



Title	イギリスにおけるミステリー黄金時代の女性作家作品の比較分析 Christie, SayersとAllingham
Author(s)	土村, 成美
Citation	言語文化共同研究プロジェクト. 2018, 2017, p. 19-29
Version Type	VoR
URL	https://doi.org/10.18910/69899
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

イギリスにおけるミステリー黄金時代の女性作家作品の比較 分析*

Christie, Sayers と Allingham

土村 成美

大阪大学大学院言語文化研究科

〒 560-0043 豊中市待兼山町 1-8

E-mail: t.naru.425@gmail.com

あらまし 本研究ではイギリスの女性ミステリー作家 Agatha Christie の作品の語彙的特徴に関して、他作家の作品との比較を通して分析を行うことを目的とする。比較対象として、Christie と同じくイギリスミステリー黄金時代に活躍した作家である Dorothy Sayers と Margery Allingham の作品を用いる。機械学習の一手法である Random Forests を用い、3 作家の作品の分類を行うと共に、各作家の特徴語を抽出した。長編・短編作品全てを用いた分類を行うと誤分類が発生するものの、分類対象を長編作品に限定すると、正確に分類を行うことが可能であった。固有名詞が分類に大きく寄与した変数として抽出され、固有名詞が分類精度を向上させている可能性も考えられたが、固有名詞を除外して分類を行なっても、正確な分類結果となった。Christie の特徴語としてまずは -ly 副詞が多く見られ、-ly 副詞を用いて Christie が登場人物の言動の様子を詳述しているのではないかと考えられる。また *yes* や *sure*, *you* のような会話文に特徴的な語が多く抽出された。Christie 作品は会話文を中心として物語が進められる作品が多く、そのことを反映した結果となっていると言える。

キーワード Agatha Christie, Random Forests, 文体, 特徴語

An Analysis on Female Writers' Works during Britain's Golden Age of Crime Fiction

Christie, Sayers, and Allingham

Narumi Tsuchimura

Graduate School of Language and Culture, University of Osaka

1-8 Machikaneyama-cho, Toyonaka, Osaka, 560-0043 Japan

Abstract This study investigates characteristic words of works by Agatha Christie, a female mystery writer in the UK, comparing with other authors' ones. The selected authors for the comparison are Dorothy Sayers and Margery Allingham. All of them are famous female mystery writers during Britain's golden age of crime fiction. This study applied Random Forests, a machine learning method, for classifying the three authors' works and extracting characteristic words from each author's works. The accuracy of classification was a little low when short stories were included in the data for the analysis, but when only long novels were used in Random Forests, all the texts were correctly classified into three different groups with an accuracy of

100%. First, the extracted characteristic words from Christie's works are -ly adverbs. Christie might use these words to describe how the characters in her works acted or said something. Second, spoken vocabularies like *yes*, *sure*, and *you* are also extracted as characteristic words of Christie's works. This result should be because Christie's novels contain a lot of conversations between characters.

Keywords Agatha Christie, Random Forests, style, characteristic words

1. はじめに

本研究では、イギリスのミステリー作家 Agatha Christie (1890–1976) 作品における使用語彙の特徴に関して、他作家の作品との比較を通して分析を行うことを試みる。比較対象の作家として、筆者がこれまで使用してきた、Christie との同時代作家である Dorothy Sayers (1893–1957) に加え、本研究では同じく同時代女性作家の Margery Allingham (1904–1966) の作品に関してもデータ化を行い、分析に用いた。

Christie 作品の計量的な分析は、語彙多様性に焦点を当てた Lancashire & Hirst (2009)、語彙多様性に加え統語にも焦点を当てた Le et al. (2011)、単一作品をデータ化したスモールコーパスにおけるキーワードをもとに、Christie の言葉の用い方に関して考察を行なった稲木 (2009, 2013) などが存在するが、Christie がアルツハイマー病であった疑いのある晩年の作品に分析対象データが集中していたり、単一作品のみを用いたりなど、分析対象作品が限定的である。本研究は Christie のミステリー作品の網羅的な分析を目的とする。

第一次世界対戦と第二次世界大戦の戦間期にあたる 1920 年代から 30 年代にかけては本格ミステリーの黄金時代と呼ばれており、それまで短編作品が中心であったミステリー作品界で次々と長編作品の名作が発表されるようになり、様々なミステリー作家が活躍した。その中でも Agatha Christie, Ngaio Marsh (1895–1982), Margery Allingham, Dorothy Sayers の 4 人はイギリスのミステリー黄金時代の 4 大女王と呼ばれる (Joannou (Ed.)(2013); 144)。本研究では Marsh を除いた 3 作家の作品を使用し、Sayers と Allingham の作品と比較した Christie 作品における語彙使用の特徴を計量的に分析することを試みる。

2. 使用データ, 分析手法

2.1. 使用データ

使用データの概要を以下の表 1 に示す。全作家長編、短編作品を含んでいる。3 作家とも複数ジャンルでの作品執筆を行っているが、ジャンルによる違いを最小限にするために、本研究ではミステリー作品のみをデータとして用いる。短編作品については、一つの短編集に既収録作品も含まれることが Christie の場合に特に多いため、短編集を 1 作品ごとに分けた上での作品数合計を示している。

*本研究は「言語研究と統計 2018」(2018 年 3 月 29 日, 30 日 於 統計数理研究所) において行なった研究発表「イギリスにおけるミステリー黄金時代の女性作家作品の比較分析—Christie, Sayers と Allingham—」に加筆・修正を行なったものである。指導教員の田畑智司先生 (大阪大学), ホドシチェク・ボル先生 (大阪大学) をはじめとする研究会のメンバーの皆さま, 「言語研究と統計 2018」における発表の際に有益なコメントをくださった前田忠彦先生 (統計数理研究所), 石川慎一郎先生 (神戸大学) に感謝の意を述べたい。なお, 本稿における誤りは全て筆者の責任である。

表 1: 分析に使用した作品数, 異なり語数, 総語数

著者	作品数	異なり語数	総語数
Christie	221	50,545	5,071,288
Sayers	55	40,098	1,375,645
Allingham	55	37,715	1,847,813

三次元散布図の描写の際に作品のラベリングを施しているが、長編作品は Christie, Sayers, Allingham それぞれ作者の名前の頭文字をとって C1, S1, A1 のように、短編作品は頭文字に S を付け加え CS1, SS1, AS1 のようにラベル付けを行なっている。

2.2. 分析手法

分析手法として、本研究では Breiman (2001) が提唱した、分類・回帰を行う機械学習の一種である Random Forests を使用した。Random Forests は、元データからブートストラップサンプリングされたデータを用いて、決定木を多数作成する。その多数の決定木の予測結果を組み合わせ、多数決により結果を得る。モデルの作成にデータの 3 分の 2 が用いられ、残りの 3 分の 1 を用いて作成したモデルのテストを行う。

Random Forests をテキスト分類や著者の識別に使用した研究の例として、金・村上 (2007) では、小説、作文、日記の 3 種類の文章を用い、Random Forests を利用した書き手の同定を行い、他の分類法と分類精度を比較した。その結果、Random Forests はバギングやサポートベクターマシンなどの他の分類法より正解率が高いことが報告されている。小林・田中・富浦 (2011) は、英語母語話者と非英語母語話者が執筆した英語学術論文について、メタ談話標識を変数として Random Forests を用いた分類を行い、母語話者と非母語話者それぞれに特徴的な表現の抽出を行なった。また、田畑 (2012) では、Charles Dickens と Wilkie Collins の作品を識別する特徴語の抽出を Random Forests を用いて行った。対数尤度比やカイ二乗値など従来の特徴語抽出の指標は、固有名詞のようにある特定のテキストのみに集中して出現する語を特徴語として抽出してしまうという問題点があるが、Random Forests を用いるとその問題の克服が可能であることを示した。

分析には統計解析環境 R Ver. 3.4.3 の randomForest パッケージを使用した。変数は全作家の全作品における高頻度語彙とし、分析に用いる変数を上位 100 語から 1000 語の間で 100 語ずつ変化させて分析を行った。

3. 分析結果, 考察

3.1. Random Forests の実行結果

3 作家の長編・短編全作品について、Random Forests を行なった結果を以下の表 2 と図 1 に示す。ここでは分類精度の高かった、高頻度語彙上位 700 語を変数とした際の分析結果を示す。

表 2: Random Forests 実行結果の一例

Call:

randomForest(formula = author ~.,

data=csa.all, proximity=T, importance=T, ntree=10000)

Type of random forest: classification

Number of trees: 10000

No. of variables tried at each split: 26

OOB estimate of error rate: 3.02%

Confusion matrix:

	Allingham	Christie	Sayers	class.error
Allingham	54	1	0	0.01818182
Christie	0	221	0	0.00000000
Sayers	0	9	46	0.16363636

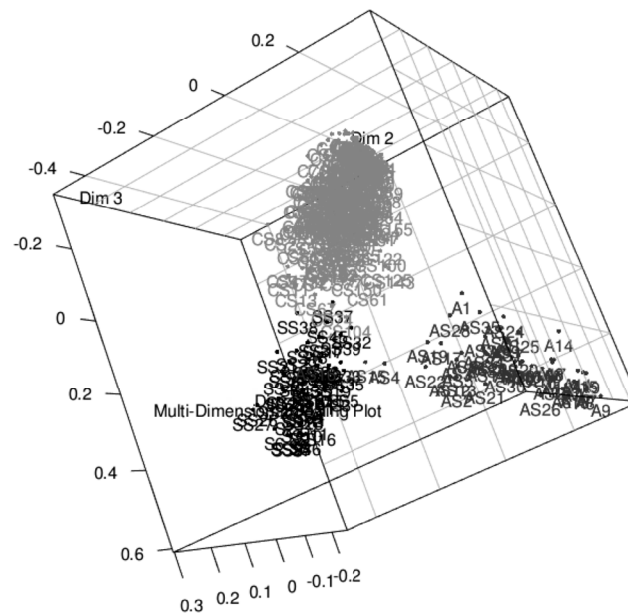


図 1: 全作品を Random Forests にかけた際の三次元散布図

分類精度は 96.98%と高く、Christie の作品は全て正しく分類出来ているものの、Sayers と Allingham は共に一部の作品が Christie 作品として誤分類されている。図 1 を見ると、図の中心付近に 3 作家の作品が混在しており、ラベルを確認すると短編作品が誤分類されている可能性が高いと考えられる。長編作品は 50,000 語を超えているものが殆どであるのに対し、短編作品は約 2,000 語から 9,000 語と語数が少ないために誤分類されている可能性が高いと考えられるため、分析対象を長編作品に限定して再度 Random Forests を行なった。長編作品に限った分析対象作品の作品数、異なり語数、総語数は表 3 の通りである。

表 3: 長編作品の分析対象作品数, 異なり語数, 総語数

著者	作品数	異なり語数	総語数
Christie	66	46,009	4,183,485
Sayers	11	36,003	1,115,019
Allingham	20	34,912	1,534,462

長編作品のみの Random Forests 実行結果を以下の表 4 と図 2 に示す。長編作品のみに関して Random Forests により分類を行うと, Christie と残りの作家との間で作品数と総語数に依然として大きな差があるものの, 分析に用いる変数を頻出語上位 700 語にした際に分類精度が 100% となった。以降はこの長編作品のみを扱った, 変数を 700 語とした際の Random Forests の結果をもとに考察を行う。

表 4: 長編作品に限定した Random Forests 実行結果の一例

Call:

randomForest(formula = author ~.,

data=csa.all, proximity=T, importance=T, ntree=10000)

Type of random forest: classification

Number of trees: 10000

No. of variables tried at each split: 26

OOB estimate of error rate: 0%

Confusion matrix:

	Allingham	Christie	Sayers	class.error
Allingham	20	0	0	0.0000000
Christie	0	66	0	0.0000000
Sayers	0	0	11	0.0000000

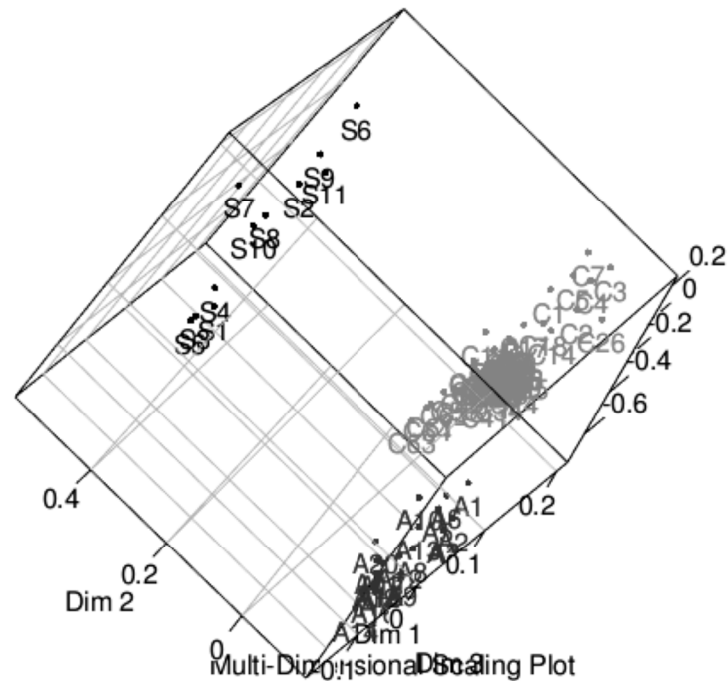


図 2: 長編作品のみを **Random Forests** にかけた際の三次元散布図

3.2. 固有名詞に関して

randomForests パッケージは、ジニ減少係数に基づいて、分類に大きく寄与した変数を求めることが可能である。ジニ係数に基づく分類に寄与した変数上位 100 語について、各作家の特徴語として分類した結果を表 5 に示す。

表 5: 3 作家の特徴語リスト

Christie
what, asked, perhaps, say, that, sure, different, kind, looked, then, yes, thoughtfully, mean, indeed, nobody, someone, do, quite, anyone, idea, have, really, are, slightly, always, chapter, believe, wrong, true, cried, shook, everyone, you, nodded, meant, sometimes, think, slowly, rather
Sayers
would, bunter, wimsey, anybody, till, if, peter, though, well, whether, be, better, might, parker, except, make, and, says, god, good, give, lord, or
Allingham
while, which, glanced, campion, appeared, became, before, its, caught, until, his, the, began, him, eyes, mouth, sighed, still, he, spoke, voice, behind, extraordinary, feet, expression, paused, were, light, was, as, round, without, smiled, towards, when, once, old, murmured

表 5 には, Sayers, Allingham それぞれの作品の主人公である *peter, wimsey, champion* のような固有名詞が特徴語として挙がってしまっている。前述した通り Random Forests は特定のテキストのみに出現する語に関しては特徴語との判断を行わないが, Sayers 作品の Peter Wimsey 卿と Allingham 作品の Mr. Campion は, 彼女らの殆どの作品において主人公として登場しているため, 特徴語として抽出されてしまったと考えられる。これら固有名詞が Random Forests による分類の精度を向上させている可能性も考えられるため, 主人公の固有名詞を変数から除去した上で再度 Random Forests を実行した。その結果以下の表 6 のように, 主人公の固有名詞を除去しても 3 作家を正確に分類することが可能であった。各作家の特徴語を分類したものを表 7 に示す。

表 6: 固有名詞を除外した Random Forests 実行結果の一例

Call:				
randomForest(formula = author ~.,				
data=csa.all, proximity=T, importance=T, ntree=10000)				
Type of random forest: classification				
Number of trees:	10000			
No. of variables tried at each split:	26			
OOB estimate of error rate:	0%			
Confusion matrix:				
	Allingham	Christie	Sayers	class.error
Allingham	20	0	0	0.0000000
Christie	0	66	0	0.0000000
Sayers	0	0	11	0.0000000

表 7: 固有名詞を除去した Random Forests における 3 作家の特徴語リスト

Christie
what, asked, perhaps, say, that, sure, different, kind, looked, then, yes, thoughtfully, mean, indeed, nobody, someone, do, quite, anyone, idea, have, really, are, slightly, always, chapter, believe, wrong, true, cried, shook, everyone, you, nodded, meant, rather, slowly, think, sometimes, stared
Sayers
would, anybody, till, if, though, well, whether, be, better, might, except, make, and, says, god, good, give, lord, or, somebody
Allingham
while, which, glanced, appeared, became, before, its, caught, until, his, the, began, him, eyes, mouth, sighed, still, he, spoke, voice, behind, extraordinary, feet, expression, paused, were, light, was, as, round, smiled, towards, old, murmured, stood, turned, face, when, without, over, once, upon, since

3.3. Sayers, Allingham と比較しての Christie 作品の特徴の考察

表7を見ると、Christie の特徴語としては-ly 副詞が挙がっている。Sayers の作品のみと比較を行った土村 (2017) においても Christie の特徴語として-ly 副詞が多く抽出されたため、Sayers 以外の作家と比較しても Christie 作品では-ly 副詞が特徴的であることが分かる。以下に特徴語として挙がっている-ly 副詞のコンコードスラインの一例を示す。

```

criminals." Superintendent Battle shook his head thoughtfully. ``I wouldn't be so C20.txt
himself out of the house. Poirot shook his head thoughtfully. The doctor had C20.txt
ch notes. Anything there?" Poirot shook his head thoughtfully. ``I hardly think C22.txt
n the end he rose to his feet, and shook his head thoughtfully. He seemed baffled C3.txt
Rich travelling Americans too." He shook his head thoughtfully. ``Wonderful C56.txt
de you hit on Eccles?" Ivor Smith shook his head thoughtfully. ``It would take too C59.txt
must have heard by now." ``Yes," said Henrietta thoughtfully. ``I suppose she's C37.txt
perhaps a note or two not more," said Sir Henry, thoughtfully. ``Did this Miss C40.txt
ns justifying that action." Poirot looked at her thoughtfully. ``That is a C13.txt
le," said the girl crisply. Poirot looked at her thoughtfully. ``You are C14.txt
Something cold " Mr Satterthwaite looked at her thoughtfully for a minute or two, C16.txt
ly, what else could I say?" Poirot looked at her thoughtfully. ``You, Madame, C17.txt
ly, ``or you wouldn't be here." He looked at her thoughtfully for a moment. C17.txt
s Miss Johnson who spoke. Mr Poirot looked at her thoughtfully. I had an idea that C19.txt
ng." She spoke impatiently. Poirot looked at her thoughtfully as he asked: ``Has C22.txt
``Sure?" ``N-no no, indeed." Luke looked at her thoughtfully. The denial, he C25.txt
Her breath came rather faster. Luke looked at her thoughtfully. ``So you feel like C25.txt
that's the best thing to do?" Luke looked at her thoughtfully, then he said: C25.txt
d who's to contradict him?" Philip looked at her thoughtfully. ``You know," he C26.txt
Underhand in her Dealings." Poirot looked at her thoughtfully for a moment. Then C27.txt
. M. Poirot, is important." Poirot looked at her thoughtfully for a minute or two. C29.txt

```

図 3: *thoughtfully* のコンコードスライン (1)

```

is. Head over ears." ``I thought so," said Jane thoughtfully. ``Why?" ``From C2.txt
``I've never been in a revolution," said Jennifer thoughtfully. ``I don't expect C51.txt
ns the cord." ``Looks like a doctor," said Jenny thoughtfully. ``Did Miss Adams C13.txt
names and dates." ``Names and dates," said Jimmy thoughtfully. ``Gerry didn't C9.txt
actly." ``I'd rather have peacocks," said Joanna thoughtfully. ``I'd like a C33.txt
is against humanity." ``I suppose," said Joanna thoughtfully, ``that that is what C33.txt
er took away." ``Another shock, eh?" said Julius thoughtfully. ``Exactly. There C2.txt
a kind of anti-climax to last night," said Julius thoughtfully. ``Yes, it does," C2.txt
next Sunday is the 29th!" ``Shucks!" said Julius thoughtfully. ``I'd almost C2.txt
hip hand of us. See?" ``Sure thing," said Julius thoughtfully. ``That's so." C2.txt
oche. ``You think he killed her," said Katherine thoughtfully. ``That is the C8.txt
oirot." ``The Comte de la Roche," said Katherine thoughtfully; 'so you really C8.txt
impression is it not so?" ``Yes," said Katherine thoughtfully. ``I suppose I did." C8.txt
e Blanche." ``Mademoiselle Blanche," said Kelsey thoughtfully. Hercule Poirot C51.txt
red?" ``Death of a Games Mistress," said Kelsey, thoughtfully. ``Sounds like the C51.txt
So she's the one who had the torch," said Kelsey thoughtfully. ``She came out here C51.txt
Hersheimer." ``Hersheimer," repeated Kramenin thoughtfully. ``I have heard that C2.txt
there." ``MacKenzie died out there," said Lance thoughtfully. ``Wasn't there a C45.txt
``She's really very useful." ``Yes," said Lance thoughtfully. ``I should think C45.txt
Chapter 8 Mr Entwhistle looked at Dr Larraby thoughtfully. He had had a C44.txt
That is possible, of course," said Father Lavignv thoughtfully. ``Did you speak to C19.txt

```

図 4: *thoughtfully* のコンコードスライン (2)

"Nobody else?" "Well er Miss Grey." He flushed **slightly** as he spoke the name. C18.txt
 on his back on his bed, smoking. Hilary flushed **slightly**. "We arrived here C46.txt
 h over to spare for your fee." Sir James flushed **slightly**. "There is no question C2.txt
 ood deal of attention, don't they?" Lucy flushed **slightly**. Scrappy remembrances C49.txt
 t Captain Marshall's knowledge?" Redfern flushed **slightly**. He said: "I don't C29.txt
 ight." "We?" "Sir Charles and I." She flushed **slightly**. Mr Satterthwaite C16.txt
 rests you? Polly, you're wonderful!" She flushed **slightly**. "I don't see what's C50.txt
 think I'm making a lot of nothing." Tom flushed **slightly**. "Not at all. You were C18.txt
 n the past. Nannie saw my expression and flushing **slightly**, she said: "When I C39.txt
 ct, he didn't see me," said Miss Marple, flushing **slightly**. "Because, you see, C10.txt
 essary." "Oh, of course." Tommy said, flushing **slightly**: "I don't like lying C30.txt
 shot or even played a decent game. A foreigner. **Slightly** amused, Hercule Poirot C32.txt
 bs. She was a woman of between thirty and forty, **slightly** haggard, with a clear- C65.txt
 ling of ideas between men looking at affairs from **slightly** different points of C56.txt
 used effect that that remark always caused. From **slightly** behind her Jason Rudd C53.txt
 He put the speaking-tube down again, and frowned, **slightly** tapping the table with C4.txt
 oors into the hall of Bertram's Hotel. He frowned **slightly**, wondering what it was C56.txt
 tation with Mr Crawford " "Crawford?" I frowned **slightly**. "Mr Crawford of Reece C58.txt
 sentiments. Hercule Poirot. Katherine frowned **slightly**. It was as though C8.txt
 dre." "Thank you, my dear boy." Poirot frowned **slightly**. Robin handed drinks to C42.txt
 k her head. "Or a Mr Basil Blake?" She frowned **slightly**. "I think I've heard C31.txt

図 5: *slightly* のコンコードンスライン

図 3 から図 5 に -ly 副詞の特徴語として挙げられている *thoughtfully* と *slightly* のコンコードンスラインの一部を示した。これらの図を見て分かるように、*thoughtfully* は会話文の直後や、*shook one's head*, *looked at* などと共に用いられている。Christie は登場人物の発言や行動を描写する際に -ly 副詞を用いて、言動の様子を詳述していることが伺える。

次に、Christie の特徴語として *yes* と *sure* も挙げられている。以下にこの 2 語のコンコードンスラインの一例を示す。

"about a young girl. A girl called Verity." "Ah **yes**. Verity Hunt." "I did not C62.txt
 say: "I am determined so to prove a villain." Ah **yes**. The Black Arrow." "Some C64.txt
 lock, you know, a piece of water, it was." "Ah **yes**, I see. And Pamela thought C64.txt
 ere was the Johnsons, of course, and there was ah **yes**, one of the constables was a C64.txt
 Crossfield. Which of the two girls was which? Ah **yes**, that was Rosamund, C44.txt
 was her last husband's, Isidore Wright's." "Ah **yes**, her last husband. Where is C53.txt
 ed then; we have only been married a year." "Ah **yes**, thank you, Madame. By the C14.txt
 e Mr Eccles. Mr Eccles looked down. "Ah yes, ah **yes**, I remember now. Yes, I C59.txt
 " Tommy obliged with his address. "Ah, yes, ah **yes**, I put the right thing on the C64.txt
 dignity. "You saw us together yesterday." "Ah, **yes**. But I'm afraid we shan't C2.txt
 been able to get down to it properly yet." "Ah **yes**, you are one of those who C42.txt
 I know he is very anxious to talk to you." "Ah! **Yes?** Shall I go and see him now?" C24.txt
 aid coldly. "You have told her yourself." "Ah, **yes**, but it is more agreeable to C12.txt
 much interest to people," said Poirot. "Ah-ha, **yes!** One knew a lot of people in C57.txt
 an air of great surprise, picked it up. "Ah-ha, **yes**. This is it. So it was you C56.txt
 looked puzzled for a moment then he said: "Aha! **yes!** your Shakespeare!" "You C46.txt
 up a pen and prepared to write. "Yes. Go ahead. **Yes**." He wrote. "What? How do C50.txt
 s orders to view and invited the waiter's aid. "Yes, sir. I know where most of C21.txt
 ell, not quite, perhaps. I had one other aim." "Yes?" "I wanted to watch the C16.txt
 mself: "Une qui aime et un qui se laisse aimer. **Yes**, I wonder too." VII C22.txt
 when?" "The following day." "Also by air?" "Yes." "You're quite sure, are C56.txt

図 6: *yes* のコンコードンスライン

ck." ``He has never said that you are that, I am sure." ``Oh, not said so. Of	C65.txt
't." ``You seem rather sure about that." ``I am sure." ``Then why did she go?"	C66.txt
vories. Exquisite! You will appreciate them, I am sure, if I succeed in acquiring	C52.txt
ab." ``Perhaps you are right. Bridge, then. I am sure he will feel that Bridge is	C37.txt
you think she hides her love letters there? I am sure she has never had a love	C51.txt
emotional but urgent, yes! And of one thing I am sure. He was in full possession	C37.txt
all we say? Well, that is a very good thing, I am sure." ``I'm starting at long	C58.txt
k earnestly into those past days. ``I think I am sure he cared for me at first.	C1.txt
s to hear all you could tell me. No, I think I am sure you are quite safe." ``If	C19.txt
Bowers to take charge of her. ``And I think I am sure that Miss Bowers stayed with	C22.txt
hink " ``It is not a question of thinking. I am sure!" That evening, students	C47.txt
t was my duty to go through. You understand, I am sure, that I shall always feel a	C50.txt
waste your time unduly, you understand." ``I am sure you will not waste my time.	C60.txt
?" ``Not yet," said Poirot. ``Not yet until I am sure that one of my little ideas	C60.txt
ongress? Any points of interest come up?" ``I am sure there will be." ``Just come	C56.txt
``That does great credit to your vigilance, I am sure. But why the little	C22.txt
ou think there is danger to our visitor?" ``I am sure of it." Mrs Pearson greeted	C7.txt
sembled students in a low thrilled voice. ``I am sure now, but absolutely sure I	C47.txt
So did Bryan Martin. This remark of his was, I am sure, equally unexplicable to us	C13.txt
irots. I disliked and distrusted him. He was, I am sure, a cruel and a dangerous	C14.txt
ot that I wish I could remember what it was. I am sure her face reminds me of	C17.txt

図 7: *sure* のコンコーダンスライン

コンコーダンスラインを見ても分かるように、*yes* は会話文で主に用いられている。*sure* についても、*I am sure*, *I'm sure* の形で会話文で多く用いられている。二人称代名詞の *you* が Christie の特徴語として含まれていることから、Christie 作品が登場人物同士の会話文を中心として構成されていることを反映していると言える。

4. おわりに

本研究では、Random Forests を用いて Christie, Sayers, Allingham 作品の分類を行い、Sayers, Allingham 作品と比較しての Christie 作品の特徴語の抽出を行なった。はじめに短編作品と長編作品全てを用いて分類を行なったところ、一部の作品が誤分類される結果となった。分類結果の作品の散布図を確認すると、短編作品が誤分類されている可能性が高かったため、長編作品のみで分類を行なった結果、分類精度が向上した。しかしこの段階では各作家の作品の主人公の固有名詞が分類に大きく寄与した語として抽出されており、これらの固有名詞が分類精度を向上させている可能性があったため、これらの語を分析対象から除外して再度分類を行った。その結果、3 作家の作品を正確に分類することが可能であった。

Random Forests による分類で分類に寄与した上位の語のうち、Christie 作品に特徴的な語を見てみると、はじめに特徴的であったのは *-ly* 副詞であった。Christie は登場人物の言動を表す表現と共に *-ly* 副詞を用いており、そのことによって登場人物の言動を詳述していると考えられる。

また、Christie の特徴語として *sure*, *yes* も挙がっており、これらの語は主に会話文で用いられていることが確認され、二人称代名詞 *you* が Christie 作品の特徴語として挙がっていることから、Christie の作品が会話文を中心として進められているという特徴が表れている結果となっている。

今後の課題として、第一にイギリスミステリー黄金時代の 4 大女王の中で Ngaio Marsh の作品のコーパス化が終了していないため、Marsh 作品のデータ化を早急に行う予定である。

第二に、今回の Random Forests による分類では、長編作品のみにおいては正確に分類を行うことが出来たが、短編作品も含めると、短編作品の語数が十分でないためか、誤分類が発生した。どの程度まで作品ごとの語数を減少させると誤分類が発生するのか分析を行うことも将来的に考えている。

文 献

- [1] Breiman, L. (2001). Random Forests. *Machine Learning*, 45: pp.5–23
- [2] 稲木昭子 (2009) 「発話動詞と共起する -ly 副詞：誤誘導のテクニック」『英語文化学会論集』18: 1–14.
- [3] 稲木昭子 (2013) 『謎解きのことば学：アガサ・クリスティの英語を楽しむ』東京：英宝社.
- [4] 金明哲・村上征勝 (2007) 「Random Forests 法による文章の書き手の同定」『統計数理』55(2): pp.255–268.
- [5] Joannou, M. (Ed.) (2013). *The History of British Women's Writing, 1920–1945: Volume Eight*. London: Palgrave Macmillan.
- [6] 小林雄一郎・田中省作・富浦洋一 (2011) 「ランダムフォレストを用いた英語科学論文の分類と評価」『情報処理学会研究報告 IPSJ SIG Technical Report』第 90 号 (2011-CH-90): 53–68.
- [7] Lancashire, I. and Hirst, G. (2009, March). *Vocabulary Changes in Agatha Christie's Mysteries as an Indication of Dementia: A Case Study*. Paper presented at the 19th Annual Rotman Research Institute Conference, Cognitive Aging: Research and Practice, Toronto.
- [8] Le, X., Lancashire, I., Hirst, G. and Jokel, R. (2011). Longitudinal detection of dementia through lexical and syntactic changes in writing: a case study of three British novelists. *Literary and Linguistic Computing*, 26(4): pp.435–461.
- [9] Malling, S. and Peters, B. (ed.) (1997). *AZ Murder Goes: Classic*. Scottsdale: Poisoned Pen Press.
- [10] 田畑智司 (2012) 「Dickens と Collins の共著作品への文体統計学的アプローチ」『情報処理学会研究報告 人文科学とコンピュータ研究会報告』93(3): pp.1-7.
- [11] 土村成美 (2017) 「Random Forests を使用した Agatha Christie 作品の文体分析：Dorothy Sayers と比較して」『言語文化共同研究プロジェクト 2016 テクストマイニングとデジタルヒューマニティーズ』pp. 43–54.