する。収集したアカウントをプロフィール情報から、大学生以外のアカウントをキーワードで取り除き、大学名と学部名がプロフィールに書き込まれているアカウントを判別し、「大学生のアカウント」と「大学生以外のアカウント」に分類する。分類したアカウントからそれぞれ投稿を自動で収集する。

調査対象となるアカウントの判別、収集と、投稿の収集は開発言語RubyとTwitter APIを用いて行う。日本語の係り受け解析にはCaboCha[5]を用いる。また、クラスター分析にはRを用いる。

5. 文末表現の出現頻度に関する分析

39の話題から自動収集した19,519アカウントを分類し、「大学生のアカウント」184アカウントから18,029ツイート、「大学生以外のアカウント」14,486アカウントから1,437,388ツイートにおいて、どのような特徴や違いがあるのか確かめるために、文末表現とその頻度に対して分析を行う。この分析によって、ツイートで使われる文末表現の特徴を視覚的にとらえる。

クラスター分析にあたり、「学生のアカウントのツイート」、「学生以外のアカウントのツイート」を変数とし、出現頻度の合計が多い文末表現30個を分析のためケースとする。両アカウントのツイートの数の差から単語数に差があるため補正が必要であり、10万語あたりの頻度に調整を行う。上位30個の文末表現の出現頻度表は表1に示す。文末表現間の非類似度計算にはキャンベラ距離を用いる。階層的クラスター分析法はワード法を用いる。「両アカウント」で使われる表現、「大学生のアカウント」で使われている表現、「大学生以外のアカウント」で多く使われている表現で分けるためクラスター数は3とする。

クラスター分析の結果、第1クラスターは「両アカ

ウントで多く使用する表現」, 第2クラスター「大学生がよく使用する表現」, 第3クラスターは「大学生以外がよく使用する表現」と区分された。分析の結果を図1に示す。

各クラスターに着目すると, まず, 第1クラスターでは, 過去や存続を意味する助動詞「た」でクラスターが形成されている。しかし, 同じ「なる」という自立語でも, 「大学生のアカウント」では「なりました」といった丁寧な表現を使っており, 「大学生以外のアカウント」では「なった」といった表現を使っている。このことから, 大学生は目上の人と交流する機会があり, 丁寧な表現の「なりました」を使う機会が多いと考えられる。「思った」という表現があり, これは両アカウントであまり差がなく使われている表現である。このことから, 両アカウントで文末に予想や推量の表現をつけ投稿する傾向があることが分かる。

次に, 第2クラスターに着目すると, 他のクラスターには出現しない「食べたい」「行きたい」「繋がりたい」といった願望を表す「たい」を含んだ文末表現がある。この結果から, 「大学生のアカウント」には願望を投稿する傾向があることが考えられる。詳しく見ていくと, 「繋がりたい」といった表現は他の表現に比べ, 普段の投稿ではあまり使う機会が多くないと考えられる。これは, Twitterの機能でハッシュタグといったものがあり, 同じ趣味を持つユーザーを探すために用いられる事が多く, そういったツイートを頻繁に投稿, RTしていることが読み取れる。また, 「食べたい」は他の表現とは別のクラスターに属しており, 同じクラスターには「しよう」「終わった」といった短文の投稿で使われる表現があり, 「食べたい」は普段からの投稿で使われていることが推測される。

最後に, 第3クラスターに着目すると, 「始まるんだよね」「泣きそうになった」「わからない」「更新しました」といった表現で形成されているクラスターがある。これは, 普段の投稿に加え, 「大学生以外のアカウント」のユーザーがRTした同じ投稿から同じ文末表現を抽出したことで, あまり使われないような表現でも出現回数が多くなり形成されたものだと考えられる。また, 「描きました」といった表現も同様に, あまり使われる頻繁に使われる表現ではなく, RTから抽出したことで「大学生以外のアカウント」の特徴的な表現だと判断されたと推測される。また, 「あった」「気づいた」「いた」といった表現があり, 「大学生以外のアカウント」では起こった出来事を報告する投稿を行う傾向があると予想される。

6. まとめ

大学生の投稿の特徴を分析するために, 文末表現に着目した。Twitterのアカウントを収集し, 「大学生のアカウント」と「大学生以外のアカウント」に分類し, 投稿に対してクラスター分析を行った。「大学生のアカウント」と「大学生以外のアカウント」の投稿から違いや共通点を見出すことができた。また, 自立語を含めて分析することで, 同じ助動詞の働きでも大学生が使用する表現, 大学生以外が使用する表現といった部

表 1. 上位 30 語の出現頻度

文末表現	大学生のアカウント	大学生以外のアカウント	文末表現	大学生のアカウント	大学生以外のアカウント
思う	26.94	29.69	思った	5.71	5.82
ある	26.48	25.78	見守ります	8.68	1.44
ない	21.46	19.97	かわいい	4.34	5.38
お願いします	16.89	14.91	描きました	1.60	7.72
思います	14.61	15.37	なった	3.42	5.72
する	12.10	11.25	わかる	4.11	4.89
なる	10.73	11.16	ありません	4.11	4.73
あります	9.82	11.92	やめて下さい	4.57	4.06
可愛い	7.99	7.13	いる	3.88	4.62
なります	7.31	7.73	欲しいだけなんです	4.34	4.03
良い	7.53	7.02	何くてもいいんです	4.34	4.03
お願い致します	9.59	4.58	やばい	4.57	3.75

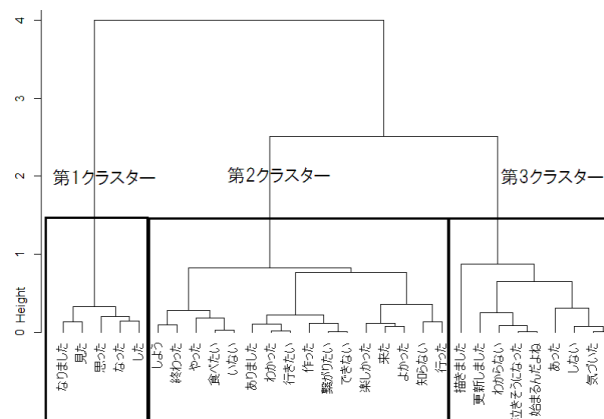


図 1. クラスター分析

分も分析することができた。今後はさらに, RTを含めない場合の分析や, 助動詞, 終助詞の働き別に分析を行い投稿の特徴を調査する必要がある。また, 大学生のアカウント数を増やし, 投稿の分析を行うことでより細かい特徴が見えてくるのではないかと考えられる。

参考文献

- [1] MRC, 【最新版】2016年のスマホ普及率を男女・地域・年代別に大公開! まさにスマホオンリー時代! マーケティングがこれからどう変わるべきか予想してみた。 , <https://marketing-rc.com/article/20160731.html?r=m>
- [2] ReseMom., 大学生が情報を得るメディアは「テレビ」「Twitter」…SNSは「LINE」利用, <http://resemom.jp/article/2016/05/19/31542.html>
- [3] 水沼友宏, 菅原真紀, 池内涼, 小山照夫, “大学生のTwitterにおける行動範囲に関する分析”, 情報社会学会誌 Vol.8, No.1, pp.23-37, 2013-05.
- [4] 松本章代, 小西達裕, 高木郎, 小山照夫, 三宅芳雄, 伊東幸宏, “文末表現を利用したウェブページの主観・客観度の判定”, 第一回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2009), pp.23-37, 2009.
- [5] 工藤拓, 松本裕治, “チャンキングの段階適用による日本語係り受け解析”, 情報処理学会論文誌, Vol.43, No.6, pp.1834-1842, 2002.