ショートノート

記事の閲覧順序がその印象に及ぼす影響の分析†

熊本 忠彦*

ポジティブやネガティブといった記事の印象を扱う研究が増えつつあるが、いずれも記事単体の印象を対象としており、記事の閲覧順序がその印象に及ぼす影響については考慮していない。そこで本論文では、ポジティブな印象の記事、ネガティブな印象の記事、どちらの印象も持たない記事の3種類の記事を組み合わせ、先行記事の印象によって後続記事の印象がどう変わるのかを調べた。その結果、多くの場合で影響なしと判断されたが、一部では印象ギャップが広がる方向に作用したことが分かった。

キーワード：センチメント分析、記事印象、文脈効果、アンケート調査

1. はじめに

近年、テキストから受ける印象に焦点を当てた研究が盛んになりつつあり、各種レビューをポジティブ、ネガティブといったクラスに分類するための研究[1]や、ニュース記事を複数の印象クラス（例えば Awesome, Heartwarming, Surprising, Sad, Useful, Happy, Boring, Angry など）に分類するための研究[2]、入力されたテキストに任意の印象語をタグとして付与するための研究[3]などが行われている。その一方で、テキストの印象を有効に利用するシステムの開発も盛んになりつつあり、いわゆる評判分析[1]だけでなく、情報可視化[4, 5]や情報推薦[6, 7]、異メディアコンテンツ生成[8, 9]、情報の信頼性評価[10]といった様々な分野で応用されている。

しかしながら、これらの研究では、テキスト単体から受ける印象を対象としており、複数テキストの閲覧順序がその印象に及ぼす影響に関してはほとんど考慮していない。そこで本論文では、新聞記事を対象に、先行記事の印象が後続記事の印象に及ぼす影響を、アンケート調査に基づいて調べ、その文脈効果について考察する。

以下、2章でアンケート調査の仕様を設計し、3章でアンケート調査の実施について述べる。4章で調査結果の分析と考察を行い、最後に5章で本論文のまとめと今後の課題について述べる。

2. アンケート調査の設計

先行記事の印象が後続記事の印象に及ぼす影響を分析するために必要なデータをアンケート調査に基づいて収集することを考え、そのために必要とされるアンケート調査の仕様について検討する。

2.1 設計コンセプト

記事の閲覧順序を考えると、何個の記事を対象にするのかがまず問題となる。記事の数が増えれば増えるほど読む順番の場合数（組み合わせ方）も増えることから、今回は、基本となる2個の記事の組合せを考えることにする。すなわち、ポジティブな印象の記事、ネガティブな印象の記事、どちらの印象も持たない記事の3種類を対象とし、全部で9通り（＝3印象×3印象）の組合せを考えることにする。但し、一人の被験者が同じ記事を複数回読んでしまうと、学習効果が生じ、調査結果に悪影響を及ぼす可能性があることから、各被験者が同じ記事を2回以上読まないように設計する必要がある。一方、先行記事の印象が後続記事の印象に及ぼす影響を分析するためには、1個の後続記事に対して印象の異なる3個の先行記事が必要となるが、「各被験者は同じ記事を2回以上読まない」という制約があるため、工夫が必要となる。

そこで、ポジティブな印象の記事（Aと表記）、ネガティブな印象の記事（Bと表記）、どちらの印象も持たない記事（Oと表記）をそれぞれ6記事ずつ用意し、表1に示したように組み合わせることにより、この2つの問題を解決する。なお、表1において、「A1⇒A2」という表記は、ポジティブな印象の記事A1を読み、印象評価を行った後、別のポジティブな印象の記事A2を読み、印象評価を行うことを意味している。

† Context Effects of Reading Order of News Articles on Their Impressions
Tadahiko KUMAMOTO

* 千葉工業大学 情報科学部 情報ネットワーク学科
Department of Information and Network Science, Faculty of Information and Computer Science, Chiba Institute of Technology

表1 記事の組み合わせ方

記事ペア	記事グループ		
	1	2	3
1	A1 ⇒ A2	O1 ⇒ A2	B1 ⇒ A2
2	B1 ⇒ A3	A1 ⇒ A3	O1 ⇒ A3
3	O1 ⇒ A4	B1 ⇒ A4	A1 ⇒ A4
4	A5 ⇒ B3	O2 ⇒ B3	B2 ⇒ B3
5	B2 ⇒ B4	A5 ⇒ B4	O2 ⇒ B4
6	O2 ⇒ B5	B2 ⇒ B5	A5 ⇒ B5
7	A6 ⇒ O4	O3 ⇒ O4	B6 ⇒ O4
8	B6 ⇒ O5	A6 ⇒ O5	O3 ⇒ O5
9	O3 ⇒ O6	B6 ⇒ O6	A6 ⇒ O6

また、記事グループが異なれば、印象評価を行う被験者グループも異なるため、一人の被験者が同じ記事を2回以上読むこともないし、さらに、1個の後続記事に対して印象の異なる3個の先行記事が組み合わせられているのもわかる。

2.2 新聞記事の選出と記事ペアの作成

アンケート調査で用いるポジティブな印象の記事(明るい記事)、ネガティブな印象の記事(暗い記事)、どちらの印象も持たない記事(明るくも暗くもない記事)を得るために、以下のような予備実験を行った。

まず、被験者900人を9つのグループ(男女50人ずつ、計100人)に分け、各グループに毎日新聞データベース2002年版(社会面)から抽出した10記事を提示した。この10記事はグループによって異なり、全部で90記事が選ばれている。各被験者は、ランダムに提示される10記事を順に読み、それぞれの印象語(「明るい」あるいは「暗い」)によって表される印象をどの程度感じるかを、「強い(1点)–割と強い(2点)–割と弱い(3点)–弱い(4点)–なし(5点)」の5段階で評価した。なお、提示された記事は、元の記事の第1段落のみであり、個人情報保護の観点から個人の特定につながる情報(個人名、組織名^{*1}、地域名^{*2})は記号(◇や○、□)と置換された。

以上の実験の結果得られたデータ(10記事×100人×2印象語×9回)を分析し、それぞれの印象語に対する記事毎の平均値を求めた。

次に、印象語「明るい」と「暗い」のそれぞれにおいて印象の強い記事(平均値が3.0以下の記事)がどのよう

*1 「大学」や「会社」など組織の種別を表す部分のみ置換されていない。

*2 「県」や「市」など地域種別を表す部分は置換されていない。

(a) 明るい印象の強い記事の分布

(b) 暗い印象の強い記事の分布

図1 印象の強い記事(平均値が3.0以下)の分布

表2 選ばれた明るい記事の概要と平均値

ID	記事概要	明るい	暗い
A1	日本人女性宇宙飛行士が女児を出産した	2.54	4.86
A2	某高校で手作りロボットのサッカー大会があった	2.65	4.86
A3	某市でさくらまつりが開幕した	2.72	4.50
A4	名誉教授(81)が自転車で680kmの走破を目指した	2.75	4.72
A5	65歳の男性がエベレストの登頂に成功した	2.76	4.82
A6	クジラとイルカが遊泳するシーンの撮影に成功した	2.82	4.72

表3 選ばれた暗い記事の概要と平均値

No.	記事概要	明るい	暗い
B1	乳児が変死した事件で父親を殺人容疑で逮捕した	4.73	2.58
B2	難民認定申請中のアフガニスタン青年が自殺した	4.67	2.58
B3	1歳の長男を餓死させたとして、母親を逮捕した	4.75	2.60
B4	交際相手の少女が産んだ新生児の遺体を放置した	4.86	2.68
B5	母親が生後10ヶ月の長男と無理心中を図った	4.89	2.70
B6	女子高生が殺害され、その容疑者が逮捕された	4.88	2.70

図2 明るくも暗くもない記事(両方の平均値が4.0以上)の分布

表4 選ばれた明るくも暗くもない記事の概要と平均値

No.	記事概要	明るい	暗い
O1	難民への支援物資を輸送した海上自衛隊の艦が帰港した	4.43	4.41
O2	某区立の小中学校で夏休み補習が始まった	4.31	4.43
O3	某政府がHP上にセックス依存度をチェックするテストを公開した	4.23	4.37
O4	雪の影響で新幹線の到着が遅れ、列車ホテルが用意された	4.43	4.19
O5	ダイヤモンド原石を腹の中に隠し持っていた男が逮捕された	4.61	4.32
O6	伊豆諸島の近海を震源とする地震があった	4.74	4.28

に分布しているかを図1に示してみる。すると、平均値が2.5～3.0の間に記事が集まっており、平均値が2.5以下の記事は少ないことがわかる。明るい記事の印象(平均値)と暗い記事の印象(平均値)はなるべく同程度である方がよいと、[「明るい」]に関しては2.54～2.82の6記事を選択し、[「暗い」]に関しては2.58～2.70の6記事を選択することにする。参考のために、選択された記事の概要と平均値を表2と表3に示す。

次に、印象語「明るい」と「暗い」の両方において印象の弱い記事(平均値が4.0以上の記事)がどのように分布しているかを図2に示してみる。すると、該当する

記事が11個あり、「明るい」に関しては4.23～4.83に、「暗い」に関しては4.03～4.41に分布していることがわかる。明るくも暗くもない記事であるためには、「明るい」と「暗い」に対する2つの平均値がなるべく同程度であることが望ましいため、本研究では、表4に示した6記事を用いることにする。

以上の方法で選ばれた明るい記事、暗い記事、明るくも暗くもない記事の計18記事を組み合わせて、表1に示したように、9個の記事ペアからなる記事グループを3つ作成した。

3. アンケート調査の実施

2章で決めた仕様にしながらアンケート調査を行った。具体的には、被験者300人(30代が150人、40代が150人でいずれも男女半々。予備実験に参加した900人とは重複しない)を3つのグループ(年齢・性別により均等に分け計100人)に分け、各グループにそれぞれ異なる記事グループ(表1参照)を割り当てた。各グループの被験者は、ランダムに提示される記事ペアの先行記事と後続記事を順に読み、各記事の印象を、「(明るい／楽しいという印象を感じる(1点)－割と感じる(2点)－やや感じる(3点)－どちらも感じない(4点)－(暗い／悲しいという印象を)やや感じる(5点)－割と感じる(6点)－感じる(7点)」という7段階の印象尺度を用いて評価した。

このアンケート調査の結果、5,400件分のデータ(9個の記事ペア×2記事/ペア×100人×3回)が得られた。

4. 調査結果の分析と考察

本章では、先行記事の印象によって後続記事の印象がどう変わるのかを分析する。具体的には、3章で得たデータを後続記事ごとに分類し、それぞれにおいて3個の後続記事の間で平均値の差の検定[11](有意水準5%)を行った。その結果、全27件(後続記事9個× ${}_3C_2$ 組)のうち3件(11.1%)において統計的に有意な差が観測された。但し、各検定の効果量 d [12]は、上から0.36, 0.29, 0.42と若干低めであった。この3件の組合せを表5に示す。

表5によれば、先行記事と後続記事との印象(平均値)の差(印象ギャップ)を大きくする方向に後続記事の印象がシフトしているのがわかる。すなわち、「 $B1 \Rightarrow A3$ 」と「 $A1 \Rightarrow A3$ 」や「 $B1 \Rightarrow A3$ 」と「 $O1 \Rightarrow A3$ 」では「 $B1 \Rightarrow A3$ 」の後続記事A3の方が「明るい／楽しい」という印象が強く、「 $A1 \Rightarrow A3$ 」や「 $O1 \Rightarrow A3$ 」の後続記事A3の印象は弱い。同様に、「 $A6 \Rightarrow O5$ 」と「 $O3 \Rightarrow O5$ 」では「 $A6 \Rightarrow O5$ 」の後続記事O5の方が「暗い／悲しい」と

表5 有意水準5%で統計的に有意な差が観測された記事ペアの組合せ

記事G 番号	先行記事			後続記事		
	記事 番号	平均 値	標準 偏差	記事 番号	平均 値	標準 偏差
1	B1	6.19	0.96	A3	2.23	1.10
2	A1	2.88	1.09	A3	2.63	1.14
1	B1	6.19	0.96	A3	2.23	1.10
3	O1	4.19	0.80	A3	2.55	1.11
2	A6	2.49	1.12	O5	4.40	0.98
3	O3	3.88	1.10	O5	4.00	0.92

いう印象が強く、「O3⇒O5」の後続記事O5の印象は弱い。

以上のことから、多くの場合において先行記事の印象は後続記事の印象に影響を与えないものと考えられるが、たまに影響を与えることがあり、そのようなときは印象ギャップが広がる方向に作用することが分かった。

5. まとめ

本論文では、300人によるアンケート調査を行い、その結果に基づいて、5,400件分のデータ(先行記事と後続記事からなる記事ペア9個(計18記事)×被験者100人×3回)を収集した。このデータを分析した結果、統計的に有意な差(有意水準5%)が出たのは、全27件のうち3件のみ(11.1%)であったが、この3件では、先行記事の印象と後続記事の印象との差(印象ギャップ)が大きくなっているのが観測された。したがって、多くの場合において先行記事の印象は後続記事の印象に影響を与えないものと考えられるが、影響を与える場合には、印象ギャップが広がる方向に作用することが分かった。

今後の課題としては、以下の点が挙げられる。まず、印象評価という主観的な行為を分析するには、記事ペアの数が9個と少なかったため、これをもっと増やしたい。また、今回は2個の記事を組み合わせたが、3個以上の記事を組み合わせることにより、文脈効果をより強調した結果が得られるかもしれない。この点についても検討したい。

謝辞

本研究の一部は、科研費MEXT/JSPS(24500134)ならびに福田将治奨学寄付金(52303)による助成の成果であり、ここに記して謝意を表す。

参考文献

- [1] Bo Pang and Lillian Lee. Opinion mining and sentiment analysis. *Found. Trends Inf. Retr.*, Vol.2, No.1-2, pp.1-135, January 2008.
- [2] Kevin Hsin-Yih Lin, Changhua Yang, and Hsin-Hsi Chen. Emotion classification of online news articles from the reader's perspective. In *Proc. of the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, pp.220-226, Washington, DC, USA, 2008. IEEE Computer Society.
- [3] 清水浩平, 萩原将文. 名詞と動詞の組み合わせに対する印象推定法. *日本感性工学会論文誌*, Vol.10, No.4, pp.505-514, 2011.
- [4] Tadahiko Kumamoto and Katsumi Tanaka. Web OpinionPoll: Extensive collection and impression-based visualization of people's opinions. In *Advances in Communication Systems and Electrical Engineering, Lecture Notes in Electrical Engineering*, Vol.4, chapter17, pp.229-243. Springer, USA, 2008.
- [5] 張建偉, 河合由起子, 熊本忠彦, 田中克己. 地域性に基づく発信者の観点差異を可視化するセンチメントマップシステムの提案. *情報処理学会論文誌データベース*, Vol.3, No.1 (TOD 45), pp.38-48, 2010.
- [6] 河合由起子, 熊本忠彦, 田中克己. 印象と興味に基づくユーザ選好のモデル化手法の提案とニュースサイトへの応用. *知能と情報*, Vol.18, No.2, pp.173-183, 2006.
- [7] 矢野絵美, 北野有亮, 末吉恵美, 篠原勲, ピンヤボン シニーナット, 加藤俊一. 消費者の感性モデルを利用したレコメンデーションシステムの構築. *情報処理学会論文誌:データベース*, Vol.44, No.SIG 8(TOD 18), pp.46-54, 2003.
- [8] 熊本忠彦, 灘本明代, 田中克己. 記事の印象を伝達するニュース番組生成システム wEE の設計と評価. *信学論 (D)*, Vol.J90-D, No.2, pp.185-195, 2007.
- [9] 石塚賢吉, 鬼沢武久, 加藤茂. 物語のシーンの印象に基づいた声楽曲の生成. *日本感性工学会論文誌*, Vol.10, No.4, pp.523-534, 2011.
- [10] 山本祐輔, 手塚太郎, アダムヤフト, 田中克己. Web-Alert: Web 情報の印象集約を利用した閲覧ページ内容に対する反対意見提示. *日本データベース学会論文誌*, Vol.7, No.1, pp.251-256, 2008.
- [11] P. G. ホーエル. 初等統計学, pp.172-176. 培風館, 東京, 1989.
- [12] Jacob Cohen, A Power Primer, *Psychological Bulletin*, Vol.112, No.1, pp.155-159, 1992.

(2012年5月31日 受付)

(2012年7月1日 採録)

[問い合わせ先]

〒275-0016 千葉県習志野市津田沼2-17-1

千葉工業大学 情報科学部 情報ネットワーク学科

熊本 忠彦

TEL: 047-478-0443

FAX: 047-478-0443

E-mail: kumamoto@net.it-chiba.ac.jp

著者紹介

くまもと ただひこ
熊本 忠彦 [正会員]

1988筑波大学第三学群情報学類卒。
1990同大学院理工学研究科修士課程了。
1990郵政省通信総合研究所(現、独立行政法人情報通信研究機構)。
2007千葉工業大学情報科学部情報ネットワーク学科准教授。
2010同大学教授。テキスト印象マイニングとその応用に関する研究に従事。
1996博士(工学)(筑波大学)。
FIT2004論文賞受賞。
電子情報通信学会、情報処理学会、人工知能学会、言語処理学会、日本データベース学会、日本感性工学会各会員。

Context Effects of Reading Order of News Articles on Their Impressions

by

Tadahiko KUMAMOTO

Abstract :

Recently, the number of studies on sentiment analysis has increased. These studies have attempted to determine whether people's impressions of news articles tend to be positive or negative. However, these studies have focused on people's impressions of individual articles and have not taken into account the effects of the reading order of articles. In this study, we conducted questionnaire surveys and examined how people's impressions of an article change in accordance with their impressions of preceding articles. The results showed that in many cases, reading order had no effect on people's impressions of news articles. However, in several cases, people's impressions of an article were enhanced after reading the preceding article.

Keywords : sentiment analysis, impressions of news articles, context effect, questionnaire survey

Contact Address : Tadahiko KUMAMOTO

Department of Information and Network Science, Faculty of Information and Computer Science, Chiba Institute of Technology
2-17-1, Tsudanuma, Narashino, Chiba, 275-0016, JAPAN
TEL : 047-478-0443
FAX : 047-478-0443
E-mail : kumamoto@net.it-chiba.ac.jp