

ツイッターに表出される対外国人感情 ——ハッシュタグで紐付けされた日本語ツイートの 量的テキスト分析から

大茂矢 由佳 筑波大学大学院博士後期課程

キーワード：ツイッター，対外国人感情，機械学習

本研究は、日本語ツイッターに表出される対外国人感情を量的テキスト分析にもとづいて解明することを目的としたものである。外国人を意味するハッシュタグとして「# 移民」，「# 外国人」，「# 外国人材」，「# 外国人労働者」，「# 技能実習生」，「# 難民」の6つを選定し，Twitter API を通じたツイートの収集を行なった。計 50,474 ツイートが収集され，ハッシュタグ毎に機械学習モデルによる感情分析，およびテキストマイニングを行なった。

分析の結果，もっとも拡散しやすいハッシュタグは「# 難民」であり，その理由として，難民の收容に対する抗議運動に多くのユーザーが共鳴し，活発な拡散行動に結びついていたことが指摘された。また，感情の方向性がハッシュタグの対象そのものではなく日本政府に向けられていた点は，「# 難民」でのみ観察された特徴である。他方，「# 外国人材」と「# 技能実習生」は外国人に対して好意的なツイートが多く，日本で働く外国人への法的サポートや，新型コロナウイルスの影響による生活困窮に関する内容が多数を占めていた。「# 移民」と「# 外国人労働者」では，停滞する賃金水準への怒りや嫌悪の矛先が日本で働く外国人に向けられており，多くの先行研究で否定されてきた経済不安仮説が，ネット空間では支持される可能性が示唆された。また，外国人関連ハッシュタグに「望ましくない他者」という意味づけがなされ，それがユーザー間で共有されていたことから，ツイッターがユーザーの排外意識の増幅に小さくない影響力を有していることが示唆された。

1 はじめに

非移民国家の日本において，外国籍人口の増加が顕在化している。とりわけ近年は，サービス業に従事する外国人が増えたことで，日常生活のなかで彼／彼女らの働く姿を見かける機会が格段に増加した。こうしたなかで，日本人の外国人に対する意識や感情は転換期を迎えている。

従来，対外国人意識に関する国内の研究は，質問紙調査を用いた統計的研究（渋谷ほか，2011；濱田，2008）や，マスメディアに映し出された外国人の姿からオーディエンスの認知を推察するメディア表象研究（船山，2008）が主流であった。しかし，前者の研究法は，研究者が用意した質問に回答

する形でしかデータを集められないことや、回答が社会的な望ましきバイアス (social desirability bias) の影響を受けやすいこと等、方法論的な課題が指摘されている (Krumpal, 2013)。また、後者のメディア表象研究の主要な関心は、メディアが作り出してきたイメージを分析することに向けられているため、オーディエンスの認知に与える影響については、あくまで推察に留まるという課題がある (竹下, 2008)。

こうした点を克服する方法として、本研究ではソーシャル・ネットワーキング・サービス (以下、「SNS」) における外国人関連の言説に着目したい。SNSユーザーは、プラットフォーム内に形成された巨大な情報共有ネットワークの構成員であり、そのなかで、情報の発信、受信、拡散を自由に行なっている。こうした双方向性の高いSNSの普及によって、オーディエンスはもはや情報を受け取るだけの受動的な存在ではなくなり、情報の発信者と受信者の境界はあいまいなものになったと指摘されている (西田, 2009)。とはいえ、古典的メディアにおけるオーディエンスが消失したわけではない。SNSユーザーは、情報の発信者であり、受信者であり、そしてそれをシェア機能によって他のユーザーに広げていく拡散者である。こうしたユーザーの複合性は、古典的メディアにおけるオーディエンスを消失させたのではなく、その性質を変化させたといえるだろう。このことから、SNS上の言説をオーディエンスの意識や感情の一部として解釈することも可能である。

無論、SNS上には政府機関や公的団体の公式アカウント等が提供する情報も含まれており、必ずしも個人ユーザーの発言のみが抽出されるわけではない。しかし、それでもなお、人々の非形式的な発言のデータベースともいえるSNSは、無修正の世論の一端として、分析対象とする意義は大きいといえる^{*1}。こうした背景から、本研究ではSNSにおける言説の分析から、日本人の対外国人感情の一端を明らかにし、日本を事例とした実証研究の蓄積に貢献することを目指す。

2 先行研究

レイシズムや排外主義に関する先行研究のなかで、排外的な意識の形成やその強化をもたらす要因が議論されてきた。主な仮説としては、経済不安仮説 (安田, 2012; 高原, 2006)、集団脅威仮説 (大槻, 2007; 永吉, 2008)、接触仮説 (眞住, 2015)、社会関係仮説 (宣ほか, 2017)、メディア影響仮説 (李, 2009; 濱田, 2013; 高, 2014, 2015a) 等が提唱され、日本を事例とした実証研究もなされている。

こうした仮説のなかでも、本研究ではとくにメディア影響仮説に着目する。メディア影響仮説とは、メディアを通じて間接的に得た情報によって、特定の集団に対する偏見や差別、排外意識が生み出されるという仮説である。先行研究において、民族的マイノリティが、犯罪やテロ等の治安の問題と結びつけて報じられやすいことが指摘されており (Dixon and Linz, 2000)、そうしたメディア情報によって、オーディエンスの排外意識が高まるのが日本においても報告されている (李, 2009)。

メディア影響仮説の理論的出自は、マスメディア効果研究における議題設定機能や培養効果にある。これらの研究領域では、直接的な経験機会が限定的な事柄について認知が形成される際に、その効果が強まることが指摘されてきた (竹下, 2008)。こうしたマスメディア効果研究の定説に依拠

するならば、外国籍人口割合が全人口の2%程度の日本では、外国人イメージの形成においてメディア情報の影響力が大きいことが推察される。

無論、日本の外国籍人口は決して少なくなく、ゆえに外国人との接触機会も豊富であるという反論もあろう。たしかに、技能実習生や留学生が事実上の移住労働者となっている現状に目を向ければ、その人口規模は欧州の移民国家にも匹敵する^{*2}。では、外国籍人口の増加が、すなわち外国人との接触機会の増加と考え得るだろうか。アメリカの黒人差別の研究で提唱されてきた回避的レイシズム (aversive racism) の理論によれば、差別は悪であるという規範をもっている人々であっても、潜在意識のレベルでは外集団に対するネガティブな感情を有しており、無意識的に内集団以外との接触を避けようとする傾向がある (Gaertner and Dovidio, 2005)。日本においても、外集団に対して非友好的な認知をもちやすいことが報告されており (唐沢, 2001; 河野, 2019), 対人接触回避が発生している可能性が指摘されている。したがって、外国籍人口の増加が著しい今日の日本においても、依然として日本人と外国人との接点は限定的であると考えられ、メディア情報から受ける影響は小さくないといえるだろう。

このメディア影響仮説は、メディア環境の急速な変化を背景に、仮説の再検証や新たな理論提唱がさかんに行なわれている。ネット空間ではマスメディアには見られなかった様々な現象が観察されており、こうした研究は、政治コミュニケーション (political communication) やコンピュータを介したコミュニケーション (Computer Mediated Communication: CMC) 等の領域で活発化している。有名な理論としては、リベラル層はよりリベラルに、保守層はより保守に傾倒していく政治的分極化 (Iyengar and Hahn, 2009) や、同じような意見をもつ人とだけコミュニケーションをとることで、その意見が強化され尖鋭化していくエコーチェンバー現象 (Sunstein, 2017) 等がある。また、こうした現象を生み出す要因として、自分にとって都合のよい情報にだけアクセスする選択的接触 (Sunstein, 2001, 2017) が提唱されている。これらの理論を検証した実証研究も多く、欧米ではSNS上の差別やヘイトスピーチに特化した研究も豊富である (Backfried and Shalunts, 2016; M. Ekman, 2019; Evolvi, 2018)。

日本におけるCMCと排外主義に関する研究でもっとも先駆的なものは、辻 (2008) であろう。同研究では、いわゆる「ネット右翼」と呼ばれる層は、現実社会でも署名活動や集会等の政治運動に参加する傾向があることを明らかにし、インターネットの登場によって、潜在的な右翼層が可視化されたと指摘している (辻, 2008)。Twitter (以下、ツイッター) を事例とした研究では、高 (2015a) が在日コリアンに対する差別的言説を分析している。この研究によれば、差別的なツイートはごく限られたアカウントによって発信されている一方、こうしたツイートは拡散しやすく、同じような意見をもつユーザーが大規模なネットワークを形成している可能性が示唆された (高, 2015a)。また、インターネットがこうした排外主義運動の安価な手段として積極的に活用されているとの指摘もある (樋口, 2014)。この他、日本語ツイッターにおけるネット右翼とエコーチェンバー現象に関する研究 (Takikawa and Nagayoshi, 2017) や、ネットメディア接触と排外的態度の二極化に関する研究 (辻・北村, 2018) 等、2010年代後半以降、日本においても関連研究が増加している。

また、必ずしもメディアとの関連で説明されるものではないが、外国や外国人に対するイメージ、

意識、感情に関する実証研究も多い。たとえば、河野（2019）は、1000人を対象としたインターネット調査で、外集団に対して抱く感情を7段階尺度で測定している。同研究では、嫌悪を感じる対象集団として、2位に「アジア3カ国」、3位に「イスラム国」、8位に「イスラム教徒」、13位に「在日」が挙げられ、ネガティブ感情（軽蔑・嫌悪・恐怖・怒り）のなかでも、とりわけ嫌悪の表出が高いことが報告されている（河野，2019）。この他、形容詞対を用いたイメージ評価の手法であるSD法（Semantic Differential Method）を用いた研究もなされている（沼田ほか，2011；渋谷ほか，2011）。しかし、同手法で用いられる形容詞対（「明るい：暗い」等）は、ポジティブ／ネガティブの二項対立であり、複数の感情への分類を可能にするものではない。

3 研究目的とリサーチクエスチョン

以上に概観してきたように、近年は日本のメディア研究や移民研究領域においても、CMCと排外意識の関係性が注目されている。しかし、実証的な研究の蓄積は未だ十分とはいえず、また、特定のエスニックグループを対象を限定した研究が多くを占める、より網羅的な研究が求められている。加えて、上述のとおり、外国人に対するイメージや感情に関する先行研究の多くは質問紙調査法を採用しており、新たな研究アプローチの必要性が高まっている。こうした課題に対し、本研究では外国人に関連するSNS上の言説を網羅的、かつ総合的に分析することによって、そこに表出されている対外国人感情の解明を試みる。

そのための方法として、本研究では日本のSNSのなかでもっとも利用率の高いツイッターに焦点をあて^{*3}、外国人関連のハッシュタグで紐付けされたツイートを分析対象とする。ハッシュタグとはSNS上で一般的に使用されているタグ機能であり、特定のキーワードやフレーズの前に「#」をつけることで投稿がタグ化される。この機能を活用することで、特定のキーワードに紐付けされたツイートを瞬時に検索・抽出することが可能になるほか、関心や意見、信条を共有する他のユーザーと繋がることも容易になる。日本のツイッターにおける外国人関連言説に詳しい高によれば、ハッシュタグの利用率は、中国関連ツイートで10.82%、コリアン関連ツイートで17.33%であり、必ずしも高いとはいえない^{*4}（高，2014，2015b）。しかし、ツイートの抽出にキーワードを用いた場合、本研究の射程にはまったく当てはまらないツイートが大量に抽出されてしまうという課題がある。たとえば、「難民」というキーワードで検索した場合、「買い物難民」や「ネットカフェ難民」等の造語も抽出され、ともすれば本来の意味での「難民」に言及しているツイートよりもデータ量が多くなる可能性が高い。こうした「外国人」とは無関係なツイートを抽出データから手作業で除外することは、作業上の困難をとまなうだけでなく、分類軸が揺らぐことによるデータの歪みや再現性の喪失等のリスクも増大する。一方で、ハッシュタグによる検索は、抽出データの純度を上げるための有効な手段であり、ヒトの手によるデータの操作が行なわれない分、再現性も担保される。ハッシュタグの検索網からこぼれ落ちるデータが一定程度あることを考慮しても、量的分析手法をとる本研究においては、ハッシュタグによるツイートの抽出が妥当な手段と考えられる。

次に、検索に用いるハッシュタグの選定であるが、日本においては「移民」という言葉が忌避さ

表1 収集されたツイート数, および固有ツイート数

	全ツイート	固有ツイート	固有ツイートの割合 (%)
# 移民	9,316	2,484	26.7
# 外国人	21,722	6,481	29.8
# 外国人材	732	461	63.0
# 外国人労働者	3,621	1,686	46.6
# 技能実習生	4,157	1,380	33.2
# 難民	10,926	1,549	14.2

出典：筆者作成

れる傾向が強く、たとえば政治的議論のなかでは「外国人材」や「外国人労働者」、「技能実習生」といった代替表現が頻繁に使用されている（大茂矢, 2020）。本研究では、これら4つの語のほか、より一般的な語である「外国人」、そして、性質が異なるものの、日本の外国籍人口の重要な構成員である「難民」を加えた計6つのハッシュタグを分析対象とする。

分析にあたり、本研究では以下の2つのリサーチクエスチョンを設定した。

RQ1：外国人関連ハッシュタグを含む日本語ツイートには、どのような感情が含まれているか。

RQ2：日本語ツイッターにおいて、外国人関連ハッシュタグはどのようなコンテキストで使用されているか。

4 分析

(1) データ

本研究では、2020年2月6日から2020年8月21日の約半年間を対象に、「# 移民」、「# 外国人」、「# 外国人材」、「# 外国人労働者」、「# 技能実習生」、「# 難民」のハッシュタグを含むツイート、およびリツイートを収集した。データの収集には、統計ソフトRのrtweetライブラリ（Kearney, 2019）を用いた。このライブラリを通じてTwitter API（Application Programming Interface）へのアクセス権限を承認することで、過去7日間のツイートを最大で18,000件収集することができる。本研究では、rtweetのsearch_tweets()関数を使用して、上述の6つのハッシュタグを含むツイートを収集した。表1は、このプロセスを経て収集されたツイート数である。表中の「全ツイート」には、リツイートや異なるアカウントによる同一内容のツイート等も含まれており、それらの重複を除外したツイート数を「固有ツイート」として示している^{*5}。

(2) 分析方法

分析の第一段階として、sentiment_ja^{*6}を用いた日本語ツイートの感情分析を行なった。sentiment_jaとは、日本語の短文に含まれる感情を分類するためのオープンソースの機械学習モデルである。このモデルのベースとなっているのは、Ekmanが提唱した6つの人間の普遍的感情、

すなわち、喜び、悲しみ、怒り、嫌悪、驚き、恐怖である (P. Ekman, 1993 ; P. Ekman and Keltner, 1997)。モデルの開発者である Shun Sugiyama は、それぞれの感情に対応すると考えられる絵文字を7つ特定し、それらの絵文字を含む日本語ツイートを各6万件収集した^{*7}。そして、ニューラルネットワークライブラリの Keras を用いて、収集されたツイートに含まれる単語と絵文字の共起パターンを学習させ、インプットされたテキストに対して感情スコアを算出する機械学習モデルを生成している。

一般的に、テキストデータの感情分析には、自然言語処理の分野で開発されてきた感情極性辞書が用いられている (徳久ほか, 2009 ; 高村ほか, 2006)。これらの辞書は、新聞記事や国会会議録のような「硬い」テキストに対しては高い精度を発揮し、優れた分析手法といえる。しかし、本研究であつかうツイッターデータには、日本語文法が不完全な発話や、いわゆる「ネットスラング」と呼ばれるようなSNS特有の言葉遣いが多数含まれている。こうした「軟らかい」テキストは従来の辞書では解釈が難しく、結果として分析の精度が低下する傾向がある。こうした理由から、本研究ではツイッターデータにもとづいて生成された sentiment_ja がより適した分析ツールであると判断し、上述の6ハッシュタグを含むツイートの感情分析を行なった。

分析の第二段階として、統計ソフト R の RMeCab ライブラリ (Ishida, 2020) を用いた形態素解析を行なった。日本語は単語ごとにスペースが空けられない膠着型の言語であるため、特定の語や品詞を抽出する前段階として、テキストを形態素 (意味のある最小の言語単位) に分割するという手順が必要になる。それを可能にする形態素解析エンジンとして 2004 年に MeCab が開発されており、本研究で用いた RMeCab は、同エンジンを R から呼び出すためのライブラリである。分析手順としては、まず、RMeCabFreq () 関数でツイートのテキストから名詞のみを抽出し、その出現頻度を算出した。分析結果は一旦エクセルファイルに書き出し、表記の揺らぎ (「わたし」と「私」等) やカタカナ語の形態素解析エラーの修正を行なった。

以上のテキストマイニング結果の図示には、R の wordcloud ライブラリ (Fellows, 2018) を用いた。ワードクラウドは、出現頻度の高い単語ほど大きく図の中央に表示されるため、分析結果を直感的にとらえるのに優れた可視化の手法である。本研究では、ハッシュタグ毎の頻出名詞上位 30 語をワードクラウドに描画した。なお、ワードクラウドの見易さを考慮し、ハッシュタグの語そのものは各テキストデータから除外している^{*8}。

(3) 分析結果

(a) 感情分析

図1は「# 移民」、「# 外国人」、「# 外国人材」、「# 外国人労働者」、「# 技能実習生」、「# 難民」のハッシュタグを含むツイートの感情分析の結果である。喜び (happiness) の割合がもっとも高かったのは「# 外国人」で 15.6%、次いで「# 外国人材」14.9%、「# 技能実習生」13.7%であった。悲しみ (sadness) の表出はすべてのハッシュタグで少ない傾向がみられ、3~7%程度であった。他方、怒り (anger) は全体的に高い割合で表出され、「# 外国人労働者」36.1%、「# 難民」30.0%、「# 移民」27.9%であった。嫌悪 (disgust) は「# 外国人」で 9.5%と低かったのに対し、「# 移民」

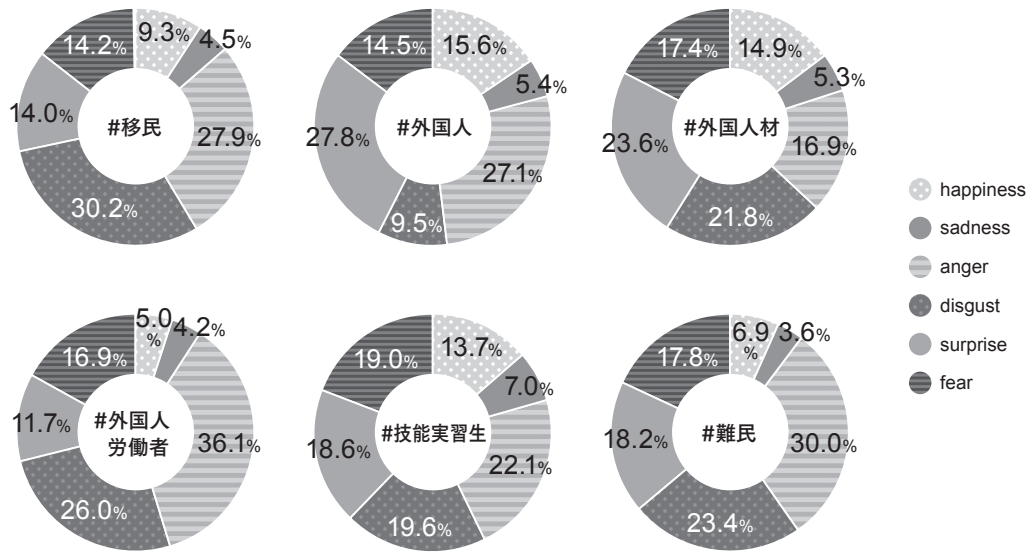


図1 各ハッシュタグを含むツイートの感情分析結果

出典：筆者作成

30.2%、「#外国人労働者」26.0%、「#難民」23.4%と、ハッシュタグによって大きなばらつきが確認された。驚き（surprise）も同様に、「#外国人労働者」の11.7%から「#外国人」の27.8%まで、表出の程度に大きな差があった。恐怖（fear）はすべてのハッシュタグで平均的に表出されており、もっとも低い「#移民」で14.2%、もっとも高い「#技能実習生」で19.0%であった^{*9}。

感情の分布に近似性が確認されたのは、「#移民」と「#外国人労働者」であった。また、怒りと嫌悪の割合が高いという点で、「#難民」もこれら2つのハッシュタグと類似傾向があるといえる。一方、「#外国人材」と「#技能実習生」には、感情の分布が比較的均等であるという共通点が確認された。驚きの割合がもっとも高い「#外国人」は、他のハッシュタグと比較しても特異な感情分布を示しているといえる。

(b) ワードクラウド

図2は「#移民」、「#外国人」、「#外国人材」、「#外国人労働者」、「#技能実習生」、「#難民」を含むツイートの最頻出名詞30語を描画したワードクラウドである。フォントサイズが大きいくほど、出現頻度が高い語であることを示している。また、図中のmfとは、ワードクラウドに描画されている語の最小出現頻度（minimum frequency）である。

まず、「#移民」のツイートにおける頻出語には、「天皇」や「左翼」等、イデオロギーに関連するものが多いことが明らかとなった。また、「政治」や「財政」といった頻出語からも、政治に関連したツイートが多いことが読み取れる。他方、4番目に出現頻度が高い「ライキン」は、ライズオブキングダムというゲーム名の略称であり、本研究が念頭に置く「移民」（現実社会で移住する人々）とは異なる意味で使用されているケースが多いことが明らかとなった^{*10}。

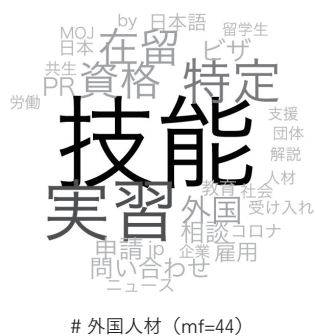
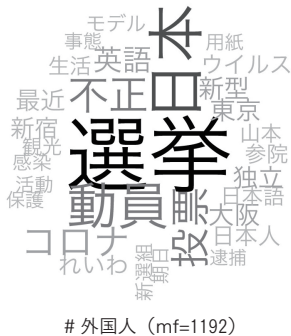


図2 各ハッシュタグを含むツイートのワードクラウド

出典：筆者作成

「# 移民」と同様に、「# 外国人」も政治的なツイートに使用される頻度が高いといえる。関連する主な頻出語は、「選挙」, 「投票」, 「参院」, 「れいわ/新選組」等であった。一方で、「生活」や「観光」等の語は、「新型/コロナ/ウイルス」に関するツイートが多かった結果であるといえる。

つぎに「# 外国人材」は、「特定/技能」や「技能/実習」, 「在留/資格」がワードクラウドの広範囲を占めている。「相談」や「問い合わせ」, 「解説」が出現していることから、日本で働く外国人材を支援する組織や団体によるツイートが中心であるといえる。また、「MOJ」 (=Ministry of Justice) という頻出語からは、政府発信のニュースのなかで「# 外国人材」が積極的に使用されていると解釈できよう。

同様に、「# 外国人労働者」においても「特定/技能」や「技能/実習」が頻出語である。ただし、「雇用」や「労働」, 「解雇」が出現していることから、在留資格等の法的地位に焦点が当てられていた「# 外国人材」とは異なるコンテキストで「# 外国人労働者」が用いられていることが示唆される。「解雇」は、「新型/コロナ/ウイルス」と深い関連性があると推察される。

この「新型/コロナ/ウイルス」による影響は、「# 技能実習生」でもっとも顕著に表れている。「生活」, 「特別/定額/給付(金)」, 「貧困」等の語からは、技能実習生の困窮した生活状況を伝えるツイートが多いことが示唆される。また、こうした問題にメディアも強い関心を寄せていることが、「news」や「密着/取材」等の頻出語から読み取れる。

最後に「# 難民」は、日本の外国人収容問題に対する抗議運動において、さかんに使用されてい

ることが明らかとなった。「#FREEUSHIKU」は、東日本入国管理センター（茨城県牛久市）への抗議運動を行なう団体のツイッターアカウント名であり、キャンペーンに用いられるハッシュタグでもある。

5 考察

本研究では、日本語ツイッターにおける対外国人感情を解明するために、「# 移民」、「# 外国人」、「# 外国人材」、「# 外国人労働者」、「# 技能実習生」、「# 難民」のハッシュタグを含むツイートを収集し、機械学習モデルによる感情分析、およびテキストマイニングを実施した。まず、各ハッシュタグの使用状況であるが、6つのハッシュタグのうち「# 外国人」の使用頻度がもっとも高く、21,722 件のツイートで使用されていた（表1）。次いで多かったのは、「# 難民」の10,926 件、「# 移民」の9,316 件であった。固有ツイートの割合がもっとも高いのは「# 外国人材」の63%であり、全ハッシュタグのなかで唯一50%を超えていた。固有ツイートの割合が高いほど、そのハッシュタグは「発信者」（最初のツイートを発信する人）によって使用されやすく、反対に、その割合が低ければ、「拡散者」（既存のツイートをリツイート機能等によって拡散させていく人）による二次的な使用が多いことが示唆される。したがって、「# 外国人材」は、「発信者」によって使用されることが多い反面、ツイートに対する他のユーザーの注目度は低く、「拡散者」のリツイート行動には結びついていないことが指摘できる。また、固有ツイートの割合が46.6%であった「# 外国人労働者」も、同様の傾向がみられたといえよう。一方、もっとも割合が低いのは「# 難民」の14.2%であり、平均すると1つのツイートが7回以上もリツイートされている計算になる。ワードクラウド（図2）から読み取れたように、「# 難民」は外国人収容問題に対する抗議運動でさかんに使用されており、こうした運動に共鳴するユーザーによって、広く拡散していったと考えられる。

次に、第3節で設定した2つのリサーチクエスチョンについて検討する。まず、sentiment_jaによる感情分析の結果からは、「# 移民」と「# 外国人労働者」、「# 外国人材」と「# 技能実習生」で、それぞれ感情の分布に近似性が確認された（図1）。前者の2ハッシュタグ（「# 移民」と「# 外国人労働者」）は、とくに嫌悪の感情が多かったことが特徴として挙げられ、怒りと嫌悪を合わせると、実に全体の60%程度を占める結果となった。その要因として示唆的であるのは、図2のワードクラウドである。「# 移民」と「# 外国人労働者」のワードクラウドを見てみると、「雇用」、「賃金」、「所得」等が頻出語であり、労働に関連したツイートが多いことが読み取れる。これは実質賃金が上がらない原因として、外国からの安価な労働力が増えていることを主張するツイートが多数抽出された結果である^{*11}。現代日本を事例とした先行研究では、経済的地位と排外意識との関連性は弱いとされ、経済不安仮説は否定されている（樋口, 2014; 永吉, 2014, 2015）。しかし、本研究の分析結果によれば、停滞する賃金水準への怒りや嫌悪の矛先が、日本で働く外国人に向けられていた。したがって、ツイッターという脱個人化（Reicher et al., 1995）が発生しやすい空間においては、経済不安仮説が支持される可能性が指摘できる。

一方、後者の2ハッシュタグ（「# 外国人材」と「# 技能実習生」）は、感情の分布が比較的均等で

あった(図1)。悲しみの表出は少ない傾向がみられたものの、他の5つの感情はほぼ同程度に表出されていた。テキストマイニングの結果からは、日本で働く外国人への法的サポートや、「新型/コロナ/ウイルス」の影響による生活困窮に関連する語が頻繁に出現していた(図2)。以上の分析結果から、「# 移民」と「# 外国人労働者」の2ハッシュタグと比較すると、「# 外国人材」と「# 技能実習生」を含むツイートは、外国人に対して好意的な内容が多かったといえよう。

また、驚きの割合がもっとも高い「# 外国人」は、他のハッシュタグと比較しても特異な感情分布であるといえる(図1)。政治的な語が多いというテキストマイニングの結果は「# 移民」との共通点であるが、「# 外国人」を含むツイートには、「外国人を動員した不正選挙」等の陰謀説ともとらえ得る内容が多く(図2)、これが感情分析で驚きの割合が高くなった要因であると推察される。

最後に「# 難民」は、極めて特異な性質をもっていた。図1をみると、「# 難民」を含むツイートの感情分布は「# 移民」や「# 外国人労働者」と類似傾向を示している。しかし、テキストマイニングを行なった結果、難民の収容問題に対する抗議ツイートがワードクラウドのほぼ全域を占めていることが明らかとなった(図2)。すなわち、「# 難民」における怒りの感情は、難民に対して向けられているのではなく、日本の難民認定制度や出入国在留管理庁に向けられているといえる。これは、ハッシュタグの対象そのものに感情の矛先が向けられていた「# 移民」や「# 外国人労働者」とは、まったく異なる性質である。こうした特殊性の背景にあるのは、第一に、公益社団法人アムネスティ・インターナショナル日本が、「外国人の長期収容に終止符を！」というオンライン署名活動^{*12}を実施し、公式アカウント(@amnesty_or_jp)を通じた広報活動を行っていたこと、第二に、ジャーナリストの志葉玲氏が自身のアカウント(@reishiva)で抗議ツイートを発信しており、それらが複数のユーザーによって繰り返しリツイートされていたことが影響している。

以上に述べてきたように、本研究では量的テキスト分析の結果にもとづきながら、日本語ツイッターユーザーの対外国人感情とその背景にあるコンテキストを考察した。これまでの排外意識研究の多くが「我々」と「彼ら」の境界に着目しているのに対して、本研究では「彼ら」のなかに含まれる多様な他者に着目し、「# 移民」、「# 外国人」、「# 外国人材」、「# 外国人労働者」、「# 技能実習生」、「# 難民」という6つのハッシュタグを手掛かりに分析を試みた。こうした語を「ラベル」と呼び、それぞれの意味づけの違いについて調査した鳥越によれば、「外国人」はヨーロッパ系白人の「望ましい他者」である一方、「移民」や「外国人労働者」は治安悪化や犯罪の原因となる「望ましくない他者」として語られる傾向があるという(鳥越, 2019)。一方、本研究の分析では、「外国人」は政治的なコンテキストに引きつけて語られる傾向が強く、「移民」と類似の意味づけがなされることが判明した。すなわち、鳥越(2019)で使用された3つのラベルは、ツイッターにおいてはおおむね排外的な言説に使用されており、すべてに「望ましくない他者」としての意味づけがなされていたといえる。

こうした知見の違いは、研究手法の違いによって説明が可能であろう。鳥越(2019)が行なった質問紙調査では、調査協力者の回答に社会的な望ましきバイアスが働きやすいことが指摘されている(Krumpal, 2013)。同調査では、「外国人」は「望ましい他者」、「移民」と「外国人労働者」は「望ましくない他者」に二分されるという分析結果が得られているが、これは提示された全ラベルを

「望ましくない他者」とする極端な回答を、調査協力者が無意識的に回避した結果である可能性が指摘できる。実際、そうしたバイアスが掛からないツイッターにおいては、すべてのラベルに排外的な意味づけがなされていることが本研究の分析結果から示された。こうした、いわゆる「本音」を解明することは従来の研究手法では難しく、非形式的な発言の集合であるツイッターを分析対象とした本研究の意義がここにも認められよう。

社会的な望ましきバイアスの不在に加えて、匿名性が高く、非言語コミュニケーションが欠如したCMCでは、脱個人化が起きやすいことも知られている (Reicher et al., 1995)。日本においては、インターネット上で排外主義的な発言を行なう「ネット右翼」は、インターネット利用者の1.3%にすぎず (辻, 2008)、ツイッターにおいても同様のことが報告されている (Takikawa and Nagayoshi, 2017)。こうしたことから、本研究のデータに含まれている過激な排外主義的ツイートは、一部のユーザーによって繰り返し発信されたツイートである可能性が高い。しかし、SNSはその名称に冠しているように、手軽で安価なネットワーキングの手段である。たとえ、排外主義的な主張がごく一部のユーザーから発信されたものであっても、そこに多くのユーザーがつながることで容易に拡散していくという特徴を有している (永吉, 2017; 高, 2015a)。こうしたSNSの特徴を踏まえれば、外国人関連ハッシュタグに「望ましくない他者」という意味づけがなされ、それがユーザー間で共有されていることは、ユーザーの排外意識の増幅に小さくない影響を与えていると考えられる。

また、「#外国人材」と「#技能実習生」は外国人に対して好意的であり、「#移民」と「#外国人労働者」は排外的であるという本研究の分析結果は、興味深いことに2018年の入管法改正をめぐる国会審議を分析した大茂矢 (2020) の知見と共通している。同研究では、与党議員と野党議員の発言を個別に分析し、外国からの労働者を指し示す語として、前者は「外国人材」を、後者は「移民」や「外国人労働者」を積極的に使用していることを明らかにしている (大茂矢, 2020)。外国人の受入れ拡大を目指す与党は好意的な語を、入管法の改正に異論を唱える野党は非好意的な語を選択的に使用していたという指摘は、本研究の分析結果からも裏付けられたといえよう。無論、一般市民がSNS上で使用する語の含意が政治の場でもそのまま引用されているのか、あるいは、政治的議論のなかでなされた意味づけが一般市民にも浸透していったのかを解明することは困難である。しかし、政治的意見が分極化しやすい「移民」というテーマにおいて、国会議員が審議のなかで使用する語と一般市民がSNS上で使用する語に、共通の含意があるという点は新たな発見といえるだろう。

6 おわりに

本研究では、日本語ツイッターにおける対外国人感情を解明することを目的に、「#移民」、「#外国人」、「#外国人材」、「#外国人労働者」、「#技能実習生」、「#難民」のハッシュタグを含むツイートの分析を行なった。その結果、「#外国人材」と「#技能実習生」は外国人に対して好意的な意味づけがなされている一方、「#移民」と「#外国人労働者」は排外的な意味づけが際立っていた。こうした分析結果は、先行研究の知見とは異なるものであり、SNSという匿名性の高い空間で発生す

る特徴的な現象として解釈することができた。また、ツイッターで使用される語と国会審議で使用される語に共通の含意があったという点も、興味深い発見であった。また、これまでの社会科学分野ではあまり活用されてこなかった機械学習やテキストマイニング等の量的分析手法を採用することで、当該研究領域に新たな研究アプローチを導入するとともに、研究の発展性を示すことができたといえよう。

本研究の限界としては、第一にSNSを分析対象としたことで、データそのものに世代的な偏りがある点が挙げられる。SNS利用率は年代間の差が大きく、高齢層ほど利用率が低い傾向がある。このことから、本研究が明らかにしてきたことは、日本語でツイッターを利用している人々の対外国人感情であり、日本人全体の傾向として一般化することはできない。第二に、データ収集上の限界が指摘できる。本研究では、上述の技術的理由から、ハッシュタグによるデータ抽出を行なった。そのため、「ハッシュタグ付きでツイートするユーザー」というセレクションバイアスが生じており、ツイッターの世論を代表しているわけではない。また、データの収集に用いた rtweet ライブラリは、過去7日間のデータにしかアクセスすることができない。そのため、筆者は毎週決められた曜日にデータ収集を実施することで、2020年2月6日から同年8月21日までの間に発信されたすべてのツイートを収集できるように努めた。しかし、Twitter APIの制約等により、アクセスが不可能だった時間帯も存在したため、本研究で使用したデータには、ごく僅かなツイートの収集漏れがある。

また、本研究で用いたデータが新型コロナウイルス感染症（COVID-19）の影響を強く受けている点にも留意が必要である。データ収集の開始時期は、日本で同感染症への危機感が高まった時期と一致しており、本稿執筆時点の2020年12月においても極めて高い状況が続いている。そのため、本研究で使用したデータ、およびその分析結果には、感染症への脅威やリスク認知のバイアスが掛かっている可能性が高い。こうした点については、同感染症の拡大前のデータやパンデミック収束後のデータと比較分析することで、コロナ禍やポストコロナ期における対外国人感情の特徴や変遷を明らかにすることができると思われる。

※ 本論文は、移民政策学会2020年度年次大会（2020年5月）での報告内容をベースに、分析手法の見直し、およびより長期的なデータを加えて再分析を行なった結果をまとめたものである。

- *1 公式アカウントから発信された情報を個人アカウントがリツイートしている場合、それは（共感／反感の何れにせよ）ユーザーの感情に触れたからこそ誘発された拡散行動であり、したがって広義には「世論の一端」と捉え得ると考えられる。
- *2 経済協力開発機構（OECD）の報告によると、日本への外国人移住者数は、2016年時点で、ドイツ、アメリカ、イギリスに次ぐ第4位となっている（OECD, 2018）。
- *3 ICT総研の調査によれば、2020年7月現在のSNS利用率は、上からLINE（以下、ライン）（77.4%）、ツイッター（38.5%）、Instagram（インスタグラム）（35.7%）である。ただし、インスタント・メッセージング・アプリであるラインは、主として個人間や小規模グループの連絡手段として利用されており、情報の公開性や拡散性は低い。そのため、SNSにおける集合感情や拡散行動の分析において、ラインは一般的に分析対象とされない。こうしたことから、SNSにおける集合感情を分析する素材として、ツイッターには一定の代表性が認められ

る（ICT総研ホームページ「2020年度 SNS利用動向に関する調査」<https://ictr.co.jp/report/20200729.html>, 2020年12月8日アクセス）。

- *4 とはいえ、日本語でハッシュタグ機能が利用できるようになったのが2011年7月であることを踏まえると、高（2014, 2015a）の調査時点と現在では、ハッシュタグの利用状況や普及度は大きく変化している可能性も指摘できる。
- *5 重複ツイートが複数収集される理由として、(1)リツイート、(2)異なるアカウントによる同一内容のツイート、(3)一つのアカウントが同一内容を繰り返しツイートしている場合が考えられる。
- *6 同モデルは以下のリンクから入手可能（https://github.com/sugiyamath/sentiment_ja, 2020年2月29日アクセス）。なお、sentimen_jaにはバージョン1とバージョン2があるが、本研究ではより精度が高いバージョン2を使用した。
- *7 嫌悪には、2種類の絵文字が割り当てられている。
- *8 たとえば「#移民」を含むツイートの場合、出現頻度をもっとも高い名詞は当然ながら「移民」となる。その出現回数が突出して多い場合、ワードクラウドの広範囲を「移民」の文字が独占してしまうため、他の名詞の重要度が読み取りにくくなる。各ハッシュタグのワードクラウドを描画するにあたってハッシュタグの語そのものを除外した理由は、こうした問題を回避するためである。
- *9 以下は、各感情の典型的、かつ広く拡散されているツイートの具体例である（「/」は改行を意味。URLリンクは省略）。

喜 び：「ベトナム技能実習生は15名になりましたが、環境を整える事とコミュニケーションに特に力を入れています。#首都圏技術交流協同組合 #技能実習生 #ベトナム語大事」

悲しみ：「帰りたいけど帰れない」名古屋の寺に続々とベトナム人が... コロナ禍で窮地の外国人に救いの手（中京テレビNEWS）／こういう場所がもっとあったら良いね。／日本政府はこういうこと政府責任でやらないんだね。子ども食堂にしても、全部善意ある民間人任せね／#外国人労働者」

怒 り：「広島に集まった左翼の集会では耳を疑うスピーチをしていたらしい。『日本を移民で埋める。そうすれば日本人の天皇が全ての外国人の象徴になれなくなる。その時、天皇制を廃止に追い込む』（続き→）#天皇 #左翼 #移民」

嫌 悪：「日本で無くなったと思われた梅毒や性感染症が再び広がっているという。汚い所で免疫がついた海外のウイルスや虫は耐性が強いので駆除が困難だ。おもてなしの心なんて言っていたら命の危険にもなる。／#移民 #おもてなし」

驚 き：「女性への性暴力がこの世のものとは思えない程、凄まじいコンゴから、必死で逃げてきた #難民 女性を、公務員である #東京入管 職員たちが裸同然にしてビデオ撮影、皆で『鑑賞会』。これが、現在、日本で起きていること。」

恐 怖：「コロナショックで大リストラ時代に突入？／罪もないのにリストラされた労働者の立場は？／無制限に入れまくった移民は？／奴等が考えるわけないか！／『コマンド通り』にやっただけだもん！／#この国は植民地／#リストラ／#移民／#ショックドクトリン」
- *10 「#移民」がゲーム名とともに使用されている理由は、ゲーム内でサーバーを移動することを俗に「移民」と呼ぶためである。ワードクラウドに描画されている「同盟」や「戦力」、「募集」、「鯖」も同様のコンテキスト内で使用されている。
- *11 ツイートの具体例としては、「財政出動をしないで、外国人労働者を入れて……／これじゃ実質賃金は永久に上がりませんよね。／名目賃金はあがるけど「実質賃金」はあがらないということは物価の上昇に賃金の上昇が追い付いていないということです。／#高度人材 #移民 #政治 #経済 #リフレ派 #リフレ」や「日本の労働力不足を #外国人労働者 移民で埋めるな。日本人の賃金が下がります。#国家は利潤を追求する企業では無い。日本国民の為の日本国です。賃金が適正水準になります。」等がある。
- *12 アムネスティ・インターナショナル日本ホームページ「外国人の長期収容に終止符を！」（https://www.amnesty.or.jp/get-involved/action/jp_202001.html, 2020年8月29日アクセス）

《参考文献》

- 李容玲, 2009 「日本人と外国人の共生を促す決定要因について——JGSS-2005 データに反映する制度と意識の

- 相関性」『日本版総合的社会調査共同研究拠点研究論文集』8 〈JGSS Research Series 5〉, 121~140 頁
- 大槻茂実, 2007 「外国人の増加に対する日本人の見解——JGSS-2003 データを用いて」『社会学論考』28 号, 1~25 頁
 - 大茂自由佳, 2020 「〈研究ノート〉日本は『移民』のタブーを克服したか——2018 年の入管法改正をめぐる国会審議の定量分析」『国際日本研究』12 号, 172~183 頁
 - 唐沢 穰, 2001 「集団の認知とステレオタイプ」唐沢 穰・池上知子・唐沢かおり・大平英樹編著『社会的認知の心理学——社会を描く心のはたらき』ナカニシヤ出版, 105~127 頁
 - 河野和明, 2019 「排斥の適応論と現代日本人の嫌悪対象集団について」『エモーション・スタディーズ』4 号 (Special Issue), 54~64 頁
 - 渋谷明子・テーシャオープン・李光鎬・上瀬由美子・萩原 滋・小城英子, 2011 「メディア接触と異文化経験と外国・外国人イメージ——ウェブ・モニター調査 (2010 年 2 月) の報告 (2)」『メディア・コミュニケーション: 慶応義塾大学メディア・コミュニケーション研究所紀要』62 号, 103~125 頁
 - 宣元錫・武田里子・山本直子・竹ノ下弘久, 2017 「多文化受容性に関する日韓比較調査研究」渡戸一郎・塩原良和・長谷部美佳・明石純一・宣元錫編著『変容する国際移住のリアリティ——「編入モード」の社会学』ハーベスト社, 220~258 頁
 - 高 史明, 2014 「日本語 Twitter ユーザーのコリアンについての言説の計量的分析」『人文研究』183 号, 131~153 頁
 - 高 史明, 2015a 『レイシズムを解剖する——在日コリアンへの偏見とインターネット』勁草書房
 - 高 史明, 2015b 「日本語 Twitter ユーザーの中国人についての言説の計量的分析——コリアンについての言説との比較」『人文学研究所報』53 号, 73~86 頁
 - 高原基彰, 2006 『不安型ナショナリズムの時代——日韓中のネット世代が憎みあう本当の理由』洋泉社
 - 高村大也・乾 孝司・奥村 学, 2006 「スピンモデルによる単語の感情極性抽出」『情報処理学会論文誌』47 巻 2 号, 627~637 頁
 - 竹下俊郎, 2008 『増補版 メディアの議題設定機能——マスコミ効果研究における理論と実証』学文社
 - 辻 大介, 2008 『インターネットにおける「右傾化」現象に関する実証研究——調査結果概要報告書』(<http://www.d-tsuji.com/paper/r04/>, 2020 年 8 月 29 日アクセス)
 - 辻 大介・北村 智, 2018 「インターネットでのニュース接触と排外主義的態度の極性化」『情報通信学会誌』36 巻 2 号, 99~109 頁
 - 徳久良子・乾健太郎・松本裕治, 2009 「Web から獲得した感情生起要因コーパスに基づく感情推定」『情報処理学会論文誌』50 巻 4 号, 1365~1374 頁
 - 鳥越千絵, 2019 「『外国人』『移民』『外国人労働者』——日本における移民ディスコースが構築する人種の階層」『西南学院大学英語英文学論集』59 巻 3 号, 155~179 頁
 - 永吉希久子, 2008 「排外意識に対する接触と脅威認知の効果——JGSS-2003 の分析から」『日本版総合的社会調査共同研究拠点研究論文集』7 〈JGSS Research Series 4〉, 259~270 頁
 - 永吉希久子, 2014 「外国籍者への権利付与意識の規定構造——潜在クラス分析を用いたアプローチ」『理論と方法』29 巻 2 号, 343~359 頁
 - 永吉希久子, 2015 「排外意識に対する社会経済的地位の効果——権威主義的態度の媒介効果の検証」筒井淳也・神林博史・長松奈美江・渡邊大輔・藤原 翔編著『計量社会学入門——社会をデータで読む』世界思想社, 208~219 頁
 - 永吉希久子, 2017 「日本の排外意識に関する研究動向と今後の展開可能性」『東北大学文学研究科研究年報』66 号, 143~164 頁
 - 西田善行, 2009 「『視聴者の反応』を分析する——インターネットから見るオーディエンス論」藤田真文・岡井崇之編『プロセスが見えるメディア分析入門——コンテンツから日常を問い直す』世界思想社, 145~169 頁
 - 沼田 潤・井上智義・朱 虹, 2011 「ネガティブなイメージを持たれている外国に対する偏見を低減する写真提示の効果」『人間環境学研究』9 巻 2 号, 53~62 頁
 - 濱田国佑, 2008 「外国人住民に対する日本人住民意識の変遷とその規定要因」『社会学評論』59 巻 1 号, 216~231 頁

- 濱田国佑, 2013 「在日ブラジル人の『社会問題』化と排外意識」 駒井 洋監修・小林真生編『レイシズムと外国人嫌悪』明石書店, 166~181 頁
- 樋口直人, 2014 『日本型排外主義——在特会・外国人参政権・東アジア地政学』名古屋大学出版会
- 船山和泉, 2008 「犯罪『行為者』としての外国人の表象に関する一考察——メディア・フレームの中の『窃盗団』」 『多文化関係学』5号, 17~31 頁
- 眞住優助, 2015 「少子高齢化時代の日本における外国人労働者の受け入れ意識を規定する要因——JGSS-2008 を用いた分析」 『日本版総合的社会調査共同研究拠点研究論文集』15 (JGSS Research Series 12), 51~61 頁
- 安田浩一, 2012 『ネットと愛国——在特会の「闇」を追いかけて』講談社
- Backfried, G. and Shalunts, G., 2016, Sentiment Analysis of Media in German on the Refugee Crisis in Europe, in Díaz P., Bellamine Ben Saoud N., Dugdale J. and Hanachi C. (eds.), *Information Systems for Crisis Response and Management in Mediterranean Countries. ISCRAM-med 2016. Lecture Notes in Business Information Processing*, vol.265, pp.234-241.
- Dixon, T. L. and Linz, D., 2000, Overrepresentation and Underrepresentation of African Americans and Latinos as Lawbreakers on Television News, *Journal of Communication* 50(2), pp.131-154.
- Ekman, M., 2019, Anti-immigration and racist discourse in social media, *European Journal of Communication* 34(6), pp.606-618.
- Ekman, P., 1993, Facial Expression and Emotion, *American Psychologist* 48(4), pp.384-392.
- Ekman, P. and Keltner, D., 1997, Universal Facial Expressions of Emotion: An Old Controversy and New Findings, in U. Segerstråle and P. Molnár (eds), *Nonverbal Communication: Where Nature Meets Culture*, Lawrence Erlbaum Associates, pp.27-46.
- Evolvi, G., 2018, Hate in a tweet: Exploring internet-based islamophobic discourses, *Religions* 9(10), pp.1-14.
- Fellows, I., 2018, wordcloud: Word Clouds (R package version 2.6.), <https://cran.r-project.org/package=wordcloud>
- Gaertner, S. L. and Dovidio, J. F., 2005, Understanding and Addressing Contemporary Racism: From Aversive Racism to the Common Ingroup Identity Model, *Journal of Social Issues* 61(3), pp.615-639.
- Ishida, M., 2020, RMeCab: interface to MeCab (R package version 1.05), <https://github.com/IshidaMotohiro/RMeCab>
- Iyengar, S. and Hahn, K. S., 2009, Red Media, Blue Media: Evidence of Ideological Selectivity in Media Use, *Journal of Communication* 59(1), pp.19-39.
- Kearney, W. M., 2019, rtweet: Collecting and analyzing Twitter data, *Journal of Open Source Software* 4(42), p.1829.
- Krumpal, I., 2013, Determinants of social desirability bias in sensitive surveys: A literature review, *Quality and Quantity* 47(4), pp.2025-2047.
- OECD., 2018, *International Migration Outlook 2018*, OECD Publishing.
- Reicher, S. D., Spears, R. and Postmes, T., 1995, A Social Identity Model of Deindividuation Phenomena, *European Review of Social Psychology* 6(1), pp.161-198.
- Sunstein, C. R., 2001, *Republic.com*, Princeton University Press.
- Sunstein, C. R., 2017, *#Republic: Divided Democracy in the Age of Social Media (NED-New)*, Princeton University Press.
- Takikawa, H. and Nagayoshi, K., 2017, Political polarization in social media: Analysis of the “Twitter political field” in Japan, *Proceedings - 2017 IEEE International Conference on Big Data, Big Data 2017*, pp.3143-3150.

Sentiment Towards “Foreigners” Among Japanese Twitter Users: A Quantitative Text Analysis of Tweets with Foreign-related Hashtags

OMOYA Yuka

University of Tsukuba

Key Words: Twitter, sentiment toward foreigners, machine learning

The aim of this study is to analyze the attitude and sentiment towards “foreigners” among Japanese Twitter users. The following six hashtags were selected for the analysis: #immigrant, #foreigner, #foreigntalent, #foreignworker, #technicaltrainee, and #refugee. Data was collected through the Twitter API and 50,474 tweets were retrieved. Sentiment analysis and text mining were conducted for all six hashtag related tweets.

From the analysis results, the hashtags #foreigntalent and #technicaltrainee were found to be used in the tweets favorable towards non-Japanese such as legal support for non-Japanese and financial difficulties due to COVID-19. In contrast, direct and indirect anger towards non-Japanese working in Japan were expressed in tweets that included #immigrant and #foreignworker, complaining that cheap labor from abroad has caused stagnation of wage levels in Japan. The economic uncertainty hypothesis has not been supported in previous studies regarding xenophobia, however, the analysis results of this paper suggests that this hypothesis may be applied to explain the anti-foreigner attitude in cyberspace in Japan. In addition, expressions such as “undesirable others” with negative connotation were observed among Japanese Twitter users in this study. Thus, one can argue that Twitter may have potential impact on the rise of xenophobia among its users in Japan. In contrast, tweets with #refugee are frequently used in the petition against the detention of asylum seekers. These tweets were commonly retweeted among those who were in concert of the campaign, and therefore, anger was not directed at refugees, but rather the Japanese government.