

2021年度（令和3年度） 卒業論文

BERTを用いた対義語の  
置き換え可否判定

電気情報系学科 卒業論文検印	
学科長	

指導教員

村田真樹  
村上仁一

鳥取大学工学部 電気情報系学科

自然言語処理研究室

B18T2044B 小西 択磨

## 概要

本研究は対義語に対して、教師あり機械学習を用いることにより、対義語の置き換え可否の判定を行うことを目指す。

対義語とは、意味が反対になっている語や意味が対称になっている語のことである。対義語には、文中のある単語を対義語に置き換える場合、対義語に沿うように文の内容を逆にしても対義語に置き換えができないものもある。例えば「足す」と「引く」という対義語の組では、「A から B を引く」とはいうが、「A から B を足す」とはいわない。これは格が変化しているからである。本研究では置き換え可能の定義は、文中のある単語 X の対義語に沿うように文を知識的な内容だけ逆に変更した時 (直前直後の付属語、係り受け関係にある文中の格助詞は変更しない)、単語 X を対義語に置き換えられる場合、置き換え可能であると定義する。荻原ら [1] は、分類語彙表に反対語 (ここでは対義語と同様に扱う) 情報を付与、またその際、置き換え可否の情報を付与している。しかし、この作業は人手で行っており置き換え可否の情報付与に関しては例文を示していない。実際に置き換えてみないと判断がつかないことも多い。よって、例文を示さずに対義語対に置き換えの可否の情報を付与することは、正確性に欠ける。また、佐々本ら [3] は機械学習の性能や素性が対義語対の使い分けに役立つと考え、機械学習の最大エントロピー法を用いて対義語対の使い分けを行い、その結果を用いて対義語対の置き換え可否を判定する研究を行った。本研究は佐々本らの研究を、機械学習の BERT[4] を用いて行う。ある対義語間での機械学習の性能が高く、より正確に使い分けを行っていた場合は、その対義語の組は機械学習で性能が高くなるほど使い方の違いが明瞭な対義語対とわかる。使い方の違いが明瞭だと対義語対は置き換え不可能だと考えられる。対して、ある対義語間での機械学習の性能が低い場合は、正確に使い分けができず、その対義語対は機械学習で判断ができないほど使い方の違いが不明瞭な対義語対とわかる。よって使い方の違いが不明瞭なので、その対義語対は置き換え可能であると考えられる。この考え方を用いて機械学習の性能を用いて対義語対の置き換え可否の判定を行う。このような実験を荻原らの研究から獲得した対義語の組を対象に行う。対義語対の使い分けした結果を、例文を示した対義語対の置き換え可

否を判定する被験者実験の結果と比較する。本研究の成果は2つある。1つ目は、機械学習の性能が高ければ置き換え可能であり、機械学習の性能が低ければ置き換え不可能であるという、機械学習の性能と対義語対の置き換え可否に逆の相関があることが、佐々本ら [3] が行った最大エントロピー法と同様に BERT でも確認できた。2つ目は、それぞれの対義語における有用な知見や対義語対の置き換え可否に関する知見などを獲得した。例えば、慣用的な表現が多く使われていると置き換え可能の値が高くなることが、佐々本ら [3] の研究と同様に確認できた。この研究の成果は日本語学習者向けの対義語学習教材作成に利用できる。

# 目次

第1章	はじめに	1
第2章	先行研究	3
2.1	『分類語彙表』に対する反対語情報付与	3
2.2	機械学習を用いた対義語の置き換え可否判定	4
第3章	問題設定と提案手法	5
3.1	問題設定	5
3.2	提案手法	6
3.3	BERT	6
3.4	素性	7
第4章	実験	9
4.1	実験データ	9
4.2	実験方法	14
4.3	検定方法	15
4.3.1	符号検定	15
4.3.2	無相関検定	16
4.3.3	相関係数の差の検定	17
4.4	使い分けの実験結果	18
4.5	被験者実験	27
第5章	考察	32
5.1	被験者実験と機械学習の性能の傾向が一致した対義語対の考察	32
5.1.1	分類分けが高で置き換え可の値が低かったもの	32
5.1.2	分類分けが低で置き換え可の値が高かったもの	37
5.2	被験者実験と機械学習の性能の傾向が一致しなかった対義語対の考察	41

5.2.1	分類分けが低だが置き換え可の値が低かったもの . . . . .	41
5.2.2	分類分けが中だが置き換え可の値が高かったもの . . . . .	43
5.2.3	分類分けが中だが置き換え可の値が低かったもの . . . . .	45
5.3	BERT と最大エントロピー法の被験者実験の考察 . . . . .	46
<b>第6章</b>	<b>おわりに</b>	<b>48</b>

# 表 目 次

3.1 対義語の判別に用いる素性 . . . . .	8
4.1 実験を行った名詞の対義語の組 . . . . .	10
4.2 実験を行った動詞の対義語の組 . . . . .	11
4.3 実験を行った形容詞の対義語の組 . . . . .	12
4.4 実験を行った副詞の対義語の組 . . . . .	13
4.5 被験者実験を行った対義語の組 . . . . .	13
4.6 被験者実験の例 . . . . .	14
4.7 対義語の分類結果 . . . . .	18
4.8 名詞の対義語対の結果 1 . . . . .	19
4.9 名詞の対義語対の結果 2 . . . . .	20
4.10 動詞の対義語対の結果 1 . . . . .	21
4.11 動詞の対義語対の結果 2 . . . . .	22
4.12 形容詞の対義語対の結果 1 . . . . .	23
4.13 形容詞の対義語対の結果 2 . . . . .	24
4.14 副詞の対義語対の結果 . . . . .	25
4.15 BERT と最大エントロピー法の再現率の平均 . . . . .	25
4.16 BERT が最大エントロピー法の再現率を上回った個数 . . . . .	26
4.17 有意差検定 . . . . .	26
4.18 BERT の被験者実験の結果 1 . . . . .	27
4.19 BERT の被験者実験の結果 2 . . . . .	27
4.20 BERT の被験者実験の結果 3 . . . . .	28
4.21 最大エントロピー法の被験者実験の結果 1 . . . . .	29
4.22 最大エントロピー法の被験者実験の結果 2 . . . . .	29
4.23 最大エントロピー法の被験者実験の結果 3 . . . . .	30
5.1 機械学習の性能「迎撃」「出撃」) . . . . .	32

5.2	機械学習が参考にした性能(「迎撃」「出撃」)	33
5.3	機械学習の性能「遅れる」「進む」)	33
5.4	機械学習が参考にした性能(「遅れる」「進む」)	34
5.5	機械学習の性能「鋭い」「鈍い」)	34
5.6	機械学習が参考にした性能(「鈍い」「鋭い」)	35
5.7	機械学習の性能「うっすら」「はっきり」)	35
5.8	機械学習が参考にした性能(「うっすら」「はっきり」)	36
5.9	機械学習の性能「男声」「女声」)	37
5.10	機械学習が参考にした性能(「男声」「女声」)	37
5.11	機械学習の性能「増える」「減る」)	38
5.12	機械学習が参考にした性能(「増える」「減る」)	38
5.13	機械学習の性能「黒い」「白い」)	39
5.14	機械学習が参考にした性能(「黒い」「白い」)	39
5.15	機械学習の性能「憎い」「いとしい」)	41
5.16	機械学習が参考にした性能(「憎い」「いとしい」)	41
5.17	機械学習の性能「固まる」「溶ける」)	42
5.18	機械学習が参考にした性能(「固まる」「溶ける」)	42
5.19	機械学習の性能「公設」「私設」)	44
5.20	機械学習が参考にした性能(「公設」「私設」)	44
5.21	機械学習の性能「閉める」「開ける」)	45
5.22	機械学習が参考にした性能(「閉める」「開ける」)	45
5.23	BERT と最大エントロピー法の相関係数	46
5.24	最大エントロピー法の無相関検定の結果	47
5.25	BERT の無相関検定の結果	47
5.26	相関係数の差の検定の結果	47

# 第1章 はじめに

対義語とは、意味が反対になっている語や意味が対称になっている語のことである。対義語には、文中のある単語を対義語に置き換える場合、対義語に沿うように文の内容を逆にしても対義語に置き換えができないものもある。例えば「引く」と「足す」で「A から B を引く」とはいうが、「A から B を足す」とはいわない。これは格が変化しているからである。本研究では置き換え可能の定義は、文中のある単語 X の対義語に沿うように文を知識的な内容だけ逆に変更した時(直前直後の付属語、係り受け関係にある文中の格助詞は変更しない)、単語 X を対義語に置き換えられる場合、置き換え可能であると定義する。また、対義語に関する研究では、荻原ら [1] が、分類語彙表に対義語情報を付与、その際、置き換え可否の情報を付与している。しかし、この作業は人手で行っており置き換え可否の情報付与に関しては例文を示していない。実際に置き換えてみないと判断がつかないものも多い。よって、例文を示さずに対義語対に置き換えの可否の情報を付与することは、正確性に欠ける。また、佐々本ら [3] は機械学習の性能や素性が対義語対の使い分けに役立つと考え、機械学習の最大エントロピー法を用いて対義語対の使い分けを行い、その結果を用いて対義語対の置き換え可否を判定する研究を行った。しかし佐々本らの研究では、機械学習の最大エントロピー法しか使用していなかった。

そこで本研究では、機械学習の BERT [4] を用いて対義語対の使い分けを行い、その結果を用いて対義語対の置き換え可否を判定する。機械学習を用いた対義語の使い分けの性能が悪い場合、対義語の使い方の違いが不明瞭であると考えられる。よって、対義語は置き換え可能であると考えられる。一方で、使い分けの結果が良い場合、対義語対の使い分けの違いが明瞭であると考えられる。よって、対義語は置き換え不可能であると考えられる。よって、機械学習の使い分けの性能は置き換えの可否と逆の相関があると考えられる。よって、機械学習の使い分けの性能は置き換えの可否と逆の相関があると考えられる。この考え方を用い、機械学習の BERT を用いて対義語の使い分けを用い、対義語対の置き換えの可否を判定した。

本研究の成果は、人手で行っていた作業を自動化し、また日本語学習者の対義語学



習などに利用できると思う。本研究では、荻原らの研究 [1] で、20 人の被験者の内置き換え 17 人以上が対義語であると判断したものを利用する。機械学習を行った対義語対の一部から例文を明示し被験者実験を行い、その結果を機械学習の性能と比較する。機械学習の性能と例文を提示し行った被験者実験の結果を比較し、機械学習の性能と被験者実験の結果に逆の相関があることを確認することを目的とする。本研究の主な主張点を以下に整理する。

- 機械学習の性能と例文を明示した被験者実験の結果を比較し、名詞や副詞では機械学習の性能と対義語対の置き換えの可否は逆の相関があることを、佐々本ら [3] が行った最大エントロピー法と同様に BERT でも得られた。
- 対義語の機械学習の素性を分析することにより、それぞれの対義語における有用な知見や対義語対の置き換え可否に関する知見などを獲得した。例えば、慣用的な表現が多く使われていたり、対義語に多義性があると置き換え可の値が高くなることが、佐々本ら [3] の研究と同様に確認できた。

本論文の構成は以下の通りである。第 2 章では、本研究に関連する研究としてどのような研究が行われてきたかを記述し、その研究と本研究との関連を説明する。第 3 章では、本研究が扱う問題の設定とそれを解決するために提案した手法について説明を行う。第 4 章では、本研究が行った機械学習の実験についての説明と、被験者実験についての説明と、それら 2 つの結果を比較した結果について記述する。第 5 章では、第 4 章の結果から素性分析による考察を行う。第 6 章ではまとめを行う。

## 第2章 先行研究

本章では、先行研究について記述する。2.1節では、荻原ら [1] が行った分類語彙表に対義語情報を付与をした研究について記述する。2.2節では、佐々本ら [3] が行った最大エントロピー法を用いた対義語の置き換え可否判定について記述する。

### 2.1 『分類語彙表』に対する反対語情報付与

荻原らは、『分類語彙表増補改訂版』(国立国語研究所 2004) (以下、分類語彙表) 中の反対語 (この論文では対義語と同様に扱う) を枚挙し、対義語情報を整備した [1]。分類語彙表は「語を意味によって分類・整理したシソーラス (類義語集)」で、意味の区切り行を入れて 101,070 件からなる。意味の区切り行というのは、例えば既婚なら「体-活動-行為-身上」というように意味ごとに区切ったものである。語句に対して、5桁の数字からなる分類番号、2桁以下の数字からなる段落番号、2桁以下の数字からなる小段落番号、2桁以下の数字からなる語番号が付与されている。分類語彙表の分類では、類義語だけでなく、反対語・対義語なども同じカテゴリに分類される。「体-自然-生物-生物」なら「有性」⇔「無性 (むせい)」や「体-活動-言語-名」なら「同性」⇔「異性」などが、反対語に相当する。しかしながら、「同性」「異性」に関しては、どちらも「両性」というカテゴリーに属していることもあり、反対語か否かの判定は作業者により揺れる。何を反対語としてとらえるかは、人によって異なる。例えば二律関係だけでなく、三律関係を反対語としてとらえる人もいる。また反対語の中にも、文脈によって置き換え不可なものもある。このような情報を評価するために、分類語彙表から反対語対候補を収集し、「反対語としての認識」「置き換えの可否」についてクラウドソーシングを用いた調査による分布情報を付与する。クラウドソーシングはYahoo!クラウドソーシングを用い、1語対あたり20人の作業者に判断を依頼した。作業者は、「対義語・反対語でない」・「対義語・反対語であるが、置き換え不可」・「対義語・反対語であり、置き換え可」の3択のいずれかを選択する。「対義語・反対語であるが、置き換え不可」の例として、格が変化する (AにBを)「加算する」、(AからBを)「減

算する」を呈示した。「対義語・反対語であり、置き換え可」の例として、「北」、「南」、「暑い」、「寒い」を呈示した。

荻原らの研究の成果として、「体-関係-空間-内外」「体-関係-空間-方向・方角」など1次元的な軸が仮定できるものは、対義語であり、かつ、置き換え可である割合が高く、「用-関係-作用増減・補充」「用-関係-作用-作用・変化」などは、対義語と認定される割合が高い一方、格標示が変化する可能性があり、置き換え不可であるものが増えていくということなどが分かった。また、何を対義語としてとらえるかは人によって異なるため、対義語か否かの二律背反的な定義は困難であることが伺えた。

この先行研究は、対義語情報を分類語彙表に付与しその際に置き換えの可否の情報も付与しているが、置き換え可否に関しては、例文を提示していないので実際に置き換えながら格の変化の確認や、言い回しの確認ができないため正確性にかけてと考えられる。

## 2.2 機械学習を用いた対義語の置き換え可否判定

佐々本ら [3] は機械学習の性能や素性が対義語対の使い分けに役立つと考え、機械学習を用いて対義語対の使い分けを行い、その結果を用いて対義語対の置き換え可否を判定する研究を行った。

佐々本らは対義語の使い分けのために機械学習の最大エントロピー法を使用し、対義語の使い分けを行う。そして、結果を分類することにより、対義語の置き換えの可否を判定する。

対義語は、荻原らの研究 [1] で、20人の被験者の内置き換え17人以上が対義語であると判断したものを利用する。

対義語対ごとに対義語の使い分けの実験を行った。入力文は、1991年～1995年、2011年～2015年の毎日新聞から獲得した、対義語対のいずれかの語を含む文である。評価は10分割のクロスバリデーションで行った。機械学習の再現率の高さごとに名詞の類義語対を、高・中・低に分類し、機械学習における素性(学習に用いる情報のこと)を分析することで類義語の使い分けに重要な情報を把握した。

佐々本らの研究の成果として、機械学習の性能が高ければ置き換え可能であり、機械学習の性能が低ければ置き換え不可能であるという、機械学習の性能と対義語対の置き換え可否に逆の相関があることが確認できた。更に、それぞれの対義語における有用な知見や対義語対の置き換え可否に関する知見などを獲得した。

## 第3章 問題設定と提案手法

本章では、本研究で扱う問題と提案手法の説明を記述する。3.1節では、本研究で扱う問題設定について記述している。3.2節では、提案手法の大きな流れについて記述している。3.3節では、本研究で使用する機械学習法であるBERTについての説明を記述している。3.4節では、機械学習で使用する素性について記述している。

### 3.1 問題設定

使い分けたい対義語の組A,Bがあるとする。語Aと語Bのことを対象語と呼ぶ。対象語のいずれかを含む文を収集する。収集した文において対象語を削除し、対象語があった箇所に対象語のうちどの語が存在したかを推定する。その結果を用い置き換えの可否を判定することが、本研究で扱う問題である。その文に元々あった方の語を選択できれば、正しく対義語を使い分けることができ、置き換えのできない対義語だと考える。逆に、元々あった方の語を選択できなければ、正しく対義語を使い分けることができず、置き換えのできる対義語だと考える。具体的な例として、対義語の組「賛成」「反対」を例に以下に示す。

第三号被保険者制度は不公平という意見に賛成です。  
自由化を前提にした農業の新政策による新たな農民の選別切り捨てに反対する。

このように対象語を含んだ文を収集する。次にこれらの文から対象語を削除する。

第三号被保険者制度は不公平という意見に名詞です。  
自由化を前提にした農業の新政策による新たな農民の選別切り捨てに名詞する。

意味のない単語として、名詞は「名詞」、動詞は「する」、形容詞は「そう」、副詞は「いい」に対象後を置き換える。名詞とした箇所に対象語のうちどちらが存在したかを

機械学習で推定する。

## 3.2 提案手法

本研究では、教師あり機械学習を利用して、対象語のうちどの語が文中にあったのかを推定し、その結果を用い置き換えの可否を判定する。対象語のいずれかを含む文を学習データとして用いる。その文が含む対象語をその文の分類先として、機械学習を用いて学習を行う。教師あり機械学習にはBERTを利用する。再現率の高さごとに結果を高（90%以上）と中（80%以上90%未満）と低（80%未満）の3つに分類し、高が置き換えしにくいもの、低が置き換えしやすいもの、中がそれらの中間として、置き換え可否の判定を行う。

分類に再現率を用いるのは、再現率は機械学習が実験データのうちどれだけ正解を認識したかという指標であるためである。

## 3.3 BERT

BERTとは、Bidirectional Encoder Representations from Transformersの略で、「Transformerによる双方向のエンコード表現」と訳され、2018年10月にGoogleのJacob Devlinらの論文で発表された自然言語処理モデルである[4]。従来の自然言語処理では、大量のラベルのついたデータを用意させ、処理を行うことで課題に取り組む。しかし従来の手法に対し、BERTは事前学習でラベルのないデータをはじめに大量に処理を行う。その後、ファインチューニングで少量のラベルの付いたデータを使用することで課題に対応させる。

BERTでは、どのような語があれば記事中に単語対が出現するかを学習する。例えば、単語対A、Bがあるとすると、A、Bを含む文を収集する。収集した文中からA、Bを削除し、Xとする。Xとした部分にA、Bどちらの語があったのかを学習結果を元に推定する。

BERTの素性分析の方法について述べる。BERTを用いてテストデータを1単語ずつに分ける。分けた1単語に対してBERTを用いることで、分けた1単語の分割、非分割に関するそれぞれ値が算出される。算出された値が分割の値が大きい場合、分割に関する素性、非分割の値が大きい場合、非分割に関する素性であると判断する。算出され

た値が分割の値が大きい場合, 分割に関する素性, 非分割の値が大きい場合, 非分割に関する素性であると判断する.

### 3.4 素性

機械学習の素性には表 3.1 のものを用いる. これらの素性を, 対象語が含まれる文から取り出す. 表 3.1 中に記述されている分類語彙表の番号とは, 分類語彙表によって与えられた語ごとの意味を表す 10 桁の番号である. 類義語の使い分けでは, 文中に存在する語から使い分けに関する情報が得られると考え, 素性 1 を設定する. その中でも対象語の前後の語に重要な情報があると考え 素性 2, 3 を設定する. また, 対象語の存在する文構造にも情報があると考え, 対象語の存在する文節の付属語, 対象語の存在する文節に係る文節, 対象語の存在する文節に係る文節の自立語と付属語をこれらの語彙情報とともに素性として設定する (素性 4-45).

表 3.1: 対義語の判別に用いる素性

番号	素性の説明
素性 1	文中の名詞
素性 2	対象語の前後 3 語
素性 3	2 の分類語彙表の番号 7,5,4,3,2,1 桁
素性 4	対象語が含まれる文節の自立語
素性 5	4 の品詞
素性 6	4 の分類語彙表の番号 7,5,4,3,2,1 桁
素性 7	対象語が含まれる文節の最初の自立語
素性 8	7 の品詞
素性 9	7 の分類語彙表の番号 7,5,4,3,2,1 桁
素性 10	対象語が含まれる文節の最後の自立語
素性 11	10 の品詞
素性 12	10 の分類語彙表の番号 7,5,4,3,2,1 桁
素性 13	対象語が含まれる文節の付属語
素性 14	13 の品詞
素性 15	13 の分類語彙表の番号 7,5,4,3,2,1 桁
素性 16	対象語が含まれる文節の最初の付属語
素性 17	16 の品詞
素性 18	16 の分類語彙表の番号 7,5,4,3,2,1 桁
素性 19	対象語が含まれる文節の最後の付属語
素性 20	19 の品詞
素性 21	19 の分類語彙表の番号 7,5,4,3,2,1 桁
素性 22	対象語が含まれる文節に係る文節の自立語
素性 23	22 の品詞
素性 24	22 の分類語彙表の番号 7,5,4,3,2,1 桁
素性 25	対象語が含まれる文節に係る文節の付属語
素性 26	25 の品詞
素性 27	25 の分類語彙表の番号 7,5,4,3,2,1 桁
素性 28	対象語が含まれる文節に係る文節の最初の自立語
素性 29	28 の品詞
素性 30	28 の分類語彙表の番号 7,5,4,3,2,1 桁
素性 31	対象語が含まれる文節に係る文節の最後の自立語
素性 32	31 の品詞
素性 33	31 の分類語彙表の番号 7,5,4,3,2,1 桁
素性 34	対象語が含まれる文節に係る文節の最初の付属語
素性 35	34 の品詞
素性 36	34 の分類語彙表の番号 7,5,4,3,2,1 桁
素性 37	対象語が含まれる文節に係る文節の最後の付属語
素性 38	37 の品詞
素性 39	37 の分類語彙表の番号 7,5,4,3,2,1 桁
素性 40	対象語が含まれる文節に係る文節の自立語
素性 41	40 の品詞
素性 42	40 の分類語彙表の番号 7,5,4,3,2,1 桁
素性 43	対象語が含まれる文節に係る文節の付属語
素性 44	43 の品詞
素性 45	43 の分類語彙表の番号 7,5,4,3,2,1 桁
素性 46	対象語が含まれる文節に係る文節の最初の自立語
素性 47	46 の品詞
素性 48	46 の分類語彙表の番号 7,5,4,3,2,1 桁
素性 49	対象語が含まれる文節に係る文節の最後の自立語
素性 50	49 の品詞
素性 51	49 の分類語彙表の番号 7,5,4,3,2,1 桁
素性 52	対象語が含まれる文節に係る文節の最初の付属語
素性 53	52 の品詞
素性 54	52 の分類語彙表の番号 7,5,4,3,2,1 桁
素性 55	対象語の類義語対が含まれる文節に係る文節の最後の付属語
素性 56	55 の品詞
素性 57	55 の分類語彙表の番号 7,5,4,3,2,1 桁

## 第4章 実験

本章では，本研究で実験を行った対義語の組を 4.1 節で説明し，本研究が行った実験方法を 4.2 節で説明し，実験結果を 4.4 節に示し，被験者実験の結果と使い分けの実験の比較を 4.5 節に示す。

### 4.1 実験データ

本研究では，荻原らの研究 [1] でまとめられた対義語データベースから被験者実験で 20 人中 17 人以上が対義語であると判断されたものを利用する。名詞，動詞，形容詞，副詞の対義語対の中から新聞記事の中に両方の対義語が 50 回以上出現したのからランダムに 50 対を抽出し，実験を行う。また副詞と形容詞に関しては条件を満たすものが副詞が 8 対，形容詞が 46 対であったので 8 対と 46 対で実験を行う。表 4.1 から表 4.4 に実験を行った対義語の組を示す。1991 年から 1995 年，2011 年から 2015 年の毎日新聞から対義語の組のいずれかの語を含む文を獲得した。データ数は学習に偏りを少なくするため対義語の組内で同数にして実験を行った。

被験者実験は，機械学習により求めた 2 つの対義語の正解率の低い方の性能をもとに対義語対を 3 種類（高 90 % 以上，中 80 % 以上 90 % 未満，低 80 % 未満）に分類し，その 3 種類から，4 つの品詞（名詞，動詞，形容詞，副詞）ごとに 3 組ずつランダムに抽出し，合計 36 対（ $3 \times 4 \times 3$ ）で実験を行う。副詞の中に分類されたものは 2 種類，低に分類されたものは 1 種類しかなかったので，33 対で実験を行った。表 4.5 に被験者実験を行った対義語の組を示す。



表 4.1: 実験を行った名詞の対義語の組

1	「増産」 「減産」
2	「充電」 「放電」
3	「勝訴」 「敗訴」
4	「白地」 「黒地」
5	「増税」 「減税」
6	「楽勝」 「辛勝」
7	「先輩」 「後輩」
8	「賛成」 「反対」
9	「一方」 「両方」
10	「当番」 「非番」
11	「悲劇」 「喜劇」
12	「和風」 「洋風」
13	「開場」 「閉場」
14	「着衣」 「脱衣」
15	「立春」 「立秋」
16	「洋裁」 「和裁」
17	「プロ」 「アマ」
18	「定数」 「変数」
19	「積」 「商」
20	「結成」 「解散」
21	「公設」 「私設」
22	「水平」 「垂直」
23	「格上げ」 「格下げ」
24	「ノンフィクション」 「フィクション」
25	「古参」 「新参」
26	「師匠」 「門弟」
27	「正室」 「側室」
28	「入社」 「退社」
29	「モノクロ」 「カラー」
30	「安眠」 「不眠」
31	「男声」 「女声」
32	「肉食」 「菜食」
33	「多弁」 「寡黙」
34	「得点」 「失点」
35	「出撃」 「迎撃」
36	「夏場」 「冬場」
37	「朗報」 「悲報」
38	「有線」 「ワイヤレス」
39	「女手」 「男手」
40	「拘禁」 「釈放」
41	「発信」 「着信」
42	「顕在」 「潜在」
43	「文語」 「口語」
44	「プラス」 「マイナス」
45	「前払い」 「後払い」
46	「是認」 「否認」
47	「実像」 「虚像」
48	「内示」 「公示」
49	「増量」 「減量」
50	「肥沃」 「不毛」

表 4.2: 実験を行った動詞の対義語の組

1	「誇る」 「恥じる」
2	「増える」 「減る」
3	「勝つ」 「負ける」
4	「会う」 「別れる」
5	「禁じる」 「許す」
6	「だます」 「引っ掛かる」
7	「出会う」 「別れる」
8	「上げる」 「落とす」
9	「迎える」 「送る」
10	「伸ばす」 「曲げる」
11	「突き当たる」 「通り抜ける」
12	「冷やす」 「沸かす」
13	「眠る」 「目覚める」
14	「溶ける」 「固まる」
15	「ささやく」 「叫ぶ」
16	「見える」 「隠れる」
17	「憎む」 「慈しむ」
18	「増す」 「減る」
19	「やる」 「もらう」
20	「温まる」 「冷える」
21	「温める」 「冷やす」
22	「褒める」 「しかる」
23	「引き上げる」 「引き下げる」
24	「進む」 「遅れる」
25	「くつつく」 「離れる」
26	「締める」 「緩める」
27	「攻める」 「防ぐ」
28	「取り付ける」 「取り外す」
29	「貸す」 「借りる」
30	「開く」 「閉まる」
31	「忘れる」 「思い出す」
32	「つなぐ」 「断ち切る」
33	「出迎える」 「見送る」
34	「近寄る」 「遠のく」
35	「詰まる」 「通り抜ける」
36	「付き合う」 「別れる」
37	「くつつける」 「放す」
38	「足りる」 「欠ける」
39	「乗る」 「降りる」
40	「押す」 「引く」
41	「尋ねる」 「答える」
42	「問う」 「答える」
43	「開ける」 「閉める」
44	「預ける」 「返す」
45	「上げる」 「下げる」
46	「見上げる」 「見下ろす」
47	「付ける」 「離す」
48	「浮く」 「沈む」
49	「上げる」 「下ろす」
50	「早める」 「遅らせる」

表 4.3: 実験を行った形容詞の対義語の組

1	「長い」 「短い」
2	「うまい」 「まずい」
3	「難しい」 「やすい」
4	「大きい」 「小さい」
5	「太い」 「細い」
6	「堅い」 「軟らかい」
7	「難しい」 「易しい」
8	「固い」 「柔らかい」
9	「明るい」 「暗い」
10	「甘い」 「厳しい」
11	「蒸し暑い」 「涼しい」
12	「多い」 「少ない」
13	「目新しい」 「古めかしい」
14	「おいしい」 「まずい」
15	「素早い」 「のろい」
16	「ふさわしい」 「そぐわない」
17	「白い」 「黒い」
18	「広い」 「狭い」
19	「濃い」 「薄い」
20	「暑い」 「寒い」
21	「硬い」 「軟らかい」
22	「美しい」 「醜い」
23	「手ぬるい」 「手厳しい」
24	「熱い」 「冷たい」
25	「鋭い」 「鈍い」
26	「男らしい」 「女らしい」
27	「心強い」 「心もとない」
28	「暖かい」 「冷たい」
29	「心強い」 「心細い」
30	「疎い」 「詳しい」
31	「高い」 「安い」
32	「遠い」 「近い」
33	「憎い」 「かわいい」
34	「固い」 「軟らかい」
35	「重い」 「軽い」
36	「深い」 「浅い」
37	「速い」 「遅い」
38	「硬い」 「柔らかい」
39	「憎い」 「いとしい」
40	「よい」 「悪い」
41	「新しい」 「古い」
42	「強い」 「弱い」
43	「高い」 「低い」
44	「おもしろい」 「つまらない」
45	「難しい」 「たやすい」
46	「厚い」 「薄い」

表 4.4: 実験を行った副詞の対義語の組

1	「はっきり」「うっすら」
2	「どっさり」「ちょっぴり」
3	「はっきり」「ぼんやり」
4	「はきはき」「ぐずぐず」
5	「うっかり」「わざと」
6	「まだ」「もう」
7	「たくさん」「少し」
8	「あっさり」「こってり」

表 4.5: 被験者実験を行った対義語の組

1	「肥沃」「不毛」
2	「潜在」「顕在」
3	「是認」「否認」
4	「格上げ」「格下げ」
5	「入社」「退社」
6	「喜劇」「悲劇」
7	「カラー」「モノクロ」
8	「フィクション」「ノンフィクション」
9	「正室」「側室」
10	「引かかる」「だます」
11	「誇る」「恥じる」
12	「迎える」「送る」
13	「尋ねる」「答える」
14	「くっつく」「離れる」
15	「眠る」「目覚める」
16	「叫ぶ」「ささやく」
17	「勝つ」「負ける」
18	「沸かす」「冷やす」
19	「難しい」「やすい」
20	「鋭い」「鈍い」
21	「熱い」「冷たい」
22	「深い」「浅い」
23	「そぐわない」「ふさわしい」
24	「速い」「遅い」
25	「憎い」「いとしい」
26	「少ない」「多い」
27	「うっすら」「はっきり」
28	「あっさり」「こってり」
29	「少し」「たくさん」
30	「はっきり」「ぼんやり」
31	「わざと」「うっかり」
32	「ちょっぴり」「どっさり」
33	「はきはき」「ぐずぐず」
34	「まだ」「もう」

## 4.2 実験方法

獲得した対義語組について、対義語の組ごとに対義語の使い分けの実験を行う。入力文は、1991年から1995年、2011年から2015年の毎日新聞から獲得した、対義語の組のいずれかの語を含む文である。評価は10分割のクロスバリデーションで行う。再現率ごとに、高中低の3種類に分類する。高が90%以上、中が80%以上90%未満、低が80%未満である。再現率が高のものは置き換えがしにくいもの、低のものは置き換えがしやすいもの、中のものはそれらの中間であるとする。

被験者実験についての説明をする。まず、対義語対の使い分けの実験を行う。機械学習で使用した文から対義語対ごとに10文ランダムに抽出し被験者に対義語の使い分けを行わせる。その際、被験者は対義語対のどちらかを選ぶのではなく対義語対のどちらかもしくは判断不可の3択の中から選ばせる。

次に、対義語対の置き換え可否の判断の実験を行う。被験者に使い分けの際、選ばなかった単語（対義語）を文章にあてはめさせる。その対義語に沿うように文を知識的な内容だけ逆に変更した時（直前直後の付属語、係り受け関係にある文中の格助詞は変更しない）、対義語に置き換えられる場合、置き換え可（Y）を、置き換えできない場合は置き換え不可（N）を選択させる。また、使い分けの際、判断不可を選んだ場合は使い方の違いがないということなので置き換え可（Y）を選ばせる。また、使い分けの選択が間違っていた場合は間違えるほど使い方の違いがないということなので置き換えできると判断し、集計する際置き換え不可を選択していても置き換え可として集計する。実際の例文を表4.6に示す。左の（）内が使い分けの実験の際に使う対義語対で、右の（Y-N）が対義語対の置き換え可否の実験に使うものである。

表 4.6: 被験者実験の例

「少し-たくさん」	
(少し-たくさん-判断不可)(Y-N)	でも早く1桁に入ることが目標」と悔し涙をぬぐった
世界の(少し-たくさん-判断不可)(Y-N)	の苦しみを自分は理解している
チョコやあめが(少し-たくさん-判断不可)(Y-N)	入っているパックを買って、小分けにしたもの
具が(少し-たくさん-判断不可)(Y-N)	詰まり、かけるだけでごちそうになるキューピーの「具のソース」
職員が(少し-たくさん-判断不可)(Y-N)	流された
さまざまな改革は選手やコーチに戸惑いを与えながらも(少し-たくさん-判断不可)(Y-N)	ずつ浸透中
嫌がらない事を確認しながら(少し-たくさん-判断不可)(Y-N)	ずつ時間を延ばしていく
私は駅から(少し-たくさん-判断不可)(Y-N)	離れ、料金もやや高いが、間違いなく置けるところを探し当てた
うどんをゆでた後、水洗いし、丼に入れ、醤油を(少し-たくさん-判断不可)(Y-N)	入れて食べます
競争相手は増えたが、(少し-たくさん-判断不可)(Y-N)	の人が研究を進め、この分野が発展するのはうれしい

## 4.3 検定方法

本研究で使用した検定方法について、符号検定は4.3.1節に、無相関検定は4.3.2節に、相関係数の差の検定は4.3.3節に記述する。

### 4.3.1 符号検定

本研究では有意差検定に、符号検定を使用する。符号検定とは、二つの相反するペアに差があるかないかを+と-の符号を付け、その符号の出現率が二項確率に従うかどうかを検定する手法である。以下、符号検定の流れを示す。

#### 1. 仮説の設定

帰無仮説 ( $H_0$ ): 「2群間に差がない」と仮定する。

対立仮説 ( $H_1$ ): 「2群間に差がある」と仮定する。

2群間に差がなければ、+と-の符号がでる確率は両方とも  $\frac{1}{2}$  である。もし2群間に差があるならば、+と-の符号がでる確率はどちらかに偏る。

#### 2. 有意水準の設定

有意水準とは、仮説を棄却するときの基準となる確率のことである。有意水準は5%設定する。

#### 3. p 値を求める

$H_0$  が正しいという前提において、極端に偏った検定量が得られる p 値を求める。

#### 4. $H_0$ を棄却する

$H_0$  が正しいという前提で求めた p 値が有意水準を下回ることで  $H_0$  の仮定を棄却し、 $H_1$  の仮説が立証することで、2群間に差があることを証明する。この時初めて「統計的に有意差があった」といえる。

### 4.3.2 無相関検定

本研究では相関係数の有意差検定に、無相関検定を使用する。無相関検定とは、ある標本の相関係数を求めた際に、その相関係数に意味があるかどうかを検定する手法である。以下、無相関検定の流れを示す。ある標本から得られた相関係数  $r$  とは別に母集団の母相関係数  $\rho$  が存在する。

- 母集団の相関係数: 母相関係数  $\rho$
- 標本の相関係数: (標本) 相関係数  $r$

#### 1. 仮説の設定

帰無仮説 ( $H_0$ ): 「 $\rho = 0$ : 母相関係数は 0 である (相関係数が存在しない)」と仮定する。  
対立仮説 ( $H_1$ ): 「 $\rho \neq 0$ : 母相関係数は 0 ではない (相関係数が存在する)」と仮定する。

#### 2. 有意水準の設定

有意水準とは、仮説を棄却するときの基準となる確率のことである。有意水準は 5% 設定する。

#### 3. p 値を求める

以下の数式で、相関係数  $r$  とデータ数  $n$  を使用して、検定統計量  $t$  を求める。

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \quad (4.1)$$

有意水準と比較する確率 p 値を計算する。p 値は t 分布において、 $|t|$  以上の値が発生する確率である。

#### 4. $H_0$ を棄却する

$H_0$  が正しいという前提で求めた p 値が有意水準を下回ることで  $H_0$  の仮定を棄却し、 $H_1$  の仮説が立証することで、2つの相関係数の間に有意差があることを証明する。この時初めて「統計的に有意差があった」といえる。

### 4.3.3 相関係数の差の検定

本研究では2つの相関係数の有意差検定に、相関係数の差の検定を使用する。相関係数の差の検定とは、2つの相関係数を求めた際に、その相関係数に有意差があるかどうかを検定する手法である。以下、相関係数の差の検定の流れを示す。

#### 1. 仮説の設定

帰無仮説 ( $H_0$ ) : 「2つの相関係数に差がない」と仮定する。

対立仮説 ( $H_1$ ) : 「2つの相関係数に差がある」と仮定する。

#### 2. 有意水準の設定

有意水準とは、仮説を棄却するときの基準となる確率のことである。有意水準は5%設定する。2つの相関係数をそれぞれ  $r_1, r_2$  とする。

#### 3. p 値を求める

以下の数式で相関係数  $r_1, r_2$  をそれぞれ,  $z_1, z_2$  (z 変換) に変換する。

$$z_1 = \frac{1}{2} \ln \left( \frac{1+r_1}{1-r_1} \right) \quad (4.2)$$

$$z_2 = \frac{1}{2} \ln \left( \frac{1+r_2}{1-r_2} \right) \quad (4.3)$$

次に, z 変換値  $z_1, z_2$  とデータ数  $n_1, n_2$  から p 値を求める。

$$= \frac{z_1 - z_2}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}} \quad (4.4)$$

$H_0$  が正しいという前提において、極端に偏った検定量が得られる p 値を求める。

#### 4. $H_0$ を棄却する

$H_0$  が正しいという前提で求めた p 値が有意水準を下回ることで  $H_0$  の仮定を棄却し,  $H_1$  の仮説が立証することで、有意差があることを証明する。この時初めて「統計的に有意差があった」といえる。



## 4.4 使い分けの実験結果

機械学習の再現率の高さごとに対義語の組を分類したものを表 4.7 に示す。それぞれの対義語ごとの結果を名詞は表 4.8 と表 4.9，動詞は表 4.10 と表 4.11，形容詞は表 4.12 と表 4.13，副詞は表 4.14 に示す。

表 4.7: 対義語の分類結果

	高	中	低
名詞	15	16	19
動詞	14	22	14
形容詞	4	23	19
副詞	2	3	1

表 4.8: 名詞の対義語対の結果 1

対義語	再現率	データ数
増産	83.32%	977
減産	84.64%	977
充電	88.00%	159
放電	83.04%	159
勝訴	67.70%	1000
敗訴	71.20%	1000
黒地	66.98%	156
白地	63.39%	156
減税	92.20%	1000
増税	90.60%	1000
楽勝	86.77%	325
辛勝	89.54%	325
後輩	70.60%	1000
先輩	67.50%	1000
賛成	78.30%	1000
反対	67.50%	1000
一方	98.30%	1000
両方	98.70%	1000
当番	83.82%	149
非番	90.55%	149
悲劇	90.40%	1000
喜劇	91.70%	1000
洋風	69.59%	365
和風	74.24%	365
閉場	86.87%	121
開場	86.93%	121
着衣	94.97%	177
脱衣	95.44%	177
立秋	68.02%	68
立春	55.82%	68
和裁	65.00%	62
洋裁	59.10%	62
アマ	88.90%	1000
プロ	89.90%	1000
変数	96.00%	51
定数	94.18%	51
商	93.50%	461
積	93.91%	461
解散	92.50%	1000
結成	96.10%	1000
公設	82.13%	375
私設	84.53%	375
垂直	85.90%	1000
水平	85.60%	1000
格下げ	90.59%	839
格上げ	91.29%	839
ノンフィクション	85.88%	545
フィクション	88.44%	545
新参	81.43%	136
古参	88.33%	136

表 4.9: 名詞の対義語対の結果 2

対義語	再現率	データ数
門弟	88.33%	60
師匠	81.66%	60
側室	59.84%	113
正室	65.67%	113
入社	91.80%	1000
退社	95.10%	1000
モノクロ	86.83%	608
カラー	82.58%	608
安眠	78.89%	90
不眠	90.00%	90
男声	45.15%	84
女声	60.73%	84
菜食	83.33%	91
肉食	81.40%	91
寡黙	70.22%	84
多弁	59.63%	84
得点	89.50%	1000
失点	90.60%	1000
迎撃	93.25%	518
出撃	95.18%	518
夏場	74.26%	952
冬場	76.47%	952
悲報	86.98%	239
朗報	86.23%	239
ワイヤレス	92.17%	117
有線	96.63%	117
女手	77.91%	68
男手	69.89%	68
拘禁	87.42%	238
釈放	89.06%	238
着信	93.10%	232
発信	91.82%	232
顕在	98.54%	611
潜在	98.03%	611
文語	68.91%	109
口語	75.97%	109
プラス	78.90%	1000
マイナス	79.10%	1000
前払い	78.46%	154
後払い	76.65%	154
否認	93.94%	364
是認	96.16%	364
虚像	72.31%	144
実像	74.95%	144
公示	96.32%	461
内示	95.87%	461
減量	82.30%	136
増量	77.22%	136
不毛	88.84%	64
肥沃	95.55%	64

表 4.10: 動詞の対義語対の結果 1

対義語	再現率	データ数
恥じる	95.24%	143
誇る	94.28%	143
増える	69.30%	1000
減る	64.40%	1000
負ける	86.00%	1000
勝つ	85.30%	1000
別れる	81.35%	296
会う	76.73%	296
許す	94.80%	1000
禁じる	95.20%	1000
だます	84.10%	64
引っ掛かる	90.77%	64
出会う	81.77%	296
別れる	85.46%	296
落とす	88.60%	1000
上げる	88.50%	1000
送る	95.50%	1000
迎える	95.50%	1000
伸ばす	79.60%	186
曲げる	84.92%	186
通り抜ける	87.17%	78
突き当たる	88.50%	78
冷やす	90.99%	112
沸かす	92.02%	112
目覚める	85.41%	247
眠る	81.36%	247
固まる	77.27%	176
溶ける	88.70%	176
ささやく	75.33%	154
叫ぶ	79.88%	154
隠れる	88.44%	294
見える	85.03%	294
憎む	77.60%	93
慈しむ	85.85%	93
減る	92.70%	1000
増す	93.30%	1000
やる	89.20%	1000
もらう	92.80%	1000
冷える	75.85%	137
温まる	81.80%	137
温める	80.00%	370
冷やす	84.05%	370
褒める	82.98%	141
しかる	78.09%	141
引き下げる	77.00%	1000
引き上げる	79.40%	1000
遅れる	92.00%	1000
進む	90.30%	1000
離れる	88.42%	112
くつつく	84.86%	112

表 4.11: 動詞の対義語対の結果 2

対義語	再現率	データ数
締める	88.24%	262
緩める	90.39%	262
攻める	97.80%	1000
防ぐ	97.60%	1000
取り外す	80.95%	105
取り付ける	78.10%	105
貸す	82.48%	622
借りる	84.41%	622
閉まる	92.80%	125
開く	88.80%	125
思い出す	89.70%	1000
忘れる	90.40%	1000
つなぐ	91.68%	444
断ち切る	93.47%	444
出迎える	93.42%	168
見送る	88.73%	168
遠のく	80.33%	68
近寄る	84.40%	68
詰まる	82.35%	85
通り抜ける	85.88%	85
別れる	79.74%	296
付き合う	76.98%	296
くっつける	80.00%	54
放す	69.82%	54
欠ける	95.95%	297
足りる	95.31%	297
乗る	91.71%	820
降りる	93.41%	820
押す	92.60%	1000
引く	92.90%	1000
尋ねる	94.00%	1000
答える	92.00%	1000
問う	94.80%	1000
答える	96.50%	1000
閉める	80.62%	283
開ける	78.39%	283
預ける	88.85%	565
返す	86.73%	565
下げる	84.60%	1000
上げる	83.00%	1000
見下ろす	85.59%	465
見上げる	82.58%	465
離す	86.79%	188
付ける	87.65%	188
浮く	88.55%	454
沈む	88.32%	454
下ろす	93.06%	620
上げる	92.58%	620
遅らせる	67.93%	474
早める	71.71%	474

表 4.12: 形容詞の対義語対の結果 1

対義語	再現率	データ数
長い	78.50%	1000
短い	76.90%	1000
まずい	83.10%	1000
うまい	83.30%	1000
難しい	96.20%	1000
やすい	96.00%	1000
大きい	81.60%	1000
小さい	80.90%	1000
細い	79.80%	1000
太い	77.40%	1000
堅い	92.35%	405
軟らかい	90.37%	405
易しい	70.53%	95
難しい	75.79%	95
柔らかい	91.22%	787
固い	89.08%	787
明るい	77.90%	1000
暗い	77.80%	1000
厳しい	89.30%	1000
甘い	87.80%	1000
涼しい	80.59%	170
蒸し暑い	83.53%	170
多い	78.90%	1000
少ない	80.70%	1000
古めかしい	72.79%	83
目新しい	78.16%	83
まずい	85.60%	1000
おいしい	90.30%	1000
のろい	91.36%	57
素早い	83.79%	57
ふさわしい	85.12%	383
そぐわない	91.39%	383
黒い	65.20%	1000
白い	68.30%	1000
狭い	73.70%	1000
広い	73.10%	1000
濃い	85.50%	1000
薄い	85.40%	1000
暑い	76.20%	1000
寒い	76.60%	1000
軟らかい	74.57%	405
硬い	70.86%	405
美しい	82.56%	195
醜い	80.51%	195
手ぬるい	89.09%	55
手厳しい	85.45%	55
冷たい	87.80%	1000
熱い	86.80%	1000
鈍い	93.82%	746
鋭い	94.65%	746

表 4.13: 形容詞の対義語対の結果 2

対義語	再現率	データ数
女らしい	68.59%	61
男らしい	78.72%	61
心強い	89.91%	229
心もとない	86.50%	229
暖かい	85.88%	914
冷たい	84.14%	914
心強い	84.43%	114
心細い	75.30%	114
詳しい	86.25%	80
疎い	86.25%	80
高い	89.80%	1000
安い	92.90%	1000
近い	91.20%	1000
遠い	91.80%	1000
かわいい	84.30%	146
憎い	77.45%	146
軟らかい	88.64%	405
固い	85.93%	405
軽い	79.90%	1000
重い	78.60%	1000
浅い	90.70%	1000
深い	89.80%	1000
速い	80.40%	1000
遅い	83.80%	1000
柔らかい	80.81%	787
硬い	80.94%	787
いとしい	77.37%	97
憎い	78.52%	97
悪い	85.70%	1000
よい	84.90%	1000
新しい	85.50%	1000
古い	85.70%	1000
強い	83.10%	1000
弱い	86.70%	1000
高い	72.20%	1000
低い	69.50%	1000
おもしろい	74.44%	626
つまらない	76.67%	626
たやすい	84.81%	104
難しい	74.76%	104
薄い	89.20%	1000
厚い	87.80%	1000

表 4.14: 副詞の対義語対の結果

対義語	再現率	データ数
うっすら	95.70%	397
はっきり	92.96%	397
ちょっぴり	84.10%	101
どっさり	82.12%	101
ぼんやり	86.93%	567
はっきり	87.15%	567
ぐずぐず	80.00%	50
はきはき	97.50%	50
うっかり	80.21%	475
わざと	86.95%	475
もう	83.50%	1000
まだ	86.40%	1000
たくさん	95.50%	1000
少し	94.60%	1000
こってり	78.46%	65
あっさり	80.00%	65

使い分けの実験の結果，表 4.7 より，名詞は格が変化しないので，文法的な使われ方の差がないと考えられ低が多くなることが予想された．予想どおり低が多く高が少ないという結果になった．動詞は格が変わることがあるので高が多くなると予想されたが，高がもっとも少なく中と低に多く分布した．形容詞は文法的な使われ方の差がないと考えられ低が多くなることが予想された．予想通り，低が少なく続いて中が多いという結果となった．副詞は実験対象の対義語対も少なく，あまり傾向が把握できない．

BERT と最大エントロピー法の再現率の平均を表 4.15 に示す．表 4.16 に対義語対ごとに，BERT が最大エントロピー法の再現率を上回った個数を示す．

表 4.15: BERT と最大エントロピー法の再現率の平均

	名詞	動詞	形容詞	副詞
BERT	83.08%	86.24%	82.60%	86.79%
最大エントロピー法	80.82%	82.33%	78.69%	85.63%



表 4.16: BERT が最大エントロピー法の再現率を上回った個数

	名詞	動詞	形容詞	副詞
BERT	40	45	42	5
合計	50	50	46	8

表 4.16 の結果を用いて, 二項分布に基づく片側検定の符号検定 (有意水準 5%) をを行い, 有意差を求める. その結果を表 4.17 に示す. p 値が 0.05 より小さければ有意差があると判断される. 表 4.17 より, 名詞, 動詞, 形容詞は BERT に有意差があると判断さ

表 4.17: 有意差検定

	名詞	動詞	形容詞	副詞
p 値	0.00001193	0.00000000	0.00000000	0.36328125

れる. 副詞は対義語対の数が 8 個と少ないので, 有意差が得られなかったと考えられる.

## 4.5 被験者実験

3人の被験者実験の結果を表4.18と表4.19と表4.20に示す。また、BERTと最大エントロピー法の使い分けの性能を比較するため、佐々本[3]の被験者実験をやり直した。その結果を表4.21と表4.22と表4.23に示す。表4.18, 表4.20, 表4.21, 表4.23中の数値は置き換え可であると判断された回数である。また、表4.18の実験した副詞の対義語対の数は高が2個, 低が1個なので, それぞれ数値を1.5倍と3倍している。同じく, 表4.21の実験した形容詞の高と副詞の低が2個なので, 数値を1.5倍している。表4.19は表4.18の結果を, 表4.22は表4.21の結果を3(分類ごとに実験した対義語対の数) × 10(1組の対義語対の問題数) × 3(被験者の人数)の90で割り, 割合にしたものである。

表 4.18: BERT の被験者実験の結果 1

	名詞	動詞	形容詞	副詞	合計
高	36	31	35	25.5	127.5
中	60	37	52	27	176
低	57	29	48	33	167

表 4.19: BERT の被験者実験の結果 2

	名詞	動詞	形容詞	副詞	合計
高	0.40	0.34	0.39	0.28	0.35
中	0.67	0.41	0.58	0.30	0.49
低	0.63	0.32	0.53	0.37	0.46

表 4.20: BERT の被験者実験の結果 3

品詞	対義語対	分類分け	置き換え可
名詞	「迎撃」「出撃」	高	12
	「変数」「定数」		11
	「解散」「結成」		13
動詞	「尋ねる」「答える」		10
	「遅れる」「進む」		8
	「許す」「禁じる」		13
形容詞	「近い」「遠い」		12
	「難しい」「やすい」		13
	「鋭い」「鈍い」		10
副詞	「少し」「たくさん」		10
	「うっすら」「はっきり」		7
名詞	「当番」「非番」		中
	「アマ」「プロ」	16	
	「公設」「私設」	25	
動詞	「温める」「冷やす」	13	
	「閉まる」「開く」	9	
	「貸す」「借りる」	15	
形容詞	「暖かい」「冷たい」	18	
	「浅い」「不快」	11	
	「速い」「遅い」	23	
副詞	「ぐずぐず」「はきはき」	3	
	「わざと」「うっかり」	7	
	「ぼんやり」「はっきり」	15	
名詞	「立秋」「立春」	低	15
	「男声」「女声」		27
	「勝訴」「敗訴」		15
動詞	「増える」「減る」		23
	「伸ばす」「曲げる」		4
	「固まる」「溶ける」		2
形容詞	「黒い」「白い」		23
	「女らしい」「男らしい」		12
	「憎い」「いとしい」		13
副詞	「あっさり」「こっそり」		11

表 4.18 と表 4.19 より，被験者実験と機械学習の性能を比較した結果，全体として

は高と低は概ね機械学習の性能と対義語対の置き換えの可否に逆の相関がある。また、名詞と副詞は高から低に近づくほど置き換え可の数値が上がり、機械学習の性能と対義語対の置き換え可否に逆の相関がある。しかし、動詞と形容詞に関しては逆の相関があると言えるが、中の値が高より高くなっていた。

表 4.21: 最大エントロピー法の被験者実験の結果 1

	名詞	動詞	形容詞	副詞	合計
高	31	11	25.5	24	91.5
中	42	31	28	32	133
低	70	33	49	24	176

表 4.22: 最大エントロピー法の被験者実験の結果 2

	名詞	動詞	形容詞	副詞	合計
高	0.34	0.12	0.28	0.27	0.25
中	0.47	0.34	0.31	0.36	0.37
低	0.78	0.37	0.54	0.27	0.49

表 4.23: 最大エントロピー法の被験者実験の結果3

品詞	対義語対	分類分け	置き換え可
名詞	「着衣」「脱衣」	高	10
	「迎撃」「出撃」		9
	「是認」「否認」		12
動詞	「問う」「答える」	高	1
	「送る」「迎える」		2
	「恥じる」「誇る」		8
形容詞	「鋭い」「鈍い」	高	8
	「難しい」「やすい」		9
	「少し」「たくさん」		11
副詞	「こっそり」「あっさり」	高	6
	「うっすら」「はっきり」		7
名詞	「入社」「退社」	中	8
	「新参」「古参」		10
	「賛成」「反対」		20
動詞	「締める」「緩める」	中	7
	「預ける」「返す」		16
	「閉まる」「開く」		8
形容詞	「悪い」「よい」	中	5
	「柔らかい」「固い」		9
	「濃い」「薄い」		14
副詞	「ちよっぴり」「どっさり」	中	8
	「わざと」「うっかり」		7
	「ぼんやり」「はっきり」		17
名詞	「後輩」「先輩」	低	24
	「文語」「口語」		22
	「増産」「減産」		23
動詞	「遅らせる」「早める」	低	15
	「増える」「減る」		14
	「固まる」「溶ける」		5
形容詞	「柔らかい」「硬い」	低	17
	「長い」「短い」		11
	「細い」「太い」		21
副詞	「ぐずぐず」「はきはき」	低	6
	「もう」「まだ」		10

表 4.21 と表 4.22 より, 被験者実験と機械学習の性能を比較した結果, 全体としては高と低は概ね機械学習の性能と対義語対の置き換えの可否に逆の相関がある. また, 名詞と副詞は高から低に近づくほど置き換え可の数値が上がり, 機械学習の性能と対義語対の置き換え可否に逆の相関がある. しかし, 副詞に関しては逆の相関があると言えるが, 中の値が高より高くなっていた.

## 第5章 考察

本章では、考察を記述する。5.1節では、被験者実験のBERTの性能の傾向が一致した対義語対の考察について記述する。5.2節では、被験者実験とBERTの性能の傾向が一致しなかった対義語対の考察について記述する。5.3節では、BERTの被験者と最大エントロピー法の被験者実験について考察する。

### 5.1 被験者実験と機械学習の性能の傾向が一致した対義語対の考察

機械学習の性能が低いと対義語対は置き換え可能であると考えられ、機械学習の性能が高いと対義語対は置き換え不可能であると考えられる。本節では、素性分析を行い、被験者実験と機械学習の性能の傾向が一致した対義語対の考察を行う。

#### 5.1.1 分類分けが高で置き換え可の値が低かったもの

ここでは分類分けが高で置き換え可の値が低かったものとして、「迎撃」「出撃」と「遅れる」「進む」と「鈍い」「鋭い」と「うっすら」「はっきり」に関して考察する。

まず、「迎撃」「出撃」に関して考察する。「迎撃」「出撃」は置き換え可と判断された割合は0.27(8/30)である。表5.1に機械学習の性能を示す。表5.2に機械学習が重要と判断した素性を示す。

表 5.1: 機械学習の性能 (「迎撃」「出撃」)

	データ数	再現率
迎撃	518	93.25%
出撃	518	95.18%

表 5.2: 機械学習が参考にした性能 (「迎撃」「出撃」)

迎撃		出撃	
素性	数値	素性	数値
素性 1:核	0.9991	素性 1:終結	0.9991
素性 1:防衛	0.9990	素性 1:太平洋	0.9991
素性 1:ミサイル	0.9988	素性 1:沖縄	0.9989

有用な素性としては、「迎撃」は、「核」「防衛」「ミサイル」など、戦争に関するものが多く見られた。以下に例を示す。

- たぶん、米国が構想している宇宙での戦略核ミサイル迎撃システムに反対し、国際規模の地上からのミサイル迎撃ミサイルを意味しているのではないか。
- 韓国の現在のミサイル防衛体制では迎撃体制では迎撃体制では名詞能力が不十分だとされ、在韓米軍が配備を検討している。

「出撃」は「太平洋」「沖縄」など、戦争の中でも第二次世界大戦に関するものが多かった。これは、新聞から文を抽出していることが原因の一つだと考える。以下に例を示す。

- 45年4月、米軍が上陸した沖縄を救うべく特攻出撃したが、わずか2時間余りの戦闘で米軍機に撃沈された。
- 太平洋戦争末期の1945年4月、米軍が上陸した沖縄に向けて出撃し、撃沈された「大和」を中心とする艦隊の戦没者を伴う。

置き換え可の値が低い理由としては、「出撃」の文中内での対義語が「迎撃」ではなく「撤退」と置き換えることもできたため、置き換え可の値が低くなったと考えられる。さらに、「迎撃」「出撃」は慣用的表現が少なかったからだと考えられる。

次に、「遅れる」「進む」に関して考察する。「遅れる」「進む」は置き換え可と判断された割合は0.27(8/30)である。表 5.3 に機械学習の性能を示す。表 5.4 に機械学習が参考にした素性を示す。

表 5.3: 機械学習の性能 (「遅れる」「進む」)

	データ数	再現率
遅れる	1000	92.00%
進む	1000	90.30%



表 5.4: 機械学習が参考にした性能 (「遅れる」「進む」)

遅れる		進む	
素性	数値	素性	数値
素性 1:半島	0.9968	素性 1:向き	0.9986
素性 1:復興	0.8993	素性 1:の	0.8402

有用な素性としては、「遅れる」は、「半島」「復興」など政治に関係するものが見られた。以下に例を示す。

- 紀伊半島の被災地は大半が山間部の過疎地で、河川の急激な増水や岩盤ごと土砂が崩れ、避難情報を伝える防災行政無線が水没して使えず、非難が遅れるなどした。
- 復興用地取得の迅速化のため復興特区法が改正されたが、精度設計に問題があり、遅れることもある。

「進む」は「動く」といった意味で使われ、「遅れる」が対義語でない場合の使われ方が見られた。以下に例を示す

- 風下に進むのは簡単だが、風上に向かうところがみそだ。
- ベストを尽くして準決勝までは進む。

置き換え可の値が低い理由としては、「進む」は「動く」や「向かう」という意味で使われ、「遅れる」の対義語として使われていなかったため置き換え可の値が低くなったと考えられる。

次に「鋭い」「鈍い」に関して考察する。「鋭い」「鈍い」は置き換え可と判断された割合は0.40(13/30)である。表 5.5 に機械学習の性能を示す。表 5.6 に機械学習が参考にした素性を示す。

表 5.5: 機械学習の性能 (「鋭い」「鈍い」)

	データ数	再現率
鈍い	746	93.82%
鋭い	746	94.65%

表 5.6: 機械学習が参考にした性能 (「鈍い」「鋭い」)

鈍い		鋭い	
素性	数値	素性	数値
素性 1:動き	0.9946	素性 1:眼光	0.9955
素性 1:人	0.9768	素性 1:知性	0.9868
素性 1:輝き	0.9288	素性 1:追及	0.9826

有用な素性としては、「鈍い」は「人」「輝き」など慣用的な表現が見られた。以下に例を示す。

- 一見彼より鈍い人が実は悪がしこく危険ということを見逃し、裏切られたのがクーデター事件だ。
- 鈍い輝きを放つ10円玉が淡い青春を彩っていたのだと思うと、どす黒くなった10円玉でさえきれいに見えてくるから不思議だ。
- 東京電力福島第1原発の汚染水漏れ事故をめぐる国会の動きが鈍い。

「鋭い」は「眼光」「知性」「追及」など慣用的な表現が見られた。以下に例を示す。

- 普通なら動揺して崩れがちな場面でも、鋭い眼光で相手を見据え、時には大声で自分を鼓舞する、たくましさが見立った。
- 「ヨーロッパでもっとも鋭い知性」とサルトルが評した思想家。
- 鋭い追求とテレビ出演の多さから参院の「顔」の一人となっている山本一太議員。

置き換え可の値が低くなった理由としては、「鈍い」は「人」「輝き」などの単語とともに使われ、「鋭い」は「知性」「追及」などの単語とともに使われた。「鋭い」「鈍い」ともに慣用的な表現がよく使われているため置き換え可の値が低くなったと考えられる。

次に「うっすら」「はっきり」に関して考察する。「うっすら」「はっきり」は置き換え可と判断された割合は0.23(7/30)である。表 5.7 に機械学習の性能を示す。表 5.8 に機械学習が参考にした素性を示す。

表 5.7: 機械学習の性能 (「うっすら」「はっきり」)

	データ数	再現率
うっすら	397	95.70%
はっきり	397	92.26%

表 5.8: 機械学習が参考にした性能 (「うっすら」「はっきり」)

うっすら		はっきり	
素性	数値	素性	数値
素性 1:染	0.9992	素性 1:結果	0.9981
素性 1:汗	0.9991	素性 1:依然	0.9984
素性 1:涙	0.9990	素性 1:態度	0.9979
素性 1:雪	0.9970	素性 1:追及	0.9826

有用な素性としては、「うっすら」は「染」「汗」「涙」「雪」などの単語と同時に出現しやすいことが見られた。以下に例文を示す。

- 低くたれ込めた雲の切れ目がうっすらと、赤く染まっていた。
- 背中にうっすらと汗が光る。

また、「うっすらと染まる」「うっすらと見える」という表現が多く見られた。「はっきり」は「結果」「依然」「態度」などの単語と同時に出現しやすいことが見られた。以下に例文を示す。

- この結果について、同連合会は「これまで言われていた不満が数字ではっきり出た」と判断。
- まず「国土強韌化」が何かを意味するかが依然としてはっきりしない点だ。
- はっきりさせる (和田会長) と態度を一層硬化させた。

また、「はっきりする」「はっきりとした」という表現が多く見られた。置き換え可の値が低くなった理由としては、「うっすら」は「光」や「雪」などのような単語とともに使われ、「うっすらと見える」「うっすらと染まる」という慣用的な表現が多く使われていた。「はっきり」は「はっきりする」という慣用的な表現がとてもよく使われており、ともに慣用的な表現がよく使われているため置き換え可の値が低くなったと考えられる。

分類分けが高で置き換え可の値が低かったものとして「迎撃」「出撃」と「進む」「遅れる」と「鋭い」「鈍い」と「うっすら」「はっきり」を考察した。その結果、分類分けが低で置き換え可の値が高かったものの特徴として、佐々本ら [3] と同様に以下のようなものが見られた。

- 慣用的な表現が多い, もしくは特定の慣用的な表現がよく使われる
- 多義性がある
- 有用な素性が多く得られる

### 5.1.2 分類分けが低で置き換え可の値が高かったもの

ここでは分類分けが低で置き換え可の値が高かったものとして、「男声」「女声」と「増える」「減る」と「黒い」「白い」に関して考察する。

まず、「男声」「女声」に関して考察する。「男声」「女声」は置き換え可と判断された割合は0.87(26/30)である。表 5.9 に機械学習の性能を示す。表 5.10 に機械学習が参考にした素性を示す。

表 5.9: 機械学習の性能 (「男声」「女声」)

	データ数	再現率
男声	84	45.15%
女声	84	60.73%

表 5.10: 機械学習が参考にした性能 (「男声」「女声」)

男声		女声	
素性	数値	素性	数値
素性 1: 兵士	0.7403	素性 1: 主婦	0.6381
素性 1: 男子	0.7324	素性 1: 女性	0.6322

「男声」は「男子校」「兵士」など、性別に関する素性が見られた。また、「女声」と同時に使われることが多いことも見られた。また、「女声」の場合も「男声」と共同様に一緒に出現しやすい。以下に例文を示す。

- 「第九」は捕虜の兵士が演奏したから、もちろん男声だけであった。
- 男声合唱を強化したいと男子校の川越高の音楽部にも参加を打診し快諾を得た。

「女声」も「男声」と同様に「女性」「主婦」など性別に関する素性が見られた。以下に例文を示す。

- 神戸市垂水区などの主婦らで作る「明舞女声コーラス」が、同市西区でチャリティコンサートを開く。
- 女性グループでの習字のお稽古、女声コーラス、講演会、バザーの実施、ボランティア活動、昼食会を定期的にこなした。

置き換え可の値が高くなった理由は、「男声」「女声」はともに慣用的な表現があまりなく、格の変化なども起こらず、多義性がないためだと考えられる。機械学習の性能が低い原因は、「男声」「女声」はともに音楽に関係する言葉なので使い方の違いがほとんど無く、有用な素性があまり得られなかったことが原因と考えられる。

次に、「増える」「減る」に関して考察する。「増える」「減る」は置き換え可と判断された割合は0.70(21/30)である。表 5.11 に機械学習の性能を示す。表 5.12 に機械学習が参考にした素性を示す。

表 5.11: 機械学習の性能(「増える」「減る」)

	データ数	再現率
増える	1000	69.30%
減る	1000	69.40%

表 5.12: 機械学習が参考にした性能(「増える」「減る」)

増える		減る	
素性	数値	素性	数値
素性 1:出張	0.9858	素性 1:党员	0.9118
素性 1:来日	0.9725	素性 1:面接	0.9050
素性 1:リハビリ	0.9677	素性 1:田舎	0.9017
素性 1:排ガス	0.9534	素性 1:年金	0.8893

素性としては、「増える」は「出張」「来日」「リハビリ」「排ガス」などが見られた。しかし、素性に共通点などは見られなかった。以下に例文を示す。

- 現在モスクには家族を含めて800を超える日本人駐在員がおり、長期出張などで訪れる日本人は増える一方。
- 先日、来日したカーター米国防衛長官は、日本の安保法制を反映した新たなガイドラインによって「米軍と自衛隊が切れ目なく協力する機会が増える。

「減る」は「党员」「面接」「田舎」「年金」などが見られたが、こちらも共通点などは見られなかった。以下に例文を示す。

- 同党幹部によると、この大統領令が施行されると、党员が二割以上減るとみており、大きな痛手をこうむることが間違いない。
- ハーグ条約加盟で、国際間の面接交流も減ることが予想される。

置き換え可の値が高くなった理由は、「増える」「減る」ともに慣用的な表現があまりなく、格の変化なども起こらず、多義性がないことだと考えられる。機械学習の性能が低い原因は、「増える」「減る」ともによく使われる言葉なので、「増える」「減る」が出現する文章内容の種類が多いことが原因だと考えられる。

次に「黒い」「白い」に関して考察する。「黒い」「白い」は置き換え可と判断された割合は0.67(20/30)である。表 5.13 に機械学習の性能を示す。表 5.14 に機械学習が参考にした素性を示す。

表 5.13: 機械学習の性能(「黒い」「白い」)

	データ数	再現率
黒い	1000	65.20%
白い	1000	68.30%

表 5.14: 機械学習が参考にした性能(「黒い」「白い」)

黒い		白い	
素性	数値	素性	数値
素性 1:石炭	0.6760	素性 1:氷原	0.9700
素性 1:噂	0.6552	素性 1:LED	0.6314
素性 1:スーツ	0.6002	素性 1:月	0.6079

素性としては、「黒い」は「石炭」「噂」「スーツ」など「黒い」色に関するものの素性が得られた。以下に例文を示す。

- 石炭は黒いダイヤと呼ばれ、産炭地の経済も潤っていた。
- 牛肉・オレンジの輸入枠に黒い噂が付きまとったのと同じ構造である。

「白い」は「歯」「氷原」「LED」「月」などが「白い」色に関するものの素性が得られた。以下に例文を示す。

- ストレート勝ちを決めると、白い歯を見せて左手を握りしめ「しぶとくいいショットを打とうというイメージだった。
- 右手のオホーツク海には沿岸から数キロのところまで白い氷原が迫る。

置き換え可の値が高くなった理由は、「黒い」「白い」ともに慣用的な表現があまりなく、格の変化なども起こらず、多義性がないことだと考えられる。機械学習の性能が低い原因は、「黒い」「白い」ともによく使われる言葉なので、「黒い」「白い」出現する文章内容の種類が多いことが原因だと考えられる。

分類分けが低で置き換え可の値が高かったものとして「男声」「女声」と「増える」「減る」と「黒い」「白い」を考察した。その結果、分類分けが低で置き換え可の値が高かったものの特徴として、佐々本ら [3] と同様に以下のようなものが見られた。

- 慣用的な表現が少ない、もしくはあまり使われない
- 多義性が少ない
- 出現する文章内容の種類が多い

被験者実験と機械学習の性能の傾向が一致した対義語対の考察をした結果、次のようなことが佐々本ら [3] と同様に確認できた。

一つ目は対義語対に慣用的な表現が少ない、もしくはあまり使われない場合、置き換え可の値が高くなることである。また、その際、慣用的な表現は機械学習が重要と判断した素性になるので、慣用的な表現が少ない、もしくはあまり使われない場合は機械学習の性能が低くなる。二つ目は対義語対のどちらかの単語に多義性が少ない場合、置き換え可の値が高くなることである。また、その際、多義性により文中のある単語 X の対義語 Y が文中に出現した時、Y はその文中では X の対義語ではなかった場合、X が出現する文や Y が X の対義語として出現する時の文と大幅に違うので機械学習でも簡単に判断できる。よってそのような文が少なくなると機械学習の性能が低くなる。

これらのことは逆のことも言える。一つ目は、対義語対に慣用的な表現が多い、もしくはよく使われている場合置き換え可の値が低くなり、有用な素性が多く得られる

ので機械学習の性能が高くなる。二つ目は対義語対のどちらかの単語に多義性が多い場合、多義性により簡単に判断できる場合が多くなり、機械学習の性能が高くなる。

また、出現する文章内容の種類が多いと学習しなければならない情報が多くなり、機械学習の性能が落ちる。このことは置き換え可否とは関係がないと考えられる。

## 5.2 被験者実験と機械学習の性能の傾向が一致しなかった対義語対の考察

機械学習の性能が低いと対義語対は置き換え可能であると考えられ、機械学習の性能が高いと対義語対は置き換え不可能であると考えられる。本節では、素性分析を行い、被験者実験と機械学習の性能の傾向が一致しなかった対義語対の考察を行う。

### 5.2.1 分類分けが低だが置き換え可の値が低かったもの

ここでは分類分けが低だが置き換え可の値が低かったものとして、置き換え可の値が低かった「固まる」「溶ける」と「憎い」「いとしい」に関して考察する。

まず、「憎い」「いとしい」に関して考察する。「憎い」「いとしい」は置き換え可と判断された割合は0.40(12/30)である。表5.15に機械学習の性能を示す。表5.16に機械学習が参考にした素性を示す。

表 5.15: 機械学習の性能 (「憎い」「いとしい」)

	データ数	再現率
憎い	97	78.52%
いとしい	97	77.37%

表 5.16: 機械学習が参考にした性能 (「憎い」「いとしい」)

憎い		いとしい	
素性	数値	素性	数値
素性 1:犯人	0.9936	素性 1:妻	0.9739
素性 2:が	0.9714	素性 1:娘	0.9559

データ数が少なかったのであまり有用な素性が得られなかった。有用な素性としては、「憎い」は直前に「が」が来て「…が憎い」という表現や「犯人」などが見られた。



- 修行で解脱できるといって巧みに若者たちを引きずり込んだオウム真理教が憎い。
- 犯人が憎いと語り、泣き崩れた。

「いとしい」に関しては、直後に名詞が来て「いとしい…」という表現や、「娘」などが見られた。以下に例文を示す。

- やや後ずさりしたいとしい妻が、後ろ手でそっと窓をあけたりする。
- 勝ち気でやる気いっばいの娘に育ててくれた母ちゃんがいとしい

などがあつた。「…が憎い」や「いとしい…」という表現は対義語に置き換えても問題無い。よって、この2つの表現はそれぞれの対義語対の使われやすい表現ではあるが慣用的な表現ではない。

置き換え可の値が低くなってしまった原因としては、多義性により、「憎い」の文中内での対義語が「いとしい」ではなく「好ましい」と判断されたものがあつたため置き換え可の値が低くなつたと考えられる。使い方に関しては、よく使われる慣用的な表現もなく、文法的な使い方の違いもなかつた。

次に「固まる」「溶ける」に関して考察する。「固まる」「溶ける」は置き換え可と判断された割合は0.10(3/30)である。表5.17に機械学習の性能を示す。表5.18に機械学習が参考にした素性を示す。

表 5.17: 機械学習の性能(「固まる」「溶ける」)

	データ数	再現率
固まる	176	77.27%
溶ける	176	88.70%

表 5.18: 機械学習が参考にした性能(「固まる」「溶ける」)

固まる		溶ける	
素性	数値	素性	数値
素性1:意見	0.8270	素性1:輝き	0.9799
素性1:考え	0.7582	素性1:疲れ	0.6145

「固まる」「意見」や「考え」が「固まる」という表現や、「集まる」の意味の「固まる」といった表現のような、「溶ける」の対義語ではない使われ方がされていた。以下に例文を示す。

- ただ、声明発表後には宗派ごとに別々に固まる様子も放映され、宗派間のわだかまるが垣間見えた。
- にもかかわらず、なぜか両車の入口付近に人が固まる。
- 日本郵政株は、09年に設立した売却凍結法で、日本郵政グループの組織再編成案が固まるまで売ることを禁じられている。

「溶ける」に関しては、「役に溶ける」「疲れが溶ける」といった慣用的な表現が得られた。

- 娘役に必要なのは、男役に溶ける輝き。
- 乳白色のお湯の中に、溶ける、溶ける、疲れがどんどん溶けてゆく。

置き換え可の値が低くなってしまった原因としては、「固まる」は「集まる」や「決まる」といった「溶ける」の対義語ではない使われ方が多くされていたためと考えられる。

分類分けが高で置き換え可の値が高かったものとして「沸かす」「冷やす」と「憎い」「いとしい」と「まだ」「もう」を考察した。その結果、分類分けが低だが置き換え可の値が低かったものの特徴として、佐々本ら[?]と同様に以下のようなものが見られた。

- 慣用的な表現が多い、もしくはよく使われる慣用的な表現がある
- どちらかの単語に多義性がある
- 出現する文章内容の種類が多い

### 5.2.2 分類分けが中だが置き換え可の値が高かったもの

ここでは、分類分けが中だが置き換え可の値が高かったものとして、最も置き換え可の値が高かった、「公設」「私設」に関して考察する。「公設」「私設」は置き換え可と判断された割合は0.67(20/30)である。表5.19に機械学習の性能を示す。表5.20に機械学習が参考にした素性を示す。

表 5.19: 機械学習の性能 (「公設」「私設」)

	データ数	再現率
公設	375	82.13%
私設	375	84.53%

表 5.20: 機械学習が参考にした性能 (「公設」「私設」)

公設		私設	
素性	数値	素性	数値
素性 1:社会党	0.9968	素性 1:応援団	0.9507
素性 1:財政	0.9962	素性 1:ヘリポート	0.8364

「公設」「私設」共に、「秘書」に関するものが多く見られた。素性としては「公設」は「社会党」「財政」など政治に関するものも多く見られた。以下に例文を示す。

- 世論の風当たりや窮屈な党財政を考えて取りやめも検討したが、各議員が公設秘書などに支払うボーナスの原資としてモチ代をあてにしていた台所事情を考慮した。
- 当時、社会党の横路孝弘の公設第一秘書だった。

「私設」は「応援団」「ヘリポート」など「秘書」に関係のない、場所についてのものが見られた。以下に例を示す。

- 投票を二週間後に控え、私設応援団のボルテージが上がってきた。
- 経営者は四年前にライセンスを取得、私設ヘリポートは許可を取っている。

置き換え可の値が高くなってしまった原因としては、慣用的な表現があまりなく、さらに文法的な使い方の違いがあまりないので、置き換え可の値が高くなったためと考えられる。文法的な使い方の違いがあまりないのに機械学習の性能がよいのは、使われている文章が「私設」「公設」ともに特徴的であり、そのことを多く学習したことが原因だと考えられる。特徴的な文章を多く学習したのは、「私設」「公設」が元々使われる場面が限定される単語なことと、新聞からデータを取得しているため収集する文章に偏りが出たことが原因だと考えられる。

### 5.2.3 分類分けが中だが置き換え可の値が低かったもの

ここでは、分類分けが中だが置き換え可の値が高かったものとして、最も置き換え可の値が低かった、「閉める」「開ける」に関して考察する。「閉める」「開ける」は置き換え可と判断された割合は0.27(8/30)である。「閉める」「開ける」について考察する。表 5.21 に機械学習の性能を示す。表 5.22 に機械学習が参考にした素性を示す。

表 5.21: 機械学習の性能 (「閉まる」「開ける」)

	データ数	再現率
閉まる	125	92.80%
開ける	125	88.80%

表 5.22: 機械学習が参考にした性能 (「閉める」「開ける」)

閉める		開ける	
素性	数値	素性	数値
素性 1:鍵	0.9645	素性 1:道	0.9363
素性 1:市場	0.9487	素性 1:幕	0.8928
素性 1:店	0.8230	素性 1:ショー	0.7464

素性としては「閉める」は、「鍵」「店」「市場」「引き戸」など終了に関するものが見られた。以下に例を示す。

- 家の鍵を閉めるとき2人で話した。
- 個人投資家から「店を閉めるのか」など問い合わせが相次いだ。

「開ける」は「幕を開ける」「道が開ける」といった慣用的な表現や、「開催する」という意味になる「閉まる」の対義語ではない使われ方がされていた。以下に例を示す。

- 井山6冠が挑戦状を獲得すると、7冠への道が開ける。
- 第65回の節目の大会は、いよいよ、26未知から甲子園球場で幕を開ける。
- 「三年前までソウルでこんなショーが開けるなんて、考えもしませんでした。

などがあった。

中に分類されたにも関わらず置き換え可の値が低くなってしまった原因としては、両方ともある程度有用な素性が得られて機械学習の性能がある程度高くなった。だが、「開ける」は多義性があり、慣用的な表現が多くなったために、置き換え可の値が低くなってしまったと考えられる。

分類分けが中だが置き換え可の値が低かったものとして、置き換え可の値が低かった、「閉まる」「閉じる」の考察を行った。置き換え可能が低くなった理由として、佐々本ら [3] と同様に以下のようなものが見られた。

- 慣用的な表現が多い、もしくは、よく使われる
- 多義性がある

被験者実験と機械学習の性能の傾向が一致しなかった対義語対の考察をした結果、それぞれの対義語対で傾向が一致しなかった原因はわかったが、全体として傾向などはつかめなかった。高中低の分類ではなく機械学習の性能ごとに置き換え可否の値と比較する考察も重要であると考えられる。

### 5.3 BERT と最大エントロピー法の被験者実験の考察

BERT と最大エントロピー法の再現率と、被験者実験の置き換え可の値との相関係数を求めた。その結果を表 5.23 に示す。

表 5.23: BERT と最大エントロピー法の相関係数

	名詞	動詞	形容詞	副詞	合計
最大エントロピー法	-0.8130	-0.6489	-0.5785	-0.0720	-0.5503
BERT	-0.5751	-0.3738	-0.5355	-0.1904	-0.5092

最大エントロピー法と BERT の合計には、逆の相関があるように見える。しかし、本当に再現率と被験者実験の置き換え可否との間に、相関があるかわからない。そこで、表 5.23 の結果を用いて、無相関検定 (有意水準 5%) を行う。まず、最大エントロピー法の合計の結果で無相関検定を行い、その結果を表 5.24 に示す。p 値が 0.05 より小さければ有意差があると判断される。

表 5.24 より、p 値が 0.05 より小さいので、有意差があると判断される。よって、最大エントロピー法の再現率と被験者実験の置き換え可の値との間に、有意な逆の相関が

表 5.24: 最大エントロピー法の無相関検定の結果

相関係数	-0.55032
データ数	34
t 値	3.728486
p 値	0.000746

あることが分かる。次に、BERT の合計の結果で無相関検定を行い、その結果を表 5.25 に示す。

表 5.25: BERT の無相関検定の結果

相関係数	-0.50915
データ数	33
t 値	3.293689
p 値	0.002478

表 5.25 より、p 値が 0.05 より小さいので、有意差があると判断される。よって、最大エントロピー法の再現率と被験者実験の置き換え可の値との間に、有意な逆の相関があることが分かる。さらに、最大エントロピー法と BERT との相関係数の間に有意差があるかを、相関係数の差の検定 (有意水準 5%) で調べる。その結果を表 5.26 に示す。

表 5.26: 相関係数の差の検定の結果

	最大エントロピー法	BERT
相関係数	-0.55032	-0.50915
データ数	34	33
z 変換値	-0.61884	-0.56158
p 値		0.223572

表 5.26 より、p 値が 0.05 より大きいので、最大エントロピー法と BERT の相関係数の間に有意差がないと判断される。したがって、BERT は最大エントロピー法と同様に、逆の相関があることが確認できた。

## 第6章 おわりに

本研究ではBERTを用いて対義語の使い分けを行い、その結果を用い対義語対の置き換え可否の判定を行った。本研究の成果は2つある。

第1の成果として、機械学習の性能が高ければ置き換え不可能であり、機械学習の性能が低ければ置き換え可能であるという、機械学習の性能と置き換え可否に逆の相関があることが、佐々本ら [3] の最大エントロピー法と同様にBERTでも確認できた。

第2の成果として、BERTの性能と被験者実験の置き換え可否の傾向が一致したものや一致しなかったものについて機械学習の素性を分析した。その結果、それぞれの対義語における有用な知見を獲得した。例えば「進む」は「向かう」の意味で使われ、「遅れる」の対義語として使われておらず、多義性があることが分かった。また、機械学習の性能と置き換え可否の被験者実験を比較した結果置き換え可の値が高くなる条件と置き換え可の値が低くなる条件が、佐々本ら [3] と同様に確認できた。置き換え可の値が高くなる条件を以下に示す。

- 慣用的な表現が多い、もしくは特定の慣用的な表現がよく使われる
- 多義性がある

また、置き換え可の値が低くなる条件を以下に示す。

- 慣用的な表現が少ない、もしくは特定の慣用的な表現がよくあまり使われない
- 多義性少ない
- 一般的によく使われていて、出現する文章内容の種類が多い

# 謝辞

また，研究の進め方や本論文の書き方など，細部にわたる御指導を頂きました，鳥取大学工学部知能情報工学科自然言語処理研究室の村田真樹教授に心から御礼申し上げます。また，本研究を進めるにあたり，御指導，御助言を頂きました，村上仁一准教授に心から御礼申し上げます。その他様々な場面で御助言を頂いた自然言語処理研究室の皆様に感謝の意を表します。



## 参考文献

- [1] 荻原亜彩美, 森山菜々美, 浅原正幸, 加藤祥, 山崎誠. 『分類語彙表』に対する反対語情報. 言語処理学会第 25 回年次大会, 2018.
- [2] 赤江涼太. 機械学習を用いた同義語の使い分けに関する知識獲得. 卒業論文, 鳥取大学工学部知能情報工学科, 2017.
- [3] 佐々本暖久. 機械学習を用いた対義語の置き換え可否判定 卒業論文, 鳥取大学工学部知能情報工学科, 2019
- [4] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.
- [5] 小学館辞典編集部. 使い方の分かる類語例解辞典. 小学館, 1994.
- [6] 山口仲美, 佐藤有紀. 「擬音語・擬態語」使い分け帳. 山海堂, 2006.
- [7] Eric Sven Ristad. Maximum entropy modeling for natural language. In *ACL/EACL Tutorial Program, Madrid*, 1997.
- [8] 村田真樹, 内山将夫, 内元清貴, 馬青, 井佐原均. 種々の機械学習手法を用いた多義解消実験. 電子情報通信学会言語理解とコミュニケーション研究会, pp. 7–14, 2001.
- [9] Masao Utiyama. Maximum entropy modeling packagen: <http://www.nict.go.jp/x/x161/members/mutiyama/software.htmlmaxent>. 2006.
- [10] Masaki Murata, Kiyotaka Uchimoto, Masao Utiyama, Qing Ma, Ryo Nishimura, Yasuhiko Watanabe, Kouichi Doi, and Kentaro Torisawa. Using the maximum entropy method for natural language processing: Category estimation, feature extraction, and error correction. Vol. 2, pp. 272–279, 2010.