

# ギャップにより面白さを生むストーリーの生成に向けた単語の自動補完

## Automatic Word Completion for Creating Interesting Story with Gaps

河村 芽依<sup>1\*</sup> 砂山 渡<sup>1</sup> 畑中 裕司<sup>1</sup> 小郷原 一智<sup>1</sup>

Mei KAWAMURA<sup>1</sup> Wataru SUNAYAMA<sup>1</sup> Yuji HATANAKA<sup>1</sup> Kazunori OGOHARA

<sup>1</sup> 滋賀県立大学工学部

<sup>1</sup> School of Engineering, The University of Shiga Prefecture

**Abstract:** In everyday life, there are many opportunities to communicate with people. Sometimes I want to make a story that is interesting and interesting to others. The purpose of this research is to create a system that presents a story to be created with gaps of fun to support story generation. The story is created by using the relevance of the word and acquiring the related word of the input word and combining it with the prepared template sentences. The texts prepared by the system were evaluated from two viewpoints, fun and betrayal of expectations, and a system evaluation experiment was performed by asking what kind of correction could be made to make it even more interesting. Experiments have confirmed that the generated story can generate a certain amount of fun. In addition, after generating a story with the system, we propose more specific words as correction candidates, and we want to aim for a more satisfactory story by selecting final words by humans.

## 1 はじめに

日常生活では、人とコミュニケーションを取る機会は多く、親しい関係の人との会話ややりとりの中では、相手に話への興味や関心を持って、面白いと思ってもらえる話を作りたいと考える。相手に自分の思うことを伝えようとする時、どのようなことをどのような流れの内容か、自分の発想力やひらめき、知識を基に考えることを求められる。しかし、相手に伝えたいことはあるが、どんな言葉で表せばよいかうまく思い浮かばず、伝えられないこともある。そもそも、相手に面白いと思わせることを、難しいと思う人もいる。そのような場面で、ストーリー生成の支援をシステムによって行いたいと考える。

そこで本研究では、予想とのギャップによって面白さを生むストーリーをテンプレート内の単語の自動補完によって作成するシステムの構築を目的とする。単語を、用意したテンプレートに自動で補完することによって、ストーリーを形成し、そのストーリーを提案することで、支援を行う。本研究では、期待はずれの場面で起こる不調和のタイプの面白さについて扱う [1]。

形成するストーリーは、面白さを生み出す内容にすることによって、相手の興味や関心を得ようとしている。また、この提案を行うシステムを作成し、実際に日常的なやりとりの中で利用したり、会話の想定や練習の機会として、利用できるようにしたい。

以下本論文では、2章で関連研究について述べる。3章で単語の自動補完によるストーリー生成システムについて述べる。4章で本システムの評価実験について述べ、5章で本論文を締めくくる。

## 2 関連研究

ボケとツッコミの対話形式で笑いを生み出す文章自動生成システムに関する研究がある [2][3]。予測される事柄をわざと誤ることでおかしさを生み出しており、誤りの単語は Google N-gram を用いて、入力文の前後を考慮している。また、文章を入力して、その一部の単語を面白みのある単語に変換することで、漫才の台本となる文章の生成を行っている。

面白いフレーズの自動生成に関する研究がある [4]。この研究では、身近にあることわざを利用して単語を置き換えることにより、ユーザの予想を裏切って笑いを生

\*連絡先：滋賀県立大学工学部 電子システム工学科 河村 芽依  
〒522-8533 滋賀県彦根市八坂町 2500  
E-mail: os23mkawamura@ec.usp.ac.jp

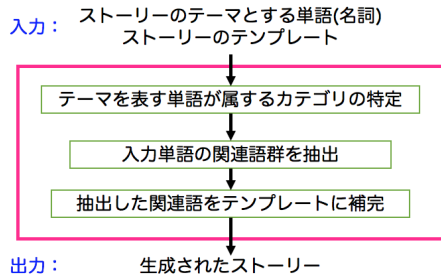


図 1: 単語の自動補完によるシステムの構成

み出している。ことわざを利用することによって、次はこうなるという予想を容易にさせることができ、音の長さや音韻の類似などを考慮して、単語を変換している。

テーマとキーワードからアイデアの発想を支援するシステムに関する研究がある [5]。この研究では、テーマやキーワードをそのまま発想支援の入力とすると、その要素に固執した発想しか得られない可能性を指摘している。これを解決するために、入力とするキーワードの共起関係を用いて関連する語を取り出し、関連性はあるが、共起度は低い要素を抽出し、発想のヒントとして提示している。このように、人のアイデアでは、発想し難い単語を取り出している。

本研究では、入力としてユーザが指定する単語のみとし、作ることができるストーリーの多様性を考慮する。また、単語の音や構造を考慮するのではなく、単語の持つ意味や状態などの情報を考慮して、意外性のある単語を取り出す手法を試みる。

### 3 単語の自動補完によるストーリー生成システム

#### 3.1 システムの構成

本研究で提案する単語の自動補完によるストーリー生成システムの構成を図 3 に示す。

まず入力として、ストーリーのテーマとする名詞の単語と、ストーリーのテンプレートを受け取る。入力された単語がどのカテゴリに属しているか特定を行い、カテゴリを決定する。これらの入力された単語と特定されたカテゴリを基に、関連語を抽出し、あらかじめ用意したテンプレートに補完することによってストーリーを生成する。

##### 3.1.1 ストーリーのテーマとする単語

入力として、ストーリーのテーマを表す 1 単語を受け取る。本研究では、話題の中での利用のしやすさを考

表 1: ストーリー生成用テンプレート

A 「[特徴 1(形容詞)][カテゴリ]って何やと思う？」 B 「一番に思いつくのは、[入力単語] やなあ。」 A 「でもそれは、[特徴 2(動詞)] ものではないねん。」 B 「なら [単語 A] やろ。」 A 「いや、[特徴 3(動詞)] もでもないねん。」 B 「じゃあそれは一体なんやねん。」 A 「実は、[単語 B] なんや。[特徴 4(形容詞)] ものやねん。 [入力単語] や [単語 A] より、[単語 B] の方が [特徴 4(形容詞)] やん。」
A 「[特徴 1(形容詞)] ものって何やと思う？」 B 「それで思いつく [入力単語のカテゴリ] は [入力単語] なんちゃう。」 A 「いや、[入力単語とは異なるカテゴリ] のことやから [単語 C] やねん。 [単語 C] こそが、[特徴 1(形容詞)][入力単語とは異なるカテゴリ] や。」
A 「友達が好きな [カテゴリ] があるらしんやけど、 その名前を忘れたみたいやねん。 いろいろ聞いても、全然名前がわからへんねんて。」 B 「じゃあ、その友達の良い [カテゴリ] を考えてあげるから、 特徴を教えてよ。」 A 「確か [特徴 1(形容詞)] もので [特徴 2(動詞)] って言った。」 B 「[入力単語] じゃない、それ？ その特徴は完全に [入力単語] や。」 A 「自分もそうやと思ってんけど、[特徴 7(形容詞)] ものらしいねん。」 B 「なら、[入力単語] ではないな。 [入力単語] は [特徴 7(形容詞)] ものじゃないもんね。」 A 「それで [特徴 5(形容詞)] もので [特徴 6(動詞)] っのも言った。」 B 「それはやっぱり [入力単語] やろ？ その特徴はやっぱり [入力単語] や。」 A 「でもな、[特徴 8(動詞)] ものらしいねんて。」 B 「なら、[入力単語] ではないな。 [特徴 7(形容詞)] ので [特徴 8(動詞)] なら、もう [単語 D] でええやろ。」

慮して、日常生活で現れやすい単語のカテゴリを 8 つ用意し、そのカテゴリのいずれかに当てはまる単語を入力する。本研究では用意したカテゴリは、洋菓子、動物、スポーツ、植物、食べ物、遊び、日用品、施設の 8 つとなっている。

##### 3.1.2 ストーリーのテンプレート

ストーリーのテーマとする入力単語と、その関連語を用いてギャップが生じるストーリーとしてテンプレートを 3 つ用意した。表 1 に用意した 3 つのテンプレートを示す。

これらのストーリーは、入力された 1 単語をテーマとして、連想ゲームでの「○○といえば」というやりとりをする内容となっている。例えば、「甘いもの」といえば何だろう、といった時、マカロンやクッキーなどが想像できる。しかし、実際は別のものだった、という形を利用して、マカロンやクッキーといった予想を裏切っている。

各テンプレートは、予想を裏切るパターンが異なり、使用する関連語および使い方によりパターンを形成している。テンプレート 1 は、単語を表す特徴に着目したやりとりを行い、テンプレート 2 は、単語のカテゴリに着目したやりとりを行い、テンプレート 3 は、入力された単語に着目したやりとりを行っている。

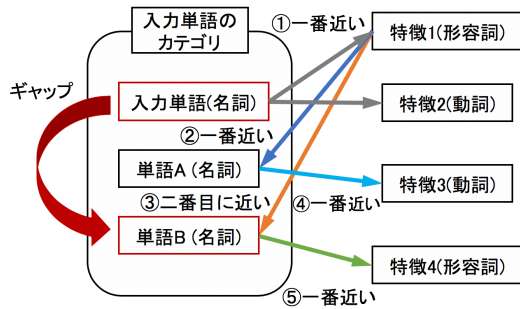


図 2: テンプレート 1 に使用する関連語

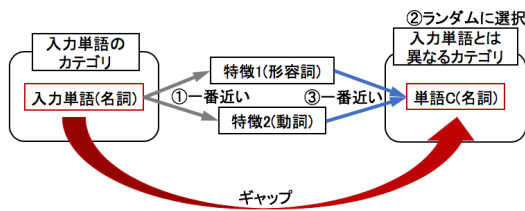


図 3: テンプレート 2 に使用する関連語

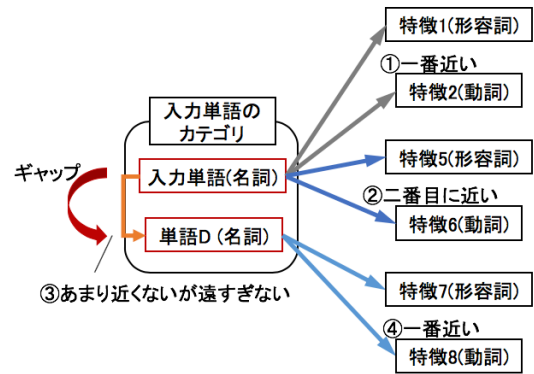


図 4: テンプレート 3 に使用する関連語

**ストーリー生成結果**

A 「【凄い】【スポーツ】って何やと思う？」  
 B 「一番に思いつくのは、【野球】やなあ。」  
 A 「でもそれは、【投げる】ものではないねん。」  
 B 「なら【カバディ】やる。」  
 A 「いや、【鍛える】ものでもないねん。」  
 B 「じゃあそれは一体なんやねん。」  
 A 「実は、【気球】なんや。【暑い】ものやねん。  
 【野球】や【カバディ】より、【気球】の方が【暑い】やん。」

図 5: 入力単語「野球」をテンプレート 1 に補完した時の生成例

### 3.2 単語の表現

本研究では、単語の分散表現を利用して単語間の関係を比較する。単語の分散表現を得る手段として fast-Text<sup>1</sup>を利用している。単語のベクトルを 300 次元とし、インターネット上での百科事典と言える Wikipedia<sup>2</sup>の全データの文章を利用して、学習を行う。

単語ベクトルを取得できることによって単語間の cos 類似度を求めることができ、単語同士の意味的な関係を捉える。

### 3.3 関連語の導出

それぞれのテンプレートに使用する関連語の関係図を図 2, 図 3, 図 4 に示し、テンプレート 1 の関連語の抽出の順序を以下に示す。

1. 入力単語との cos 類似度が最も大きい形容詞 (特徴 1) および動詞 (特徴 2)
2. 入力単語のカテゴリ内の単語で、特徴 1 の形容詞との cos 類似度が一番大きい単語 (単語 A)
3. 特徴 1 の形容詞との cos 類似度が 2 番目に大きい単語 (単語 B)
4. 単語 A との cos 類似度が最も大きい動詞 (特徴 3)

<sup>1</sup><https://fasttext.cc>

<sup>2</sup><https://ja.wikipedia.org/wiki/メインページ>

5. 単語 B との cos 類似度が最も大きい形容詞 (特徴 4)

この順序で関連語を全て抽出することができたら、テンプレートの決められた箇所にそれぞれの単語を当てはめて、補完を行う。また、赤色で囲まれた単語 (入力単語, 単語 B) を裏切ることによってギャップを生み出している。

### 3.4 システムの出力例

図 5 に、入力単語を「野球」としてテンプレート 1 に補完した時のストーリー生成結果を示す。野球という単語から、カテゴリを「スポーツ」と特定している。ギャップを生み出しているのは、「野球」ではなく「気球」だったという単語の裏切りである。

### 3.5 システムのインターフェース

この単語の自動補完によるストーリー生成システムのインターフェースは Python で扱っている。本システムを作成するにあたり、Python からインターフェースを

構築するための標準ライブラリである Tkinter を使用している。単語を入力および、テンプレートを選択すると、用意したカテゴリに入力単語が含まれている場合、カテゴリを特定する。この場合、含まれるカテゴリが複数ある場合は、ユーザーに選択してもらう。入力単語を基に、形容詞および動詞の cos 類似度が高い 10 単語を表示させる。そして、入力単語から各関連語を抽出し、選択されたテンプレートに抽出した関連語を当てはめることによって生成されたストーリーも表示させる。

## 4 ストーリーの生成に向けた単語の自動補完システムの評価実験

本研究で提案したシステムの評価をするため、評価実験を行った。その方法と実験結果を示す。

### 4.1 実験方法

評価実験の方法として、まず単語を自動補完しストーリーを生成するシステムを用いて、3つのテンプレートを使ってそれぞれ5つずつ、全部で15個のテキストを用意した。そして、被験者として大学生、大学院生11名にそれぞれのテキストを読んでもらい、5段階の評価を行ってもらった。評価してもらう項目は、AとBの会話のやりとりを面白いと感じたか、AとBのどちらかが予想を裏切られているかの2つである。これらの項目から、本システムでギャップを生むことができているか、それによりおもしろさを生むことができているのかを検証する。また、用意されたテキストをより良くするための修正を1単語単位で行ってもらった。これにより、さらに面白いストーリーに作り変える形で、実際にストーリー生成に向けた支援ができているのか、検証する。

### 4.2 面白さと裏切りの評価についての結果

表2に、評価実験に用いた15のテキストの面白さおよび裏切りの評価の順位を示す。この表より、面白さの評価について、全テキストにおいては、テンプレート1,3のテキストが上位を占めているのに対して、テンプレート2のテキストは下位を占めていることがわかる。裏切りの評価についても同様に、テンプレート1,3のテキストが上位を占めているのに対して、テンプレート2のテキストは下位を占めていることがわかる。特にテンプレート2は文章量も他の二つに比べて少ないため、情報量や予想を裏切る度合いも小さく感じられたことも考えられる。面白さの評価に比べて、裏切りの評価の方が全体的に高くなっている。よって、裏切りの評価が高いにも関わらず、面白さの評価が低いものもある。こ

表 2: 15 の文章の面白さおよび裏切りの評価の順位

面白さの評価			裏切りの評価		
順位	テキスト	評価点	順位	テキスト	評価点
1	3-4	3.82	1	1-1	4.55
2	1-3	3.55	2	1-2	4.36
2	1-5	3.55	3	1-3	4.27
4	3-1	3.45	4	2-1	4.27
5	3-3	3.36	5	2-4	3.82
6	2-1	3.27	6	3-1	3.82
7	3-5	3.09	7	3-4	3.82
8	1-1	2.91	8	2-2	3.73
8	3-2	2.91	9	3-3	3.73
10	1-2	2.82	10	3-5	3.64
10	1-4	2.82	11	1-5	3.55
10	2-4	2.82	12	1-4	3.18
13	2-2	2.64	13	3-2	3.18
13	2-5	2.64	14	2-3	2.45
15	2-3	2.55	15	2-5	2.36

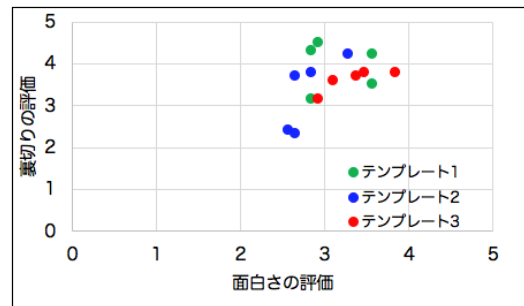


図 6: 面白さと裏切りの評価の関連性

れらより、ただ裏切るだけではなく、ギャップを生む名詞の単語間の関係性は一定の距離を考慮するための新たな抽出条件が必要と考えられる。

図6に、面白さと裏切りの評価の分布を示す。本システムは、ギャップを生み出して面白さを感じさせるため、面白さと裏切りの評価の間には、比例の関係である必要がある。また、本実験での二つの評価に対する相関係数は0.36であった。このことから、面白さと裏切りの評価の間には弱い正の相関があるといえる。本研究で用いたテンプレートでは、予想が裏切られたテキストは面白さを生み出しているといえる。

実験で使用したテキストの中で、最も裏切りが面白さにつながった例を、図7に、面白さの評価と裏切りの評価を実験者ごとに棒グラフで示す。このテキストは表3に示すテキスト3-3となっている。実験者1, 8, 9を除いた実験者は、裏切りの面白さの評価が同等であることがわかる。「おもしろい」「遊ぶ」「転がす」「楽しい」の特徴から「ピリヤード」しかない、形容詞と動詞の関連語により、ギャップに関する名詞の単語を強く予想できていると考えられ、面白さにつながったと言える。また、「雪合戦」は「遊ぶ」「楽しい」といったピ

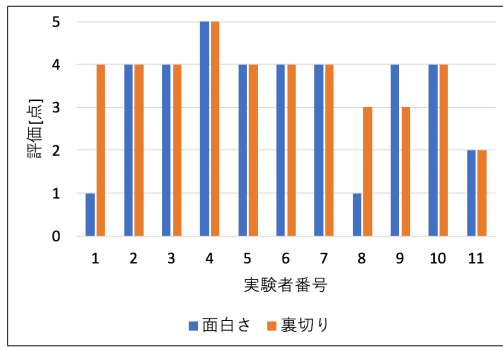


図 7: 裏切りが面白さにつながった例

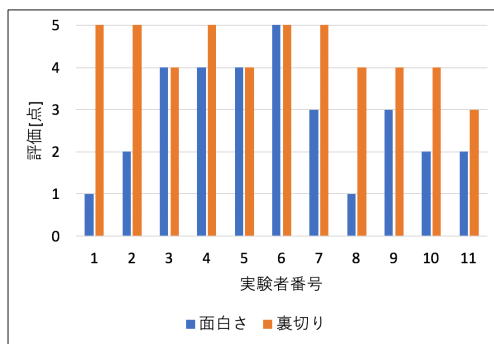


図 8: 裏切りが面白さにつながらなかった例

リヤードの特徴にも当てはまっているため、ビリヤードと雪合戦は遠い関係ではないといえる。

実験で使用したテキストの中で、最も裏切りが面白さにつながらなかった例を、図 8 に、面白さの評価と裏切りの評価を実験者ごとに棒グラフで示す。このテキストは表 3 に示すテキスト 1-2 となっている。裏切りの評価は全体的に高いが、面白さの評価は低いものが多いことがわかる。丸い動物で連想をしていき、特徴を否定されていくというストーリーだが、最後にジャガーときて裏切りはあるものの、はじめの「丸い動物」に対して違和感を感じて面白さの評価が低くなってしまったと考える。

よって、これらのことより、ギャップに関する二つの単語は一定の距離の関連度を考慮して抽出することが重要と考えられる。本システムにおいて、関連語を抽出する際、cos 類似度が大きいといったルールだけではなく、近すぎず遠すぎないといった、抽出条件を増やす必要があると考えられる。

表 3: 評価実験に用いたテキスト 1-2, テキスト 3-3

テキスト	
1-2	A 「丸い 動物 って何やと思う？」 B 「一番に思いつくのは、アザランやなあ。」 A 「でもそれは、泳ぐ ものではないねん。」 B 「なら、ブタ やろ。」 A 「いや、飼う もでもないねん。」 B 「じゃあそれは一体なんやねん。」 A 「実は、ジャガー なんや。速い ものやねん。」 アザラン や ブタ より、ジャガー の方が速い やん。」
3-3	A 「友達が好きな スポーツ があるらしんやけど、その名前を忘れたみたいやねん。いろいろ聞いても、全然名前がわからへんねん。」 B 「じゃあ、その友達の好きな スポーツ を考えてあげるから、特徴を教えてよ。」 A 「確か おもしろい もので遊ぶ って言った。」 B 「ビリヤード じゃない、それ？その特徴は完全にビリヤード や。」 A 「自分もそうやと思ってんけど、痛い ものらしいねん。」 B 「なら、ビリヤード ではないな。」 ビリヤード は 痛い ものじゃないもんね。」 A 「それで 転がす もので 楽しい ってのも言った。」 B 「それはやっぱり ビリヤード やろ？その特徴はやっぱり ビリヤード や。」 A 「でもな、滑る ものらしいねん。」 B 「なら、ビリヤード ではないな。」 痛い ので 滑る なら、もう 雪合戦 でええやろ。」

### 4.3 文章の修正の結果

用意したテキストを面白くするための修正をしてもう場合と、裏切るための修正をしてもう場合の二つを行った。修正されたパターンとして、同じ品詞の同カテゴリの単語、同じ品詞のより具体的な単語、の 2 つに分けることができる。表 4 に、面白くするための修正案、および裏切るための修正案を 2 つのパターンに分けた平均件数を示す。

面白くするための修正案と裏切るための修正案の件数が最も多いのは、テンプレート 2 であることがわかる。テンプレート 2 の評価において、面白さ、裏切り共に平均も低かったため、修正案も多くなったと考えられる。全体的に面白さの観点の方が、裏切りの観点からの修正案の件数が多かった。これは、面白さの評価が裏切りの評価より低かったため、面白くするための修正案を促されたと考えられる。

修正する観点について、関連語を同じ品詞の別の単語に言い換える修正が多いことがわかる。また、変換された名詞に関しては、より具体的な単語に言い換える修正が多く、ストーリーが具体的に変わり、確かに面白くなるといえる。具体的な単語に置き換えることによって、よりイメージしやすく、裏切りが強く感じられる。これらより、本システムにおいて、関連語を抽出し

表 4: 修正案の平均件数

修正する観点 テンプレート	面白さ			裏切り		
	1	2	3	1	2	3
同カテゴリの単語 [件]	3.8	4.2	5.6	3.4	4.4	2.2
具体的な単語 [件]	3.4	4.4	2.2	2.0	3.0	2.4

た後, さらにギャップに関する関連語の名詞の具体的な単語を, 取り出すことを可能にしたいと考える.

## 5 おわりに

日常的な会話やメッセージのやりとりにおいて, 面白さを生むストーリーの生成に向けた単語の自動補完により支援を行うシステムを作成した.

入力単語 1 単語から, 関連語を取り出し用意したテンプレートに当てはめてストーリーを作成するシステムを作成した. 作成したテキストが面白いと感じられるか, また裏切られていると感じられるかを評価するための実験を行った. その結果, 単語の自動補完により, ギャップによる面白さを生むストーリーを提示できることが確認できた. また, 単語を裏切る際には, 二つの単語間の意味は近すぎず遠すぎない単語を抽出する条件を増やす必要があると考えられる. さらに予想を裏切るために, 具体的な単語を修正候補として提案する機能を追加していきたいと考える.

これらのことから, 単語の関連語を抽出する際の条件の追加, ストーリー生成後にさらに具体的な単語を修正候補として提案する機能を追加, テンプレートおよびカテゴリの追加による, システムの改善をすることでより多様性のあるシステム作成を目標としたい.

## 参考文献

- [1] 志水 彰: 笑い-その異常と正常, 勁草書房, (2000)
- [2] 吉田 祐介, 萩原 将文: 漫才形式の対話文自動生成システム, 日本感性工学会論文誌, Vol. 11, No.2, pp.265-272(2012)
- [3] 真下 遼, 梅谷 智弘, 北村 達也, 灘本 明代: 文の感情を考慮した漫才ロボット台本自動生成 手法の提案, DEIM, Forum, F4-4(2015)
- [4] 山根 宏彰, 萩原 将文: 笑いを生むことわざかしの自動生成システム, 日本知能情報ファジィ学会誌, Vol.24, No.2, pp.671-679(2012)

- [5] 深津 春文, 伊藤 毅志: テーマとキーワードからアイデアの発想を支援するシステム, 情報 処理学会 研究報告 (Web), Vol.2016, No.11, HCI, pp.167-175(2016)