

A Quantitative Analysis of Japanese Twitter Users' Public Opinions of Koreans

TAKA Fumiaki

Keywords: Ethnic Prejudice, Social Media, Internet Right Wings,
Twitter, Quantitative Text Analysis

Abstract

This study gathered Japanese tweets (statements made on Twitter) about Koreans and conducted a quantitative analysis on them. Most of the statements regarding Koreans were negative, and a few hatemongers produced large numbers of tweets that gained them many followers. The tweets about Koreans tended to be accompanied by intentions to disseminate them. There were many cases where dissemination took place. In the light of several characteristics of the statements, quantitative text analysis was conducted to uncover evidence of both old-fashioned and modern racism (McConahay, 1986; Taka, 2013; Taka & Amemiya, 2013). We further examined users' distrust of mass media, their reliance on the anonymous BBS site "2channel" and its curator sites ("2channel-matome-blogs"), and their beliefs about the actual "truth" being hidden. In addition, we explored accusations against "anti-Japan" agents and users' intentions to either diffuse or share information about Koreans.

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¹⁾

高 史 明²⁾

本研究は、Twitter における日本語でのコリアンについての発言を収集し、計量的な分析を行った。コリアンについての発言の多くはネガティブなものであり、少数の差別的なユーザーが多くの投稿を行い、多くのフォロワーを獲得していた。コリアンについての発言は“拡散”を意図してなされることが多く、実際“拡散”されやすかった。また、古典的/現代的レイシズム (McConahay, 1986; 高, 2013; 高・雨宮, 2013) との関連、マスメディアへの不信、匿名掲示板“2ちゃんねる”と“2ちゃんねるまとめブログ”への傾倒、“真実”が隠されているという信念、何らかの“反日的な”ものに対する非難、情報を拡散ないし共有しようという意図などに注目して計量テキスト分析を行った。

キーワード：民族偏見、ソーシャル・メディア、ネット右翼、ツイッター、計量テキスト分析

1) 本研究は、東京大学大学院人文社会系研究科博士課程に提出した学位申請論文『在日コリアンに対するレイシズムの研究：現代的レイシズム理論に着目して』の研究1を再編集したものである。また、日本グループ・ダイナミクス学会第60回大会において、『Twitterにおけるコリアンについての言説の分析(2)』として発表したものである。なお、本研究の分析に際し、2013年度に龍谷大学国際社会文化研究所からの研究助成を受けた。

2) 研究実施時の所属は、東京大学大学院人文社会系研究科。著者への連絡用 E-mail アドレスは、FumiakiTaka@gmail.com。

序論

近年の日本では、コリアン³⁾や中国人に対する偏見・差別が表面化し、重要な社会問題となっている。例えば書店には、韓国・朝鮮や中国を侮蔑し中傷する書籍を置くコーナーが登場するようになっている（守，2014）。また、インターネット上では、特に在日コリアンに対する差別的言説が相当数見られ、中にはインターネット上で仲間を集め、排外主義のデモや示威行動を行う団体も登場している（安田，2012）。2013年には、ユーキャン新語・流行語大賞のトップテンに“ヘイトスピーチ”が入選し（自由国民社，2013）、警察庁の発行する“治安の回顧と展望”にも、“右派系市民グループをめぐる動向”の項目が初めて立てられた（警察庁警備局，2013）。

こうした状況を理解し解消の糸口を探るため、本研究では差別的な言説の主なターゲットとなっているコリアンに注目し、彼らがソーシャル・メディアでどのように語られているのかを分析する。

ソーシャル・メディアでの言説を取り上げて分析するのは、ソーシャル・メディアが2000年代のレイシズム⁴⁾の伸張に大きな役割を果たしたからである。レイシズム言説が最初に目立つようになったのは匿名掲示板の“2ちゃんねる”であったし、排外デモや街宣を行うヘイトグル

-
- 3) コリアンの呼称としては、“韓国人”“朝鮮人”なども用いられるが、それらはしばしば、特定の国家や地域と結びつけて用いられる。そこで本論文では、それらの総称として、コリアンの語を用いることとした。分析にあたっては、国籍国・地域や居住国・地域を限定せず、“韓国人”“朝鮮人”などの語で言及される対象一般についての言説を収集することとした。
- 4) 本研究では、民族偏見・人種偏見の総称として、“レイシズム (racism)”の語を用いる。“人種 (race)”と“民族 (ethnicity)”を生物学的基板の有無という点で異なるものとする用法もあるが、その区別は、人間間の差異も生物学的基板を欠いているという科学的事実から言っても、一般人の持つ素人理論の観点から言っても、さほど有益ではない（Zárate, 2008）。実際、The New Oxford American Dictionary などにおいても、race の語は ethnicity も含むと定義されている（“race,” 2005/2008）。

ープが支持を拡大したのは、動画共有サイト、ブログ、Twitter などの活用によるところが大きい（安田，2012）。また、東日本大震災後に典型的に見られたように、ソーシャル・メディアでは、コリアンなどの外国籍住民についての悪質なデマがしばしば流行する（荻上，2011）。したがって、ソーシャル・メディアでのレイシズムの扇動という問題は、現在の日本におけるレイシズムを考えると、避けて通れない問題である。

レイシズムの扇動をするものを離れてより一般的な人口について考えるときにも、ソーシャル・メディアは有益な分析対象である。ソーシャル・メディアでは、多くの個人や団体が情報発信をし、同時に様々な情報を受信し、影響を受け、再びアウトプットしている。したがって、ソーシャル・メディアは、現在では日本の全人口の約 60% を占めるインターネット・ユーザー（総務省統計研修所，2010）の信念や態度を映す鑑となる。

本研究ではソーシャル・メディアの中でも特に Twitter を取り上げ、分析を行う。Twitter に着目した理由として、まず、Twitter が日本で最も利用されているソーシャル・メディアの一つであることが挙げられる。2013 年 3 月時点で、国内でパーソナル・コンピュータを用いてインターネットにアクセスするユーザーのおよそ 22.3%、1,315 万人が Twitter にアクセスしている。これは、1,715 万人の Facebook について多い（関根，2013）。また、パーソナル・コンピュータを用いてアクセスするユーザーと一部重複するが、スマートフォンなどを用いてアクセスするユーザーも多く、2012 年 12 月時点で iPhone ユーザーの 88.0% が Twitter を利用している（許，2013）。したがって Twitter の分析は、インターネットのユーザーの少なくない部分の分析を可能にする。

また、Twitter が様々なサイト・サービスと連携していることにも、分析を行う理由がある。ユーザーは、様々なニュースサイトやブログから簡単な操作で Twitter に記事名や URL を投稿したり、他の Web サイト上の記事に“埋め込み”されたツイートを閲覧したりすることができる。このため、Twitter はインターネット上の情報のいわば結節点となっている。

さらに、ツイートの多くが公開情報であることも、理由の一つに挙げられる。Twitter にも許可したユーザー以外に自身の投稿を講読させない機能はあるが、そうした機能を利用する必要性は、主として身内とのコミュニケーションを目的とする SNS に比べて低い。例えば大学生に限ると、常に投稿を非公開にしている Twitter ユーザーは 21.2% に過ぎず、逆に常に公開しているユーザーは 71.2% であった（水沼・菅原・池内, 2013）。そのため、研究者がユーザーに許諾を求めることなく大部分の発言を捕捉できるというメリットがある。

以上のことから、本研究では Twitter 上でのコリアンに関する言説を収集し、分析を行うこととした。

分析の手法としては、主に計量テキスト分析を用いる。この手法は、テキストデータに対してある種の数値化操作を加えることで、計量的に把握することを可能にするものである（樋口, 2012）。日本語のものだけでも一日に 5,000 万件以上とも言われる膨大なツイート（NEC ビッグロブ, 2012）のできる限り多くを活用して分析を行うためには、質的な分析を行う前に情報を量的なものに変換して処理することが欠かせない。本研究では、過去の研究を参考に指針を構築して量的な分析を行い、それを元に質的な解釈を行う。

方法

2012 年 11 月 5 日から 2013 年 2 月 16 日にかけてツイートを収集した。収集には、当時 Twitter が公式にサポートしていた RSS 機能を用いた⁵⁾。RSS とは、Web サイトの更新情報を取得するための機能であり、Twitter の場合には、リクエストを送信した時点から遡って最新のツイートを 15 件まで取得することができた。データの取得には Thunderbird (Mozilla, Inc) を使い、15 分毎に自動でリクエストを送信した。つまり、この方法では、15 分以内に 16 件以上の投稿があれば最新の 15 件を、それ以下の投稿しかなければ全ての投稿を、取得することができた。したがってこの方法で取得されたデータは、全数でもないし無作為抽出でもないことを、予め断っておく必要がある、総投稿数が多い日中～夜間の特定のツイートが抽出される確率は相対的に低く、総投稿数が少ない深夜～早朝の特定のツイートが抽出される確率は相対的に高くなる。しかしながら、このような時間的な偏りが投稿内容の偏りをもたらすとは考えにくいため、この方法を用いた。なお、機材のメンテナンスのため、上記期間中にも不定期に抽出を停止していた時間がある。

ツイートの一部を抽出したのは、全てのツイートを使用すると、ハードウェア・ソフトウェア的に扱うのが困難なデータ・サイズに数日で達してしまうが、短期間で取得したデータでは、少数のツイートが爆発的にリツイート（他人の投稿したツイートを自分のフォロワー＝講読者にも配信されるようにする、特殊な投稿形式）されたものや“炎上”と呼ばれる単一のトピックについて短期間に投稿が集中したものが多数を占

5) この機能は 2013 年 3 月に廃止されたため、現在同一の方法を用いてデータを取得することは不可能である。

めることがあり、一時的なトレンドではなく Twitter での言説一般を分析するには不適切なことによる。また、外交上の事件などにより一時的にツイートの内容が左右される可能性があることも、理由の一つである。後者の点については、時期の要因も加えた分析を行うという方法も考えられたのだが、本研究では、3ヶ月に渡って収集したデータをまとめて分析することで、時期の影響をできるだけ排除することを試みた。これは、コリアンについての言説を分析する研究の端緒として、ベースラインとなる言説の特徴を明らかにすることを目指したからである⁶⁾。

抽出に用いられたのは、“在日 or 韓国人 or 朝鮮人 or チョン⁷⁾”の検索クエリである。この場合、いずれかの単語を用いているツイート最大15件が一度に取得されることになる。上記の4単語のいずれも用いていないツイートは、文脈上明らかにコリアンに言及するものであったとしても、本研究では取得されていない。

また、比較対象として、日本語ツイートのワイルドカード (“lang=ja&q=”) の検索結果も、同様の期間・方法で取得した。この検索クエリで抽出対象となるのは、ユーザーの使用する言語が日本語に設定されている全てのツイートである。

なお、RSSで収集したデータでは、公式リツイートの場合で、リツイートされたものであることを示す文字列である“RT @ (Twitter ID)”を冒頭に加えた時の文字数が140文字を超過したときには、超過した終末部分が省略されるという制約があった。また、RSSでデータ

6) なお、データを収集した時期としては、2012年8月の韓国李明博大統領の竹島（独島）上陸を受けて、官民両レベルにおいて日韓関係が冷却していた時期である。また、2012年11月16日に野田・民主党内閣が解散、12月16日に総選挙が行われ、韓国・朝鮮民主主義人民共和国に対して強硬な姿勢を取る安倍総裁率いる自民党が大勝した。

7) コリアンを指す蔑称であり、インターネット上ではしばしば用いられるものである。本研究ではその目的上、差別語やスラングなどの社会的に望ましくない用語も用いられているが、研究の資料的価値を鑑み、極力編集は加えずにそのまま掲載した。

を集めた場合には、発言者のデモグラフィック変数は取得できないという限界もあった。それらを取得するためには、回答者の許諾を伴う調査を行う必要があるが、その場合にはデータを網羅的に取得し言説の総体を分析するという目的を達することは不可能になる。そこで、種々の制約があるとはいえ、RSS によるデータ取得が最も適していると考えた⁸⁾。

結果と考察

コリアン関係ツイートとして、113,189 件のツイートを取得した。このうち、以下の基準に従い、分析対象に加えるべきではないと思われるツイート 3,600 件を削除した。(1) “朝鮮人” “韓国人” “チョン” のいずれも含まず、“在日アメリカ” “在日米” “在日ブラジル” “在日中国” “在日ペルー” “在日ファンク” “大使館” を含むツイートは、フィルタにより削除した。なお、“在日外国人” (1,640 件) は、暗に在日コリアンを指すと思われるツイートが大半だったため、削除せずに残した。(2) “在日” “韓国人” “朝鮮人” のいずれも含まず、“チョン・” “ジフン” “ダヨン” “ウソン” “ジヒョン” “テセ” “首チョン” “チョン切” を含むものは、フィルタにより削除した。(3) “チョンと” を含むツイートは一件ずつ確認し、“チョン” が擬音語、擬態語、固有人名と思われるものを削除した。その結果、109,589 件のツイートが分析対象となった。なお、リツイート (48,934 件) は、独立したツイートとして扱われている。

ワイルドカード検索で抽出されたのは 114,932 件で、その全てを分析した。リツイートは 10,517 件であった。

8) 参考のため、Twitter ユーザー一般（本研究に用いられたデータに含まれているとは限らない日本のユーザー）の特徴を記す。性別の分布は、男性 42%、女性 58% と、女性の方がやや多い。年齢の分布は、10 代 12%、20 代 14%、30 代 20%、40 代 26%、50 代以上 29% となっている（関根、2013）。

コリアン関係ツイートとワイルドカードのツイートでの、総発言数に占めるリツイートの比率を計算すると、前者では44.7%、後者では10.4%であり、有意差があった ($\chi^2(1) = 35672.1, p < .001$)。したがって、コリアンに関するツイートは一般のツイートに比べて“拡散”されやすいと言える。

発言者の性質

コリアン関係ツイートの投稿者として捕捉されたユーザーは43,619IDであった⁹⁾・このうち77.6%は捕捉されたツイートは1件のみであったが、捕捉されたツイートが100件を超えるユーザーも47IDあった。これらのユーザーはいずれも明確に差別的な投稿を大量にしており、また投稿時のクライアントはtwittbot (バンプアップ株式会社) などの自動投稿可能なものであった¹⁰⁾ ため、bot と呼ばれる自動投稿用プログラムであると考えられる。これら47IDの2013年3月22日時点のフォロワー数は182~15,499IDであり、中央値は2,135IDであった(“凍結¹¹⁾”もしくは削除されていたアカウントを除く)。これは、フォロワー数が200ID以下のユーザーが過半数を占める一般的なユーザー(facenavi, 2012) に比べ、遥かに多い。

他方、ワイルドカード検索で得られたサンプルでは、捕捉されたアカウント数は107,145IDで、87.7%のアカウントは捕捉されたツイート数

9) 取得期間中にアカウント名(半角・全角文字の両方が使用可能な、いつでも自由に変更可能な識別子)を変更した場合でも一つのIDとして扱うことができているが、ID(半角文字のみの、容易に変更できない識別子)を変更したユーザーがいた場合、別個のユーザーとして扱われている。一人のユーザーが複数のIDを使用している可能性もある。

10) Twitterでは、公式のWebサイト (<https://twitter.com/>) にブラウザでアクセスして利用する以外に、外部の開発者によるアプリケーションを用いてアクセスすることもできる。これらのアプリケーションの中には、決まったルールに従い自動的に投稿を繰り返すことができるものもある。

11) スпам行為などを理由に、アカウントの一時的な利用停止処分を受けること。

が1件のみであった。また、捕捉されたツイート数が最大のアカウントでも、15件のツイートが捕捉されたのみであった。

投稿数が多い上位25IDが全体の投稿数に占めるシェアは、コリアン関係ツイートでは10.1%に昇ったが、ワイルドカードツイートにおいては0.16%に留まった。上位50IDでは、コリアン関係ツイートでは12.8%であったが、ワイルドカードツイートでは0.27%であった。したがって、コリアン関係ツイートの方が、プレゼンス（投稿数およびフォロワー数）の極端に高い少数のユーザーによって寡占されていることが分かる。差別的な一部のアカウントに多くのユーザーが繋がる形のネットワーク構造が示唆される。

頻出語の分析

以下の分析には KHCoder 2.beta.30f（樋口, 2001/2013）を用い、形態素解析エンジンには Mecab 0.996（工藤, 2006/2013）を用いた。分析手順としては、まず KHCoder により分析対象ファイルの文字化け等を自動で修正した後に形態素解析を実行し、頻繁に登場する複合語や特殊記号を強制抽出語に加えて再度形態素解析するプロセスを繰り返した。

まず、用いられたツイート数が多かった後のうち、名詞（“人” “国” “韓” “日” など漢字1文字で通常複合語に用いられる名詞を除く）および解釈の可能な記号・文字列（“RT” および “@” を除く¹²⁾）のみの上位60件を、表1に示す。

抽出に用いられた語のうち、差別語である“チョン”の語は相対的に

12) “RT” は、他人のツイートをリツイートしたり引用したりしていることを示すために、ツイート冒頭及びツイート中で用いられる文字列。“@” は、フォロー関係に関わらず他人の画面に自分のツイートを表示させる機能（“リプライ”と呼ばれ、会話などに用いられる）を使ったり、リツイートや引用であることを示したりするために用いられる記号である。これらは、発言の内容を分析する上では有用ではないため、分析から除外した。

表 1 頻出語とそれを含む発言数

順位	単語	発言数	順位	単語	発言数	順位	単語	発言数
1	韓国人	50360	21	生活保護	3971	41	逮捕	2462
2	http://	48909	22	中国	3883	42	原発	2355
3	在日	43263	23	安倍	3681	43	歴史	2339
4	朝鮮人	28960	24	選挙	3550	44	AKB	2326
5	日本	25399	25	デモ	3463	45	北朝鮮	2314
6	韓国	24982	26	報道	3280	46	マスコミ	2308
7	日本人	20426	27	慰安婦	3179	47	経済	2267
8	#	18421	28	速報	3048	48	世界	2237
9	チョン	8550	29	自民党	2977	49	嘘	2170
10	朝鮮	7776	30	事件	2960	50	民族	2141
11	反日	6653	31	竹島	2937	51	反対	2114
12	拡散	6348	32	人権	2909	52	政権	1989
13	外国人	6251	33	参政権	2798	53	嵐	1983
14	政治	5115	34	特権	2746	54	KPOP	1961
15	民主党	4679	35	女性	2676	55	国家	1903
16	問題	4496	36	自分	2597	56	2ch	1880
17	差別	4413	37	法案	2548	57	廃止	1869
18	中国人	4344	38	通名	2540	58	工作	1823
19	犯罪	4321	39	国民	2516	59	時代	1794
20	ニュース	4152	40	捏造	2472	60	関係	1791

は最も少ないとはいえ、抽出されたツイートの7.8%で用いられており、決して無視できないボリュームであった。

抽出に用いられた単語以外では、“日本” (25,399件)、“日本人” (20,426件)、“反日” (6,653件) など、日本に何らかの形で言及するものが多く見られた。

以下でコリアンについてのレイシズムを分析する枠組みとしては、古典的レイシズムと現代的レイシズムに着目する。これは、米国における黒人に対するレイシズムの分類 (McConahay, 1986; Sears, 1988) を、高 (2013)、高・雨宮 (2013) が在日コリアンに対するレイシズムの分

析に援用したものである。古典的レイシズムは、マイノリティは劣っているという信念に基づくあからさまなものであるが、現代的レイシズムは、(1) 偏見や差別は既に存在せず、(2) したがって存在する格差はマイノリティの努力の欠如によるものであり、(3) にもかかわらずマイノリティは差別に抗議し過剰な要求を行い、(4) 本来得るべきもの以上の特権を得ているという、4つの信念に基づく、より隠微なものである (McConahay, 1986; Sears, 1988)。コリアンに対するヘイトスピーチは、コリアンの劣等性を主張するものである場合もあるし、コリアン (在日コリアン) が不当な“特権”を得ているということを糾弾するという形を取ることもある (安田, 2012)。

そこでまず古典的レイシズムに関して、コリアンの劣等性に何らかの形で言及するときに用いられると考えられる語に注目すると、“犯罪” (4,321 件)、“事件” (2,960 件)、“逮捕” (2,462 件) など、犯罪、つまり道徳的劣等性に関する語が多く見られた。なお、“女性” (2,676 件) の語も、日本人女性における韓流ブームに言及したものも多かったが、コリアン男性が日本人女性に性的暴行を加えているといった内容での投稿も多く見られた。

次に現代的レイシズムに関して、コリアンの地位や権利に関わる語を検討すると、“生活保護” (3,971 件)、“人権” (2,909 件)、“参政権” (2,798 件)、“特権” (2,746 件)、“通名” (2,540 件) などの語が見られた。

ただしこれらのツイートは、コリアンに対して否定的なもののばかりではない。コリアンへの差別を批判するものなども計上されているため、これらの数は何らかの形でレイシズムに関係するツイートで、そのうち関連する語を含むものの数である。なお、今回収集された発言が主にコリアンに対してネガティブなものであったことについては、後述する。

“政治” (5,115 件)、“民主党” (4,679 件)、“安倍” (3,681 件)、“選挙”

(3,550 件)、“自民党” (2,977 件)、“法案” (2,548 件)、“政権” (1,989 件) など、政治と政治家に関する語も多く見られている。これは、ある程度まではデータ取得期間中の 2012 年 12 月に衆議院選挙があったことの影響もあると思われるが、在日コリアンを巡る法的問題や韓国・朝鮮民主主義人民共和国との外交問題が話題になりやすいことも関わっていると考えられる。

また、コリアンに関する差別的言説は、しばしばマスコミへの不信感を伴って表出される。安田 (2012) は、ヘイトグループの関係者の多くが、新聞、雑誌、テレビなどが自分たちから“真実”を隠しているという被害者感情を抱いているとしている。そこでマスコミに関する語に注目すると、“ニュース” (4,152 件)、“報道” (3,280 件)、“マスコミ” (2,308 件) などの語が見られた。ただしこのうち、“ニュース”の語は、後述する 2 ちゃんねるまとめブログのブログ名の場合も多かった。また表 1 には含まれていないが、マスコミに侮蔑的に言及するときのスラングである“マスゴミ”の語も 1,290 件見られた。

次に、“真実が隠されている”という被害者感情に注目して、“何らかの隠された真実がある”ことに言及するときに用いられるような単語に目を向けると、“捏造” (2,472 件)、“嘘” (2,170 件) などが上位に位置していることが分かる。

マスコミ以外のメディアに目を向けると、“2ch” (1,880 件) の語が多く見られた。これは、多くは“#2ch”の形で、ハッシュタグと呼ばれる Twitter の機能を利用するために用いられていた。ハッシュタグとは、“#”を先頭につけた文字列を投稿した場合に同じ文字列を含むツイートが容易に検索可能になる、話題を共有するための機能およびそのための文字列の呼び名である。この場合には、匿名掲示板の 2 ちゃんねるもしくは、“2 ちゃんねるまとめブログ”と呼ばれる、2 ちゃんねるの投稿を

編集してブログ形式で投稿するサイトのどちらかに関するものであることを示すために用いられていたと考えられる。なお“#2ch”を含む投稿が多かったのは、2ちゃんねるまとめブログからツイートボタンを経由して投稿した場合、自動的にこのハッシュタグを含むように設定されている場合が多いためであったようである。また、2ちゃんねるまとめブログの名称が“2ch”の文字列を含んでいる場合も多かった。これに加えて、“速報” (3,048 件) の語も、そのほとんどは2ちゃんねるまとめブログの名称の一部であった。

また、“慰安婦” (3,179 件)、“竹島” (2,937 件)、“歴史” (2,339 件) など、歴史問題・外交問題に関する語も多く見られている。

この他に、一部の文字列・記号が多く見られていることも特徴的である。上述した話題の共有を行うためのハッシュタグ機能が用いられていることを示す“#”の記号は、18,421 件ものツイートで見られている。また、“拡散”の語も多く見られているが、これは“【拡散】”“【拡散希望】”などの文字列をツイートの冒頭につけることで、リツイートし“拡散”するように訴えるために用いられていた。また、何らかのサイトに向けてリンクが貼られていることを示す“http://”の文字列は 48,909 件と、抽出に用いられた語を除けば最も多くのツイートに含まれていた。これらのリンク先には、ニュースサイト、2ちゃんねるまとめブログ、その他のブログ、動画共有サイト、他のツイートの個別のページ、画像アップローダーなど、様々なウェブサイトが含まれていた。

なお、“AKB” (2,326 件)、“嵐” (1,983 件)、“KPOP” (1,961 件) など、芸能人に関する語も多く見られた。これらのほとんどは、それらの芸能人に直接言及するものではなく、一般の関心が高く使用頻度の高い芸能人に関するハッシュタグを差別的なツイートに付け加えることで、“拡散”されやすくしたものと思われた。

コードに基づく分析

以上のように頻出語からもある程度言説の内容をうかがい知ることができるのだが、より詳細な分析を行うために、関連する単語同士を同じコードに分類し、ツイートのテーマを分析するために用いる。用いるコードと分類に用いる単語は、表2のとおりである。各コードについて、表中のいずれかの単語を含んでいる場合は、そのコードに分類した。一つのツイートが複数のコードに分類されることもあった。

“コリアン”コードには、抽出に用いた語を含め、コリアンやコリアンの国・言語を指す語を含めた。“古典的レイシズム”“現代的レイシズム”コードは、上述のようにコリアンの劣等性や“特権”などに関わる語である。これらのコードを含むツイートの中には、“コリアンの犯罪者が逮捕された”というような“事実の摘示”を行っているものもあるが、しかしそのような事実の摘示がコリアンについてのみ頻繁に行われる傾向があるのであれば、レイシズムを反映していると考えることができる。

“歴史問題”にはコリアン（国民、住民、および国家）に対する日本の加害行為に言及するものを分類した。“虐殺”（1,266件）の語は、コリアンに対する虐殺よりも“コリアンが日本人を虐殺した”とするものの方が多かったため、このコードには含めていない。“外交問題”には、韓国又は朝鮮民主主義人民共和国との間に抱える問題に関する語を分類した。“マスコミ”“政治”“中国”には、それぞれのテーマに関わる語を分類した。また、頻出語リストに見られた“反日”に加えて、“売国”“売国奴”などの、何者かが日本に対して敵対的な行為・認知をしているとする際に用いられる語を、“反日系”に分類した。

“真実系”としては、何らかの隠された真実がある、それを明らかにする、とするツイートで用いられる語を分類した。“侮辱語”としては、

表 2 ツイート分類コード¹³⁾

コード	分類に用いられた単語
コリアン	韓国人、朝鮮人、在日、チョン、韓国、朝鮮、韓、朝鮮語、ハングル
現代的レイシズム	特権、生活保護、受給、人権、通り名、通名、参政権、年金
古典的レイシズム	犯罪、事件、逮捕、悪事、凶悪、強姦、レイプ、犯す、襲う、強盗、整形、劣る、劣等、ヤクザ、火病
歴史問題	歴史、大日本帝国、植民地、統治、併合、支配、侵略、加害、強制連行、連行、慰安婦、従軍、太平洋戦争、戦争、戦前、戦後、敗戦、関東大震災、謝罪、賠償、補償
外交問題	竹島、独島、領土、東海、日本海、領土、ミサイル、拉致、条約
マスコミ	マスコミ、マスゴミ、ニュース、テレビ、テレビ局、新聞、放送、報道、記事、記者
政治	政治、政党、政府、政権、議員、首相、総理、大統領、民主党、民主、自民党、自民、維新、安倍、選挙、法律、法案、政策、憲法
中国	中国、中国人、中国語、シナ、支那、シナ人、支那人
反日系	反日、売国、売国奴
真実系	実態、真実、本性、正体、発覚、疑惑、暴露、嘘、ウソ、嘘八百、捏造、自演、歪曲、隠蔽、矛盾、陰謀
侮辱語	馬鹿、バカ、アホ、ボケ、カス、クズ、ゴミ、ゴキブリ、寄生虫、ウジ、ウジ虫、ハエ、害虫、ブタ、鬼畜、糞、ガン、嘘つき、童貞、底辺、乞食、変態、デブ、気違い、基地外、マジキチ
2ちゃんねる	2ch、速報、ニュー速、v 速、嫌儲、鬼女、および 14 個の代表的な 2 ちゃんねるのまとめブログ名、2 ちゃんねるのまとめブログの管理人の TwitteID (5 つ)
拡散呼びかけ	拡散、RT 希望、RT を、RT 望、RT 願
話題の共有	#、#

- 13) 本表において、以下のスラング・蔑称が記載されている。“火病”はネット・スラングで、コリアンに特有の精神疾患とされているもの。“支那”、“シナ”は中国の蔑称。“基地外”は“気違い”のネット・スラング、“マジキチ”は“マジで気違い”を省略したネット・スラング。“ニュー速”、“v 速”、“嫌儲”、“鬼女”は、それぞれ 2 ちゃんねるの、“板”と呼ばれる掲示板群の小分類の呼称。

他者を激しく侮辱する際に用いられる多くの語を含めた。“2ちゃんねる”のコードには、2ちゃんねるそのものと、2ちゃんねるまとめブログに関する語・文字列を分類した。これらの中には、代表的なまとめブログおよびコリアンや中国人に関しては話題にのぼることの多いまとめブログの、サイト名および管理者の Twitter アカウントの ID を含めた。

これらに加えて、ハッシュタグに用いられる記号を“話題の共有”コードに、“拡散”を呼びかけるときに用いられる語・文字列を“拡散呼びかけ”のコードに分類した。

以上のようにコーディングを行い、それらに該当するツイート数と総抽出ツイートに占める出現率（％）を示したものが、表3である。

コリアン関連ツイートで“コリアン”コードが100％ではないのは、抽出語が何らかの複合語の一部として分類されているためであると考え

表3 コード別出現ツイート数と出現率

コード	コリアン関連ツイート		ワイルドカード	
	ツイート数	出現率（％）	ツイート数	出現率（％）
コリアン	108,794	99.27	225	0.20
現代的レイシズム	13,365	12.20	60	0.05
古典的レイシズム	11,778	10.75	349	0.30
歴史問題	12,336	11.26	253	0.22
外交問題	4,920	4.49	80	0.07
マスコミ	9,922	9.05	767	0.67
政治	22,654	20.67	767	0.67
中国	9,851	8.99	147	0.13
反日系	8,280	7.56	25	0.02
真実系	8,964	8.18	361	0.31
侮辱語	5,523	5.04	1,197	1.04
2ちゃんねる	5,721	5.22	317	0.28
拡散呼びかけ	6,348	5.79	461	0.40
話題の共有	18,991	17.33	7,257	6.31
コード無し	246	0.22	103,889	90.39

られる。その他のコードに関しては、コリアン関連ツイートではいずれも 4.49%~20.67% を占めているのに対して、ワイルドカードのツイートではほとんどが 1% を下回っており、これらのコードがコリアンに対する言説の特徴を捉えていることが分かる。

次に、コリアン関連ツイートについて、“コリアン”を除いた各コードの共起関係を検討することで、それぞれのテーマがどのように語られていたかを分析する。

図 1 には、コード間の共起ネットワークを示す。これは、一つのツイートが複数のコードに分類された場合にコードが共起したと考え、その関連の強さを示すものである。図の可読性に配慮して、以下の 2 通りを作図した。(a) 全てのコードを用い、Jaccard 係数が上位 10 本のパスを用いて作図した図、(b) もともとの生起率が極端に高い“政治”“話題の共有”を除いて上位 15 本を作図した図、

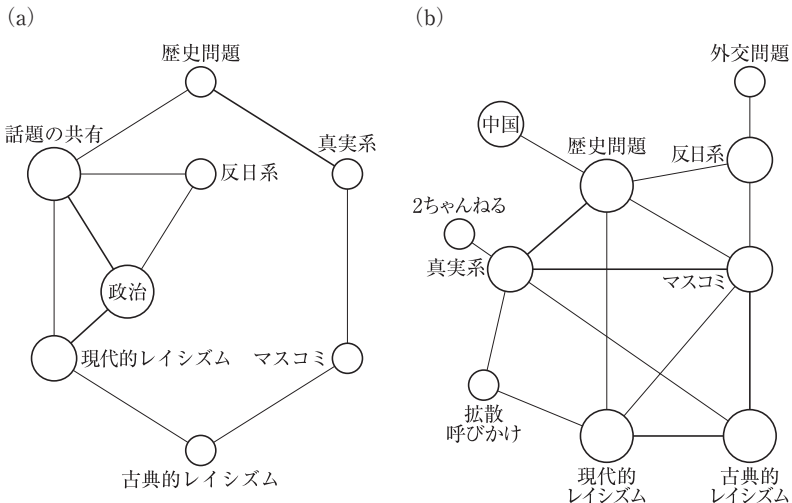


図 1：コード間の共起ネットワーク

(a) は全コードを用い Jaccard 係数の上位 10 本を作図。(b) は“政治”“話題の共有”を除いて上位 15 本を作図した。

円の大きさが出現率に、線の太さが Jaccard 係数の大きさに対応する。

題の共有”を除いた残りのコードのみを用い、Jaccard 係数が高い 15 本のパスを用いた図、である。

図 1a を見ると、“政治”と“話題の共有”同士も共起しやすいのだが、これらの語は“現代的レイシズム”“反日系”とクラスタを作っていることが分かる。これらの共起関係は、政府・政治家の在日コリアンの権利に対する姿勢（特定の政治家が在日コリアンの利益のために働いているなど）や“反日的な”姿勢への非難が、他者と共有しようとする意図を持って投稿されていたことを示している。

図 1b では、“現代的レイシズム”“真実系”のコードが“拡散呼びかけ”のコードと共起しており、これらの話題が積極的に他者に広めようという意図を伴って発言されていたことを示している。

この“真実系”は“古典的レイシズム”や“歴史問題”とも共起しており、これらについて隠された“真実”を明らかにするという主旨の発言がなされていたことを示している。また“真実系”は“マスコミ”と“2ちゃんねる”という 2 種類のメディアのコードとも共起していたのだが、内容を仔細に検討すると、マスコミは“真実”を隠蔽する主体として言及されることが多かったのに対して、2ちゃんねるは情報の出典が 2ちゃんねるもしくは 2ちゃんねるまとめブログであることを示すためにハッシュタグやまとめブログ名として言及されることが多かった。したがって、コリアンの劣等性や歴史問題についてマスコミが真実を隠しているという被害者意識があり、2ちゃんねるを情報源として攻撃を仕掛けているという構図が見て取れる。

“マスコミ”は、“反日系”“歴史問題”との間にクラスタを作っており、歴史問題についてマスコミが日本の利益に反する報道を行っているという言説が多いことを示している。

また、“現代的レイシズム”は“マスコミ”“歴史問題”と共起してお

り、これらは、マスコミが在日コリアンに特別な配慮をしているという言説や、在日コリアンが“特権”を得るに至った歴史的経緯が不当であるとする発言などが多いことを反映していた。

“中国”のコードは、“歴史問題”と共に起しており、歴史問題を抱えているという共通点から、コリアンに関する発言で中国も言及されていたことが分かる。

発言の感情価

最後に、投稿がコリアンに対してポジティブなものであるネガティブなものであるかを分析した。分析対象のコリアン関連ツイートから 150 件を無作為に抽出し、訓練された二人の評定者が、コリアンに対してネガティブ (1) / ニュートラルもしくは判断できない (2) / ポジティブ (3) の 3 段階で評定した。カッパ係数は .54 であり、“だいたい一致する”結果であった。一致しなかったものについては、両者の平均点をそのツイートの感情価とした。

その結果、ネガティブな発言が 70.0% (95%CI=62.6%~77.3%) と最も多く、次にポジティブ 17.3% (95%CI=11.2%~23.4%)、ニュートラルもしくは判断できない 12.7% (95%CI=7.34~18.0%) と続いた。したがって、コリアンについての言及の大部分は、コリアンに対するネガティブな態度が表れたものであった。

まとめ

本研究では、Twitter に投稿されるコリアンについてのツイートを計量的に分析した。

その結果、レイシズムに関する語や侮辱的な語が、決して少なくない

ツイートに見られることが明らかになった。コリアンについてなされるツイートの多くはネガティブなものであるが、その少なくない割合は、プレゼンスが極端に高いごく少数の差別的なユーザーによってなされていた。これらの差別扇動家の果たしている役割は、無視できないものであろう。

コリアンに関するツイートは、“拡散”や“共有”を指向するものが多く、また実際にリツイートされることも多かった。これらのツイートの話題は、現代的レイシズムや歴史問題、何者かが“反日的である”という非難や隠された“真実”を明らかにするというものなどであった。

また、安田（2012）が指摘するように、マスコミへの不信と、それに代わる2ちゃんねる・2ちゃんねるまとめブログへの依存も広範に見られた。

なお本研究では、ワイルドカード検索で収集したツイートとの比較を行ったのだが、コリアンについての言説の特徴をより明確にするためには、中国人などの他の人種・民族集団についての言説との比較が欠かせないであろう。この点は今後の検討に委ねるものとする。

引用文献

- facenavi. (2012). Twitter 日本人ユーザー・データ調査 フォロワー数、ツイート数など.
 facenavi. Retrieved January 16, 2014, from <http://facebook.boj.jp/twitter-user-survey-2012>
- 樋口耕一. (2012). KH Coder 2.X チュートリアル.
- 自由国民社. (2013). 「現代用語の基礎知識」選 ユーキャン新語・流行語大賞 全受賞記録.
 Retrieved January 05, 2014, from <http://singo.jiyu.co.jp/>
- 警察庁警備局. (2013). 治安の回顧と展望 (2013 年版). Retrieved from http://www.npa.go.jp/keibi/biki/kaiko_to_tenbou/H25/honbun.pdf
- 許直人. (2013). 【スマホ版】 mixi, Twitter, Facebook, LINE, カカオ 2012 年 12 月ソーシャルメディア最新視聴率調査 【VRI】 [in the loop]. in the loop. Retrieved from <http://media.looops>

net/naoto/2013/02/12/vri-sm3-2012-12/

- McConahay, J. B. (1986). Modern racism, ambivalence, and the modern racism scale. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, Discrimination, and Racism* (pp. 91-125). Orlando, FL: Academic Press.
- 水沼友宏・菅原真希・池内淳. (2013). 大学生の Twitter における行動規範に関する分析. 情報社会学会誌, 8 (1), 23-37.
- 守真弓. (2014). 売れるから「嫌中憎韓」書店に専用棚／週刊誌、何度も扱う. 朝日新聞. Retrieved February 28, 2014, from http://digital.asahi.com/articles/DA3S10972937.html?requesturl=articles/DA3S10972937.html&ieref=comkiji_txt_end_s_kjid_DA3S10972937
- NEC ビッグロブ. (2012). ツイッター分析：BIGLOBE が 11 月の Twitter 利用動向を発表：企業の公式アカウントによる“心に直接呼びかける”ツイートが話題に、ついつぶるトレンド. Retrieved January 02, 2014, from <http://tr.twipple.jp/info/bunseki/201211.html>
- 荻上チキ. (2011). 検証 東日本大震災の流言・デマ. 東京：光文社.
- race. (2005/2008). In *The New Oxford American Dictionary*. Oxford University Press.
- Sears, D. O. (1988). Symbolic racism. In P. A. Katz & D. A. Taylor (Eds.), *Eliminating racism: Profiles in controversy* (pp. 53-84). Plenum Press: New York & London.
- 関根健介. (2013). mixi, Twitter, Facebook, Google+, LinkedIn 2013 年 3 月最新ニールセン調査. in the loop. Retrieved December 27, 2013, from <http://media.loops.net/sekine/2013/05/08/neilsen-netview-201303/>
- 総務省統計研修所. (2010). 日本の統計 2010.
- 高史明. (2013). 在日コリアンに対する古典的／現代的レイシズム尺度の確認的因子分析と基礎的な検討. 人文研究, 180, 69-86.
- 高史明・雨宮有里. (2013). 在日コリアンに対する古典的／現代的レイシズムについての基礎的検討. 社会心理学研究, 28 (2), 67-76.
- 安田浩一. (2012). ネットと愛国：在特会の「闇」を追いかけて (kindle 版.). 東京：講談社.
- Zárate, M. A. (2008). Racism in the 21th Century. In T. Nelson (Ed.), *Handbook of Prejudice, Stereotyping, and Discrimination* (pp. 387-406). New York: Psychological Press.