

キャラクターのセリフに表れる音変化の傾向の可視化

南出 瞭馬[†] 山西 良典^{†,a} 松下 光範^{†,b}

[†] 関西大学総合情報学部

a) ryama@kansai-u.ac.jp b) m.mat@kansai-u.ac.jp

概要 Web 上で漫画の配信・蓄積が可能になった今、閲覧可能な漫画の数は増加し続けている。膨大な量の漫画の中から自らの興味や関心に沿った漫画を探し出すことは難しい。読者は、好みの要素を多く含んだキャラクターに深く共感したり、感情移入したりすることも多い。そのため、キャラクターを主軸にした作品検索は、読者の好みに合った漫画検索を実現できる可能性がある。本稿では、キャラクターのセリフを「内容」と「表現」に分割し、「表現」が類似したキャラクターを発見するための手法を提案する。キャラクターごとのセリフで使用される「表現」をベクトル化し、キャラクター間のベクトル類似度を算出することで、セリフ中の「表現」が類似しているキャラクターを検索可能にする。検証の結果、話し方の「表現」に着目してキャラクターの類似度を算出できる可能性が示唆された。

キーワード キャラクター特性, セリフの分析, 音変化表現

1 はじめに

漫画は、セリフや説明文などのテキスト情報と登場人物や背景などの画像情報を用いて物語を伝えるエンタテインメントである。Web 上で漫画の配信・蓄積が可能になったことに後押しされ、閲覧可能な漫画の数は増加し続けており、膨大な量の漫画の中からユーザが自らの興味や関心に沿った漫画を探し出すことは難しい。漫画を対象とした従来の検索技術は、書誌情報やジャンル情報など事前に付与されたメタデータや、コンテンツの 1 部分から抽出したテキストや画像を用いて情報検索や情報推薦を実現してきた。しかし、これらの技術では漫画の内容に踏み込んだ検索や推薦は難しく、ユーザの多様なニーズに応えるには至っていない。

読者が物語世界へ没入しているとき、しばしば作中の登場人物になったかのような体験をしたり、登場人物の感情を共有したりすることがあると小山内らは指摘している [1]。その中でも読者は、好みの要素を多く含んだキャラクターに深く共感したり、感情移入したりすると考えられる。そのため、似たようなキャラクターが登場する作品は、その読者の好みに合う可能性が高い。また、キャラクターは漫画のストーリー構成に深く関わっており、似たような状況でも、キャラクターの性格によってその後のストーリー展開は大きく変化する。こうしたことから、キャラクターの性格が好みの漫画を見つけ出す際のひとつの基準となると想定し、キャラクターの性格に着目した漫画検索システムの実現を目指す。

本研究の最終目的は、セリフからキャラクターの性格を推定する事が可能なシステムを構築し、キャラクターや漫

画の検索支援に貢献することである。この目的を達成するために、キャラクターのセリフを、発話行為によって相手に伝達する事柄である「内容」とキャラクターを特徴付ける言い回しや語尾である「表現」の 2 種類の観点に分け、キャラクター特有の「内容」や「表現」からキャラクターの性格を推定する。本稿では「表現」に着目し、キャラクターを特徴付ける言い回しや語尾といった「表現」を検出し、多次元ベクトルとして表現する手法を提案する。キャラクター間でベクトル類似度を算出することで、セリフ内に使用される「表現」が似ているキャラクターを発見可能にする。このために、アニメの字幕データから抜き出したキャラクターのセリフを記載したデータセットを構築した。セリフの表現に着目したキャラクター検索の可能性を検討する。

2 関連研究

2.1 漫画キャラクターの性格分類に関する研究

漫画キャラクターの性格分類に関する研究として、朴らの研究 [2] がある。朴らは、登場キャラクターの性格タイプを考慮した検索システムの実現を目指している。性格診断ツールとして広く利用されているエゴグラム (東大式エゴグラム第 2 版 (以下, 新版 TEG2) [3]) を利用し、漫画キャラクターの性格分類を試みている。エゴグラム [4] は、交流分析理論に基づく性格診断法であり、質問紙により性格の類型を判別する。エゴグラムによる分析結果は、客観性や信頼性、妥当性が高いとされており、教育現場や医療現場など様々な場で活用されている [5]。

漫画キャラクターの性格のデータとして、漫画やアニメのファンの主観データの集合知蓄積されたファンサイトの情報を参照した報告もある。中島らは、キャラクターの

顔画像の特徴点間のバランスから、キャラ属性王国¹の基礎性格・性格タグの推定を試みた。また、キャラ属性王国のデータを情報源としたキャラクタのデザインと性格の関係性の分析 [6, 7] も行っている。

これらの研究では、キャラクタの性格を表現した文章やタグを Web 上から収集、抽出しているため、キャラクタ本人が作品内で有する特徴を分析しているわけではない。本稿では、漫画キャラクタのセリフに着目する。

2.2 漫画キャラクタの発話の役割に関する研究

漫画キャラクタの発話の役割に関する研究として森らの研究 [8] がある。森らは、対話相手に対する一連の発話役割分布をベクトルで表現する方法を提案し、その有用性について検討している。従来の漫画の検索サイトや漫画の理解を促すシステム [9] においては、漫画内での登場人物の関係性を示す上で、恋人、友人、ライバル、幼馴染などのように、登場人物 2 者間関係の有無だけを表す無向関係が用いられている。森らは、無向関係の「友達以上恋人未満」と表現するよりも、A にとっての「片想いの相手」と B にとっての「友達」といったように登場人物 2 者それぞれにとっての関係を表す有向関係で表現する方が詳細で的確に表現できると考えている。無向関係に比べて有向関係は、登場人物それぞれがお互いにどのように捉えているのかを示すことができるようになるため、漫画のメタデータ表現や利用 [10] においても、登場人物間関係性をより詳細に表現可能になる。登場人物間の有向関係について、「人間関係は会話内容に現れる」という Matsumoto らの指摘 [11] に基づいて発話から関係性の推定を試みている。人間の会話は、客観的な事柄を表す命題と命令や聞き手に対する発話者の態度の 2 つで構成され、話者の態度の中には発話意図が含まれていると言われている [12]。発話は発話者から話相手への感情や意図の伝達といった有向の役割がある。このような観点から、森らは登場人物間の有向関係の表現に、発話の役割を応用する手法を提案している。森らの手法では、発話の役割ラベルに西原らの研究 [13] で使用された発話の役割を参考に設定されている。西原らの研究では発話文の役割を 23 種類に分類し、それを用いて人間関係の仲の良さや上下関係を推定している。感情に関与する発話の役割は“心情”という役割のみで定義されているが、森らの研究では、対話相手に対しての感情の表出を発話の役割として表現する必要があるため、“心情”ラベルを“好意”“嫉妬”“挑発”のように細分化している。細分化する際、感情音声に関わる役割ラベル [14] を参照し、漫画の中での出現頻度が高かった“好意”“嫉妬”“挑発”を、漫画の登場人物間の有向関係を理解するために必要な役割ラベルとして位置づけている。また、

榎本の研究 [15] では、友人への感情は、“同調感”“自尊・独立感”“親密感”“信頼感”“不信感”“ライバル”の 6 つの感情的側面に分けられることが述べられている。このようにして、発話をあらかじめ定めた役割に分類 (38 種類) し、その役割を基底としてある対話相手に対する一連の発話の役割分布をベクトルで表現している。

発話の役割への注目はキャラクタ同士の関係性を表現するうえで有用な観点と考えられる。しかし、登場人物の関係性を表現するだけでは、キャラクタに着目した漫画検索は実現できない。本稿は、個々のキャラクタの性格に着目して、セリフからのキャラクタの性格のモデル化を目指す。

2.3 漫画キャラクタの発話の表現に関する研究

漫画キャラクタの発話の表現に関する研究として、金水 [16] の研究がある。金水は、特定の人物像と結びついた話し方の類型を役割語と呼び、「老人語」「幼児語」「お嬢様言葉」など、どのようなキャラクタがどのような表現を使うのかや文法的な特徴を示した。

宮崎らの研究 [17, 18] では、文末表現をはじめとする機能語の語彙選択に着目し、バリエーション豊かな言語表現を生成する手法を提案している。具体的には、「これはひどいな」という発話を「これはひどいわね」のように、発話を構成する各文節の機能語を付与したいキャラクタ性に適した表現に書き換えるを行っている。さらに、宮崎らの他の研究 [19] では、キャラクタの発話に現れる音変化表現 (元の形から一部の音が変化して派生したと思われる表現) を収集し、それらを基に音変化表現を人為的に発生させるための知識を整理している。宮崎らの研究では、小説や漫画における発話文の話者 (キャラクタ) を推定する実験に音変化表現のパターンの情報を利用することで、推定性能が向上するキャラクタが存在することが確認されていた。音変化表現のパターン一覧が用意されているため、キャラクタらしさを特徴付ける音変化表現を抽出するシステムの構築ができればキャラクタの性格推定に利用することができると考える。

柳らの研究 [20] では、文末述語解析システム Panzer/U について述べている。入力された発話文を形態素解析し、文末述語の範囲を決定した。そして、その範囲を内容語と機能語列に変換している。機能語には音変化表現が含まれており、音変化表現を表示するために検出規則を用意している。

本稿では、これらの音変化表現の検出規則を応用する。キャラクタのセリフ全文を解析して音変化表現をベクトル化することで、セリフの表現に着目したキャラクタ検索を実現する。

¹<https://chara-zokusei.jp/>

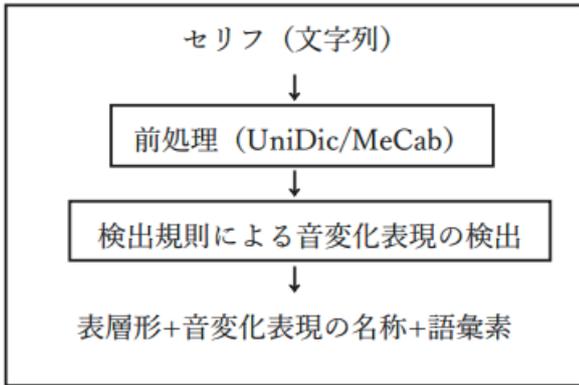


図 1 音変化表現の検出フロー

3 データの準備

本稿では、アニメキャラクターを分析対象とし、アニメ初登場話のセリフに着目する。ここで、キャラクターのセリフ情報を収集するために、映像作品ストリーミング配信サービス Netflix² に収録されているアニメの字幕を参照した。字幕データを取得するアニメタイトルは、キャラ属性王国に掲載されているキャラクターが登場する作品としている。今回のデータセットの作成で字幕データを取得したアニメタイトルは、Netflix で 2022 年に放送されたアニメとした。これらのアニメに付与される字幕データを手作業で収集し、1 キャラクターずつ初登場話のセリフを画面上に表れる 1 文を 1 発話として抜き出し、記録した。結果として、表 1 に示す 45 作品、87 キャラクターのデータセットが作成された。表中には、セリフを抜き出した作品のキャラクターと各キャラクターの発話数を示している。

4 音変化の検出手法と音変化に着目した類似キャラクター群の抽出

提案手法では、入力したセリフ内の単語に音変化表現があった場合、音変化表現ごとの頻度を要素とする多次元ベクトルを獲得する。得られた多次元ベクトルは次元削減したうえでベクトル間のコサイン類似度を算出し、使用された音変化表現の傾向が類似するキャラクターを発見可能にする。

入力した文章に音変化表現があった場合、該当する単語にどのような音変化が起きているのかを検出する。図 1 に示すように、入力されたセリフへの前処理を行ったうえで、音変化表現の検出規則を参照し、セリフから音変化表現を検出する。

4.1 音変化表現の検出規則

音変化表現を検出するためには、参照する音変化表現の一覧とその検出規則が必要となる。本稿では、柳ら [20] の「音変化表現の検出規則一覧」を参照する。柳

活用形 = 終止形 - 一般

$$\left(\begin{array}{l} \text{表層形末尾} = \text{え} \\ \text{基本形末尾} = \text{A い} \end{array} \right) \Rightarrow \text{アエ交替}$$

例：うるせえ

図 2 音変化表現の検出規則の一例

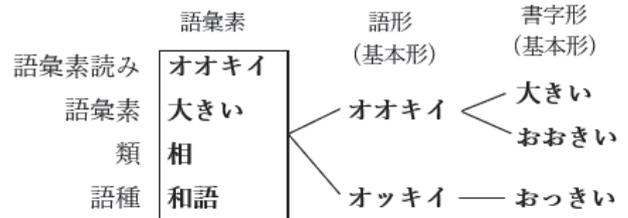


図 3 語彙素の構成

らは、図 2 のように、ある単語に指定された活用形があり、括弧内の条件が満たされる場合、音変化表現の名称が決定されるといった検出規則を提案している。柳らは、話し言葉に見られるくだけた表現にも比較的広く対応した辞書「現代話し言葉 UniDic³」を形態素解析に利用し、その解析結果で得られる情報から音変化表現を検出している。現代話し言葉 UniDic は、音変化表現に対して他の辞書に比較して頑健な性能を示すことが報告されている。現代話し言葉 UniDic を利用した形態素解析では、品詞大分類～細分類、活用型、活用形、語彙素読み、語彙素、書字形出現形・基本形、発音形出現形・基本形、仮名形出現形・基本形、語形基本形といった情報が得られる。ここで、語彙素とは言語学における形態論の単位である。図 3 に、語彙素の構成を示す。異なる形態であるが同じ語であると考えられるものからなる語の集合を語彙素と呼ぶ。

柳らの「音変化表現の検出規則一覧」は、一部の音変化表現のみ (27 パターン) を対象としている。そこで、音変化表現の検出規則として、柳らの「音変化表現の検出規則一覧」の基となる宮崎ら [19] の「音変化表現のパターンの一覧」についても採用した。検出規則を構築する際には、宮崎らの「音変化表現のパターンの一覧」に付記されたパターンの説明と柳らの「音変化表現の検出規則一覧」にある音変化表現の検出規則を参考にした。これにより、現代話し言葉 UniDic を使用した形態素解析から得られる情報を組み合わせて、134 パターンの音変化表現の検出を可能にした。

4.2 音変化表現の検出アルゴリズム

提案手法では、以下の検出アルゴリズムで音変化表現を検出する。

²<https://www.netflix.com>

³<https://clrd.ninjal.ac.jp/unidic/download.html>

第 10 回コミック工学研究会予稿集

表 1 セリフを抜き出した作品一覧

| 作品名 | キャラクタ名 | 発話数 | 作品名 | キャラクタ名 | 発話数 |
|-----------------|--------------|-----|----------------|--------------|-----|
| BASTERD!!暗黒の破壊神 | ターク・シュナイダー | 64 | ハコヅメ～交番女子の逆襲～ | 川合麻依 | 155 |
| BASTERD!!暗黒の破壊神 | ティア・ノート・ヨーコ | 95 | ハコヅメ～交番女子の逆襲～ | 藤聖子 | 113 |
| BLEACH | 朽木ルキア | 109 | ブルーロック | 潔世一 | 121 |
| BLEACH | 黒崎一護 | 146 | リコリス・リコイル | 井ノ上たきな | 100 |
| Dr.STONE | 大木大樹 | 137 | リコリス・リコイル | 錦木千束 | 206 |
| Dr.STONE | 石神千空 | 144 | 七つの大罪 | エリザベス | 79 |
| Engage Kiss | キサラ | 107 | 七つの大罪 | メリオダス | 82 |
| Engage Kiss | 夕桐アヤノ | 57 | 僕のヒーローアカデミア | オールマイト | 44 |
| Engage Kiss | 緒方シュウ | 150 | 僕のヒーローアカデミア | 爆豪勝己 | 32 |
| SPY × FAMILY | ロイド・フォージャー | 235 | 僕のヒーローアカデミア | 緑谷出久 | 145 |
| SPY × FAMILY | アーニャ・フォージャー | 83 | 古見さんは、コミュ症です。 | 古見硝子 | 9 |
| TIGER&BUNNY | ワイルドタイガー | 142 | 可愛いだけじゃない式守さん | 式守 | 99 |
| ULTRAMAN | 早田進次郎 | 77 | 名探偵コナン ゼロの日常 | 安室透 | 65 |
| うる星やつら | ラム | 68 | 名探偵コナン ゼロの日常 | 榎本梓 | 59 |
| うる星やつら | 三宅しのぶ | 56 | 名探偵コナン 犯人の犯沢さん | 犯沢真人 | 70 |
| うる星やつら | 諸星あたる | 190 | 弱虫ペダル | 今泉俊輔 | 106 |
| かぐや様は告らせたい | 四宮かぐや | 156 | 弱虫ペダル | 小野田坂道 | 120 |
| かぐや様は告らせたい | 白銀御行 | 116 | 彼女、お借りします | 木ノ下和也 | 213 |
| かぐや様は告らせたい | 藤原千花 | 72 | 彼女、お借りします | 水原千鶴 | 141 |
| からかい上手の高木さん | 西片 | 125 | 後宮の烏 | 柳寿雪 | 92 |
| からかい上手の高木さん | 高木さん | 56 | 機動戦士ガンダム 水星の魔女 | スレッタ・マーキュリー | 118 |
| ぼっち・ざ・ろっく! | 伊地知虹夏 | 132 | 機動戦士ガンダム 水星の魔女 | ミオリネ・レンブラン | 77 |
| ぼっち・ざ・ろっく! | 山田リョウ | 37 | 無職転生 | ルーデウス・グレイラット | 238 |
| ぼっち・ざ・ろっく! | 後藤ひとり | 253 | 無職転生 | ロキシー・ミグルディア | 50 |
| ようこそ実力至上主義の教室へ | 堀北鈴音 | 48 | 盾の勇者の成り上がり | 北村元康 | 64 |
| ようこそ実力至上主義の教室へ | 榎田栞穂 | 68 | 盾の勇者の成り上がり | 天木錬 | 39 |
| ようこそ実力至上主義の教室へ | 綾小路清隆 | 136 | 盾の勇者の成り上がり | 岩谷尚文 | 389 |
| ヴァニタスの手記 | ヴァニタス | 99 | 盾の勇者の成り上がり | 川澄樹 | 42 |
| オーバーロード | アウラ・ベラ・フィオーラ | 35 | 組長娘と世話係 | 桜樹八重花 | 36 |
| オーバーロード | アルベド | 37 | 賢者の弟子を名乗る賢者 | ミラ | 102 |
| オーバーロード | マーレ・ベロ・フィオーレ | 24 | 賭ケグルイ双 | 早乙女芽亜里 | 162 |
| オーバーロード | モモンガ | 232 | 賭ケグルイ双 | 花手毬つづら | 103 |
| カッコウの許嫁 | 天野エリカ | 157 | 進撃の巨人 | アルミン・アルレルト | 45 |
| キングダム | 信 | 204 | 進撃の巨人 | エレン・イエーガー | 92 |
| サマータイムレンダ | 小舟滯 | 93 | 進撃の巨人 | ミカサ・アッカーマン | 20 |
| サマータイムレンダ | 網代慎平 | 185 | 阿波連さんははかれない | 阿波連れいな | 60 |
| デリシャスパーティープリキュア | 和実ゆい | 143 | 魔入りました!入間くん | アスモデウス・アリス | 32 |
| シャドーハウス | エミリコ | 181 | 魔入りました!入間くん | 鈴木入間 | 140 |
| シャドーハウス | ケイト | 116 | 魔法科高校の劣等生 | 千葉エリカ | 35 |
| チェンソーマン | デンジ | 189 | 魔法科高校の劣等生 | 司波深雪 | 92 |
| テニスの王子様 | 竜崎桜乃 | 78 | 魔法科高校の劣等生 | 司波達也 | 139 |
| テニスの王子様 | 越前リョーマ | 48 | 魔法科高校の劣等生 | 柴田美月 | 31 |
| パリピ孔明 | 月見英子 | 136 | ジョジョの奇妙な冒険 6部 | 空条徐倫 | 176 |
| パリピ孔明 | 諸葛孔明 | 187 | | | |

1. セリフの形態素解析
2. 音変化表現検出規則との照合による語彙素の獲得
3. 音変化表現パターンごとの集計

まずは入力したセリフを形態素解析し、単語に分割する。それぞれの単語の表層形、語彙素、書字形基本形、品詞、品詞細分類、活用型、活用形といった7種類の情報を取得する。取得された7種類の情報を基に、条件式として記述した検出規則と照合し、規則に該当する単語があった場合、表層形と音変化表現の名称、語彙素を獲得する。このとき、複数の検出規則に該当した場合、規則ごとに表層形と音変化表現の名称、語彙素を獲得する。どの検出規則にも該当しない単語については無視する。そして、音変化表現パターンごとに語彙素を集計する。

表層形と音変化表現の名称、語彙素の3種類の情報を提示することで、元の単語から起きた音変化の理解を助ける。検出対象とした音変化表現パターンは、宮崎らの「音変化表現のパターンの一覧」のパターン 34, 48, 49 を除く 134 パターンである。これらのパターンを整理したところ、112 種類の検出規則としてまとめられた。例えば、パターン 24 の“意志活用語尾の「う」が脱落す

る”とパターン 25 “「だろう (でしょう)」の「う」が脱落する”は現代話し言葉 UniDic を利用した形態素解析結果の活用形で“意志推量形であり語尾が「う」でない”を検出する規則としてまとめられる。本稿では、この 112 種類の検出規則への該当を音変化表現のベクトル要素として扱う。音変化を現代話し言葉 UniDic の参照では検出できなかった音変化表現のパターンは、検出対象から除外した。具体的には、パターン 34 の“「いうような」の「よう」が脱落し、「いうな」となる”, パターン 48 の“「さ」「し」「す」「せ」「そ」が「ちゃ」「ち」「ちゅ」「ちえ」「ちょ」に交替する”, パターン 49 の“「つ」が「ちゅ」に交替する”の3種類であった。パターン 34 では、「いうな」の「な」を「終助詞な」としか判断できず、助動詞の体言接続と判別できなかった。また、パターン 48 と 49 では、変化後の「ちゃ」「ち」「ちゅ」「ちえ」「ちょ」が含まれるセリフからは、本来これらの変換前の単語が「サ行」や「つ」が使われていることを確認できないため検出不可能であった。

検出アルゴリズムに従った出力例を以下に示す。ここで、i が入力、o が出力を表す。

1. 「すごい」の音変化表現の検出

i: すんげえ

o: すんげえ **オエ**交替**凄い**, すんげえ **撥音挿入****凄い**

2. 「あるじゃないか」の音変化表現の検出

i: あるじゃねえか

o: ある **有る**, じゃ**だ**, ねえ **アエ**交替**無い**, **かか**

各キャラクターのセリフの音変化表現は多次元ベクトルによって表現する. キャラクター c の音変化表現の検出規則 i の出現数を N_i とすると, 音変化表現の検出規則 i に対応する音変化表現ベクトルの要素 x_i は式 (1) で表現される.

$$x_i = \frac{N_i}{\sum_{i=1}^{112} N_i}. \quad (1)$$

キャラクター c の音変化表現ベクトル $v(c)$ は, 式 (1) で算出される x_i を用いて, 式 (2) で表現される.

$$v(c) = (x_1, x_2, \dots, x_{112}). \quad (2)$$

式 (2) より, $v(c)$ は, キャラクター c のセリフにおける音変化表現の各検出規則の相対的な占有率を示すことになる.

4.3 音変化表現ベクトルによる
類似キャラクター群の抽出

3 節で作成したデータセット内のキャラクターごとに音変化表現ベクトル $v(c)$ を獲得した. ここで, $v(c)$ は 112 次元のベクトルで表されるため, 音変化表現ベクトルの要素 x_i は 0 の値が多く, 疎なデータになってしまう. そこで, 主成分分析 (以下, PCA) を用いて, 得られた音変化表現ベクトルを次元削減し, 密なデータに変形させた.

本稿では, PCA の累積寄与率が 98% 以上となる 21 次元の主成分ベクトルを得た. 次元削減された音変化表現ベクトルについてキャラクター間でコサイン類似度を算出し, セリフの表現が類似したキャラクター群を抽出した.

5 分析結果と考察

表 2 に, 分析対象とした 87 キャラクターの音変化表現ベクトルの類似度上位 15 組を示す. 以下, いくつかのペアについてセリフの音変化表現を参照しながら考察する.

コサイン類似度の値が最も高いペアは, 「BLEACH」の「黒崎一護」と「キングダム」の「信」で類似度は 0.9422 であった. この 2 キャラクターは共に, 文末に促音が挿入

表 2 コサイン類似度の値が高い上位 15 組

| キャラクター a | キャラクター b | 類似度 |
|----------|----------|--------|
| 黒崎一護 | 信 | 0.9422 |
| 竜崎桜乃 | 鈴木入間 | 0.9162 |
| 堀北鈴音 | ケイト | 0.9045 |
| 緑谷出久 | 鈴木入間 | 0.9003 |
| アルペド | 司波深雪 | 0.8969 |
| 安室透 | 川澄樹 | 0.8890 |
| 黒崎一護 | 竜崎桜乃 | 0.8864 |
| 白銀御行 | 司波深雪 | 0.8762 |
| 白銀御行 | 諸葛孔明 | 0.8683 |
| 黒崎一護 | 鈴木入間 | 0.8628 |
| 竜崎桜乃 | 緑谷出久 | 0.8547 |
| 信 | 鈴木入間 | 0.8544 |
| ヴァニタス | 柳寿雪 | 0.8475 |
| 安室透 | 柴田美月 | 0.8329 |
| 信 | 竜崎桜乃 | 0.8318 |

される「文末促音挿入」という音変化表現が最も多く使われており, 「うるさい」というセリフが「うるせえ」と変化するような「アエ交替」という音変化表現が 2 番目に多く使われていた. セリフを詳細に見てみると, 「黒崎一護」と「信」は, 戦闘描写が多く描かれているため, 「うあああっ!」や「ウオーッ!」といった攻撃時の掛け声に「文末促音挿入」が多く使用されていた. また, 「黒崎一護」と「信」は共に, 「テニスの王子様」の「竜崎桜乃」や, 「魔入りました! 入間くん」の「鈴木入間」とも類似度が高い結果となった. 「竜崎桜乃」と「鈴木入間」は, さらに「僕のヒーローアカデミア」の「緑谷出久」とも類似度が高いことが分かった. 「竜崎桜乃」「鈴木入間」「緑谷出久」の 3 キャラクターは, 「黒崎一護」や「信」と同じく「文末促音挿入」が最も多く使われていた. しかし, 「竜崎桜乃」と「鈴木入間」と「緑谷出久」の「文末促音挿入」は「ええーっ!」といった驚きや, 「あっ…」といった言い淀む場合などで使われており, 「黒崎一護」や「信」の「文末促音挿入」の使われ方とは異なっていた.

「かぐや様は告らせたい」の「白銀御行」と「魔法科高校の劣等生」の「司波深雪」は類似度が高い上位 15 組に 2 度出現したキャラクターである. 「白銀御行」は「司波深雪」との類似度が 0.8762 と 8 番目に高く, 「パリピ孔明」の「諸葛孔明」との類似度は 0.8683 と 9 番目に高くなっていた. また, 「司波深雪」は「オーバーロード」の「アルペド」との類似度が 0.8969 と 5 番目に高かった. 「白銀御行」「司波深雪」「アルペド」「諸葛孔明」は全員, 「食べん」のような活用形が終止形-撥音便で活用

型が助動詞-ヌである「終止音便形」という音変化表現が最も多く使われていた。セリフを詳細に見てみると、「司波深雪」「アルペド」「諸葛孔明」は「申し訳ありません」といった丁寧な謝罪や、「～じゃありません」のような丁寧な否定で「終止音便形」が多く使われていた。一方で、「白銀御行」は「やらんでもないがな」や「気にせんが」といった偉そうな言葉遣いとして使われることが多かった。

特に、「黒崎一護」と「信」のように直感的に類似していると思われるキャラクター同士に高い類似度が認められたことから、音変化表現を用いることでセリフ中の「表現」が似ているキャラクターの類似度を計算できる可能性が示唆された。一方で、「黒崎一護」や「信」に対しての「竜崎桜乃」「鈴木入間」「緑谷出久」のように、コンテキストによって表現されるパラ言語の違い [21] (気合を入れる「はあっ!」と口ごもる「あっ…」の「っ」の違い) はテキスト情報からは判別できず、直感にそぐわない類似キャラクターが抽出されてしまうことも確認された。音変化表現としては同一であっても、どのように発話されるかや発話意図といった情報までは分析できないため、セリフのテキスト以外の特徴量や会話の流れなどのコンテキスト情報の導入が課題となる。

6 おわりに

本稿では、セリフからのキャラクターの性格推定を目指して、セリフの「表現」に着目した音変化表現ベクトルを提案した。音変化表現ベクトルの類似度からセリフの表現が類似したキャラクター群を抽出した。抽出結果の考察から、セリフの表現が類似しているキャラクター群を抽出し可視化できる可能性が示唆された。セリフの内容に着目したキャラクター間の類似度の算出については、今後の課題として取り組んでいきたい。

謝辞

本研究は、一部、科研費 22K12338 および 20K12130 の支援のもと行われた。記して謝意を表す。

参考文献

- [1] 小山内秀和, 楠見 孝: 物語世界への没入体験, 心理学評論, Vol. 56, No. 4, pp. 457–473 (2013).
- [2] 朴 柄宣, 居林香奈枝, 松下光範: エゴグラムに基づいたコミックキャラクターの性格分類, 第 32 回人工知能学会全国大会論文集, pp. 1J302–1J302 (2018).
- [3] 和田迪子, 渡部麻美, 市村美帆, 松井 豊: Web 調査による新しいエゴグラムの尺度開発, 筑波大学心理学研究, Vol. 53, pp. 63–71 (2017).
- [4] Dusay, J.: Egograms and the constancy hypothesis, *Transactional Analysis Journal*, Vol. 2, No. 3, pp. 37–41 (1972).
- [5] 山崎勝之: 日本における性格研究の動向と展望, 教育心理学年報, Vol. 41, pp. 73–83 (2002).
- [6] Nakashima, F., Fujimoto, N. and Yamanishi, R.: Do Appearances Represent Personality of Created Characters?, *Proc. of the 27th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems* (2023).
- [7] 中島楓華, 藤本直樹, 横井 優, 西明和樹, 山西良典: ファンサイト上のキャラクターの外見属性と内面属性の関係性の基礎分析, 電子情報通信学会第 9 回コミック工学研究会 (2023).
- [8] 森理緒奈, 山西良典, 松下光範: 発話の役割ベクトルによる登場人物間の有向関係表現の有用性検証, 電子情報通信学会第 5 回コミック工学研究会, pp. 55–61 (2021).
- [9] Murakami, H., Nagaoka, Y. and Kyogoku, R.: Creating Character Networks with Kinship Relations from Comics, *International Journal of Service and Knowledge Management*, Vol. 4, No. 1, pp. 1–26 (2020).
- [10] 福田ちひろ, 三原鉄也, 永森光晴: マンガのシーン抽出のためのコマの連続性評価手法—人物出現パターンに着目して—, 第 3 回コミック工学研究会予稿集, pp. 13–19 (2020).
- [11] Matsumoto, K., Minato, J., Ren, F. and Kuroiwa, S.: Estimating human emotions using wording and sentence patterns, *Proc. IEEE International Conference on Information Acquisition*, Vol. 1, pp. 421–426 (2005).
- [12] 益岡隆志: モダリティの文法, くろしお出版 (1991).
- [13] 西原陽子, 砂山 渡, 谷内田正彦: 発話テキストからの人間の仲の良さと上下関係の推定, 電子情報通信学会論文誌, Vol. J91-D, No. 1, pp. 78–88 (2008).
- [14] 永岡 篤, 森 大毅, 有本泰子: 感情音声コーパス共通化のための新たな感情ラベル推定における既存感情ラベル併用の効果, 日本音響学会誌, Vol. 73, No. 1, pp. 682–693 (2017).
- [15] 榎本淳子: 青年期における友人との活動と友人に対する感情の発達の变化, 教育心理学研究, Vol. 47, No. 2, pp. 180–190 (1999).
- [16] 金水 敏: 役割語と日本語教育, 日本語教育, Vol. 150, pp. 34–41 (2011).
- [17] Miyazaki, C., Hirano, T., Higashinaka, R., Makino, T. and Matsuo, Y.: Automatic Conversion of Sentence-end Expressions for Utterance Characterization of Dialogue Systems., *Proc. the 29th Pacific Asia Conference on Language*, pp. 307–314 (2015).
- [18] 宮崎千明, 平野 徹, 東中竜一郎, 牧野俊朗, 松尾義博, 佐藤理史: 文節機能部の確率的書き換えによる言語表現のキャラクター性変換, 人工知能学会論文誌, Vol. 31, No. 1, pp. DSF-E.1–9 (2016).
- [19] 宮崎千明, 佐藤理史: 発話テキストへのキャラクター性付与のための音変化表現の分類, 自然言語処理, Vol. 26, No. 2, pp. 407–440 (2019).
- [20] 柳 将吾, 佐藤理史, 夏目和子, 宮田 玲, 小川浩平: 話し言葉を対象とした文末解析と表現文型の同定, 言語処理学会 第 27 回年次大会 発表論文集, pp. 1708–1712 (2021).
- [21] 清野陽平, 山西良典, 辻野雄大, 松村耕平: 「はあ」データセットの構築と音声特徴の基礎分析, 第 69 回情報処理学会エンタテインメントコンピューティング研究会, No. 7, pp. 1–6 (2023).