

新聞社のニュース説明文における記事の要約特徴分析

川上沙耶加, 吉井理乃, 藤代裕之

法政大学

sayaka.kawakami.5f@stu.hosei.ac.jp, rino.yoshii.7q@stu.hosei.ac.jp, fujisiro@hosei.ac.jp

概要 SNS(ソーシャル・ネットワーキング・サービス)の普及により, 新聞社は SNS 上でニュースを配信するようになった。そのため, SNS 上で記事を紹介する投稿について, 記者は端的に記事の要約をするだけでなく, より多くの読者に記事を SNS の読者に届けるために記事要約の工夫を行う必要に迫られている。しかし, 従来の記事要約は文意を保ちながら記事を短くするものであり, 記者の届ける工夫は考慮されていない。そこで, 本研究では, 朝日新聞社のニュースサイトの記事と, その対となるツイートを比較することで, 記者の届ける工夫を明らかにした。その結果, 説明文には, ニュース記事を最小限に説明した部分と読者の目を引く表現部分があることが明らかになった。

キーワード SNS, ニュース, 文章要約, 新聞

1 はじめに

これまで新聞やテレビなど, マスメディアによって届けられてきたニュースは, SNS(ソーシャル・ネットワーキング・サービス)の普及により, 伝達方法が変化した。そのため, 新聞社は読者にニュースを選んでもらうため, SNS 上でニュース記事を的確に説明するための説明文(以降, ニュース説明文とする)の作成が必要となった。実際に, 朝日新聞のソーシャルメディア担当者は, ニュース説明文の作成について, 読者の興味を引くために, あえて結論を書かない, インパクトのある言葉を選ぶ, などさまざまな届ける工夫を行っていると述べている。

このように, 新聞社はニュース説明文の作成において, SNS 上での拡散, リンクから記事本文へのアクセス誘導をするなど, よりニュース記事本文を読者に読んでもらうことを考えている。従来のニュース記事の要約は, 記事の内容を端的に伝えることを求められていたが, それでは, 新聞社の考えているニュース説明文の作成は困難である。

そこで本研究では, 新聞社がより読者に訴求するニュース説明文の生成を最終目標とする。本論文では, ニュース説明文の作成における記事本文の要約に, SNS 上での拡散性や, リンクから記事本文へのアクセスを誘導するために記者が施している工夫に何かしらの傾向・法則があるのではないかと仮説から, ニュース説明文の作成において記者がどのように記事本文から要約しているのか, その工夫を明らかにすることを目指す。

2 先行研究

文章要約に関する研究は内容を端的に示す手法が検討されてきた。山田らは, ニュース記事には報道する話題特有の単語や句が用いられることが多いという特徴を考慮し, 本文中で話題特有の表現が使われている部

分を重要文として抽出し, 要約を生成する手法を提案した[3]。畑山らは, 文の中で最も重要な意味をもつ動詞を選定する際, 日本語の新聞記事によく見られる表現を考慮した文要約手法を提案した[4]。

これらの研究は, ニュース記事の特徴を用いて, 記事の重要箇所を特定しているが, 記者がどのように記事本文を要約しているか, その工夫については考慮されていない。

興梠らは拡散を考慮した要約について研究を行った。ニュースサイト「ハフィントン・ポスト日本版」のツイート調査から, 拡散されるニュース説明文の法則を明らかにしている。しかし, 興梠らの手法では, 同じニュース記事に関するニュース説明文同士の fav(お気に入り)数, RT(リツイート)数の比較により, どのような表現がより読者の関心を惹起するのかわかるということ, また記者がどのように要約をしているのかわかるということ, は明らかになっているが, 記者がニュースを届けるための工夫がどのように行われているのかは明らかになっていない。

そこで本研究では, SNS 上でより多くの読者にニュースを届けるため, ニュース説明文作りにおいて, 記者がどのように要約をしているのか, その工夫を明らかにする。

3 調査

本研究では記者の届ける工夫を明らかにするため, ニュース記事とその記事に対応するツイートを比較し, ニュース説明文が記事本文からどのように要約されているか, 調査を行った。

3.1 調査手法

調査には, 朝日新聞社のニュースサイトである「朝日

新聞デジタル¹」の記事、及び朝日新聞社の公式 Twitter アカウント@asahi のツイートを用いた。

収集は、2016 年 12 月 1 日から同月 18 日の間に毎日 13 時に行い、朝日新聞デジタルのトップニュースに掲載されたニュース記事と、対応するニュース説明文を収集した。

それぞれのニュース記事は、新聞社がカテゴリ別に分けており、今回使用した 145 記事の内訳は、政治 14 記事(9.6%)、経済 11 記事(7.5%)、国際 29 記事(20%)、社会 81 記事(55.8%)、スポーツ 5 記事(3.4%)、その他のカテゴリのものが 6 記事(4.1%)あった。また、社会のカテゴリに区分される記事はジャンルが幅広く、主な内訳としては、事件・事故・災害に関する記事が 39.5%、その他・話題に関する記事が 33.3%で、その他にはカルチャーや皇室、医療・健康・福祉、教育・子育てなどの多様な記事が含まれている。

3.2 調査結果

調査の結果、ニュース説明文には、そのニュースがどのような内容なのかを最小限に説明した従来の要約文と、その他の文、もしくは単語が組み合わされていた。ここでは従来の要約文のことを概要とする。

ニュース説明文がニュース記事の端的な要約である概要ばかりのものが全体の 20%であった。80%は概要以外の文もしくは、単語が概要に組み合わされていた。

概要以外の文もしくは単語の例としては、「1380 円の品も」や「30 年で 4 割減」などのように、具体的な数値を含んだ表現、また具体的な数値を用いなくても「首位」や「長年」などの程度を表現しているものが、全体で 13.7%あった。これらのカテゴリ別の内訳は、社会 55%、経済 20%、スポーツ 15%、国際・その他がそれぞれ 5%であった。

また、インタビュー記事でない記事のニュース説明文であっても「何の恨みが…」や「逃走の男、知らない」などの具体的な発言の表現を用いているものが、全体で 17.2%あった。内訳としては、社会 68%、国際 20%であった。

まず、ニュース説明文がニュース記事の端的な要約であるものが、全体の 20%であった。カテゴリ別の内訳は、国際 27.5%、経済 27.2%、社会 20%、政治 7.1%となっている。また、社会のカテゴリに区分される記事は幅広く、その内訳を詳しく見てみると、事件・事故・災害に関する記事が 88.2%であった。

4 考察

また、国際、経済、社会の中でも、事件・事故・災害といった社会に大きく影響のあるニュースに関しては、ニ

ュース説明文がニュース記事の要約のみである傾向がある。これは、ニュース説明文という短い文の中で、伝えるべき情報が多い記事など、ニュースの内容に応じて、読者により惹起することより、ニュース記事を的確に説明するための説明文という、ニュース説明文としての役割を優先していることがわかる。

概要以外の箇所における、記者工夫としては、記事の中から具体的な数値や程度、発言などの中でもインパクトのあるものをニュース説明文に用いていることがわかった。

これらのことから、記者はニュース説明文という短い文の中で、ニュースの内容に応じて必要最小限でニュース説明文本来の役割を果たしつつ、インパクトのある表現を用いて読者への訴求を図っていることが明らかになった。

5 終わりに

本研究では、ニュース記事と対応するニュース説明文を照らし合わせ、記者はニュース説明文の作成において記事本文からどのように要約しているのか、その工夫を明らかにした。これにより、記者は本文を要約する際、「概要:ニュース記事の内容を、最小限に説明した箇所」とそれ以外の箇所が含まれていることがわかった。

今後、概要以外の箇所で、どのような表現がされているのか、より詳しく調査することで、ニュース説明文における記者の記事要約における工夫を明らかにし、既存の自動要約技術と組み合わせることで、記者の工夫を反映したニュース説明文の自動生成に貢献できると考える。

参考文献

- [1] 興梠紗和, 木村昭悟, 藤代裕之, 西川仁: SNS 上で拡散する web ニュース説明文の調査と自動選択, 電子情報通信学会論文誌(D), vol.J99-D, no.4, pp.403-414, 2016.
- [2] 奥村学: ソーシャルメディアを対象としたテキストマイニング, 電子情報通信学会 基礎・境界サイエンス, vol.6, no.4, pp.285-293, 2012.
- [3] 山田一郎, 柴田正啓: ニュース記事の定型性を利用した話題要約の検討, 情処学自然言語処理研資, pp45-50, 2002.
- [4] 畑山満美子, 松尾義博, 白井論: 重要語句抽出による新聞記事自動要約, 自然言語処理, vol.9, no.4, pp55-pp73, 2002.
- [5] 奥村学, 難波英嗣: テキスト自動要約に関する研究動向(巻頭言に代えて), 自然言語処理, vol.6, no.6, pp1-26, 1999.

¹<https://www.asahi.com/>