

複数の手がかりを利用した小説発話の話者推定

石川和樹¹ 佐藤理史¹ 宮田玲¹ 小川浩平¹¹ 名古屋大学大学院工学研究科

ishikawa.kazuki.h8@s.mail.nagoya-u.ac.jp

概要

人間は、小説中のセリフの話者が誰なのかを苦もなく理解することができる。これは話者を推測するための様々な手がかりが、テキスト中に存在するためである。本論文では、複数の手がかりを利用した話者推定システムを提案する。本システムは2つのモジュールから構成されており、前者の話者候補リスト作成では、セリフの周囲に出現する人物名を抽出して話者候補リストを作成したのち、直前直後の文における話者の明記、話者交替、一人称の3つの手がかりを利用して話者候補を絞り込む。後者の口調に基づく話者推定では、セリフの口調を手がかりに話者を推定する。

1 はじめに

小説は、脚本とは異なり各セリフの話者が明記されない。しかし、我々読者は、話者が誰なのかを苦もなく理解することができる。

これは、テキスト中に話者を推測するための多くの手がかりが存在するためである。それらのうち、主要なものに次の3種類がある。

- セリフの直前直後の文で、話者が明示される場合がある。
- 連続するセリフでは、話者が交替する。
- セリフの内容や使われている口調から、話者を推測できる。

英語の小説の場合は、1と2の手がかりのみでほとんどのセリフの話者を推定することができる。これは、“replied Kitty”のように話者が明記される場合や、“said her father”のようにセリフの直後に代名詞で話者が記される場合が多いためである。Heら[1]は、英語小説の“Pride and Prejudice”を対象に調査を行った結果、前者の割合は約25%、後者の割合は約15%であったと報告している。

一方日本語の小説では、セリフの話者が直前直後

の地の文で明記される割合は全体の数%程度しかなく、1と2の手がかりのみで話者を推定できるセリフは限定される。

我々は昨年、3の手がかり、すなわち口調に基づく話者推定の方法を提案した[2]。この方法ではセリフを口調ベクトルに変換する変換器を実現し、少数の訓練例(話者が既知のセリフ集合)から算出した各話者の代表口調ベクトルとの類似度に基づき、セリフの話者を決定する。

今回、この方法を小説テキストに適用可能にするために、各セリフに対して話者候補リストを生成するモジュールをその前段として追加した。このモジュールは、セリフの周囲に出現する人物名を抽出して候補リストを作成したのち、直前直後の文における話者の提示(手がかり1)、話者交替(手がかり2)、一人称表現の3つの手がかりを用いて候補を絞り込む。

本論文で提案するような話者の自動推定システムは、セリフに話者を付与したデータベースの作成支援に応用できる。これは、小説のセリフを題材とした研究(セリフの話者推定や、話者の性別推定[3]、性格推定[4]など)や、キャラクター性を付与したセリフ生成[5]の実現のための基礎データとなる。さらに、音声合成と組み合わせることにより、話者ごとに音声を変えたオーディオブックの自動作成も可能となる。

2 関連研究

英語の小説を対象とした研究では、地の文に提示されている人名(手がかり1)や、話者交替のルール(手がかり2)といった表層に現れている手がかりを使って話者推定を行っている。英語の小説はこれらの手がかりのみで十分な推定が可能であり、例えば、Heらは地の文の情報を特徴量としたSVMを用いて話者推定を行い、86.5%の精度を達成した[1]。

英語小説を対象とした研究には他にも、Muznyら[6]のルールベースの手法や、遠田ら[7]の教師なし

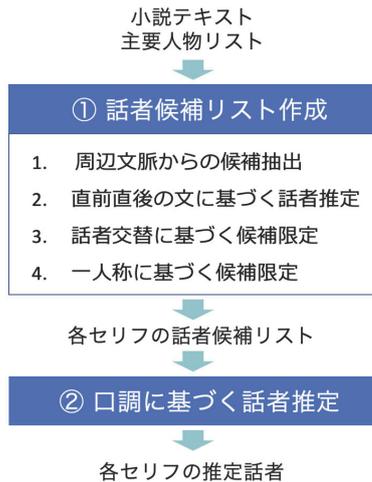


図1 システム構成

機械学習を用いた手法が存在する。どちらも、1節で示した手がかり1と2を主要な手がかりとして扱ったものであり、手がかり3はほとんど使われていない。

3 話者推定システム

3.1 問題設定

本システムの入出力は以下の通りである。

- 入力：小説テキストと主要人物リスト
- 出力：各セリフの推定話者

さらに、以下の制約を設けている。

- 推定対象作品：日本語のエンタメ小説およびライトノベル
- 推定対象のセリフ：主要人物のセリフのみ
- 推定先の候補：主要人物のみ

なお、主要人物とは、発話数の多い人物5名程度を指す。主要人物リストをシステムの入力として与える設定のため、解くべきタスクは分類問題(各セリフに対して主要人物リストから1人選択する問題)となる。

3.2 システム構成

図1に提案システムの構成を示す。本システムは、「話者候補リスト作成(3.3節)」と「口調に基づく話者推定(3.4節)」の2つのモジュールから構成されている。

3.3 話者候補リスト作成

話者候補リスト作成は、以下の4ステップで構成されている。

1. 周辺文脈からの候補抽出
2. 直前直後の文に基づく話者推定
3. 話者交替に基づく候補限定
4. 一人称に基づく候補限定

3.3.1 ステップ1: 周辺文脈からの候補抽出

まず、各セリフに対して、そのセリフの周辺に出てくる人名を抽出することによって話者候補リストを作成する。具体的な手順を以下に示す。

1. 対象のセリフの N_b 文前から N_a 文後を候補抽出範囲とする。ただし候補抽出範囲にセリフが含まれる場合、そのセリフを候補抽出範囲には含めるが、 N のカウントには含めない。
2. 候補抽出範囲に名前が出てくる主要人物を話者候補とする。ただし、一人称小説の場合は、主人公は必ず話者候補に含める。

なお、システムに与える主要人物のリストでは、それぞれの人物に対して姓と名を与えるため、姓または名だけでも、人物を同定できる。

3.3.2 ステップ2: 直前直後の文に基づく話者推定

次に、セリフの直前直後の地の文が、話者名と発話動詞を含む特定パターンにマッチしている場合、そのセリフの話者を一意に定める。その詳細を付録Aに示す。

例えば、以下の例の場合、 $U1$ の話者をAに決定する。

$N1$: Aは、
 $U1$: 「こんにちは」
 $N2$: と言った。

3.3.3 ステップ3: 話者交替に基づく候補限定

地の文を挟まずに連続するセリフでは、話者が交替するというルールを利用して話者候補を限定する。具体的には、以下の手続きを適用する。

あるセリフの話者が S で、かつ、その直前または直後がセリフだった場合に、それらのセリフの話者候補から S を除外する。

なお、この手続きにより新たに話者候補が1人とな

表1 一人称として認定する語

俺	おれ	オレ
僕	ぼく	ボク
私	わたし	ワタシ
	わたくし	ワタクシ
	あたし	アタシ

るセリフが得られた場合は、そのセリフに対してもこの手続きを適用する。

例えば、以下の例の場合、直前直後の文に基づく話者推定(3.3.2節)によってU1の話者候補がAのみになるため、U2の話者候補からAを除外する。

N1: Aはお礼を言った。

U1: 「ありがとう」

U2: 「どういたしまして」

3.3.4 ステップ4: 一人称に基づく候補限定

最後に、各話者が使う一人称を収集し、その情報を用いて話者候補を限定する。以下に手順を示す。

1. 話者が一意に定まったセリフを用いて、それぞれの話者が使う一人称を決定する。複数の一人称を使用していた場合は、最も使用頻度が高い一人称を、その人物の一人称とする。
2. セリフ中に一人称 X が含まれる場合、そのセリフの話者候補から、 X 以外の一人称を使う人物を除外する。

表1に、一人称として認定する語の一覧を示す。

例えば、以下の例の場合、U1の話者候補からAを除外する。

各人物の一人称: {A: 俺, B: 私, C: 私}

U1: 「私はこれがいいです」

3.4 口調に基づく話者推定

各セリフに対して話者候補リストを作成した後、話者候補が複数であるセリフに対して口調に基づく話者推定を適用する。この時点で、すでに話者が一意に決定されているセリフは、その結果をそのまま出力すると同時に、以下の手続きの代表ベクトルの算出に用いる。セリフを口調ベクトルに変換する方法には、すでに提案した方法[2]を用いるが、口調ベクトルの次元数は64を用いる。

1. 代表口調ベクトルの算出

それぞれの主要人物に対し、その人物の既知のセリフを口調ベクトルに変換し、その平均を代表口調ベクトル(話者の口調を表すベクトル)とする。

表2 口調に基づく話者推定の結果

	セリフ数	recall	precision	F1
A	321	77.9%	71.4%	74.5%
B	317	92.4%	78.8%	85.1%
話者 C	285	58.6%	81.5%	68.2%
D	208	75.5%	69.8%	72.5%
E	64	53.1%	81.0%	64.2%
macro-F1				72.9%
micro-F1				75.4%

2. 話者推定

各推定対象のセリフ(話者候補が複数のセリフ)に対して、そのセリフの口調ベクトルと各主要人物の代表口調ベクトルとの類似度を計算し、その値が最も高い話者を推定話者とする。

4 実験

4.1 実験条件

以下の条件で実験を行った。

- 対象作品: 米澤穂信『氷菓』及び『愚者のエンドロール』(一人称小説)
- 主要人物: 5人
- 全セリフのうち、主要人物のセリフが占める割合: 69.9%
- 周辺文脈からの候補抽出での候補収集範囲: セリフの前10文~後ろ3文 ($N_b = 10, N_a = 3$)

なお、対象作品の『氷菓』と『愚者のエンドロール』は同一シリーズの作品であり、主要人物も共通している。そのため、これら2作品を繋げ、1つの作品とみなし実験を行った。

4.2 口調に基づく話者推定の結果

表2に、口調に基づく話者推定の、主要人物5人の結果(recall, precision, F1)と、全体の結果(macro-F1, micro-F1)を示す。

先の研究[2]—各主要人物のセリフの10%を代表口調ベクトルの算出に用い、残り90%を推定対象セリフとした場合—のmacro-F1は61.4、micro-F1は68.3であったが、今回は、それぞれ72.9と75.4に向上した。さらに、先の研究では各話者のF1の値はばらついた(28.2%~77.6%)が、今回は64.2%~85.1%とばらつきも抑えられた。この結果より、口調以外の手がかりを用いて一部のセリフの話者を確定し、それらのセリフを用いて代表口調ベクトルを算出する方法で、口調に基づく話者推定が実行でき

表3 話者候補リスト作成の各ステップの話者候補数と話者推定結果

		ステップ1 候補抽出	ステップ2 直前直後の文	ステップ3 話者交替	ステップ4 一人称
話者候補リスト作成	平均候補者数	3.14人	3.00人	2.98人	2.76人
	正解含有率	98.8%	98.4%	98.2%	98.0%
	候補が1人の精度	73.7% (14/19)	90.3% (102/113)	90.4% (113/125)	94.9% (259/273)
口調に基づく話者推定	平均候補者数	-	3.17人	3.17人	3.16人
	話者推定精度	-	76.5% (1037/1355)	77.2% (1037/1343)	75.4% (901/1195)
システム全体	話者推定精度	-	77.6% (1139/1468)	78.3% (1150/1468)	79.0% (1160/1468)

ることが示された。

4.3 話者候補リスト作成の詳細

次に、話者候補リスト作成の各ステップで、セリフの話者候補リストがどのように推移しているかを調べた。この結果を表3に示す。

ステップ1(候補抽出)では、平均3.14人の候補が得られ、その中に正解が含まれる割合は98.8%であった。ただし、話者が一意に決定できるセリフの数は19と少ない。

ステップ2(直前直後の文の利用)では、新たに94セリフの話者が一意に決定され、この段階で113セリフ(全1468セリフ中の7.7%)の話者が一意に決定された。その精度は $102/113=90.3\%$ である。この時点の結果を(つまり、ステップ3,4を省略して)後段の話者推定に入力した場合、口調による話者推定の精度は76.5%、システム全体の精度は77.6%となる。なお、口調に基づく話者推定では、話者候補が複数存在するセリフのみを対象とするため、対象セリフの平均候補者数は、話者候補リスト作成時の平均候補者数とは一致しない。

ステップ3(話者交替の利用)を追加すると、話者が一意に決定されるセリフが12増加した。このステップで候補が削減できるのは、セリフが連続し、かつ、それらのいずれかのセリフの話者が確定している場合のみである。加えて、1つのセリフから削減できる候補数も1候補である。このため、このステップの効果は限定的である。

ステップ4(一人称の利用)を追加すると、話者が一意に決定されるセリフが148増加し、一意に決定されるセリフの精度も94.9%に向上した。このステップで候補が削減できるのは、一人称が含まれるセリフで、かつ、そのセリフから複数の話者候補を除外することができる。たとえば、実験対象の作品の主要人物が用いる一人称は、Aは「俺」、BとDは

「わたし」、Cは「僕」、Eは「私」であり、もしあるセリフに「俺」が含まれる場合は、そのセリフの話者をAと確定できる。エンタメ小説では、主要人物が使う一人称は異なる場合が多いと考えられ、このステップはエンタメ小説一般に対して高い候補削減効果を持つと考えられる。

5 今後の課題

現状のシステムの課題として、以下の2点が挙げられる。

1. 主要人物のみのセリフを推定対象としており、それ以外の話者のセリフを対象外としている。
2. 周辺文脈からの候補抽出での候補収集範囲は、作品によって適切な範囲が異なるが、固定した値を用いている。

1つ目の課題に対する解決策として、口調に基づく話者推定の際に、推定話者だけでなく、各話者候補の確信度を同時に出力することを考えている。もしも全ての候補の確信度が低ければ、その中に正解の話者はいないと判断し、話者は主要人物以外の人物であると推定することが可能となる。実際にシステムを使用する際には、各セリフが主要人物のセリフかどうかは事前には分からないため、全てのセリフを推定対象とする必要がある。そのため、少なくとも話者が「主要人物以外の話者」とであると判定する機能が不可欠である。

2つ目の課題に対しては、小説の単語数に対する人名数の割合を元に、収集範囲を自動決定することを検討している。この割合が小さい場合は収集範囲を広く取り、割合が大きい場合は収集範囲を狭く取ることで、候補が集まらないといった問題や過剰に集めてしまうといった問題を解決できると考えている。

謝辞

本研究は JSPS 科研費 JP21H03497 の助成を受けたものです。

参考文献

- [1] Hua He, Denilson Barbosa, and Grzegorz Kondrak. Identification of speakers in novels. In **Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)**, 2013.
- [2] 石川和樹, 宮田玲, 小川浩平, 佐藤理史. 口調ベクトルを用いた小説発話の話者推定. Technical report, 情報処理学会第 253 回自然言語処理研究会, 2022.
- [3] 中川翔太, 孫昊, 金明哲. 小説会話文における発話者の性別推定. 日本行動計量学会大会抄録集, Vol. 47, pp. 362–365, 2019.
- [4] 津崎誠也, 山本博史. キャラクタの性格推定. 2017 年度 情報処理学会関西支部 支部大会 講演論文集, 2017.
- [5] 米田智美, 佐藤理史, 夏目和子, 宮田玲, 小川浩平. 話者の性格を反映した発話文の生成. 言語処理学会第 28 回年次大会発表論文集, 2022.
- [6] Grace Muzny, Michael Fang, Angel Chang, and Dan Jurafsky. A two-stage sieve approach for quote attribution. In **Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers**, 2017.
- [7] 遠田哲史, 吉永直樹. 文学作品における教師なし話者同定. 言語処理学会第 25 回年次大会発表論文集, 2019.
- [8] 国立国語研究所. 分類語彙表増補改訂版データベース (ver.1.0), 2004.
- [9] Mecab: Yet another part-of-speech and morphological analyzer. <https://taku910.github.io/mecab/>.
- [10] Yuji Matsumoto Taku Kudo. Japanese dependency analysis using cascaded chunking. In **CoNLL 2002: Proceedings of the 6th Conference on Natural Language Learning 2002 (COLING 2002 Post-Conference Workshops)**, pp. 63–69, 2002.

A 直前直後の文に基づく話者推定

直前直後の文に基づく話者推定では、各セリフに対して、以下に示す4つのパターンにマッチするかどうかを順に確認する。もし、いずれかのパターンにマッチしたならば、そのパターン中のAに対応する人物をそのセリフの話者だと推定する。

なお、発話動詞とは、「言った・答えた」などの発話を示唆する動詞である。発話動詞としては、国立国語研究所の「分類語彙表増補改訂版データベース」[8]から収集した940語を使用する。

パターン1

N1: (任意の地の文)

U1: (推定対象のセリフ)

N2: と… A… (発話動詞 V)…

パターン2

N1: …A…

U1: (推定対象のセリフ)

N2: と… (発話動詞 V)…

パターン3

N1: (任意の地の文)

N2: … A… (発話動詞 V)…

U1: (推定対象のセリフ)

パターン4

U1: (推定対象のセリフ)

N2: … A… (発話動詞 V)…

N1: (任意の地の文)

いずれのパターンにおいても以下の条件を満たす必要がある。

条件1

Aの係り先に発話動詞Vが存在する。

条件2

発話動詞Vの後ろに否定を表す助動詞が付かない。

この条件の判定で必要となる形態素解析と係り受け解析には、MeCab (IPAdic) [9] と CaboCha[10] を用いる。