

実際のツイートを収集して感情との関連を検討する方法の提案 —KJ法および最適尺度法を用いて—

Proposal on a method to consider the relationship between actual tweets collected and emotions : By KJ method and Optimal Scaling

堀切大器*・谷島弘仁**

Taiki HORIKIRI, Hirohito YAJIMA

要旨：本研究の目的は、ツイート頻度の客観性を担保するために、ツイートを一つの単位としたテキストデータの分析を行うこと。また、なるべく多くの情報を捨象せずにテキストデータを数量化し、感情変数との関連を検討することで、探索的にツイート内容とネガティブ感情の関連を検討すること、およびその手法の提案である。25人の過去2週間のツイート（234ツイート）をラベル化し、KJ法式のグループ編成とカテゴリカル主成分分析およびクラスター分析を用いてグルーピングを行った。その結果、【イベントツイート】【情緒的ツイート】【ユーモアツイート】の3つが生成された。非線形正準相関分析により、大学生用ストレス反応尺度（情動反応）「抑うつ」「不安」「怒り」との関連を検討したところ、【ユーモアツイート】と「抑うつ」「不安」の関連、および【情緒的ツイート】と「怒り」の関連が示唆された。

キーワード：Twitter, KJ法, 最適尺度法, カテゴリカル主成分分析, ユーモアツイート

問題と目的

Twitterは140文字までの短いメッセージを投稿（以下：ツイート）するサービスであり、その時考えていることやしていること等のメッセージや写真などを他者と共有することで様々な情報の交換を行なうことができる。とりわけ若年層を中心に利用されており、登場から10年以上が経った今日においても、日常に最も浸透しているSNSの1つである。先行研究では、Twitterによるネガティブ感情のツイートについて検討されている。北村・佐々木・河合（2016）では「気分が楽になる」と感じてツイートする内容として、腹立たしい・悲しい・疲れた・さびしいといったネガティブ感情のツイートが上位を占めた。また遠藤・藤（2015）や藤・遠藤（2017）では、

*ほりきり たいき 文教大学人間科学部臨床心理学科 非常勤講師

**やじま ひろひと 文教大学人間科学部

ネガティブ感情（抑うつ・不安・怒り）に関するツイートが長期的なネガティブ感情の持続・増大に繋がることを示した。堀切（2018）では、中程度のツイート頻度が低頻度群、高頻度群に比べてネガティブ感情（抑うつ・不安・怒り）の得点が低いことが示された。以上のようにツイートとネガティブ感情の関連は示唆されるものの結果のコンセンサスを得ていない。その理由として、ツイート内容とツイート頻度の測定の違いが考えられる。遠藤・藤（2015）や藤・遠藤（2017）では、ツイート内容を感情的側面と事実に側面の2側面のみを想定しており、あくまでネガティブ感情の共有に焦点を当てている。北村他（2016）では多様なツイート内容を想定しているが、両者ともツイート頻度をリッカート尺度により測定していることから頻度の客観性が担保されていない。特に遠藤・藤（2015）や藤・遠藤（2017）では統計手法としてパラメトリック検定を用いているが、北村他（2016）のツイート頻度の割合を見るに正規分布に従うか疑わしく、堀切（2018）ではツイート頻度の正規性は示されなかった。そこで本研究では、正確なツイート頻度を測定するため堀切（2018）で収集したツイートのログデータを用い、かつ、ネガティブ感情のツイートに限らず多様なツイート内容を想定し感情変数との関連を検討する事を目的とする。そのためにはツイートのログデータに含まれる情報をできるだけ捨象せずに数量化する必要がある。

ツイートのログデータは写真等の画像データを除けばテキストデータとなる。テキストデータを分析する手段として、日本ではKJ法が有名である。KJ法とは川喜多二郎によって開発された発想法であり、おおまかに「ラベルづくり」「グループ編成」「図解化」「叙述化」の4つのステップでテキストデータを分析する（田中，2012）。また近年ではテキストマイニングといった手法も隆盛している。テキストマイニングとは、言語データを対象に、データの背後に潜む有益な情報を採り出す、理論・技術の総称である（小林，2017）。

テキストマイニングがKJ法に勝る点として、KJ法の分類結果が研究者の解釈によって異なるという信頼性の問題を解決している点と、扱えるデータ数が多いという点が挙げられる。一方で基本的な手法として形態素解析を用いることから、複数の単語で構成されて初めて意味をなすテキストまで分解してしまう（藤井・小杉・李，2005）だけでなく、頻度の測定が形態素数に依存するため、本研究の目的であるツイート頻度の正確な測定が行えない。KJ法の場合、意味のまとまりを1つの単位で分析が可能であることから、ツイート単位での頻度の測定が可能で、かつ、テキストデータの文脈をつかみやすいという利点がある。以上のことから、KJ法のように意味のまとまり（ここではツイート）を1つの単位としてテキストデータを分析しつつ、テキストマイニングのように信頼性を担保し、かつ、テキストデータを量的データに変換する手法を考案することで、正確なツイート頻度の測定と感情変数との関連が検討できる。

KJ法を数量的な手法で代用しようとする試みは、KJ法創案当時から行われている。中村・沢崎・吉川・近藤・小池・立花（1970）は、大量データのグループ編成と、グループ間の空間配置を簡便化するとともに、KJ法のラベル集めの過程で用いる「情念」という曖昧な基準をより明確にする手法を考案した。具体的には、ラベルの類似性の評価として数人の評定者による一対比較法を用い、かつ、空間配置の方法として数量化Ⅳ類を用いて、KJ法の手順を機械的な処理で代用している。また田中・大鹿（1978）も、一対比較法を用いる点では中村他（1970）と同様だが、空間配置の手法として最小次元解析という手法を用いている。この手法はラベルの相互間の類似性を多段階で評定し、それを低次元空間に布置することが可能である。そして布置されたラベル相互間の距離を類似度とし、クラスター分析を用いてグループ編成を行っている。これによりグループ編成の過程を簡便化するとともに、類似度に基づく合理的な分類を可能にしている。

しかしこれらの研究で用いられた一対比較法による分析は、精度の高い類似性の基準を算出できるが、それは統計解析の段階の話であり、類似性そのものの評定においては、やはり明確な基準があるわけではない。すなわち一対のラベルの類似性を評定する際、ラベル同士が似ているかどうかについての判断は「情念」と同じく曖昧なものである。

そもそも何をもってラベル同士が類似していると言えるのだろうか。葛西（2008）は認知言語学の観点から KJ 法の類似性について論じている。これは KJ 法のグループ編成の手順を「カテゴリー化」のプロセスと同一視した発想であろう。人間には事物から何らかの類似性や一般性を抽出することで、事物間にあるまとまりを認識し分類することのできる能力が備わっていると考えられる。このような事物をグループにまとめる認識上のプロセスをカテゴリー化という。そしてこのカテゴリー化の理論として支持されているものにプロトタイプ理論がある。プロトタイプとは「カテゴリーの最も典型的な成員の持つ特徴の抽象的合成物もしくは集合体」である。そして人間が事物をカテゴリー化する場合、そのプロトタイプを核とし、その周りにさまざまな成員を位置付けることで、全体を構造化している（河上，1996）。すなわちカテゴリーの成員全てに共通する属性があるのではなく、そのカテゴリーの理想的なプロトタイプの属性とどの程度合致しているかによってカテゴリーは構造化されているのである。では、なぜカテゴリーの成員全てに共通する属性がないにもかかわらず、成員同士が相互に関連しあい1つのまとまりを形成し得るのだろうか。この説明を与えるのが「家族的類似性」という考え方である。家族的類似性とは Wittgenstein（1953 藤本・坂井訳 1968）によって提唱された概念であり「すべての対象に共通する特性があるのではなく、さまざまな類似性が重なり合い、交差して、全体としてゆるい類縁関係を結んでいる状態（Nisbet, 2003 村本 2004 訳注より引用）」とされる。このように事物は家族的類似性により1つのカテゴリーを成している（鷲見，2013）。

以上のようなプロトタイプ理論に則り、KJ 法式のグループ編成は、家族的類似性に基づくカテゴリー化のプロセスと同一視することで、一対比較法による類似性の評定よりも人間の認知処理の仕方を考慮に入れた際、より妥当な手法であると言える。実際に一般的なプロトタイプ研究では Shaver, Schwartz, Kirson, & Connor (1987) や Li & Fisher (2004) のように、対象となる 100 語前後の記述をカードに印刷し、調査協力者が自由に分類する方法を用いる。この手続きは KJ 法のラベル集めと類似している。以上のことから、本研究では認知言語学におけるカテゴリー化の観点から、個々のツイートの家族的類似性に基づくグルーピングが妥当であると判断し、KJ 法式のグループ編成の方法を採用する。しかしこれだけでは KJ 法の課題を解決したとは言えない。カテゴリーの構造について、プロトタイプを中心として周辺へ拡がるという関係は、抽象度の階層関係においても同様のことが言える（河上，1996）。カテゴリーの階層関係は大きく3つのレベルに階層化され、抽象化のレベルの高い順に見ると、上位カテゴリー（例：動物）、基礎レベルカテゴリー（例：犬）、下位カテゴリー（例：柴犬）に分類できる。特に基礎レベルカテゴリーが特権的な地位を持っており、言語形式も短く単純で、最もイメージしやすい。基礎レベルカテゴリーが他の抽象レベルに比べ特権的地位にあるのは、カテゴリー内での高い類似性が保たれていると同時に、異なるカテゴリー間で高い非類似性が保たれているためである。このことから KJ 法式のグループ編成を考える際、基礎レベルカテゴリー内の下位カテゴリー同士は高い類似性を持つため、家族的類似性に基づくグルーピングが可能である。しかし、ラベル集めのサイクルを数回繰り返したあたりで、抽象度が上がると、そこで生成されるカテゴリーは、基礎レベルカテゴリーとなる。すると、基礎レベルカテゴリー同士は「差異の多さ」が顕在化する

ため、家族的類似性に基づくグルーピングが困難になると予測できる。つまり基礎レベルカテゴリーを包括する上位カテゴリーの生成には、家族的類似性とは別の基準を設ける必要がある。

そこで、テキストマイニングの知見を応用できる。テキストマイニングでは、調査対象者ごとの各形態素の出現頻度を算出し、調査対象者×形態素の対応表を作成することで多変量解析を行うことができる。この対応表を用いて主成分分析をする事で、抽出された成分負荷をもとに形態素をユークリッド空間に布置できる。そして抽出された成分負荷、すなわち布置された形態素の座標を用いてクラスター分析をする事で、似通った形態素を数量的に分類できる（藤井他, 2005）。この手法をKJ法に応用すると、調査対象者×生成されたカテゴリーの対応表を作成することで、同様の分析が可能となる。つまり下位カテゴリーから基礎レベルカテゴリーへは、家族的類似性に基づくグルーピングが人間の認知処理的に妥当であったが、基礎レベルカテゴリーから上位カテゴリーへのグルーピングは、基礎レベルカテゴリー同士の非類似性が高いことから家族的類似性よりも数量的なグルーピングの方が適している。以上のことから、本研究ではツイートのテキストデータを他のカテゴリーとの差異が目立つ基礎レベルカテゴリーになるまでKJ法におけるグループ分けをし、基礎レベルカテゴリーから上位カテゴリーへのグルーピングは数量的な方法を用いることとする。

しかし数量的データに変換する際に留意点がある。主成分分析や因子分析は、本来正規分布に従うデータに適用となるパラメトリック法である（対馬, 2010）。堀切（2018）ではツイート頻度の正規性が示されなかったため、ツイートデータに対して主成分分析や因子分析は適用できない。このような場合、代替案として「最適尺度法」を用いた分析が有効である（坂本・一門, 2013）。最適尺度法とは、交互最小二乗法と呼ばれる反復法を用いて、各尺度変数に対する最適な数量化を行うものである。これによって、名義尺度や順序尺度などのカテゴリカルデータを適切に分析できる。また標準的なデータ解析は、観測数が少なすぎる場合や、変数が多すぎる場合、1変数あたりの値が多すぎる場合などにデータセットがうまく働かないことがあるが、最適尺度法を用いることでこれらの問題を回避できる（IBM, 2017）。最適尺度法のうち、カテゴリカル主成分分析、非線形正準相関分析、多重応答分析、コレスポンデンス分析の4つは次元縮尺と呼ばれる多変量データ分析の一般分野に分類される。いずれの手法も抽出される成分負荷や数量化得点を用いてユークリッド空間に変数を布置することで、視覚的にデータを把握できる。つまり最適尺度法を用いることで正規性が仮定されていないデータに対しても、最適な空間布置や変数同士の関連の検討できる。このような最適尺度法とKJ法を組み合わせた例はいくつか存在する。

加藤・柏木（2000）は25名に対するインタビュー調査で得られた逐語録をもとにKJ法を実施している。そして「調査対象者が、生成されたカテゴリーに属する語りを何回行ったか」というエピソードの出現頻度を算出して対応表を作成し、カテゴリカル主成分分析を用いて解析している。また鈴木・和田・村上（2005）では、自由記述によって得られたテキストデータをKJ法によって分析し「調査対象者がカテゴリーに属する記述を行っていたか否か」の二値データの対応表を作成して多重応答分析を用いて解析している。しかしこれらの研究では、軸の解釈や空間布置までにとどまっており、外的変数との関連は検討されていない。

外的変数との関連を検討する方法として、同じ最適尺度法の1つである非線形正準相関分析が挙げられる。非線形正準相関分析は、主成分分析のような変数自体の関係ではなく、変数の2つ以上の組の関係を分析できる（IBM, 2017）。すなわち複数の変数をグループ化して1つの変数群として定義し、その変数群同士の関連を分析できる（石村・加藤・劉・石村, 2013）。また非

線形正準相関分析では、数量化された変数とオブジェクトスコアの相関係数（成分負荷）を算出し、二次元空間にプロットすることで、変数同士の関係を視覚的に表現できる。

以上をまとめ本研究の具体的手続きを総括する。第一に KJ 法式のグループ編成を用い、ツイートのテキストデータから基礎レベルカテゴリーを生成する。この際、いたずらに抽象度を上げ過ぎないように、葛西（2008）が提唱した関連性法定質的分析に用いられる「抽象度を上げないラベル付け」を援用する。そして「調査対象者が基礎レベルカテゴリーに属するツイートを何回行ったか」という調査対象者×基礎レベルカテゴリーの対応表を作成する。

第二に、対応表は度数データであるためカテゴリカル主成分分析を行う。抽出する次元数は鈴木他（2005）を参考に、固有値と Cronbach の α 係数を基準とし、基礎レベルカテゴリーの成分負荷を降順で並び替えた際に解釈が可能なものまでを採用する。そして得られた成分負荷を用いて基礎レベルカテゴリーを空間布置する。次に基礎レベルカテゴリーの成分負荷、すなわち布置された基礎レベルカテゴリーの座標をもとに階層クラスター分析を行い上位カテゴリーを生成する。

第三に、調査対象者×上位カテゴリーの対応表を作成する。するとテキストデータは豊富な情報を損なわずに信頼性を保ちながら数量化できるため、非線形正準相関分析を用いたツイートの上位カテゴリーと感情変数の関連を検討する。

方法

調査対象者 本研究では堀切（2018）で収集されたデータを用いた。文教大学の大学生および大学院生のうち現在 Twitter を利用している者 44 名（男性 16 名、女性 28 名）が対象者となった。

手続き 2017 年の 6 月から 10 月にかけて、授業等で対象者を募集し、協力が得られたものに対し、後日個別で再度説明を加え、同意を得られた者に対して調査を実施した。調査内容は質問紙の回答と、調査日から過去 1 ヶ月間（調査日から過去 30 日間）のツイートの収集とした。ツイートの収集の手順は、著者が作成した研究用の Twitter アカウントのフォロー申請を承認してもらい、投稿を検索、copy & paste で Word に添付するという方法で行った。尚、本調査は文教大学人間科学部倫理審査委員会により承認された上で実施した。

質問項目（1）大学生用ストレス反応尺度（情動反応）。本尺度は、新名・坂田・矢富・本間（1990）によって作成された心理的ストレス反応尺度のうち、情動的反応の 15 項目を用いた。新名他（1990）はこの 15 項目を因子分析によって「抑うつ気分」「不安」「不機嫌」「怒り」の 4 因子構造としたが、尾関（1990）では「不機嫌」と「怒り」が 1 つの因子にまとまったため「抑うつ」「不安」「怒り」の 3 因子構造となった。さらに尾関・原口・津田（1991）でも 3 因子構造を採用しているため、本研究でも「抑うつ」「不安」「怒り」の 3 因子構造を採用した。本研究では、情動的反応のみを使用したため「ここ 1 週間のあなたの心の状態を、よく表すものとして当てはまるものに、○を付けてください。」という教示を与え「あてはまらない（0 点）」～「非常に当てはまる（3 点）」の 4 件法で回答を求めた。本尺度は新名他（1990）、尾関（1990）ともに Cronbach の α 係数は高い値を示しており、尾関他（1991）では MPI との相関が示されたことから、本尺度の信頼性・妥当性が示されている。（2）ツイート動機についての自由記述。ツイートの動機について「Twitter でつぶやくとき、あなたの場合、どのような気持ちのときにつぶやきますか」という教示を与え、自由記述による回答を求めた。

結果

44人の全調査対象者より得られた、調査日から過去1ヶ月間のツイートの総数は、1546ツイートであった。しかし過去1週間の感情状態を説明する仮説を立てる際、1ヶ月という期間は長すぎると判断し、以下の分析では過去2週間のツイートに限定した。また画像および他のSNSやサイトへのリンクが添付されていたツイートに関しては、言語的メッセージ以外の情報が多く含まれることから以下の分析から除外した。以下より堀切(2018)における「中頻度群」のデータのみを扱う。理由は度数のばらつきが大きく、極値のデータが分析に影響を及ぼす可能性を排除するためである。中頻度群は「1週間に1回以上、1日に2回未満」のツイート頻度の者であり、サンプル数は25であった。尚、分析はSPSS statistics 23を用いた。

KJ法によるグループ化を用いた基礎レベルカテゴリー(ツイートカテゴリー)の生成 中頻度群の対象者から収集した調査日から過去2週間までのTwitterの投稿のログデータを用い、グループKJ法を実施した。分析は心理学を専攻する大学院生2~6人で数日に分けて実施した。全てのラベルの枚数は234枚であった。ラベル集め・表札づくりの作業を3ラウンド繰り返したところで、各グループ間の差異が顕在化したため、抽象度が基礎レベルカテゴリーに達したと判断し作業を終了した。第3段階の時点でカテゴリー数は22であった。また第3段階目のラベル集めで、元ラベルが他のラベルとグルーピングされずに残ったものは4枚であった。それらは以下の分析からは除外した。またカテゴリーMは、調査対象者の1人のみのツイートによって構成されていたため、以下の分析では除外した。以下、KJ法によってまとめられたツイートのグループを総称して「ツイートカテゴリー」と記載する。ツイートカテゴリーおよびツイートの具体例を表1に示す。

上位カテゴリー(ツイートクラスター)の生成 上位カテゴリーを生成するため、まずカテゴリカル主成分分析を実施した。分析によって抽出する成分数を決定するために、21成分を抽出したところ固有値1以上は8成分であった。そこで1成分から再度カテゴリカル主成分分析を行

表1 KJ法によるツイートカテゴリーと具体例の結果

記号	ツイートカテゴリー	具体例	記号	ツイートカテゴリー	具体例
A	楽しかった出来事や充実した体験の感想や思い出を語るつぶやき	「昨日の〇〇会楽しかった」	L	困惑や疑問を投げ出すつぶやき	「なんか今日は足の裏が痛いんです……なんだろう」
B	つらいことや怒りの気持ちを軽く言ったりおもしろく言ったりしているつぶやき	「〇〇の冷房は寒すぎる説」	M	調子が悪いことを伝えるつぶやき	「体調悪い、気持ち悪い」
C	趣味などについての感想や感動、興奮を表したつぶやき	「時々すぐ見たくなくて画像欄から選ってます、幻想的で好きです」	N	夢について語っているつぶやき	「久しぶりに長編の夢を見ました」
D	自虐的ユーモアを語るつぶやき	「今日傘ささない友達をかばっていたら、シャツの右半分だけびしょ濡れで左半分と比べると芸術的なコントラストだった」	O	嬉しさや喜びを表したつぶやき	「念願のPS4をてにいれたぞ！」
E	自分の現状や願望について語るつぶやき	「〇〇くんはずっと一緒にいれますように」	P	何かを気にしたり、不安に思ったり、心配したりするつぶやき	「連絡待ち～生きてるのかな」
F	おもしろいことや驚いたことを語るつぶやき	「常温用の生キャラメルを冷蔵したらキャラメルになった」	Q	これからすることを報告するつぶやき	「ライブ楽しみメドレーでもいいからもう全曲歌って欲しい」
G	最近の日常や感じたことについて語るつぶやき	「最近キンモクセイの香りがして、秋を感じる」	R	自分を励まし、鼓舞するつぶやき	「とりあえず今日の目標は11時就寝！」
H	怒りや不平不満を表すつぶやき	「しめほどむかつくんだけだ」	S	Twitterの機能上での行動の補足など、道具的な伝達をしているつぶやき	「〇〇中の同窓会アカウント作りました」
I	不安や焦り、迷いを表しているつぶやき	「だめた決められない」	T	これから取り組む大変な事を、反語的かつユーモラスに表現しているつぶやき	「俺、今日の練習が終わったら結婚するんだ…くらい身体の調子」
J	これからすることについて嘆いたり、駄々をこねたりしているつぶやき	「学校行こうか悩む行くことは確定」	U	食べ物がおいしかったことや、好きであることを伝えるつぶやき	「ワインとチーズうまうま」
K	ガッカリしたり、落ち込んだりした気持ちを表したつぶやき	「ライブ…。せめてグッズ買いに行きたかったな。またツアーの時は絶対行く！」	V	持論や偏見を展開するつぶやき	「目に見えるものも見えないもの何重もの構造があって全ての本質は矛盾にあるとおもう」

い、成分負荷得点の降順にカテゴリーの並び替えた結果、4成分抽出したところで解釈が困難となったため、成分数を3成分に決定した。最終的に3成分を抽出するように分析した結果、固有値5.22の第1成分（Cronbachの $\alpha = .85$ ）、固有値3.32の第2成分（Cronbachの $\alpha = .73$ ）および固有値2.61の第3成分（Cronbachの $\alpha = .65$ ）が得られた（表2）。続いてカテゴリカル主成分分析によって得られた3つの成分負荷を用いて、ツイートの上位カテゴリーを合理的にグルーピングするため、Ward法の階層クラスター分析を行った。デンドログラムを参考にクラスター数を4に決定した（図1）。

表2 カテゴリカル主成分分析の結果

ツイートカテゴリー	成分負荷量		
	I	II	III
J これからすることについて嘆いたり、駄々をこねたりしているつぶやき	.95	.09	-.15
Q これからすることを報告するつぶやき	.95	.16	-.02
L 困惑や疑問を投げ出すつぶやき	.92	.03	-.09
C 趣味などについての感想や感動、興奮を表したつぶやき	.81	.25	-.31
A 楽しかった出来事や充実した体験の感想や思い出を語るつぶやき	.69	.25	-.14
R自分を励まし、鼓舞するつぶやき	.66	-.13	.08
U食べ物がおいしかったことや、好きであることを伝えるつぶやき	.52	-.22	.16
G最近の日常や感じたことについて語るつぶやき	.51	-.06	.50
E自分の現状や願望について語るつぶやき	-.12	.82	-.12
H怒りや不平不満を表すつぶやき	-.16	.74	-.18
P何かを気にしたり、不安に思ったり、心配したりするつぶやき	-.13	.74	.16
Bつらいことや怒りの気持ちを軽く言ったりおもしろく言ったりしているつぶやき	-.04	.60	.19
O嬉しさや喜びを表したつぶやき	-.15	.53	-.49
S Twitterの機能上での行動の補足など、道具的な伝達をしているつぶやき	-.24	.47	-.18
Tこれから取り組む大変な事を、反語的かつユーモラスに表現しているつぶやき	-.03	.27	.72
N夢について語っているつぶやき	-.03	.02	.58
D自虐的なユーモアを語るつぶやき	.07	.05	.57
Fおもしろいことや驚いたことを語るつぶやき	.42	.41	.57
I不安や焦り、迷いを表しているつぶやき	.32	.06	-.47
Kガッカリしたり、落ち込んだりした気持ちを表したつぶやき	.23	-.42	-.21
V持論や偏見を展開するつぶやき	-.15	-.20	-.12

注)太線枠は成分負荷量.5以上で採用された項目の基準値

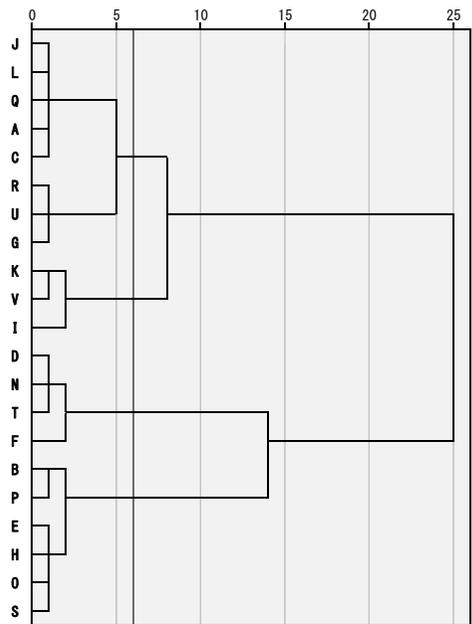


図1 クラスター分析の結果

第1クラスターはA、C、G、Uなどのように、現実場面での出来事に関するカテゴリーによって構成されており、第1成分の負荷量が高いカテゴリーと概ね一致したため、第1クラスターを【イベントツイート】と命名した。第2クラスターはH、O、Eなどのように感情や願望に関するカテゴリーによって構成されており、第2成分負荷量が高いカテゴリーと概ね一致した。しかし、Sのみ他のツイートカテゴリーと違い感情や願望との関連が薄いと思われた。表2を見ると、Sの第2成分の成分負荷量は.47と低かったため以下の分析からは除外し、第2クラスターを【情緒的ツイート】と命名した。第3クラスターはD、F、Tなどのように、ユーモアに関するカテゴリーによって構成されており、第3成分の負荷量が高いカテゴリーと概ね一致したため、第3クラスターを【ユーモアツイート】と命名した。第4クラスターの解釈は困難である。I、K、Vの3カテゴリーによって構成されているが、解釈するにはややまとまりがない。表2を見ると、どの成分負荷に対しても低い値を示しており、本研究のサンプルでは説明しきれなかったカテゴリー群であると解釈し以下の分析からは除外した。尚、この分析によってグルーピングされた変数を総称して、以下「ツイートクラスター」と記載する。

ツイートクラスターと感情変数との関連 各ツイートクラスターにおけるツイート頻度と感情変数との関連を検討するため、非線形正準相関分析を行なった（表3）。ツイートクラスターの3変数を1つの変数群、大学生用ストレス反応尺度「抑うつ」「不安」「怒り」の3因子を1つの変数群、計2つの変数群を投入した。その際、Shapiro-Wilkの正規性の検定にて正規分布が仮定されなかった変数は線形関係を仮定しない順序付きカテゴリカルデータとして扱う必要がある（坂本・一門, 2013）。ツイートクラスターおよび大学生用ストレス反応尺度の3因子の正規性が確認できなかったため、順序付きカテゴリカルデータとして分析を行った。各変数の成分負荷に基づいて結果を二次元プロットに示した（図2）。

その結果、ツイートクラスターはそれぞれ離れた位置にプロットされたことから各クラスターの個別性が高いことが示された。また【ユーモアツイート】と「抑うつ」「不安」の変数が近くにプロットされ関連が示された。さらに【情緒的ツイート】と「怒り」の変数も比較的近くにプロットされた。一方【イベントツイート】はどの感情変数とも距離が遠く関連は見られなかった。

表3 非線形正準相関分析の成分負荷表

変数群	成分負荷量	
	I	II
ツイートカテゴリー		
イベントツイート	-.79	-.17
情緒的ツイート	-.21	.84
ユーモアツイー	.46	-.31
大学生用ストレス反応尺度		
抑うつ	.46	-.51
不安	.69	-.38
怒り	-.07	.29

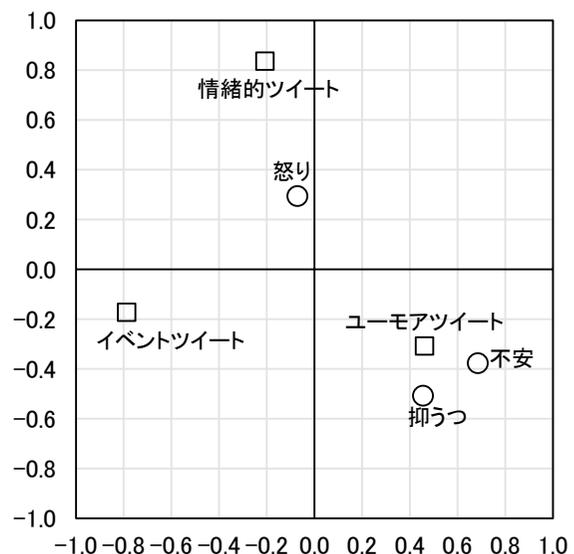


図2 非線形正準相関分析の結果

考察

本研究の目的は、ツイート頻度の客観性を担保するために、ツイートを1つの単位としたテキストデータの分析を行うこと。また、なるべく多くの情報を捨象せずにテキストデータを数量化し、感情変数との関連を検討することで、探索的にツイート内容とネガティブ感情の関連を検討すること、およびその手法の提案であった。

その結果、ツイートクラスターとして【イベントツイート】【情緒的ツイート】【ユーモアツイート】の3つが生成された。北村他(2016)では、テキストマイニングの手法を用い、ツイート内容を出来事系、占い系、お得系、季節系の4つに分類していたが、本研究の結果と概ね一致したと思われるのは【イベントツイート】のみである。これはTwitter自体が、その時考えていることやしていること等のメッセージを投稿するツールであるためと考えられ、実際に本研究でも【イベントツイート】のカテゴリー数が最も多かった。一方、本研究では占い系、お得系、季節系といったグルーピングは行われず【情緒的ツイート】【ユーモアツイート】といったクラス

ターが生成された。これは形態素レベルの分析ではなく、ツイートを1つの単位として投稿者の意図や文脈を考慮した分析を行ったことによる結果と考えられる。例えば【ユーモアツイート】のカテゴリ T の具体例「俺、今日の練習が終わったら結婚するんだ…くらいの身体の調子」というツイートは、形態素レベルでは「結婚」「調子」などの単語が抽出される可能性があるが、このツイートの「～したら結婚するんだ…」というのは、いわゆる「死亡フラグ」と呼ばれるネットスラングであり、要は練習が終わったら死んでしまうのではないかとというほど体調が悪いということを意味している。このように、ツイートにはスラングやユーモア、メタファー等の表現が多く含まれており、そうした投稿者の意図を踏まえて分析することが重要と考えられる。

感情変数との関連は【情緒的ツイート】と「怒り」感情、【ユーモアツイート】と「抑うつ」「不安」感情に関連が見られた。【情緒的ツイート】と「怒り」感情に関連が見られたことは、藤・遠藤(2017)の結果と一致している。一方で【ユーモアツイート】と「抑うつ」「不安」の関連が見られたことは本研究による新たな知見と言える。ユーモアにはストレスを低減させる効果がありストレスコーピングの1つとしても用いられる(Martin & Lefcourt, 1983)。投稿動機を尋ねた自由記述では「不満を笑いに変える時」「笑える失敗をした時」にツイートをするという回答が見られた事からも、ネガティブ感情を和らげるためにユーモラスなツイートがされていた可能性が示唆される。しかし非線形正準相関分析の結果の出力は成分負荷を元に類似度を示すだけであり、統計的に有意な関連が示されるわけではない。また因果関係を想定した検定ではないのでユーモラスなツイートが実際にネガティブ感情を低減する効果あるとも言い切れない。あくまで探索的な研究であり今後は仮説検証的な検討を行う必要がある。

本研究の分析手法は、テキストデータを数量化し、リッカート尺度の外的変数との関連を検討するというものであり、量的研究と質的研究の狭間を埋める試みとなった。一方で以下のような課題も散見された。1つ目は本調査のサンプルでは説明できないクラスターが生まれたことである。クラスター分析によって4つのクラスターが抽出されたが、低い成分負荷量でまとまった第4クラスターおよびツイートカテゴリ S の解釈がしがたく分析から除外した。なるべく多くの情報を捨象せずに数量化することを目指したが、次元分解をする都合上どうしてもこのようなノイズが生まれてしまう点は再考の余地がある。また最適尺度法を利用した心理学の研究が少なくレビューが限られてしまった。しかし、あらゆる変数にリッカート尺度・パラメトリック検定を用いることは実際の数値や分布と解離を生み出してしまう恐れがある。今後は最適尺度法やそれに類する数理化理論等の分析法を心理学研究にも取り入れることで新たな発展に期待したい。

引用文献

- 遠藤寛子・藤桂(2015). ネガティブツイートは何をもたらすのか?—SNS 上における社会的共有の逆説的效果— 日本心理学会第 79 回大会発表論文集, 275.
- 藤桂・遠藤寛子(2017). 怒り, ツイートし, その果てに—SNS 上における社会的共有の効果に関する縦断的研究— 日本心理学会第 81 回大会発表論文集, 206.
- 藤井美和・小杉考司・李政元(2005). 福祉・心理・看護のテキストマイニング入門 中央法規出版.
- 堀切大器(2018). Twitter において発信されるメッセージの探索的検討および, そのメッセージと発信者の感情や動機との関連について—KJ 法と最適尺度法を用いて— 文教大学大学院人間科学研究科臨床心理学専攻 平成 29 年度 修士論文.
- IBM(2017). IBM SPSS Categories 25. 日本アイ・ビー・エム株式会社, https://www.ibm.com/docs/en/SSLVMB_25.0.0/pdf/ja/IBM_SPSS_Categories.pdf (2022 年 11 月 3 日).
- 石村貞夫・加藤千恵子・劉晨・石村友二郎(2013). SPSS によるカテゴリカルデータ分析の手順 東京図書.

- 葛西俊治 (2008). 関連性評定質的分析による逐語録研究—その基本的な考え方と分析の実際— 札幌大学院人文学会紀要, 83, 61-100.
- 加藤千恵子・柏木恵子 (2000). 成人期前期の日本男性の結婚観・仕事観—インタビューおよびKJ法・最適尺度法による— 発達研究, 15, 51-77.
- 河上誓作 (1996). 認知言語学の基礎 研究社出版.
- 北村智・佐々木裕一・河井大介 (2016). ツイッターの心理学—情報環境と利用者行動— 誠信書房.
- 小林雄一郎 (2017). Rによるやさしいテキストマイニング オーム社.
- Li, J., & Fischer, K. W. (2007). Respect as a positive self-conscious emotion in European Americans and Chinese. In J. L. Tracy, R. W. Robins, & J. P. Tangney (Eds.), *The self-conscious emotions: Theory and research*. New York: Guilford Press. 224-242.
- Martin, R. A., & Refcort, H. M. (1983). Sense of Humor as a Moderator of the Relation Between Stressors and Moods. *Journal of Personality & Social Psychology*, 45(6), 1313-1324.
- 中村弘・沢崎俊幸・吉川浩・近藤恭一・小池将貴・立花希尹子 (1970). 発想法 (KJ法) への数量化理論の適用, 経営科学, 14 (3), 168-183.
- 新名理恵・坂田成輝・矢富直美・本間昭 (1990). 心理的ストレス反応尺度の開発 心身医学, 30, 29-38.
- Nisbet, R. E. (2003). *The geography of thought: how Asians and Westerners think differently ... and why.* (ニスベット, R. E. 村本由紀子 (訳) (2004). 木を見る西洋人 森を見る東洋人—思考の違いはいかにして生まれるか— ダイヤモンド社)
- 尾関友佳子 (1990). 大学生のストレス自己評価尺度—質問紙構成と質問紙短縮について— 久留米大学院紀要, 比較文化研究, 1, 9-32.
- 尾関友佳子・原口雅浩・津田彰 (1991). 大学生の生活ストレス、コーピング、パーソナリティとストレス反応 健康心理学研究, 4 (2), 1-9.
- 坂本裕・一門恵子 (2013). 特別支援学校教員のバーンアウトへの関与要因についての探索的研究 特殊教育研究, 51, (3), 261-267.
- Shaver, P. R., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology*, 52, 1061-1086.
- 鈴木郁子・和田真雄・村上隆 (2005). KJ法および多重対応分析を用いた自由記述型応答の数量化 名古屋大学大学院教育発達科学研究科紀要, 心理発達科学, 52, 135-152.
- 田中博晃 (2012). KJ法入門—発想や仮説を得るには—, 竹内理・水本篤 (編), 外国語教育研究ハンドブック—研究手法のより良い理解のために— (pp258-284), 松柏社.
- 田中信一・大鹿譲 (1978). KJ法手順の統計数理学的解明—MDA-ORとクラスター分析の適用— 大阪工業大学紀要, 理工篇, 23 (1), p95-107.
- 対馬栄輝 (2010). 医療系研究論文の読み方・まとめ方—論文のPECOから正しい統計的判断まで—, 東京図書.
- 鷺見幸美 (2013). カテゴリー化とプロトタイプ, 森雄一・高橋英光 (編), 認知言語学—基礎から最前線へ— (pp27-52), くろしお出版.
- Wittgenstein, L. (1953). *Logisch-philosophische Abhandlung.* (ヴィトゲンシュタイン, L. 藤本隆志・坂井秀寿 (訳) (1968). 論理哲学論考 法政大学出版局).