

## 日本語事実性解析課題の経験的分析

成田 和 弥<sup>†1</sup> 水野 淳 太<sup>†2,†1</sup> 乾 健 太 郎<sup>†1</sup>

事実性とは、文中のある事象が実際に起こったことなのか、あるいは起こる可能性を述べただけなのか、に関する情報である。このような事実性の解析は、情報抽出や含意関係認識などに有用であるが、内在する問題の分析・整理が不十分であり、いまだ十分な解析性能は実現されていない。

本稿では、高度な日本語事実性解析器の実現に向けて、事実性に影響を与える手がかりとなる表現に着目し、それら語彙知識の組み合わせによって、どの程度の実事性解析課題を解決できるのかを分析した。具体的には、事象の成立に影響を与える表現を手がかりとして、事実性の解析を行う Sauri (2008) の解析モデルをもとに、日本語事実性解析器を構築した。そして、構築した解析器をウェブ上の数千文に対して適用し、誤り分析を行った。

### An Empirical Analysis of Issues in Japanese Factuality Analysis

KAZUYA NARITA,<sup>†1</sup> JUNTA MIZUNO<sup>†2,†1</sup>  
and KENTARO INUI<sup>†1</sup>

Event factuality is information about whether events mentioned in natural language correspond to actual events that have occurred in the real world or to events that are of uncertain interpretation. Factuality analysis is useful for information extraction and recognizing textual entailment, among others, but the problems inherent to factuality analysis are not yet fully-understood, and sufficient performance has not yet been realized.

In this paper, we make efforts toward a high-performance Japanese factuality analyzer by focusing on "factuality markers" (Sauri 2008), expressions that can influence the factuality of events, and determine the level of performance that can be obtained through a compositional approach using these markers. We construct a Japanese factuality analyzer following the approach of Sauri (2008) and conduct error analysis on several thousand sentences from the Web.

### 1. はじめに

事実性とは、文中のある事象が実際に起こったことなのか、あるいは起こる可能性を述べただけなのか、に関する情報である。本稿では、事象は行為、出来事、状態の総称であると考え<sup>14)</sup>。しかし、その事象が真に起こったことかはわからないため、著者および文章中の登場人物による事象の成否の判断として事実性を解釈する。

- (1) a. 彼はさきほど部屋を出た。
- b. もう遅いから、彼は先に帰ったんだらう。
- c. 問題が発生するのを防いだ。

例えば、(1a)の事象「出る」は実際に起こったことだと著者は判断している、と解釈できる。(1b)の事象「帰る」は「だらう」という表現により、起こった可能性が高いことだと著者は判断している、と解釈できる。(1c)の事象「発生する」は「防いだ」という表現により、実際には起こっていないことだと著者は判断している、と解釈できる。事象に対する事実性の解析は、情報抽出や含意関係認識などに有用であるが、いまだ十分な解析性能は実現されていない。その原因の一つに、内在する問題の分析・整理が不十分であることがあげられる。

本研究の最終的な目的は、高精度な日本語事実性解析器の実現である。上記のように、事実性に影響を与える表現は多数存在し、その組み合わせによって事実性を決定できる場合が少なからず存在している。そこで我々は、事実性に影響を与える語彙知識に焦点を当て、その語彙知識の組み合わせによって解決できる問題、できない問題を明らかにする。本稿では、その細分類として、現在利用可能な語彙知識によってどれだけの問題がカバーできるのか、カバーできなくても語彙知識の拡充および拡張によって解決できるのか、語彙知識の曖昧性はどの程度問題となるのか、局所的な語彙知識だけで解決できない問題はあるのか、といった論点に着目する。語彙知識の組み合わせに基づく日本語事実性解析器を構築することで、これら論点に基づいた問題の分析を行った。

具体的には、Sauri (2008)<sup>6)</sup>の英語を対象とした事実性解析モデルをもとに、日本語事実性解析器を構築し、その誤り分析を行った。Sauriの解析モデルは、事実性に影響を与える表現を手がかりとし、それらの組み合わせによって事実性解析を行う。このモデルは、語彙知識の組み合わせによって事実性を解析できる構成性をもち、後述する確信度と肯否極性に

<sup>†1</sup> 東北大学 Tohoku University

<sup>†2</sup> 奈良先端科学技術大学院大学 Nara Institute of Science and Technology

事実性を分割して、個別に分析を行えるといった特長がある。そのため、問題の分析という本研究の目的に適している。

## 2. 関連研究

事実性およびその周辺情報をマークアップするためのタグ体系およびその付与基準の関連研究として、Sauríらによる FactBank<sup>7),9)</sup> や松吉ら (2010)<sup>14)</sup>、川添ら (2011)<sup>17),18)</sup> などがある。Sauríらは、事象を対象とし、事実らしさに対する modality (certainty) とその確信の方向を表す polarity の組によって事実性を定義した。そして、事象とその時制、肯否、モダリティが付与された TimeML<sup>8)</sup> に対して、modality (certainty) と polarity を態度表明者 (source) ごとに付与する枠組みを提案し、FactBank と呼ばれるコーパスを構築した。本稿では、modality (certainty) を確信度、polarity を肯否極性と呼ぶ。松吉らは、態度表明者、相対時、仮想、態度、真偽判断、価値判断の6項目からなる拡張モダリティタグ体系を設計し、それを付与したコーパスを構築した。彼らは拡張モダリティの項目の組み合わせによって事実性を表現している。川添らは、様相表現・否定表現・条件表現を「確実性」の側面から分析・分類し、それに従ったアノテーションスキーマの作成を行った。近年このような研究は、上で述べた純粋な自然言語処理の分野ばかりでなく、生物医学分野においても行われている<sup>3),4),10)</sup>。

解析手法としては、機械学習に基づく手法<sup>2)-4),11),12),16)</sup> や、人手で構築した語彙的・統語的知識に基づくパターンベースの手法<sup>3),6)</sup> があげられる。例えば江口ら (2010)<sup>12)</sup> は、拡張モダリティにおける項目間、および事象間の依存関係を考慮できる条件付確率場を利用した解析システムを提案した。松吉ら (2011)<sup>16)</sup> は、最大エントロピーモデルを用いた拡張モダリティ解析システムを試作し、その中の1つのタグである「態度」に着目して誤り分析を行った。

Saurí (2008)<sup>6)</sup> は、事象の成立に影響を与える手がかり表現を利用し、確信度と肯否極性で表される事実性を、依存構造木の根から伝搬させて解析するアルゴリズムを提案した。このモデルは、手がかり表現の情報を持つ語彙知識が与える事実性を組み合わせて、事象の事実性解析を行うという構成性をもつため、どのような解析過程でその結果が出力されたのかが明確である。また、事実性を確信度と肯否極性という2軸に分割して解析を行うため、それぞれを個別に分析することができる。

事実性に影響を与える表現として、「～ない」「～だろう」などの機能表現がある。このような機能表現を網羅的に集めた辞書として、日本語機能表現辞書「つつじ」<sup>15)</sup> が利用でき

る。これは、日本語の機能表現の表層形約 17,000 種 に対して、その ID、意味、文法的機能、音韻的变化などが網羅的に収録された辞書であり、機能表現の意味として、89 種類のラベルが定義されている。これらの機能表現は表層による曖昧性を持つため、その意味を同定する研究も行われている<sup>13),19)</sup>。

また、CoNLL-2010 Shared Task において *Learning to detect hedges and their scope in natural language text* というテーマが扱われる<sup>1)</sup> など、否定表現およびヘッジのスコープを同定する研究も近年盛んに行われている<sup>5)</sup>。

## 3. 日本語事実性解析器

日本語事実性解析における問題を分析・整理するため、Saurí(2008)<sup>6)</sup> の英語を対象とした事実性解析モデルをもとに、日本語に対する事実性解析器を構築した。本節では、3.1 で事実性の定義、3.2 で使用する語彙知識、3.3 で解析モデルについて述べる。

### 3.1 事実性の定義

Sauríは確信度を Certain (CT)・Probable (PR)・Possible (PS)・Underspecified (U) の4段階、肯否極性を positive (+)・negative (-)・underspecified (u) の3値として扱い、これらの組み合わせによって事実性を定義している。例えば (1a) の事象「出る」は、著者から見て実際に起こったことだと解釈でき、その事実性は CT+ と表される。同様に (1b) における事象「帰る」は、著者から見て起こった可能性が高いことだと解釈でき、その事実性は PR+、(1c) における事象「発生する」は、著者から見て実際には起こっていないことだと解釈でき、その事実性は CT- と表される。

英語では、PR は *probable*、PS は *possible* といった表現によって解釈しているが、日本語では表現が多彩であるため、PR と PS の境界が曖昧で、その区別は容易でないことが予想される。

(2) 彼が無事に着いたと私は信じている。

例えば (2) において、事象「着く」に対する著者の確信度は明らかに CT でも U でもないが、PR であるのか PS であるのかは解釈できない。そこで我々は、PR と PS を1つの値 PR としてまとめ、それ以外は Sauríの定義を利用した。即ち、確信度を Certain (CT)・Probable (PR)・Underspecified (U) の3段階、肯否極性を positive (+)・negative (-)・underspecified (u) の3値として扱い、これらの組み合わせによって事実性を表す。

Sauríは著者以外の態度表明者 (source) から見た事実性も考慮しているが、タスクを簡略化し分析をしやすくするため、本稿では著者の事実性判断を認識する問題に限定し、これ以

表 1 機能表現辞書の意味クラス例と事実性への影響

機能表現辞書中の意味クラス	該当する機能表現	事実性への影響
否定	ない ぬ わけでない	肯否極性: + → -, - → +
推量	だろう かもしれない らしい	確信度: CT → PR
疑問	か かな ではないか	確信度: CT → U, PR → U 肯否極性: + → u, - → u

表 2 手がかり表現辞書の見出し語例と対応する事実性

見出し語	下位事象の事象形式	上位事象の肯否環境	上位事象の真偽判断	上位事象に対する事実性
防ぐ	スル形	肯定	不成立	CT-
	スル形	否定	成立	CT+
忘れる	タ形	肯定	-	-
	タ形	否定	-	-
忘れる	スル形	肯定	不成立	CT-
	スル形	否定	成立	CT+
忘れる	タ形	肯定	成立	CT+
	タ形	否定	成立	CT+

降は態度表明者を省略する。

### 3.2 使用する語彙知識

Sauriのモデルは、確信度と肯否極性の組で表される事実性を、factuality marker と呼ばれる事実性に影響を与える表現を利用して解析する。例えば *not* は肯否極性を反転させる factuality marker, *may* は確信度を下げる factuality marker である。機能語だけでなく述語についても factuality marker が考えられ、例えば *know that* という表現は *that* 節の内容が成立していることを前提としているので、*know* は *that* 節内の事象を CT+ とする factuality marker と考えることができる。このような語彙的な factuality marker だけでなく、目的節や関係詞節等の統語的な factuality marker も考慮している。

日本語においても「～ない」は肯否極性の反転、「～だろう」は確信度の減少、というように、同等の表現が存在する。このような表現を集めた語彙知識として、日本語機能表現辞書「つつじ」(機能表現辞書)<sup>15)</sup> およびモダリティ解析手がかり表現辞書(手がかり表現辞書)<sup>12)</sup> を利用した。

機能表現辞書は、文の構成に関わる要素である機能表現を、意味、文法的機能、音韻的变化などに応じて網羅的に収録した辞書である。機能表現辞書の意味クラスおよびその表現の例、それらの事実性に対する影響を表 1 に示す。この辞書は意味クラスによって表現を分類しており、それぞれの表現に対して活用や表記の異なりなどを考慮して構成されている。この中で「否定」と分類される表現は肯否極性を反転、「推量」と分類される表現は確信度を減少、「疑問」と分類される表現は確信度を U、肯否極性を u に割当、といった役割を持つ factuality marker として利用できる。このような事実性に影響を与える手がかりとなる機能表現として、機能表現辞書の一部である 5,345 表現を利用した。ここでは「否定」「推量」「疑問」「意志」「願望」などの事実性に影響を与える意味クラスに基づいて表現を選出した。

手がかり表現辞書は、拡張モダリティに影響を与える 3,692 述語を収録しているが、これは事実性とも大きな関わりをもっている。手がかり表現辞書の項目の例を表 2 に示す(実際の辞書には態度・真偽判断・価値判断の 3 つ組が記述されているが、表では事実性と関わりの深い真偽判断に関する項目のみを示した)。この辞書は、各述語が直前の事象に与える影響を、直前の事象の時制(未来:スル形「～する」、非未来:タ形「～した/していた/している」)および肯否環境ごとに収録している。例えば(1c)の「防いだ」という述語は、肯定環境下では不成立(CT-)、否定環境下では成立(CT+) というように、事象「発生する」の肯否極性を反転させる。(3)の「忘れる」は、述語の直前の事象の時制を考慮した例である。

(3) a. 彼は発言したのを忘れている。

b. 彼は発言するのを忘れている。

「発言する」に対して、(3a)では「発言する」の事実性は CT+ となるが、(3b)では「発言する」ことが実際には起こっておらず、事実性は CT- となる。

### 3.3 解析モデル

前節であげた事実性に影響を与える手がかりとなる表現を集めた語彙知識を利用し、確信度と肯否極性を依存構造木の根から伝搬させて事実性を解析する。Sauriの英語を対象とした解析モデルでは、単語をノードとした依存構造木を走査するが、日本語では文節を意味の区切りと考えるのが一般的であるため、文節をノードとした依存構造木を走査する。

このアルゴリズムを Algorithm 1 に示す。ここでは、文全体で伝搬させる事実性を contextual factuality、各事象に割り当てる事実性を event factuality と呼ぶ。入力是一文の構文解析結果、出力は各 event factuality であり、contextual factuality は event factuality を伝搬させるための変数である。

まず、多くの事象の事実性は CT+ であるため、初期値として contextual factuality に

**Algorithm 1** 日本語事実性解析システム

```

1: chunk = root node in DEPENDENCY TREE
2: set initial value (CT+) of contextual-factuality
3: function traverseDependencyTree(chunk)
4: # Step1: 機能表現による更新
5: if chunk has functional expressions then
6:   update contextual-factuality with functional expressions dictionary
7: end if
8: # Step2: 事象をもつかの確認
9: if chunk has an event then
10:  get event-factuality from contextual-factuality
11: end if
12: # Step3: 手がかり表現による更新
13: if chunk has key expressions then
14:  update contextual-factuality with key expressions dictionary
15: end if
16: for all chunks c adjacent to chunk not visited do
17:  traverseDependencyTree(c)
18: end for
19: end

```

CT+を割り当て (line 2), 文末の文節から依存構造木の走査を開始する。各文節で行うことは, 機能表現辞書を用いた更新 (Step1: line 4-7)・事象がその文節内に存在するかの確認 (Step2: line 8-11)・手がかり表現辞書による更新 (Step3: line 12-15) の3ステップである。以下では, それぞれのステップについて述べる。

Step1 (line 4-7) では, 機能表現辞書と着目している文節とを照らし合わせ, 機能表現が見つかった場合はそれに応じて contextual factuality を更新する。例えば「～ない」のような否定表現があれば肯否極性を反転させ, 「～ではないか」のような疑問表現があれば確信度を U, 肯否極性を u に更新する。Step2 (line 8-11) では, その文節内に事実性を割り当てるべき事象が存在するかどうかを確認し, 存在するならば現在の contextual factuality をその事象の event factuality として割り当てる。Step3 (line 12-15) では, 手がかり表現辞書と着目している文節内の述語とを照らし合わせ, 手がかり表現が見つかった場合は, その文節よりも前の文節に事実性を伝搬させるために contextual factuality を更新する。この順番で処理を行うのは, 文節内の事象は機能表現に影響を受け, その事象自体がそれ自身に影響を与えることはなく, 現在見ている文節より前の文節に影響を及ぼすためである。

Algorithm lines	依存構造木	手がかり表現	contextual factuality	event factuality
2			CT+	
(4-7) 8-11 12-15	知らない	polarity: -(ない) modality: CT (知る) polarity: + (知る)	CT- CT+	知る: CT-
(4-7) (8-11) (12-15)	相手は ことを		CT+	
(4-7) 8-11 12-15	断念した	polarity: -(断念する)	CT+ CT-	断念する: CT+
(4-7) 8-11 (12-15)	彼が 出場を		CT- CT-	出場: CT-

図 1 (4) に対する事実性解析アルゴリズムの動き

例として, (4) に対するアルゴリズムの動きを図 1 に示す。

(4) 彼が出場を断念したことを相手は知らない。

まず contextual factuality の初期値 CT+を割り当て, 依存構造木を文末から走査していく。最初の文節「知らない」に対して, Step1 では極性を反転させる否定の機能表現「ない」が存在するので, contextual factuality を CT- に更新する。Step2 において「知る」は事実性を割り当てるべき事象であるため, そのときの contextual factuality (CT-) を「知る」の event factuality として出力する。Step3 では「知る」という述語が手がかり表現辞書内に存在するので, それに基づいて contextual factuality を CT+に更新する。各文節についてこれを繰り返し, 「知る」の事実性が CT-, 「断念する」の事実性が CT+, 「出場」の事実性が CT- という出力が最終的に得られる。

#### 4. 実験

日本語事実性解析における問題を整理・分析するために, 構築した解析器をウェブ上の数千文に対して適用し, その評価を行った。

実験では, 拡張モダリティタグ付与とコーパス<sup>14)</sup>の一部である, Yahoo!知恵袋 (OC) に含まれる 6,404 文 (形態素数: 110,649) に対してアルゴリズムを適用した。このコーパスは現

表 3 真偽判断ラベルと事実性との対応

確信度 \ 肯否極性	+	-	u
CT	成立 不成立から成立 (CT+)	不成立 成立から不成立 (CT-)	0 (Uu)
PR	高確率 低確率から高確率 (PR+)	低確率 高確率から低確率 (PR-)	
U			

代日本語書き言葉均衡コーパス (BCCWJ)\*<sup>1</sup> に対して、6 項目からなる拡張モダリティを付与したものである。

この中の真偽判断は、肯定か否定か、断定か推量かの組み合わせを表しているものであり、我々の事実性に相当する。態度や相対時といったラベルも事実性に影響を与える場合があるが、今回の実験では考慮しない。表 3 に真偽判断ラベルと我々の事実性との対応を示す。事実性と真偽判断ラベルとの対応を考え、CTu, PRu, U+, U-, Uu は全て 0 に対応づけた。また、我々の事実性には、「成立から不成立」や「高確率から低確率」などの変化を伴うラベルに相当するものは存在しない。これら変化を含意するラベルは、変化後が主要要素であると考え、変化後のラベルに対応づけた（「成立から不成立」→ CT-）。

(5) 支給が停止されました。

例えば (5) の事象「支給」の真偽判断ラベルは「成立から不成立」とされる。現在は「支給」されていないという解釈ができるので、CT- に対応づけた。

本実験で対象とする事象は、態度表明者が著者となっているものの中で、限定修飾・機能表現のように「対象外」が付与されていない 14,917 事象（約 2.33 事象/文）である。正解の形態素情報を入力し、構文解析を行った結果を、構築した事実性解析器の入力とする。尚、事実性を付与すべき事象の同定は、あらかじめ正解を与えた。

評価尺度としては、それぞれのラベル (CT+・PR+・PR-・CT-・Uu) に対する二値分類と考える。例えば CT+ に対しては、CT+ を正しく出力できているものを TP (True Positive), CT+ を出力しているが正解は異なるものを FP (False Positive), 正解には CT+ が付いているが正しく CT+ を出力できていないものを FN (False Negative) とする。他のラベルも同様にし、TP, FP, FN の数と、各ラベルに対する以下の 3 値を表 4 に示す。

表 4 それぞれのラベルに対する評価

	CT+	PR+	PR-	CT-	Uu	Micro-Average
TP	7,140	141	7	631	1,893	(9,812)
FP	1,467	286	71	583	2,698	(5,105)
FN	2,791	794	168	408	944	(5,105)
<b>Precision</b>	0.83	0.33	0.09	0.52	0.41	0.66
<b>Recall</b>	0.72	0.15	0.04	0.61	0.67	0.66
<b>F<sub>1</sub></b>	0.77	0.21	0.06	0.56	0.51	0.66

表 5 Confusion Matrix

正解 \ 出力	CT+	PR+	PR-	CT-	Uu	Total
CT+	<b>7,140</b>	237	25	408	2,121	9,931
PR+	528	<b>141</b>	8	23	235	935
PR-	47	6	<b>7</b>	67	48	175
CT-	97	1	16	<b>631</b>	294	1,039
Uu	795	42	22	85	<b>1,893</b>	2,837
Total	8,607	427	78	1,214	4,591	14,917

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

また、各正解ラベルに対してどのような出力が与えられたかを表す Confusion Matrix を表 5 に示す。表 4 には、多値分類と考えたときの精度（表 5 の太字部分の和を Total で割った値=Micro-Average）も併記する。これを見ると、特に PR- は確信度と肯否極性の両方に対して更新を行わなければならない、難易度が高いことがわかる。また、確信度および肯否極性の 2 軸を個別に分析するため、確信度および肯否極性を個別に評価し、その Confusion Matrix を作成した。これを表 6~9 に示す。肯否極性の方がわずかに結果が良くなり、確信度の判定は人間でも難しい場合がある、という直感に即している。

## 5. 誤り分析

3 節で述べた解析器は、確信度と肯否極性を独立して認識し、その組み合わせによって事実性を表している。2 つのうち肯否極性の解析誤りは、成立と不成立のような真逆の誤りにつながるため、事実性を利用するアプリケーションにおいて、確信度の誤りよりも深刻であ

\*1 <http://www.tokuteicorpus.jp/>

表 6 確信度に関する評価

	CT	PR	U	Micro-Average
Precision	0.84	0.32	0.41	0.69
Recall	0.75	0.15	0.67	0.69
F1	0.80	0.20	0.51	0.69

表 7 確信度に関する Confusion Matrix

正解 \ 出力	CT	PR	U	Total
CT	<b>8,276</b>	279	2,415	10,970
PR	665	<b>162</b>	283	1,110
U	880	64	<b>1,893</b>	2,837
Total	9,821	505	4,591	14,917

表 8 肯否極性に関する評価

	+	-	u	Micro-Average
Precision	0.89	0.56	0.41	0.71
Recall	0.74	0.59	0.67	0.71
F1	0.81	0.58	0.51	0.71

表 9 肯否極性に関する Confusion Matrix

正解 \ 出力	+	-	u	Total
+	<b>8,046</b>	464	2,356	10,866
-	151	<b>721</b>	342	1,214
u	837	107	<b>1,893</b>	2,837
Total	9,034	1,292	4,591	14,917

る場合が多い。そこで、肯否極性に関する誤り事例の中から、ランダムにおよそ 200 事例を選出し、分析を行った。ここで議論したい目的としては、以下の 3 点があげられる。

- I. 現在利用可能な語彙知識（機能表現辞書・手がかり表現辞書）によってどれだけの問題がカバーできるのか。カバーできなくても語彙知識の拡充および拡張によって解決できるのか。
  - II. 語彙の手がかりが曖昧性を持つ場合は多いのか。
  - III. 局所的な語彙知識の組み合わせだけで解決できない問題は存在するのか。
- 以上の論点に基づいて誤り分析を行った結果、以下のように誤りを分類することができた。
- i. 語彙知識のカバレッジの問題（内容語 16 %，機能語 8 %）
  - ii. 語彙知識としては存在しているが、表記揺れなど、辞書引きが難しい問題（2 %）
  - iii. 語彙的手がかりの多義性の問題（9 %）

iv. 手がかりとなる表現のスコープの問題（37 %）

v. その他（28 %）

本節では、これらの誤りについて例文とともに詳述する。

i. 語彙知識のカバレッジの問題：語彙知識のカバレッジの問題としては、(6) のような例がある。これは内容語のカバレッジに関する誤りの例である。

(6) サーバーは接続を解除しました。（正解：CT-，出力：CT+）

解析器は「接続」の肯否極性を+と出力したが、正解は-である。これは、事象の肯否極性を反転させる述語である「解除する」が手がかり表現辞書に存在しないことが原因である。

ii. 語彙知識としては存在しているが、表記揺れなど、辞書引きが難しい問題：割合としては少ないが、例えば表記揺れのように、語彙知識として存在はしているが、うまく引くことができない問題も存在し、(7) はその例である。

(7) 関東だから違うかもしれない。（正解：PR+，出力：PR-）

「違う」の肯否極性の正解は+であるが、解析器は-と出力してしまった。辞書には、推量の機能表現「かもしれない」は載っているが、口語的な表現「かもしれない」としては引けないためである。この種類の誤りは少なく、機能表現辞書では表記の異なりも考慮して収録できているため、多くの表現をカバーできていることがわかった。

以上の問題は、対応する語彙知識を充実させ、その表現を正しく認識できれば解決できる誤りである。しかし、語彙知識を充実させれば全ての問題が解けるわけではない。語彙知識の単純な適用だけでは難しい問題として、語彙的手がかりの多義性の問題や、手がかりとなる表現のスコープの問題が存在する。これらの問題は、語彙知識のカバレッジの問題よりも、割合の大きな重要な問題となっている。

iii. 語彙的手がかりの多義性の問題：語彙的手がかりの多義性に起因する誤りの例を(8)に示す。

(8) 歩けば交通費はかからないか、いい運動にもなるしね！（正解：CT-，出力：Uu）

この例では「ないか」が勧誘の「ないか」として認識されてしまい、Uuを出力するが、実際にはこれは勧誘の意味ではなく、否定の「ない」と自己確認の「か」との組み合わせである。「ないか」が勧誘の意味ではないことは、「交通費がかかる」が意思的な行為ではないことからわかる。このような局所的な文脈により、意味を同定するようなアプローチがまず考えられる<sup>13),19)</sup>。しかし、「ないか」が勧誘の意味ではないことと、「ない」が否定の意味であることは判別できても、「か」が並立の意味なのか自己確認の意味なのかを判別することは、文外の情報に依存するため、容易ではない。局所的な文脈だけを用いて意味を同定するアプ

ローチによって、どこまでの問題が解決でき、そこから先の問題はどの程度残るのかを明らかにする必要がある。

iv. 手がかりとなる表現のスコープの問題：(9) のようなスコープの問題は多く存在し、全体の誤りの 4 割近くを占めている。

(9) 車が見えたがぶつかるのを避け切れなかった。(正解：CT+, 出力：CT-)

(9) では「避け切れなかった」のスコープが問題となる。「避け切れなかった」のは「ぶつかる」であり「見える」ではないので「見える」の事実性は「避け切れなかった」に依存しないはずである。しかし、現在のアルゴリズムでは、係り元すべてをスコープとしているため、余計に伝搬がなされてしまい、間違った出力がされてしまう。スコープ外には伝搬させないようにする必要はあるが、スコープを完全に特定するのは容易ではない。まず「避け切れなかった」の項に関する知識が必要となる。そして、この項の範囲を特定する必要があり、このためにも世界知識が必要となる。単純に項に関するマーカーを特定すればよいのではなく、その項がどこからどこまでの範囲なのかを特定する必要はある、難しい問題である。

v. その他：その他の誤りとして、構文解析誤りやアノテーション誤り、連体修飾部の問題などが存在する。

(10) 信用できない人には貸しません。(正解：CT-, 出力：CT+)

例えば (10) は、連体修飾部が問題となる誤りである。「貸しません」の肯否極性 - が連体修飾部内の事象に対しても伝搬されてしまうため、間違った出力がされてしまう。このような限定修飾の場合の事実性などについて、判断可能であるかはそもそも問題である。

また、肯否極性と同様に、確信度においても同様の誤りが見られた。例えば語彙の手がかりの多義性に起因する誤りの例を (11) に示す。(12) は語彙知識のカバレッジ、スコープ、アルゴリズムの拡張と、多くの問題が含まれる例である。

(11) 何分かおきに音で知らせてくれるような機械です。(正解：CT+, 出力：PR+)

(12) おそらく、親戚は自慢したいだけの見栄っ張りです。(正解：PR+, 出力：CT+)

(11) では、本来は比況の意味である機能表現「よう」を誤って推量と認識してしまったため「知る」の確信度を PR と判定してしまっている。(12) では「見栄っ張り」に対して「おそらく」がその確信度を下げるが、今回利用した語彙知識では副詞はカバーできていない。このような確信性に影響する副詞は、川添ら (2011)<sup>17),18)</sup> によって体系化されており、ある程度カバーできる。また、カバレッジの問題を解決できても、「おそらく」は「見栄っ張り」だけでなく「親戚」や「自慢したい」に係る解釈も可能であるので、副詞が持つスコープを特定する必要がある。さらに、現在のアルゴリズムでは依存構造木の根から伝搬させて

いくので、副詞のように後続の文節に影響を与える手がかり表現にも対応できるよう、アルゴリズムの拡張も必要となる。

## 6. おわりに

本稿では、日本語事実性解析において、事実性に影響を与える語彙知識に焦点を当て、語彙知識の組み合わせに関する種々の問題を分析・整理することを目的として、Saurí (2008) のモデルをもとに日本語事実性解析システムを構築し、実験・誤り分析を行った。その結果、語彙知識のカバレッジの問題・語彙知識としては存在しているが、表記揺れなど、辞書引きが難しい問題・語彙の手がかりの多義性の問題・手がかりとなる表現のスコープの問題などが存在することがわかった。

この分析結果に基づき、今後、語彙知識の整備や語義曖昧性解消、およびスコープの特定といった種々の課題に取り組む必要がある。例えば語彙知識の整備は、現在の辞書に不足している表現を、現在の辞書をシードとして獲得していくことが考えられる。また、本稿では考慮しなかった、著者以外の態度表明者に関する事実性や、文外の談話構造を考慮した事実性などの解析も考えるべき事項である。今後は更なる分析とともにこれらの課題を進め、語彙知識および言語現象を加味した、高度な事実性解析器の構築を行う。

## 謝 辞

本研究を遂行するにあたり、多大な助力をいただきました山梨大学の松吉俊氏、ならびに NLP 若手の会 第 6 回シンポジウムにてコメントいただいた方々に心より感謝いたします。

## 参 考 文 献

- 1) R.Farkas, Veronika Vincze, G.Móra, J.Csirik, and G.Szarvas. The CoNLL-2010 shared task: learning to detect hedges and their scope in natural language text. *Proceedings of the Fourteenth Conference on Computational Natural Language Learning - Shared Task*, pp. 1–12. Association for Computational Linguistics, 2010.
- 2) Kentaro Inui, Shuya Abe, Kazuo Hara, Hiraku Morita, Chitose Sao, Megumi Eguchi, Asuka Sumida, Koji Murakami, and Suguru Matsuyoshi. Experience Mining: Building a Large-Scale Database of Personal Experiences and Opinions from Web Documents. In *the 2008 IEEE/WIC/ACM International Conference on Web Intelligence*, pp. 314–321, 2008.
- 3) Marc Light, XinYing Qiu, and Padmini Srinivasan. The language of bioscience: Facts, speculations, and statements in between. In *Proceedings of BioLink 2004*

- workshop on linking biological literature, ontologies and databases*, pp. 17–24, 2004.
- 4) Ben Medlock and Ted Briscoe. Weakly supervised learning for hedge classification in scientific literature. In *Annual Meeting of Association of Computational Linguistics*, Vol.45, pp. 992–999, 2007.
  - 5) Roser Morante, Sarah Schrauwen, and Walter Daelemans. Corpus-based approaches to processing the scope of negation cues: an evaluation of the state of the art. In *Proc. of IWCS 2011*, pp. 350–354, 2011.
  - 6) Roser Saurí. *A factuality profiler for eventualities in text*. PhD thesis, Brandeis University, 2008.
  - 7) Roser Saurí. FactBank 1.0 Annotation Guidelines, 2008.
  - 8) Roser Saurí, Jessica Littman, Robert Gaizauskas, Andrea Setzer, and James Pustejovsky. TimeML Annotation Guidelines, Version 1.2.1, 2006.
  - 9) Roser Saurí and James Pustejovsky. FactBank: a corpus annotated with event factuality. *Language resources and evaluation*, Vol.43, No.3, pp. 227–268, 2009.
  - 10) György Szarvas, Veronika Vincze, Richárd Farkas, and J.Csirik. The BioScope corpus: annotation for negation, uncertainty and their scope in biomedical texts. In *Proceedings of the Workshop on Current Trends in Biomedical Natural Language Processing*, pp. 38–45, 2008.
  - 11) 原一夫, 乾健太郎. 事態抽出のための事実性解析. 情報処理学会研究報告, 2008-FI-89, 2008-NL-183, pp. 75–80, 2008.
  - 12) 江口萌, 松吉俊, 佐尾ちとせ, 乾健太郎, 松本裕治. モダリティ、真偽情報、価値情報を統合した拡張モダリティ解析. 言語処理学会第 16 回年次大会発表論文集, pp. 852–855, 2010.
  - 13) 今村賢治, 泉朋子, 菊井玄一郎, 佐藤理史. 述部機能表現の意味ラベルタガー. 言語処理学会第 17 回年次大会発表論文集, pp. 308–311, 2011.
  - 14) 松吉俊, 江口萌, 佐尾ちとせ, 村上浩司, 乾健太郎, 松本裕治. 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, Vol. J93-D, No.6, pp. 705–713, 2010.
  - 15) 松吉俊, 佐藤理史, 宇津呂武仁. 日本語機能表現辞書の編纂. 自然言語処理, Vol.14, pp. 123–146, 2007.
  - 16) 松吉俊, 佐尾ちとせ, 乾健太郎, 松本裕治. 拡張モダリティタグ付与コーパスの設計と構築. 言語処理学会第 17 回年次大会発表論文集, pp. 147–150, 2011.
  - 17) 川添愛, 齊藤学, 片岡喜代子, 崔榮殊, 戸次大介. 確実性判断のためのアノテーション済みコーパスの構築. 言語処理学会第 17 回年次大会発表論文集, pp. 143–146, 2011.
  - 18) 川添愛, 齊藤学, 片岡喜代子, 崔榮殊, 戸次大介. 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, Department of Information Science, Ochanomizu University, OCHA-IS 10-4, 2011.
  - 19) 鈴木敬文, 阿部佑亮, 宇津呂武仁, 松吉俊, 土屋雅稔. 大規模階層辞書と用例を用いた日

本語機能表現の解析. 『現代日本語書き言葉均衡コーパス』完成記念講演会予稿集, pp. 105–110, 2011.