

叫喚表現化によるツイート印象の変化の分析

熊本 忠彦

千葉工業大学 情報科学部 情報ネットワーク学科

kumamoto@net.it-chiba.ac.jp

概要 マイクロブログの一つである Twitter では、突発的な感情の発露を表すために、ツイートを「勝ったあああああああ」や「やめてええええ」というように叫喚表現化することがある。先行研究では、ツイートの叫喚表現化された部分を正規形（「勝った」や「やめて」）に変換することで、既存の様々な辞書（lexicon）を利用できるようにするための手法が提案されているが、叫喚表現化によりツイートの感情や印象がどう変化するのかについては触れていない。そこで本論文では、ツイートを読んだ人が受ける印象に焦点を当て、叫喚表現を含むツイート（叫喚ツイート）の印象と叫喚表現を正規化したツイート（非叫喚ツイート）の印象を 267 人の Twitter ユーザーが参加するアンケート調査に基づいて調べ、比較することで、その特徴を明らかにする。

キーワード SNS, Twitter, 叫喚, Word lengthening, ツイート印象

1 はじめに

近年、Twitter や Facebook, Instagram といった様々なタイプの SNS (Social Networking Service) が盛り上がりを見せている。特に、Twitter は、日本国内でも数多くのユーザが利用しており、(1) 1 回の投稿で入力可能な文字数が 140 字と少なく、気軽に投稿できる、(2) いつでもどこからでも簡単に投稿でき、リアルタイム性が高いといった点に特徴がある。このような Twitter の特徴を利用した研究は数多くなされており、トレンド分析 [1] やニュース記事推薦 [2]、リアルタイムイベント抽出 [3]、世論分析 [4]、群衆行動分析 [5] などに応用されている。一方、ツイートを読んだ人がそのツイートからどのような印象を感じるかに着目した研究もあり、ツイートの印象を定量的に算出する手法 [6] や任意のユーザの印象選好（そのユーザが普段、どのような話題の、どのような印象のツイートを見たり投稿したりしているか）を可視化するシステム [7] が提案されている。

Twitter では、突発的な感情の発露を表すために、ツイートの一部（文末や文中の母音、長音、記号など）を繰り返すことで、「勝ったああああああああ」や「やめてええええ」というような叫喚を表現することがある。このような叫喚表現は、ほとんどの場合、辞書（lexicon）に登録されていないため、形態素解析や単語出現頻度の算出といった各種テキスト処理の妨げとなる。そのため、先行研究 [8][9] では、ツイートの叫喚表現化された部分を正規形に変換することで、既存の様々な辞書を利用できるようにするための手法が提案されている。例えば、Brody らは、英語のツイートを対象に、単語の一部が 3 回以上繰り返された「niiice」や「reallly」のような叫喚表現を検出し、正規形（「nice」や「really」）に変換す

表 1 Twitter アカウントの有無（予備調査-Q1）

選択肢	回答者数 (%)
はい (=持っている)	9,439 (42.9%)
いいえ (=持っていない)	11,953 (54.3%)
ツイッターが何かわからない	618 (2.8%)
合計	22,010 (100%)

表 2 閲覧頻度別の回答者数（予備調査-Q2）

閲覧頻度	回答者数	閲覧頻度	回答者数
4 回以上/日	1,924	2~3 回/週	784
2~3 回/日	1,166	1 回/週	433
1 回/日	1,028	たまに	1,703
4~5 回/週	462	ほぼゼロ	1,939
		合計	9,439

表 3 本調査に参加した回答者の性別・年齢構成

	A 群			B 群		
	男性	女性	計	男性	女性	計
20 代	23	22	45	22	22	44
30 代	22	22	44	22	22	44
40 代	22	24	46	22	22	44
計	67	68	135	66	66	132

る手法 [8] を提案している。浅井らは、日本語のツイートを対象に、語尾の母音が 3 回以上繰り返された「うわあああああ」や「ねむいいいいい」のような叫喚表現を検出し、正規形（「うわあ」や「ねむい」）に変換する手法 [9] を提案している。一方、高橋らは、感情的なツイートでどのような叫喚表現が用いられているかを実験的に調べ、ひらがなとカタカナの母音あるいは記号「！」が繰り返されているツイートを検出すれば、最も高い F 値で感情的なツイートを抽出できることを示している [10]。しかしながら、いずれの先行研究も、叫喚表現化によりツイートの感情や印象がどう変化するのかについては触れていない。

表 4 本調査における設問項目 (A セット)

設問	設問内容
Q1	帰宅。お布団きもちいいい
Q2	5000 個限定は!?
Q3	寝不足つらあああああああああああい
Q4	電車がないぞ
Q5	ひええ照れるううう
Q6	みんな賢い言葉の準備はいいか
Q7	またメールの返信先間違てるうううううう
Q8	新しい靴履いたら足いたいん
Q9	自分の描いた絵が誰かのアイコンになるって興奮するうううううううう
Q10	今日は早くかえれる
Q11	お腹減った だれかごはんんんん
Q12	てか燃料すら危ないんだけど
Q13	もう仕事嫌だ仕事嫌だ仕事嫌だ稼がなきゃああああ
Q14	あー音ゲー意欲高まる
Q15	つっぱり棒落ちた死ぬほどびっくりしたあああ
Q16	明日の卒検不安だなあ
Q17	退勤 帰る
Q18	全力で転げたし、、はっず
Q19	全部かっこいいよね!!!! お金あれば全部ほしい
Q20	めーっちゃホッとした

そこで本論文では、ツイートを読んだ人が受ける印象に焦点を当て、叫喚表現を含むツイート（叫喚ツイート）の印象と叫喚表現を正規化したツイート（非叫喚ツイート）の印象を 267 人の Twitter ユーザが参加するアンケート調査に基づいて調べ、比較することで、叫喚表現化によりツイートの印象がどう変化するのかを分析し、その特徴を明らかにする。なお、本論文で対象とする印象は、国語学者の中村明が提唱した 10 種類の基本感情「喜、好、安、哀、厭、怖、怒、恥、昂、驚」[11] をベースとし、アンケート回答者には評価しやすいように、「喜び、好き、安心、悲しい、嫌い、怖れ、怒り、恥ずかしい、高揚、驚き」の 10 語を印象項目として提示する。

2 予備調査による Twitter ユーザの抽出

本章では、一定以上の頻度で Twitter を閲覧しているヘビーユーザを抽出するために行ったアンケート調査（予備調査）について述べる。

2.1 予備調査の概要

まず、20 代から 50 代のインターネットユーザを対象とする予備調査を実施し、男性 13,804 人、女性 8,206 人からなる 22,010 人分の回答データを収集した。具体的には、設問 Q1 で Twitter アカウントの有無を尋ね、「はい」、「いいえ」、「ツイッターが何かわからない」の三択の中から「はい」と回答した人には、設問 Q2 に進んでもらい、Twitter の閲覧頻度を尋ねた。Q2 に対する選択肢は、「ほぼ毎日（1 日に 4 回以上）」、「ほぼ毎日（1 日に 2 回～3 回）」、「ほぼ毎日（1 日に 1 回）」、「週 4～5 日」、「週 2～3 日」、「週 1 日」、「たまに閲覧している」、

表 5 本調査における設問項目 (B セット)

設問	設問内容
Q1	帰宅。お布団きもちいい
Q2	5000 個限定はあああああああああああ！
Q3	寝不足つらい
Q4	電車がないぞおおおおお
Q5	ひええ照れる
Q6	みんなああああ 賢い言葉の準備はいいかああああ
Q7	またメールの返信先間違てる
Q8	新しい靴履いたら足いたいんん
Q9	自分の描いた絵が誰かのアイコンになるって興奮する
Q10	今日は早くかえれるうううう
Q11	お腹減った だれかごはん
Q12	てか燃料すら危ないんだけどおおお
Q13	もう仕事嫌だ仕事嫌だ仕事嫌だ稼がなきゃあああ
Q14	あーーーー音ゲー意欲高まる
Q15	つっぱり棒落ちた死ぬほどびっくりした
Q16	明日の卒検不安だなああああ
Q17	退勤 帰るうううう
Q18	全力で転げたし、、はっずううう
Q19	全部かっこいいよね！お金あれば全部ほしい
Q20	めーっちゃホッとしたあああああ

「ほとんど閲覧していない」の 8 つであり、最も近いものを 1 つ選んでもらった。それぞれの集計結果を表 1 と表 2 に示す。

2.2 ヘビーユーザの抽出

表 1 と表 2 によれば、22,010 人のインターネットユーザのうち 9,439 人 (42.9%) が Twitter アカウントを持っており、その 9,439 人のうち「ほとんど閲覧していない」と回答した 1,939 人を除く 7,500 人 (79.5%) が Twitter を日常的に閲覧していることがわかる。本研究では、ツイートを読んだ人が受ける印象を対象にしていることから、Twitter を日常的に閲覧している 7,500 人をヘビーユーザと定義し、50 代を除く、20 代から 40 代の男女 5,190 人 (55.0%) を抽出した。

3 本調査に基づく印象評価データの取得

本章では、叫喚ツイートから受ける印象と非叫喚ツイートから受ける印象をアンケート調査（本調査）に基づいて調べる。

3.1 本調査の概要

まず、予備調査の結果に基づいて選出された Twitter を日常的に閲覧している 20 代から 40 代の男女を対象に本調査を実施し、267 人分の印象評価データを取得した。すなわち、叫喚ツイートと非叫喚ツイートからなる 20 個のツイートを 2 セット用意し、A セットを A 群の 135 人に、B セットを B 群の 132 人に読んでもらい、それぞれのツイートの印象を評価してもらった。本調査に参加した回答者 267 人の性別・年齢構成を表 3 に示す。

本調査の実施手順は以下の通りである。まず、叫喚ツ

表 6 ツイートの印象値

印象 項目	非叫喚ツイート			叫喚ツイート		
	最強	平均	最弱	最強	平均	最弱
喜	2.95	3.58	3.95	2.82	3.59	3.99
好	2.98	3.51	3.78	2.93	3.57	3.89
安	2.71	3.60	3.89	2.81	3.67	4.02
哀	3.11	3.59	4.11	3.07	3.62	3.98
厭	3.08	3.57	3.95	3.02	3.41	3.77
怖	3.30	3.69	4.11	3.27	3.67	3.98
怒	3.46	3.72	4.16	3.41	3.69	3.96
恥	3.11	3.59	3.90	2.90	3.43	3.80
昂	3.12	3.60	3.92	3.18	3.54	3.77
驚	2.94	3.48	3.86	3.14	3.49	3.79
平均	3.08	3.59	3.96	3.06	3.57	3.89

イートを 20 個, Twitter 上で収集した。浅井らは叫喚フレーズを「語尾の母音が 3 回以上繰り返して付加されている。母音は大文字, 小文字を区別しない。母音はひらがな, カタカナの大小文字すべて」[9] と定義したが, 本論文では叫喚表現の自動抽出が目的ではないので, 人が見て叫喚表現だと思うものを対象とした。具体的には, 文末や文中の母音や長音, 記号などの繰り返しを叫喚表現とした。次に, 叫喚ツイート内の叫喚表現を正規化することで, 20 個の非叫喚ツイートを得た。この 40 個のツイートを適当に分け, 表 4 に示した A セットと表 5 に示した B セットを作成した。但し, 表 4 と表 5 を見比べてみればわかるように, 同じ設問では, 一方に叫喚ツイートが, もう一方に対応する非叫喚ツイートが割り当てられている。また, 印象評価の際には, 国語学者の中村明が提唱した 10 種類の基本感情「喜, 好, 安, 哀, 厳, 怖, 怒, 恥, 昂, 驚」[11] をベースに, アンケート回答者が評価しやすいように、「喜び, 好き, 安心, 悲しい, 嫌い, 怖れ, 怒り, 恥ずかしい, 高揚, 驚き」の 10 語を印象項目として提示した。それぞれのツイートの印象に関して、「あなたがフォローしている人が以下のようないツイートをしたとします。そのツイートを読んで、感じる印象をお答えください。」と尋ね, 当該ツイートを提示した後, 10 語の印象項目のそれぞれについて「当てはまる, やや当てはまる, どちらともいえない, あまり当てはまらない, 当てはまらない」の 5 段階評価の中から最も近いものを 1 つ選択してもらった。

3.2 印象評価データの取得

まず, ツイートの印象を数値化するために、「当てはまる」に 1 点, 「やや当てはまる」に 2 点, 「どちらともいえない」に 3 点, 「あまり当てはまらない」に 4 点, 「当てはまらない」に 5 点を割り当て, ツイートごとにそれぞれの印象項目において回答者の印象評価データの平均値を算出した。この平均値をそのツイートの当該印象項目における印象値とする。以上の結果を, 叫喚ツイートと非叫喚ツイートに分けて, 表 6 にまとめる。表 6 には, そ

表 7 ツイートの印象値 (年齢別)

印象 項目	非叫喚ツイート			叫喚ツイート		
	最強	平均	最弱	最強	平均	最弱
喜	2.93	3.49	3.93	2.91	3.48	3.98
好	2.96	3.41	3.73	2.87	3.50	3.82
安	2.68	3.52	3.98	2.76	3.58	4.00
哀	3.07	3.50	3.98	2.98	3.51	3.86
厭	3.02	3.44	3.70	2.64	3.27	3.68
怖	3.18	3.52	3.98	3.22	3.56	3.84
怒	3.27	3.63	4.00	3.20	3.54	3.80
恥	2.96	3.49	3.84	2.82	3.33	3.77
昂	3.11	3.52	3.93	2.96	3.45	3.70
驚	2.93	3.36	3.73	2.95	3.39	3.75
平均	3.01	3.49	3.88	2.93	3.46	3.82

(a) 20 代の場合

印象 項目	非叫喚ツイート			叫喚ツイート		
	最強	平均	最弱	最強	平均	最弱
喜	2.89	3.60	4.05	2.75	3.64	4.02
好	2.93	3.49	3.91	2.80	3.58	3.91
安	2.82	3.58	3.98	2.73	3.66	4.02
哀	3.30	3.67	4.16	3.27	3.72	4.11
厭	3.05	3.60	3.98	3.07	3.51	4.00
怖	3.34	3.71	4.07	3.39	3.73	4.14
怒	3.41	3.72	4.11	3.43	3.79	4.11
恥	3.16	3.61	3.93	2.95	3.51	3.84
昂	3.11	3.62	3.98	3.27	3.59	3.91
驚	2.95	3.54	3.98	3.23	3.57	3.95
平均	3.10	3.62	4.01	3.09	3.63	4.00

(b) 30 代の場合

印象 項目	非叫喚ツイート			叫喚ツイート		
	最強	平均	最弱	最強	平均	最弱
喜	2.86	3.65	4.07	2.80	3.65	4.15
好	2.95	3.63	3.98	3.02	3.64	4.02
安	2.52	3.70	4.13	2.70	3.75	4.07
哀	2.86	3.61	4.36	2.96	3.62	4.07
厭	3.07	3.67	4.16	2.98	3.46	3.78
怖	3.33	3.78	4.41	3.16	3.77	4.07
怒	3.41	3.80	4.36	3.37	3.74	4.15
恥	3.11	3.65	4.07	2.77	3.46	3.86
昂	2.95	3.67	3.96	2.98	3.58	3.91
驚	2.80	3.53	3.87	3.00	3.52	3.93
平均	2.99	3.67	4.14	2.97	3.62	4.00

(c) 40 代の場合

印象 項目	非叫喚ツイート			叫喚ツイート		
	最強	平均	最弱	最強	平均	最弱
喜	2.86	3.65	4.07	2.80	3.65	4.15
好	2.95	3.63	3.98	3.02	3.64	4.02
安	2.52	3.70	4.13	2.70	3.75	4.07
哀	2.86	3.61	4.36	2.96	3.62	4.07
厭	3.07	3.67	4.16	2.98	3.46	3.78
怖	3.33	3.78	4.41	3.16	3.77	4.07
怒	3.41	3.80	4.36	3.37	3.74	4.15
恥	3.11	3.65	4.07	2.77	3.46	3.86
昂	2.95	3.67	3.96	2.98	3.58	3.91
驚	2.80	3.53	3.87	3.00	3.52	3.93
平均	2.99	3.67	4.14	2.97	3.62	4.00

それぞれの印象項目における全ツイートの印象値の平均値に加え, 当該印象が最も強かったツイートの印象値(最強値)と最も弱かったツイートの印象値(最弱値)も示す。さらに, 回答者を年齢別に分けた場合と男女別に分けた場合のツイートの印象値を算出し, それぞれの平均値, 最強値, 最弱値を表 7 と表 8 に示す。

4 印象評価データの比較・分析

4.1 節では, 前章で取得した叫喚ツイートと非叫喚ツイートの印象評価データを比較することで, 叫喚表現化によりツイートの印象がどう変化するのかを調べ, 4.2 節では, 叫喚ツイートと非叫喚ツイートのそれぞれの印

表 8 ツイートの印象値（男女別）
(a) 男性の場合

印象 項目	非叫喚ツイート			叫喚ツイート		
	最強	平均	最弱	最強	平均	最弱
喜	2.76	3.44	3.85	2.70	3.45	3.79
好	2.86	3.36	3.58	2.93	3.42	3.71
安	2.65	3.46	3.74	2.80	3.54	3.82
哀	2.91	3.45	3.95	2.99	3.45	3.77
厭	2.83	3.45	3.88	2.79	3.29	3.55
怖	3.14	3.56	3.95	3.02	3.54	3.83
怒	3.32	3.58	4.00	3.29	3.55	3.75
恥	2.79	3.49	3.79	2.71	3.32	3.67
昂	2.98	3.46	3.82	2.92	3.40	3.72
驚	2.74	3.40	3.73	3.14	3.39	3.76
平均	2.90	3.46	3.83	2.93	3.44	3.74

(b) 女性の場合

印象 項目	非叫喚ツイート			叫喚ツイート		
	最強	平均	最弱	最強	平均	最弱
喜	2.94	3.72	4.17	2.94	3.73	4.20
好	3.11	3.66	4.11	2.93	3.72	4.09
安	2.68	3.74	4.23	2.68	3.80	4.21
哀	3.22	3.74	4.26	3.15	3.79	4.20
厭	3.28	3.69	4.02	2.97	3.53	4.00
怖	3.24	3.81	4.26	3.44	3.81	4.23
怒	3.47	3.86	4.32	3.41	3.83	4.19
恥	3.06	3.68	4.02	2.94	3.54	4.03
昂	3.26	3.74	4.09	3.21	3.67	4.06
驚	3.06	3.56	3.98	3.12	3.60	4.02
平均	3.13	3.72	4.14	3.08	3.70	4.12

表 9 t 検定の結果（叫喚ツイート vs 非叫喚ツイート）

	喜	好	安	哀	厭
F 検定—片側確率	0.46	0.37	0.50	0.31	0.26
t 検定—両側確率	0.92	0.43	0.56	0.76	0.02
	怖	怒	恥	昂	驚
F 検定—片側確率	0.49	0.38	0.42	0.34	0.20
t 検定—両側確率	0.83	0.58	0.05	0.27	0.82

象評価データを分析することで、叫喚ツイートや非叫喚ツイートに対する印象の感じ方が年齢別あるいは男女別でどう異なるのかを調べる。

4.1 叫喚ツイートと非叫喚ツイートの印象比較

まず、表 6 に示した印象評価データに対し、有意水準 5% で F 検定を行い、分散が等しいとみなせる場合は分散が等しいときの t 検定を、みなせない場合は分散が異なるときの t 検定（Welch 法）を行った。F 検定を行ったときの片側確率と t 検定を行ったときの両側確率を表 9 に示す。

表 9 によれば、「厭」と「恥」において有意水準 5% で統計的に有意な差があることがわかる。また、表 6 から、この 2 つの印象項目では叫喚表現化によりツイートの印象が強くなっているのがわかる。そこで、それぞれの対応する叫喚ツイートと非叫喚ツイートの印象値に対し、F 検定と t 検定を行ってみたところ、有意水準 5% で統

表 10 t 検定の結果（年齢別、叫喚ツイート vs 非叫喚ツイート）

	(a) 20 代の場合				
	喜	好	安	哀	厭
F 検定—片側確率	0.39	0.39	0.39	0.47	0.22
t 検定—両側確率	0.92	0.18	0.55	0.93	0.02
	怖	怒	恥	昂	驚
F 検定—片側確率	0.23	0.34	0.33	0.38	0.26
t 検定—両側確率	0.49	0.10	0.08	0.28	0.74

	(b) 30 代の場合				
	喜	好	安	哀	厭
F 検定—片側確率	0.37	0.48	0.21	0.37	0.17
t 検定—両側確率	0.71	0.37	0.47	0.48	0.30
	怖	怒	恥	昂	驚
F 検定—片側確率	0.45	0.49	0.41	0.40	0.32
t 検定—両側確率	0.84	0.33	0.17	0.56	0.69

	(c) 40 代の場合				
	喜	好	安	哀	厭
F 検定—片側確率	0.47	0.25	0.33	0.11	0.34
t 検定—両側確率	0.98	0.96	0.68	0.88	0.00
	怖	怒	恥	昂	驚
F 検定—片側確率	0.31	0.44	0.41	0.48	0.37
t 検定—両側確率	0.92	0.39	0.03	0.27	0.95

計的に有意な差があり、かつ叫喚表現化によりツイートの印象が強くなったのは、「厭」の 5 つ（Q2, Q3, Q7, Q9, Q11）と「恥」の 4 つ（Q1, Q2, Q3, Q9）、「怒」の 1 つ（Q2）の計 10 個であり、ネガティブな印象の中でも特に「厭」と「恥」が強くなっているのがわかる。逆に、有意水準 5% で統計的に有意な差があり、かつ叫喚表現化によりツイートの印象が弱くなったのは、「喜」の 1 つ（Q9）と「好」の 1 つ（Q9）、「安」の 1 つ（Q2）の計 3 個であり、少数ではあるが、ポジティブな印象が弱くなっているのがわかる。

次に、叫喚表現化による印象の変化の仕方が年齢によって異なるのかを調べるために、表 7 に示した年齢別の印象評価データに対し、同じ要領で F 検定と t 検定を行った。F 検定を行ったときの片側確率と t 検定を行ったときの両側確率を表 10 に示す。

表 10 によれば、有意水準 5% で統計的に有意な差が観測されたのは、20 代では「厭」のみ、30 代では該当なし、40 代では「厭」と「恥」の 2 つであり、年齢によって叫喚表現化による印象の変化の仕方に違いがあることがわかる。なお、表 7 から、この 3 つの印象項目では叫喚表現化によりツイートの印象が強くなっているのがわかる。

さらに、叫喚表現化による印象の変化の仕方が男女によって異なるのかを調べるために、表 8 に示した男女別の印象評価データに対し、同じ要領で F 検定と t 検定を行った。F 検定を行ったときの片側確率と t 検定を行っ

表 11 t 検定の結果（男女別，叫喚ツイート vs 非叫喚ツイート）

(a) 男性の場合					
	喜	好	安	哀	厭
F 検定—片側確率	0.45	0.31	0.48	0.28	0.43
t 検定—両側確率	0.93	0.36	0.41	0.96	0.02
	怖	怒	恥	昂	驚
F 検定—片側確率	0.37	0.46	0.33	0.46	0.01
t 検定—両側確率	0.78	0.57	0.04	0.32	0.90

(b) 女性の場合					
	喜	好	安	哀	厭
F 検定—片側確率	0.40	0.44	0.49	0.35	0.17
t 検定—両側確率	0.91	0.52	0.69	0.61	0.06
	怖	怒	恥	昂	驚
F 検定—片側確率	0.42	0.39	0.34	0.47	0.35
t 検定—両側確率	0.91	0.67	0.17	0.37	0.64

表 12 t 検定の結果（年代間，叫喚ツイートの場合）
(a) 20 代と 30 代の比較

(a) 20 代と 30 代の比較					
	喜	好	安	哀	厭
F 検定—片側確率	0.34	0.18	0.21	0.46	0.16
t 検定—両側確率	0.15	0.39	0.46	0.01	0.01
	怖	怒	恥	昂	驚
F 検定—片側確率	0.45	0.17	0.16	0.45	0.39
t 検定—両側確率	0.00	0.00	0.04	0.03	0.02

(b) 20 代と 40 代の比較					
	喜	好	安	哀	厭
F 検定—片側確率	0.20	0.11	0.19	0.12	0.25
t 検定—両側確率	0.15	0.14	0.15	0.21	0.01
	怖	怒	恥	昂	驚
F 検定—片側確率	0.29	0.09	0.35	0.19	0.20
t 検定—両側確率	0.00	0.00	0.15	0.06	0.08

(c) 30 代と 40 代の比較					
	喜	好	安	哀	厭
F 検定—片側確率	0.34	0.38	0.47	0.15	0.05
t 検定—両側確率	0.92	0.54	0.49	0.27	0.54
	怖	怒	恥	昂	驚
F 検定—片側確率	0.33	0.35	0.26	0.23	0.29
t 検定—両側確率	0.53	0.51	0.55	0.97	0.59

たときの両側確率を表 11 に示す。

表 11 によれば、有意水準 5%で統計的に有意な差が観測されたのは、男性では「厭」と「恥」の 2つであったが、女性では該当なしであり、表 8 に示した印象評価データと合わせると、叫喚表現化によりツイートのネガティブな印象（特に「厭」と「恥」）をより強く感じる傾向が男性にはあることがわかる。

4.2 叫喚ツイートと非叫喚ツイートの印象分析

まず、20 代と 30 代の間、20 代と 40 代の間、30 代と 40 代の間で印象の感じ方に違いがあるのかを調べるために、表 7 に示した年齢別の印象評価データを用いて、叫喚ツイートと非叫喚ツイートのそれぞれにおいて F 検

表 13 t 検定の結果（年代間、非叫喚ツイートの場合）

(a) 20 代と 30 代の比較					
	喜	好	安	哀	厭
F 検定—片側確率	0.36	0.13	0.39	0.38	0.21
t 検定—両側確率	0.29	0.28	0.54	0.03	0.03
	怖	怒	恥	昂	驚
F 検定—片側確率	0.22	0.28	0.21	0.34	0.45
t 検定—両側確率	0.02	0.12	0.13	0.11	0.04

(b) 20 代と 40 代の比較					
	喜	好	安	哀	厭
F 検定—片側確率	0.15	0.21	0.15	0.01	0.30
t 検定—両側確率	0.17	0.01	0.16	0.36	0.00
	怖	怒	恥	昂	驚
F 検定—片側確率	0.04	0.15	0.39	0.26	0.29
t 検定—両側確率	0.00	0.01	0.07	0.04	0.06

(c) 30 代と 40 代の比較					
	喜	好	安	哀	厭
F 検定—片側確率	0.24	0.37	0.09	0.01	0.39
t 検定—両側確率	0.67	0.12	0.36	0.56	0.34
	怖	怒	恥	昂	驚
F 検定—片側確率	0.15	0.31	0.14	0.15	0.33
t 検定—両側確率	0.38	0.25	0.60	0.51	0.95

(a) 叫喚ツイートの場合					
	喜	好	安	哀	厭
F 検定—片側確率	0.11	0.07	0.07	0.33	0.14
t 検定—両側確率	0.01	0.00	0.03	0.00	0.01
	怖	怒	恥	昂	驚
F 検定—片側確率	0.33	0.22	0.07	0.17	0.01
t 検定—両側確率	0.00	0.00	0.02	0.00	0.01

(b) 非叫喚ツイートの場合					
	喜	好	安	哀	厭
F 検定—片側確率	0.20	0.04	0.07	0.41	0.38
t 検定—両側確率	0.01	0.00	0.02	0.00	0.00
	怖	怒	恥	昂	驚
F 検定—片側確率	0.16	0.13	0.25	0.22	0.45
t 検定—両側確率	0.00	0.00	0.03	0.00	0.06

定と t 検定を行った。それぞれの結果を表 12 と表 13 に示す。

表 12 によれば、有意水準 5%で統計的に有意な差が観測されたのは、20 代と 30 代の間では「哀、厭、怖、怒、恥、昂、驚」の 7 つであり、20 代と 40 代の間では「厭、怖、怒」の 3 つであったが、30 代と 40 代の間では該当なしであった。一方、表 13 によれば、有意水準 5%で統計的に有意な差が観測されたのは、20 代と 30 代の間では「哀、厭、怖、恥、驚」の 4 つであり、20 代と 40 代の間では「好、厭、怖、怒、昂」の 5 つであったが、30 代と 40 代の間では該当なしであった。統計的に有意な差が観測された場合は、いずれの印象項目でも 20 代の方が強い印象を感じていたことから、20 代の人は他の年代

に比べ、叫喚表現を含むか否かに関係なくツイートの印象をよりネガティブに感じる傾向にあることがわかる。

次に、男性と女性の間で印象の感じ方に違いがあるのかを調べるために、表 8 に示した男女別の印象評価データを用いて、叫喚ツイートと非叫喚ツイートのそれぞれにおいて F 検定と t 検定を行った。それぞれの結果を表 14 に示す。

表14によれば、叫喚ツイートに対してはすべての印象項目において、非叫喚ツイートに対しても「驚」以外のすべての印象項目において、有意水準5%で統計的に有意な差が観測されている。表8の印象評価データと合わせてみると、叫喚ツイートに対しても非叫喚ツイートに対しても男性の方が強い印象を感じていることがわかる。

5 まとめ

本論文では、Twitter 上のツイートを読んだ人が受けた印象に焦点を当て、叫喚表現を含むツイート（叫喚ツイート）の印象と叫喚表現を正規化したツイート（非叫喚ツイート）の印象を 267 人の Twitter ユーザが参加するアンケート調査に基づいて調べ、比較した。但し、本論文で対象とした印象は、国語学者の中村明が提唱した 10 種類の基本感情「喜、好、安、哀、厭、怖、怒、恥、昂、驚」[11] であり、アンケート調査時には「喜び、好き、安心、悲しい、嫌い、怖れ、怒り、恥ずかしい、高揚、驚き」の 10 語を印象項目として提示した。

アンケート調査の結果、叫喚表現化によりツイートに対するネガティブな印象（特に「厭」と「恥」）が強くなる傾向にあることがわかった。但し、20代から40代の年齢別の調査では、「厭」と「恥」が強くなったのは40代のみであり、20代では「厭」のみが強くなり、30代では印象の変化は見られなかった。また、男女別の調査では、「厭」と「恥」が強くなったのは男性のみであり、女性には印象の変化は見られなかった。一方、叫喚ツイートと非叫喚ツイートの印象をそれぞれ分析したところ、20代の人は他の年代の人に比べ、叫喚表現を含むか否かに関係なくツイートの印象をよりネガティブに感じる傾向にあることがわかった。男性と女性を比べた場合は、男性の方が、叫喚表現を含むか否かに関係なくすべての印象項目において強く感じる傾向にあることがわかった。

今後の課題として、以下を挙げる。本論文では、ツイートの印象値の比較や分析を年齢別あるいは男女別で行ったが、回答者数の制約により年齢と性別の両方を考慮した比較・分析は行っていない。アンケート調査の規模をより大規模化することで、年齢と性別の両方を考慮したより詳しい比較・分析を行いたい。また、本研究

の最終的な目的は、非叫喚ツイートだけでなく、叫喚ツイートに対しても、その印象値を精度よく算出する手法を開発することである。そのためには、非叫喚ツイートの印象値と叫喚ツイートの印象値の関係をより詳しく分析し、定式化することが必要とされる。

謝辞

本研究は JSPS 科研費 26330347 ならびに福田将治奨学寄付金による研究助成の成果であり、本研究の一部（ツイートの収集やアンケート調査の実施）は千葉工業大学情報ネットワーク学科を平成 28 年 3 月に卒業した小林薫氏と市原新一郎氏による卒業研究の一環として行われたものである。ここに記して感謝の意を表す。

参考文献

- [1] Michael Mathioudakis, and Nick Koudas, Twitter-Monitor: Trend Detection over the Twitter Stream, Proc. of the 2010 ACM SIGMOD International Conference on Management of Data , Indianapolis, Indiana, USA, pp. 1155–1158, 2010.
 - [2] Fabian Abel, Qi Gao, Geert-Jan Houben, and Ke Tao, Twitter-Based User Modeling for News Recommendations, Proc. of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, pp. 2962–2966, 2013.
 - [3] Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo, Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors, Proc. of the 19th International Conference on World Wide Web, Raleigh, North Carolina, USA, pp. 851–860, 2010.
 - [4] 那須野薰, 奥山晶二郎, 中西鏡子, 松尾豊, Twitterにおける候補者の選挙地盤に着目した国政選挙の当選者予測, 情報処理学会論文誌, Vol. 56, No. 10, pp. 2044–2053, 2015.
 - [5] 李龍, 若宮翔子, 角谷和俊, Tweet 分析による群衆行動を用いた地域特徴抽出, 情報処理学会論文誌データベース, Vol. 5, No. 2, pp. 36–52, 2012.
 - [6] 熊本忠彦, 河合由起子, 張建偉, ユーザ印象評価データの分析に基づく印象マイニング手法の設計と評価, 情報処理学会論文誌データベース, Vol. 6, No. 2, pp. 1–15, 2013.
 - [7] 熊本忠彦, 鈴木智也, Twitter ユーザの印象選好を可視化するシステムの設計と評価, 電子情報通信学会論文誌, Vol. J98-D, No. 5, pp. 788–801, 2015.
 - [8] Samuel Brody and Nicholas Diakopoulos, Cooooooooooooooollllllllllllll!!!!!!: Using Word Lengthening to Detect Sentiment in Microblogs, Proc. of the Conference on Empirical Methods in Natural Language Processing, Edinburgh, United Kingdom, pp. 562–570, 2011.
 - [9] 浅井洋樹, 秋岡明香, 山名早人, きたあああああああああああああああああ!!!! 11: マイクロブログを用いた教師なし叫喚フレーズ抽出, 第5回データ工学と情報マネジメントに関するフォーラム, A4-4, 2013.
 - [10] 高橋雄太, 片岡義雅, 浅井洋樹, 山本祐輔, 秋岡明香, 山名早人, 繰り返し表現を含んだ感情的なツイートの抽出, 第4回データ工学と情報マネジメントに関するフォーラム, C3-6, 2012.
 - [11] 中村明, 感情表現辞典, 東京堂出版, 東京, 1993.