

令和 6 年 5 月 9 日現在

機関番号：10101

研究種目：基盤研究(C)（一般）

研究期間：2021～2023

課題番号：21K12007

研究課題名（和文）文脈情報を考慮したユーモアの面白さの評価のための標準データベースの開発

研究課題名（英文）Development of standard database for evaluating the fun of humor considering contextual information

研究代表者

荒木 健治（ARAKI, Kenji）

北海道大学・情報科学研究院・特任教授

研究者番号：50202742

交付決定額（研究期間全体）：（直接経費） 3,100,000円

研究成果の概要（和文）：本研究課題の成果は、高齢者単身世帯の急増に伴う話し相手の需要の急増、ユーモアの医療への応用といった問題を解決するために、ユーモア研究に対する基盤的な研究環境を整備したことであり、ユーモアの面白さが文脈に依存することは自明である。本研究課題では文脈を考慮したユーモアの面白さを評価するための標準データベースの開発を行った。このことにより他分野の研究者の本研究分野への進出が容易になり、多くの研究者がユーモアの研究に取り組むようになるものと考えられる。今後は、あらゆるユーモアに関する工学的な研究を統合的に行うユーモア工学を確立することを考えている。

研究成果の学術的意義や社会的意義

ユーモアの面白さが文脈に依存することは自明であるが、文脈まで考慮したユーモアの面白さの評価手法は未だに確立されていなかった。本研究成果の学術的意義はこの問題を解決するために文脈を考慮したユーモアの面白さを評価するための標準データベースの開発を行ったことである。また、生成AIの急速な進歩によって社会は急激に進化している。本研究成果の社会的意義は、本研究課題で開発した大規模な文脈付きユーモアデータベースを生成AIの学習に使用することにより生成AIのユーモア能力が進化することである。また、生成AIに最も重要な大規模なデータベースの開発も本データベースの開発方法が貴重な知識になるものと思われる。

研究成果の概要（英文）：The outcome of this research project is the development of a basic research environment for humor research in order to solve problems such as the rapidly increasing demand for talking partners due to the rapid increase in the number of elderly single-person households and the application of humor to medicine. It is obvious that humor depends on context. In this research project, we have developed a standard database for evaluating humor in context. We believe that this will facilitate the entry of researchers from other fields into this research field and encourage more researchers to engage in humor research. In the future, we plan to establish humor engineering that integrates engineering research on all kinds of humor.

研究分野：自然言語処理

キーワード：ユーモア 文脈 データベース 面白さ ユーモア工学

1. 研究開始当初の背景

研究開始当初の背景としては、高齢者単身世帯の急増に伴う話し相手の需要の急増、ユーモアの医療への応用といった問題が未解決であったことがあげられる。この問題を解決するためにはユーモア研究に対する基盤的な研究環境を整備することが必要であった。また、ユーモアの面白さが文脈に依存することは自明であるが、文脈まで考慮したユーモアの面白さの評価手法が確立されていなかった。

2. 研究の目的

本研究の目的は、文脈を考慮したユーモアの面白さを評価するための標準データベースの開発を行うことである。このような標準データベースが開発されることにより他分野の研究者が容易に本研究分野へ進出できるようになり、ユーモア研究の活性化が図られる。最終的には、あらゆるユーモアに関する工学的な研究を統合的に行うユーモア工学を確立することやユーモアに関する研究を統合的に発表、討論する場としてユーモア工学学会の設立が考えられる。

3. 研究の方法

本研究課題では対話の文脈によるユーモアの面白さへの影響を分析するために、Twitter上に存在する26,739件の駄洒落を含む対話を収集した。収集したもののうち、駄洒落を含む対話12,000件に対して対話の面白さの評価を行った。ここで、評価にかかる時間的コストの高さと評価者の駄洒落を含む対話に対する飽きを考慮し、クラウドソーシングを用いて16名の評価者を雇用した。評価者16名には1件の対話に対して、3名の評価が付与されるように1,000-3,000件の作業を割り当てた。

4. 研究成果

評価の結果として、第一に評価者16名のスコアの91.3%は1点、2点及び3点であることが確認された。この要因として、Twitter上の対話には第三者から面白さを評価される前提がないことが考えられる。第二に、人によって異なる面白さの度合いについて分析を行った。評価者間の一致度をコーエンの重み付き係数を用いて算出した結果、係数0.29となり、おおむね一致していることが確認された。それに加えて、評価者組間の一致数と割合についても分析を行った。その結果、2者一致しているものが約5割、3者一致しているものが約2割、3者一致していないものが約2割であることが確認された。次に、面白さに影響を及ぼす対話文脈がどのような特徴を持つかについて、駄洒落に対する話者の反応に着目した分析を行った。ここで、一般的な駄洒落に対する話者の反応をもとに、高度なノリ、単純なノリ、無反応、その他の4種類の話者反応を定義した。これらの定義をもとに、面白さが増加及び減少した対話2,000件に対して話者反応の分類を複数人で行い、多数決によって定めた話者反応の割合を分析した。ここで、アノテータ間で完全に意見が分かれたものを分析対象から除外した。その結果、面白さが増加した対話918件のうち、67.6%が高度なノリ、単純なノリの話者反応を含む対話となった。このことから、駄洒落を含む発話に対して、関連性のある表現または駄洒落を返すような対話文脈は駄洒落の面白さを増加させることが考えられる。また、面白さが減少した対話905件のうち、89.0%は無反応、その他の話者反応を含む対話となった。このことから、対話中の駄洒落に対して反応がないまたは対話

内容の解釈が困難な対話文脈は駄洒落の面白さを減少させることが考えられる。

今後の課題として、はじめに駄洒落データベース中の10,001番以降の駄洒落を用いた駄洒落を含む対話の追加収集を行うことが挙げられる。次に、第三者から面白いと感じられる対話をより多く収集するために、駄洒落が好きな人に駄洒落を含む対話を作成していただくことが挙げられる。これにより、第三者に面白さを与える意図がある対話の収集を行った上で、これらの対話をもとに、対話文脈を含む駄洒落の面白さを評価する標準データベースを構築することである。

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件/うち国際共著 0件/うちオープンアクセス 2件）

1. 著者名 1. 浅野歴, 谷津元樹, 森田武史	4. 巻 34
2. 論文標題 キャプション及び物体認識ラベルを用いた描画内容に即した駄洒落文の選択手法	5. 発行年 2023年
3. 雑誌名 画像ラボ	6. 最初と最後の頁 59 ~ 65
掲載論文のDOI (デジタルオブジェクト識別子) 10.11501/3348745	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Ryoma Hanabusa and Kenji Araki	4. 巻 2
2. 論文標題 Construction and Analysis of Contextual Pun Database in Japanese Using Twitter	5. 発行年 2022年
3. 雑誌名 Proceedings of the Twenty-Seventh International Symposium on Artificial Life and Robotics	6. 最初と最後の頁 1233-1237
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 花房竜馬, 荒木健治	4. 巻 35
2. 論文標題 対話文脈付き駄洒落データベース構築のための駄洒落を含む対話の面白さの評価及び分析	5. 発行年 2023年
3. 雑誌名 知能と情報 (日本知能情報ファジィ学会誌)	6. 最初と最後の頁 683 ~ 692
掲載論文のDOI (デジタルオブジェクト識別子) 10.3156/jssoft.35.3_683	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

〔学会発表〕 計12件（うち招待講演 0件/うち国際学会 0件）

1. 発表者名 浅野歴, 谷津元樹, 森田武史
2. 発表標題 画像キャプション生成に基づく描画内容に即した駄洒落文の選択手法
3. 学会等名 2022年度人工知能学会全国大会 (第36回)
4. 発表年 2022年

1. 発表者名 浅野歴，谷津元樹，森田武史
2. 発表標題 キャプション及び物体認識ラベルを用いた画像に関連する駄洒落文の選択手法
3. 学会等名 情報処理学会 第252回自然言語処理研究会
4. 発表年 2022年

1. 発表者名 浅野歴，谷津元樹，森田武史，江上周作，鶴飼孝典，福田賢一郎
2. 発表標題 画像キャプション生成及び物体認識を用いた駄洒落文ランキング手法の評価
3. 学会等名 言語処理学会第29回年次大会
4. 発表年 2023年

1. 発表者名 志方脩，谷津元樹，森田武史
2. 発表標題 常識的知識グラフ及び単語埋め込みを用いた重畳型駄洒落ユーモア認識
3. 学会等名 言語処理学会第29回年次大会
4. 発表年 2023年

1. 発表者名 谷津元樹，稲垣ルイ，田羅間大晴，森田武史
2. 発表標題 汎用的言語理解フレームワークに基づく統合型駄洒落雑談対話システムの構築と評価
3. 学会等名 人工知能学会第2種研究会 第71回ことば工学研究会
4. 発表年 2023年

1. 発表者名 貞誉志樹, 谷津元樹, 森田武史
2. 発表標題 常識的知識グラフに基づく理解容易性及びユーモア性を考慮した直喩表現生成
3. 学会等名 人工知能学会第2種研究会 第71回ことば工学研究会
4. 発表年 2023年

1. 発表者名 花房竜馬, 荒木健治
2. 発表標題 駄洒落を含む対話における可読性と面白さの相関分析
3. 学会等名 言語処理学会 第29回年次大会 発表論文集
4. 発表年 2023年

1. 発表者名 井上蒼一朗, 谷津元樹, 森田武史
2. 発表標題 重畳型駄洒落ユーモアにおける常識的知識グラフを用いた潜在表現抽出
3. 学会等名 言語処理学会第28回年次大会
4. 発表年 2022年

1. 発表者名 王格格, 荒木健治
2. 発表標題 深層格解析を用いた駄洒落質問応答ペアの自動生成
3. 学会等名 第12回言語獲得と理解研究会報告
4. 発表年 2022年

1. 発表者名 原直大, 荒木健治
2. 発表標題 ChatGPT を利用したユーモア応答生成の分析
3. 学会等名 電子情報通信学会技術研究報告, 言語理解とコミュニケーション研究会, NLC2023-8, pp.41-46
4. 発表年 2023年

1. 発表者名 内田ゆず, 高丸圭一, 乙武北斗, 木村泰知
2. 発表標題 母親向け情報サイトにおける検索語と投稿文を対象とした感情分析
3. 学会等名 第 39 回ファジィシステムシンポジウム 講演論文集, pp.779-783
4. 発表年 2023年

1. 発表者名 佐山公一, 荒木健治
2. 発表標題 一瞬笑える, または面白いと感じる駄洒落で答える言語生成AI: 個人・グループ専用データベースの駄洒落と似てはいるが異なる駄洒落を事前学習データとWebから引き出す
3. 学会等名 第74回ことば工学研究会(人工知能学会第2種研究会)資料, SIG-LSE-C403, pp.27-34
4. 発表年 2024年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	佐山 公一 (SAYAMA Kohichi) (90271733)	小樽商科大学・商学部・教授 (10104)	

6. 研究組織（つづき）

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	内田 ゆず (UCHIDA Yuzu) (80583575)	北海学園大学・工学部・教授 (30107)	
研究分担者	谷津 元樹 (YATSU Motoki) (30805015)	青山学院大学・理工学部・助教 (32601)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関