

お題に対してユーモアを生起する回答文選択の検討

伊勢崎 隆司¹ 小林 明美² 有賀 玲子¹ 山田 智広¹

概要: 近年, ユーザとインタラクションが可能なエージェントやロボットの日常生活への普及が進んでいる。これらのシステムにユーモアを生起させるインタラクションが実装できれば, よりユーザフレンドリーなシステムや, ユーモア生起を通じたユーザへの有益な効果が期待される。本研究では, 言語インタラクションに着目し, 不適合理論に基づいてユーモアを生起させる文選択 AI を開発した。構築した文選択 AI をダウンゴ社主催のニコニコ超会議にて 2 日間の展示を通じた実証実験を行い, 288 個のお題と回答のセットに対する観客の感想を取得し, ユーモア率を算出した。一部のお題と回答のセットにおいて高いユーモア率を取得できていたことから, 文選択 AI の有効性が示唆された。一方で, 安定した文選択を実現するために, データベースの拡張や個人最適化の必要性を確認できた。

Selection of Answer Sentences in which Humor Occurs to Input Sentences

TAKASHI ISEZAKI¹ AKEMI KOBAYASHI² REIKO ARUGA¹ TOMOHIRO YAMADA¹

1. まえがき

人工知能技術や計算基盤技術の進歩に伴い, Apple 社の Siri のような身体性を持たないエージェントや, ヴェストン社の Sota のような身体性を持つロボットの普及が進んでおり, ユーザがこれらのシステムとインタラクションする機会が増加してきている。エージェントやロボットのインタラクションをより知的にする試みとしてユーモアを持たせる取り組みがある [1], [2], [3], [4]。ユーモア性を持たせることにより, ユーザ親和性を高める効果や, ユーモアによって生じた笑いを通じてユーザの脳機能・運動機能の向上といった効果が期待される [5]。インタラクションには言語行為と非言語行為の 2 つが存在する [6]。ユーモアの生起においては言語情報と非言語情報の両方が重要であり, それぞれの側面において研究が進んでいる [1], [2], [3], [4], [7]。本研究ではエージェントとロボットの両者が共通で持つ言語行為インタラクション, すなわち対話に着目する。

ユーモアを生起する対話については「不適合理論」が有力である [8]。不適合理論は, 通常はまったく異種であり, 関連がないと思われる思考や概念や状況が意外な方法で結

び合わされることがユーモアを生起させるとする理論である [9]。不適合理論については, 不適合のみでユーモアが生起する不適合モデルと [10], 不適合を説明する認知的ルールを探し出すことによって解消しようとする問題解決活動による不適合-解決モデル [11] がある。本研究では対話でのユーモア生起を考えるため, 不適合モデルや不適合-解決モデルの区別は行わず, 入力文に対して不適合理論が成立するような文の選択を考える。

入力文に対して不適合理論が成立するような文の生成については, 単語の一部を一定のルールに基づいて変更することで実現している例 [2],[12] や, ユーザの発言の一部をわざと間違えて聞き返すことでユーモアを生起させる試みがある [7]。不適合理論を成立させるには文の単語の変更だけでなく, 関連が無いと思われる思考・概念・状況などに基いて文全体を変更することも考えられる。しかしながら, すべてのケースに対応可能なルールを定義するのは困難であると考えられる。本研究では, ルールを明示的に定義するのではなく, 不適合理論が成立する入力文とそれに続く文のセットを複数用意し, 深層学習を通じて関係性を学習するアプローチを採る。このアプローチによってルールを明示的に定義することなく, 多様なユーモア生起が可能であると考えられる。

¹ NTT サービスエボリューション研究所

² NTT コミュニケーション科学基礎研究所

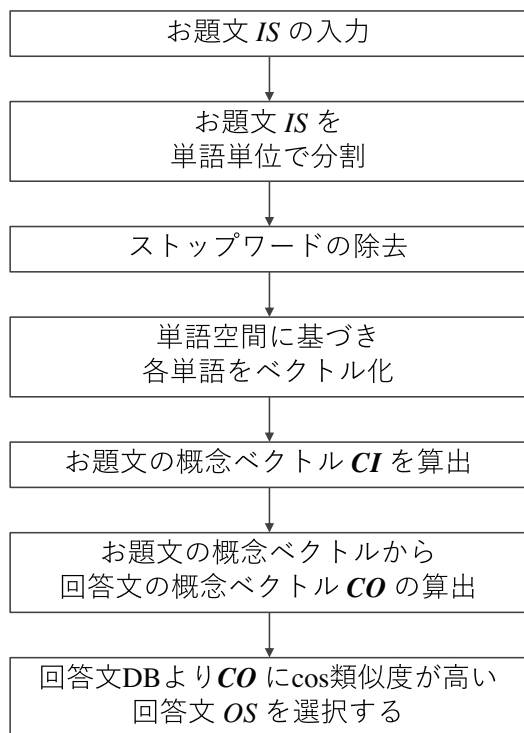


図 1 文選択 AI の処理フロー

Fig. 1 Concept for sentence selection

本研究は入力文に対してユーモアが生起する文の選択を目的とし、ルールを明示的に定めずにユーモアが生起する入力文と出力文のペア群から関係性を学習する方針を採る。

2. 開発システム

コンテキストの文とそれに続く文のセットを用意するにあたり、最もシンプルなケースとして大喜利に着目する。大喜利ではお題に対して不適合理論が成立する文を回答する娯楽であると考えられる。不適合理論に即したお題と回答のセットの関係性を深層学習を用いて学習し、任意のお題に対して適切な文を選択する AI を実装する。

単語や文を数値化するにあたり、日本語の意味空間を構築する必要がある。単語の意味や単語間の Wikipedia の本文データ (2.48GB) に対して、オープンソース形態素解析エンジン MeCab を用いて単語の分かち書きを行い、オープンソース Word2Vec を用いて単語空間を構築した。Word2Vec は 2 層から成るニューラルネットワークで、コーパスを入力として単語の特徴量ベクトル (単語ベクトル) を出力する。図 1 に文選択 AI の処理フローを示す。お題文 IS から回答文 OS を選択する機能を持つ。 IS を単語に分かち書きし、ストップワードを除去したものを $[w_0, w_1, \dots, w_n]$ とする。本研究では助詞と助動詞をストップワードとした。 n は IS の単語数である。 $[w_0, w_1, \dots, w_n]$ の各単語に対して単語空間を用いてベクトル化し、 $n \times 200$ 次元から成る文ベクトル $[w_0, w_1, \dots, w_n]$ を得る。本稿では単語長は 10 に固定し、満たない場合は 0 をパディングした。

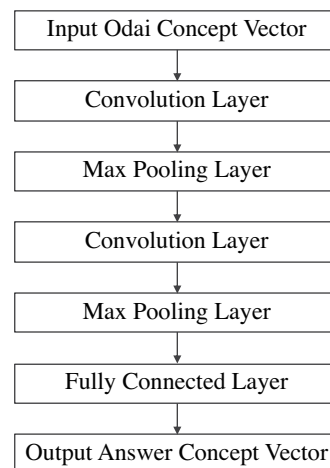


図 2 お題文概念ベクトルから回答文概念ベクトルを算出するモデル

Fig. 2 Model for calculating concept vectors of answer sentences from concept vectors of odai sentences

概念ベクトルの算出は、学習データベースに格納されているお題文と回答文の文ベクトルと回答文データベースに格納されている回答文の文ベクトルを用い、2 層の Encoder と 2 層の Decoder から成る LSTM 方式で AutoEncoder の学習を行い、Encoder の部分を用いることで実装する。Encoder のモデルを用いて文ベクトル $[w_0, w_1, \dots, w_n]$ の概念ベクトル CI (500 次元) を得る。

お題概念ベクトルから回答文概念ベクトルを算出するモデルは、学習データベースに格納されているお題文と回答文の概念ベクトルを用い、図 2 に示す畳み込みニューラルネットワークによる学習を行った。概念ベクトル CI を元に回答文概念ベクトル CO (500 次元) を推定する。

回答文データベースの中から、推定した回答文概念ベクトル CO に近い回答文を検索し選択する。回答文データベースはお題と対になっている回答文群と、我々が保持する文章データから 10000 件の文を抽出して構築した。また、各次元数は経験的に定めた。1 つのお題に対して複数の回答文を出力する際には、複数の回答文ベクトルを結合させて学習させる。すなわち、3 つの回答文を出力させたい場合は、500 次元の回答文概念ベクトルを結合して 1500 次元のベクトルを出力するように学習させる。学習データベースの構築はお笑い構成作家と筆者らで行い、731 件のお題と回答文のセットを得た。

3. ニコニコ超会議への展示を通じた実験

本研究で開発したシステムがユーモアを生起する回答文選択が行えていることを評価するために、株式会社ドワンゴが主催する「ニコニコ超会議 2017」にて展示した。本展示では、図 3 に示すようにガイストン社の Sota5 台を用い、お題に対してロボットが面白い回答をする大喜利形式のデモとして実施した。多様なお題を入力とするために、お題については観客に任意の入力を依頼した。観客にお題全体



図 3 実験環境

Fig. 3 Experimental environment



図 4 感想が印字された 6 面体サイコロ

Fig. 4 A hexahedron dice with impressions printed

を考案してもらうには負担が大きすぎてしまい、お題入力に時間がかかってしまうと予想された。本デモにおいては、「こんな ooo 嫌だ。どんな ooo?」というお題テンプレートを用意した。観客は ooo に入力する言葉を入力することで任意のお題入力を入力することとした。

本展示では 1 つのお題に対して 3 つの回答文をロボットを通じて出力する展示であったため、提案手法において出力回答文は 3 つとなるように学習した。また、一般来場者への展示形式での実験であったため、倫理的に問題のある文章の出力は避ける必要があった。したがって、入力お題に対して文選択 AI を通じて 3 つの回答文概念ベクトルを算出し、それぞれの算出した概念ベクトルに対して cos 類似度の高い順に 10 文ずつ回答文 DB より選択して 30 個の回答候補文を得た後、30 個の回答候補文の中から倫理的に問題が無い文章を実験者が判断して一つ選択する方針を採った。

従来研究におけるエージェントやロボットを用いたユーモアの研究においては、被験者の主観評価に基づいた検討が主な評価手法の一つである [1], [3], [4]。本研究ではお題に対するユーモアを生起する文選択 AI についての検討であるため、従来研究で用いられていた主観評価項目に該当する部分のみを用いることとし、「面白さ」「意外性」「わかりやすさ」の観点で評価項目を設定した。本実験では、表 1 に示す 6 個の感想を観客からのフィードバックとして得ることとした。ポジティブな感想とネガティブな感想を 3 つずつ用意し、「爆笑!」と「つまらん」のように、それぞれが対となるように感想文言を設定した。上記の感想を取得するために図 4 に示すような各面に感想が印字された 6 面体サイコロを用いた。感想の入力は該当文言が記載されている面を上にして台に置くことで、内部に搭載した

表 1 主観評価項目と感想の対応

Table 1 Relationship between evaluation points of view and impressions

	ポジティブ	ネガティブ
面白さ	爆笑!	つまらん
意外性	そうきたか!	単純すぎ!
わかりやすさ	わかる~	わかりにくい

Raspberry PI と加速度センサの処理を通じて上面情報がサーバに送られる。

本展示は約 15 分ほどの実演と 5 分の休憩で構成され、ニコニコ超会議の二日間で連続実演を行った。一回の実演では観客から得られた任意のお題 1 つに対して 3 つの回答文を出力する試行を 3 試行行う。したがって、一回の実演では 3 つのお題に対してそれぞれ 3 つの回答文を出力するため、9 セットのお題と回答文が観客に対して実演される。観客は一つのお題と回答のセットが出力された後に感想をサイコロを通じ、当てはまる感想を複数回答有りて入力した。本展示では 6 面体サイコロを 10 個用意し、観客一人に対して一つのサイコロを割り当てたため、一回の実演で得られるデータは 10 人分のデータとなる。

お題入力は観客の発話を通じてを取得し、スピーカを通じてお題の確認と回答の発話を実施したが、音声認識と音声合成は NTT の技術を用いた。ロボットの動作や 6 面体サイコロの計測、音声認識、音声合成、文選択 AI は NTT のクラウド対応型インタラクション制御技術「R-env: 連舞™」をとして用いて連携制御を実現した。実験結果の解析で用いる 6 面体サイコロの計測情報とお題、回答内容については、R-env に一括で保持されるログを解析することで行った。

4. 結果

本実験では 32 回の実演を実施したため、96 回のお題入力機会と 288 セットのお題と回答に対する 10 人分のデータを得た。図 5 に入力されたお題の内容と頻度を示す。本実験ではお題は「こんな oo 嫌だ。どんな oo?」のテンプレートに制限していたため、「oo」の単語のみ図示している。観客からの自由入力を許したため、複数回入力されているお題もいくつかある。展示システムに係わる「ロボット」などの単語や、開催イベントに係わる単語「超会議」「ニコニコ超会議」などが複数回入力される傾向にあった。

一つのお題と回答のセットに対する 10 人分の観客のポジティブの感想の入力数 p とネガティブの感想の入力数 n を用いて下記の計算を通じて得られるユーモア率 r を算出した。

$$r = \frac{p - n}{p + n} \quad (1)$$

r の算出結果を図 6 に結果を示す。 r が 1 に近いほどユーモアが生起されており、-1 に近いほどユーモアが生起されて

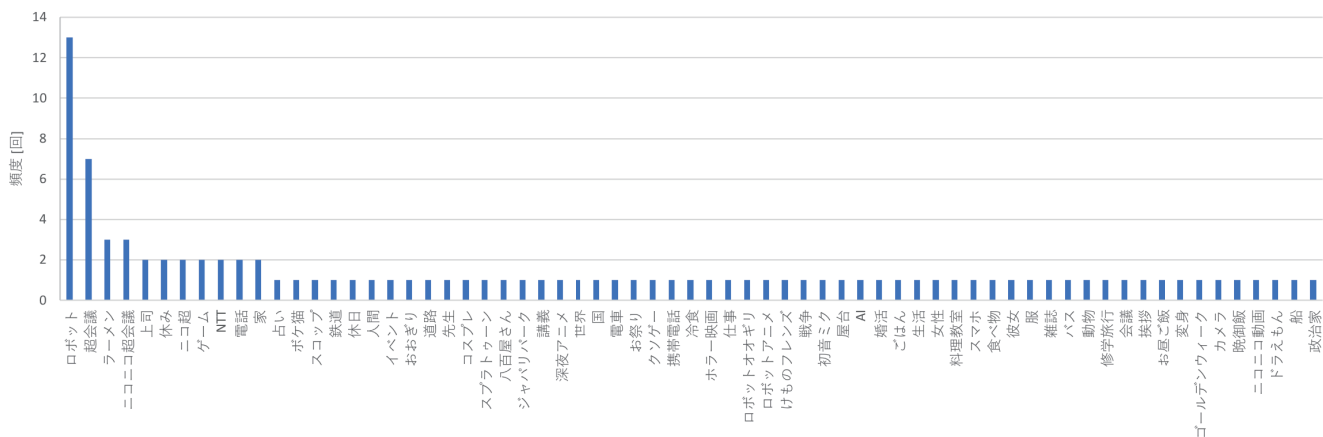


図 5 お題の入力頻度

Fig. 5 Histogram of odoi input times

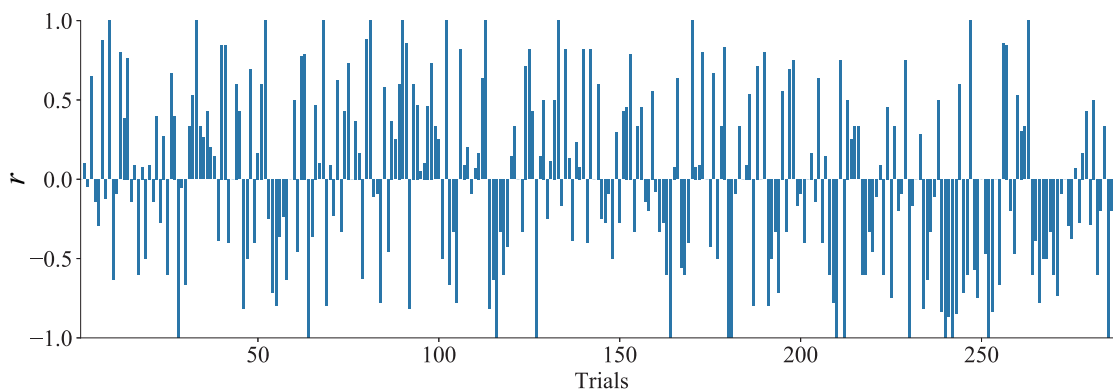


図 6 実験結果

Fig. 6 Result

いないと考えられる。ニコニコ超会議では合計 288 セットのお題と回答を観客に対して提示して感想を得たため、288 個の r を算出した。複数個の r において 1 や -1 に近い値を取っていることが確認できる。また、 r の平均値は 0.002 (標準偏差: 0.561) であった。

r の値が 1 に近かったお題と回答例としては下記のようなものが挙げられる。

- お題：こんな人間嫌だ。どんな人間？
 - 回答：胡散臭い
- お題：こんな雑誌嫌だ。どんな雑誌？
 - 回答：人肌で発火する
- お題：こんなロボット嫌だ。どんなロボット？
 - 回答：酒臭い

r の値が -1 に近かったお題と回答例としては下記のようなものが挙げられる。

- お題：こんなロボット嫌だ。どんなロボット？
 - 回答：縮尺が変
- お題：こんな休日嫌だ。どんな休日？
 - 回答：中身が生肉

- お題：こんな婚活嫌だ。どんな婚活？
 - 回答：水着がスケルトン

5. 考察

図 5 で示す通り、最もお題として入力が多かったのは「こんなロボット嫌だ。どんなロボット？」であった。このお題に対して、表 1 の各感想の得票数の上位 5 つの回答文を表 2 のように集計した。面白さのネガティブ表現「つまらん」と意外性のネガティブ表現「単純すぎ！」の感想において最も得票数が多かったのは「ビームが出る」で共通していた。これは、アニメや SF における「ロボット」の機能として「ビームが出る」は適合しており不適合の関係性とはならなかったことが原因として考えられる。「わかりやすさ」のポジティブ表現においても「ビームが出る」が上位 5 つに入っていたことからもお題文と回答文が適合の関係性であったことが考えられる。本実験で用いた多面体の感想入力機器を用いて、適切なお題文と回答文の評価が可能であることが示唆される。

お題文と回答文の関係性を多面体の感想入力機器で評価

できる可能性が示唆されたが、一度の試行結果を元にお題と回答の関係性を定義してしまうのは適切ではない。例えば、「見た目がゴキブリ」という回答文は、開催期間中に3回「こんなロボット嫌だ。どんなロボット？」のお題に対して回答として出力したが、1試行目と2試行目は面白さのネガティブ表現「つまらん」が最高得票数であった一方で、3試行目は面白さのポジティブ表現「爆笑！」が最高得票であった。このような違いは、前の回答文との相対的なユーモアの大きさの差異、観客の笑いに対する特性の差異の影響等が考えられる。観客の笑いに対する特性の差異や前の回答文との相対的なユーモアの大きさの差異が感想入力に対してどのように影響を与えるかは詳細な検討が必要である。また、他観客が入力する感想による影響も否定できない。他者の情報が入らない環境での評価も今後求められる。

図6の結果において r の値が-1から1の範囲で分散していたことから、文選択AIの評価が偏りなく実施されている点において、本展示で用いた6面体サイコロを用いた観客の感想の取得は有効であると考えられる。また、一部のお題に対する回答については r が1に近い値を得られていることから、本展示で構築した文選択AIでユーモアが生起できていると考えられる。一方で、 r の平均が0に近く、分散も大きい結果であったことから、構築した文選択AIが安定して観客全員に対して適切な文選択が行えていないと考えられる。この結果の要因の一つとして、文選択AIの文選択精度が低いことが考えられる。安定して適切な文選択を行うためには、様々なトピックに対応できる網羅的なデータベースを用意する必要がある。構築した文選択AIが安定して観客全員に対して適切な文選択が行えていない別の要因としては、観客間の感想が分散している可能性がある。すなわち、観客間でユーモア生起の特性が異なるため一様のデータベースに基づく文選択AIでは、ユーモアが生起する観客と生起しない観客に別れてしまうと考えられる。特定の観客に対して安定的に適切な文選択を実現するためには個人最適化を行う必要があると考えられる。

r の値が1に近かったお題と回答例においては、不適合理論に基づいた文選択が適切に行われていると考えられ、その結果 r の値が高く得られたと考えられる。一方で r の値が-1に近かったお題と回答例においては、お題に対して不適当な回答はできているが、解釈として解決するに至らなかったことが原因だと考えられる。

6. まとめ

本研究では不適合理論に基づいて作成したお題と回答文データベースを元に、深層学習を用いたユーモアを生起する文選択AIを構築した。構築した文選択AIをドワンゴ社主催のニコニコ超会議にて2日間の展示を通じた実証実験を行い、288個のお題と回答のセットに対する観客の感想

を取得し、ユーモア率を算出した。一部のお題と回答のセットにおいて高いユーモア率を取得できていたことから、文選択AIの有効性が示唆された。一方で、安定した文選択を実現するために、データベースの拡張や個人最適化の必要性を確認できた。今後は、データベースの拡張や個人最適化を中心に取り組み、文選択AIの精度を高めていく。

参考文献

- [1] 灘本明代：漫才ロボット：Webを用いた漫才本自動生成，システム制御情報学会研究発表講演会講演論文集，Vol. 60, p. 4p (オンライン)，入手先<<http://ci.nii.ac.jp/naid/40020836203/>> (2016).
- [2] 吉田裕介，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] 前田実香，鬼沢武久：単語の関連性とおもしろさを取り入れたなぞかけ生成，感性工学研究論文集，Vol. 5, No. 3, pp. 17-22 (2005).
- [4] 林宏太郎，神田崇行，宮下敬宏，石黒 浩，萩田紀博：ロボット漫才：社会的受動メディアとしての二体のロボットの利用，日本ロボット学会誌，Vol. 25, No. 3, pp. 381-389 (2007).
- [5] 松阪崇久：ヒトはなぜ笑うのか？：行動学の視点から，笑い学研究，Vol. 21, pp. 5-18 (オンライン)，入手先<<http://ci.nii.ac.jp/naid/110009840886/>> (2014).
- [6] ENOMOTO, M. and NAKANO, Y.: Verbal/Nonverbal Behavior Allocation Rules in Social Interactions between Users and Conversational Agents, *Journal of Japan Society for Fuzzy Theory and Intelligent Informatics*, Vol. 20, No. 4, pp. 540-556 (online), DOI: 10.3156/jsoft.20.540 (2008).
- [7] 鈴木 奨，呉 健朗，瀧田航平，堀越 和，中辻 真，宮田章裕：ボケて返す対話型エージェントの基礎検討，技術報告3，日本大学文理学部，日本大学文理学部，日本大学文理学部，日本大学文理学部，NTT レゾナント株式会社，日本大学文理学部 (2017).
- [8] 伊藤大幸：ユーモア経験に至る認知的・情動的過程に関する検討：不適合理論における2つのモデルの統合へ向けて，認知科学，Vol. 14, No. 1, pp. 118-132 (オンライン)，DOI: 10.11225/jcss.14.118 (2007).
- [9] 上野行良：ユーモア現象に関する諸研究とユーモアの分類化について，社会心理学研究，Vol. 7, No. 2, p. 112120 (1992).
- [10] Nerhardt, G.: Humor and inclination to laugh: emotional reactions to stimuli of different divergence from a range of expectancy, *Scandinavian journal of psychology*, Vol. 11, No. 3, p. 18595 (1970).
- [11] Suls, J. M.: A two-stage model for the appreciation of jokes and cartoons: An information-processing analysis, *The psychology of humor: Theoretical perspectives and empirical issues*, Elsevier, pp. 81-100 (online), DOI: 10.1016/B978-0-12-288950-9.50010-9 (1972).
- [12] Fujikura, S., O. Y. and Kikuchi, H.: Humor Utterance Generation for Non-task-oriented Dialogue Systems, *Proceedings of the 3rd International Conference on Human-Agent Interaction. ACM*, p. 171173 (online), DOI: 10.1145/2814940.2814966 (2015).

表 2 お題「ロボット」における, 各感想項目での得票数上位 5 つの回答文

Table 2 Top 5 answer sentences in each impression when the odai is “Robot”.

爆笑!	つまらん	そうきたか!	単純すぎ!	わかる～	分かりにくい!
酒臭い	ビームが出る	ジャリジャリしている	ビームが出る	トゲトゲがすごく鋭利	中身がグチャグチャ
締め付けてくる	縮尺が変	ムキムキマツチョ	見た目がゴキブリ	再現性が無い	パイロットがバイト
パイロットがバイト	慈悲の衣揚げ定食。	武器がゴム鉄砲	熱風が出る	ムキムキマツチョ	幽霊が映る
加齢臭がひどい	肩に落ち武者がいる	トゲトゲがすごく鋭利	酒臭い	ビームが出る	嫌なものなんて無い
トゲトゲがすごく鋭利	見た目がゴキブリ	パイロットがバイト	加齢臭がひどい	パイロットがバイト	熱風が出る