

## 構文情報に着目したキャッチコピーの自動生成

西村 百之輔<sup>b</sup>    木村 優平<sup>b</sup>    片倉 多智<sup>b</sup>    戎 淳<sup>b</sup>    早坂 絵央<sup>b</sup>    延澤 志保<sup>\*</sup>  
<sup>b</sup>東京都市大学知識工学部                      <sup>\*</sup>東京都市大学情報工学部

### 1 研究背景

キャッチコピー自動生成を目的として、山根らは対象物とそれに付随する性質の品詞が違う2単語を入力して、Web上からそれらがいった文章を抽出することで候補を生成する手法を提案した[1]。山根らの手法では、日本語彙大全を使って入力した対象物と性質の2単語とより関連度の高い単語を抽出し、それらの単語の入れ替えによってキャッチコピーの候補の出力をしている。また山根らは、言葉の使われ方が日々変化することに着目し、Web情報を用いてよりタイムリーな語句を抽出してキャッチコピーに適用している[1]。

幅野らは、映画のあらすじに基づいてキャッチコピーの型を選択すると同時に、レビュー文から単語を抽出し、キャッチコピー型と単語を虫食い算のように組み合わせてキャッチコピーを出力する手法を提案している[2]。

これらの先行研究のように、キャッチコピーの生成では、キャッチコピーの型と語句の距離とを活用することが有効と考えられる。

### 2 キャッチコピー自動生成

キャッチコピーでは、伝えたい内容だけでなく、その印象が重要である。本研究ではキャッチコピーの型に着目したキャッチコピーの自動生成を目指す。

#### 2.1 キャッチコピーの特徴

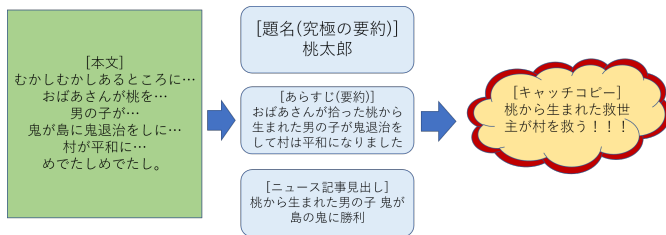


図 1: 題名と要約, 見出し, キャッチコピーの比較

文章の特徴を示す最も短い文のものは題名だと考えられる。例えば昔話「桃太郎」では、主人公の名前自体が「桃から生まれた男児」という特徴を示すことからこれだけで題名とされるが、これはキャッチコピーとしては短すぎる(図1)。それに対して、あらすじはキャッチコピーとしては文が長い。またこれをニュース記事の見出しの形に要約すると、事実関係の列挙が基本となり、キャッ

**Catchphrase Alchemy Considering the Sentence Structure.**  
 Momonosuke Nishimura<sup>b</sup>, Yuhei Kimura<sup>b</sup>, Taichi Katakura<sup>b</sup>,  
 Bo Rong (戎 ボツ)<sup>b</sup>, Kaio Hayasaka<sup>b</sup>, and Shiho Hoshi Nobesawa<sup>\*</sup>.  
<sup>b</sup> Faculty of Knowledge Engineering, Tokyo City University  
<sup>\*</sup> Faculty of Information Technology, Tokyo City University

チコピーとしては弱い。キャッチコピーは、図1に示すように、より人を惹きつけるような文章とする必要がある。本研究では、このようなキャッチコピーの特殊性、独創性を再現することを目的とする。

#### 2.2 キャッチコピーの型

キャッチコピーがどのような品詞の構成なのかを調査するためキャッチコピー5,689個の構成要素を調べた(表1)。表1に挙げたとおり、名詞が約3割、動詞が約1割、

表 1: キャッチコピーの構成要素

名詞	動詞	その他	計
18,201(26.6%)	10,096(14.7%)	40,075(58.7%)	68,372

その他形容詞等が約6割であることがわかった。名詞が動詞の3倍使用されている。句読点や数字が多く含まれているため、その他が多くなっている。

#### 2.3 提案手法の概要

本研究では、キャッチコピー生成の元とする文章(説明文)を入力としてとり、名詞と動詞の2方向からアプローチしたキャッチコピーを出力する(図2)。提案手法では、対象の説明文から名詞句と動詞句を抽出し、名詞句からキャッチコピーに挿入すべき特徴語句を、動詞句部分から文の型となるテンプレートを選定、そのテンプレートに特徴語句を適用してキャッチコピーを生成する。事前に行ったキャッチコピーの分析から、名詞が3割に上ることがわかっている(表1)。またどの品詞よりも割合が多い。このことからキャッチコピーを作る上で名詞が重要であると仮定し、提案手法では名詞句に着目する。まず、説明文を入力すると、名詞句と動詞句をそ

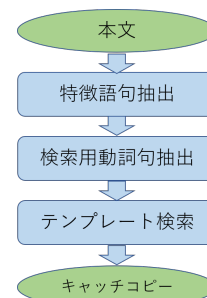


図 2: キャッチコピー生成の流れ

れぞれ抽出する。キャッチコピー型を選択するための基本語句は動詞句とし、特徴語句として名詞句を用いる。

#### 2.4 説明文からの特徴語句の抽出

提案手法では、要約のため、MeCabを用いて説明文の形態素解析を行い、頻出語句を特徴語句とする。

2.5 キャッチコピーテンプレートの選択

キャッチコピーは動詞句を基に選択する。頻出の動詞がその文章の中で、キャッチコピーにより有効だと判断し、それをを用いる。ここでは Web からジャンルをランダムに集めてきた 5,689 個のキャッチコピーを使用する。例として、表 2 に 2 つ挙げる

表 2: 実験に用いるキャッチコピー群

キャッチコピー	名詞句	動詞句
くすり指がドキドキ <span style="border: 1px solid black;">した</span> 。	1	1
ただのスポーツクラブでは <span style="border: 1px solid black;">ない</span> 、リゾートでも <span style="border: 1px solid black;">ある</span> 。	3	2

表 2 の下線部は名詞句、四角で囲った部分は動詞句を示す。表 2 に示す通り、名詞と動詞が 1 つずつのもあれば、複数個の組み合わせもある。この個数にも着目してみていきたい。

2.6 キャッチコピーテンプレートの適用

特徴語句は虫食いの状態にあるテンプレートに頻出順に先頭から挿入していく。先頭に重要な名詞を配置することで、よりキャッチーになる可能性を考慮している。

3 実験結果

本研究では研究室紹介文章 (図 3) を元にキャッチコピー生成を行った。図 3 に示したとおり、ここで入力とする

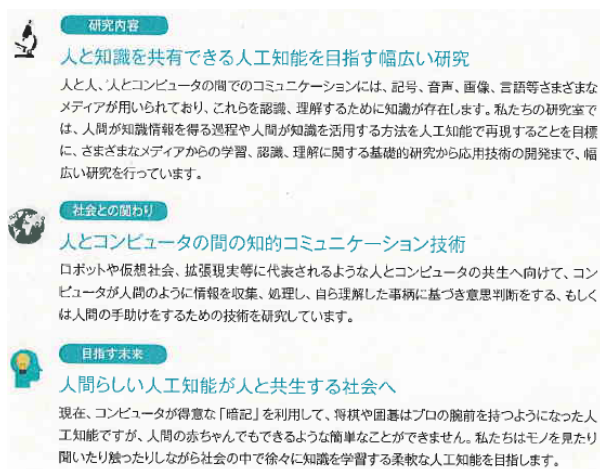


図 3: 研究室紹介の例

研究室紹介文は文章量が多いという特徴がある。これをよりキャッチーで、より簡潔なキャッチコピーの自動生成が目標である。

この本文では、動詞は「する」が一番多く使われている。あてはまるテンプレートの一部を表 3 に挙げる。

表 3: キャッチコピーのテンプレート例

キャッチコピー	名詞句	動詞句
ひとをあたためて、美しく <span style="border: 1px solid black;">する</span> 。	1	1
受験は家族を強く <span style="border: 1px solid black;">する</span> 。	2	1
窓は日常を文学に <span style="border: 1px solid black;">する</span> 。	3	1

これらを用いてキャッチコピー自動生成実験を行った結果を表 4 に示す。結果としては意味不明な文は出てこ

表 4: 実験結果

キャッチコピー	名詞句	動詞句
人工知能をあたためて、美しく <span style="border: 1px solid black;">する</span> 。	1	1
人工知能は人を強く <span style="border: 1px solid black;">する</span> 。	2	1
人工知能は人を人間に <span style="border: 1px solid black;">する</span> 。	3	1

なかった。また、「人工知能は人を強くする」といった、ポジティブなものも生成することができた。形容詞をそのまま残しているの、意外な組み合わせの文章が出てきていてよりキャッチーになった。

表 4 の 3 文目のようにテンプレートの名詞の数が多くなればなるほど、意味が通りづらくなる印象がある。これに、読点を含む 2 文のようなテンプレートの時の対処法も考えたい。

4 まとめ

キャッチコピーは伝えたい内容だけでなくその印象が重要である。本研究ではキャッチコピーの型に着目したキャッチコピーの自動生成を目指す。提案手法では、対象の説明文を要約し、動詞句部分からテンプレートを選定、名詞句に特徴語句を適用してキャッチコピーを生成する。事前に行ったキャッチコピーの分析から、名詞が約 6 割、動詞が約 3 割、その他形容詞等が約 1 割であることがわかっている。このことからキャッチコピーを作る上で名詞が重要であると仮定し、提案手法では名詞句に着目した。

参考文献

- [1] 山根宏彰, 萩原将文, “Web 情報を用いたキャッチコピー自動生成システム,” 日本知能情報ファジィ学会ファジィシステムシンポジウム, vol.28, pp.1008–1013, 2012.
- [2] 幅野裕貴, 浦谷則好, “あらすじとレビューを用いた映画のキャッチコピー支援手法の提案,” 言語処理学会第 20 回年次大会, pp.864–867, 2014.