

# 日本語における否定の焦点アノテーション

松吉 俊 (山梨大学大学院医学工学総合研究部)<sup>†</sup>

大槻 諒 (山梨大学工学部)

福本 文代 (山梨大学大学院医学工学総合研究部)

## Annotation of Focus of Negation in Japanese Text

Suguru Matsuyoshi (Interdisciplinary Graduate School of Medicine  
and Engineering, University of Yamanashi)

Ryo Otsuki (Faculty of Engineering, University of Yamanashi)

Fumiyo Fukumoto (Interdisciplinary Graduate School of Medicine  
and Engineering, University of Yamanashi)

### 1. はじめに

「誰がいつどこで何を」するという、主に述語項構造で表現される**事象**の末尾に、「ない」や「ん」、「ず」などの語が付くと、いわゆる否定文となる。否定文では、一般に、その事象が成立しないことが表現される。否定文において、否定の働きが及ぶ範囲を**スコープ**、その中で特に否定される部分を**焦点** (フォーカス) と呼ぶ (日本語記述文法研究会 2007)。本論文では、日本語における否定の焦点をアノテーションする枠組みを提案し、現在構築中である、否定の焦点コーパスについて報告する。

『現代日本語書き言葉均衡コーパス』(BCCWJ)<sup>\*1</sup>から抽出した、否定のスコープと焦点の例を以下に示す。ここでは、注目している否定を表す表現を太字にしており、そのスコープを角括弧で囲み、焦点の語句に下線を付している<sup>\*2</sup>。

(1) だが、[学校での] 子どもの様子はわから **ない**から、それだけでうれしい。 [PN1a\_00002]

(2) [十七日まで] 選手にも協会関係者にも明かさ **ない**。 [PN2f\_00002]

(3) [力を出し切って] 敗れた **わけではない**。 [PN2f\_00003]

(4) [WHOは] 五月十八日、ジュネーブで開いた総会で台湾の総会へのオブザーバー参加問題を議題としないことを決め、[オブザーバー参加を認め] **なかった**。 [PN4g\_00001]

文(1)は、家庭訪問を受けた母親の発言の一部である。「ない」のスコープは、「学校での子どもの様子はわから」で表現される事象である。家庭での子どもの様子は分かると考えられるので、焦点は「学校での」とするのが妥当であると思われる。

文(2)は、最終登録選手に関する監督の発言の一部である。「ない」のスコープは、「十七日まで選手にも協会関係者にも明かさ」で表現される事象である。十七日かそれ以降に登録選手を明かすことが期待できるので、焦点は「十七日まで」と考える。

<sup>†</sup> sugurum at yamanashi. ac. jp

<sup>\*1</sup> [http://www.ninjal.ac.jp/corpus\\_center/bccwj/](http://www.ninjal.ac.jp/corpus_center/bccwj/)

<sup>\*2</sup> 例文の後の“PN”から始まる文字列は、その例文を抽出したBCCWJ内のファイル名を表す。

文(3)は、試合に敗れた選手に関する報道記事の一部である。否定の複合辞「わけではない」のスコープは「力を出し切って敗れた」であり、否定の焦点は「力を出し切って」であると解釈できる。

文(4)は、WHO 総会に関する報道記事の一部である。「なかつ」のスコープは、「WHOはオブザーバー参加を認め」で表現される事象である。この例文においては、(前後の文脈を考慮しても、)スコープの中に特に否定される部分はないように思われる。本研究では、このような場合に、「なかつ」の焦点は、無しとせず、便宜上、スコープ全体であると考え、紙面が煩雑になるのを避けるため、焦点がスコープ全体である場合には、例文に下線を付けない。

英語や日本語を対象として、否定とその焦点について言及した言語学的研究に、加藤ほか(2010)や日本語記述文法研究会(2007)がある。

これまでに、否定のスコープや焦点を対象としたアノテーションコーパスがいくつか構築されている。BioScope (Vincze et al. 2008) は、生医学分野における英語文章を対象として、否定のスコープをアノテーションしたコーパスであり、自然言語処理の分野において、これを利用して否定のスコープを自動解析する研究 (Morante et al. 2008, Li et al. 2010) が進められている。川添ほか (2011) は、日本語の新聞を対象として否定のスコープのアノテーションを進めている。否定のスコープを対象としたこれらの研究に比べ、否定の焦点を対象とした研究は少ない。Blanco and Moldovan (2011a) は、PropBank (Babko-Malaya 2005) における述語-項関係の上に否定の焦点をアノテーションし、コーパスを構築した。これを利用して否定の焦点を自動検出する研究 (Blanco and Moldovan 2011b, Rosenberg and Bergler 2012) も行われている。日本語では、松吉ほか (2010) が、拡張モダリティの1項目として、コーパスにおいて否定の焦点を扱っているが、この研究で実際にアノテーションした事例は非常に少ない。

本論文は、以下のように構成される。まず、2章において、否定の焦点アノテーションの基本指針について述べる。続く3章で、与えられた日本語文章に否定の焦点をアノテーションする枠組みを説明する。4章で、現在構築中である、2つのジャンルのコーパスについて報告する。5章はまとめである。

## 2. 否定の焦点アノテーションの基本指針

文章に存在する否定を検出し、その焦点にラベルを付け、コーパスを構築する。言語学的利用のみでなく、自然言語処理への応用も考慮して、アノテーションの基本指針を定める。

### 2.1 焦点の部分を除いた事象が成立すること

否定の焦点がスコープ全体でない場合、スコープの事象が成立しないだけでなく、焦点の部分を除いた事象は成立することが推測できる (日本語記述文法研究会 2007, Blanco and Moldovan 2011a)。例えば、1章の文(3)において、「力を出し切って敗れた」ことは否定されるが、「力を出し切って」の部分に否定の焦点があることが分かれば、「敗れた」ことは成立することが推測できる。同様に、1章の文(2)において、「十七日まで」の部分に否定の焦点があることが分かれば、監督はずっと明かさないのでなく、十七日かそれ以降に「選手にも協会関係者にも明かす」ことが成立することが推測できる。Blanco and Moldovan (2011a) は、こ

の考え方にに基づき、否定の焦点がスコープ全体でない場合の表現法を提案し、アノテーションコーパスを構築した。我々も、基本指針の1つとしてこの考え方を取り入れる。

## 2.2 否定要素

本論文では、文中において否定を表す表現のことを**否定要素**と呼ぶ。本研究では、次の3種類の語群をまとめたものを否定要素と定める。

**否定辞** 助動詞「ない」と「ず」、接尾辞「ない」、接頭辞「非」、「不」、「無」、「未」、「反」、「異」  
**非存在の内容語** 形容詞「無い」、名詞「無し」

**否定を表す複合辞** 「のではない」、「わけではない」、「わけにはいかない」など

否定辞のみでなく、非存在の内容語まで含める理由は、「無い」は、存在の内容語「ある」の丁寧な否定「ありません」と同等と思われるからである。否定辞「ん」が使用されている「ありません」は対象とし、「無い」は内容語なので対象としないのは、不合理であると思われる。

言語学の文献((森田・松木 1989)など)において、否定を表す複合辞とされる表現は、1形態素の否定辞と異なる性質を持つと思われるので、区別して扱う。

接頭辞「非」や「不」は、直後の語を否定する働きを持つのみであり、これらに対して焦点を判断する必要はないと思われがちである。しかしながら、次の例のように、「ない」や「ん」と同様に、接頭辞もスコープの一部に焦点を持つことがあるので、対象とした。

(5) 九十年代の「失われた十年」ではつきりしたのは、[もはや 民間まかせでは 過剰債務処理] **不** [可能] ということだ。[PN1b\_00004]

これは、前の文脈から、過剰債務処理には政府の介入が必要であることが読み取れる例であり、否定の焦点は「民間まかせでは」であると考えられる。

## 2.3 否定要素としない語句

否定辞か非存在の内容語を含む2形態素以上の慣用表現は、全体を1語とし、焦点判断の対象としないこととする。これらの慣用表現は、大きく分けると、次の2種類からなる。

**複合語** 「物足りない」、「仕方がない」、「思わず」など

**否定以外の意味を持つ複合辞** 「なければならない」、「かもしれません」、「だけでなく」など  
上記の複合語に相当するかどうかは、次の2点から判断する。

- 肯定形(例えば、「仕方がない」に対する「仕方がある」)が、通常、使用されるか
- 国語辞典(松村・小学館『大辞泉』編集部 1998, 西尾ほか 2000)に見出しが立っているか  
複合辞であるかどうかの判断は、言語学の文献((森田・松木 1989)など)を参考にし、前節で述べたように、否定を表す複合辞とされる表現は、否定要素として扱う。

助動詞「ない」か接尾辞「ない」、もしくは、形容詞「無い」を使った単純な否定表現に言い換えられない否定の接頭辞は、否定要素とはしない。例えば、「不十分」は、「十分でない」ことであるので、焦点判断の対象とする。一方、「不気味」は、「気味が悪い」ことであり、「気味がない」や「気味でない」に言い換えられないので、対象としない。

## 2.4 否定要素と呼応する程度・頻度の副詞

以下の例文のように、否定要素と呼応する、程度の副詞や頻度の副詞が用いられることがある。ここでは、注目している否定要素を太字にし、程度の副詞や頻度の副詞に下線を付ける。

- (6) ボールを回すくらいで、そんなにハードな練習じゃなかった。 [PN2f\_00002]
- (7) 市街地では、街灯やライトアップによる“光害”で夜空の星が なかなか 見えない。  
[PN2g\_00004]
- (8) 価格は1万円前後で、「いつもは ぜいたくできないけれど、お正月くらい、という方が多いようです」。 [PN3b\_00004]

文(6)で述べられていることは、「全くハードな練習ではなかった」ことではなく、ハードな練習ではあったが、その程度が想定されるよりも高くなかったということである。同様に、文(7)では、星は全く見えないのではなく、見える程度や頻度が低いということが述べられている。文(8)の該当箇所は、いわゆる部分否定であり、「ぜいたくできる」ことが全く成り立たないわけではなく、たまには成り立つことが読み取れる。

本研究では、便宜上、呼応するこのような副詞を否定の焦点とみなす。自然言語処理における含意認識 (Dagan et al. 2005) というタスクにおいては、程度や頻度はともかく、ある事象が成立するかどうかを計算機で自動的に判定することが求められる。例えば、上の文(7)から、「市街地で夜空の星が見える」ことが成立するかどうかを判定することが問われる。文(7)において、「なかなか」に否定の焦点があることが分かれば、2.1節の考え方をを用いて、機械的に、「市街地では、街灯やライトアップによる“光害”で夜空の星が見える」\*<sup>3</sup>ことが導出でき、「見える」と正しく答えることができる。

否定と呼応する、「そんなに」や「なかなか」のような程度・頻度の副詞は、厳密には、否定の焦点ではなく、含意認識で利用したいのならば、別の枠組みを用意すべきであるかもしれない。しかしながら、これらの副詞は、「多くは(持てない)」や「速くは(走れない)」のような形容詞連用形 + 「は」や、「頻繁には(通えない)」のような形状詞 + 「には」と同様に用いられる。このような形容詞や形状詞を否定の焦点として扱うことは自然であることから、本研究では、これらに連続するものとして、上の副詞も否定の焦点とみなす。

含意認識への応用という観点から、「全然」や「絶対に」、「決して」のような完全否定を表す副詞は、否定の焦点とはみなさないこととする。

### 3. 否定の焦点アノテーションの枠組み

この章では、まず、否定の焦点を判断する基準について述べる。そして、否定要素とその焦点に対して定めたアノテーション項目と、そこに付与するラベルについて説明する。

#### 3.1 否定の焦点の判断基準

1章で述べたように、否定要素によって特に否定される部分が否定の焦点である。これを安定して判断するために、2.1節の考え方に基づいて、我々は次のような判断基準を定めた。

1. ある文の否定の焦点を判断する時には、その文だけでなく、周りの文脈も広く参照する
2. 対象とする文から、一部の表現と否定要素を除外した事象を生成する。その事象が成立

---

\*<sup>3</sup> もちろん、この単純な書き換えは正しくない。人手で正確に書き換えると、「市街地では、街灯やライトアップによる“光害”があっても、夜空の星は、程度や頻度はともかくとして、見える」のようになるであろう。しかしながら、含意認識というタスクにおいて、上の問いに答えるためには、ここまで複雑な書き換えは必要でない。

することが推測できれば、除外した表現の部分を否定の焦点と判断する

3. 解釈に複数の可能性が考えられる場合は、否定の焦点はスコープ全体であるとする

- 例えば、一部に焦点があると考えられることもできるし、スコープ全体が焦点であると考えられることもできる場合
- 例えば、A という部分に焦点があると解釈することもできるし、B という部分に焦点があると解釈することもできる場合

基準 3. は、判断する人間の思い込みを最大限排除するために設けたものである。

### 3.2 項目とラベル

否定要素に対して、以下のアノテーション項目を定める。

**表層文字列** 文に出現した否定要素の表層文字列。出現形で記述する

**形態素 ID** 否定要素の形態素の ID

**品詞** 助動詞、接尾辞、接頭辞、形容詞、名詞、否定複合辞のいずれか (2.2 節参照)

**最終更新日** “YYYYMMDD” という形式で記述された最終更新日

プログラムを用意すれば、これらは自動付与が可能である。ただし、形態素解析辞書 UniDic\*4では、助動詞ではない「ない」は、すべて「形容詞, 非自立可能」と解析されるため、これらを半自動的に「形容詞」と「接尾辞」に分類する必要がある。

否定の焦点に対して、以下のアノテーション項目を定める。

**代表表層文字列** 焦点の表層文字列。ただし、後述する代表形態素のみを記述する

**代表形態素 ID** 焦点の代表形態素の ID

**項・節の種類** ガ格、ヲ格、デ格、副詞、ノの項、ナの項、テ節、ト節など、焦点の統語的分類。複数記述可

**とりたて詞の有無** 「しか」や、数量語に付く「も」が存在するか

**意味分類** 制限-時間、制限-場所、制限-対象、付加-連用修飾、付加-連体修飾、付加-アスペクトなど、意味解釈に基づいた、否定されている語句の分類

**判断の根拠** その箇所を焦点であると判断するに至った根拠。自由記述

**手がかり語句** 文章中に存在する、焦点判断の手がかりとなった語句。複数記述可

コーパスにおいて否定の焦点は代表 1 形態素にラベル付けする。このように決めた理由は、否定の焦点の自動検出システムを評価する際に、正解とシステムの出力の比較が容易になるからである。代表 1 形態素は、次のように定める。

- 内容語
- 複合語の場合、接尾辞を除く末尾の語
- 修飾語が存在する場合、それが係る末尾の語

表層的な格助詞や接続助詞などに基づく分類が、「項・節の種類」であり、焦点の語句が表す意味に基づく分類が、「意味分類」である。例えば、「意味分類」の“制限-場所”は、場所を表す語句に否定の焦点があり、そこではない場所をうまく選べば、対象事象が成立することを表す。「意味分類」の“付加-連用修飾”は、程度の副詞や頻度の副詞に対して付与する。

\*4 <https://www.tokuteicorpus.jp/dist/>

次の例のように、とりたて詞「しか」は、必ず否定要素と共起する。

(9) [普段は 決まったものしか 料理し] **ない**ので、おけいこ感覚で。 [PN3b\_00004]

「しか」が付く項が否定の焦点となり、文に述べたこの場合には事象は成立するが、これ以外の場合には成立しないことを表す。「しか」が存在する事例には、2.1 節の考え方を適用できないので、特別なマークを付けて、「しか」が存在することを明示する。

数量語に付く「も」が否定要素と共起すると、「その概数には届かない」という意味と、「書き手はそれを少ない・低いと捉えている」ことが表現される (日本語記述文法研究会 2009)。これは、累加の「も」にはない性質である。例を以下に示す。

(10) [出場者ランキングの 二十位にも 入って] **なかった** [2年生・高平慎士] が、晴れの舞台上で堂々と高校3傑入り。 [PN1e\_00003]

2.1 節の考え方の適用外ではないが、自然言語処理における評判分析・感情解析タスクに有用であると思われるので、特別なマークを付けて、数量語に付く「も」が存在することを明示する。

### 3.3 データ構造

否定の焦点コーパスは、図1のようなXMLによって表現する。この図は、1章の文(1)に対するアノテーション結果である。XMLにおいて、<wsb:negation>要素を用いて否定要素の情報を記述し、<wsb:focus>要素と<wsb:description>要素、<wsb:clue>要素を用いて否定の焦点の情報を記述する。ここで、“wsb”は、植物のわさびから名付けた名前空間である。

<wsb:negation>要素は、1文もしくは文の断片を表す<sentence>要素の直接の子要素として記述する。<wsb:negation>要素の属性に、前節で述べたアノテーション項目を記述する。<wsb:focus>要素は、<wsb:negation>要素の直接の子要素として記述する。否定の焦点がスコープ全体である場合は、<wsb:focus>要素に対して、1という値を記述した@wsb:scope属性のみを指定する。否定の焦点がスコープの一部である場合、<wsb:description>要素と<wsb:clue>要素を、<wsb:focus>要素の直接の子要素として記述する。アノテーション項目の多くは、<wsb:focus>要素の属性に記述するが、「判断の根拠」は<wsb:description>要素として、「手がかり語句」は<wsb:clue>要素として記述する。

### 3.4 否定のスコープ

本来ならば、否定の焦点をアノテーションする前に、否定のスコープを明示的にアノテーションすべきである。既存の述語項構造解析の技術を用いれば、ある程度は自動的に否定のスコープを認識することができるが、対象が整った文章でない場合、人間による修正作業が多く発生する。本研究では、人的コストの関係から、否定のスコープをアノテーションしない。人間が否定の焦点を判断する時には、対象となる否定要素のスコープを目で確認するに留める。

## 4. 否定の焦点コーパス

前章で説明したアノテーションの枠組みに基づき、次の2つのテキストデータを対象として、否定の焦点コーパスの構築を進めている。

```

:
</superSentence>
<sentence wsb:sID="73">
  <wsb:negation wsb:orthToken="ない" wsb:morphID="13250" wsb:POS="助動詞"
wsb:lastupdate="20130112">
  <wsb:focus wsb:orthToken="学校" wsb:morphID="13170" wsb:argTypes="ノの項;テ格"
wsb:class="制限-場所" wsb:NumOfCandidates="pl">
  <wsb:description>家での様子は分かる</wsb:description>
  <wsb:clue wsb:sID="62" wsb:orthTokens="主婦" wsb:morphIDs="12020" />
  </wsb:focus>
</wsb:negation>
<LUW B="S" SL="vf" l_lemma="だが" l_lForm="ダガ" l_wType="和" l_pos="接続詞"
l_formBase="ダガ">
  <SUW orderID="13140" lemmaID="22916" lemma="だ" lForm="ダ" wType="和" pos="助動詞"
cType="助動詞-ダ" cForm="終止形-一般" formBase="ダ" orthBase="だ" pron="ダ"
start="19490" end="19500">だ
  </SUW>
:

```

図1 コーパスにおけるXMLファイルの例 [PN1a-00002]

1. 楽天データ\*5の楽天トラベル: レビューデータ
2. BCCWJ におけるコアデータ内の新聞 (PN)

#### 4.1 楽天トラベル: レビューデータ

楽天トラベル: レビューデータのうち、先行研究において小池ほか (2012) が使用したものと同一レビュー集合を対象とした。彼らは、まず、宿泊施設に対するレビュー数の分布を調査し、90% 以上の宿泊施設はレビュー数が1から58の範囲にあることを明らかにした。そして、その結果に基づき、レビュー数が10から58の範囲の宿泊施設の全体から、無作為に40の宿泊施設を抽出し、独自の文分割規則により半自動的にそのレビュー集合を文分割した。

このコーパスには、5,178文が含まれており、形態素の品詞情報のみに基づいて抽出した否定要素の候補は、1,246個であった。以下、このコーパスを「レビュー」と表記する。

#### 4.2 BCCWJ コアデータの新聞

BCCWJ 全体の約1/100のデータがコアデータに指定されており、このデータは、その他の部分と比較して高い精度で解析が施されている。コアデータの一部に言語学的情報を付与する場合、国立国語研究所が定めたファイル優先順位に従うことが推奨される。我々は、コアデータ内の新聞340ファイルのうち、優先順位が1から54までの“A”グループを対象とした。

このコーパスには、1文もしくは文の断片を表す、XMLの<sentence>要素が2,708個含まれており、否定要素の候補は、406個であった。以下、このコーパスを「新聞」と表記する。

#### 4.3 コーパスの分析

2つのコーパス「レビュー」と「新聞」における、否定要素候補の分布を表1に示す。2つのコーパスにおいて、否定要素はそれぞれ1,023個と304個であり、いずれのコーパスでも、

\*5 <http://travel.rakuten.co.jp/>

表2 スコープ全体でない焦点の分布

	レビュー	新聞	計
副詞	141	18	159
ガ格	30	5	35
ヲ格	7	6	13
ニ格	49	11	60
デ格	17	6	23
マデ格	5	4	9
カラ格	3	2	5
ト格	3	1	4
その他の格	1	2	3
ノの項	20	7	27
連体の述語	8	8	16
接頭辞「全」	1	0	1
テ節	1	2	3
ト節	1	0	1
アスペクト	14	0	14
計	301	72	373

表1 否定要素候補の分布

	レビュー	新聞	計
助動詞	637	173	810
接尾辞	116	33	149
接頭辞	19	34	53
形容詞	211	53	264
名詞	28	6	34
否定複合辞	12	5	17
(上記小計)	(1,023)	(304)	(1,327)
複合語	94	30	124
その他複合辞	121	72	193
解析誤り	8	0	8
(上記小計)	(223)	(102)	(325)
計	1,246	406	1,652

助動詞「ない」と「ず」が全体の過半数を占めることが分かる。

2つのコーパスにおいて、否定の焦点がスコープ全体でないものは、それぞれ301個と72個であった。「レビュー」では、29%(301/1,023)の否定要素が、「新聞」では、24%(72/304)の否定要素が、スコープの一部に焦点を持つことが分かる。これらの焦点の「項・節の種類」の分布を表2に示す。図1に例示されるような、ある格と“ノの項”が同時に付与されている事例は、この表では、“ノの項”として集計した。「レビュー」には、焦点が副詞である否定要素が多いことが分かる。「新聞」のデータ数が少ないので、確定的なことは言えないが、どの格が焦点になりやすいかも、2つのコーパスで異なる傾向があるようである。

焦点である部分に付いていたとりたて詞の数を表3に示す。2つのコーパスを合わせ、35%(129/373)の焦点に何らかのとりたて詞が付いていたことが分かる。とりたて詞「は」は、焦点である箇所の手がかりとして利用できそうに見えるが、「は」は、特に主題を表す「は」として、スコープ全体が焦点である事例にも多く出現するので、注意が必要である。3.2節で述べたように、スコープの中に「しか」が付く項が存在する場合、それが否定の焦点となる。

焦点の語句が表す意味に基づく分類結果を表4に示す。「レビュー」には、焦点が副詞である否定要素が多いため、“付加-連用修飾”が多いことが見て取れる。「レビュー」は宿泊施設のレビュー集合であるので、場所を表す語句に否定の焦点がある“制限-場所”が、「新聞」に比べ、著しく多いことが分かる。

#### 4.4 アノテーション作業

独自のプログラムにより、3.3節で説明したXML形式のファイルから、人間が見やすいHTML形式に自動変換可能である。作業者は、ブラウザ上でHTMLファイルを確認しながら、テキストエディターにおいてXMLファイルを更新する。作業にかかる時間は、100個の



表4 焦点の意味分類結果

	レビュー	新聞	計
制限-動作主	13	5	18
制限-対象	27	12	39
制限-時間	10	9	19
制限-場所	40	3	43
制限-数量	10	5	15
制限-範囲	43	12	55
付加-連用修飾	125	15	140
付加-連体修飾	19	11	30
付加-アスペクト	14	0	14
計	301	72	373

表3 焦点に付いていたとりたて詞

	レビュー	新聞	計
「は」	66	13	79
「しか」	34	7	41
「も」	7	1	8
「だけ」	0	1	1
計	107	22	129

否定要素候補に対して3時間程度である。

2人の作業者が独立に「新聞」に対してアノテーション作業を行い、2人の作業結果において焦点の場所がどれほど一致するかを調査した。全304個の否定要素のうち、103個が不一致であったが、2時間ほど2人で議論することにより、これらの不一致を解消することができた。不一致の主な原因は、以下の3点であった。

- スcopeが明示されていないことによる勘違い
- 作業者のうち1名は、広く文脈を参照していなかった
- とりたて詞「だけ」が持つ限定の意味に引っ張られた

## 5. おわりに

本論文では、日本語における否定の焦点をアノテーションする枠組みを提案し、現在構築を進めている否定の焦点コーパスについて報告した。今後は、BCCWJの新聞以外のレジスタに対してもアノテーション作業を進めることを考えている。このコーパスを用いて試作した、否定の焦点自動検出システムについては、別稿(大槻ほか2013)を参照されたい。

構築したコーパスは、BCCWJおよび楽天データとの差分形式で、論文末に示すURLにて無償で一般公開する予定である。

## 謝 辞

本研究の一部は、科研費若手研究(B)「高精度モダリティ解析のための言語資源構築に関する研究」(課題番号: 23700176、代表: 松吉俊)の支援を受けている。

## 文 献

- 日本語記述文法研究会(編)(2007).『現代日本語文法3』くろしお出版。  
 日本語記述文法研究会(編)(2009).『現代日本語文法5』くろしお出版。  
 Babko-Malaya, Olga (2005). *PropBank Annotation Guidelines*, ACE (Automatic Content Extraction) Program. <http://verbs.colorado.edu/~mpalmer/projects/ace/PBguidelines.pdf>

- Blanco, Eduardo, and Dan Moldovan (2011a). “Semantic Representation of Negation Using Focus Detection.” *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, pp. 581–589.
- Blanco, Eduardo, and Dan Moldovan (2011b). “Some Issues on Detecting Negation from Text.” *Proceedings of the 24th International Florida Artificial Intelligence Research Society Conference*, pp. 228–233.
- Dagan, Ido, Oren Glickman, and Bernardo Magnini (2005). “The PASCAL Recognising Textual Entailment Challenge.” *Proceedings of the PASCAL Challenges Workshop on Recognising Textual Entailment*.
- Li, Junhui, Guodong Zhou, Hongling Wang, and Qiaoming Zhu (2010). “Learning the Scope of Negation via Shallow Semantic Parsing.” *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)*, pp. 671–679.
- Morante, Roser, Anthony Liekens, and Walter Daelemans (2008). “Learning the Scope of Negation in Biomedical Texts.” *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 715–724.
- Rosenberg, Sabine, and Sabine Bergler (2012). “UConcordia: CLaC Negation Focus Detection at \*Sem 2012.” *Proceedings of the First Joint Conference on Lexical and Computational Semantics: SemEval’12*, pp. 294–300.
- Vincze, Veronika, György Szarvas, Richárd Farkas, Gydotorgy Móra, and János Csirik (2008). “The BioScope Corpus: Biomedical Texts Annotated for Uncertainty, Negation and their Scopes.” *BMC Bioinformatics*, pp. 1–9.
- 加藤泰彦、吉村あき子、今仁生美 (編) (2010). 『否定と言語理論』 開拓社.
- 小池惇爾、松吉俊、福本文代 (2012). 「評価視点別レビュー要約のための重要文候補抽出」 言語処理学会第 18 回年次大会論文集, pp. 1188–1191.
- 松吉俊、江口萌、佐尾ちとせ、村上浩司、乾健太郎、松本裕治 (2010). 「テキスト情報分析のための判断情報アノテーション」 電子情報通信学会論文誌. D, 情報・システム, 93:6, pp. 705–713.
- 松村明、小学館『大辞泉』編集部 (編) (1998). 『大辞泉 (増補・新装)』 小学館.
- 森田良行、松木正恵 (1989). 『日本語表現文型用例中心・複合辞の意味と用法』 アルク.
- 西尾実、岩淵悦太郎、水谷静夫 (編) (2000). 『岩波国語辞典第六版』 岩波書店.
- 川添愛、齊藤学、片岡喜代子、崔榮殊、戸次大介 (2011). 『言語情報の確実性に影響する表現およびそのスコープのためのアノテーションガイドライン Ver.2.4』, Technical Report of Department of Information Science Ochanomizu University.
- 大槻諒、松吉俊、福本文代 (2013). 「否定の焦点コーパスの構築と自動検出器の試作」 言語処理学会第 19 回年次大会論文集 C6-1.

#### 関連 URL

否定の焦点コーパス: <http://cl.cs.yamanashi.ac.jp/nldata/negation/>