

超短編小説のフリオチパターンの分類とオチの自動生成 Classification of Furi-Ochi Patterns and Automatic Generation of Ochi in Very Short Stories

大島 一海¹⁾ 小川 浩平¹⁾ 佐藤 理史¹⁾
Kazumi Oshima Kohei Ogawa Satoshi Sato

1 はじめに

物語生成は長年取り組まれてきた研究課題である。近年では、大規模言語モデルが登場し、AI のべりすと [1] や ChatGPT [2] により、流暢で破綻の少ない物語の生成が可能になってきている。しかし、生成される物語は、必ずしも面白い物語であるとは限らない。

そこで、本研究では面白い物語の生成を目指す。具体的には、面白さの要因の一つであるフリオチ構造を持つ物語の生成を試みる。

本研究の資料として「54 字の物語」シリーズを用いる。「54 字の物語」とは、縦 9 マス横 6 行の 54 字で綴る物語である。2022 年までにシリーズ全 9 巻が出版されている。表 1 にシリーズのタイトルと収録件数を示す。なお、巻数は出版順に独自に割り振ったもので、書籍に明示されていない。

表 1 「54 字の物語」シリーズのタイトルと収録件数

巻数	タイトル	収録件数
1	54 字の物語 [3]	90
2	54 字の物語 怪 [4]	90
3	54 字の物語 参 [5]	90
4	54 字の物語 史 [6]	90
5	54 字の物語 ZOO [7]	90
6	54 字の百物語 [8]	100
7	旅する 54 字の物語 [9]	82
8	54 字の物語 ∞ [10]	88
9	54 字の物語 Q [11]	80

本研究では、「54 字の物語」シリーズ全 9 巻のうち、1 巻から 8 巻を対象とする。9 巻は内容が物語ではなくクイズであるため、対象外とする。

2 「54 字の物語」のフリオチパターン

2.1 フリとオチ

「54 字の物語」の具体例を以下に示す。

- (1) 「ただいま」と言えば「お帰りなさい」と返ってくる新生活が始まった。家賃も安いし、こんな一人暮らしも悪くない。(第 1 巻 4 話)

この物語の「～始まった。」までの前半部は「二人暮らし」を予想させるが、「家賃も～」の後半部で「一人暮らし」であることを明らかにし、幽霊の存在を示唆する。このように「54 字の物語」の多くは、読者に期待・予想させる前半部と、予想を裏切る後半部で構成される。本研究では、このような前半部を「フリ」、後半部を「オチ」と呼ぶ。

1) 名古屋大学大学院工学研究科 Graduate School of Engineering, Nagoya University

2.2 フリオチパターンの分類

「54 字の物語」のフリオチ構造には、類似するものが存在する。そこで、類似するフリオチ構造を整理し、12 種類のフリオチパターンを設定した。作成したフリオチパターンとその説明、および「54 字の物語」1 巻から 8 巻に含まれる件数を表 2 に示す。この表に示すように、720 件中 364 件の物語が 12 種類のいずれかのパターンに分類できた。

物語のオチを分類する研究として、Murai による星新一のショートショート作品のオチパターン分類 [12] がある。Murai のオチパターンと本研究のフリオチパターンを比較すると、Murai の「評価」「真実」「パロディ」と本研究の「評価逆転」「真実暴露」「パロディ」の 3 パターンは類似していると考えられる。一方、それ以外の、Murai の 6 パターン(「主客」「行動」「目的」「現実-非現実」「原因」「状況悪化」と本研究の 9 パターン(「視点のすり替え」「対象・役割のすり替え」「非現実オチ」「ナンセンス」「斜め上」「ブラック」「不幸」「夢」「言葉遊び」)は類似していないと考えられる。

類似しないものが多い要因として、物語の長さが異なることが考えられる。星新一のショートショートは 1600-5600 字からなる作品が大半を占める [13]。一方、「54 字の物語」は原則 54 字からなる。例えば、Murai の「主客」「行動」などのパターンは、54 字という長さでは成立しえない。その一方で、本研究の「言葉遊び」などのパターンは短い物語でのみ可能である。

3 大規模言語モデルを用いたオチの自動生成

3.1 オチの生成方法

事前学習済み言語モデル GPT-2 をファインチューニングした後、フリを入力し後続文を生成することで、フリオチ構造を持つ非常に短い物語の生成を試みる。この生成方法は以下の 3 ステップで構成される。

- 以下の手順で、フリ/オチのラベルが付いた訓練データを作成する。
 - フリオチ構造を持つ短い物語のテキストを用意する。
 - 物語をフリとオチに分割する。
 - フリ/オチのそれぞれの先頭に「フリ:」「オチ:」のラベルを付与する。
- 1 の訓練データを用いて、日本語 GPT-2 をファインチューニングする。モデルには rinna japanese-gpt2-medium [14] を用いる。
- ファインチューニングした日本語 GPT-2 に、フリの前後に「フリ:」と「オチ:」を付加したものを入力し、出力トークン数を最大 30 に制限して、後続文を生成させる。

以下の実験では、訓練データとして、「54 字の物語」

表 2 フリオチパターンの件数と説明、具体例。括弧内の件数は実験での訓練データ数。

タグ	フリオチパターン	件数	説明と具体例
V	評価逆転	3 (1)	説明 人・ものに対する評価が逆転するオチ 例 私は雨女。初めての海外旅行なのにずっと雨で最悪。そんな私の横で、遊牧民たちが四年ぶりに降る雨に歓喜している。(第 8 巻 37 話)
S	視点のすり替え	52 (19)	説明 視点が期待とは異なる人・ものだと判明するオチ 例 1 あなたはいつも、絵本の途中で寝てしまう。今日は私も疲れたから、このまま一緒に寝ようかな。おやすみ、お母さん。(第 3 巻 75 話) 例 2 異様に安い物件を見つけた。都心で家賃が月一万円。事故物件ではないらしい。残念だ。仲間が見つかると思ったのに。(第 2 巻 2 話)
O	対象・役割のすり替え	18 (10)	説明 対象や役割が期待と異なる人・ものだと判明するオチ 例 手術の日を迎えた。看護師は皆口々に「そんな不安な顔をしないで。大丈夫。簡単な手術ですよ」と励ます。執刀医を。(第 8 巻 76 話)
T	真実暴露	29 (16)	説明 期待とは異なる事実が判明するオチ 例 1 先日研究室に送ってくれた大きなエビ、おいしかったよ。話は変わるが、例の新種生命体のサンプルはいつ届くのかね？(第 1 巻 2 話) 例 2 温暖化が進み、海も汚染されてしまった。見切りをつけて別の星に移住するしかない。人類は三つ目の定住地を探した。(第 2 巻 9 話)
U	非現実オチ	27 (16)	説明 フリは普通(現実的)で、オチが非現実的なもの 例 1 慣れない靴を履いていたから、足にタコができてしまった。タコはじわじわと足を侵食し、やがて僕は八本足になった。(第 8 巻 18 話) 例 2 休日、猫をブラッシング。季節の変わり目なので大量の毛が取れる。どんどん取れる。いつの間にか、猫は消えていた。(第 5 巻 23 話)
N	ナンセンス	73 (31)	説明 無意味、滑稽、期待はずれなオチ 例 1 相手の考えていることが視える眼鏡を手に入れたが、誰で試してもこう映し出されるのだった。「何だこの変な眼鏡は」(第 1 巻 80 話) 例 2 白蛇の抜け殻を買って財布に入れた。大きな招き猫も家に置いた。フクロウも飼った。なぜ私はお金が貯まらないんだ！(第 5 巻 50 話)
E	斜め上	23 (5)	説明 能力・効果が期待を上回るオチ 例 これまで不治の病に冒されていた彼も、とうとう自由の身となった。今はどこへでも行けるし、壁さえすりぬけられる。(第 6 巻 22 話)
B	ブラック	16 (3)	説明 ブラックユーモア。倫理に反するオチや不謹慎なオチ 例 「職場に座敷童が出たって？良かったじゃないか。座敷童のいるお店は繁盛するんだよ」「私の職場、火葬場なんです」(第 6 巻 32 話)
M	不幸	30 (14)	説明 近い将来、不幸(死など)が訪れる(と予想される)オチ 例 1 ある男が呪いの人形を拾ったが、彼の身には何も起こらなかったそう。就職も結婚も、あらゆることが、一生、何も。(第 2 巻 10 話) 例 2 時空郵便にて十年後の自分に手紙を出してみると、すぐに返信が来た。何か書いてあるな。「宛所に尋ねあたりません」(第 8 巻 38 話)
D	夢	2 (1)	説明 現実と夢を混同するオチ 例 うわっ、びっくりした。俺が重大事故を引き起こす夢を見てしまった。ホッとした俺はあくびをしてハンドルを握り直す。(第 3 巻 80 話)
P	パロディ	23 (11)	説明 既存の物語がもとになっているもの 例 蟻がせっせと働くのを横目に音楽に興じるキリギリスは、いつしか超一流のバイオリニストになり巨万の富を得ました。(第 8 巻 26 話)
W	言葉遊び	68 (20)	説明 ダブルミーニング、ダジャレ、漢字などを使ったもの 例 1 「やあ、北岳くん。やっぱり二位だと知名度イマイチだよ。日本一の僕に比べたらさ」「くそっ、マウントフジ……」(第 7 巻 32 話) 例 2 僕は全てを捨てて、この群馬に来た。でもまさか君が突然いなくなるなんて。残された羊と馬をどうしたらいいんだい？(第 3 巻 30 話)
他	その他	356 (78)	以上のパターンに当てはまらないもの
	計	720 (224)	

シリーズ 1 巻から 3 巻中の物語全 270 件のうち、文の区切りでフリとオチに分離できた 224 件を用いた。表 2 の括弧内の数字は、各フリオチパターン毎の訓練データに用いた件数を表す。ファインチューニングは、PyTorch で実装し、主要なハイパーパラメータとして、学習率 5×10^{-5} 、バッチサイズ 1、エポック数 10 を用いた。

3.2 既知のフリからのオチ生成

まず、訓練データに用いた物語に含まれるフリを入力し、オチが生成されるか、元のオチと異なる内容やパターンのおチが生成できるかを調べた。具体的には、「54 字の物語」1 巻と 3 巻の物語のフリ全 153 件から、ランダムにフリ 100 件を抽出し、フリ 1 件につき後続文 40 件を生成した (全 4000 件)。これらを目視で調査し、オチとして認められるかどうかを主観で判定し、114 件 (2.85%) をオチと認定した。

得られたオチの例を以下に示す。それぞれ入力したフリ、a. 生成されたオチ、b. 元のオチの順に示す。

- (2) 「あの葉っぱが落ちたら私は死ぬの」(第 1 巻 46 話)
 - a. 「お願いだから葉っぱを食べないで」/ ナンセンス (N)
 - b. 病室の窓から外を見つめる彼女の目の前には青々とした広大な森が広がっていた。/ ナンセンス (N)
- (3) 畑に強力な成長促進剤をまいた。(第 3 巻 17 話)
 - a. 長い年月をかけて畑はどんどん大きくなり、とうとう畑はなくなってしまった。/ ナンセンス (N)
 - b. 野菜は大きく育っているものの、最近巨大な鳥の目撃情報が多発しているので心配だ。/ 斜め上 (E)
- (4) 「私は神だ。お前の願いを何でも一つ叶えよう」「では、十年後の自分と話がしたい」(第 3 巻 18 話)
 - a. 十年後の自分と会話するにはパスワードを入力して「送信」ボタンを押す必要がある / その他 (他)
 - b. 「残念ながら、それはできない」/ 不幸 (M)
- (5) ちょっと一服。「すみません、ここは禁煙です」世知辛い世の中になったものだ。(第 3 巻 29 話)
 - a. 喫煙所を求め、俺は迷路をさまよっていた。/ ナンセンス (N)
 - b. 喫煙所を求めて、地球を後にする。/ 対象・役割のすり替え (O)

内容的に元のオチと類似したオチは、生成したオチ全体の約 2-3 割を占めた。予想に反して、元のオチと異なる内容のオチのほうが多かった。

オチと認定した 114 件をフリオチパターンに分類した結果を表 3 に示す。元のフリオチパターンと同じパターンに分類されたのは 114 件中 66 件 (58%)、異なるパターンに分類されたのは 48 件 (42%) であった。これ

表 3 既知のフリから生成したオチの分類結果。太字が元の物語と同じパターンの件数。太字以外の数字が元の物語と異なるパターンの件数。空欄は 0 件。

		生成した物語のフリオチパターン												
		V	S	O	T	U	N	E	B	M	D	P	W	他
元の物語のフリオチパターン	V													1
	S					1								1
	O			3			1			2			1	1
	T					1								
	U				1	15	1		3					1
	N				1		27			4				3
	E			2	1		1							
	B													1
	M									12				4
	D													
	P				1	1	5					1		3
	W													1
	他		3	1			1							8
	計			3	6	4	18	36		3	18		1	1

より、元のフリオチパターンと異なるパターンのオチを生成できることが明らかになった。

3.3 未知のフリからのオチ生成

次に、訓練データに含まれないフリを入力し、どのようなオチが生成されるかを調べた。入力に用いるフリとしては、「54 字の物語」4 巻から 8 巻の中から、フリオチパターン毎にランダムに抽出した計 43 件 (表 4) を用いた。それぞれのフリ 1 件に対し、後続文 40 件を生成し (全 1720 件)、主観でオチとして認められるかどうかを評価し、28 件 (1.6%) をオチと認定した。

表 4 フリオチパターンと抽出したフリ件数

フリオチパターン	フリ件数
評価逆転	2
視点のすり替え	4
対象・役割のすり替え	4
真実暴露	4
非現実オチ	4
ナンセンス	4
斜め上	4
ブラック	4
不幸	4
夢	1
パロディ	4
言葉遊び	4
計	43

得られたオチの例を以下に示す。前節と同様に、a. が生成されたオチである。

- (6) ゾウは鼻を人間の手のように器用に扱える。(第 5 巻 24 話)
 - a. ゾウは家に帰ると早速自分の鼻を洗った。/ 対象・役割のすり替え (O)
 - b. 小さなものをつまむこともできるし、このように原稿を書くこともできる。/ 視点のすり替え (S)

- (7) 「職場に座敷童が出たって？良かったじゃないか。座敷童のいるお店は繁盛するんだよ」(第 6 巻 32 話)
- 「うちの従業員は全員座敷童だから安心してください」/斜め上 (E)
 - 「私の職場、火葬場なんです」/ブラック (B)
- (8) 私は雨女。初めての海外旅行なのにずっと雨で最悪。(第 8 巻 37 話)
- 翌日から私は海坊主にされてしまった。/ナンセンス (N)
 - そんな私の横で、遊牧民たちが四年ぶりに降る雨に歓喜している。/評価逆転 (V)
- (9) 「死ぬまでに使い切れないほどの金をくれ」夢に出てきた悪魔にそう願った昨晚。(第 8 巻 49 話)
- 翌日、私はこの世を去っていた。/不幸 (M)
 - 今朝起きたら、枕元に百円玉が一枚。/不幸 (M)

既知のフリを用いるときよりも、オチとして認められる数は少ないが、未知のフリに対してもオチを生成できること、および、内容的にも既存のオチとは異なるオチを生成できることを確認した。

オチと認定した 28 件をフリオチパターンに分類した結果を表 5 に示す。12 種類のフリオチパターンのうち、7 種類のパターン(「視点のすり替え (S)」「対象・役割のすり替え (O)」「非現実オチ (U)」「ナンセンス (N)」「斜め上 (E)」「ブラック (B)」「不幸 (M)」)は生成されたが、それ以外の 5 種類のパターン(「評価逆転 (V)」「真実暴露 (T)」「パロディ (P)」「夢 (D)」「言葉遊び (W)」)は生成されなかった。その要因として以下の 2 点が考えられる。

- 訓練データ数が少ない。
「評価逆転 (V)」「夢 (D)」パターンは、他のパターンと比べてファインチューニングに用いたデータ数が少ないため、生成されにくいと考えられる。
- 生成されにくいパターンがある。
「パロディ (P)」「言葉遊び (W)」パターンは、フリから想起される文脈以外の外部知識(昔話、音韻など)を必要とするため、本研究の手法では生成されにくいと考えられる。

4 おわりに

本研究では、超短編小説「54 字の物語」の分類を試み、12 種類のフリオチパターンを設定した。さらに、事前学習済み大規模言語モデル GPT-2 をファインチューニングしてフリを入力しオチを生成する実験を行った。その結果、オチを生成できること、多様な内容・フリオチパターンのオチを生成できることを確認した。

その一方で、オチが生成される割合は少ないため、ファインチューニングしたモデルのオチ生成能力は、かなり限定されていると考えるのが妥当であろう。

表 5 未知のフリから生成したオチの分類結果。太字が元の物語と同じパターンの件数。太字以外の数字が元の物語と異なるパターンの件数。空欄は 0 件。

生成した物語のフリオチパターン

	V	S	O	T	U	N	E	B	M	D	P	W	他
元の物語のフリオチパターン							1						
V													
S		2	1		1				1				1
O					1	2							3
T								2					1
U						2							
N						1							
E							1		1				
B							1						4
M									2				
D													
P													
W													
計		2	1		2	6	2	2	4				9

参考文献

- [1] Bit192. AI のベリすと. <https://ai-novel.com/>, (参照 2023-05-17).
- [2] OpenAI. ChatGPT: Optimizing Language Models for Dialogue, 2022. <https://openai.com/blog/chatgpt/>, (参照 2023-02-21).
- [3] 氏田雄介 (著/文), 佐藤おどり (イラスト). 意味がわかるとゾクゾクする超短編小説 54 字の物語. PHP 研究所, 2018.
- [4] 氏田雄介 (著/文 | イラスト). 意味がわかるとゾクゾクする超短編小説 ゾク編 54 字の物語 怪. PHP 研究所, 2018.
- [5] 氏田雄介 (著/文 | イラスト), 武田侑大 (イラスト). みんなでつくる 意味がわかるとゾクゾクする超短編小説 54 字の物語 参. PHP 研究所, 2019.
- [6] 氏田雄介 (著/文), 西村創 (著/文), 武田侑大 (イラスト). 超短編小説で学ぶ日本の歴史 54 字の物語 史. PHP 研究所, 2019.
- [7] 氏田雄介 (著/文 | 編集), 今泉忠明 (監修), 武田侑大 (イラスト). 超短編小説で読む いきもの図鑑 54 字の物語 ZOO. PHP 研究所, 2019.
- [8] 氏田雄介 (著/文 | 編集), 武田侑大 (イラスト). 意味がわかるとゾクゾクする超短編小説 54 字の百物語. PHP 研究所, 2020.
- [9] 氏田雄介 (著/文 | 編集), 武田侑大 (イラスト). 超短編小説で読む 47 都道府県 旅する 54 字の物語. PHP 研究所, 2021.
- [10] 氏田雄介 (著/文 | 編集), 武田侑大 (イラスト). みんなでつくる 意味がわかるとゾクゾクする超短編小説 54 字の物語 ∞. PHP 研究所, 2022.
- [11] 氏田雄介 (著/文 | 編集), 武田侑大 (イラスト). 意味がわかるとスッキリする超短編クイズ 54 字の物語 Q. PHP 研究所, 2022.
- [12] Hajime Murai. Automatic Extraction of Reversal-Type Punch Lines in Shin'ichi Hoshi's Flash Fictions. *Journal of the Japanese Association for Digital Humanities*, Vol. 2, No. 1, pp. 31-47, 2017.
- [13] 佐藤知恵, 村井源, 往住彰文. 文学作品群の特征的語彙と概念カテゴリーの抽出. *情報知識学会誌*, Vol. 19, No. 2, pp. 132-137, 2009.
- [14] 趙天雨, 沢田慶. 日本語自然言語処理における事前学習モデルの公開. *人工知能学会研究会資料 言語・音声理解と対話処理研究会*, Vol. 93, pp. 169-170, 2021.